Fondements des Mathématiques II

J. Melleray, Ch. Mercat

Université Lyon I
Semestre de printemps 2021

Avertissement. Ces notes contiennent certainement leur lot d'erreurs, typographiques ou autres. Si vous en trouvez, n'hésitez pas à me contacter pour me les signaler.

Table des matières

Première partie

Algèbre

Chapitre 1

Calcul matriciel

Dans tout ce chapitre la lettre $\mathbb{K}$ désignera le corps $\mathbb{Q}, \mathbb{R}$, ou $\mathbb{C}$.

Les matrices sont des tableaux rectangulaires de nombres. Par exemple l'ensemble des notes des étudiants du cursus prépa sur l'année pour toutes les évaluations (en négligeant les abscences), ou bien un sondage : la réponse à une question est codée sous forme numérique, une réponse individuelle donne une colonne et toute l'enquête donne une grosse matrice.

1.1 Systèmes et point de vue matriciel

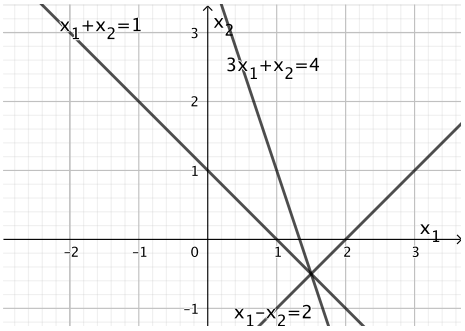
Un système d'équations linéaires à n équations et m inconnues $x_1, \dots, x_m$ dans $\mathbb{K}$ est défini par des constantes $(a_{i,j})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant m}} \in \mathbb{K}^{n \times m}$ et $(b_i)_{1 \leqslant i \leqslant n} \in \mathbb{K}^n$ et se présente sous la forme

$$\begin{cases} a_{1,1}x_1 + \ldots + a_{1,m}x_m &= b_1, \\ a_{2,1}x_1 + \ldots + a_{2,m}x_m &= b_2, \\ \qquad\vdots & \qquad\vdots \\ a_{n,1}x_1 + \ldots + a_{n,m}x_m &= b_n. \end{cases}$$

Exemple.

$n = 3$, $m = 2$, $a_{1,1} = a_{1,2} = a_{2,1} = 1$, $a_{2,2} = -1$, $a_{3,1} = 3$, $a_{3,2} = 1$, $b_1 = 1$, $b_2 = 2$, $b_3 = 4$:

$$\begin{cases} x_1 + x_2 &= 1, \\ x_1 - x_2 &= 2, \\ 3\,x_1 + x_2 &= 4. \end{cases}$$



Notons qu'une droite $\mathcal{D}$ dans $\mathbb{R}^2$ a une forme cartésienne : on trouve trois nombres $a, b, c \in \mathbb{R}$ tels que $\mathcal{D} = \{(x_1, x_2) \in \mathbb{R}^2 : ax_1 + bx_2 = c\}$ ou plus simplement $\mathcal{D} : ax_1 + bx_2 = c$ en sous-entendant l'ensemble des points solutions de cette équation. Le *produit scalaire* de deux vecteurs $\vec{n} = \begin{pmatrix} a \\ b \end{pmatrix}$ et $\vec{m} = \begin{pmatrix} a' \\ b' \end{pmatrix}$ est le scalaire (nombre, en opposition à vecteur) $\vec{n} \cdot \vec{m} = aa' + bb' \in \mathbb{R}$. Dans un repére orthonormé, on comprend l'équation cartésienne comme décrivant la droite *perpendiculaire* au vecteur $\vec{n}$ passant par un point M tel que $\overrightarrow{OM} \cdot \vec{n} = c$. Si $c = 0$, c'est une droite linéaire, qui passe par l'origine ; si $c > 0$, M est du côté de $\vec{n}$; si $c < 0$, M est à l'opposé de la direction de $\vec{n}$. Ceci se généralise en dimension supérieure, le produit scalaire des vecteurs de coordonnées $(a_i)_{1 \leqslant i \leqslant n}, (b_i)_{1 \leqslant i \leqslant n} \in \mathbb{R}^n$ est $\sum_{i=1}^n a_i b_i \in \mathbb{R}$ et cela permet de définir des "hyper-plans". Par exemple, en dimension trois, un plan perpendiculaire au vecteur de coordonnées $\vec{n} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$, dans un repère

orthonormé, a une équation cartésienne de la forme $ax + by + cz = d$ où d est la valeur commune du produit scalaire $\overrightarrow{OM} \cdot \vec{n}$ pour tout point M du plan. Nous n'élaborerons pas ici sur ce sujet, qui sera étudié en L2, mais ce produit scalaire n'a un sens géométrique que quand le système est **orthonormé**. Alors, $\vec{n} \cdot \vec{m} = 0$ est équivalent au fait que les vecteurs sont *orthogonaux*.

Trouver les solutions (s'il y en a) du système représenté plus haut, revient à trouver l'image inverse de $\{b_1, \dots, b_n\}$ par une certaine fonction associée au système : la fonction $f\colon \mathbb{K}^m \to \mathbb{K}^n$ définie par

$$f(x_1, \dots, x_m) = (a_{1,1}x_1 + \dots + a_{1,n}x_m, \dots, a_{n,1}x_1 + \dots a_{n,m}x_m).$$

En effet, dire que $(x_1, \dots, x_m)$ est solution du système revient à dire que $f(x_1, \dots, x_m) = (b_1, \dots, b_n)$, ou encore que $(x_1, \dots, x_m) \in f^{-1}\big(\{(b_1, \dots, b_n)\}\big)$.

Dans notre exemple 1.1, la solution est $(x_1, x_2) = (\frac{3}{2}, -\frac{1}{2})$, qu'on peut déterminer graphiquement comme intersection des droites et vérifier numériquement comme solution de chacune des trois équations. Nous allons voir un algorithme de résolution d'un tel système.

Toute l'information sur la fonction f est contenue dans les valeurs de $a_{1,1}, \dots, a_{1,m}, \dots, a_{n,1}, \dots, a_{n,m}$; on regroupe ces nombres dans un tableau appelé une *matrice*.

Définition 1.1. Une *matrice* à n lignes et m colonnes à coefficients dans $\mathbb{K}$ est un tableau de la forme

$$\begin{pmatrix} a_{1,1} & \dots & a_{1,m} \\ \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,m} \end{pmatrix}.$$

Dans notre exemple, c'est $\begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 3 & 1 \end{pmatrix}$. Quand $n = 1$ on dit que A est une *matrice ligne* ; quand $m = 1$ on dit que A est une *matrice colonne*. Enfin, quand $n = m$ on dit que la matrice est une *matrice carrée*.

On note $M_{n,m}(\mathbb{K})$ l'ensemble des matrices à n lignes et m colonnes à coefficients dans $\mathbb{K}$. L'ensemble des matrices carrées à n lignes et n colonnes est simplement noté $M_n(\mathbb{K})$.

Ces objets sont très utiles pour mener à bien des calculs pratiques ; souvent, on est en fait intéressé par un système ou une fonction associée à la matrice comme ci-dessus.

1.2 Opérations sur les matrices

Les opérations sur les fonctions se traduisent en opérations sur les matrices ; notons que si f, g sont deux fonctions de $\mathbb{K}^m$ dans $\mathbb{K}^n$ alors on peut considérer leur somme $f + g\colon x \in \mathbb{K}^m \mapsto f(x) + g(x) \in \mathbb{K}^n$.

Définition 1.2. Soit $n, m \in \mathbb{N}^*$, et $A, B \in M_{n,m}(\mathbb{K})$, deux matrices de même taille, leur somme $A + B$ est la matrice dont le coefficient sur la i-ième ligne et la j-ième colonne est égal à $a_{i,j} + b_{i,j}$; et leur différence $A - B$ est la matrice de coefficients $a_{i,j} - b_{i,j}$.

Par exemple, $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 4 & 4 \end{pmatrix}$. La somme de $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ et $\begin{pmatrix} 12 \\ 21 \end{pmatrix}$ n'est pas définie.

On peut aussi multiplier une matrice par un *scalaire*, c'est-à-dire un élément de $\mathbb{K}$.

Définition 1.3. Soit $n, m \in \mathbb{N}^*$, $\lambda \in \mathbb{K}$ et $A \in M_{n,m}(\mathbb{K})$. Alors λA est la matrice de coefficients $\lambda a_{i,j}$.

Par exemple, $3 \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 3 & 6 \\ 9 & 12 \end{pmatrix}$.

Notons que l'addition des matrices est commutative : pour tout $A, B \in M_{n,m}(\mathbb{K})$ on a $A + B = B + A$. Elle est également associative : $(A + B) + C = A + (B + C)$, et en général on l'écrit simplement $A + B + C$. Enfin, pour tout $\lambda \in \mathbb{K}$ on a $\lambda(A + B) = \lambda A + \lambda B$.

Le produit entre deux matrices est une opération plus complexe ; il correspond à la *composition* des fonctions associées aux matrices. On peut composer deux fonctions $f\colon \mathbb{K}^m \to \mathbb{K}^n$ et $g\colon K^{m'} \to \mathbb{K}^{n'}$ et considérer la fonction $f \circ g$ exactement quand l'espace d'arrivée de g est égal à l'espace de définition de f, autrement dit quand $n' = m$. En termes de matrices, le produit de deux matrices AB sera donc défini exactement quand B a autant de *lignes* que A a de *colonnes*.

Définition 1.4. Soit $n, m, n', m' \in \mathbb{N}^*$, $A \in M_{n,m}(\mathbb{K})$ et $B \in M_{n',m'}(\mathbb{K})$. Alors le produit AB est défini si, et seulement si, $m = n'$, et dans ce cas c'est la matrice à n lignes et m' colonnes dont le coefficient sur la ligne i et la colonne j est égal à

$$(AB)_{i,j} = \sum_{k=1}^{m} a_{i,k} b_{k,j} \ .$$

Autrement dit, le coefficient pour une ligne i et une colonne j données est ce qu'on appelle le *produit scalaire* de la ligne $A_{i,\cdot}$ avec la colonne $B_{\cdot,j}$. Attention, signalons que l'interprétation géométrique du produit scalaire, rencontré au lycée, se fait dans une *base orthonormée*, en particulier, le produit scalaire nul entre une ligne et une colonne ne s'interprète comme leur orthogonalité que si la base est orthonormale ! L'étude de ces questions n'est pas à l'ordre du jour de ce cours et sera abordée l'année prochaine.

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,k} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,k} & \cdots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots & \iddots & \vdots \\ a_{i,1} & a_{i,2} & \cdots & a_{i,k} & \cdots & a_{i,m} \\ \vdots & \vdots & \iddots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,k} & \cdots & a_{n,m} \end{pmatrix} \begin{pmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,j} & \cdots & b_{1,m'} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,j} & \cdots & b_{2,m'} \\ \vdots & \vdots & \ddots & \vdots & \iddots & \vdots \\ b_{k,1} & b_{k,2} & \cdots & b_{k,j} & \cdots & b_{k,m'} \\ \vdots & \vdots & \iddots & \vdots & \ddots & \vdots \\ b_{m,1} & b_{m,2} & \cdots & b_{m,j} & \cdots & b_{m,m'} \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{k=1}^m a_{1,k}b_{k,1} & \sum_{k=1}^m a_{1,k}b_{k,2} & \cdots & \sum_{k=1}^m a_{1,k}b_{k,j} & \cdots & \sum_{k=1}^m a_{1,k}b_{k,m'} \\ \sum_{k=1}^m a_{2,k}b_{k,1} & \sum_{k=1}^m a_{2,k}b_{k,2} & \cdots & \sum_{k=1}^m a_{2,k}b_{k,j} & \cdots & \sum_{k=1}^m a_{2,k}b_{k,m'} \\ \vdots & \vdots & \ddots & \vdots & \iddots & \vdots \\ \sum_{k=1}^m a_{i,k}b_{k,1} & \sum_{k=1}^m a_{i,k}b_{k,2} & \cdots & \sum_{k=1}^m a_{i,k}b_{k,j} & \cdots & \sum_{k=1}^m a_{i,k}b_{k,m'} \\ \vdots & \vdots & \iddots & \vdots & \ddots & \vdots \\ \sum_{k=1}^m a_{n,k}b_{k,1} & \sum_{k=1}^m a_{n,k}b_{k,2} & \cdots & \sum_{k=1}^m a_{n,k}b_{k,j} & \cdots & \sum_{k=1}^m a_{n,k}b_{k,m'} \end{pmatrix}$$

Pourquoi cette formule ? Pensons que A est la matrice associée à une application $f\colon \mathbb{K}^m \to \mathbb{K}^n$, et B est la matrice associée à une application $g\colon \mathbb{K}^{m'} \to \mathbb{K}^m$. Alors AB doit être la matrice associée à l'application $f \circ g\colon \mathbb{K}^{m'} \to \mathbb{K}^n$. Pour $x = (x_1, \ldots, x_{m'})$, calculons :

$$\begin{aligned} f \circ g(x_1, \ldots, x_{m'}) &= f\big(g(x_1, \ldots, x_{m'})\big) \\ &= f\left(\sum_{j=1}^{m'} b_{1,j}x_j, \ldots, \sum_{j=1}^{m'} b_{k,j}x_j, \ldots, \sum_{j=1}^{m'} b_{m,j}x_j\right) \\ &= \left(\sum_{k=1}^{m} a_{1,k}(\sum_{j=1}^{m'} b_{k,j}x_j), \ldots, \sum_{k=1}^{m} a_{i,k}(\sum_{j=1}^{m'} b_{k,j}x_j), \ldots, \sum_{k=1}^{m} a_{n,k}(\sum_{j=1}^{m'} b_{k,j}x_j)\right) \\ &= \left(\sum_{j=1}^{m'}(\sum_{k=1}^{m} a_{1,k}b_{k,j})x_j, \ldots, \sum_{j=1}^{m'}(\sum_{k=1}^{m} a_{i,k}b_{k,j})x_j, \ldots, \sum_{j=1}^{m'}(\sum_{k=1}^{m} a_{n,k}b_{k,j})x_j\right) \end{aligned}$$

L'application $f \circ g$ a donc pour matrice la matrice à n lignes et m' colonnes dont le coefficient sur la ligne i et la colonne j est égal à $\sum_{k=1}^m a_{i,k}b_{k,j}$, ce qui correspond bien à notre définition du produit de matrices.

Avec ces conventions, si x est l'élément de $\mathbb{K}^m$ de coordonnées $x_1, \ldots, x_m$, et $f\colon \mathbb{K}^m \to \mathbb{K}^n$ a pour matrice A, et qu'on considère le vecteur colonne à m lignes $X = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$, la matrice colonne à n lignes

$$AX = x_1 \begin{pmatrix} a_{1,1} \\ \vdots \\ a_{n,1} \end{pmatrix} + x_2 \begin{pmatrix} a_{1,2} \\ \vdots \\ a_{n,2} \end{pmatrix} + \cdots + x_m \begin{pmatrix} a_{1,m} \\ \vdots \\ a_{n,m} \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^m a_{1,k}x_k \\ \vdots \\ \sum_{k=1}^m a_{n,k}x_k \end{pmatrix}$$

est la combinaison linéaire des colonnes de A dont les coefficients sont donnés par les x_k et peut être vue comme écrivant *en colonne* l'image $f(x)$. Toujours en écrivant les vecteurs en colonnes, on voit donc que résoudre un système revient à résoudre une équation de la forme $AX = B$, d'inconnue X, où B et X sont deux vecteurs colonnes (X a le même nombre m de lignes que d'inconnues du systèmes, tandis que le nombre de lignes de B est le nombre n d'équations du système).

Exemple. Dans notre exemple de départ 1.1, la solution du système s'écrit $\begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} \frac{3}{2} \\ -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}$.

Ce produit peut paraître surprenant (notez les espaces différents !) :

$$\begin{pmatrix} 2 & 3 & -1 \end{pmatrix} \begin{pmatrix} 4 \\ -2 \\ 3 \end{pmatrix} = \begin{pmatrix} 8-6-3 \end{pmatrix} = (-1) \in M_1(\mathbb{R}) \ ;$$

$$\text{ou encore } \begin{pmatrix} 0 & -1 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} 2 & -3 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \text{ alors que } \begin{pmatrix} 2 & -3 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 0 & 5 \end{pmatrix} = \begin{pmatrix} 0 & -17 \\ 0 & 0 \end{pmatrix}.$$

Le dernier exemple ci-dessus nous montre deux phénomènes inhabituels : on peut avoir $AB = 0$ (la matrice nulle) alors que $A \neq 0$ et $B \neq 0$; et on peut avoir $AB \neq BA$. On a tout de même l'associativité du produit de matrices.

Exercice 1.5. Montrer que si A, B, C sont trois matrices telles que AB et BC sont définis, alors $A(BC)$ et $(AB)C$ sont bien définis et $A(BC) = (AB)C$. Autrement dit, le produit matriciel est associatif.

On retrouve la règle habituelle de distributivité : pour toutes matrices $A \in M_{n,m}(\mathbb{K})$ et $B, C \in M_{m,p}(\mathbb{K})$ on a $A(B + C) = AB + AC$.

Définition 1.6. Soit n, m deux entiers. La *matrice nulle* $0_{n,m} \in M_{n,m}(\mathbb{K})$ est la matrice dont tous les coefficients valent 0. Quand elle est carrée, on la note 0_n et s'il n'y a pas de risque de confusion, on la note parfois simplement 0.

Notons que pour tout $A \in M_{n,m}(\mathbb{K})$ on a $0_{n,m} + A = A + 0_{n,m} = A$; et si $B \in M_{m,p}(\mathbb{K})$ alors $0_{n,m}B$ est bien défini et $0_{n,m}B = 0_{n,p}$.

Définition 1.7. Soit $n \in \mathbb{N}^*$. La matrice identité I_n est l'élément de $M_n(\mathbb{K})$ dont le coefficient sur la i-ième ligne et la j-ième colonne vaut 0 si $i \neq j$, et 1 si $i = j$. En notant $\delta_{i,j} = 1$ si $i = j$, 0 si $i \neq j$, le coefficient (i, j) de I_n est donc $\delta_{i,j}$.

Proposition 1.8. *Pour toute matrice $A \in M_n(\mathbb{K})$ on a $I_n A = AI_n = A$.*

Démonstration. Par définition, AI_n est l'élément de $M_n(\mathbb{K})$ dont le coefficient sur la ligne i et la colonne j vaut

$$\sum_{k=1}^{n} a_{i,k}\delta_{k,j} = a_{i,j}\delta_{j,j} = a_{i,j} \ .$$

Donc on a comme attendu $AI_n = A$. De même $I_n A$ a pour coefficient (i, j)

$$\sum_{k=1}^{n} \delta_{i,k}a_{k,j} = \delta_{i,i}a_{i,j} = a_{i,j} \ .$$

□

En général, on ne peut pas multiplier une matrice A par elle-même : le seul cas où c'est possible est celui où A a autant de lignes que de colonnes, c'est-à-dire le cas où A est une matrice carrée.

Définition 1.9. Soit $n \in \mathbb{N}^*$ et $A \in M_n(\mathbb{K})$. Les *puissances* de A sont les matrices $A^i \in M_n(\mathbb{K})$ définies par récurrence par

$$A^0 = I_n \quad ; \quad \forall i \in \mathbb{N}, \ A^{i+1} = A \cdot A^i \ .$$

On vérifie facilement (par exemple, par récurrence sur j) que pour tout $i, j \in \mathbb{N}$ on a $A^{i+j} = A^i A^j$.

Une matrice et ses puissances peuvent modéliser différents phénomènes, comme les suites récurrentes linéaires, les graphes ou les processus de MARKOV.

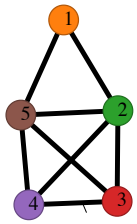
Exemple. Soit la suite *récurrence linéaire d'ordre deux* $(u_n)_{n\in\mathbb{N}}$, définie à l'aide de ses deux premiers termes $u_0, u_1 \in \mathbb{K}$ et $a, b \in \mathbb{K}$, par $\forall n \in \mathbb{N}, \quad u_{n+2} = au_{n+1} + bu_n$. On a pour tout $n \in \mathbb{N}$,

$$\begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix} = \begin{pmatrix} a & b \\ 1 & 0 \end{pmatrix}^n \begin{pmatrix} u_1 \\ u_0 \end{pmatrix}.$$

Exemple. Une matrice $\sigma \in M_n(\mathbb{K})$ de *permutation* est une matrice carrée ne contenant que des zéros, excepté quelques $\sigma_{i\ j} = 1$, un sur chaque ligne i et un sur chaque colonne j. Elle modélise une permutation où le j-ième terme est mis à la place du i-ème. Par exemple $\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ décrit la permutation circulaire $(1, 2, 3) \mapsto (2, 3, 1)$ notée aussi $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$.

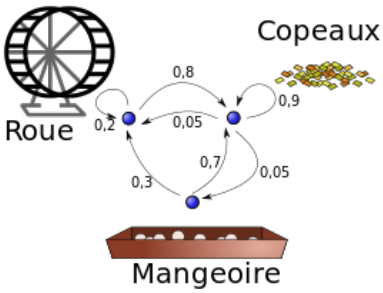
Exemple. Un graphe peut être codé par sa matrice d'adjacence carrée, qui compte une colonne et une ligne par sommet et un nombre entier codant l'adjacence de deux sommets. Pour un graphe non orienté, la matrice est symétrique, le coefficient en ligne i et colonne j est égal au coefficient en ligne j et colonne i.

Par exemple le graphe ci-contre est codé par la matrice $A = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{pmatrix}$. La puissance A^n compte le nombre de façons d'aller d'un sommet à un autre en un chemin continu de n arêtes. Par exemple $A^3 = \begin{pmatrix} 2 & 7 & 4 & 4 & 7 \\ 7 & 8 & 9 & 9 & 9 \\ 4 & 9 & 6 & 7 & 9 \\ 4 & 9 & 7 & 6 & 9 \\ 7 & 9 & 9 & 9 & 8 \end{pmatrix}$, on lit qu'il y a 2 façons d'aller du sommet 1 à lui-même en 3 étapes, c'est en tournant autour du triangle supérieur, dans un sens ou dans l'autre.



Exemple. (D'après Wikipedia) Doudou le hamster ne connaît que trois endroits dans sa cage : les copeaux où il dort, la mangeoire où il mange et la roue où il fait de l'exercice. Ses journées sont assez semblables les unes aux autres, et son activité se représente aisément par *une chaîne de* MARKOV. Toutes les minutes, il peut soit changer d'activité, soit continuer celle qu'il était en train de faire. On modélise son comportement par un processus sans mémoire.

- Quand il dort, il a 9 chances sur 10 de ne pas se réveiller la minute suivante.
- Quand il se réveille, il y a 1 chance sur 2 qu'il aille manger et 1 chance sur 2 qu'il parte faire de l'exercice.
- Le repas ne dure qu'une minute, après il fait autre chose.
- Après avoir mangé, il y a 3 chances sur 10 qu'il parte courir dans sa roue, mais surtout 7 chances sur 10 qu'il retourne dormir.
- Courir est fatigant pour Doudou ; il y a 8 chances sur 10 qu'il retourne dormir au bout d'une minute. Sinon il continue en oubliant qu'il est déjà un peu fatigué.



On représente sa probabilité d'occupation d'un état par un vecteur ligne (copeaux, mangeoire, roue) et la transition d'un état à un autre par la matrice $P = \begin{pmatrix} 0,9 & 0,05 & 0,05 \\ 0,7 & 0 & 0,3 \\ 0,8 & 0 & 0,2 \end{pmatrix}$. Ses puissances permettent d'étudier le comportement moyen du hamster, en l'occurrence son état stationnaire est $q = \begin{pmatrix} 0,884 & 0,0442 & 0,0718 \end{pmatrix}$, il passe 88% de son temps à dormir, 4% à manger et 7% à faire de l'exercice.

Définition 1.10. Soit $n \in \mathbb{N}^*$ et $A \in M_n(\mathbb{K})$. On dit que A est *inversible* s'il existe $B \in M_n(\mathbb{K})$ telle que $AB = BA = I_n$. On dit alors que B est l'*inverse* de A et on note $B = A^{-1} \in GL_n(\mathbb{K})$ le groupe linéaire.

Seule une matrice carrée peut être inversible ; notons que l'inverse, s'il existe, est unique : si $AB = BA = I_n$ et $AC = CA = I_n$ alors on a d'une part $C(AB) = CI_n = C$ et $C(AB) = (CA)B = I_nB = B$, d'où $B = C$.

On verra un peu plus bas que, quand $A, B \in M_n(\mathbb{K})$, avoir $AB = I_n$ entraîne nécessairement que $BA = I_n$; mais pour l'instant on n'a pas les moyens de prouver cela.

Proposition 1.11. *1. Si A est inversible alors A^{-1} est inversible et $(A^{-1})^{-1} = A$.*

2. Si $A, B \in M_n(\mathbb{K})$ sont inversibles alors AB est inversible et $(AB)^{-1} = B^{-1}A^{-1}$.

3. Si A est inversible alors pour tout $m \in \mathbb{N}$ A^m est inversible et $(A^m)^{-1} = (A^{-1})^m$; on note simplement cette matrice A^{-m}.

Démonstration. Le premier point est immédiat : on a $A^{-1}A = AA^{-1} = I_n$, ce qui montre à la fois que A^{-1} est inversible et que son inverse est A.

Le deuxième point découle de l'associativité du produit de matrices :

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AA^{-1} = I_n \ ; \ (B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}B = I_n \ .$$

Le troisième point est aussi une conséquence de l'associativité et se montre par exemple par récurrence : le résultat est clair pour $m = 0, 1$. S'il est vrai au rang m alors on a

$$A^{m+1}(A^{-1})^{m+1} = A(A^m(A^{-1})^m)A^{-1} = AA^{-1} = I_n \ .$$

□

1.3 Matrices élémentaires et opérations élémentaires sur les lignes et les colonnes

Une méthode pour résoudre les systèmes d'équations est de faire des opérations sur ces équations pour se ramener à un système plus simple. Ces opérations sont de la forme : multiplier une équation L_i par λ, un scalaire non nul ; remplacer l'équation L_i par la combinaison $L_i + \lambda L_j$, où $j \neq i$ et λ est un scalaire ; et permuter les équations L_i et L_j. Quand on code un système par des équations en ligne, ces trois opérations sont appelées *opérations élémentaires sur les lignes*.

Exemple. Dans notre système de départ 1.1, $L_2 - L_1 \to L_2$ et $L_3 - 3L_1 \to L_3$ donnent $\begin{cases} x_1 + x_2 & = 1 \\ -2x_2 & = 1 \\ -2x_2 & = 1 \end{cases}$ puis $-\frac{1}{2}L_2 \to L_2$ et $L_1 - L_2 \to L_1$ donnent la solution $x_1 = \frac{3}{2}$, $x_2 = -\frac{1}{2}$.

Ces transformations peuvent être interprétées en terme de produit matriciel.

Définition 1.12. Soit $n \in \mathbb{N}^*$.

1. Pour $i \in \{1, \ldots, n\}$ et $\lambda \in \mathbb{K}^*$, la matrice $E_i(\lambda)$ est obtenue en multipliant par λ la i-ième ligne de I_n (et en ne touchant pas aux autres lignes). Par exemple (pour $n = 5$)

$$E_3(12) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 12 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

2. Pour $i, j \in \{1, \ldots, n\}$, la matrice $E_{i,j}$ est obtenue en échangeant la i-ème et la j-ième ligne de I_n. Par exemple (toujours pour $n = 5$) :

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

3. Enfin, pour $i, j \in \{1, \dots, n\}$ et $\lambda \in \mathbb{K}$, la matrice $E_{i,j}(\lambda)$ est la matrice obtenue en ajoutant λ fois la j-ième ligne de I_n à sa i-ième ligne. Par exemple (sempiternellement pour $n = 5$) :

$$E_{2,1}(3) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Ces notations n'ont rien de normalisé. Remarquons que les matrices élémentaires sont celles qu'on obtient en appliquant à I_n une opération élémentaire sur les lignes ; on pourrait aussi définir les opérations élémentaires sur les colonnes, et on vérifie facilement que les matrices élémentaires sont également celles qui sont obtenues en appliquant à I_n une opération élémentaire sur les colonnes.

Le lien entre matrices élémentaires et opérations élémentaires est le suivant : multiplier A *à gauche* par une matrice élémentaire revient à appliquer à A la même opération élémentaire qui a servi à définir la matrice élémentaire en question. Plus précisément :

Proposition 1.13. *Soit n, m deux entiers, et $A \in M_{n,m}(\mathbb{K})$. Alors :*

1. *Pour $i \in \{1, \dots, n\}$ et $\lambda \in \mathbb{K}^*$, la matrice $E_i(\lambda)A$ est la matrice obtenue en appliquant à A l'opération $\lambda L_i \to L_i$.*
2. *Pour $i, j \in \{1, \dots, n\}$, la matrice $E_{i,j}A$ est la matrice obtenue en permutant la i-ième ligne et la j-ième ligne de A.*
3. *Pour $i, j \in \{1, \dots, n\}$ et $\lambda \in \mathbb{K}$, la matrice $E_{i,j}(\lambda)A$ est la matrice obtenue en appliquant à A l'opération $L_i + \lambda L_j \to L_i$.*

Cette propriété se vérifie facilement et on laisse la preuve en exercice. Notons aussi que multiplier A *à droite* par une matrice élémentaire revient à appliquer une opération élémentaire *sur les colonnes* de A.

Proposition 1.14. *les matrices élémentaires sont inversibles ; et on a les formules :*

$$(E_i(\lambda))^{-1} = E_i(\frac{1}{\lambda}) \; ; \quad (E_{i,j})^{-1} = E_{i,j}; \quad (E_{i,j}(\lambda))^{-1} = E_{i,j}(-\lambda) \; .$$

Définition 1.15.

— *L'image* $\mathrm{Im}(A) \in \mathbb{K}^n$ d'une matrice $A \in M_{n,m}(\mathbb{K})$ est l'ensemble $\{v \in \mathbb{K}^n / \exists (\lambda_k)_{1 \leqslant k \leqslant m} \in \mathbb{K}^m, v = \sum_{k=1}^m \lambda_k A_{.,k}\}$ des vecteurs combinaisons linéaires des colonnes de la matrice, c'est-à-dire que $v_i = \sum_{k=1}^m \lambda_k A_{i,k}$ pour tout $1 \leqslant i \leqslant n$, ou encore $v = Au$ pour $u = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix}$.

— *Le noyau* $\ker(A) \in \mathbb{K}^m$ d'une matrice est l'ensemble $\{u \in \mathbb{K}^m / Au = 0_{n,1}\}$ des vecteurs annulés par cette matrice c'est-à-dire que $\sum_{k=1}^n A_{i,k} u_k = 0$ pour tout $1 \leqslant i \leqslant n$.

Nous verrons que ces ensembles sont des *sous-espaces vectoriels*, en particulier, les vecteurs nuls sont dans l'image et le noyau : $0_n \in \mathrm{Im}(A)$, $0_m \in \ker(A)$ et la dimension de l'image est le rang de la matrice.

Exemple. Pour la matrice $A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 3 & 1 \end{pmatrix}$, l'image est le plan $\mathrm{Im}(A) = \{\begin{pmatrix} u+v \\ u-v \\ 3u+v \end{pmatrix} : u, v \in \mathbb{R}\}$ $= \{\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : 4x + 2y - 2z = 0\}$ (on peut vérifier cette forme cartésienne mais elle n'est pas immédiate à trouver) et le noyau est réduit à $\ker(A) = \left\{\begin{pmatrix} 0 \\ 0 \end{pmatrix}\right\}$.

Proposition 1.16. *Les opérations élémentaires sur les lignes préservent le noyau, les opérations élémentaires sur les colonnes préservent l'image d'une matrice.*

Preuve: Soit $E \in GL_n(\mathbb{K})$ une matrice inversible, en particulier obtenue par opération élémentaire sur les lignes de $A \in M_{n,m}(\mathbb{K})$. L'appartenance au noyau de A est équivalente à $u \in \ker(A) \iff Au = 0_m \iff EAu = 0_m \iff u \in \ker(EA)$.

Soit $E \in GL_m(\mathbb{K})$ une matrice inversible, en particulier obtenue par opération élémentaire sur les colonnes de $A \in M_{n,m}(\mathbb{K})$. Si un vecteur est combinaison linéaire $v = \sum_{k=1}^{m} \lambda_k A_{\cdot,k}$ des colonnes de A, $(\mu_k)_{1 \leqslant k \leqslant m} = E^{-1}(\lambda_k)_{1 \leqslant k \leqslant m} \in \mathbb{K}^m$ définit une combinaison linéaire des colonnes de AE telle que $AE \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_m \end{pmatrix} = AEE^{-1} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix} = A \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix} = v$ et $v \in \mathrm{Im}(AE)$. □

Ces notions de noyau et d'image prennent tout leur sens pour l'application linéaire associée. Notons que, si jamais une matrice A est inversible et qu'on connaît son inverse A^{-1}, il est très facile de résoudre le système $AX = B$: en multipliant par A^{-1} des deux côtés, on a $X = A^{-1}B$.

1.4 La méthode de Gauss : forme échelonnée réduite d'une matrice.

La méthode du pivot de Gauss, appliquée à un système, revient à appliquer des opérations élémentaires sur les équations afin de se ramener à un système plus simple, sous forme *échelonnée*, ce qui revient sur les matrices associées à appliquer des opérations élémentaires sur les lignes ; ci-dessous on va définir ce qu'est une matrice échelonnée en lignes, mais on pourrait aussi définir les matrices échelonnées en colonnes (en échangeant le rôle des lignes et des colonnes dans les définitions). Il faudra savoir échelonner une matrice en ligne et en colonnes : on a vu qu'échelonner en lignes préserve le *noyau* de la matrice (ou plutôt, de l'application linéaire associée ; on verra ces notions plus loin dans le cours) tandis qu'échelonner en colonnes préserve l'*image*.

Définition 1.17. Soit $A \in M_{n,m}(\mathbb{K})$. Le *pivot* d'une ligne $L_i = A_{i,\cdot}$ non nulle est son premier terme $A_{i,k}$ non nul, $\forall \ell < k, A_{i,\ell} = 0$.

Une matrice A est *échelonnée*(en lignes) si elle possède les propriétés suivantes :

- Ses lignes nulles sont celles d'indice plus grand qu'un certain nombre r, qu'on appelle *son rang* : pour tout $i > r$, $L_i = 0$;
- La suite des indices des pivots des lignes non nulles est strictement croissante.

Si de plus chaque pivot est égal à 1, et les autres coefficients dans la colonne d'un pivot sont nuls, alors on dit que la matrice est *échelonnée réduite*.

Définition 1.18. Soit $n, m \in \mathbb{N}^*$ et $A, B \in M_{n,m}(\mathbb{K})$. On dit que A et B sont *équivalentes en lignes* si l'on peut passer de A à B en appliquant un nombre fini d'opérations élémentaires sur les lignes.

En termes de produits de matrices, deux matrices A, B sont équivalentes en lignes s'il existe des matrices élémentaires $E_1, \ldots, E_k$ telles que $B = E_1 \ldots E_k A$.

La méthode du pivot de Gauss consiste à appliquer des opérations élémentaires sur les lignes d'une matrice pour se ramener à une matrice échelonnée réduite. Le théorème suivant assure que c'est bien possible.

Théorème 1.19. *Toute matrice est équivalente en lignes à une matrice échelonnée réduite.*

Preuve. On va raisonner par récurrence sur le nombre m de colonnes de la matrice. Si $m = 1$, alors soit la matrice est nulle (donc échelonnée), soit elle a un coefficient non nul. Quitte à permuter deux coefficients (opération élémentaire) on peut assurer que ce coefficient soit sur la première ligne. Quitte à multiplier le premier coefficient par une constante non nulle (encore une opération élémentaire) on peut faire en sorte qu'il vaille 1. Et en remplaçant L_j par $L_j - \alpha_k L_1$ (toujours une opération élémentaire) pour un α_j bien choisi pour tout $j \geqslant 2$, on peut faire en sorte que tous les coefficients à partir de la deuxième ligne soient nuls.

On a donc montré le résultat désiré pour $m = 1$; supposons qu'il soit vrai jusqu'au rang m, et considérons une matrice A avec n lignes et $m + 1$ colonnes. Si la première colonne de A est nulle, on peut regarder la matrice obtenue en oubliant cette colonne et lui appliquer l'hypothèse de récurrence pour la mettre sous forme échelonnée réduite, et donc mettre A sous forme échelonnée réduite.

Sinon, un coefficient sur la première colonne est non nul ; en appliquant les mêmes opérations élémentaires que dans le cas $m = 1$, on peut faire en sorte que ce coefficient vaille 1 et que le premier coefficient de chaque

ligne à partir de la deuxième soit nul. Regardons la sous-matrice de A obtenue en oubliant la première ligne et la première colonne de A : par hypothèse de récurrence on peut la mettre sous forme échelonnée en appliquant des opérations élémentaires sur les lignes. Ceci signifie que, par une suite d'opérations élémentaires sur les lignes, on peut faire en sorte que : la première colonne de A soit $\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}$, et la sous-matrice obtenue B en oubliant la première ligne et la première colonne est échelonnée réduite. Cette matrice est échelonnée mais pas encore échelonnée réduite : dans la colonne d'un pivot de B, le terme situé sur la première ligne de A pourrait être non nul. Si cela se produit, disons que le pivot est sur la j-ième ligne de A, l'opération $L_1 \to L_1 - \alpha_j L_j$, où α_j est le coefficient au-dessus du pivot sur la première ligne de A, appliquée autant de fois qu'il y a de pivots dans B, résoud le problème. On s'est bien ramené à une matrice échelonnée réduite. □

Notons qu'en fait si une matrice est équivalente en ligne à une classe de matrices échelonnées différentes, elle est équivalente à une *unique* matrice échelonnée *réduite*. Nous allons admettre ce fait dans ce cours (en pratique on n'aura pas besoin de cette unicité de la forme échelonnée réduite). Pour cette raison on parlera de *la* forme échelonnée réduite (en lignes) d'une matrice.

Exemple. Dans notre système de départ 1.1, les opérations élémentaires sur les lignes se traduisent en

$$E_{1,2}(-1)E_2(-\frac{1}{2})E_{3,2}(-1)E_{3,1}(-3)E_{2,1}(-1)\begin{pmatrix}1&1\\1&-1\\3&1\end{pmatrix} = \begin{pmatrix}1&0\\0&1\\0&0\end{pmatrix}$$

et comme $E_{1,2}(-1)E_2(-\frac{1}{2})E_{3,2}(-1)E_{3,1}(-3)E_{2,1}(-1)\begin{pmatrix}1\\2\\4\end{pmatrix} = \begin{pmatrix}\frac{3}{2}\\-\frac{1}{2}\\0\end{pmatrix}$, on retrouve la solution.

On voit une matrice échelonnée comme le produit d'une matrice inversible par la matrice à échelonner, $C = EA$. Cette matrice inversible E est construite pas à pas comme produit d'opérations élémentaires. La forme échelonnée C et cette matrice inversible E contiennent toutes les informations dont on a besoin pour analyser A.

Travaillons sur un exemple à appliquer la méthode de Gauss à la matrice

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

pour la mettre sous forme échelonnée en lignes ; on cherche la première colonne avec des coefficients non nuls (ici, la deuxième) et (si nécessaire) on permute deux lignes pour que ce coefficient non nul soit sur la première ligne. Ici, par exemple, on échange L_1 et L_2 pour arriver à

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

Ensuite on utilise le pivot pour faire en sorte que les coefficients sur les lignes en-dessous soient nulles ; ici on remplace L_3 par $L_3 - L_1$:

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

Ensuite, $L_3 - L_2 \to L_3$ nous donne

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

Cette matrice est échelonnée, mais pas encore échelonnée réduite ; pour éliminer les coefficients au-dessus de la diagonale on utilise les opérations $2L_2 + 3L_3 \to L_2$ et $2L_1 + L_3 \to L_1$ pour aboutir à

$$\begin{pmatrix}0&6&0&0\\0&0&1&0\\0&0&0&-2\end{pmatrix}.$$

En multipliant chaque ligne par le bon scalaire, on voit que la forme échelonnée réduite est

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

Définition 1.20. Soit $n, m \in \mathbb{N}^*$ et $A \in M_{n,m}(\mathbb{K})$. Alors sa *transposée* est la matrice $A^t \in M_{m,n}(\mathbb{K})$ dont le coefficient sur la ligne i et la colonne j est égal à $a_{j,i}$.

Par exemple, $\begin{pmatrix} 1 & 2 & 3 \end{pmatrix}^t = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$; $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^t = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$.

Proposition 1.21. *Pour toute matrice A on a $(A^t)^t = A$; et si AB sont deux matrices telles que AB est défini alors B^tA^t est bien défini et $B^tA^t = (AB)^t$.*

On vérifie que la transposée d'une matrice élémentaire est encore une matrice élémentaire ; ainsi, en transposant, on échange les opérations élémentaires sur les lignes appliquées à A en opérations élémentaires sur les colonnes appliquées à A^t. Par conséquent, lorsqu'on a montré que toute matrice était équivalente en lignes à une matrice échelonnée réduite en lignes, on a aussi montré que toute matrice est équivalente en colonnes à une matrice échelonnée réduite en colonnes et ces matrices échelonnées réduites sont égales (*a priori* à une permutation des lignes et des colonnes près). Ceci définit le rang d'une matrice, indépendamment de son échelonnage en lignes ou en colonnes : Le nombre de colonnes non nulles dans la forme échelonnée réduite en ligne est le même que le nombre de lignes non nulles de la forme échelonnée réduite en colonne.

1.5 Une méthode de calcul de l'inverse d'une matrice

Théorème 1.22. *Soit $n \in \mathbb{N}^*$ et $A \in M_n(\mathbb{K})$. Les propriétés suivantes sont équivalentes :*

1. *A est inversible.*
2. *Pour tout vecteur colonne $Y \in M_{n,1}(\mathbb{K})$, l'équation $AX = Y$ a une solution unique dans $M_{n,1}(\mathbb{K})$.*
3. *L'équation $AX = 0_n$ a 0_n pour seule solution dans $M_{n,1}(\mathbb{K})$.*
4. *A est de rang n.*
5. *A est équivalente par lignes à I_n.*
6. *A est un produit de matrices élémentaires de $M_n(\mathbb{K})$.*

Démonstration. On a déjà vu que si A est inversible alors le système $AX = B$ a pour unique solution $X = A^{-1}B$; il est clair que la deuxième condition est plus forte que la troisième. Pour voir que la troisième condition implique la quatrième, utilisons le fait que A est équivalente par lignes à une matrice échelonnée réduite A'. Alors les systèmes $AX = 0$ et $A'X = 0$ ont les mêmes solutions (une opération élémentaire sur les lignes ne change pas les solutions d'un système) ; donc A' est une matrice carrée échelonnée telle que le système $A'X = 0$ a une unique solution, ce qui impose qu'aucune colonne de A' n'est nulle. Comme A' est une matrice carrée, ceci n'est possible que si le rang est n ce qui, pour une matrice échelonnée *réduite* signifie $A' = I_n$.

Supposons ensuite que A est équivalente par lignes à I_n ; en utilisant le lien vu plus haut entre opérations élémentaires sur les lignes et multiplication à droite par une matrice élémentaire, ceci revient à dire qu'il existe des matrices élémentaires $F_1, \ldots, F_k$ dans $M_n(\mathbb{K})$ telles que $A = F_1 \ldots F_k$. Finalement, comme les matrices élémentaires sont inversibles et que tout produit de matrices inversibles est inversible, on déduit de cette propriété que A est inversible. □

Notons une conséquence de ce théorème : On dit que A est *inversible à gauche* s'il existe $B \in M_n(\mathbb{K})$ telle que $BA = I_n$

Proposition 1.23. *Si A est inversible à gauche, alors A est inversible et son inverse à gauche est son inverse ; de même, si elle est inversible à droite.*

Preuve: Soit $A \in M_n(\mathbb{K})$ telle qu'il existe $B \in M_n(\mathbb{K})$ avec $BA = I_n$. Alors $AX = 0_{n,1}$ implique, en composant à gauche par B, $B0_{n,1} = B0_{n,1} = B(AX) = (BA)X = X$, donc le système $AX = 0$ a $X = 0$ pour solution

unique. Par conséquent, le théorème précédent nous dit que A est inversible ; et alors de $BA = I_n$ on déduit, en mutlipliant par A^{-1} à droite, que $B = A^{-1}$.

Si A est inversible à droite, disons $AB = I_n$, alors le même raisonnement appliqué à B nous permet d'abord de conclure que B est inversible puis que $A = B^{-1}$, donc A est inversible et $B = A^{-1}$. □

Le théorème 1.22 a aussi une application pratique : une méthode effective de calcul de l'inverse d'une matrice quand elle existe et d'une matrice caractéristique intéressante si elle n'est pas inversible. Partant d'une matrice $A \in GL_n(\mathbb{K})$ inversible, on agit successivement sur les lignes du produit de matrices $AA^{-1} = I_n$, donnant $F_k AA^{-1} = F_k I_n = F_k$, jusqu'à obtenir $A^{-1} = F_1 F_2 \dots F_k$ quand on a complètement réduit A à l'identité. Pratiquement, considérons la matrice rectangulaire *augmentée* $B \in M_{n,2n}(\mathbb{K})$ définie par $B = (A|I_n)$; autrement dit, les n premières colonnes de B sont celles de A, les n autres colonnes de B sont celles de I_n. Si on applique à B les mêmes opérations élémentaires qu'à A on obtient après réduction une nouvelle matrice $B' = (I_n|A^{-1})$. Si la matrice n'était pas inversible, le processus ne se terminera pas avec la partie gauche de la matrice B transformée en l'identité mais dans une matrice échelonnée réduite de rang plus faible, en particulier si la matrice n'est pas carrée.

Exemple. Appliquons cette technique sur un exemple concret : on cherche à déterminer si la matrice $A = \begin{pmatrix} 1 & 2 & 1 \\ 4 & 0 & -1 \\ -1 & 2 & 2 \end{pmatrix}$ est inversible, et le cas échéant à calculer son inverse. On commence par former la nouvelle matrice

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 4 & 0 & -1 & 0 & 1 & 0 \\ -1 & 2 & 2 & 0 & 0 & 1 \end{array}\right) .$$

Puis on applique l'algorithme de Gauss : on commence par remplacer L_2 par $L_2 - 4L_1$ et L_3 par $L_3 + L_1$ pour aboutir à la matrice

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & -8 & -5 & -4 & 1 & 0 \\ 0 & 4 & 3 & 1 & 0 & 1 \end{array}\right) .$$

Puis remplacer L_3 par $2L_3 + L_2$ nous amène à

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & -8 & -5 & -4 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 & 2 \end{array}\right) .$$

Ensuite on remplace L_2 par $-\frac{1}{8}L_2 + \frac{5}{8}L_3$:

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & \frac{7}{4} & -\frac{3}{4} & -\frac{5}{4} \\ 0 & 0 & 1 & -2 & 1 & 2 \end{array}\right) .$$

Pour finir, remplacer L_1 par $L_1 - 2L_2 - L_3$ nous donne la matrice

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{7}{4} & -\frac{3}{4} & -\frac{5}{4} \\ 0 & 0 & 1 & -2 & 1 & 2 \end{array}\right) .$$

On conclut que A est inversible, et que

$$A^{-1} = \begin{pmatrix} -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{7}{4} & -\frac{3}{4} & -\frac{5}{4} \\ -2 & 1 & 2 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} -2 & 2 & 2 \\ 7 & -3 & -5 \\ -8 & 4 & 8 \end{pmatrix} .$$

Remarquons que, si on appliquait la même méthode mais avec des opérations élémentaires sur les colonnes, on arriverait aussi à calculer l'inverse de A : en effet, on aurait ainsi trouvé des matrices élémentaires $G_1, \dots, G_\ell$ telles que $AG_1 \dots G_\ell = I_n$, par conséquent $G_1 \dots G_\ell = A^{-1}$. Par contre, si on appliquait cette méthode en alternant opérations élémentaires sur les lignes et sur les colonnes, on obtiendrait des matrices élémentaires $F_1, \dots, F_k$ et $G_1, \dots, G_\ell$ telles que $F_1 \dots F_k A G_1 \dots G_\ell = I_n$, mais il n'y a aucune raison que $F_1 \dots F_k G_1 \dots G_\ell$

soit l'inverse de A ! Moralité : on peut appliquer cette méthode en utilisant des opérations élémentaires sur les lignes ou sur les colonnes, mais au cours d'un même calcul il ne faut pas alterner entre lignes et colonnes - si on commence par des opérations élémentaires sur des lignes, par exemple, tout le calcul doit être mené en utilisant des opérations sur les lignes.

Exemple. Sur le cas $\begin{pmatrix}1 & 1\\ 1 & -1\\ 3 & 1\end{pmatrix}$ en l'augmentant d'une matrice identité 3x3 à droite, on obtient, après échelonnement et réduction,

$$\left(\begin{array}{cc|ccc}1 & 1 & 1 & 0 & 0\\ 1 & -1 & 0 & 1 & 0\\ 3 & 1 & 0 & 0 & 1\end{array}\right) \simeq \left(\begin{array}{cc|ccc}1 & 1 & 1 & 0 & 0\\ 0 & -2 & -1 & 1 & 0\\ 0 & -2 & -3 & 0 & 1\end{array}\right) \simeq \left(\begin{array}{cc|ccc}1 & 0 & \frac{1}{2} & \frac{1}{2} & 0\\ 0 & 1 & \frac{1}{2} & -\frac{1}{2} & 0\\ 0 & 0 & -2 & -1 & 1\end{array}\right)$$

c'est-à-dire que $\begin{pmatrix}1 & 1 & 0\\ 1 & -1 & 0\\ -4 & -2 & 2\end{pmatrix}\begin{pmatrix}1 & 1\\ 1 & -1\\ 3 & 1\end{pmatrix} = 2\begin{pmatrix}1 & 0\\ 0 & 1\\ 0 & 0\end{pmatrix}$ ce qui permet d'affirmer que le rang de la matrice est 2, que son noyau est donc réduit à $\{0_{2,1}\}$ et qu'une inverse à gauche est $\frac{1}{2}\begin{pmatrix}1 & 1 & 0\\ 1 & -1 & 0\end{pmatrix}$. Ceci nous apprend également que le système $\begin{cases}x+y=b_1\\ x-y=b_2\\ 3x+y=b_3\end{cases}$ aura comme solution $\begin{cases}x=(b_1+b_2)/2,\\ y=(b_1-b_2)/2,\\ \text{ssi } -4b_1-2b_2+2b_3=0.\end{cases}$

En échelonnant sur les colonnes, on obtient

$$\left(\begin{array}{cc}1 & 1\\ 1 & -1\\ 3 & 1\\ \hline 1 & 0\\ 0 & 1\end{array}\right) \simeq \left(\begin{array}{cc}1 & 0\\ 1 & -2\\ 3 & -2\\ \hline 1 & -1\\ 0 & 1\end{array}\right) \simeq \left(\begin{array}{cc}1 & 0\\ 0 & 1\\ 2 & 1\\ \hline \frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & -\frac{1}{2}\end{array}\right) \text{ c'est-à-dire } \begin{pmatrix}1 & 1\\ 1 & -1\\ 3 & 1\end{pmatrix}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix} = 2\begin{pmatrix}1 & 0\\ 0 & 1\\ 2 & 1\end{pmatrix},$$

ce qui signifie que l'image de cette matrice est le plan linéaire généré par les deux vecteurs $\begin{pmatrix}1\\ 0\\ 2\end{pmatrix}$ et $\begin{pmatrix}0\\ 1\\ 1\end{pmatrix}$ qui sont liés aux deux colonnes de la matrice.

En constatant que $\begin{pmatrix}2 & 1 & -1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & 1\\ 2 & 1\end{pmatrix} = \begin{pmatrix}2 & 1 & -1\end{pmatrix}\begin{pmatrix}1 & 1\\ 1 & -1\\ 3 & 1\end{pmatrix}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}/2 = \begin{pmatrix}0 & 0\end{pmatrix}$ (on verra plus tard comment trouver un tel vecteur par le produit vectoriel et ses généralisations), on obtient l'équation cartésienne de ce plan : $2x+y-z=0$.

1.6 Forme échelonnée : ce qu'elle nous apprend

Nous venons de voir que mettre une matrice carrée inversible sous forme échelonnée permet d'en calculer l'inverse. Mais, quand la matrice n'est pas inversible, en particulier quand elle n'est pas carrée, on obtient beaucoup d'informations en calculant sa forme échelonnée.

Mettre une matrice sous une forme échelonnée réduite par lignes, signifie trouver une matrice E inversible, construite comme produit de transformations élémentaires, telle que EA soit échelonnée réduite.

Tout l'intérêt d'une forme échelonnée réduite est que son noyau et son image sont simples à déterminer : Comme les dernières lignes sont nulles, $\mathrm{Im}(EA)$ est clairement engendrée par les colonnes contenant un pivot. Pour $A \in M_{pq}(\mathbb{K})$, nommons $C_A \subset \{1,\dots,q\}$ l'ensemble des indices des colonnes des pivots, alors tout vecteur $v \in \mathrm{Im}(EA) \in \mathbb{K}^p$ s'écrit comme l'image $v = EA\begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_q\end{pmatrix}$ d'un vecteur $\begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_q\end{pmatrix} \in \mathbb{K}^q$ dont les seules valeurs non nulles sont associées aux colonnes des pivots, $\lambda_j \neq 0 \Rightarrow j \in C_A$. Mais tout vecteur $w \in \mathrm{Im}(A)$ peut s'écrire

$w = E^{-1}v$ pour un $Ew = v \in \mathrm{Im}(EA)$, donc $w = A\begin{pmatrix}\lambda_1\\ \vdots \\ \lambda_q\end{pmatrix}$ est une combinaison linéaire unique des colonnes de A associées aux pivots de EA : $\mathrm{Im}(A) = \mathrm{Vect}_{\mathbb{K}}\left((A_{ij})_{1\leqslant i\leqslant p}/j \in C_A\right)$.

Et les colonnes de EA qui ne sont pas des pivots sont associées à des éléments du noyau de EA, en effet, pour $j \notin C_A$, par définition des pivots, les seuls coefficients non nuls de cette j-ème colonne de $EA = R$, qui est $(R_{ij})_{1\leqslant i\leqslant p}$, sont sur les lignes des pivots d'indice inférieur à j. Mais alors, cette colonne est combinaison linéaire de ces colonnes pivots et nous donne un élément du noyau de EA, donc du noyau de A car, pour $u \in \mathbb{K}^q$, $EAu = 0_p \iff Au = 0_p$ en composant par E^{-1}.

Par exemple la forme échelonnée de $A = \begin{pmatrix}1&2&2&-3&2&3\\2&4&1&0&-5&-6\\4&8&5&-6&-1&0\\5&10&7&-9&1&3\\-4&-8&-5&6&1&0\end{pmatrix}$ est $R = \begin{pmatrix}1&2&0&1&-4&-5\\0&0&1&-2&3&4\\0&0&0&0&0&0\\0&0&0&0&0&0\\0&0&0&0&0&0\end{pmatrix}$ ce qui nous permet de voir que $\mathrm{Im}(A) = \mathrm{Vect}\left\{\begin{pmatrix}1\\2\\4\\5\\-4\end{pmatrix}, \begin{pmatrix}2\\1\\5\\7\\-5\end{pmatrix}\right\}$ les première et troisième colonnes de A, indices des colonnes des pivots de R. De plus, les autres colonnes permettent de calculer le noyau $\ker(A) = \mathrm{Vect}\left\{\begin{pmatrix}2\\-1\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\-2\\-1\\0\\0\end{pmatrix}, \begin{pmatrix}-4\\3\\0\\0\\-1\\0\end{pmatrix}, \begin{pmatrix}-5\\4\\0\\0\\0\\-1\end{pmatrix}\right\}$. Remarquez que le rang de A, c'est-à-dire la dimension de l'image est 2, la dimension du noyau est 4, ce qui fait que $\dim\ker A + \mathrm{rang}(A) = 2 + 4 = 6$, la dimension de l'espace de départ, c'est-à-dire le nombre de colonnes de A.

De même, en échelonnant sur les colonnes, on obtient une matrice inversible E telle que AE est échelonnée réduite en colonnes. Donc les vecteurs correspondant aux dernières colonnes nulles de AE forment clairement le noyau de AE, mais comme E est inversible, cela signifie que ce sont les dernières colonnes de E qui forment le noyau de A. De plus, l'image de A est égale à l'image de AE, qui est clairement engendrée par les colonnes non nulles de AE.

Par exemple, $AE = R$ avec $A = \begin{pmatrix}1&-2&1&3&1\\1&1&1&-3&-2\\5&-1&5&-3&-4\\0&0&3&0&0\end{pmatrix}$, $E = \frac{1}{3}\begin{pmatrix}1&2&-1&1&1\\-1&1&0&2&1\\0&0&1&0&0\\0&0&0&1&0\\0&0&0&0&1\end{pmatrix}$ et $R = \begin{pmatrix}1&0&0&0&0\\0&1&0&0&0\\2&3&0&0&0\\0&0&1&0&0\end{pmatrix}$ nous indique que les première colonnes de R génèrent l'image, $\mathrm{Im}(A) = \mathrm{Vect}\left\{\begin{pmatrix}1\\0\\2\\0\end{pmatrix}, \begin{pmatrix}0\\1\\3\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\end{pmatrix}\right\}$ et que les dernières colonnes de E génèrent le noyau $\ker(A) = \mathrm{Vect}\left\{\begin{pmatrix}1\\2\\0\\1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\0\\0\\1\end{pmatrix}\right\}$. Encore une fois, le rang est 3, complété par la dimension du noyau, 2, donne la dimension de l'espace de départ, le nombre de colonnes de A : 5.

1.7 Un peu de vocabulaire sur les matrices carrées

Définition 1.24. $A = (a_{i,j}) \in M_n(\mathbb{K})$ est *triangulaire supérieure* si $a_{i,j} = 0$ pour tout $i > j$ (graphiquement : tous les coefficients en-dessous de la diagonale sont nuls). Elle est *triangulaire inférieure* si $a_{i,j} = 0$ pour tout $j > i$.

Un intérêt de ces matrices : les systèmes dont la matrice associée est triangulaire sont très faciles à résoudre !

Définition 1.25. $A \in M_n(\mathbb{K})$ est *diagonale* si elle est à la fois triangulaire supérieure et inférieure ; autrement dit, si $a_{i,j} = 0$ dès que $i \neq j$.

Définition 1.26. A est *symétrique* si $A = A^t$, c'est-à-dire $a_{i,j} = a_{j,i}$ pour tout i, j ; elle est *antisymétrique* si $A = -A^t$, c'est-à-dire $a_{i,j} = -a_{j,i}$ pour tout i, j (en particulier, les coefficients sur la diagonale sont nuls !)

Attention à ne pas confondre une matrice symétrique de la matrice d'une symétrie (voir 3.41).

1.8 Résolution matricielle d'un système linéaire

Pour conclure ce chapitre, formulons en termes matriciels la méthode du pivot de Gauss pour résoudre un unique système d'équations linéaires, sans pour autant inverser la matrice associée.

Définition 1.27. Soit $\mathcal{S}$ le système d'équations linéaires suivant :

$$\begin{cases} a_{1,1}x_1 + \ldots + a_{1,m}x_m &= y_1, \\ a_{2,1}x_1 + \ldots + a_{2,m}x_m &= y_2, \\ \vdots & \vdots \\ a_{n,1}x_1 + \ldots + a_{n,m}x_m &= y_n. \end{cases}$$

On peut écrire $\mathcal{S}$ sous la forme $AX = Y$. La *matrice augmentée* est la matrice obtenue en ajoutant la colonne B tout à droite des colonnes de A.

Par exemple, considérons le système suivant :

$$\begin{cases} x_1 + x_2 + 7x_3 = -1 \\ 2x_1 - x_2 + 5x_3 = -5 \\ -x_1 - 3x_2 - 9x_3 = -5 \end{cases}$$

Sa matrice augmentée est

$$\begin{pmatrix} 1 & 1 & 7 & -1 \\ 2 & -1 & 5 & -5 \\ -1 & -3 & -9 & -5 \end{pmatrix}.$$

La méthode du pivot de Gauss consiste à mettre la matrice augmentée sous forme échelonnée réduite ; chaque colonne de la matrice (sauf la dernière !) correspond à une des inconnues du système, et dans la forme échelonnée réduite de la matrice il y a deux types de colonnes : celles qui contiennent un pivot, et les autres. Chaque inconnue dont la colonne ne contient pas un pivot (peut-être qu'il n'y en a pas !) est une inconnue *libre* : on lui donne un nom ; ensuite, on exprime les inconnues dont la colonne contient un pivot en fonction des inconnues libres (ou on obtient des formules contradictoires si le système n'a pas de solution), et cela nous donne une description des solutions du système (s'il y en a !).

Revenons à l'exemple du système. En faisant les opérations $L_2 \to L_2 - 2L_1$ et $L_3 \to L_3 + L_1$ sur la matrice augmentée, on arrive à

$$\begin{pmatrix} 1 & 1 & 7 & -1 \\ 0 & -3 & -9 & -3 \\ 0 & -2 & -2 & -6 \end{pmatrix}$$

Puis on peut par exemple faire $L_2 \to -\frac{1}{3}L_2$ et $L_3 \to L_3 + 2L_2$ pour arriver à

$$\begin{pmatrix} 1 & 1 & 7 & -1 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 4 & -4 \end{pmatrix}.$$

Ensuite, les opérations $L_3 \to \frac{1}{4}L_3$, puis $L_1 \to L_1 - 7L_3$ et $L_2 \to L_2 - 3L_3$ nous amènent à

$$\begin{pmatrix} 1 & 1 & 0 & 6 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & -1 \end{pmatrix}.$$

Une dernière opération pour se ramener à une matrice échelonnée réduite : après $L_1 \to L_1 - L_2$ on aboutit à la matrice

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

Finalement, notre système de départ est équivalent au système suivant (dans ce cas, il n'y a pas d'inconnues libres) :

$$\begin{cases} x_1 = 2 \\ x_2 = 4 \\ x_3 = -1 \end{cases}$$

En pratique, on n'a pas toujours besoin d'aller jusqu'à obtenir une forme échelonnée réduite : on échelonne la matrice jusqu'à aboutir à une matrice pour laquelle le système est facile à résoudre.

En regardant les systèmes associés aux matrices échelonnées réduites (cas auquel on peut toujours se ramener, comme on l'a vu plus haut), on voit que trois situations peuvent se produire :

1. Le système a exactement une solution (cas où la matrice est inversible),
2. Le système n'a pas de solutions (une ligne nulle dans la matrice échelonnée réduite alors que le b_i correspondant n'est pas nul),
3. Le système a une infinité de solutions.

Chapitre 2

Espaces vectoriels

2.1 Corps commutatifs

Formellement, la définition d'un espace vectoriel nécessite de préciser au-dessus de quel *corps* l'espace vectoriel est défini. Précisons la définition d'un corps.

Définition 2.1. Un corps $(\mathbb{K}, +, \times)$ est la donnée d'un ensemble $\mathbb{K}$, et de deux opérations $+, \times$ avec les propriétés suivantes :

1. Pour tout $x, y, z \in \mathbb{K}$ on a $x + (y + z) = (x + y) + z$ (associativité), $x + y = y + x$ (commutativité), il existe un élément $0 \in K$ *neutre* pour l'addition : tel que $0 + x = x + 0 = x$ pour tout $x \in K$, et pour tout $x \in \mathbb{K}$ il existe $y \in \mathbb{K}$ tel que $x + y = y + x = 0$, un tel y est nécessairement unique, on le note $y = -x$, *l'opposé* de x.
2. Pour tout $x, y, z \in \mathbb{K}$ on a $x \times (y \times z) = (x \times y) \times z$ (associativité), $x \times y = y \times x$ (commutativité), il existe un élément $1 \in K \setminus \{0\}$ *neutre* pour la multiplication : tel que $1 \times x = x \times 1 = x$ pour tout $x \in K$, et pour tout $x \in \mathbb{K} \setminus \{0\}$ il existe $y \in \mathbb{K} \setminus \{0\}$ tel que $x \times y = y \times x = 1$, un tel y est nécessairement unique, on le note $y = x^{-1}$, *l'inverse* de x.
3. Pour tout $x, y, z \in \mathbb{K}$ on a $x \times (y + z) = x \times y + x \times z$ et $(x + y) \times z = x \times z + y \times z$ (distributivité de la multiplication sur l'addition).

Dans ce cours, les seuls corps que nous allons rencontrer sont $\mathbb{Q}$, $\mathbb{R}$ et $\mathbb{C}$ munis de leurs opérations usuelles ; à chaque fois qu'on écrira « soit $\mathbb{K}$ un corps », on aura en tête un de ces trois exemples.

2.2 Espaces vectoriels ; combinaisons linéaires

Définition 2.2. Soit $\mathbb{K}$ un corps. Un $\mathbb{K}$-*espace vectoriel* $(E, +, \cdot)$ est un ensemble non vide E muni d'une loi d'addition $+ : E \times E \to E$ et d'une loi externe (multiplication d'un élément de E par un élément de $\mathbb{K}$) $\cdot : \mathbb{K} \times E \to E$ avec les propriétés suivantes :

1. L'addition est *commutative* (pour tout $u, v \in E,\ u + v = v + u$) et *associative* (pour tout $u, v, w \in E$, $u + (v + w) = (u + v) + w$).
2. L'addition admet un *élément neutre*, le *vecteur nul*, noté 0_E, qui est tel que pour tout $u \in E$ on a $u + 0_E = 0_E + u = u$; et pour tout u il existe v tel que $u + v = 0_E$ (notons qu'un tel v est unique et qu'on a aussi $v + u = 0_E$; on note $v = -u$).
3. L'addition et la multiplication sont compatibles, c'est-à-dire que pour tout $\lambda, \mu \in \mathbb{K}$ et tout $u, v \in E$ on a :
   (a) $\lambda \cdot (u + v) = \lambda \cdot u + \lambda \cdot v$,
   (b) $(\lambda +_{\mathbb{K}} \mu) \cdot u = \lambda \cdot u +_E \mu \cdot u$,
   (c) $\lambda \cdot (\mu \cdot u) = (\lambda \times \mu) \cdot u$,
   (d) $0_{\mathbb{K}} \cdot u = 0_E$,
   (e) $1_{\mathbb{K}} \cdot u = u$.

On appellera les éléments de E des *vecteurs* et les éléments de $\mathbb{K}$ des *scalaires*. On n'indicera les nombres zéro et un, le vecteur nul ou les opérations seulement s'il y a possibilité de confusion et on omettra souvent les opérateurs de multiplication $\cdot$ et $\times$.

Ce qu'il faut vraiment retenir : on peut ajouter des vecteurs entre eux, et multiplier les vecteurs par des constantes ; et les règles de calcul ci-dessus ont les conséquences attendues, par exemple $0_{\mathbb{K}} \cdot u = 0_E$ et $(-1) \cdot u = -u$ pour tout $u \in E$. Vérifions ces deux propriétés : pour tout u on a $0_{\mathbb{K}} \cdot u = (0_{\mathbb{K}} + 0_{\mathbb{K}}) \cdot u = 0_{\mathbb{K}} \cdot u + 0_{\mathbb{K}} \cdot u$. En soustrayant $0_{\mathbb{K}} \cdot u$ on obtient $0_E = 0_{\mathbb{K}} \cdot u$. On a aussi $u + (-1) \cdot u = (1 + -1) \cdot u = 0_{\mathbb{K}} \cdot u = 0_E$, ce qui montre que $(-1) \cdot u = -u$.

La définition d'un espace vectoriel peut paraître très lourde au premier abord ; mais elle regroupe des propriétés partagées par beaucoup de structures mathématiques importantes, que vous avez déjà rencontrées. Par exemple, sont des espaces vectoriels :

- L'espace $\mathbb{K}^n$ pour tout $n \in \mathbb{N}^*$, avec ses opérations usuelles. C'est l'exemple fondamental d'espace vectoriel, sur lequel les propriétés des espaces vectoriels sont modelées.
- L'espace des polynômes à coefficients dans $\mathbb{K}$, noté $\mathbb{K}[X]$, ainsi que l'espace des polynômes de degré $\leqslant n$ $\mathbb{K}_n[X]$.
- L'espace des fonctions d'un ensemble X à valeurs dans $\mathbb{K}$, avec l'addition et la multiplication usuelles des fonctions : $(f+g)(x) = f(x) + g(x)$, $(\lambda f)(x) = \lambda f(x)$.
- Pour tout $n, m \in \mathbb{N}^*$, $M_{n,m}(\mathbb{K})$, muni des opérations vues au chapitre précédent.

Définition 2.3. Soit E un $\mathbb{K}$-espace vectoriel, $n \in \mathbb{N}^*$ et $u_1, \ldots, u_n, v$ des vecteurs de E.

1. S'il existe $\lambda \in \mathbb{K}$ tel que $v = \lambda u_1$ on dit que v est *colinéaire* à u_1.
2. S'il existe $(\lambda_1, \ldots, \lambda_n) \in \mathbb{K}^n$ tels que
$$v = \sum_{i=1}^{n} \lambda_i u_i$$
on dit que v est une *combinaison linéaire* de $u_1, \ldots, u_n$.

Notons qu'avec ces définitions, 0_E est colinéaire à tout vecteur de E ; et si u, v sont non nuls et v est colinéaire à u, alors $v = \lambda u$ se récrit sous la forme $u = \frac{1}{\lambda} v$, et u est colinéaire à v. Souvent, on dira que u et v sont colinéaires en traitant à part le cas où l'un est nul.

Définition 2.4. Soit $u_1, \ldots, u_n$ des vecteurs de E. On dit que la famille $(u_1, \ldots, u_n)$ est *libre* si pour tout $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ on a l'implication
$$\lambda_1 u_1 + \ldots + \lambda_n u_n = 0_E \Rightarrow \forall i, \lambda_i = 0_{\mathbb{K}} \ .$$

Quand une famille de vecteurs n'est pas libre on dit qu'elle est *liée*, c'est-à-dire qu'il existe une combinaison linéaire de coefficients non tous nuls qui se somme au vecteur nul :
$$(u_i) \in E^n \text{ liée } \iff \exists (\lambda_i) \in \mathbb{K}^n, \lambda_1 u_1 + \ldots + \lambda_n u_n = 0_E \text{ et } \exists i, 1 \leqslant i \leqslant n, \lambda_i \neq 0_{\mathbb{K}}.$$

Proposition 2.5. *La famille $(u_1, \ldots, u_n)$ est libre si, et seulement si, pour tout $i \in \{1, \ldots, n\}$, le vecteur u_i ne peut pas être écrit comme une combinaison linéaire des vecteurs de la famille $(u_k)_{k \neq i}$.*

Démonstration. Si jamais la famille n'est pas libre, c'est qu'on a des scalaires $\lambda_1, \ldots, \lambda_n$ non tous nuls mais tels que $\lambda_1 u_1 + \ldots + \lambda_n u_n = 0_E$; par hypothèse on peut trouver i tel que λ_i est non nul et on a
$$\lambda_i u_i = -\sum_{\substack{k=1 \\ k \neq i}}^{n} \lambda_k u_k \quad \text{donc} \quad u_i = -\sum_{\substack{k=1 \\ k \neq i}}^{n} \frac{\lambda_k}{\lambda_i} u_k \ .$$
Réciproquement, si pour un certain i on peut écrire $u_i = \sum_{\substack{k=1 \\ k \neq i}}^{n} \lambda_k u_k$ alors en posant $\lambda_i = -1$ on a $\sum_{k=1}^{n} \lambda_k u_k = 0_E$ et les λ_k ne sont pas tous nuls. □

2.3 Sous-espaces vectoriels

Souvent, les structures qu'on considère sont contenues dans un espace dont on sait déjà que c'est un espace vectoriel ($\mathbb{R}^n$, un espace de fonctions, de polynômes, de suites...) et on voudrait utiliser les lois de l'espace vectoriel ambiant pour en faire un espace vectoriel. Ceci nous amène à la notion suivante, qui est primordiale.

Définition 2.6. Soit $\mathbb{K}$ un corps, E un $\mathbb{K}$-espace vectoriel et F un sous-ensemble de E. On dit que F est un *sous-espace vectoriel* si :

1. F est non vide.
2. Pour tout $u, v \in F$, $u + v \in F$.
3. Pour tout $u \in F$ et tout $\lambda \in \mathbb{K}$, $\lambda u \in F$.

Il est important de vérifier que F est non vide ! souvent, on le fait en vérifiant que $0_E \in F$; de toute façon si F est non vide il doit contenir 0_E, puisque si l'on prend un u quelconque dans F (ce qui est possible dès que F est non vide) et qu'on le multiplie par 0 on obtient 0_E, qui doit donc être dans F puisque F est stable par multiplication par un scalaire. Par ailleurs, la deuxième et la troisième condition pourraient être regroupées en une seule condition :

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

C'est-à-dire : une combinaison linéaire d'éléments de F est encore un élément de F ou bien encore que, pour deux vecteurs de F, le plan vectoriel engendré

On écrira souvent simplement 0 au lieu de 0_E pour désigner le vecteur nul.

Exemple.

1. Dans $\mathbb{R}^2$, l'ensemble $\{(x, y) \colon ax + by = 0\}$ est un sous-espace vectoriel pour tout $a, b \in \mathbb{R}$; si jamais $a = b = 0$ c'est $\mathbb{R}^2$ tout entier, sinon c'est la droite passant par 0 et de vecteur directeur $(-b, a)$, c'est-à-dire orthogonale au vecteur (a, b).
2. Dans $\mathbb{R}^3$, l'ensemble $\{(x, y, z) \colon ax + by + cz = 0\}$ est un sous-espace vectoriel pour tout $a, b, c \in \mathbb{R}^3$; s'ils sont tous les trois nuls c'est $\mathbb{R}^3$ tout entier, sinon c'est le plan contenant 0 et orthogonal au vecteur (a, b, c) (notion qui n'a de sens que dans le cas d'une base orthogonale).
3. Plus généralement, vérifions que pour tout corps $\mathbb{K}$, et toute matrice $A \in M_{n,m}(\mathbb{K})$, l'ensemble $\ker(A) = \{X \in \mathbb{K}^m \colon AX = 0_n\}$ est un sous-espace vectoriel de $\mathbb{K}^m$. Puisque $A0_m = 0_n$, $0_m \in \ker(A)$; et si $X, Y \in \ker(A)$ et $\lambda \in \mathbb{K}$ alors on a

   $$A(\lambda X + Y) = \lambda AX + AY = 0_n + 0_n = 0_n\ .$$

   Notons que l'ensemble $\ker(A)$ décrit ci-dessus pourrait de manière équivalente être vu comme l'ensemble de toutes les solutions $(x_1, \ldots, x_m)$ d'un système linéaire homogène à n lignes et m inconnues ; et en fait on verra dans le chapitre sur les applications linéaires que tout sous-espace vectoriel de $\mathbb{K}^m$ peut-être mis sous cette forme.
4. Dans l'espace vectoriel formé par toutes les suites de nombres réels, noté $\mathbb{R}^{\mathbb{N}}$, l'espace des suites convergentes est un sous-espace vectoriel : la suite nulle est convergente, et une combinaison linéaire de suites convergentes est encore une suite convergente. L'espace formé par les suites qui convergent vers 0 est-il un sous-espace vectoriel de $\mathbb{R}^{\mathbb{N}}$? Et celui formé par les suites qui tendent vers 1 ?
5. Toujours dans $\mathbb{R}^{\mathbb{N}}$, l'ensemble des suites satisfaisant une relation de *récurrence linéaire d'ordre deux* (par exemple) est un sous-espace vectoriel. Plus explicitement : fixons $a, b \in \mathbb{R}$ et considérons l'espace formé E par toutes les suites (u_n) telles que

   $$\forall n \in \mathbb{N}, \quad u_{n+2} = au_{n+1} + bu_n\ .$$

   Alors la suite nulle appartient bien à E ; et si u, v sont deux éléments de E, et $\lambda \in \mathbb{R}$, vérifions que $w = \lambda u + v \in E$: pour tout $n \in \mathbb{N}$, on a

   $$\begin{aligned} w_{n+2} &= \lambda u_{n+2} + v_{n+2} \\ &= \lambda(au_{n+1} + bu_n) + (av_{n+1} + bv_n) \\ &= (\lambda a u_{n+1} + bv_{n+1}) + (\lambda a u_n + bv_n) \\ &= w_{n+1} + w_n. \end{aligned}$$

   Ceci montre bien qu'une combinaison linéaire d'éléments de E reste un élément de E, donc que E est un sous-espace vectoriel de $\mathbb{R}^{\mathbb{N}}$. La même preuve montre que les suites à coefficients dans $\mathbb{K}$ satisfaisant une relation de récurrence linéaire d'ordre 2 forme un sous-espace vectoriel de $\mathbb{K}^{\mathbb{N}}$ pour tout corps $\mathbb{K}$.

Proposition 2.7. *Soit $\mathbb{K}$ un corps, E un espace vectoriel et $(F_i)_{i \in I}$ une famille de sous-espaces vectoriels de E. Alors $F = \bigcap_i F_i$ est un sous-espace vectoriel de E.*

Démonstration. Puisque $0 \in F_i$ pour tout $i \in I$, on voit que $0 \in \bigcap_i F_i = F$. Si maintenant $u, v \in F$ et $\lambda \in \mathbb{K}$, alors pour tout $i \in I$ on a $u, v \in F_i$ donc $\lambda u + v \in F_i$ puisque F_i est un sous-espace vectoriel de E ; donc $\lambda u + v \in \bigcap F_i = F$, ce qui montre que F est bien un sous-espace vectoriel. □

Définition 2.8. Soit $\mathbb{K}$ un corps, E un $\mathbb{K}$-espace vectoriel et $x_1, \dots, x_m$ des vecteurs de E. Le *sous-espace vectoriel engendré par* $x_1, \dots, x_m$ est l'intersection de tous les sous-espaces vectoriels de E qui contiennent $x_1, \dots, x_m$. On le note $\mathrm{Vect}_{\mathbb{K}}(x_1, \dots, x_m)$. On étend la notion à $\mathrm{Vect}(P)$ engendré par $P \subset E$, une partie de E.

Quand $\mathrm{Vect}_{\mathbb{K}}(x_1, \dots, x_m) = E$ on dit que $(x_1, \dots, x_m)$ est une *famille génératrice* de E, c'est-à-dire que tout vecteur de E peut s'écrire comme une combinaison linéaire des $(x_i)_{1 \leqslant i \leqslant m}$.

Autrement dit : le sous-espace vectoriel engendré par $x_1, \dots, x_m$ est le plus petit sous-espace vectoriel de E qui contient tous les vecteurs $x_1, \dots, x_m$. Par exemple, si $x \neq 0$, le sous-espace vectoriel engendré par x est la droite passant par 0 et de vecteur directeur x.

La description ci-dessus de l'espace vectoriel engendré par $x_1, \dots, x_m$ est souvent trop abstraite pour être utilisable en pratique. On utilisera la caractérisation suivante (qu'on pourrait prendre comme définition alternative).

Proposition 2.9. *Soit $\mathbb{K}$ un corps, E un $\mathbb{K}$-espace vectoriel et $x_1, \dots, x_m$ des vecteurs de E. Alors on a*

$$\mathrm{Vect}_{\mathbb{K}}(x_1, \dots, x_m) = \{u \in E \colon u \text{ est combinaison linéaire de } x_1, \dots, x_m\} = \left\{ \sum_{i=1}^{m} \lambda_i x_i \colon (\lambda_i)_{1 \leqslant i \leqslant m} \in \mathbb{K}^m \right\} .$$

Démonstration. Si F est un sous-espace vectoriel contenant $x_1, \dots, x_m$ alors F doit contenir toutes leurs combinaisons linéaires, ce qui montre l'inclusion de droite à gauche dans l'égalité ci-dessus. Pour montrer l'inclusion de gauche à droite, il nous suffit de vérifier que

$$F = \{u \in E \colon u \text{ est combinaison linéaire de } x_1, \dots, x_m\}$$

est un sous-espace vectoriel de E (il contient bien sûr $x_1, \dots, x_m$). Pour cela, notons tout d'abord que $0_E = 0_{\mathbb{K}} x_1 + \dots + 0_{\mathbb{K}} x_m \in F$.

Ensuite, si $u = \lambda_1 x_1 + \dots + \lambda_m x_m \in F$ et $\lambda \in \mathbb{K}$, alors $\lambda u = \lambda\lambda_1 x_1 + \dots + \lambda\lambda_m x_m \in F$.

Enfin, si $u = \lambda_1 x_1 + \dots + \lambda_m x_m \in F$ et $v = \mu_1 x_1 + \dots + \mu_m x_m \in F$ alors $u + v = (\lambda_1 + \mu_1) x_1 + \dots + (\lambda_m + \mu_m) x_m \in F$.

On a vérifié tous les points de la définition d'un sous-espace vectoriel, ce qui achève la démonstration. □

Exemple. L'exemple fondamental est donné par $x_1, \dots, x_m \in \mathbb{K}^n$, une famille de m vecteurs colonnes de taille n. En construisant la matrice $A \in M_{n,m}$ obtenue en rangeant ces m vecteurs en colonne, on a

$$\mathrm{Vect}_{\mathbb{K}}(x_1, \dots, x_m) = \mathrm{Im}(A) \subset \mathbb{K}^n.$$

Proposition 2.10. *Soit $\mathbb{K}$ un corps, E un $\mathbb{K}$-espace vectoriel et $(x_1, \dots, x_m)$ une famille* libre *d'éléments de E. Alors $y \in \mathrm{Vect}_{\mathbb{K}}(x_1, \dots, x_m)$ si, et seulement si, la famille $(x_1, \dots, x_m, y)$ est liée.*

On se sert souvent de cette proposition sous la forme négative suivante : si $(x_1, \dots, x_m)$ est libre, alors $y \notin \mathrm{Vect}(x_1, \dots, x_m)$ est équivalent à $(x_1, \dots, x_m, y)$ est libre.

Démonstration. Soit $(x_1, \dots, x_m)$ une famille libre, et $y \in E$. Si $y \in \mathrm{Vect}(x_1, \dots, x_m)$, alors on a l'existence de $\lambda_1, \dots, \lambda_m \in \mathbb{K}$ tels que $\lambda_1 x_1 + \dots + \lambda_m x_m = y$, donc la combinaison linéaire $y - \lambda_1 x_1 - \dots - \lambda_m x_m = 0$ alors que ses coefficients ne sont pas tous nuls (le coefficient de y vaut 1), ce qui montre que $(x_1, \dots, x_m, y)$ est liée.

Réciproquement, supposons que $(x_1, \dots, x_m, y)$ soit liée ; alors on a $\lambda_1, \dots, \lambda_m$ et μ non tous nuls tels que $\lambda_1 x_1 + \dots + \lambda_m x_m + \mu y = 0$. Si μ n'est pas nul alors on peut écrire $y = -\frac{\lambda_1}{\mu} x_1 - \dots \frac{\lambda_m}{\mu} x_m$, par conséquent $y \in \mathrm{Vect}(x_1, \dots, x_m)$. Si $\mu = 0$ alors on a $\lambda_1 x_1 + \dots + \lambda_m x_m = 0$, donc, comme la famiile $(x_i)_{1 \leqslant i \leqslant m}$ est libre, $\lambda_1 = \dots = \lambda_m = 0$. Par conséquent μ ne peut pas être nul et $y \in \mathrm{Vect}(x_1, \dots, x_m)$. □

2.4 Bases et dimension

Définition 2.11. Soit $\mathbb{K}$ un corps et E un $\mathbb{K}$-espace vectoriel. On dit qu'une famille $(e_1, \dots, e_n)$ est une *base* de E si cette famille est à la fois libre et génératrice.

Proposition 2.12. *Soit $\mathbb{K}$ un corps et E un $\mathbb{K}$-espace vectoriel. Alors $(e_1, \dots, e_n)$ est une base de E si, et seulement si, tout $u \in E$ s'écrit de manière unique sous la forme $u = \lambda_1 e_1 + \cdots + \lambda_n e_n$. On dit alors que $\lambda_1, \dots, \lambda_n$ sont les* coordonnées *de u dans la base $(e_1, \dots, e_n)$.*

Démonstration. Supposons que $(e_1, \dots, e_n)$ est une base de E, et $u \in E$. On sait déjà que u s'écrit sous la forme $u = \lambda_1 e_1 + \dots \lambda_n e_n$ puisque $(e_1, \dots, e_n)$ est génératrice ; montrons que cette décomposition est unique : si on a aussi $u = \lambda'_1 e_1 + \dots \lambda'_n e_n$ alors on obtient $u - u = 0_E = (\lambda_1 - \lambda'_1)e_1 + \cdots + (\lambda_n - \lambda'_n)e_n$, donc comme $(e_1, \dots, e_n)$ est libre on déduit $\lambda_1 - \lambda'_1 = \dots, = \lambda_n - \lambda'_n = 0_{\mathbb{K}}$.

Pour montrer la réciproque, notons que si tout vecteur s'écrit de manière unique comme combinaison linéaire de $(e_1, \dots, e_n)$ alors la famille est en particulier génératrice, et on doit simplement vérifier qu'elle est libre. Pour ce faire, supposons que $\lambda_1, \dots, \lambda_n$ vérifient $\lambda_1 e_1 + \cdots + \lambda_n e_n = 0_E = 0_{\mathbb{K}} e_1 + \dots + 0_{\mathbb{K}} e_n$; par hypothèse d'unicité on doit avoir $\lambda_1 = 0_{\mathbb{K}}, \dots, \lambda_n = 0_{\mathbb{K}}$, ce qui montre comme attendu que notre famille est libre. □

Dans $\mathbb{K}^n$, la *base canonique* est formée par les vecteurs $e_1, \dots, e_n$, dont toutes les coordonnées de e_i sont nulles, sauf la i-ième qui vaut 1. Tout vecteur $x = (x_1, \dots, x_n)$ s'écrit $x = x_1 e_1 + \cdots + x_n e_n$, et d'autre part si $\lambda_1 e_1 + \cdots + \lambda_n e_n = x$, alors en regardant la i-ième coordonnée on obtient $\lambda_i = x_i$ pour tout i, ce qui nous montre que la base canonique est bien une base de $\mathbb{K}^n$ (ouf !).

Lemme 2.13 (Lemme de Steinitz)**.** *Soit $v_1, \dots, v_m$ des vecteurs linéairement indépendants d'un $\mathbb{K}$-espace vectoriel E engendré par des vecteurs $w_1, \dots, w_n$. Alors $m \leqslant n$ et soit $(v_1, \dots, v_m)$ est une base de E soit on peut trouver $i_1, \dots, i_k$ tels que la famille $(v_1, \dots, v_m, w_{i_1}, \dots, w_{i_k})$ soit une base de E.*

Attention, les $(w_j)_{1 \leqslant j \leqslant n}$ forment une famille génératrice mais ils ne sont pas forcément libres ! En particulier, si cette famille est liée, on aura $m + k < n$.

Démonstration. Nous allons montrer par récurrence sur $k \leqslant m$ la propriété $\mathcal{P}_k$: $(v_1, \dots, v_k, w_{k+1}, \dots, w_n)$ est génératrice. En fait nous allons le montrer à un renumérotage des indices de la famille $(w_j)_{1 \leqslant j \leqslant n}$ près.

Initialisation : Comme $(w_j)_{1 \leqslant j \leqslant n}$ est génératrice, il existe des scalaires $(\lambda_j)_{1 \leqslant j \leqslant n} \in \mathbb{K}^n$ tels que $v_1 = \sum_{j=1}^n \lambda_j w_j$. Comme $v_1 \neq 0_E$ car la famille $(v_j)_{1 \leqslant j \leqslant m}$ est libre, les scalaires ne sont pas tous nuls. Supposons, quitte à renuméroter, que $\lambda_1 w_1 \neq 0$. Alors on a

$$w_1 = \frac{1}{\lambda_1}(v_1 - \lambda_2 w_2 - \dots - \lambda_n w_n) \ .$$

Par conséquent, toute combinaison linéaire de (w_i) impliquant w_1 peut être écrite comme une combinaison linéaire impliquant v_1 à la place et la famille $(v_1, w_2, \dots, w_n)$ est génératrice, ce qui prouve $\mathcal{P}_1$.

Hérédité : Supposons $\mathcal{P}_{k-1}$, c'est-à-dire que $(v_1, \dots, v_{k-1}, w_k, \dots, w_n)$ soit génératrice, pour un certain $1 < k \leqslant m$. Alors, $v_k = \sum_{j=1}^{k-1} \mu_j v_j + \sum_{j=k}^n \mu_j w_j$. Comme v_k n'est pas combinaison linéaire des $(v_i)_{1 \leqslant i \leqslant k-1}$, il existe $j \geqslant k$ avec $\mu_j \neq 0$, disons $j = k$, ce qui prouve également que $k \leqslant j \leqslant n$. Alors, de la même manière, w_k peut-être remplacé dans la famille génératrice par v_k et $(v_1, \dots, v_k, w_{k+1}, \dots, w_n)$ est génératrice, c'est-à-dire $\mathcal{P}_k$.

Conclusion : Par récurrence, $(v_1, \dots, v_m, w_{m+1}, \dots, w_n)$ est génératrice et $m \leqslant n$.

En considérant la famille $(v_1, \dots, v_m, w_{m+1}, \dots, w_k)$ pour des k croissants, on peut éliminer des $(w_j)_{m+1 \leqslant j \leqslant n}$ les vecteurs qui sont liés aux précédents et fabriquer ainsi une base de l'espace vectoriel. L'argument qui suppose qu'à chaque pas, c'est le w_k de plus petit indice qu'on remplace et de l'élimination des vecteurs liés, se formalisent dans l'écriture d'un sous-ensemble d'indices $(i_\ell)_{1 \leqslant \ell \leqslant k}$.

□

Théorème 2.14. *Soit $\mathbb{K}$ un corps et E un $\mathbb{K}$-espace vectoriel. Si $(e_1, \dots, e_n)$, $(f_1, \dots, f_m)$ sont deux bases de E alors $n = m$.*

Démonstration. Du lemme précédent et du fait que $(e_1, \dots, e_n)$ est libre et $(f_1, \dots, f_m)$ est génératrice on déduit que $n \leqslant m$; et comme $(f_1, \dots, f_m)$ est libre et $(e_1, \dots, e_n)$ est génératrice on a aussi $m \leqslant n$, donc $m = n$. □

Définition 2.15. Soit $\mathbb{K}$ un corps et E un $\mathbb{K}$-espace vectoriel. Si E admet une partie génératrice finie, on dit que E est de *dimension finie*. Si $E = \{0\}$ on dit que E est de dimension 0 ; sinon on appelle *dimension* de E le cardinal d'une base de E. On note cet entier $\dim(E)$.

Par exemple, $\dim(\mathbb{R}) = 1$, $\dim(\mathbb{R}^2) = 2$ et plus généralement $\dim(\mathbb{K}^n) = n$ pour tout corps $\mathbb{K}$ et tout entier $n \in \mathbb{N}^*$: en effet, on a vu que la base canonique est une base de $\mathbb{K}^n$ avec n éléments.

Un autre exemple important : pour tout entier n l'espace $\mathbb{K}_n[X]$ admet pour base la famille $(1, X, \dots, X^n)$, et est donc de dimension $n+1$.

Le théorème suivant est une conséquence immédiate du lemme de Steinitz (en fait, c'est ce qu'on a commencé par montrer dans la démonstration du lemme de Steinitz).

Théorème 2.16. *Soit $\mathbb{K}$ un corps, E un $\mathbb{K}$-espace vectoriel, L une partie libre finie de E et G une partie génératrice finie de E telles que $L \subseteq G$. Alors il existe une base B de E telle que $L \subseteq B \subseteq G$.*

Ce théorème reste vrai si l'on ne suppose pas L, G finies, modulo l'utilisation de l'*axiome du choix* ; dans ce cours on va éviter d'entrer dans ce type de considérations.

Corollaire 2.17 (Théorème de la base incomplète)**.** Soit $\mathbb{K}$ un corps, et E un $\mathbb{K}$-espace vectoriel de dimension finie $n \geqslant 1$.

1. Pour toute famille libre $(x_1, \dots, x_m)$ on peut trouver $x_{m+1}, \dots, x_n$ tels que $(x_1, \dots, x_n)$ soit une base de E (en particulier, $m \leqslant n$).
2. Pour toute famille génératrice $(x_1, \dots, x_m)$ on peut trouver $i_1 < \dots < i_n$ tels que $(x_{i_1}, \dots, x_{i_n})$ soit une base de E (en particulier $m \geqslant n$).

Démonstration. Le premier point est à nouveau une conséquence immédiate du lemme de Steinitz.

Le deuxième point se montre de manière similaire : puisque E est non réduit à $\{0\}$ il doit exister un plus petit indice i_1 tel que x_{i_1} soit non nul, et la famille (x_{i_1}) est libre ; puisque $(x_{i_1}, \dots, x_m)$ est génératrice, on peut trouver une base B telle que $x_{i_1} \in B \subseteq \{x_{i_1}, \dots, x_m\}$. □

Le théorème de la base incomplète nous dit en particulier que toute partie génératrice contient une base ; comment la trouver en pratique ? Etant donnés des vecteurs $w_1, \dots, w_n$, comment déterminer une base de l'espace vectoriel qu'ils engendrent ? La méthode utilisée dans la preuve du théorème de la base incomplète se résume ainsi : on prend le premier vecteur qui n'est pas nul, et il fera partie de notre base ; puis on prend le premier qui ne lui est pas colinéaire, et on l'ajoute, et ainsi de suite : à chaque étape, on cherche le premier vecteur qui n'est pas dans l'espace vectoriel engendré par les vecteurs qu'on a déjà mis de côté, et on l'ajoute à notre famille. Quand ce procédé s'arrête, on a trouvé notre base.

Une autre méthode est basée sur l'algorithme du pivot de Gauss. On commence par former une matrice dont les *colonnes* sont les vecteurs $w_1, \dots, w_n$. Quand on applique une opération élémentaire sur les colonnes de la matrice, l'espace engendré par les colonnes ne change pas : Dire qu'un vecteur $u \in \mathrm{Im}(A) \subset \mathbb{K}^n$ est dans l'image d'une matrice $A \in M_{nm}(\mathbb{K})$, c'est-dire qu'il existe une combinaison linéaire $\lambda = (\lambda_i) \in \mathbb{K}^m$ telle que $u = A\lambda$. Mais une opération élémentaire $E \in M_m(\mathbb{K})$ agissant sur les colonnes de A définit la combinaison linéaire $\mu = E^{-1}\lambda$ telle que $u = AE\mu \in \mathrm{Im}(AE)$ et $\mathrm{Im}(A) = \mathrm{Im}(AE)$.

On peut donc se ramener à une matrice échelonnée réduite (en colonnes) dont les colonnes engendrent le même espace que l'espace de départ, et les colonnes non nulles dans cette matrice nous donnent la base qu'on était en train de chercher.

Exemple. Appliquons cette méthode sur un exemple : dans $\mathbb{R}^4$, cherchons une base de l'espace vectoriel engendré par les vecteurs $(1,0,1,1)$, $(0,1,0,1)$, $(-1,1,1,-1)$ et $(0,2,2,1)$. On commence par former la matrice

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

Par des opérations élémentaires sur les colonnes, on se ramène à une matrice échelonnée en colonnes :

$$\underset{(C_3 \to C_3 + C_1)}{\simeq} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 2 \\ 1 & 0 & 2 & 2 \\ 1 & 1 & 0 & 1 \end{pmatrix} \underset{(C_3 \to C_3 - C_2 \text{ et } C_4 \to C_4 - 2C_2)}{\simeq} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 2 & 2 \\ 1 & 1 & -1 & -1 \end{pmatrix}.$$

Enfin, $C_4 \to C_4 - C_3$ nous amène à $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 1 & 1 & -1 & 0 \end{pmatrix}$, qui est échelonnée en colonnes.

Dans une matrice échelonnée en colonnes, les colonnes non nulles forment une famille libre ; comme elles engendrent le même sous-espace vectoriel que les vecteurs de départ, une base du sous-espace vectoriel engendré par $(1,0,1,1)$, $(0,1,0,1)$, $(-1,1,1,-1)$ et $(0,2,2,1)$ est donc $(1,0,1,1), (0,1,0,1), (0,0,2,-1)$. En particulier, ce sous-espace vectoriel est de dimension 3.

Deux remarques : tout d'abord, il est inutile de chercher à se ramener à une forme échelonnée réduite ! Le fait que la matrice soit échelonnée nous suffit. Ensuite, différentes combinaisons sur les colonnes pourraient nous donner des bases différentes : c'est normal, un espace vectoriel non réduit à $\{0\}$ admet une infinité de bases différentes...

Il suit des théorèmes précédents que, si E est de dimension finie, alors la dimension de E est égale au plus grand cardinal d'une famille libre d'élements de E, et au plus petit cardinal d'une famille génératrice d'eléments de E. Le résultat de l'exercice suivant est aussi souvent utile.

Exercice 2.18. Soit $\mathbb{K}$ un corps et E un $\mathbb{K}$-espace vectoriel de dimension finie $n \geqslant 1$. Soit B une partie de E. Montrer que les conditions suivantes sont équivalentes :

1. B est une base de E.
2. B est libre et $\text{card}(B) = n$.
3. B est génératrice et $\text{card}(B) = n$.

Proposition 2.19. *Soit $\mathbb{K}$ un corps, E un $\mathbb{K}$-espace vectoriel de dimension finie et F un sous-espace vectoriel de E. Alors F est de dimension finie, $\dim(F) \leqslant \dim(E)$ et $F = E$ si, et seulement si, $\dim(F) = \dim(E)$.*

Démonstration. Si $F = \{0\}$ il n'y a rien à démontrer. Par conséquent on suppose que $\dim(F) \geqslant 1$ (et donc aussi $\dim(E) \geqslant 1$).

Soit $n = \dim(E)$. Alors pour tout $(x_1, \ldots, x_{n+1})$ de F, la famille $(x_1, \ldots, x_{n+1})$ est liée dans E, donc aussi dans F ; appelons m le plus grand entier tel qu'il existe $(x_1, \ldots, x_m)$ qui forment une famille libre dans F (on vient de justifer que m existe et $m \leqslant n$). Alors, pour tout $y \in F$, la famille $(x_1, \ldots, x_m, y)$ est liée ; comme $(x_1, \ldots, x_m)$ est libre on doit avoir $y \in \text{Vect}(x_1, \ldots, x_m)$. Par conséquent $(x_1, \ldots, x_m)$ est une base de F, et on a donc montré à la fois que F est de dimension finie et $\dim(F) \leqslant \dim(E)$.

Bien sûr, si $F = E$ alors $\dim(F) = \dim(E)$. Pour voir la réciproque, supposons que $F \subsetneq E$; on peut trouver une base $(x_1, \ldots, x_m)$ de F et $y \in E \setminus F$; en particulier y n'est pas dans $\text{Vect}(x_1, \ldots, x_m)$, donc $(x_1, \ldots, x_m, y)$ est libre, ce qui montre que $\dim(E) \geqslant m + 1 > m = \dim(F)$. □

Revenons à un exemple vu plus haut : les suites de nombres complexes satisfaisant une relation de récurrence linéaire d'ordre 2.

Proposition 2.20. *Soit $(a, b) \in \mathbb{C}^2 \setminus \{0, 0\}$ et E le sous-espace vectoriel de $\mathbb{C}^{\mathbb{N}}$ formé par les suites satisfaisant la relation de récurrence linéaire d'ordre 2 suivante :*

$$\forall n \in \mathbb{N}, \quad u_{n+2} = a u_{n+1} + b u_n \ .$$

Alors E est un espace vectoriel de dimension 2 ; et on peut en décrire une base explicitement :

1. *Si le polynôme $P(X) = X^2 - aX - b$ a deux racines complexes distinctes α, β alors les suites $(\alpha^n)_{n\in\mathbb{N}}$ et $(\beta^n)_{n\in\mathbb{N}}$ forment une base de E.*
2. *Si P a une racine double α dans $\mathbb{C}$, alors les suites $(\alpha^n)_{n\in\mathbb{N}}$ et $(n\alpha^n)_{n\in\mathbb{N}}$ forment une base de E.*

Démonstration. Commençons par noter qu'un élément de E est uniquement déterminé par ses deux premiers termes. Montrons le de deux manières, par les matrices et par récurrence forte. En construisant la suite de vecteurs $U_n = \begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix}$ pour $n \in \mathbb{N}$, la condition d'appartenance à E se traduit par $U_{n+1} = AU_n$ où $A = \begin{pmatrix} a & b \\ 1 & 0 \end{pmatrix} \in M_2(\mathbb{C})$. Donc, pour tout n, on a $U_n = A^n U_0$ et u_n est uniquement déterminé par u_0 et u_1.

Par récurrence directe, si u, v sont deux éléments de E tels que $u_0 = v_0$ et $u_1 = v_1$ alors on montre que $u_n = v_n$ pour tout $n \in \mathbb{N}$: (Initialisation) Cette égalité est vraie pour $n = 0, 1$. (Hérédité) Supposons que pour un certain $n \geqslant 1$, l'égalité $u_k = v_k$ soit vraie pour tout $k \leqslant n$; alors on a

$$\begin{aligned} u_{n+1} &= au_n + bu_{n-1} \\ &= av_n + bv_{n-1} \qquad \text{(par hypothèse de récurrence)} \\ &= v_{n+1} \ . \end{aligned}$$

Donc par récurrence, (Conclusion) $u_n = v_n$ pour tout $n \in \mathbb{N}$.

De plus, la condition d'appartenance à E est linéaire : si $u, v \in E$ alors, $\lambda u + v \in E$ pour tout $\lambda \in \mathbb{C}$. C'est donc un sous-espace vectoriel de dimension 2 car une base canonique associée à ces deux premiers termes en est donnée par les suites $e_1 = (1, 0, b, ab, a^2b + b^2, a^3b + 2ab^2, \ldots)$ et $e_2 = (0, 1, a, a^2 + b, a^3 + 2ab, \ldots)$. Leurs deux premiers termes assurent qu'elles ne sont pas liées. Mais quels sont leurs termes généraux explicites ?

Notons que si α est une racine de P alors on a pour tout $n \in \mathbb{N}$ l'égalité

$$\alpha^{n+2} - a\alpha^{n+1} - b\alpha^n = \alpha^n(\alpha^2 - a\alpha - b) = \alpha^n P(\alpha) = 0 \ .$$

Donc la suite $(\alpha^n)_{n\in\mathbb{N}}$ est bien un élément de E. Supposons qu'on est dans le cas où P a deux racines distinctes α, β, et montrons que la famille (u, v) définie par, pour tout n, $u_n = \alpha^n$ et $v_n = \beta^n$, est également une base de E.

Tout d'abord, on doit montrer qu'elle est libre, autrement dit que u, v ne sont pas colinéaires : par l'absurde, si $\lambda u = v$ alors en évaluant en $n = 0, 1$ on a à la fois $\lambda u_0 = v_0$ et $\lambda u_1 = v_1$, c'est à dire $\lambda = 1$ et $\lambda\alpha = \beta$, impossible puisque $\alpha \neq \beta$.

Reste à vérifier que (u, v) est génératrice ; soit $w \in E$. On cherche $\lambda, \mu \in \mathbb{C}$ tels que $w = \lambda u + \mu v$. Encore en évaluant en $0, 1$, λ et μ doivent être solutions du système

$$\begin{cases} \lambda \ + \mu & = w_0 \\ \lambda\alpha + \mu\beta & = w_1 \end{cases}$$

ce système se résoud facilement : on obtient

$$\lambda = \quad \frac{w_1 - \beta w_0}{\alpha - \beta} \text{ et } \quad \mu = \frac{\alpha w_0 - w_1}{\alpha - \beta} \ .$$

Alors, $\lambda u + \mu v$ est un élément de E qui prend les mêmes valeurs que w pour $n = 0, 1$: par la remarque faite au début de la preuve, on a $\lambda u + \mu v = w$. Ceci montre bien que (u, v) est génératrice et que le terme général de $w = \lambda u + \mu v$ est

$$w_n = \quad \frac{w_1 - \beta w_0}{\alpha - \beta}\alpha^n + \frac{\alpha w_0 - w_1}{\alpha - \beta}\beta^n.$$

C'est-à-dire que le terme général des deux premières suites est

$$e_{1n} = \frac{\alpha\beta^n - \beta\alpha^n}{\alpha - \beta} \text{ et } e_{2n} = \frac{\alpha^n - \beta^n}{\alpha - \beta}.$$

Le cas où $\alpha \neq 0$ est racine double se traite par exemple « à la physicienne » en considérant a et b comme variables, en faisant un développement limité en $\alpha - \beta$ proche de 0. Alors $\beta^n = \big(\alpha(1 + \frac{\beta-\alpha}{\alpha})\big)^n \underset{0}{=} \alpha^n(1 + n\frac{\beta-\alpha}{\alpha}) + o(\beta - \alpha) = \alpha^n + n(\beta - \alpha)\alpha^{n-1} + o(\beta - \alpha)$ ce qui nous invite à tester la suite $\big(n\alpha^{n-1}\big)_{n\in\mathbb{N}}$ qui fait son office, définissant le plan de solution $E = \mathrm{Vect}\big((\alpha^n)_{n\in\mathbb{N}}, (n\alpha^n)_{n\in\mathbb{N}}\big)$. Remarquez qu'on aurait pu prendre pour deuxième vecteur de base $\big(n\alpha^{n-1}\big)_{n\in\mathbb{N}}$ ou plus généralement $\big((n + k)\alpha^{n+\ell}\big)_{n\in\mathbb{N}}$ pour $k, \ell \in \mathbb{Z}$ fixés, elles définissent le même plan. □

Exercice 2.21. Décrire les suites de nombres réels satisfaisant une relation de récurrence linéaire d'ordre 2. On pourra distinguer les cas où le polynôme P introduit ci-dessus a deux racines réelles, une racine double (forcément réelle) ou deux racines complexes conjuguées.

2.5 Sommes de sous-espaces, sommes directes, supplémentaires

Définition 2.22. Soit $\mathbb{K}$ un corps, E un $\mathbb{K}$-espace vectoriel et F, G deux sous-espaces vectoriels de E. Alors on note

$$F + G = \{f + g \colon f \in F \text{ et } g \in G\} \ .$$

Exemple. Dans $\mathbb{R}^3$, prenons $E_1 = \mathrm{Vect}\big((1,0,0)\big)$, $E_2 = \mathrm{Vect}\big((-1,1,0)\big)$, $E_3 = \{(x,y,z) \in \mathbb{R}^3 : x+y+z = 0\}$. Alors $E_1 + E_2 = \mathrm{Vect}\big((1,0,0),(-1,1,0)\big) = \{(x,y,z) \in \mathbb{R}^3 : z = 0\}$, $E_1 + E_3 = \mathbb{R}^3$, $E_2 + E_3 = E_3$, $(E_1 + E_2) + E_3 = \mathbb{R}^3$, $(E_1 + E_2) \cap E_3 = E_2$.

Proposition 2.23. *Soit $\mathbb{K}$ un corps, E un $\mathbb{K}$-espace vectoriel et F, G deux sous-espaces vectoriels de E. Alors $F + G$ est un sous-espace vectoriel de E.*

Notons que tout sous-espace vectoriel de E qui contient à la fois F et G doit contenir $F + G$: de manière équivalente, on aurait pu définir $F + G$ comme l'intersection de tous les sous-espaces vectoriels de E qui contiennent à la fois F et G.

Démonstration. Puisque $0_E \in F$ et $0_E \in G$, on a $0_E = 0_E + 0_E \in F + G$. Si maintenant $u = f_1 + g_1 \in F + G$, $v = f_2 + g_2 \in F + G$ et $\lambda \in \mathbb{K}$, alors

$$\lambda u + v = (\lambda f_1 + f_2) + (\lambda g_1 + g_2) \in F + G \ .$$

Donc $F + G$ contient 0_E et est stable par combinaisons linéaires : c'est un sous-espace vectoriel de E. □

Définition 2.24. Soit $\mathbb{K}$ un corps, E un $\mathbb{K}$-espace vectoriel de dimension finie et F, G deux sous-espaces vectoriels de E. On dit que F, G sont *en somme directe* si $F \cap G = \{0_E\}$. Dans ce cas on note $F + G = F \oplus G$.

Définition 2.25. Soit $\mathbb{K}$ un corps, E un $\mathbb{K}$-espace vectoriel de dimension finie et F, G deux sous-espaces vectoriels de E. On dit que F et G sont *supplémentaires* si $F \oplus G = E$ tout entier.

Proposition 2.26. *Soit $\mathbb{K}$ un corps, E un $\mathbb{K}$-espace vectoriel de dimension finie et F, G deux sous-espaces vectoriels de E. Si $F \cap G = \{0_E\}$ alors* $\dim(F) + \dim(G) = \dim(F + G)$.

Démonstration. Soit $(f_1, \ldots, f_n)$ une base de F et $(g_1, \ldots, g_m)$ une base de G. La famille $(f_1, \ldots, f_n, g_1, \ldots, g_m)$ est une famille génératrice de $F + G$: ce sont des éléments de $F + G$, et tout élément de $F + G$ s'écrit comme combinaison linéaire de ces vecteurs. On va montrer que cette famille est libre.

Pour ce faire, supposons que $\lambda_1, \ldots, \lambda_n, \mu_1, \ldots, \mu_m$ sont des scalaires tels que

$$\lambda_1 f_1 + \ldots + \lambda_n f_n + \mu_1 g_1 + \ldots + \mu_m g_m = 0_E \ .$$

On peut récrire cela sous la forme

$$\lambda_1 f_1 + \ldots + \lambda_n f_n = -\mu_1 g_1 - \ldots - \mu_m g_m \ .$$

Le vecteur à gauche de cette égalité appartient à F, celui de droite appartient à G ; par conséquent le fait que $F \cap G = \{0_E\}$ entraîne que ce vecteur est nul, autrement dit

$$\lambda_1 f_1 + \ldots \lambda_n f_n = 0_E = \mu_1 g_1 + \ldots + \mu_m g_m \ .$$

Comme $f_1, \ldots, f_n$ et $g_1, \ldots, g_m$ sont libres, on obtient que tous les λ_i et tous les μ_j sont nuls, ce qu'on voulait démontrer. □

Proposition 2.27. *Soit $\mathbb{K}$ un corps, E un $\mathbb{K}$-espace vectoriel de dimension finie et F un sous-espace vectoriel de E. Alors F admet un supplémentaire.*

Démonstration. Si $F = E$ alors F et $\{0_E\}$ sont supplémentaires ; de même si $F = \{0_E\}$ alors F et E sont supplémentaires. Sinon, considérons une base $f_1, \ldots, f_n$ de F. Cette famille est libre dans F, donc peut être complétée en une base $(f_1, \ldots, f_n, g_1, \ldots, g_m)$ de E. En notant $G = \mathrm{Vect}(g_1, \ldots, g_m)$ on a $F + G = \mathrm{Vect}(f_1, \ldots, f_n, g_1, \ldots, g_m) = E$; et on vérifie par un raisonnement analogue à celui employé ci-dessus que $F \cap G = \{0_E\}$. □

Théorème 2.28 (Formule de Grassmann)**.** *Soit $\mathbb{K}$ un corps, E un $\mathbb{K}$-espace vectoriel de dimension finie et F, G deux sous-espaces vectoriels de E. Alors on a*

$$\dim(F+G) = \dim(F) + \dim(G) - \dim(F \cap G) \ .$$

Démonstration. $F \cap G$ est un sous-espace vectoriel de G, donc on peut lui trouver un supplémentaire dans G, c'est-à-dire un sous-espace vectoriel $G_1 \subseteq G$ tel que $(F \cap G) \oplus G_1 = G$. On a alors :

- $\dim(F \cap G) + \dim(G_1) = \dim(G)$;
- $F \oplus G_1 = F + G$.

Le premier point ci-dessus est l'application de (2.26), et il nous reste à justifier le deuxième. Soit $x \in F+G$, alors il existe $f \in F$ et $g \in G$ avec $x = f + g$; et on a aussi $g = a + g_1$ avec $a \in F \cap G$ et $g_1 \in G_1$. D'où $x = (a+f) + g_1 \in F + G_1 \supseteq F + G$. D'autre part, $F \cap G_1 \subseteq (F \cap G) \cap G_1 = \{0_E\}$; finalement on a bien $F \oplus G_1 = F + G$. En écrivant les dimensions, on obtient :

$$\dim(F) + \dim(G_1) = \dim(F+G) \ , \text{ autrement dit}$$

$$\dim(F+G) = \dim(F) + \dim(G) - \dim(F \cap G) \ .$$

□

Proposition 2.29. *Soit $\mathbb{K}$ un corps, E un $\mathbb{K}$-espace vectoriel de dimension finie et F, G deux sous-espaces vectoriels de E. Alors F et G sont en somme directe si, et seulement si,* $\dim(F+G) = \dim(F) + \dim(G)$.

Démonstration. F et G sont en somme directe si, et seulement si $F \cap G = \{0_E\}$, ce qui est équivalent à dire que $\dim(F \cap G) = 0$, et la formule de Grassmann nous dit que ceci est équivalent à $\dim(F) + \dim(G) = \dim(F+G)$. □

Proposition 2.30. *Soit $\mathbb{K}$ un corps, E un $\mathbb{K}$-espace vectoriel de dimension finie et F, G deux sous-espaces vectoriels de E. Les propriétés suivantes sont équivalentes :*

1. *F et G sont supplémentaires.*
2. $\dim(F) + \dim(G) = \dim(E)$ *et* $F \cap G = \{0_E\}$.

Chapitre 3

Applications linéaires

3.1 Définition ; noyau et image d'une application linéaire

Définition 3.1. Soit $\mathbb{K}$ un corps et E, F deux $\mathbb{K}$-espaces vectoriels. Une fonction $\varphi\colon E \to F$ est une *application linéaire* si on a :

- $\varphi(0_E) = 0_F$.
- Pour tout $u, v \in E$ $\varphi(u+v) = \varphi(u) + \varphi(v)$.
- Pour tout $u \in E$ et tout $\lambda \in \mathbb{K}$ $\varphi(\lambda u) = \lambda\varphi(u)$.

On note $\varphi \in \mathcal{L}_{\mathbf{K}}(E;F) = \mathrm{Hom}_{\mathbf{K}}(E;F)$ une application linéaire, ou *morphisme d'espaces vectoriels* de E vers F et $\varphi \in \mathcal{L}_{\mathbf{K}}(E) = \mathrm{End}_{\mathbf{K}}(E)$ si $E = F$, un *endomorphisme* de E. Une application linéaire bijective est appelée un *isomorphisme* de E, $\varphi \in \mathrm{Isom}_{\mathbf{K}}(E;F)$. Un endomorphisme bijectif est un *automorphisme* $\varphi \in \mathrm{GL}_{\mathbf{K}}(E) = \mathrm{Aut}_{\mathbf{K}}(E)$.

Cette définition pourrait s'énoncer de manière beaucoup plus condensée, comme la définition d'un sous-espace vectoriel : $\varphi\colon E \to F$ est une application linéaire si, et seulement si,

$$\forall u, v \in E,\ \forall \lambda \in \mathbb{K}, \quad \varphi(\lambda u + v) = \lambda\varphi(u) + \varphi(v)\ .$$

Notons que si $\varphi, \psi\colon E \to F$ sont des applications linéaires et $\lambda \in \mathbb{K}$ alors $\varphi + \lambda\psi$ est encore une application linéaire. On voit ainsi que l'espace des applications linéaires est un sous-espace vectoriel de l'espace vectoriel formé par toutes les fonctions de E dans F. Ci-dessous, on utilisera simplement le fait que si φ, ψ sont des applications linéaires alors $\varphi - \psi$ est aussi une application linéaire.

Vous avez déjà rencontré énormément d'applications linéaires !

Exemple.

1. Les applications $x \mapsto ax$ (de $\mathbb{R}$ dans $\mathbb{R}$) ; $(x, y) \mapsto ax + by$ (de $\mathbb{R}^2$ dans $\mathbb{R}$) ou $(x_1, \ldots, x_n) \mapsto a_1x_1 + \ldots a_nx_n$ (de $\mathbb{R}^n$ dans $\mathbb{R}$) sont des applications linéaires.
2. Plus généralement, si $A \in M_{n,m}(\mathbb{K})$ alors on a une application linéaire $\varphi\colon \mathbb{K}^m \to \mathbb{K}^n$ définie par

$$\varphi(x_1, \ldots, x_m) = \left(\sum_{k=1}^{m} a_{1,k}x_k, \ldots, \sum_{k=1}^{m} a_{n,k}x_k \right)\ .$$

   On verra dans ce chapitre que toute application linéaire de $\mathbb{K}^m$ dans $\mathbb{K}^n$ est de cette forme, associée à une matrice $A = (a_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant m}}$.
3. La dérivation de polynômes peut être vue comme une application linéaire $D\colon \mathbb{K}[X] \to \mathbb{K}[X]$. Notons que cette application est surjective : pour tout polynôme P il existe un polynôme Q tel que $Q' = P$. Mais elle n'est pas injective : on a $D(1) = D(0) = 0$ mais $1 \neq 0$.
4. Sur l'espace des suites $\mathbb{R}^{\mathbb{N}}$, on peut considérer l'application S qui décale les indices d'une suite : $S(u)$ est la suite définie par $S(u)(n) = u(n+1)$. De nouveau cette application est surjective mais pas injective.
5. Toujours sur $\mathbb{R}^{\mathbb{N}}$, on peut considérer l'application T qui à une suite u associe la suite $T(u)$ définie par $T(u)(0) = 0$, et $T(u)(n) = u(n-1)$ pour $n \geqslant 1$. Cette fois T est injective mais n'est pas surjective :

une suite v telle que $v(0) \neq 0$ n'est pas dans l'image de T. Notons que pour toutes suites u, v on a $S \circ T(u) = u$; pourtant,

$$T \circ S(u) = n \mapsto \begin{cases} 0 & \text{si } n = 0, \\ u(n) & \text{sinon.} \end{cases}$$

donc $T \circ S(u)$ est différent de u alors que $S \circ T(u) = u$!

6. Si on note E l'espace vectoriel formé par toutes les fonctions continues de $\mathbb{R}$ vers $\mathbb{R}$, l'application qui à f associe

$$I(f) \colon x \mapsto \int_0^x f(t)dt$$

est une application linéaire de E dans E. Cette application est injective (si deux fonctions continues ont une primitive commune, elles sont égales) mais pas surjective (toutes les fonctions continues ne sont pas dérivables).

Les exemples 3 à 6 ci-dessus illustrent des phénomènes qui ne peuvent se produire quand on a affaire à des endomorphismes d'espaces vectoriels de dimension finie ; dans ce cours on va continuer à se concentrer sur le cas de la dimension finie.

Définition 3.2. Soit $\mathbb{K}$ un corps, E, F deux $\mathbb{K}$-espaces vectoriels et $\varphi \colon E \to F$ une application linéaire.

- Le *noyau* de φ, noté $\ker(\varphi)$ (pour *kernel*, signifiant noyau en anglais), est défini par

$$\ker(\varphi) = \{u \in E \colon \varphi(u) = 0_F\} \subseteq E \ .$$

- On note $\mathrm{Im}(\varphi)$ l'image de E par φ,

$$\mathrm{Im}(\varphi) = \{v \in F \colon \exists u \in E,\ \varphi(u) = v\} \subseteq F \ .$$

Proposition 3.3. *Soit $\mathbb{K}$ un corps, E, F deux $\mathbb{K}$-espaces vectoriels et $\varphi \colon E \to F$ une application linéaire. Alors $\ker(\varphi)$ est un sous-espace vectoriel de E, et $Im(\varphi)$ est un sous-espace vectoriel de F.*

Démonstration. Commençons par le noyau : il est non vide puisque $\varphi(0_E) = 0_F$ donc $0_E \in \ker(\varphi)$. Ensuite, soit $u, v \in \ker(\varphi)$ et $\lambda \in \mathbb{K}$. On a

$$\varphi(\lambda u + v) = \lambda\varphi(u) + \varphi(v) = \lambda 0_F + 0_F = 0_F \ .$$

Donc $\lambda u + v \in \ker(\varphi)$, et $\ker(\varphi)$ est un sous-espace vectoriel.

Passons à $\mathrm{Im}(\varphi)$: de nouveau il est non vide puisque $0_F = \varphi(0_E) \in \mathrm{Im}(\varphi)$. Si v_1, v_2 appartiennent à $\mathrm{Im}(\varphi)$ et $\lambda \in \mathbb{K}$, alors il existe $u_1, u_2 \in E$ tels que $\varphi(u_1) = v_1$ et $\varphi(u_2) = v_2$. Alors on a

$$\lambda v_1 + v_2 = \lambda\varphi(u_1) + \varphi(u_2) = \varphi(\lambda u_1 + u_2) \in \mathrm{Im}(\varphi) \ .$$

□

Définition 3.4. Soit $\mathbb{K}$ un corps, E, F deux $\mathbb{K}$-espaces vectoriels de dimension finie, et $\varphi \colon E \to F$. Le *rang* de φ est la dimension de $\mathrm{Im}(\varphi)$.

Nous avons vu que le rang d'une matrice $A \in M_{n,m}(\mathbb{K})$ défini par Définition 1.17 (par échelonnage en ligne) est le même que le rang de sa transposée (c'est-à-dire le rang par échelonnage en colonne) qui est donc le même que le rang de l'application linéaire de $\mathcal{L}(\mathbb{K}^m; \mathbb{K}^n)$ associée aux bases canoniques.

Proposition 3.5. *Soit $\mathbb{K}$ un corps, E, F deux $\mathbb{K}$-espaces vectoriels et $\varphi \colon E \to F$ une application linéaire. Alors pout tout $u_1, \ldots, u_n \in E$ on a*

$$\varphi\big(\mathrm{Vect}(u_1, \ldots, u_n)\big) = \mathrm{Vect}\big(\varphi(u_1), \ldots, \varphi(u_n)\big) \ .$$

En particulier, si $(u_1, \ldots, u_n)$ est une famille génératrice de E, alors $\big(\varphi(u_1), \ldots, \varphi(u_n)\big)$ est une famille génératrice de $Im(\varphi)$.

Démonstration. Notons pour commencer que les vecteurs $\varphi(u_1), \dots, \varphi(u_n)$ appartiennent bien à l'espace $\varphi\big(\mathrm{Vect}(u_1, \dots, u_n)\big)$. Reste à montrer que ces vecteurs en forment une partie génératrice.

Soit $v \in \varphi\big(\mathrm{Vect}(u_1, \dots, u_n)\big)$. On peut trouver $u \in \mathrm{Vect}(u_1, \dots, u_n)$ tel que $\varphi(u) = v$; et il existe $\lambda_1, \dots, \lambda_n$ tels que $u = \sum_{i=1}^{n} \lambda_i u_i$. Alors on a

$$\begin{aligned} v &= \varphi(u) \\ &= \varphi\Big(\sum_{i=1}^{n} \lambda_i u_i\Big) \\ &= \sum_{i=1}^{n} \lambda_i \varphi(u_i) \ . \end{aligned}$$

On vient de montrer que $v \in \mathrm{Vect}\big(\varphi(u_1), \dots, \varphi(u_n)\big)$, cette famille est donc bien génératrice de l'espace vectoriel image $\mathrm{Im}(\varphi) = \varphi\big(\mathrm{Vect}(u_1, \dots, u_n)\big)$. □

Exercice 3.6. Soit $\mathbb{K}$ un corps, E, F deux $\mathbb{K}$-espaces vectoriels de dimension finie et $\varphi\colon E \to F$ une application linéaire. Montrer que les propriétés suivantes sont équivalentes :

1. φ est surjective.
2. Pour toute famille génératrice $(u_1, \dots, u_n)$ de E, $\big(\varphi(u_1), \dots, \varphi(u_n)\big)$ est une famille génératrice de F.
3. Il existe une famille génératrice $(u_1, \dots, u_n)$ de E telle que $\big(\varphi(u_1), \dots, \varphi(u_n)\big)$ est une famille génératrice de F.

Proposition 3.7. *Soit $\mathbb{K}$ un corps, E, F deux $\mathbb{K}$-espaces vectoriels de dimension finie, et $\varphi\colon E \to F$. Alors on a* $\mathrm{rang}(\varphi) \leqslant \dim(E)$.

En effet, si $n = \dim(E) \geqslant 1$, alors E a une famille génératrice à n vecteurs ; d'après la propriété précédente, l'image de cette partie est une partie génératrice de $\mathrm{Im}(\varphi)$ contenant n vecteurs, donc la dimension de $\mathrm{Im}(\varphi)$ doit être inférieure à n. Et si $\dim(E) = 0$, alors $E = \{0_E\}$ et $\mathrm{Im}(\varphi) = \{0_F\}$ est aussi de dimension 0.

Proposition 3.8. *Soit $\mathbb{K}$ un corps, E, F deux $\mathbb{K}$-espaces vectoriels et $\varphi\colon E \to F$ une application linéaire. Alors φ est injective si, et seulement si,* $\ker(\varphi) = \{0_E\}$.

Démonstration. Supposons φ injective, et que $u \in E$ est tel que $\varphi(u) = 0_F$. Alors on a $\varphi(u) = \varphi(0_E)$, et comme φ est injective on en déduit $u = 0_E$. Par conséquent si φ est injective on a $\ker(\varphi) = \{0_E\}$.

Réciproquement, supposons $\ker(\varphi) = \{0_E\}$, et que $u, v \in E$ sont tels que $\varphi(u) = \varphi(v)$. Alors on a $0_F = \varphi(u) - \varphi(v) = \varphi(u - v)$, par conséquent $u - v \in \ker(\varphi)$ et donc $u - v = 0_E$, c'est-à-dire $u = v$. on vient de montrer que φ est injective. □

Proposition 3.9. *Soit $\mathbb{K}$ un corps, E, F deux $\mathbb{K}$-espaces vectoriels de dimension finie et $\varphi\colon E \to F$ une application linéaire. Alors les propriétés suivantes sont équivalentes :*

1. *φ est injective.*
2. *Pour toute famille libre $(u_1, \dots, u_n)$ de E la famille $\big(\varphi(u_1), \dots, \varphi(u_n)\big)$ est libre dans F.*
3. *Il existe une base $(u_1, \dots, u_n)$ de E telle que la famille $\big(\varphi(u_1), \dots, \varphi(u_n)\big)$ est libre dans F.*

Démonstration. Supposons que φ est injective et que $u_1, \dots, u_n$ est une famille libre dans E. Soit $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ tels que $\lambda_1\varphi(u_1) + \dots \lambda_n\varphi(u_n) = 0_F$.

On utilise tout d'abord la linéarité de φ pour voir que $\varphi(\lambda_1 u_1 + \dots + \lambda_n u_n) = 0_F$; puis comme φ est injective on en déduit que $\lambda_1 u_1 + \dots + \lambda_n u_n = 0_E$. Enfin, comme $(u_1, \dots, u_n)$ est libre, on conclut que $\lambda_1 = \dots = \lambda_n = 0_{\mathbb{K}}$. On vient de montrer que $\big(\varphi(u_1), \dots, \varphi(u_n)\big)$ est libre. Il est clair que la deuxième propriété est plus forte que la troisième ; reste à montrer que la troisième implique la première. Soit $(u_1, \dots, u_n)$ une base de E telle que $\big(\varphi(u_1), \dots, \varphi(u_n)\big)$ soit libre, et $u \in \ker(\varphi)$. Il existe $\lambda_1, \dots, \lambda_n$ tels que $u = \sum_{i=1}^{n} \lambda_i u_i$. Et on a $0_F = \varphi(u) = \sum_{i=1}^{n} \lambda_i \varphi(u_i)$.

Comme $\big(\varphi(u_1), \dots, \varphi(u_n)\big)$ est libre, on en déduit $\lambda_1 = \dots = \lambda_n = 0$, donc $u = 0_E$. On vient de montrer que $\ker(\varphi) = \{0_E\}$: φ est injective. □

Corollaire 3.10. Soit $\mathbb{K}$ un corps, E, F deux $\mathbb{K}$-espaces vectoriels de dimension finie, et $\varphi\colon E \to F$ une application linéaire. Si φ est injective alors l'image d'une base de E est une base de $\mathrm{Im}(\varphi)$, et en particulier $\mathrm{rang}(\varphi) = \dim(E)$. Réciproquement, si $\mathrm{rang}(\varphi) = \dim(E)$ alors φ est injective.

Démonstration. On sait que l'image d'une partie génératrice de E est une partie génératrice de $\mathrm{Im}(\varphi)$; et si φ est injective alors l'image d'une famille libre est encore libre. Par conséquent, si φ est injective l'image d'une base de E est une partie à la fois libre et génératrice de $\mathrm{Im}(\varphi)$, autrement dit une base de $\mathrm{Im}(\varphi)$.

Réciproquement, supposons que $\mathrm{rang}(\varphi) = \dim(E) = n$ et considérons une base $(u_1, \dots, u_n)$ de E : la famille $\big(\varphi(u_1), \dots, \varphi(u_n)\big)$ est génératrice de $\mathrm{Im}(\varphi)$, et a n éléments ; comme $n = \mathrm{rang}(\varphi) = \dim\big(\mathrm{Im}(\varphi)\big)$, $\big(\varphi(u_1), \dots, \varphi(u_n)\big)$ doit être une base de $\mathrm{Im}(\varphi)$. En particulier on vient de trouver une famille libre dont l'image est une famille libre, donc φ est injective. □

Définition 3.11. Soit $\mathbb{K}$ un corps, E, F deux $\mathbb{K}$-espaces vectoriels, et $\varphi\colon E \to F$ une application linéaire. Si φ est bijective on dit que φ est un *isomorphisme* de E sur F. Si de plus $E = F$ alors on dit que φ est un *automorphisme*.

Théorème 3.12. *Soit $\mathbb{K}$ un corps, E, F deux $\mathbb{K}$-espaces vectoriels de dimension finie, et $\varphi\colon E \to F$. Si φ est un isomorphisme de E sur F alors* $\dim(E) = \dim(F)$.

Si $\dim(E) = \dim(F)$ *alors les propriétés suivantes sont équivalentes :*

1. *φ est bijective.*
2. *φ est injective.*
3. *φ est surjective.*

Toutes ces propriétés sont aussi équivalentes à dire que φ envoie une base de E sur une base de F. Ce théorème peut être vu comme un cas particulier du théorème du rang qu'on verra dans la prochaine section.

Démonstration. Si φ est un isomorphisme, alors φ est injective et donc $\mathrm{rang}(\varphi) = \dim(E)$; mais φ est aussi surjective donc $\mathrm{rang}(\varphi) = \dim(F)$. Par conséquent, on doit bien avoir $\dim(E) = \dim(F)$ s'il existe un isomorphisme de E sur F. Supposons maintenant que $\dim(E) = \dim(F)$ et montrons que les propriétés (1), (2) et (3) sont équivalentes. Manifestement, (1) est plus forte que (2) et (3). Notons dans la suite $n = \dim(E) = \dim(F)$.

Supposons φ injective, et soit $(u_1, \dots, u_n)$ une base de E. Alors, d'après 3.9, on sait que $\big(\varphi(u_1), \dots, \varphi(u_n)\big)$ est une famille libre de F, et cette famille libre a $n = \dim(F)$ éléments. C'est donc une base. On vient de montrer que l'image d'une base de E est une base de F : φ est bijective.

Reste à montrer que si φ est surjective alors φ est bijective ; le raisonnement est très proche du précédent : si $(u_1, \dots, u_n)$ est une base de E alors 3.6, on sait que $\big(\varphi(u_1), \dots, \varphi(u_n)\big)$ est une famille génératrice de F, et cette famille génératrice a $n = \dim(F)$ éléments. C'est donc une base. On vient de montrer que l'image d'une base de E est une base de F : φ est bijective. □

En particulier : il ne peut y avoir d'isomorphisme entre $\mathbb{K}^n$ et $\mathbb{K}^m$ que si $n = m$; et une application linéaire de $\mathbb{K}^n$ dans lui-même est bijective ssi elle est injective ssi elle est surjective.

3.2 Le théorème du rang

Lemme 3.13. *Soit $\mathbb{K}$ un corps, E, F deux $\mathbb{K}$-espaces vectoriels, et $\varphi\colon E \to F$ une application linéaire. Supposons que $E = \ker(\varphi) \oplus H$. Alors la restriction de φ à H induit un isomorphisme de H sur $Im(\varphi)$.*

Démonstration. Commençons par montrer que la restriction de φ à H est injective : si u appartient à son noyau, alors on a à la fois $\varphi(u) = 0_F$ et $u \in H$; autrement dit, $u \in \ker(\varphi) \cap H = \{0_E\}$ puisque H et $\ker(\varphi)$ sont en somme directe.

Pour montrer que l'application est surjective de H sur $\mathrm{Im}(\varphi)$, considérons $v \in \mathrm{Im}(\varphi)$; il existe $u \in E$ tel que $\varphi(u) = v$, et puisque $\ker(\varphi) \oplus H = E$ on peut écrire $u = u_1 + u_2$ avec $u_1 \in \ker(\varphi)$ et $u_2 \in H$. Mais alors on a

$$\varphi(u) = \varphi(u_1 + u_2) = \varphi(u_1) + \varphi(u_2) = 0 + \varphi(u_2) = \varphi(u_2)\ .$$

Par conséquent v appartient à l'image de H par φ, ce qu'on souhaitait démontrer. □

Théorème 3.14 (Théorème du rang)**.** *Soit $\mathbb{K}$ un corps, E, F deux $\mathbb{K}$-espaces vectoriels de dimension finie, et $\varphi\colon E \to F$ une application linéaire. Alors on a*

$$\operatorname{rang}(\varphi) + \dim\big(\ker(\varphi)\big) = \dim(E) \ .$$

En particulier, si $\dim(E) < \dim(F)$ alors φ ne peut pas être surjective ; et si $\dim(E) > \dim(F)$ alors son noyau ne peut pas être réduit à $\{0_E\}$, donc φ ne peut pas être injective.

Démonstration. Comme E est de dimension finie et $\ker(\varphi)$ est un sous-espace vectoriel de E, on peut trouver un supplémentaire H de $\ker(\varphi)$ dans E. Alors la formule de Grassmann nous dit que

$$\dim\big(\ker(\varphi)\big) + \dim(H) = \dim(E) \ .$$

Mais puisque φ induit un isomorphisme de H sur $\operatorname{Im}(\varphi)$, on a $\dim(H) = \dim\big(\operatorname{Im}(\varphi)\big) = \operatorname{rang}(\varphi)$, et on obtient bien

$$\operatorname{rang}(\varphi) + \dim\big(\ker(\varphi)\big) = \dim(E) \ .$$

□

3.3 Définition d'une application linéaire à partir de ses valeurs sur une base

Proposition 3.15. *Soit $\mathbb{K}$ un corps, E, F deux $\mathbb{K}$-espaces vectoriels et $(u_1, \dots, u_n)$ une base de E. Alors pour toute famille $(v_1, \dots, v_n)$ de vecteurs de F il existe une unique application linéaire $\varphi\colon E \to F$ telle que $\varphi(u_1) = v_1, \dots, \varphi(u_n) = v_n$.*

Démonstration. Commençons par montrer qu'une telle application, si elle existe est unique : si φ, ψ sont deux applications linéaires telles que $\varphi(u_i) = \psi(u_i) = v_i$ pour tout $i \in \{1, \dots, n\}$, alors $(\varphi - \psi)(u_i) = 0_F$ pour tout $i \in \{1, \dots, n\}$; par conséquent $\ker(\varphi - \psi)$ est un sous-espace vectoriel de E qui contient tous les u_i : c'est E tout entier puisque la famille $(u_1, \dots, u_n)$ est génératrice. Donc $(\varphi - \psi)(u) = 0_F$ pour tout $u \in E$, autrement dit $\varphi = \psi$.

Cet argument nous dit que, pour définir φ, on n'a pas beaucoup de choix : tout vecteur u s'écrit de manière unique sous la forme $u = \lambda_1 u_1 + \dots \lambda_n u_n$, et on pose

$$\varphi(u) = \lambda_1 \varphi(u_1) + \dots + \lambda_n \varphi(u_n) \ .$$

Reste à vérifier que φ ainsi définie est linéaire ; pour $u, v \in E$ et $\lambda \in \mathbb{K}$, on a des scalaires $\lambda_1, \dots, \lambda_n$ et $\mu_1, \dots, \mu_n$ tels que

$$u = \sum_{i=1}^{n} \lambda_i u_i \quad \text{et} \quad v = \sum_{i=1}^{n} \mu_i u_i \ .$$

Donc

$$\lambda u + v = \sum_{i=1}^{n} (\lambda\lambda_i + \mu_i) u_i \ .$$

Par définition de φ, on obtient

$$\begin{aligned} \varphi(\lambda u + v) &= \varphi\left(\sum_{i=1}^{n} (\lambda\lambda_i + \mu_i) u_i \right) \\ &= \sum_{i=1}^{n} (\lambda\lambda_i + \mu_i) \varphi(u_i) \\ &= \lambda \sum_{i=1}^{n} \lambda_i \varphi(u_i) + \sum_{i=1}^{n} \mu_i \varphi(u_i) \\ &= \lambda\varphi(u) + \varphi(v) \ . \end{aligned}$$

□

Corollaire 3.16. Soit $\mathbb{K}$ un corps et E un $\mathbb{K}$-espace vectoriel de dimension $n \geqslant 1$. Alors E est isomorphe à $\mathbb{K}^n$.

Démonstration. Soit $(e_1, \dots, e_n)$ la base canonique de $\mathbb{K}^n$, et $(u_1, \dots, u_n)$ une base de E. Par la proposition précédente, il existe une unique application linéaire de $\mathbb{K}^n$ sur E telle que $\varphi(e_i) = u_i$; par construction, cette application envoie une base de $\mathbb{K}^n$ sur une base de E et est donc un isomorphisme. □

On voit donc que deux $\mathbb{K}$-espaces vectoriels de même dimension finie sont isomorphes. Dans la suite, on prendra la convention que $\mathbb{K}^0 = \{0_\mathbb{K}\}$.

Proposition 3.17. *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n, et F un sous-espace vectoriel de dimension k. Alors il existe une application linéaire $\varphi\colon E \to \mathbb{K}^{n-k}$ tel que $F = \ker(\varphi)$*

(Notons qu'on convient que $\mathbb{K}^0 = \{0_\mathbb{K}\}$, ce qui est cohérent avec le fait que $\dim(\mathbb{K}^n) = n$)

Démonstration. Si $k = n$, autrement dit si $F = E$, l'application $E \to \mathbb{K}^0 = \{0_\mathbb{K}\}$ qui envoie tout $u \in E$ sur $0_\mathbb{K}$ convient. Sinon, partons d'une base $(e_1, \dots, e_k)$ de F et complétons-la (grâce au théorème de la base incomplète) en une base $(e_i)_{i \leqslant n}$ de E. Soit $f_1, \dots, f_{n-k}$ la base canonique de $\mathbb{K}^{n-k}$, et φ l'unique application linéaire telle que $\varphi(e_i) = 0$ pour tout $i \leqslant k$, et $\varphi(e_{k+i}) = f_i$ pour tout $i \in \{1, \dots, n-k\}$. Alors il est clair que $F \subseteq \ker(\varphi)$, et que $\mathrm{Im}(\varphi) = \mathbb{K}^{n-k}$. Le théorème du rang nous donne $\dim\big(\ker(\varphi)\big) = k = \dim(F)$, donc $F = \ker(\varphi)$. □

3.4 Applications linéaires et matrices

Si $(e_1, \dots, e_n)$ est une base de E, alors d'après la proposition 3.15 toute l'information sur φ est contenue dans la valeur des vecteurs $\varphi(e_1), \dots, \varphi(e_n)$; on peut regrouper cette information dans une matrice.

Définition 3.18. Soit $\mathbb{K}$ un corps, E, F deux $\mathbb{K}$-espaces vectoriels de dimension finie et $\varphi\colon E \to F$ une application linéaire. Etant données une base $B = (e_1, \dots, e_m)$ de E et une base $B' = (f_1, \dots, f_n)$ de F, on peut écrire pour tout j

$$\varphi(e_j) = \sum_{i=1}^{n} a_{i,j} f_i \ .$$

La *matrice de φ dans les bases B, B'* est la matrice

$$\mathrm{Mat}_{B',B}(\varphi) := (a_{i,j})_{\substack{1 \leqslant j \leqslant m \\ 1 \leqslant i \leqslant n}} .$$

Autrement dit : la matrice de φ dans les bases B, B' est la matrice dont la j-ième colonne correspond aux coordonnées du vecteur $\varphi(e_j)$ dans la base $(f_1, \dots, f_n)$. Le nombre de colonnes de cette matrice correspond à la dimension de l'espace de départ, et le nombre de lignes correspond à la dimension de l'espace d'arrivée.

Quand $E = F$ et $B = B'$, on parle simplement de la matrice de φ dans la base B.

Donnons quelques exemples.

Exemple.

1. Considérons la rotation d'angle $\pi/2$ dans $\mathbb{R}^2$, dont on note la base canonique (e_1, e_2) : elle envoie e_1 sur e_2 et e_2 sur $-e_1$, et sa matrice dans la base canonique est donc $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.
2. Soit l'application $\varphi\colon \mathbb{R}^4 \to \mathbb{R}^2$ définie par $\varphi(x, y, z, t) = (x + y + t, z - y)$. En notant (e_1, e_2, e_3, e_4) la base canonique de $\mathbb{R}^4$ et (f_1, f_2) la base canonique de $\mathbb{R}^2$ on a : $\varphi(e_1) = (1, 0) = f_1$; $\varphi(e_2) = (1, -1) = f_1 - f_2$; $\varphi(e_3) = (0, 1) = f_2$, et $\varphi(e_4) = (1, 0) = f_1$. La matrice de φ dans les bases canoniques est donc $\begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & -1 & 1 & 0 \end{pmatrix}$.
3. Considérons l'espace $\mathbb{R}_4[X]$ muni de la base $(1, X, X^2, X^3, X^4)$ et l'application linéaire $P \mapsto P - XP'$. Alors on a $\varphi(1) = 1$, $\varphi(X) = 0$, $\varphi(X^2) = -X^2$, $\varphi(X^3) = -2X^3$ et $\varphi(X^4) = -3X^4$. La matrice de φ dans cette base est donc
$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & -3 \end{pmatrix}.$$

Une application linéaire n'est associée à des matrices différentes si on change les bases de départ et d'arrivée, de même que deux applications linéaires peuvent être associées à la même matrice dans des bases différentes ! Par exemple, si $\varphi\colon E \to E$ est un isomorphisme d'un espace de dimension finie E sur lui-même, que $(e_1, \dots, e_n) = B$ est une base de E, alors $B' = \big(\varphi(e_1), \dots, \varphi(e_n)\big)$ est aussi une base de E ; et la matrice $\mathrm{Mat}_{B',B}(\varphi)$ est la matrice identité I_n (où $n = \dim(E)$). Retenez bien que l'association entre une application linéaire et une matrice n'a de sens que si on spécifie les bases de départ et d'arrivée.

Le calcul fait dans la preuve de la Proposition 3.15 nous permet d'expliciter comment utiliser des matrices pour calculer les images de vecteurs par une application linéaire en utilisant leurs coordonnées dans la base de départ, donnant alors leurs coordonnées dans la base d'arrivée.

Proposition 3.19. *Soit $\mathbb{K}$ un corps, E, F deux $\mathbb{K}$-espaces vectoriels de dimension finie et $\varphi\colon E \to F$ une application linéaire, $B = (e_1, \dots, e_n)$ une base de E et $B' = (f_1, \dots, f_n)$ une base de F, $A = \mathrm{Mat}_{B',B}(\varphi)$ la matrice de φ de la base B dans la base B'. Pour tout $u \in E$, on définit $X = \mathrm{Mat}_B(u) = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$, où $x_1, \dots, x_m \in \mathbb{K}$ sont les coordonnées de u dans la base B. Alors, si $\varphi(u) = a_1 f_1 + \ldots + a_n f_n$, on a*

$$AX = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \mathrm{Mat}_{B'}\big(\varphi(u)\big) = \mathrm{Mat}_{B',B}\big(\varphi\big)\,\mathrm{Mat}_B(u).$$

Autrement dit : calculer le vecteur $\varphi(u)$ revient à calculer le vecteur colonne AX, car se sont ses coordonnées dans la base B'.

En particulier, déterminer le noyau de φ revient matriciellement à résoudre l'équation $AX = 0_n$ dans $M_{m,1}(\mathbb{K})$; comme on l'a vu au chapitre précédent, une façon de résoudre ce système est d'échelonner $A \in M_{n,m}(\mathbb{K})$ en lignes (en termes d'application linéaires : remplacer φ par $\psi \circ \varphi$, où ψ est un automorphisme de F, ne change pas le noyau de φ).

On identifie le *noyau* $A \in M_{n,m}(\mathbb{K})$ (un sous-ensemble de $\mathbb{K}^m$) avec le noyau de l'application linéaire φ dont la matrice dans les bases canoniques de $\mathbb{K}^m$ et $\mathbb{K}^n$ est égale à A.

Rappelons aussi que $A \in M_n(\mathbb{K})$ est inversible si, et seulement si, le système $AX = 0_n$ n'a que 0_n comme solution ; d'après le résultat ci-dessus cela arrive si et seulement si l'application linéaire φ dont A est la matrice dans la base canonique a un noyau réduit à $\{0_n\}$, autrement dit si et seulement si φ est un automorphisme de $\mathbb{K}^n$.

Dans le même ordre d'idées, déterminer l'image de φ revient à trouver tous les $Y \in M_{n,1}(\mathbb{K})$ tels que le système $AX = Y$ a une solution, ou encore l'espace vectoriel engendré par les colonnes de A ; cette fois on a vu qu'on pouvait faire cela en échelonnant A en colonnes (remplacer φ par $\varphi \circ \psi$, où ψ est un automorphisme de E, ne change pas son image).

Démonstration. Rappelons que le coefficient $a_{i,j}$ de A est défini par l'égalité

$$\varphi(e_j) = \sum_{i=1}^{n} a_{i,j} f_i \ .$$

Si les coordonnées de u dans la base B sont $x_1, \dots, x_n$, c'est-à-dire si on a $u = \sum_{k=1}^{m} x_k e_k$ alors :

$$\begin{aligned}
\varphi(u) &= \varphi\left(\sum_{k=1}^{m} x_k e_k\right) \\
&= \sum_{k=1}^{m} x_k \varphi(e_k) \\
&= \sum_{k=1}^{m} x_k \left(\sum_{i=1}^{n} a_{i,k} f_i\right) \\
&= \sum_{i=1}^{n} \left(\sum_{k=1}^{m} a_{i,k} x_k\right) f_i.
\end{aligned}$$

Pour $i \in \{1, \dots, n\}$, la i-ième coordonnée de $\varphi(u)$ est donc égale à $\sum_{k=1}^{m} a_{i,k}x_k$, qui est bien le coefficient sur la i-ème ligne de AX. □

Rappelons que, étant données deux matrices $A \in M_{n,m}(\mathbb{K})$ et $B \in M_{m,p}(\mathbb{K})$, si on appelle $B_1, \dots, B_p$ les colonnes de B, alors AB est la matrice de $M_{n,p}(\mathbb{K})$ dont les colonnes sont $AB_1, \dots, AB_p$. Au cas où ce ne serait pas clair, refaisons le calcul : pour tout $j \in \{1, \dots, p\}$, B_j est la matrice à m lignes et 1 colonne $\begin{pmatrix} b_{1,j} \\ \vdots \\ b_{m,j} \end{pmatrix}$; donc AB_j est la matrice à n lignes et 1 colonne dont le coefficient sur la ligne i est égal à $\sum_{k=1}^{m} a_{i,k}b_{k,j}$. On voit donc que, comme annoncé, AB_j est la j-ième colonne de AB.

Proposition 3.20. *Soit $\mathbb{K}$ un corps, E_1, E_2, E_3 trois espaces vectoriels de dimension finie et B_1, B_2, B_3 des bases de E_1, E_2, E_3 respectivement. Alors, pour toutes applications linéaires $\varphi\colon E_2 \to E_3$ et $\psi\colon E_1 \to E_2$ on a*

$$\mathrm{Mat}_{B_3,B_1}(\varphi \circ \psi) = \mathrm{Mat}_{B_3,B_2}(\varphi)\,\mathrm{Mat}_{B_2,B_1}(\psi)\ .$$

On retrouve ici le lien entre composition d'applications et produit de matrices qui nous a servi à introduire le produit de matrices...

Démonstration. Notons $B = (u_1, \dots, u_{n_1})$. Soit n_i la dimension de E_i. Alors $\mathrm{Mat}_{B_3,B_1}(\varphi \circ \psi)$ est la matrice à n_3 lignes et n_1 colonnes dont la j-ième colonne est formée des coordonnées de $\varphi \circ \psi(u_j)$ dans la base B_3.

Quant à $\mathrm{Mat}_{B_3,B_2}(\varphi)\,\mathrm{Mat}_{B_2,B1}(\psi)$, c'est la matrice à n_3 lignes et n_1 colonnes dont la j-ième colonne est égale au produit $\mathrm{Mat}_{B_3,B_2}(\varphi)G_j$, où G_j est la j-ième colonne de $\mathrm{Mat}_{B_2,B_1}(\psi)$, c'est-à-dire la matrice colonne formée des coordonnées de $\psi(u_j)$ dans la base B_2. La j-ième colonne de $\mathrm{Mat}_{B_3,B_2}(\varphi)\,\mathrm{Mat}_{B_2,B_1}(\psi)$ est donc formée des coordonnées de $\varphi\big(\psi(u_j)\big)$ dans la base B_3, ce qui nous démontre bien l'égalité attendue. □

Définition 3.21. Soit $\mathbb{K}$ un corps, E un $\mathbb{K}$-espace vectoriel de dimension finie $n \geqslant 1$, et $B = (e_1, \dots, e_n), B' = (f_1, \dots, f_n)$ deux bases de E. La *matrice de passage* $P_B^{B'}$ de B à B' est la matrice $P_B^{B'} := \mathrm{Mat}_{B,B'}(\mathrm{id}_E)$; autrement dit, c'est la matrice de $M_n(\mathbb{K})$ dont la j-ième colonne donne les coordonnées de f_j exprimés dans la base B.

Attention aux conventions ! La matrice de passage de B à B' est bien définie comme la matrice de l'identité *de B' vers B* ! Chacune de ses colonnes correspond à un vecteur de B' et chaque ligne à un vecteur de B. Pour cette raison je recommanderais d'utiliser la notation lourde mais explicite $\mathrm{Mat}_{B,B'}(\mathrm{id}_E)$ plutôt que la notation $P_B^{B'}$. Ci-dessous on va tout de même énoncer quelques propriétés des matrices de passage en utilisant la notation $P_B^{B'}$, qui est la plus communément utilisée.

Proposition 3.22. *Soit $\mathbb{K}$ un corps, E un $\mathbb{K}$-espace vectoriel de dimension finie $n \geqslant 1$, et B, B' deux bases de E. Alors la matrice de passage $P_B^{B'}$ est inversible, et son inverse est la matrice de passage $P_{B'}^{B}$ de B' à B.*

Démonstration. Si on raisonne en termes d'applications linéaires, ce résultat est immédiat :

$$P_B^{B'}P_{B'}^{B} = \mathrm{Mat}_{B,B'}(\mathrm{id}_E)\,\mathrm{Mat}_{B',B}(\mathrm{id}_E) = \mathrm{Mat}_{B,B}(\mathrm{id}_E) = I_n\ .$$

□

Réciproquement, toute matrice inversible peut être vue comme une matrice de passage : si $A \in GL_n(\mathbb{K})$ est inversible, ses vecteurs colonnes sont les coordonnées de vecteurs qui forment une base B' de $\mathbb{K}^n$, lui-même muni de sa base canonique B. Donc A est la matrice de passage de B à B'. Comprendre cette quasi-tautologie est le problème essentiel des matrices de passage.

Proposition 3.23. *Soit $\mathbb{K}$ un corps, E un $\mathbb{K}$-espace vectoriel de dimension finie $n \geqslant 1$, et B, B' deux bases de E. Si $x \in E$, et X, X' désignent les vecteurs colonnes donnant les coordonnées de x dans les bases B, B' respectivement, alors on a $X = P_B^{B'}X'$.*

Démonstration. Par définition, $P_B^{B'}X' = \mathrm{Mat}_{B,B'}(\mathrm{id}_E)X'$ est la matrice colonne listant les coordonnées de $\mathrm{id}(x) = x$ dans la base B : c'est X. □

Proposition 3.24. *Soit $\mathbb{K}$ un corps, E, F deux $\mathbb{K}$-espaces vectoriels de dimension finie $\geqslant 1$, et $\varphi\colon E \to F$ une application linéaire. Soit B_E, B'_E deux bases de E ; B_F, B'_F deux bases de F, et P, Q les matrices de passage de B_E à B'_E et B_F à B'_F respectivement. Si M est la matrice de φ de B_E dans B_F, et M' est la matrice de φ de B'_E dans B'_F, alors on a*

$$M' = Q^{-1}MP\ .$$

Démonstration. Les définitions de P et Q sont : $P = \mathrm{Mat}_{B_E,B'_E}(\mathrm{id}_E)$ et $Q = \mathrm{Mat}_{B_F,B'_F}(\mathrm{id}_F)$. On écrit

$$\begin{aligned} M' &= \mathrm{Mat}_{B'_F,B'_E}(\varphi) \\ &= \mathrm{Mat}_{B'_F,B_F}(\mathrm{id})\,\mathrm{Mat}_{B_F,B_E}(\varphi)\,\mathrm{Mat}_{B_E,B'_E}(\mathrm{id}) \\ &= Q^{-1}MP. \end{aligned}$$

□

En particulier, quand $E = F$, que M désigne la matrice de φ dans la base B, M' la matrice de φ dans la base B' et P la matrice de passage de B à B' alors on a

$$M' = P^{-1}MP\ .$$

Par exemple, considérons l'application linéaire de $\mathbb{R}^3$ dans $\mathbb{R}^3$ définie par

$$\varphi(x, y, z) = (9x - y - 6z, 5x + 3y - 6z, x - y + 2z)\ .$$

Cette application est linéaire, et sa matrice dans la base canonique (e_1, e_2, e_3) est

$$\begin{pmatrix} 9 & -1 & -6 \\ 5 & 3 & -6 \\ 1 & -1 & 2 \end{pmatrix}\ .$$

Calculons la matrice de φ dans une autre base par exemple la famille de vecteurs $(f_1, f_2, f_3) = \big((1,1,1),(1,-1,1),(1,1,0)\big)$. On a vu dans le chapitre sur les matrices une méthode basée sur le pivot de Gauss qu'on pourrait appliquer pour montrer que la matrice de passage de la base canonique e à la base f

$$P_e^f = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

est inversible, et calculer son inverse. Il y a d'autres méthodes : par exemple, pour montrer que (f_1, f_2, f_3) est une base il nous suffit de montrer qu'elle est génératrice car c'est une famille de 3 vecteurs de $\mathbb{R}^3$, qui est de dimension 3. Autrement dit peut-on exprimer (e_1, e_2, e_3) comme combinaisons linéaires de (f_1, f_2, f_3) ? Les coefficients de ces combinaisons linéaires nous donneront alors la matrice P_f^e. Essayons d'exprimer (e_1, e_2, e_3) en fonction de (f_1, f_2, f_3), sachant qu'on a

$$\begin{cases} f_1 &= e_1 + e_2 + e_3 \\ f_2 &= e_1 - e_2 + e_3 \\ f_3 &= e_1 + e_2. \end{cases}$$

En considérant $L_1 - L_2$, on obtient que $2e_2 = f_1 - f_2$ soit $e_2 = \frac{1}{2}f_1 - \frac{1}{2}f_2$; et avec $L_1 - L_3$ on arrive à $e_3 = f_1 - f_3$. Il est alors facile d'obtenir

$$e_1 = f_3 - \frac{f_1 - f_2}{2} = -\frac{1}{2}f_1 + \frac{1}{2}f_2 + f_3\ .$$

On voit donc que l'espace vectoriel engendré par (f_1, f_2, f_3) contient e_1, e_2, e_3 et est donc égal à $\mathbb{R}^3$. Donc (f_1, f_2, f_3) forment une base de $\mathbb{R}^3$; de plus on a calculé la matrice de passage de (f_1, f_2, f_3) à (e_1, e_2, e_3), autrement dit la matrice $P_f^e = {P_e^f}^{-1}$, qui est égale à

$$P_f^e = \begin{pmatrix} -\frac{1}{2} & \frac{1}{2} & 1 \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ 1 & 0 & -1 \end{pmatrix}.$$

Pour calculer la matrice M' de φ dans la base (f_1, f_2, f_3) il nous reste à calculer

$$\begin{aligned} M' &= \mathrm{Mat}_f(\varphi) = \mathrm{Mat}_{f,e}(\mathrm{id}) Mat_e(\varphi) \mathrm{Mat}_{e,f}(\mathrm{id}) = P_f^e M P_e^f \\ &= (P_f^e M) P_e^f \\ &= \begin{pmatrix} -1 & 1 & 2 \\ 2 & -2 & 0 \\ 8 & 0 & -8 \end{pmatrix} P_e^f \\ &= \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 8 \end{pmatrix} . \end{aligned}$$

Ceci revient à dire que $\varphi(f_1) = 2f_1$, $\varphi(f_2) = 2f_2$ et $\varphi(f_3) = 4f_3$; notons que la matrice M' est diagonale, et qu'il est beaucoup plus simple de comprendre les propriétés de φ en regardant M' que M. Par exemple, on voit immédiatement que φ est surjective et donc que φ est un automorphisme de $\mathbb{R}^3$.

Définition 3.25. Soit $n, m \in \mathbb{N}^*$, $\mathbb{K}$ un corps et $A, B \in M_{n,m}(\mathbb{K})$. On dit que A et B sont *équivalentes* s'il existe deux matrices inversibles $P \in GL_n(\mathbb{K})$ et $Q \in GL_m(\mathbb{K})$ telles que $A = PBQ$.

Si $n = m$ et qu'on peut écrire $A = P^{-1}BP$ pour une matrice inversible $P \in GL_n(\mathbb{K})$, alors on dit que A et B sont *semblables*.

Si on raisonne en termes d'applications linéaires : deux matrices équivalentes dans $M_{n,m}(\mathbb{K})$ peuvent être vues comme représentant la même application de $\mathbb{K}^m$ dans $\mathbb{K}^n$ quitte à changer les bases au départ et à l'arrivée ; deux matrices semblables dans $M_n(\mathbb{K})$ peuvent être vues comme représentant la même application linéaire de $\mathbb{K}^n$ dans $\mathbb{K}^n$ quitte à changer de base (avec la même base au départ et à l'arrivée).

3.5 Rang d'une matrice et de sa transposée

Définition 3.26. Soit $A \in M_{n,m}(\mathbb{K})$. Le *rang* de A est la dimension du sous-espace vectoriel de $\mathbb{K}^n$ engendré par les colonnes de A.

Autrement dit : si $\varphi \in \mathcal{L}(\mathbb{K}^m; \mathbb{K}^n)$ est l'application dont la matrice dans les bases canoniques de $\mathbb{K}^m$ et $\mathbb{K}^n$ est A, alors $\mathrm{rang}(A) = \mathrm{rang}(\varphi)$ suivant Définition 3.4.

Proposition 3.27. *Soit $\mathbb{K}$ un corps, $n, m \in \mathbb{N}^*$ et $A, B \in M_{n,m}(\mathbb{K})$ deux matrices équivalentes. Alors* $\mathrm{rang}(A) = \mathrm{rang}(B)$.

Démonstration. Une première preuve s'appuie sur la forme échelonnée réduite d'une matrice de rang r donné, en échelonnant suivant les lignes A, on construit une matrice inversible $E \in \mathrm{GL}_n(\mathbb{K})$ telle que $EAS = \begin{pmatrix} I_r & 0_{r,m-r} \\ 0_{n-r,r} & 0_{n-r,m-r} \end{pmatrix}$ à une permutation $S \in \mathrm{GL}_m(\mathbb{K})$ des colonnes près et de même, en échelonnant B suivant les colonnes, on obtient $F \in \mathrm{GL}_m(\mathbb{K})$ telle que $S'BF$ est la même matrice, avec $S' \in \mathrm{GL}_n(\mathbb{K})$ une permutation des lignes. Ainsi, $EAS = S'BF$ et $A = (E^{-1}S')B(FS^{-1})$ sont équivalentes.

Raisonnons en termes d'applications linéaires : si $\varphi \colon \mathbb{K}^m \to \mathbb{K}^n$ est une application linéaire et $\psi_1 \colon \mathbb{K}^n \to \mathbb{K}^n$, $\psi_2 \colon \mathbb{K}^m \to \mathbb{K}^m$ sont des automorphismes, alors le noyau de $\psi_1 \circ \varphi$ et celui de φ sont égaux, donc $\mathrm{rang}(\psi_1 \circ \varphi) = \mathrm{rang}(\varphi)$. Et l'image de $\psi_1 \circ \varphi \circ \psi_2$ et celle de $\psi_1 \circ \varphi$ sont égales, donc

$$\mathrm{rang}(\psi_1 \circ \varphi \circ \psi_2) = \mathrm{rang}(\psi_1 \circ \varphi) = \mathrm{rang}(\varphi) .$$

□

Théorème 3.28. *Soit $\mathbb{K}$ un corps $n, m \in \mathbb{N}^*$ et $A, B \in M_{n,m}(\mathbb{K})$ deux matrices. Alors* $\mathrm{rang}(A) = \mathrm{rang}(B)$ *si, et seulement si, A et B sont équivalentes.*

Démonstration. On vient de voir une implication ; si $\mathrm{rang}(A) = \mathrm{rang}(B) = 0$ alors A et B sont nulles et donc équivalentes. Il nous suffit maintenant de montrer que, si $\mathrm{rang}(A) = r \geqslant 1$ alors A est équivalente C de $M_{n,m}(\mathbb{K})$ de coefficients

$$c_{i,j} = \begin{cases} 1 & \text{si } i = j \leqslant r, \\ 0 & \text{sinon.} \end{cases}$$

Pour cela, considérons l'application linéaire φ dont la matrice dans les bases canoniques $B_m = (e_1, \dots, e_m)$, $B_n = (f_1, \dots, f_n)$ de $\mathbb{K}^m$ et $\mathbb{K}^n$ est égale à A. Comme $\operatorname{rang}(\varphi) = r$, on sait qu'il existe $(e_{i_1}, \dots, e_{i_r})$ tels que $\big(\varphi(e_{i_1}), \dots, \varphi(e_{i_r})\big)$ forment une base de $\operatorname{Im}(\varphi)$. En notant $f'_j = \varphi(e_{i_j})$, on peut compléter (si $r < n$; sinon on ne fait rien) la famille $(f'_1, \dots, f'_r)$ en une base B'_n $(f'_1, \dots, f'_n)$ de $\mathbb{K}^n$.

En notant E le sous-espace vectoriel engendré par $(e_{i_1}, \dots, e_{i_r})$, on sait que φ est injective sur E (elle envoie une base de E sur une famille libre) donc $E \cap \ker(\varphi) = \{0\}$. Comme de plus $\dim(E) + \dim\big(\ker(\varphi)\big) = \operatorname{rang}(\varphi) + \dim\big(\ker(\varphi)\big) = m$, on en déduit que E et $\ker(\varphi)$ sont supplémentaires. Si $\ker(\varphi) \neq \{0\}$, on peut donc compléter $(e_{i_1}, \dots, e_{i_r}) = (e'_1, \dots, e'_r)$ en une base $B'_m = (e'_1, \dots, e'_m)$ de $\mathbb{K}^m$. Par construction, on a $\varphi(e'_i) = f'_i$ si $1 \leqslant i \leqslant r$, et $\varphi(e'_i) = 0$ sinon. Autrement dit, la matrice de φ dans les bases B'_m, B'_n est égale à C ; si on note P la matrice de passage de B_m à B'_m et Q la matrice de passage de B_n à B'_n, on vient de montrer que $C = Q^{-1}AP$, autrement dit A et C sont équivalentes.

□

Proposition 3.29. *Soit $\mathbb{K}$ un corps, $n, m \in \mathbb{N}^*$ et $A \in M_{n,m}(\mathbb{K})$. Alors le rang de A et celui de A^t sont égaux.*

Démonstration. Soit $r = \operatorname{rang}(A)$. En notant comme ci-dessus C la matrice de $M_{n,m}(\mathbb{K})$ de coefficients

$$c_{i,j} = \begin{cases} 1 & \text{si } i = j \leqslant r, \\ 0 & \text{sinon.} \end{cases}$$

(qui est la matrice nulle si $r = 0$) on a des matrices inversibles $P \in M_m(\mathbb{K})$ et $Q \in M_n(\mathbb{K})$ telles que $A = QCP$. En transposant, cela nous donne $A^t = P^t C^t Q^t$. Comme P^t et Q^t sont inversibles, on en déduit que A^t est équivalente à C^t ; et manifestement C^t est de rang r, donc $\operatorname{rang}(A^t) = r$. □

Exercice 3.30. Soit $n, m, p \in \mathbb{N}^*$, $\mathbb{K}$ un corps et $A \in M_{n,m}(\mathbb{K})$, $B \in M_{m,p}(\mathbb{K})$. Montrer que $\operatorname{rang}(AB) \leqslant \max\big(\operatorname{rang}(A), \operatorname{rang}(B)\big)$.

3.6 Trace d'une matrice et d'un endomorphisme

Définition 3.31. Soit $\mathbb{K}$ un corps, $n \in \mathbb{N}^*$ et $A = (a_{i,j})_{1 \leqslant i,j \leqslant n} \in M_n(\mathbb{K})$. La *trace* de A, notée $\operatorname{tr}(A)$, est le scalaire $\displaystyle\sum_{i=1}^{n} a_{i,i}$.

Proposition 3.32 (Circularité de la trace)**.** *Soit $n \in \mathbb{N}^*$ et $A, B \in M_n(\mathbb{K})$. Alors* $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

Démonstration. Le coefficient sur la i-ème ligne et la i-ième colonne de AB est égal à

$$\sum_{k=1}^{n} a_{i,k} b_{k,i}.$$

On en déduit l'égalité

$$\operatorname{tr}(AB) = \sum_{i=1}^{n} \left(\sum_{k=1}^{n} a_{i,k} b_{k,i} \right) .$$

En échangeant le rôle de A et B dans cette formule, on obtient

$$\operatorname{tr}(BA) = \sum_{i=1}^{n} \left(\sum_{k=1}^{n} b_{i,k} a_{k,i} \right) .$$

En permutant les deux sommes, on arrive bien à

$$\operatorname{tr}(BA) = \sum_{k=1}^{n} \left(\sum_{i=1}^{n} a_{k,i} b_{i,k} \right) = \sum_{i=1}^{n} \left(\sum_{k=1}^{n} a_{i,k} b_{k,i} \right) = \operatorname{tr}(AB) .$$

□

Proposition 3.33. *Soit $n \in \mathbb{N}^*$, $\mathbb{K}$ un corps et $A, B \in M_n(\mathbb{K})$ deux matrices semblables. Alors* $\operatorname{tr}(A) = \operatorname{tr}(B)$.

Démonstration. Par définition, il existe une matrice inversible $P \in GL_n(\mathbb{K})$ telle que $A = P^{-1}BP$. On peut alors écrire

$$\mathrm{tr}(A) = \mathrm{tr}\big(P^{-1}(BP)\big) = \mathrm{tr}((BP)P^{-1}) = \mathrm{tr}(B)\ .$$

□

Proposition 3.34. *Soit $\mathbb{K}$ un corps, E un espace vectoriel de dimension finie, et φ un endomorphisme de E. Alors pour toutes bases B, B' on a*

$$\mathrm{tr}\big(\mathrm{Mat}_B(\varphi)\big) = \mathrm{tr}\big(\mathrm{Mat}_{B'}(\varphi)\big)\ .$$

On appelle cette valeur la trace de φ et on la note $\mathrm{tr}(\varphi)$.

Démonstration. C'est une conséquence immédiate de la question précédente, puisque si P désigne la matrice de passage de B à B' on sait que $\mathrm{Mat}_{B'}(\varphi) = P^{-1}\,\mathrm{Mat}_B(\varphi)P$. □

Remarque 3.35. Nous avons donc deux invariants de matrices, à côté bien-sûr des dimensions,

- *le rang*, un entier qui est caractéristique de deux matrices *équivalentes* A et B avec $A = QBP$, Q et P inversibles,
- et *la trace*, un scalaire, qui est un invariant partagé par deux matrices *semblables* avec $A = P^{-1}BP$, mais qui ne leur est pas caractéristique.

3.7 Sous-espaces vectoriels et endomorphismes de $\mathbb{R}^n$

On a vu plus haut que, pour tout sous-espace vectoriel E de dimension k dans $\mathbb{R}^n$, il existe une application linéaire surjective $\varphi\colon \mathbb{R}^n \to \mathbb{R}^{n-k}$ telle que $E = \ker(\varphi)$. En d'autre termes il existe une matrice $A \in M_{n-k,n}(\mathbb{R})$ de rang $n-k$ et telle que E est égal à l'ensemble

$$\{(x_1, \dots, x_n) \in \mathbb{R}^n \colon \forall i \in \{1, \dots, n-k\}\ \sum_{j=1}^{n} a_{i,j}x_j = 0\}\ .$$

On appelle l'identité ci-dessus une *équation cartésienne* de E ; vous avez déjà vu de telles équations. Par exemple, un espace vectoriel de dimension 1 dans $\mathbb{R}^2$ est une droite D, et en donner une équation cartésienne revient à trouver deux réels non nuls a, b tels que

$$D = \{(x, y) \in \mathbb{R}^2 \colon ax + by = 0\}.$$

Dans $\mathbb{R}^3$, il va nous falloir deux équations (non proportionnelles !) pour définir une droite D, qu'on décrira comme l'ensemble des solutions du système

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 &= 0, \\ a_{2,1}x_1 + a_{2,2}x_2 + a_{2,3}x_3 &= 0. \end{cases}$$

En général, un sous-espace de dimension k dans $\mathbb{R}^n$ peut donc se décrire comme l'ensemble des solutions de système de $n-k$ équations à n inconnues. Un cas particulier important est celui où il n'y a qu'une équation.

Définition 3.36. Soit $\mathbb{K}$ un corps et E un $\mathbb{K}$-espace vectoriel de dimension finie $\geqslant 1$. Un sous-espace vectoriel F de E est un *hyperplan* si $\dim(F) = \dim(E) - 1$.

Par exemple, un hyperplan de $\mathbb{R}^2$ est une droite ; un hyperplan de $\mathbb{R}^3$ est ce qu'on a l'habitude d'appeler un plan.

Finissons ce chapitre en décrivant les matrices d'applications linéaires bien connues (?) dans des bases bien choisies.

Définition 3.37. Soit E, F deux sous-espaces vectoriels de $\mathbb{R}^n$ tels que $E \oplus F = \mathbb{R}^n$. La *projection* sur E parallèlement à F est l'unique application linéaire p telle que $p(e) = e$ pour tout $e \in E$ et $p(f) = 0$ pour tout $f \in F$.

Si $E = \mathbb{R}^n$ et $F = \{0\}$, la projection définie ci-dessus est simplement l'identité ; si $E = \{0\}$ et $F = \mathbb{R}^n$, c'est l'application nulle. Dans le cas où ni E ni F n'est réduit à $\{0\}$, notons que, si $(e_1, \dots, e_k)$ est une base de E et $(f_{k+1}, \dots, f_n)$ est une base de F, alors $(e_1, \dots, e_k, f_{k+1}, \dots, f_n)$ est une base de $\mathbb{R}^n$, et la matrice de la projection sur E parallèlement à F dans cette base est la matrice diagonale qui s'écrit par blocs

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

Définition 3.38. Soit $n \in \mathbb{N}^*$, et p une application linéaire telle que $p^2 = p$. On dit que p est un *projecteur*.

Proposition 3.39. *Soit $n \in \mathbb{N}^*$, et p un projecteur de $\mathbb{R}^n$. Alors p est la projection sur $Im(p)$ parallèlement à* $\ker(p)$.

Démonstration. On doit en particulier montrer que $E = \mathrm{Im}(p) \oplus \ker(p)$. Si $x \in \ker(p) \cap \mathrm{Im}(p)$ alors on a à la fois $p(x) = 0$ et $x = p(y)$ pour un certain $y \in \mathbb{R}^n$, donc

$$0 = p(x) = p\big(p(y)\big) = p(y) = x \ .$$

On vient de montrer que $\ker(p) \cap \mathrm{Im}(p) = \{0\}$; comme $\dim\big(\ker(p)\big) + \dim\big(\mathrm{Im}(p)\big) = n$, ceci suffit à montrer que $\ker(p)$ et $\mathrm{Im}(p)$ sont supplémentaires.

On aurait tout aussi bien pu montrer directement que $\mathbb{R}^n = \ker(p) + \mathrm{Im}(p)$ puis que cette somme est directe : pour tout $x \in \mathbb{R}^n$ on peut écrire $x = \big(x - p(x)\big) + p(x)$, et $p\big(x - p(x)\big) = p(x) - p^2(x) = 0$. Donc x est la somme d'un élément de $\ker(p)$ et d'un élément de $\mathrm{Im}(p)$.

Ensuite, il est clair que $p(x) = 0$ pour tout $x \in \ker(p)$; et pour tout $x \in \mathrm{Im}(p)$ on a déjà vu que $p(x) = x$. Donc p est bien la projection sur $\mathrm{Im}(p)$ parallèlement à $\ker(p)$. □

Exercice 3.40. Montrer que la trace d'un projecteur est égale à son rang.

Ces projecteurs sont particulièrement utiles dans le cas de sommes directes $E \oplus F = G$, alors $\mathrm{id}_G = \pi_E + \pi_F$, avec π_E le projecteur sur E parallèlement à F, respectivement π_F sur F parallèlement à E, décomposent tout vecteur $u \in G$ en la somme $u = \pi_E(u) + \pi_F(u)$ où $\pi_E(u) \in E$ est la composante suivant le sous-espace vectoriel E et $\pi_F(u) \in F$ la composante suivant le sous-espace vectoriel F.

Une base $(e_1, \dots, e_n)$ d'un espace vectoriel E permet de décomposer tout vecteur suivant cette base à l'aide des projections π_i sur $\mathrm{Vect}(e_i)$ parallèlement à $\mathrm{Vect}(e_j, j \neq i)$. Ainsi, tout vecteur $u \in E$ s'écrit $u = \sum_{i=1}^n \pi_i(u)$ avec $\pi_i(u) = \lambda_i e_i$ et les coordonnées de u dans la base sont données par ces nombres $(\lambda_i)_{1 \leqslant i \leqslant n}$.

Définition 3.41. Soit E, F deux sous-espaces vectoriels tels que $E \oplus F = \mathbb{R}^n$. La *symétrie* par rapport à E parallèlement à F est l'unique application linéaire S telle que $S(e) = e$ pour tout $e \in E$ et $S(f) = -f$ pour tout $f \in F$.

Si $E = \mathbb{R}^n$ et $F = \{0\}$, la symétrie définie ci-dessus est simplement l'identité ; si $E = \{0\}$ et $F = \mathbb{R}^n$, c'est $-id$. Dans le cas où ni E ni F n'est réduit à $\{0\}$, on a comme précédemment que si $(e_1, \dots, e_k)$ est une base de E et $(f_{k+1}, \dots, f_n)$ est une base de F, alors $(e_1, \dots, e_k, f_{k+1}, \dots, f_n)$ est une base de $\mathbb{R}^n$, et la matrice de la symétrie sur E parallèlement à F dans cette base est la matrice diagonale qui s'écrit par blocs

$$\begin{pmatrix} I_k & 0 \\ 0 & -I_{n-k} \end{pmatrix} .$$

Proposition 3.42. *Soit S un endomorphisme de $\mathbb{R}^n$ tel que $S^2 = id$. On dit que S est une* symétrie *et S est la symétrie par rapport à* $\ker(S - id)$ *parallèlement à* $\ker(S + id)$.

Démonstration. Soit $E = \ker(S - id)$ et $F = \ker(S + id)$. Par définition, $S(e) = e$ pour tout $e \in E$ et $S(f) = -f$ pour tout $f \in F$, donc on doit simplement montrer que E et F sont supplémentaires. Pour cela, considérons $x \in E \cap F$: on a alors à la fois $S(x) = x$ puisque $x \in E$, et $S(x) = -x$ puisque $x \in F$, donc $x = -x$ et par conséquent $x = 0$.

Pour montrer que $E + F = \mathbb{R}^n$, on remarque que pour tout $x \in \mathbb{R}^n$ on a $x - S(x) \in F$ et $x + S(x) \in E$: en effet,

$$S(x - S(x)) = S(x) - S^2(x) = S(x) - x = -(x - S(x)) \quad \text{et} \quad S(x + S(x)) = S(x) + S^2(x) = S(x) + x \ .$$

A partir de l'égalité

$$x = \frac{1}{2}\left((x + S(x)) + (x - S(x))\right)$$

on conclut donc que tout élément de $\mathbb{R}^n$ appartient à $E + F$, donc $E \oplus F = \mathbb{R}^n$. □

Exercice 3.43. Montrer que si p est un projecteur de $\mathbb{R}^n$ alors $x \mapsto 2p(x) - x$ est la symétrie par rapport à $\mathrm{Im}(p)$ parallèlement à $\ker(p)$.

Montrer que si S est une symétrie de $\mathbb{R}^n$ alors $p = \frac{S+id}{2}$ est la projection sur $\ker(S - id)$ parallèlement à $\ker(S + id)$.

En plus des projections et des symétries, rappelons la forme d'une matrice de *rotation* dans $\mathbb{R}^2$ muni de la base orthonormée directe canonique.

Dans $\mathbb{R}^2$, la rotation d'angle θ envoie le vecteur $(1, 0)$ sur $(\cos(\theta), \sin(\theta))$ et le vecteur $(0, 1)$ sur $(-\sin(\theta), \cos(\theta))$; sa matrice dans la base canonique est donc égale à $\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$.

Attention, ces notions n'ont de sens que dans un repère orthonormé direct, ce qui est le cas de la base canonique. Plus généralement, pour décrire les matrices de rotations dans $\mathbb{R}^3$, il nous faudrait introduire la notion d'orthogonalité de vecteurs de $\mathbb{R}^3$, ce qu'on ne va pas faire ici.

Chapitre 4

Fractions rationnelles

Dans ce chapitre, $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.

4.1 Résultats théoriques

Définition 4.1. Une *fraction rationnelle* est une fraction de la forme $F(X) = \dfrac{P(X)}{Q(X)}$, où $P, Q \in \mathbb{K}[X]$ et $Q \neq 0$.

C'est-à-dire une classe d'équivalence de couples de polynomes et on écrit $\dfrac{P_1}{Q_1} = \dfrac{P_2}{Q_2}$ quand $P_1Q_2 = P_2Q_1$, exactement comme quand on travaille avec les rationnels. On peut additionner et multiplier des fractions rationnelles :

$$\frac{P_1}{Q_1} + \frac{P_2}{Q_2} = \frac{P_1Q_2 + P_2Q_1}{Q_1Q_2} \quad \text{et} \quad \frac{P_1}{Q_1} \cdot \frac{P_2}{Q_2} = \frac{P_1P_2}{Q_1Q_2} \ .$$

On peut également les dériver, avec la formule

$$\left(\frac{P}{Q}\right)' = \frac{P'Q - Q'P}{Q^2} \ .$$

On note $\mathbb{K}(X)$ l'espace formé par les fractions rationnelles sur $\mathbb{K}$, qui est clairement un $\mathbb{K}$-espace vectoriel. Comme les rationnels de $\mathbb{Q}$ les fractions rationnelles ont une forme irréductible.

Définition 4.2. On dit que $F = \dfrac{P}{Q} \in \mathbb{K}(X)$ est sous *forme irréductible* si P et Q sont premiers entre eux, et Q est unitaire.

Autrement dit, comme on travaille sur $\mathbb{R}$ ou $\mathbb{C}$: $F = \dfrac{P}{Q}$ est sous forme irréductible si et seulement si Q est unitaire et P et Q n'ont pas de racine commune dans $\mathbb{C}$.

Une fraction rationnelle a une forme irréductible unique et on va principalement travailler avec ces formes irréductibles ; en pratique, il faut simplement s'habituer à simplifier le quotient $\dfrac{P}{Q}$ par d'éventuels diviseurs communs non constants de P et Q. Une fraction rationnelle a une forme irréductible unique.

Définition 4.3. Si $F = \dfrac{P}{Q} \in \mathbb{K}(X)$, on note $\deg(F) = \deg(P) - \deg(Q) \in \mathbb{Z} \cup \{-\infty\}$.

Par divison euclidienne, on peut écrire $\dfrac{P}{Q} = P_1 + \dfrac{P_2}{Q}$, où $P_1, P_2 \in \mathbb{K}[X]$ avec $\deg(P_2) < \deg(Q)$. Notons qu'une telle décomposition est unique. On l'appelle P_1 la *partie entière* de $\frac{P}{Q}$.

Notons que le degré d'une fraction rationnelle est bien défini : si $\dfrac{P_1}{Q_1} = \dfrac{P_2}{Q_2}$, alors on a $P_1Q_2 = P_2Q_1$, donc $\deg(P_1) + \deg(Q_2) = \deg(P_2) + \deg(Q_1)$ et, comme $\deg(Q_1), \deg(Q_2)$ sont des entiers puisque Q_1 et Q_2 sont non nuls, on en déduit $\deg(P_1) - \deg(Q_1) = \deg(P_2) - \deg(Q_2)$. De même, on montre que le degré du produit est la somme des degrés, $\deg(F \times G) = \deg F + \deg G$, et le degré est décroissant par combinaison linéaire : $\deg(\alpha F + \beta G) \leqslant \sup(\deg F, \deg G)$, pour $F, G \in \mathbb{K}(X)$, $\alpha, \beta \in \mathbb{K}$.

Définition 4.4. Soit $F = \dfrac{P}{Q}$ une fraction rationnelle irréductible dans $\mathbb{K}(X)$. Un *zéro* de F est un élément α de $\mathbb{K}$ tel que $P(\alpha) = 0$; et un *pôle* de F est un élément α de $\mathbb{K}$ tel que $Q(\alpha) = 0$.

Notons que F a un ensemble fini (éventuellement vide !) A de pôles ; et que F définit une fonction de $\mathbb{K} \setminus A$ dans $\mathbb{K}$. Deux fractions rationnelles qui définissent la même fonction doivent être égales : en effet, si $F = \dfrac{P_1}{Q_1}$ et $G = \dfrac{P_2}{Q_2}$ définissent la même fonction, alors pout tout $x \in \mathbb{K}$ qui n'est ni un pôle de F ni un pôle de G on a

$$\frac{P_1(x)}{Q_1(x)} = \frac{P_2(x)}{Q_2(x)} \quad \text{donc} \quad P_1(x)Q_2(x) - P_2(x)Q_1(x) = 0 \ .$$

En particulier, $P_1Q_2 - P_2Q_1$ doit avoir une infinité de racines [i], donc $P_1Q_2 - P_2Q_1 = 0$, autrement dit $F = G$.

Définition 4.5. Un *élément simple* est une fraction rationnelle de la forme $\dfrac{P}{Q^k}$, où Q est un polynôme unitaire irréductible de $\mathbb{K}[X]$, $k \geqslant 1$ et $\deg(P) < \deg(Q)$.

Comme on l'a dit précédemment, dans ce cours on ne s'intéresse qu'à $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Vous avez vu au premier semestre que les polynômes irréductibles sur $\mathbb{C}$ sont les polynômes de degré 1 ; et sur $\mathbb{R}$ les polynômes irréductibles sont les polynômes de degré 1 et les polynômes de degré 2 dont le discriminant est strictement négatif. Par conséquent, les éléments simples ont la forme suivante :

1. Sur $\mathbb{C}$, un élément simple est de la forme $\dfrac{a}{(X-\alpha)^k}$, où $a, \alpha \in \mathbb{C}$ et $k \in \mathbb{N}^*$.
2. Sur $\mathbb{R}$, un élément simple est de la forme $\dfrac{a}{(X-\alpha)^k}$, où $a, \alpha \in \mathbb{R}$ et $k \in \mathbb{N}^*$, ou bien de la forme $\dfrac{aX+b}{(X^2+\alpha X+\beta)^k}$, où $a, b, \alpha, \beta \in \mathbb{R}$ et $\alpha^2 - 4\beta < 0$ (de manière équivalente : le polynôme $X^2 + \alpha X + \beta$ n'a pas de racine réelle).

Notre but est de *décomposer une fraction rationnelle en éléments simples* ; on va expliquer pourquoi c'est possible de le faire dans $\mathbb{C}(X)$(en fait on peut le faire sur tout corps, et nos arguments ci-dessous s'adaptent au contexte général, mais ici on essaie d'éviter des énoncés trop généraux).

Proposition 4.6. *Soit $n \geqslant 1$, $Q \in \mathbb{C}_n[X]$ un polynôme unitaire, et*

$$Q(X) = (X-\alpha_1)^{n_1} \dots (X-\alpha_k)^{n_k}$$

la décomposition de Q en produit de facteurs irréductibles unitaires. Alors l'espace $E_Q = \{P/Q \colon P \in \mathbb{C}_{n-1}[X]\}$ est un sous-espace vectoriel de $\mathbb{C}(X)$ de dimension n, et les fractions rationnelles de la forme $\dfrac{1}{(X-\alpha_i)^j}$ pour $i \in \{1, \dots, k\}$ et $j \in \{1, \dots, n_i\}$ en forment une base.

Démonstration. Il est immédiat que $0 = \dfrac{0}{Q} \in E_Q$, et qu'une combinaison linéaire d'éléments de E_Q est un élément de E_Q, qui est donc bien un sous-espace vectoriel de $\mathbb{C}(X)$. Il est tout aussi clair que la famille $\left(\dfrac{1}{Q}, \dots, \dfrac{X^{n-1}}{Q}\right)$ en forme une base et donc E_Q est de dimension n.

La famille de E_Q formée par tous les $\dfrac{1}{(X-\alpha_i)^j}$ pour $i \in \{1, \dots, k\}$ et $j \in \{1, \dots, n_i\}$ a $n_1 + \dots + n_k = n$ éléments ; pour montrer que c'est une base il nous suffit donc de vérifier qu'elle est libre. Soit donc des coefficients $\lambda_{i,j}$ tels que

$$F(X) = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \lambda_{i,j} \frac{1}{(X-\alpha_i)^j} = 0 \ .$$

L'idée est qu'on peut retrouver les valeurs des $\lambda_{i,j}$ en estimant les valeurs des dérivées de

$$G_p(X) = (X-\alpha_p)^{n_p} F(X)$$

i. Ici il est important de se rappeler que les corps qu'on considère dans ce cours sont infinis : $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$...

en α_p ; en effet, on a pour tout $p \in \{1, \dots, k\}$:

$$G_p(X) = (X-\alpha_p)^{n_p} F(X) = \sum_{j=1}^{n_p} \lambda_{p,j} (X-\alpha_p)^{n_p-j} + (X-\alpha_p)^{n_p} \sum_{i=1, i\neq p}^{k} \sum_{j=1}^{n_k} \lambda_{i,j} \frac{1}{(X-\alpha_i)^j} \ .$$

On voit alors que $G_p^{(q)}(\alpha_p) = q! \lambda_{p,n_p-q}$ pour tout $p \in \{1, \dots, k\}$ et tout $q \in \{0, \dots, n_p - 1\}$, et donc de $F(X) = 0$ on tire $\lambda_{i,j} = 0$ pour tout i, j, montrant que la famille est libre. □

Théorème 4.7 (Décomposition en éléments simples dans $\mathbb{C}(X)$)**.** *Tout élément de $\mathbb{C}(X)$ admet une unique* décomposition en élements simples. *Plus précisément : si $F = \dfrac{P}{Q}$ est une fraction rationnelle irréductible dans $\mathbb{C}(X)$ et $Q(X) = (X-\alpha_1)^{n_1} \dots (X-\alpha_k)^{n_k}$, alors on peut écrire*

$$F(X) = P_1(X) + \sum_{i=1}^{k} \sum_{j=1}^{n_k} \frac{a_{i,j}}{(X-\alpha_i)^j} \ .$$

où P_1 est un polynôme et les $a_{i,j}$ sont des nombres complexes. De plus cette décomposition est unique.

Démonstration. On a déjà mentionné l'unicité de la partie entière d'un polynôme ; P_1 dans l'identité ci-dessus doit être la partie entière de F. Alors $F - P_1$ s'écrit sous la forme $\frac{P}{Q}$, où $\deg(P) < \deg(Q)$; autrement dit $F - P_1$ appartient à l'espace E_Q défini dans la proposition 4.6, et le résultat de cette proposition nous dit que $F - P_1$ s'écrit de manière unique sous la forme

$$F - P_1 = \sum_{i=1}^{k} \sum_{j=1}^{n_k} \frac{a_{i,j}}{(X-\alpha_i)^j} \ .$$

□

Notons que pour calculer les coefficients d'une décomposition en éléments simples, la méthode de la preuve de 4.6 nous dit qu'on peut toujours procéder en multipliant $F(X)$ par $(X-\alpha_i)^{n_i} F(X)$ et en évaluant en α_i les dérivées successives de $(X-\alpha_i)^{n_i} F(X)$. En pratique, cette méthode est souvent assez lourde en calculs et on essaiera si possible de faire plus simple.

On verra dans la prochaine section des exemples de calcul de décomposition en éléments simples ; pour l'instant, voyons à quoi ressemble cette décomposition dans $\mathbb{R}(X)$.

Théorème 4.8 (Décomposition en éléments simples dans $\mathbb{R}(X)$)**.** *Tout élément de $\mathbb{R}(X)$ admet une unique décomposition en élements simples. Plus précisément : si $F = \dfrac{P}{Q}$ est une fraction rationnelle irréductible dans $\mathbb{R}(X)$ et $Q(X) = (Q_1(X))^{n_1} \dots (Q_n(X))^{n_k}$ est la factorisation de Q en produits de polynômes irréductibles unitaires, alors on peut écrire*

$$F(X) = P_1(X) + \sum_{i=1}^{k} \sum_{j=1}^{n_k} \frac{P_{i,j}(X)}{(Q_i(X))^j} \ .$$

où P_1 est un polynôme et $\deg(P_{i,j}) < \deg(Q_i)$. De plus cette décomposition est unique.

Notons que chaque Q_i est de degré 1 ou 2, et donc chaque $P_{i,j}$ sera soit un polynôme constant soit de degré 1. On ne va pas donner ici la preuve de ce théorème ; l'argument d'algèbre linéaire qu'on a donné dans $\mathbb{C}(X)$ s'adapterait sans difficulté notable (ou on pourrait commencer par décomposer F en éléments simples dans $\mathbb{C}(X)$, et en déduire la décomposition dans $\mathbb{R}(X)$, comme on le verra sur un exemple dans la prochaine section).

L'unicité de la décomposition en élément simples nous permet aussi d'utiliser des symétries de la fonction pour calculer plus vite des coefficients.

4.2 Calculs pratiques

Insistons sur le fait qu'on peut utiliser des nombres complexes pour décomposer une fraction rationnelle en éléments simples sur $\mathbb{R}$! En particulier, si deux fractions rationnelles F_1, F_2 sont égales dans $\mathbb{R}(X)$ alors elles

sont égales dans $\mathbb{C}(X)$: on peut évaluer une fraction rationnelle en un nombre complexe pour la décomposer sur $\mathbb{R}$. Et si jamais on a décomposé $F = \dfrac{P}{Q} \in \mathbb{R}(X)$ en éléments simples dans $\mathbb{C}(X)$, alors on peut trouver la décomposition dans $\mathbb{R}(X)$ en regroupant deux à deux les éléments simples correspondant à une racine complexe non réelle de Q et à son conjugué.

Voyons un exemple simple : considérons la fraction rationnelle $F(X) = \dfrac{1}{X(X^2+1)}$. Sa décomposition en éléments simples dans $\mathbb{R}(X)$ sera de la forme

$$F(X) = \frac{a}{X} + \frac{bX+c}{X^2+1} \ .$$

Pour calculer a, on regarde $XF(X)$, qui vaut à la fois $\dfrac{1}{X^2+1}$ et $a + X\dfrac{bX+c}{X^2+1}$. En évaluant en 0, on obtient

$$\frac{1}{0^2+1} = a \text{ donc } a = 1 \ .$$

Et pour calculer b, c on regarde $(X^2+1)F(X)$, qu'on évalue en i, pour obtenir

$$\frac{1}{i} = bi + c \text{ donc } -i = bi + c \ .$$

Comme b, c sont réels, on en déduit $b = -1$ et $c = 0$. Finalement, on est arrivé à

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

On aurait pu procéder autrement pour calculer b, c : comme $F(X) = -F(-X)$, et que

$$\begin{aligned} F(X) &= \frac{a}{X} + \frac{bX+c}{X^2+1}, \\ -F(-X) &= \frac{a}{X} + \frac{bX-c}{X^2+1}. \end{aligned}$$

on déduit de l'unicité de la décomposition en éléments simples que $c = -c$, donc $c = 0$. Pour calculer a, la méthode la plus simple reste de multiplier par X et d'évaluer en 0 pour obtenir $a = 1$. Pour calculer b, outre l'évaluation en i, on pourrait évaluer en 1 et obtenir :

$$F(1) = \frac{1}{2} = a + \frac{b}{2} \text{ donc } b = -1 \ .$$

On aurait aussi pu calculer b en regardant la limite de $xF(x)$ quand $x \to +\infty$:

$$\lim_{x\to+\infty} xF(x) = \lim_{x\to+\infty} \frac{1}{x^2+1} = 0 \ ;$$

d'autre part, puisque $F(X) = \dfrac{a}{X} + \dfrac{bX}{X^2+1}$ on voit que la limite de $xF(x)$ en $+\infty$ est égale à $a+b$, d'où le fait que $a + b = 0$ et donc $b = -1$.

Encore une autre façon : remarquer que

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

pour obtenir directement la décomposition en éléments simples.

Finalement, on aurait pu passer par le calcul de la décomposition en élement simples dans $\mathbb{C}(X)$, qui doit être de la forme

$$F(X) = \frac{\alpha}{X} + \frac{\beta}{X-i} + \frac{\gamma}{X+i} \ .$$

En évaluant $XF(X)$ en 0, on trouve $\alpha = 1$. Puis en évaluant $(X-i)F(X) = \frac{1}{X(X+i)}$ en i, on obtient $\beta = \frac{1}{2i^2} = -\frac{1}{2}$. Pour calculer γ, on peut évaluer $(X+i)F(X)$ en $-i$; ou bien utiliser le fait que la fraction est réelle, égale à

sa conjuguée, $F(X) = \overline{F}(X)$ (où tous les scalaires complexes sont remplacés par leurs conjugués)) pour déduire que $\beta = \bar{\gamma}$, donc $\gamma = -\frac{1}{2}$. Ceci nous donne la décomposition en éléments simples dans $\mathbb{C}(X)$:

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

Et on peut se servir de cela pour retrouver la décomposition en élément simples sur $\mathbb{R}$:

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

On voit que la décomposition en élements simples sur $\mathbb{C}$ contient plus d'informations que la décomposition en élements simples sur $\mathbb{R}$: passer par le calcul de toute la décomposition dans $\mathbb{C}(X)$ pour en déduire la décomposition dans $\mathbb{R}(X)$ n'est pas en général la méthode la plus efficace...

Le bilan à tirer de cet exemple : une fois qu'on a correctement identifié la forme a priori de la décomposition en éléments simples d'une fraction rationnelle, tous les coups sont permis pour identifier les coefficients : calcul de dérivées, utilisation de symétries de la fonction, utilisation de limites, évaluation en un point bien choisi... Réduire au même dénominateur, identifier les coefficients pour obtenir un système et résoudre le système correspondant n'est pas recommandé !

Un autre exemple : décomposons sur $\mathbb{R}(X)$ la fraction rationnelle

$$F(X) = \frac{4X}{(X-1)^2(X^2+1)^2} \ .$$

On voit dès le départ que la décomposition va être pénible, à cause des deux termes au carré dans le dénominateur. Le numérateur est de degré strictement inférieur au numérateur, qui est déjà factorisé, et on écrit donc :

$$F(X) = \frac{a}{X-1} + \frac{b}{(X-1)^2} + \frac{cX+d}{(X^2+1)} + \frac{eX+f}{(X^2+1)^2} \ ,$$

où a, b, c, d, e, f sont 6 coefficients réels à déterminer.

Pour calculer b, on forme

$$G(X) = (X-1)^2 F(X) = \frac{4X}{(X^2+1)^2}$$

et on évalue en 1 : $b = \dfrac{4}{4} = 1$. De même pour évaluer e et f, on multiplie par $(X^2+1)^2$ et on évalue en i, ce qui donne

$$ei + f = \frac{4i}{(i-1)^2} = \frac{4i}{-2i} = -2.$$

Comme e, f sont réels, on déduit que $e = 0$ et $f = -2$.

Les autres coefficients sont plus compliqués à calculer ; on essaie d'extraire des informations en faisant le moins de calculs possible.

Par exemple, on peut évaluer en 0 pour obtenir

$$0 = -a + b + d + f \quad \text{donc } d - a = 1 \ .$$

Ou regarder la limite de $xF(x)$ quand x tend vers $+\infty$ pour obtenir

$$0 = a + c \ .$$

On voit que, si on arrive à calculer a, on aura tous les autres coefficients : il nous manque juste une information pour conclure. Si on aime calculer des dérivées, on utilise le fait que $a = G'(1)$ et que

$$G'(X) = \frac{4(X^2+1)^2 - 4X \times 4X(X^2+1)}{(X^2+1)^4} \quad \text{donc } G'(1) = -1 \ .$$

Ceci nous donne $a = -1$, donc $d = 0$ et $c = 1$.

On a fini notre calcul ; une autre méthode si on est sujet aux erreurs lors des calculs de dérivées : on évalue en un point ne donnant pas des calculs trop compliqués, par exemple ici -1, pour obtenir

$$-\frac{1}{2}a + \frac{1}{4}b + \frac{-c+d}{2} + \frac{-e+f}{4} = \frac{-4}{4.4} = -\frac{1}{4} \ .$$

En remplaçant b, e, f par leurs valeurs, c par $-a$ et d par $1+a$ on obtient

$$a\left(-\frac{1}{2} + \frac{1}{2} + \frac{1}{2}\right) + \frac{1}{4} + \frac{1}{2} - \frac{1}{2} = -\frac{1}{4}$$

donc

$$\frac{1}{2}a = -\frac{1}{2} \text{ d'où } a = -1 \ .$$

Un risque avec le fait qui consiste à utiliser des évaluations en des points où les calculs sont simples est qu'on peut obtenir des informations redondantes : par exemple, si au lieu d'évaluer en -1 on avait évalué en 0, on aurait obtenu $F(0) = 0 = -a + b + d + f = -a + 1 + 1 + a - 2$, ou encore $0 = 0$, ce qui n'est pas très utile !

Au final, on a obtenu (de haute lutte) l'identité

$$\frac{4X}{(X-1)^2(X^2+1)^2} = \frac{-1}{X-1} + \frac{1}{(X-1)^2} + \frac{X}{X^2+1} - \frac{2}{(X^2+1)^2} \ .$$

Deuxième partie

Analyse

Chapitre 5

Bref retour sur les nombres réels

Les notions de ce chapitre, et certaines des notions du chapitre suivant, ont été vues au premier semestre ; n'hésitez pas à reprendre vos notes de cours pour compléter ce cours-ci.

5.1 Majorants, minorants ; borne sup, borne inf

Définition 5.1. Soit A une partie de $\mathbb{R}$. On dit que $M \in \mathbb{R}$ est un *majorant* de A si

$$\forall x \in A,\ x \leqslant M\ .$$

Et on dit que m est un *minorant* de A si

$$\forall x \in A,\ m \leqslant x\ .$$

Définition 5.2. Soit A une partie de $\mathbb{R}$, et $M \in \mathbb{R}$.

1. On dit que M est la *borne supérieure* de A si :
   (a) M est un majorant de A.
   (b) Pour tout majorant M' de A, on a $M \leqslant M'$.
2. On dit que m est la *borne inférieure* de A si :
   (a) m est un minorant de A.
   (b) Pour tout minorant m' de A, on a $m' \leqslant m$.

Autrement dit : la borne supérieure est le plus petit des majorants ; la borne inférieure est le plus grand des minorants. Borne supérieure et borne inférieure sont intimement liées, comme le montre l'exercice suivant.

Exercice 5.3. Soit A une partie de $\mathbb{R}$ admettant une borne supérieure M. Montrer que $-A$ admet $-M$ comme borne inférieure.

La propriété fondamentale de $\mathbb{R}$, sur laquelle sont basées toutes les preuves des théorèmes d'analyse que vous verrez cette année, est la suivante :

Toute partie de $\mathbb{R}$ non vide et majorée admet une borne supérieure.

(Et bien sûr, cela entraîne que toute partie de $\mathbb{R}$ non vide et minorée admet une borne inférieure).

On appelle cette propriété l'*axiome de la borne supérieure*. Au premier semestre, vous avez par exemple utilisé cette propriété pour montrer que $\mathbb{Q}$ est dense dans $\mathbb{R}$: dans tout intervalle ouvert non vide de $\mathbb{R}$, il y a un nombre rationnel.

On va se servir de l'axiome de la borne supérieure pour retrouver quelques propriétés des suites vues au premier semestre. Pour cela, on aura besoin de la caractérisation suivante.

Proposition 5.4. *Soit $A \subset \mathbb{R}$ une partie de $\mathbb{R}$. Alors M est la borne supérieure de A si, et seulement si, M est un majorant de A et pour tout $\varepsilon > 0$ il existe $a \in A$ tel que $M - \varepsilon < a$.*

Démonstration. Supposons que M est la borne supérieure de A. Alors M est un majorant de A ; de plus, pour tout $\varepsilon > 0$ $M - \varepsilon < M$ donc $M - \varepsilon$ ne peut être un majorant de A, par conséquent il existe $a \in A$ tel que $M - \varepsilon < a$.

Réciproquement, si M est un majorant de A ayant la propriété ci-dessus, montrons que aucun $x < M$ ne peut être un majorant de A : si $x < M$ il existe $\varepsilon > 0$ tel que $x < M - \varepsilon < M$, et on peut par hypothèse trouver $a \in A$ tel que $M - \varepsilon \leqslant a$, en particulier il existe $a \in A$ tel que $x < a$ et donc x n'est pas un majorant de A. Par conséquent M est le plus petit des majorants de A, autrement dit, la borne supérieure de A. □

La borne supérieure M de A est donc caractérisée par

$$\begin{cases} \forall x \in A, x \leqslant M, & M \text{ est un majorant} ; \\ \forall \varepsilon > 0, \exists a \in A, M - \varepsilon < a, & M - \varepsilon \text{ n'est pas un majorant.} \end{cases}$$

A partir de maintenant, pour simplifier les notations, on va autoriser les sup et les inf à être infinis : autrement dit, si A est non vide et n'est pas majoré on écrira $\sup(A) = +\infty$ et si A n'est pas minoré on écrira $\inf(A) = -\infty$. Le seul cas où on évitera de parler de sup ou d'inf est quand $A = \emptyset$.

Proposition 5.5. *Soit A une partie de $\mathbb{R}$. Alors A est un intervalle si, et seulement si, pour tout $x, y, z \in \mathbb{R}$ on a l'implication*

$$(x \in A \text{ et } y \in A \text{ et } x < z < y) \Rightarrow z \in A \ .$$

Démonstration. Il n'y a rien à démontrer si $A = \emptyset$ (on convient que l'ensemble vide est bien un intervalle, par exemple l'intervalle ouvert $]0, 0[$). Clairement, les intervalles ont la propriété ci-dessus. Supposons que A est non vide et a cette propriété. Soit $a = \inf(A)$ et $b = \sup(A)$. Montrons que pour tout $z \in \mathbb{R}$ tel que $a < z < b$ on a $z \in A$: Soit donc z tel que $a < z < b$. Par définition d'une borne inférieure, z ne peut pas être un minorant de A : il doit donc exister $x \in A$ tel que $x < z$. De même, z ne peut pas être un majorant de A, ce qui nous donne $y \in A$ tel que $z < y$. On a alors $x < z < y$, et puisque $x, y \in A$ on en conclut que $z \in A$. Ceci nous montre que $]a, b[\subseteq A$, et alors on est dans un des quatre cas suivants :

1. $a \notin A$ et $b \notin A$; alors $A =]a, b|$.
2. $a \in A$ et $b \notin A$; alors $A = [a, b[$.
3. $a \notin A$ et $b \in A$; alors $A =]a, b]$.
4. $a \in A$ et $b \in A$; alors $A = [a, b]$ (et on dit que A est un segment).

□

Exercice 5.6. Soit I, J deux intervalles de $\mathbb{R}$ tels que $I \cap J \neq \emptyset$. Montrer que $I \cup J$ est un intervalle.

5.2 Suites convergentes. Suites extraites

Définition 5.7. Soit $(u_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ une suite de nombre réels, c'est-à-dire une valeur $u_n \in \mathbb{R}$ pour tout $n \in \mathbb{N}$. On note cette suite indifféremment u et $(u_k)_{k\in\mathbb{N}}$. On dit que u *converge vers* $l \in \mathbb{R}$ si

$$\forall \varepsilon > 0, \ \exists N \in \mathbb{N}, \ \forall n \geqslant N, \ |u_n - l| \leqslant \varepsilon \ .$$

Autrement dit : pour n suffisamment grand, u_n devient arbitrairement proche de l.

Exercice 5.8. Soit $(u_n)_{n\in\mathbb{N}}$, $(v_n)_{n\in\mathbb{N}}$ deux suites convergentes de nombres réels. Montrer que $u + v$ et uv sont des suites convergentes et exprimer leur limite en fonction de celles de u et de v.

On pourrait aussi parler de convergence vers $\pm\infty$; on dira que $(u_n)_{n\in\mathbb{N}}$ tend vers $+\infty$ si elle satisfait la condition suivante :

$$\forall M \in \mathbb{R} \ \exists \mathbb{N} \in \mathbb{N} \ \forall n \geqslant N u_n \geqslant M \ .$$

Exercice 5.9. Écrire avec des quantificateurs la définition de la phrase « $(u_n)_{n\in\mathbb{N}}$ tend vers $-\infty$ ».

Exercice 5.10. Soit A une partie non vide de $\mathbb{R}$. Montrer qu'il existe une suite d'éléments $(a_n)_{n\in\mathbb{N}}$ de A telle que $(a_n)_{n\in\mathbb{N}}$ tend vers $\sup(A)$, et une suite d'éléments $(b_n)_{n\in\mathbb{N}}$ de A telle que $(b_n)_{n\in\mathbb{N}}$ tend vers $\inf(A)$ (on rappelle qu'on autorise les possibilités $\sup(A) = +\infty$, $\inf(A) = -\infty$).

Définition 5.11. Une suite de nombre réels $(u_n)_{n\in\mathbb{N}}$ est *bornée* s'il existe $M \in \mathbb{R}$ tel que $|u_n| \leqslant M$ pour tout $n \in \mathbb{R}$.

Proposition 5.12. *Toute suite convergente de nombre réels est bornée.*

Démonstration. Soit $(u_n)_{n\in\mathbb{N}}$ une suite convergente de nombre réels. Il existe $N \in \mathbb{N}$ tel que pour tout $n \geqslant N$ on ait $|u_n - l| \leqslant 1$, autrement dit $u_n \in [l-1, l+1]$. Soit m le minimum des nombres $u_0, u_1, \ldots, u_N, l-1$, et M le maximum des nombres $u_0, u_1, \ldots, u_N, l+1$. Alors pour tout $n \in \mathbb{N}$ on a $m \leqslant u_n \leqslant M$. Par conséquent $(u_n)_{n\in\mathbb{N}}$ est bornée. □

La réciproque de cette proposition n'est pas vraie : il existe beaucoup de suites bornées non convergentes, par exemple la suite $((-1)^n)_{n\in\mathbb{N}}$.

Théorème 5.13. *Soit $(u_n)_{n\in\mathbb{N}}$ une suite croissante majorée de nombres réels. Alors $(u_n)_{n\in\mathbb{N}}$ est convergente.*

Démonstration. L'ensemble des valeurs prises par la suite $(u_n)_{n\in\mathbb{N}}$ est majoré ; soit M sa borne supérieure. Montrons que $(u_n)_{n\in\mathbb{N}}$ converge vers M. Pour cela, fixons $\varepsilon > 0$; d'après la proposition 5.4, il existe $N \in \mathbb{N}$ tel que $M - \varepsilon \leqslant u_N$. Comme $(u_n)_{n\in\mathbb{N}}$ est croissante, on doit aussi avoir $M - \varepsilon \leqslant u_n$ pour tout $n \geqslant N$; comme de plus $u_n \leqslant M$ pour tout $n \in \mathbb{N}$, on en déduit que pour tout $n \geqslant N$ on a $|u_n - M| = M - u_n \leqslant \varepsilon$, ce qui montre comme attendu que $(u_n)_{n\in\mathbb{N}}$ converge vers M. □

Définition 5.14. Soit $(u_n)_{n\in\mathbb{N}}$ une suite de nombres réels. Une *sous-suite*, ou *suite extraite*, de $(u_n)_{n\in\mathbb{N}}$ est une suite v obtenue en choisissant une suite $(n_k)_{k\in\mathbb{N}}$ strictement croissante d'entiers et en posant

$$\forall k \in \mathbb{N},\ v_k = u_{n_k}\ .$$

Rappelons que si $(n_k)_{k\in\mathbb{N}}$ est une suite strictement croissante d'entiers, alors $n_k \geqslant k$ pour tout k. En effet, $n_0 \in \mathbb{N}$ donc $n_0 \geqslant 0$; et si $n_k \geqslant k$ alors on a $n_{k+1} > n_k$ et donc $n_{k+1} > k$, d'où $n_{k+1} \geqslant k+1$. On conclut à l'aide du principe de récurrence.

Exercice 5.15. Montrer que, si $(u_n)_{n\in\mathbb{N}}$ est une suite convergente de nombres réels, alors toute suite extraite $(u_{n_k})_{k\in\mathbb{N}}$ converge, vers la même limite que $(u_n)_{n\in\mathbb{N}}$.

Vous avez-vu au premier semestre que de toute suite on peut extraire une sous-suite monotone ; en particulier, on a le résultat suivant.

Théorème 5.16 (Théorème de Bolzano–Weierstrass)**.** *Soit $(u_n)_{n\in\mathbb{N}}$ une suite bornée de nombre réels. Alors $(u_n)_{n\in\mathbb{N}}$ admet une sous-suite convergente.*

Démonstration. La suite $(u_n)_{n\in\mathbb{N}}$ admet une sous-suite monotone, qui est bornée puisque $(u_n)_{n\in\mathbb{N}}$ l'est, et donc convergente en tant que suite monotone et bornée. □

Exercice 5.17. Soit $(u_n)_{n\in\mathbb{N}}$ une suite de nombre réels. Montrer que :

1. Si $(u_{2n})_{n\in\mathbb{N}}$ et $(u_{2n+1})_{n\in\mathbb{N}}$ convergent toutes les deux vers une même limite l, alors $(u_n)_{n\in\mathbb{N}}$ converge vers l.
2. Si $(u_{2n})_{n\in\mathbb{N}}$, $(u_{2n+1})_{n\in\mathbb{N}}$ et $(u_{3n})_{n\in\mathbb{N}}$ sont toutes convergentes alors $(u_n)_{n\in\mathbb{N}}$ converge.

Proposition 5.18. *Soit $(x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}}$ deux suites de nombres réels telles qu'il existe $M \in \mathbb{R}$ satisfaisant*

$$\forall n \in \mathbb{N},\ |x_n| \leqslant M \text{ et } |y_n| \leqslant M\ .$$

Alors il existe une suite strictement croissante d'entiers $(n_k)_{k\in\mathbb{N}}$ telles que les suites $(x_{n_k})_{k\in\mathbb{N}}$ et $(y_{n_k})_{k\in\mathbb{N}}$ convergent toutes les deux.

Démonstration. La suite $(x_n)_{n\in\mathbb{N}}$ est une suite bornée de nombres réels : par le théorème de Bolzano–Weierstrass, on peut en extraire une suite convergente $(x_{a_k})_{k\in\mathbb{N}}$, où $(a_k)_{k\in\mathbb{N}}$ est une suite strictement croissante d'entiers. La suite $(y_{a_k})_{k\in\mathbb{N}}$ est une suite bornée de nombre réels, donc on peut en extraire une sous-suite convergente $(y_{a_{b_k}})_{k\in\mathbb{N}}$. Si on pose, pour tout k, $n_k = a_{b_k}$, alors $(n_k)_{k\in\mathbb{N}}$ est une suite strictement croissante d'entiers, et

- $(x_{n_k})_{k\in\mathbb{N}}$ est une suite extraite de la suite $(x_{a_k})_{k\in\mathbb{N}}$ qui est convergente ; donc $(x_{n_k})_{k\in\mathbb{N}}$ converge.
- $(y_{n_k})_{k\in\mathbb{N}}$ est convergente.

□

Voyez-vous pourquoi, dans le raisonnement ci-dessus, on a d'abord extrait une sous-suite de (x_n), avant d'en réextraire une nouvelle suite ?

5.3 Limite d'une fonction en un point

Si A est une partie de $\mathbb{R}$, on dira que $x \in \mathbb{R}$ est *adhérent à* A s'il existe une suite d'éléments de A qui converge vers x.

Exercice 5.19. Soit $A \subseteq \mathbb{R}$ et $x \in \mathbb{R}$. Montrer que $x \in \mathbb{R}$ est adhérent à A si, et seulement si

$$\forall \varepsilon > 0,\ \exists a \in A,\ |x - a| \leqslant \varepsilon\ .$$

Définition 5.20. Soit A une partie de $\mathbb{R}$, $f\colon A \to \mathbb{R}$ une fonction et $x \in \mathbb{R}$ un point adhérent à A. On dit que f admet l pour limite en x si la condition suivante est satisfaite :

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall a \in A,\ |a - x| \leqslant \delta \Rightarrow |f(a) - l| \leqslant \varepsilon\ .$$

On note alors $\lim\limits_{a \to x} f(a) = l$.

On dit que f tend vers $+\infty$ en x si

$$\forall M \in \mathbb{R},\ \exists \delta > 0,\ \forall a \in A,\ |a - x| \leqslant \delta \Rightarrow f(a) \geqslant M\ .$$

Enfin, f tend vers $-\infty$ en x si $-f$ tend vers $+\infty$, ou encore

$$\forall M \in \mathbb{R},\ \exists \delta > 0,\ \forall a \in A,\ |a - x| \leqslant \delta \Rightarrow f(a) \leqslant M\ .$$

Notons que, dans la définition ci-dessus, on ne suppose pas a priori que x appartient à A ; et si on change le domaine de définition de f (c'est-à-dire, si on change A) alors l'existence de la limite peut être affectée. Par exemple, si on considère la fonction $f\colon \mathbb{R} \to \mathbb{R}$ définie par $f(0) = 1$ et $f(x) = 0$ pour tout $x \neq 0$, alors f n'admet pas de limite en 0 ; mais si on considère la restriction de f à $\mathbb{R} \setminus \{0\}$, alors elle admet une limite en 0, qui vaut 0.

L'existence d'une limite pour une fonction f en un point se traduit en une propriété de l'image par f des suites qui convergent vers ce point.

Exercice 5.21. Soit A une partie de $\mathbb{R}$, $f\colon A \to \mathbb{R}$ une fonction, $x \in \mathbb{R}$ un point adhérent à A et $l \in \mathbb{R} \cup \{-\infty, +\infty\}$. Montrer que $\lim\limits_{a \to x} f(a) = l$ si, et seulement si, pour toute suite (a_n) d'éléments de A qui converge vers x on a $\lim\limits_{n \to +\infty} f(a_n) = l$.

En particulier : s'il existe une suite (a_n) d'éléments de A qui tend vers x et telle que $(f(a_n))$ n'est pas convergente alors f n'a pas de limite en x !

Parfois on se contente de regarder si f a une limite à gauche ou à droite en x (c'est surtout intéressant quand A est un intervalle ; en pratique, nous allons nous concentrer sur le cas de fonctions définies sur des intervalles de $\mathbb{R}$.)

Définition 5.22. Soit A une partie de $\mathbb{R}$, $f\colon A \to \mathbb{R}$ une fonction et $x \in \mathbb{R}$ un point adhérent à A. On dit que f admet l pour limite à droite en x si la condition suivante est satisfaite :

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall a \in A,\ 0 \leqslant a - x \leqslant \delta \Rightarrow |f(a) - l| \leqslant \varepsilon\ .$$

On note alors $\lim\limits_{a \to x^+} f(a) = l$.

De même, on dit que f admet l pour limite à gauche en x si la condition suivante est satisfaite :

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall a \in A,\ 0 \leqslant x - a \leqslant \delta \Rightarrow |f(a) - l| \leqslant \varepsilon\ .$$

On note alors $\lim\limits_{a \to x^-} f(a) = l$.

On peut bien sûr aussi donner, comme ci-dessus, des définitions pour le fait d'avoir une limite à gauche ou à droite égale à $\pm\infty$, et on laisse le lecteur écrire cette définition ; de même pour la définition de la notion de limite d'une fonction en $\pm\infty$. Vous êtes invité(e) à relire vos notes de cours du premier semestre si vous ne vous sentez pas au point sur ces notions... Souvent c'est plus facile de raisonner avec des suites pour démontrer l'existence (ou l'inexistence) de limites de fonctions. Par exemple, démontrons le résultat suivant, que vous avez vu au premier semestre.

Proposition 5.23. *Soit A, B deux parties de $\mathbb{R}$, $f\colon A \to \mathbb{R}$ et $g\colon B \to \mathbb{R}$ deux fonctions et x un point adhérent à B. On suppose que $g(B) \subseteq A$, $\lim\limits_{b\to x} g(b) = L$ et $\lim\limits_{a\to L} f(a) = l$. Alors*

$$\lim_{b\to x} f \circ g(b) = l \ .$$

Notons que ci-dessus on autorise les valeurs $\pm\infty$ pour l et L.

Démonstration. Soit (b_n) une suite d'éléments de B qui converge vers x. Alors $g(b_n)$ est une suite d'éléments de A qui tend vers L puisque $\lim\limits_{b\to x} g(x) = L$; donc $f(g(b_n))$ tend vers l puisque $\lim\limits_{a\to L} f(a) = l$. On vient de montrer que, pour toute suite (b_n) d'éléments de B qui converge vers x la suite $f \circ g(b_n)$ tend vers l, autrement dit

$$\lim_{b\to x} f \circ g(x) = l \ .$$

□

Chapitre 6

Continuité et dérivabilité de fonctions réelles

6.1 Continuité : théorèmes fondamentaux

Définition 6.1. Soit I un intervalle de $\mathbb{R}$, et $x \in I$. On dit que f est *continue en x* si $\lim_{y\to x} f(y) = f(x)$.
On dit que f est continue sur I si elle est continue en x pour tout $x \in I$.
Avec des quantificateurs : f est continue en $x \in I$ si et seulement si la condition suivante est vérifiée :

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall y \in I,\ |x-y| \leqslant \delta \Rightarrow |f(x) - f(y)| \leqslant \varepsilon\ .$$

La première fois qu'on rencontre cette formulation, on ne la comprend pas, *c'est normal, tout va bien.* Voici une explication en bande-dessinée tout à fait pertinente sur le site Images des Mathématiques par Jean-Paul Mohsen qui présente cette définitions comme un jeu, ce qui est peu ou prou ce que fait cet exercice WIMS.

Comme vous l'avez vu au premier semestre, cette définition se reformule à l'aide de suites.

Proposition 6.2. *Soit I un intervalle de $\mathbb{R}$, et $x \in I$. Alors f est continue en x si, et seulement si, pour toute suite (x_n) d'éléments de I qui converge vers x on a $\lim_{n\to+\infty} f(x_n) = f(x)$.*

En particulier, si I est un intervalle de $\mathbb{R}$, $a \in I$, f est une fonction continue de $I \setminus \{a\}$ dans $\mathbb{R}$, et $f(x)$ admet une limite quand x tend vers a, alors on peut prolonger f en a en posant

$$f(a) = \lim_{x\to a, x\in I\setminus\{a\}} f(x)\ .$$

La fonction obtenue est continue, et on l'appelle *prolongement par continuité* de f à I.

Exercice 6.3. On reprend les notations ci-dessus. Montrer que le prolongement par continuité de f est bien une fonction continue sur I.

Vous avez vu le résultat suivant au premier semestre, et on laisse sa preuve en exercice.

Proposition 6.4. *Soit I, J deux intervalles de $\mathbb{R}$.*
- *Si $f, g\colon I \to \mathbb{R}$ sont des fonctions continues en $x \in I$ alors fg et $f+g$ sont continues en x. Si de plus g ne s'annule pas sur I alors $\dfrac{f}{g}$ est continue en x.*
- *Si $g\colon J \to \mathbb{R}$ est continue en $x \in J$, $g(J) \subseteq I$ et $f\colon I \to \mathbb{R}$ est continue en $g(x)$, alors $f \circ g$ est continue en x.*

Théorème 6.5 (Théorème des valeurs intermédiaires)**.** *Soit I un intervalle de $\mathbb{R}$ et $f\colon I \to \mathbb{R}$ une fonction continue. S'il existe $a, b \in I$ tels que $f(a) < 0$ et $f(b) > 0$ alors il existe $x \in I$ tel que $f(x) = 0$.*
Plus généralement : l'image d'un intervalle par une fonction continue est un intervalle.

Démonstration. Supposons par exemple qu'il existe $a < b$ tels que $f(a) < 0$ et $f(b) > 0$ (quitte à remplacer f par $-f$ on peut se ramener à ce cas). Alors par continuité de f en a on a $\delta > 0$ tel que $f(y) < 0$ pour tout $y \in [a, a+\delta[$; par conséquent

$$E = \{x > a \colon \forall y \in [a, x[,\ f(y) < 0\}$$

est non vide, et majoré par b : il admet une borne supérieure x. Il existe une suite strictement croissante (e_n) d'éléments de E tels que $f(e_n)$ tend vers x ; on a $f(e_n) < 0$ donc $f(x) \leqslant 0$. Si jamais $f(x) < 0$, alors comme précédemment on sait que par continuité on a $\delta > 0$ tel que $f(x+y) < 0$ pour tout $y < \delta$, par conséquent $x + \delta \in E$, ce qui contredit la définition de x. Donc $f(x) = 0$.

Pour conclure que l'image de I par f est un intervalle, supposons que $x < y$ appartiennent à $f(I)$, et soit $z \in \mathbb{R}$ tel que $x < z < y$. On peut trouver $a, b \in I$ tels que $f(a) = x$ et $f(b) = y$; par conséquent, la fonction continue $f - z$ prend une valeur < 0 en a, et > 0 en b : d'après ce qu'on vient de démontrer elle doit donc s'annuler sur I, par conséquent $f(z) \in I$. □

Corollaire 6.6 (Théorème de la bijection)**.** Soit $a < b$ deux réels et $f \colon [a, b] \to \mathbb{R}$. Si f est strictement monotone et continue, alors f réalise une bijection de $[a, b]$ sur l'intervalle fermé J d'extrémités $f(a)$ et $f(b)$. De plus la fonction réciproque f^{-1} est continue sur J.

Démonstration. Supposons par exemple f strictement croissante, et soit alors $J = [f(a), f(b)]$. Puisque $f(a), f(b)$ appartiennent tous deux à l'image de f, qui est un intervalle d'après le théorème des valeurs intermédiaires, on a $J \subseteq f([a, b])$. Réciproquement, si $x \in f([a, b])$ alors il existe $y \in [a, b]$ tel que $f(y) = x$, et puisque f est croissante on a $f(a) \leqslant x \leqslant f(b)$, donc $x \in J$.

On vient de montrer que $f([a, b]) = J$, autrement dit f réalise une surjection de $[a, b]$ sur J. Pour montrer qu'elle est aussi injective, considérons $x \neq y \in [a, b]$. Si $x < y$ alors $f(x) < f(y)$ puisque f est strictement croissante, de même si $y < x$ on a $f(y) < f(x)$. Donc si $x \neq y$ on doit avoir $f(x) \neq f(y)$, autrement dit f est injective.

Reste à montrer que f^{-1} est continue sur J ; soit (y_n) une suite d'élements de J qui converge vers $y \in J$, et $x_n = f^{-1}(y_n)$. Si (x_n) ne converge pas vers $x = f^{-1}(y)$, il existe une sous-suite (x_{n_k}) et $\varepsilon > 0$ tels que

$$\forall k \in \mathbb{N},\ |x_{n_k} - x| \geqslant \varepsilon\ .$$

Mais comme (x_{n_k}) est une suite d'élements du segment $[a, b]$, on peut en extraire une sous-suite qui converge vers $x' \in [a, b]$; ceci nous donne une sous-suite (x_{m_k}) de (x_n) qui converge vers x' et telle que

$$\forall k \in \mathbb{N},\ |x_{m_k} - x| \geqslant \varepsilon\ .$$

En particulier, $|x' - x| \geqslant \varepsilon$; mais comme f est continue on a

$$f(x') = \lim_{k \to +\infty} f(x_{m_k}) = \lim_{k \to +\infty} y_{m_k} = y$$

puisque (y_{m_k}) est une sous-suite de la suite (y_n) qui converge vers y. On a donc $f(x') = y = f(x)$ mais $x' \neq x$, ce qui est impossible puisque f est injective. Donc $f^{-1}(y_n)$ converge vers $f^{-1}(y)$, ce qui prouve que f^{-1} est continue sur J. □

La monotonie stricte nous a permis de garantir l'injectivité de f ; cette condition était en fait nécessaire (sous l'hypothèse qu'on travaille avec des fonctions continues !), comme le montre l'exercice suivant.

Exercice 6.7. Soit I un segment de $\mathbb{R}$ et $f \colon I \to \mathbb{R}$ une application injective et continue. Montrer que f est strictement monotone.

Théorème 6.8. *Soit f une fonction continue sur un segment. Alors f est bornée et atteint ses bornes.*

Rappelons qu'un *segment* est un intervalle $[a, b]$ avec $a \leqslant b \in \mathbb{R}$. En combinaison avec le théorème des valeurs intermédiaires, le théorème ci-dessus montre que l'image d'un segment par une application continue est encore un segment.

Démonstration. Soit $[a, b]$ un segment et $f \colon [a, b] \to \mathbb{R}$ une fonction continue sur $[a, b]$. On va d'abord montrer que f est majorée et admet un maximum, c'est-à-dire qu'il existe $x \in [a, b]$ tel que $f(y) \leqslant f(x)$ pour tout $y \in [a, b]$. Pour cela, notons

$$M = \sup\{f(x) \colon x \in [a, b]\} \in \mathbb{R} \cup \{+\infty\}$$

et $(x_n) \in [a,b]^{\mathbb{N}}$ une suite telle que $\big(f(x_n)\big)_{n\in\mathbb{N}}$ converge vers M. D'après le théorème de Bolzano-Weierstrass, il existe une sous-suite (x_{n_k}) de (x_n) qui converge, et sa limite x appartient encore à $[a,b]$. Par continuité de f en x, la suite $f(x_{n_k})$ converge vers $f(x)$; mais puisque c'est une sous-suite d'une suite convergeant vers M, elle converge également vers M : on a $f(x) = M$, ce qui montre à la fois que M est fini et qu'il est atteint.

Pour montrer que f est minorée et admet un minimum, on pourrait suivre le même raisonnement que précédemment en remplaçant $\sup\{f(x)\colon x \in [a,b]\} \in \mathbb{R} \cup \{+\infty\}$ par $\inf\{f(x)\colon x \in [a,b]\} \in \mathbb{R} \cup \{-\infty\}$; ou bien appliquer à $-f$ le résultat qu'on vient de démontrer. □

Revenons sur la définition d'une fonction continue avec des quantificateurs : f est continue sur I si et seulement si :

$$\forall \varepsilon > 0,\ \forall x \in I,\ \exists \delta > 0,\ \forall y \in I\ |x-y| \leqslant \delta \Rightarrow |f(x) - f(y)| \leqslant \varepsilon\ .$$

Le point crucial est que δ ci-dessus dépend à la fois de x et de ε. L'ordre des quantificateurs est important !

Exercice 6.9. Soit I un intervalle de $\mathbb{R}$.

1. Disons que $f\colon I \to \mathbb{R}$ est *vraiment très continue* si :

$$\exists \delta > 0,\ \forall \varepsilon > 0,\ \forall x \in I,\ \forall y \in I,\ |x-y| \leqslant \delta \Rightarrow |f(x) - f(y)| \leqslant \varepsilon\ .$$

   (autrement dit, le même δ marche pour tout ε et pour tout x).

2. Disons que $f\colon I \to \mathbb{R}$ est *assez peu continue* si :

$$\forall \varepsilon > 0,\ \forall x \in I,\ \forall y \in I,\ \exists \delta > 0,\ |x-y| \leqslant \delta \Rightarrow |f(x) - f(y)| \leqslant \varepsilon\ .$$

   (cette fois, δ dépend à la fois de ε, x, et y)

Montrer que les seules fonctions vraiment très continues sur I sont les fonctions constantes ; et que toute fonction de I dans $\mathbb{R}$ est assez peu continue.

Les définitions dans l'exercice ci-dessus ne sont bien sûr pas à retenir et ne sont là que pour illustrer le fait que l'ordre dans lequel les quantificateurs sont écrits est fondamental pour le sens d'un énoncé mathématique ; il faut s'habituer à y faire très attention.

Les énoncés de l'exercice ci-dessus ont été obtenus en bougeant le « $\exists\delta$ »de la définition d'une fonction continue tout à gauche de l'énoncé, ou tout à droite ; une possibilité intermédiaire existe et donne naissance, contrairement à ses consœurs, à une notion mathématique importante.

Définition 6.10. Soit I un intervalle de $\mathbb{R}$. Une fonction $f\colon I \to \mathbb{R}$ est *uniformément continue* si :

$$\forall \varepsilon > 0,\quad \exists \delta > 0,\ \forall x, y \in I,\ |x-y| \leqslant \delta \Rightarrow |f(x) - f(y)| \leqslant \varepsilon\ .$$

Autrement dit : δ ne dépend plus que de ε et plus du point x. Manifestement, cette définition est plus forte que celle de la continuité et est en fait, en général, strictement plus forte.

Exemple.

1. La fonction $x \mapsto x^2$ n'est pas uniformément continue sur $\mathbb{R}$. En effet, si elle l'était, alors (en prenant $\varepsilon = 1$ dans la définition de la continuité uniforme) on pourrait trouver $\delta > 0$ tel que pour tout $x \in \mathbb{R}$ on ait $(x+\delta)^2 - x^2 \leqslant 1$. Mais pour $x = \frac{1}{\delta}$ on a

$$(x+\delta)^2 - x^2 = 2\delta x + \delta^2 \geqslant 2\ .$$

2. De même, la fonction $x \mapsto \frac{1}{x}$ n'est pas uniformément continue sur $]0,1]$. En effet, si elle l'était, alors on pourrait comme expliqué ci-dessus trouver $\delta > 0$, qu'on peut supposer plus petit que $\frac{1}{2}$, tel que pour tout $x \in]0, \frac{1}{2}]$ on ait

$$\frac{1}{x} - \frac{1}{x+\delta} \leqslant 1\ .$$

   Mais pour $x = \delta$ on obtient

$$\frac{1}{x} - \frac{1}{x+\delta} = \frac{\delta}{x(x+\delta)} = \frac{1}{2\delta} \geqslant 1\ .$$

Exercice 6.11. Déterminer tous les polynômes $P \in \mathbb{R}[X]$ tels que $x \mapsto P(x)$ soit uniformément continue sur $\mathbb{R}$.

On voit que, sur un intervalle non borné, même des fonctions très sympathiques comme $x \mapsto x^2$ ne vont pas être uniformément continues ; et qu'une fonction qui a une limite infinie en 0 ne risque pas d'être uniformément continue sur $]0,1]$ (pourquoi ?). En utilisant les mêmes idées, on peut se convaincre que, quand l'intervalle I n'est pas un segment, il y a des fonctions continues sur I qui n'y sont pas uniformément continues. Remarquablement, cela n'arrive pas quand I est un segment.

Théorème 6.12 (Théorème de HEINE)**.** *Soit I un segment de $\mathbb{R}$ et $f\colon I \to \mathbb{R}$ une fonction continue. Alors f est uniformément continue sur I.*

Démonstration. On va démontrer la contraposée de l'implication ci-dessus, c'est-à-dire qu'on va supposer que f n'est pas uniformément continue sur I et en déduire qu'il existe $x \in [a,b]$ tel que f n'est pas continue en x. Dire que f n'est pas uniformément continue sur I, c'est affirmer que f a la propriété suivante :

$$\exists \varepsilon > 0,\ \forall \delta > 0,\ \exists x, y \in I,\ |x-y| \leqslant \delta \text{ et } |f(x)-f(y)| \geqslant \varepsilon\ .$$

Fixons ε témoignant que la propriété ci-dessus est vraie ; alors pour tout $n \in \mathbb{N}^*$ on peut trouver $x_n, y_n \in I$ tels que

$$|x_n - y_n| \leqslant \frac{1}{n} \text{ et } |f(x_n) - f(y_n)| \geqslant \varepsilon\ .$$

Comme I est un segment, (x_n) et (y_n) sont en particulier bornées et on a vu en 5.18 qu'il existe des sous-suites (x_{n_k}) et (y_{n_k}) qui convergent toutes les deux, disons vers x, y ; comme I est un segment x, y appartiennent à I. De plus, on sait que $n_k \geqslant k$ et donc $|x_{n_k} - y_{n_k}|$ tend vers 0 : $x = y$. Si f était continue en x, $(f(x_{n_k}))$ et $(f(y_{n_k}))$ devraient toutes deux converger vers $f(x)$; ce n'est pas le cas puisque $|f(x_{n_k}) - f(y_{n_k})| \geqslant \varepsilon$ ne tend pas vers 0. Par conséquent f n'est pas continue en x et le théorème de HEINE est démontré. □

6.2 Dérivabilité

Définition 6.13. Soit I un intervalle de $\mathbb{R}$, et $x \in I$. On dit que f est *dérivable* en x si

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

existe ; dans ce cas on note cette limite $f'(x)$.

On dit que f est dérivable sur I si elle est dérivable en x pour tout $x \in I$.

Notons que, dans la limite ci-dessus, on suppose implicitement que $y \in I$; ainsi, si $I = [a,b]$, quand on regarde si f est dérivable en a on ne regarde que des $y \geqslant a$, tandis que pour la dérivée en b on ne regarde que des $y \leqslant b$. Il est parfois plus naturel de ne se préoccuper que de la dérivées de fonctions définies sur un intervalle ouvert ; aux extrémités des intervalles on préférera souvent utiliser la notion de dérivée à gauche ou à droite.

Définition 6.14. Soit I un intervalle de $\mathbb{R}$, et $x \in I$. On dit que f est *dérivable à gauche* en x si

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

existe ; dans ce cas on note cette limite $f'_g(x)$.

De même, quand la limite

$$\lim_{y \to x^+} \frac{f(x) - f(y)}{x - y}$$

existe on dit que f est *dérivable à droite* et on note cette limite $f'_d(x)$.

Une autre façon de reformuler la définition de la dérivée : f est dérivable en x de dérivée l si et seulement si il existe une fonction $\varepsilon\colon I \setminus \{x\} \to \mathbb{R}$ telle que pour tout $y \in I$ on ait

$$f(y) = f(x) + l(x-y) + (x-y)\varepsilon(y) \text{ et } \lim_{y \to x} \varepsilon(y) = 0\ .$$

En particulier, $f(y)$ tend vers $f(x)$ quand y tend vers x : si f est dérivable en x alors elle est nécessairement continue en x. La réciproque est fausse : par exemple $x \mapsto |x|$ n'est pas dérivable en 0. Cela dit, elle y admet une dérivée à droite et une dérivée à gauche, ce n'est donc pas un exemple très convaincant ; en fait, il existe des fonctions continues sur $\mathbb{R}$ qui n'ont pas de dérivée à droite, ni à gauche, en tout point de $\mathbb{R}$...

Proposition 6.15. *Soit $n \in \mathbb{Z}$ et $g_n \colon x \mapsto x^n$ (définie sur $\mathbb{R}$ tout entier). Alors :*

1. *Si $n > 0$ alors g_n est dérivable sur $\mathbb{R}$ et pour tout $x \in \mathbb{R}$ on a $g_n'(x) = nx^{n-1}$.*
2. *Si $n < 0$ alors g_n est dérivable sur $\mathbb{R}^*$ et pour tout $x \in \mathbb{R}^*$ on a $g_n'(x) = nx^{n-1}$.*

Bien sûr, si $n = 0$ la fonction est constante et de dérivée nulle.

Démonstration. Commençons par le cas où $n > 0$. Soit $x \in \mathbb{R}$; d'après la formule du binôme de Newton, on a pour tout h :

$$\begin{aligned}(x+h)^n - x^n &= \left(\sum_{k=0}^{n} \binom{n}{k} h^k x^{n-k}\right) - x^n \\ &= nhx^{n-1} + \sum_{k=2}^{n} \binom{n}{k} h^k x^{n-k}\end{aligned}$$

En divisant par h et en faisant tendre h vers 0, on obtient comme attendu que

$$\lim_{h\to 0} \frac{(x+h)^n - x^n}{h} = nx^{n-1} \ .$$

Quand $n = -m$ avec $m > 0$, on considère $x \neq 0$, $h \in \mathbb{R}$ et on réduit au même dénominateur :

$$\begin{aligned}(x+h)^n - x^n &= \frac{1}{(x+h)^m} - \frac{1}{x^m} \\ &= \frac{x^m - (x+h)^m}{x^m (x+h)^m}\end{aligned}$$

Quand on divise par h et qu'on fait tendre h vers 0, on a vu que le numérateur tend vers $-mx^{m-1}$; et le dénominateur tend vers x^{2m}. On obtient donc l'égalité

$$\lim_{h\to 0} \frac{(x+h)^n - x^n}{h} = \frac{-mx^{m-1}}{x^{2m}} = -mx^{-m-1} = nx^{n-1} \ .$$

□

On retrouve les formules classiques sur la dérivation d'un produit ou d'une composée.

Proposition 6.16. *Soit I un intervalle de $\mathbb{R}$ et $f, g \colon I \to \mathbb{R}$ deux fonctions dérivables en $x \in I$. Alors fg est dérivable, et on a*

$$(fg)'(x) = f'(x)g(x) + f(x)g'(x) \ .$$

Par conséquent, si f et g sont dérivables sur I alors fg l'est aussi et $(fg)' = f'g + fg'$.

Démonstration. On a des fonctions $\varepsilon_1, \varepsilon_2 \colon I \to \mathbb{R}$ qui tendent vers 0 en x et telles que pour tout $y \in I$ on ait

$$f(y) = f(x) + (y-x)f'(x) + (y-x)\varepsilon_1(y) \text{ et } g(y) = g(x) + (y-x)g'(x) + (y-x)\varepsilon_2(y) \ .$$

On obtient donc pour tout $y \in Y$

$$f(y)g(y) = f(x)g(x) + (y-x)(f'(x)g(x) + f(x)g'(x)) + (y-x)\varepsilon(y) \ .$$

Ci-dessus, on a posé

$$\varepsilon(y) = f(x)\varepsilon_2(y) + (y-x)f'(x)g'(x) + (y-x)f'(x)\varepsilon_2(y) + \varepsilon_1(y)g(x) + (y-x)\varepsilon_1(y)g'(x) + (y-x)\varepsilon_1(y)\varepsilon_2(y) \ .$$

Cette fonction tend vers 0 quand y tend vers x et on a obtenu

$$f(y)g(y) = f(x)g(x) + (y-x)(f'(x)g(x) + f(x)g'(x)) + (y-x)\varepsilon(y) \ .$$

Ceci montre à la fois que fg est dérivable en x et que sa dérivée y est égale à $f'(x)g(x) + f(x)g'(x)$. □

Proposition 6.17. *Soit I, J deux intervalles de $\mathbb{R}$, $f\colon I \to \mathbb{R}$ et $g\colon J \to \mathbb{R}$ telle que $g(J) \subseteq J$. Si $x \in J$ est tel que g est dérivable en x et f est dérivable en $g(x)$, alors $f \circ g$ est dérivable en x et on a*

$$(f \circ g)'(x) = f'(g(x))g'(x) \ .$$

Démonstration. Comme ci-dessus, Il existe des fonctions $\varepsilon_1, \varepsilon_2\colon I \to \mathbb{R}$ telle que ε_1 tend vers 0 en $g(x)$, ε_2 tend vers 0 en x et telles que l'on ait

$$\forall y \in I,\ f(y) = f(g(x)) + (y - g(x))f'(g(x)) + (y - g(x))\varepsilon_1(y) \text{ et } \forall z \in J,\ g(z) = g(x) + (z - x)g'(x) + (z - x)\varepsilon_2(z) \ .$$

En particulier, on obtient que pour tout $z \in J$ on a

$$\begin{aligned} f(g(z)) &= f\left(g(x) + (z - x)g'(x) + (z - x)\varepsilon_2(z)\right) \\ &= f(g(x)) + (z - x)g'(x)f'(g(x)) + +(z - x)\varepsilon_2(z)g'(x) + (z - x)\varepsilon_1(g(z)) \\ &= f(g(x)) + (z - x)g'(x)f'(g(x)) + (z - x)\varepsilon(z) \end{aligned}$$

où ε tend vers 0 quand z tend vers x. □

Cette formule permet de retrouver les dérivées des fonctions composées f^n, $\cos(f)$, $\sin(f)$, $\ln(f)$, etc. que vous connaissez déjà.

Corollaire 6.18. Soit I un intervalle de $\mathbb{R}$, $f\colon I \to \mathbb{R}$ une fonction dérivable en un certain $x \in I$.

1. Pour $n \in \mathbb{N}$, f^n est dérivable en et on a $(f^n)'(x) = nf'(x)(f(x))^{n-1}$.
2. $\cos(f)$, $\sin(f)$ sont dérivables en x et $(\cos(f))' = -f'\sin(f)$, $(\sin(f))' = f'\cos(f)$.
3. Si $f(x) \neq 0$, $\dfrac{1}{f}$ est dérivable en x et $\left(\dfrac{1}{f}\right)'(x) = -\dfrac{f'(x)}{(f(x))^2}$.
4. Si $f(x) > 0$, $\ln(f)$ est dérivable en x et $(\ln(f))'(x) = \dfrac{f'(x)}{f(x)}$.

Démonstration. Montrons par exemple le troisième point ci-dessus. On sait que $g\colon y \mapsto \dfrac{1}{y}$ est dérivable en $f(x)$ puisque $f(x) \neq 0$, et on peut alors appliquer le théorème de dérivation des fonctions composées pour conclure que $\dfrac{1}{f} = g \circ f$ est dérivable en x, de dérivée égale à

$$g'(f(x))f'(x) = -\frac{1}{(f(x))^2}f(x) = -\frac{f'(x)}{(f(x))^2} \ .$$

□

Corollaire 6.19. Soit I un intervalle de $\mathbb{R}$ et $f, g\colon I \to \mathbb{R}$ deux fonctions dérivables en $x \in I$. Si $x \in I$ est tel que $g(x) \neq 0$, alors $\dfrac{f}{g}$ est dérivable en x et on a $\left(\dfrac{f}{g}\right)'(x) = \dfrac{f'(x)g(x) - f(x)g'(x)}{(g(x))^2}$.

Démonstration. Le quotient $\dfrac{f}{g}$ n'est rien d'autre que le produit de f et de $\dfrac{1}{g}$, qui sont toutes deux dérivables en x. Par conséquent, le théorème de dérivabilité des produits nous donne que $\dfrac{f}{g}$ est dérivable en x, de dérivée

$$\begin{aligned} \left(\frac{f}{g}\right)'(x) &= f'(x)\left(\frac{1}{g}\right)(x) + \left(\frac{1}{g}\right)'(x)f(x) \\ &= \frac{f'(x)}{g(x)} - \frac{g'(x)f(x)}{(g(x))^2} \\ &= \frac{f'(x)g(x) - f(x)g'(x)}{(g(x))^2} \end{aligned}$$

□

Parfois, pour décider si une fonction est dérivable en un point, on est obligé de revenir à la définition : considérons par exemple le cas de la fonction $f\colon \mathbb{R} \to \mathbb{R}$ définie par

$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right) & \text{si } x > 0 \\ 0 & \text{si } x = 0 \end{cases}$$

Notons que cette fonction est continue en 0 : comme $\left|\sin\left(\frac{1}{x}\right)\right| \leqslant 1$ pour tout $x > 0$, on a bien

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

Manifestement, f est dérivable sur $]0, +\infty[$ et sur $]-\infty, 0[$ et sa dérivée en un point $x \neq 0$ est égale à $f'(x) = 2x \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right)$. Notons que $f'(x)$ n'a pas de limite quand x tend vers 0^+ (pourquoi ?). Cela ne signifie pas, a priori, que f n'est pas dérivable en 0 ! Pour décider si elle est dérivable, on forme le taux d'accroissement :

$$\frac{f(x) - f(0)}{x - 0} = \frac{f(x)}{x} = x \sin\left(\frac{1}{x}\right) \ .$$

On conclut que $\dfrac{f(x) - f(0)}{x - 0}$ tend vers 0 quand x tend vers 0 : f est bien dérivable en 0, et $f'(0) = 0$. Mais f' n'est pas continue en 0.

6.3 Théorème de Rolle et des accroissements finis.

Définition 6.20. Soit I un intervalle de $\mathbb{R}$ et $f\colon I \to \mathbb{R}$ une fonction. On dit que $a \in I$ est un :

- *maximum* de f sur I si pour tout $x \in I$ on a $f(x) \leqslant f(a)$;
- *minimum* de f sur I si pour tout $x \in I$ on a $f(x) \geqslant f(a)$;
- *extremum* de f sur I si a est un minimum ou un maximum de f sur I.

Définition 6.21. Soit I un intervalle *ouvert* de $\mathbb{R}$, $f\colon I \to \mathbb{R}$ et $a \in I$. On dit que a est un :

- *maximum local* de f sur I s'il existe $\delta > 0$ tel que pour tout $x \in I$ on ait

$$|x - a| \leqslant \delta \Rightarrow f(x) \leqslant f(a) \ .$$

- *minimum local* de f sur I s'il existe $\delta > 0$ tel que pour tout $x \in I$ on ait

$$|x - a| \leqslant \delta \Rightarrow f(x) \geqslant f(a) \ .$$

- *extremum local* de f sur I si a est un minimum local ou un maximum local de f sur I.

Proposition 6.22. *Soit I un intervalle ouvert de $\mathbb{R}$, $f\colon I \to \mathbb{R}$ une fonction et x un extremum local de f. Si f est dérivable en x alors on doit avoir $f'(x) = 0$.*

Démonstration. Supposons que $x \in I$ est tel que $f'(x) \neq 0$, et montrons que x ne peut être un extremum local pour f. Supposons par exemple $f'(x) > 0$. Comme $f'(x)$ est la limite du taux d'accroissement $\dfrac{f(y) - f(x)}{y - x}$ quand y tend vers x, celui-ci doit être > 0 pour y suffisamment proche de x, autrement dit :

$$\exists \delta > 0,\ \forall y \in I,\ 0 < |y - x| \leqslant \delta \Rightarrow \frac{f(y) - f(x)}{y - x} > 0 \ .$$

Par conséquent, pour tout $y \in I$ tel que $0 < y - x \leqslant \delta$, on a $f(y) > f(x)$, ce qui montre que x n'est pas un maximum local de f ; et pour tout $y \in I$ tel que $-\delta \leqslant y - x < 0$, on a $f(y) < f(x)$ donc x n'est pas non plus un minimum local.

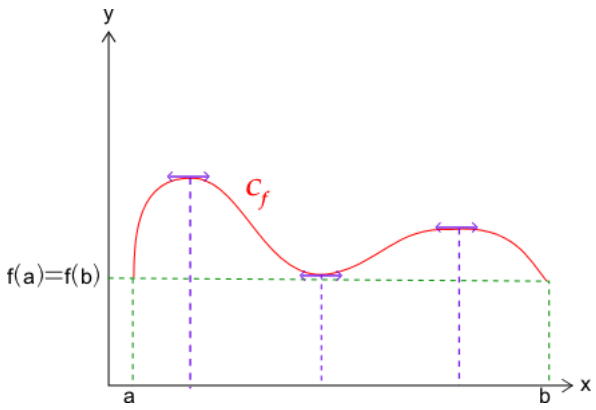
Le cas $f'(x) < 0$ se traite de la même façon, ou se déduit en appliquant ce qu'on vient de démontrer à $-f$. □

Notons par contre que la condition $f'(x) = 0$ n'est pas suffisante pour conclure que x est un extremum local de f ! Par exemple, si on considère la fonction $f : x \mapsto x^3$, alors $f'(0) = 0$ mais 0 n'est pas un extremum local de f. Les développements limités nous permettront bientôt de mieux étudier le comportement local d'une fonction en un point où $f'(x) = 0$.

Théorème 6.23 (Théorème de Rolle)**.** *Soit $a < b$ deux réels, et f une fonction continue sur $[a, b]$ et dérivable sur $]a, b[$ telle que $f(a) = f(b)$. Alors il existe $c \in]a, b[$ tel que $f'(c) = 0$.*

Démonstration. Si f est constante sur $[a, b]$ il n'y a rien à faire ; sinon, on sait que f admet un maximum M et un minimum m sur $[a, b]$, et au moins l'un des deux doit être différent de $f(a)$. Disons par exemple que $M > f(a)$ et soit $c \in]a, b[$ tel que $f(c) = M$. Alors $f(c)$ est le maximum de f sur $]a, b[$: c est un maximum local de f, donc $f'(c) = 0$. □

Ci-dessous une image [i] illustrant le théorème de Rolle ; sur le dessin il y a trois c satisfaisant l'égalité $f'(c) = 0$.

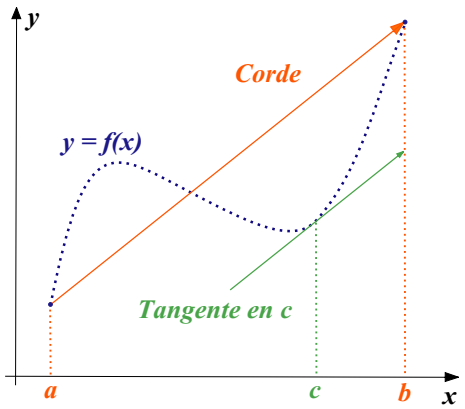


Théorème 6.24 (Égalité des accroissements finis)**.** *Soit $a < b$ deux réels, et $f : [a, b] \to \mathbb{R}$ une fonction continue sur $[a, b]$ et dérivable sur $]a, b[$. Alors il existe $c \in]a, b[$ tel que*

$$f(b) - f(a) = f'(c)(b - a) \ .$$

Notons que $\dfrac{f(b) - f(a)}{b - a}$ est le coefficient directeur de la droite reliant $(a, f(a))$ à $(b, f(b))$ (qu'on appelle une *corde* entre deux points du graphe de f) tandis que $f'(c)$ est le coefficient directeur de la tangente en c. Le théorème des accroissements finis nous dit donc que, étant donné une corde reliant deux points sur le graphe de f, on peut trouver quelque part entre ces deux points une tangente au graphe qui est parallèle à la corde en question.

C'est illustré par le dessin suivant [ii]



i. Issue de l'article Wikipedia sur le théorème de Rolle `https://fr.wikipedia.org/wiki/Théorème_de_Rolle`

ii. version à peine modifiée du dessin figurant sur la page Wikipedia `https://fr.wikipedia.org/wiki/Théorème_des_accroissements_finis`

Démonstration. On considère la fonction g définie sur $[a,b]$ par

$$g(x) = f(x) - \frac{f(b)-f(a)}{b-a}(x-a) \ .$$

Alors g est continue sur $[a,b]$, dérivable sur $]a,b[$, et on a de plus $g(a) = f(a)$ et $g(b) = f(a)$. Le théorème de Rolle nous dit donc qu'il existe $c \in]a,b[$ tel que $g'(c) = 0$, c'est-à-dire :

$$f'(c) - \frac{f(b)-f(a)}{b-a} = 0 \ .$$

□

Une application fondamentale de l'égalité des accroissements finis est de nous donner un lien entre signe de la dérivée et sens de variation de la fonction, qui justifie l'utilisation de la dérivée pour dresser des tableaux de variations.

Proposition 6.25. *Soit I un intervalle de $\mathbb{R}$ et $f\colon I \to \mathbb{R}$ une fonction dérivable. Alors f est croissante sur I si, et seulement si, on a $f'(x) \geqslant 0$ pour tout $x \in I$; et f est décroissante sur I si, et seulement si, on a $f'(x) \leqslant 0$ pour tout $x \in I$. De plus, si $f'(x) > 0$ pour tout $x \in I$ alors f est strictement croissante, et si $f'(x) < 0$ pour tout $x \in I$ alors f est strictement décroissante.*

Démonstration. Si f est croissante et dérivable sur I, alors pour tout $x \in I$ et tout $y \neq x$ on a $f(y) - f(x)$ est du même signe que $y - x$, et donc

$$f'(x) = \lim_{y \to x} \frac{f(y)-f(x)}{y-x}$$

est une limite de nombres positifs donc $f'(x) \geqslant 0$. Réciproquement, si $f'(x) \geqslant 0$ pour tout $x \in I$ alors considérons $x < y \in I$. Par l'égalité des accroissements finis, il existe $c \in]x,y[$ tel que

$$f(y) - f(x) = f'(c)(y-x) \ .$$

On voit donc que $f(y) \geqslant f(x)$: f est croissante ; et si on avait supposé $f'(x) > 0$ sur I alors on aurait déduit $f(y) > f(x)$ dès que $y > x$, c'est-à-dire que f est strictement croissante.

Le cas des fonctions décroissantes se traite de la même façon, ou se déduit du cas des fonctions croissantes en remplaçant f par $-f$. □

On voit aussi que si $f'(x) = 0$ pour tout $x \in I$ alors f est constante ; et, pour conclure que f est strictement monotone sur un intervalle $[a,b]$, l'égalité des accroissements finis nous dit qu'il suffit de vérifier que $f'(x) > 0$ pour tout $x \in]a,b[$.

Citons une autre conséquence fréquemment utilisée de l'égalité des accroissements finis.

Proposition 6.26. *Soit I un intervalle de $\mathbb{R}$, et $f\colon I \to \mathbb{R}$ une fonction continue. Si de plus $x \in I$ est tel que f est dérivable sur $I \setminus \{x\}$, et $\lim\limits_{y \to x, y \neq x} f'(y) = l$ existe dans $\mathbb{R}$, alors f est dérivable en x et $f'(x) = l$.*

Démonstration. Soit (y_n) une suite d'éléments de I, différents de x, qui converge vers x. Alors pour tout n, l'égalité des accroissements finis nous assure de l'existence de $c_n \in]x, y_n|$ tel que $f(x) - f(y_n) = (x - y_n)f'(c_n)$. De plus, par hypothèse, $f'(c_n)$ converge vers l. Mais alors

$$\frac{f(x)-f(y_n)}{x-y_n} = f'(c_n) \to l \ .$$

Autrement dit, on a comme espéré, puisque la suite (y_n) était quelconque :

$$\lim_{y \to x} \frac{f(x)-f(y)}{x-y} = l \ .$$

□

Souvent, pour étudier des fonctions et calculer des limites, on a besoin d'établir des inégalités. L'égalité des accroissements finis (et sa généralisation, la *formule de* Taylor–Lagrange, qu'on verra plus tard dans ce cours) nous fournit une méthode utile.

Théorème 6.27 (Inégalité des accroissements finis)**.** *Soit $I = [a, b]$ un segment de $\mathbb{R}$, et f une fonction continue sur $[a, b]$ et dérivable sur $]a, b[$. Si $k = \sup_{c\in]a,b[} |f'(c)| < +\infty$, alors pour tout $x, y \in [a, b]$ on a*

$$|f(x) - f(y)| \leqslant k|x - y| \ .$$

Démonstration. Soit $x < y \in I$. On peut appliquer l'égalité des accroissements finis sur $[x, y]$ et obtenir $c \in]x, y[$ tel que $f(x)-f(y) = f'(x)(x-y)$. Par définition de k on a $|f'(c)| \leqslant k$, et donc aussi $|f(x)-f(y)| \leqslant k|x-y|$. □

Notons ici un point important : quand on étudie des fonctions à valeurs dans $\mathbb{R}^n$, on peut généraliser la notion de dérivée (en dérivant chaque coordonnée séparément ; ci-dessous on va brièvement parler du cas des fonctions à valeurs dans $\mathbb{C} \cong \mathbb{R}^2$) et l'*inégalité* des accroissements finis se généralise à ce contexte, alors que l'*égalité* des accroissements finis ne s'y généralise pas ! Un autre fait que vous verrez peut-être un jour est que l'inégalité des accroissements finis reste vraie même sous des hypothèses plus faibles sur f, où on ne demande pas que f soit dérivable sur $[a, b]$ tout entier.

Exercice 6.28. Montrer que pour tout $x \in \mathbb{R}$ on a $|\sin(x)| \leqslant x$.

Notons que les fonctions dérivées partagent des propriétés des fonctions continues, même si elles ne le sont pas nécessairement (plus haut dans ces notes on a vu un exemple de fonction dérivée dont la dérivée n'était pas une fonction continue).

Théorème 6.29 (Théorème de Darboux)**.** *Les fonctions dérivées satisfont la conclusion du théorème des valeurs intermédiaires : si I est un intervalle de $\mathbb{R}$, et $f\colon I \to \mathbb{R}$ est dérivable sur I, alors $f'(I)$ est un intervalle.*

En particulier, si une dérivée ne s'annule pas, elle ne peut pas changer de signe, et la fonction étudiée est strictement monotone.

Démonstration. Soit $a < b \in I$ et λ compris entre $f'(a)$ et $f'(b)$; on doit trouver c dans I tel que $f'(c) = \lambda$. Quitte à remplacer f par $-f$ on suppose $f'(a) > \lambda > f'b)$. En remplaçant f par $x \mapsto f(x) - \lambda x$ on se ramène au cas où $f'(a) > 0$, $f'(b) < 0$ et on cherche c tel que $f'(c) = 0$. Comme $f'(a) > 0$, $f(x) > f(a)$ pour $x > a$ suffisamment proche de a, et donc a n'est pas un maximum de f sur $[a, b]$; de même, comme $f'(b) < 0$ on sait que b n'est pas un maximum de f sur $[a, b]$. Mais f, étant continue sur le segment $[a, b]$, doit y admettre un maximum c, qui appartient à $]a, b[$ par ce qui précède et est donc un maximum local de f : d'où $f'(c) = 0$, et on a trouvé le c qu'on cherchait. □

6.4 Fonctions circulaires réciproques et leurs dérivées

Proposition 6.30. *Soit I un intervalle de $\mathbb{R}$, et $f\colon I \to \mathbb{R}$ une fonction continue, strictement monotone sur I et dérivable en $x \in I$. Alors $f(I) = J$ est un intervalle, f est une bijection de I sur J, et $f^{-1}\colon J \to I$ est continue sur J et dérivable en $f(x)$ si et seulement si $f'(x) \neq 0$. Dans ce cas on a la formule*

$$(f^{-1})'(f(x)) = \frac{1}{f'(x)} \ .$$

En particulier, si f est dérivable sur I et $f'(x) \neq 0$ pour tout $x \in I$, alors f est une bijection de I sur J, la fonction réciproque de f est dérivable sur J et pour tout $y \in J$ on a

$$(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))} \ .$$

Démonstration. Soit f une fonction satisfaisant les conditions ci-dessus. On sait déjà par 6.6 que $f(I) = J$ est un intervalle, f est une bijection de I sur J et $g = f^{-1}$ est continue sur J. Reste à comprendre quand g est dérivable en $f(x)$ et calculer le cas échéant la valeur de $g'(x)$. Pour cela, on considère une suite (y_n) d'éléments de J qui converge vers $y = f(x)$ avec $y_n \neq y$ pour tout $n \in \mathbb{N}$, et on cherche à déterminer la limite de $\dfrac{g(y_n) - g(y)}{y_n - y}$. On sait que la suite $g(y_n)$ converge vers $g(y) = x$ puisque g est continue ; puisque f est dérivable en x, on peut écrire

$$f(g(y_n)) = f(x) + (g(y_n) - x)f'(x) + (g(y_n) - x)\varepsilon(g(y_n)) \ ,$$

où $\varepsilon\colon I \to \mathbb{R}$ est une fonction qui tend vers 0 en x. Puisque $f(g(y_n)) = y_n$ et $x = g(y)$, on obtient en définissant ε' sur J par $\varepsilon'(z) = \varepsilon(g(z))$ que

$$y_n = y + (g(y_n) - g(y))f'(x) + (g(y_n) - g(y))\varepsilon'(y_n) \ .$$

Ainsi,

$$\lim_{n\to+\infty} \frac{y_n - y}{g(y_n) - g(y)} = f'(x) \ .$$

Autrement dit,

$$\lim_{z\to y} \frac{z - y}{g(z) - g(y)} = f'(x) \ .$$

On voit donc que la limite en y du taux d'accroissement $\dfrac{g(z) - g(y)}{z - y}$ existe si et seulement si $f'(x) \neq 0$, et vaut $\dfrac{1}{f'(x)}$ dans ce cas. □

La fonction arcsin**.**

En utilisant le fait que $\cos(x) > 0$ pour tout $x \in]-\frac{\pi}{2}, \frac{\pi}{2}[$ et $\sin' = \cos$, on obtient le tableau de variations suivant pour la fonction sin sur cet intervalle :

x	$-\dfrac{\pi}{2}$		0		$\dfrac{\pi}{2}$
$\cos(x)$		$+$	1	$+$	
$\sin(x)$	-1	↗	0	↗	1

On voit que la fonction sin réalise une bijection, strictement croissante, de $[-\frac{\pi}{2}, \frac{\pi}{2}]$ sur $[-1, 1]$. Elle admet donc une bijection réciproque, notée arcsin, définie sur $[-1, 1]$ et à valeurs dans $[-\frac{\pi}{2}, \frac{\pi}{2}]$. Pour tout $t \in [-1, 1]$ et tout $x \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ on a l'équivalence suivante :

$$(\sin(x) = t) \Leftrightarrow (x = \arcsin(t)) \ .$$

En particulier, puisque la fonction sin est impaire, on a pour tout $t \in [-1, 1]$ et tout $x \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ que

$$\begin{aligned} \arcsin(t) = x &\Leftrightarrow \sin(x) = t \\ &\Leftrightarrow \sin(-x) = -t \\ &\Leftrightarrow -x = \arcsin(-t) \ . \end{aligned}$$

Autrement dit, la fonction arcsin est impaire. Puisque $\sin' = \cos$ ne s'annule pas sur $]-\frac{\pi}{2}, \frac{\pi}{2}[$, arcsin est dérivable sur $\sin\left(]-\frac{\pi}{2}, \frac{\pi}{2}[\right) =]-1, 1[$ et on a

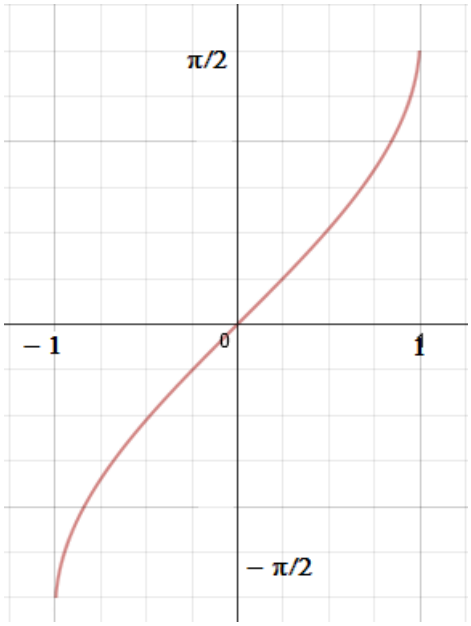
$$\forall t \in]-1, 1[, \quad \arcsin'(t) = \frac{1}{\cos(\arcsin(t))} \ .$$

Cette expression se simplifie : pour $t \in [-1, 1]$, en notant $x = \arcsin(t)$, on sait que $\sin(x) = t$, $x \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ donc $\cos(x) \geqslant 0$, et $\cos^2(x) + \sin^2(x) = 1$, donc $\cos^2(x) + t^2 = 1$. On vient de montrer que $\cos(\arcsin(t)) = \sqrt{1 - t^2}$ pour tout $t \in [-1, 1]$, et on a donc la formule

$$\forall t \in]-1, 1[, \quad \arcsin'(t) = \frac{1}{\sqrt{1 - t^2}} \ .$$

Ci-dessous, le tableau de variation et le graphe de arcsin sur $[-1, 1]$.

x	-1		0		1
$\frac{1}{\sqrt{1-x^2}}$	‖	+	1	+	‖
$\arcsin(x)$	$-\frac{\pi}{2}$	↗	0	↗	$\frac{\pi}{2}$



La fonction arccos**.**

De la même façon, on obtient le tableau de variation suivant pour la fonction cos sur $[0, \pi]$:

x	0		$\frac{\pi}{2}$		π
$-\sin(x)$		−	−1	−	
$\cos(x)$	1	↘	0	↘	-1

On peut donc définir la bijection réciproque arccos: $[-1, 1] \to [0, \pi]$ (attention à l'ensemble d'arrivée de arccos !). De la formule $\cos(\pi - x) = \cos(x)$ on déduit la relation

$$\forall t \in [-1, 1], \quad \arccos(-t) = \pi - \arccos(t) \ .$$

Comme précédemment, à partir de la relation $\cos^2 + \sin^2 = 1$ et du fait que $\sin(\arccos(t)) \geqslant 0$ pour tout $t \in [-1, 1]$ on déduit la relation $\sin(\arccos(t)) = \sqrt{1 - t^2}$ pour tout $t \in [-1, 1]$; et on obtient que la fonction arccos est dérivable sur $]-1, 1[$, avec

$$\forall t \in]-1, 1[, \quad \arccos'(t) = -\frac{1}{\sqrt{1 - t^2}} \ .$$

Vous avez peut-être remarqué que $\arccos'(t) + \arcsin'(t) = 0$ sur $]-1, 1[$, ce qui entraîne que la fonction $\arccos + \arcsin$ est constante ; comme $\arccos(0) = \frac{\pi}{2}$ et $\arcsin(0) = 0$, cette constante vaut $\frac{\pi}{2}$ et on vient d'établir la relation suivante :

$$\forall t \in [-1, 1], \ \arccos(t) + \arcsin(t) = \frac{\pi}{2} \ .$$

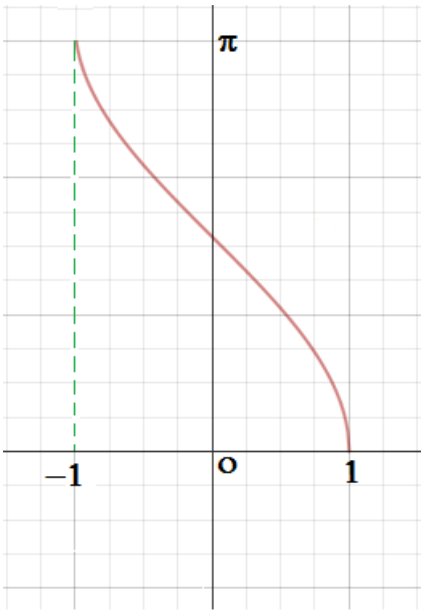
Cette relation peut bien sûr se démontrer en utilisant des formules de trigonométrie : soit $t \in [-1,1]$ et $x = \arcsin(t)$. Alors $x \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ donc $\frac{\pi}{2} - x \in [0,\pi]$. Et on a

$$\cos\left(\frac{\pi}{2} - x\right) = \sin(x) = t\ .$$

On obtient donc comme attendu $\frac{\pi}{2} - x = \arccos(t)$, autrement dit $\frac{\pi}{2} - \arcsin(t) = \arccos(t)$.

Ci-dessous, le tableau de variation et le graphe de arccos sur $[0,\pi]$.

x	-1		0		1
$-\frac{1}{\sqrt{1-x^2}}$	‖	$-$	-1	$-$	‖
$\arccos(x)$	π	↘	1	↘	0



La fonction arctan

La fonction circulaire réciproque qui joue le rôle le plus important dans les exercices est la réciproque de la fonction tan ; c'est lié à la formule pour sa dérivée qu'on va obtenir ci-dessous. Pour l'instant, rappelons-nous que la fonction tan est définie sur $]-\frac{\pi}{2}, \frac{\pi}{2}[$ et que sur cet intervalle on a

$$\tan'(x) = 1 + \tan^2(x)\ .$$

La fonction tan est donc strictement croissante sur $]-\frac{\pi}{2}, \frac{\pi}{2}[$, de plus elle est impaire, et sa limite en $-\frac{\pi}{2}$ vaut $-\infty$ tandis que sa limite en $\frac{\pi}{2}$ vaut $+\infty$. On obtient le tableau de variations suivant.

x	$-\frac{\pi}{2}$		0		$\frac{\pi}{2}$
$1+\tan^2(x)$	‖	$+$	1	$+$	‖
$\tan(x)$	$-\infty$	↗	0	↗	$+\infty$

La fonction réciproque arctan réalise donc une bijection (strictement croissante) de $]-\frac{\pi}{2},\frac{\pi}{2}[$ sur $\mathbb{R}$; pour $t\in]-\frac{\pi}{2},\frac{\pi}{2}[$ et $x\in\mathbb{R}$ on a l'équivalence

$$(\tan(x)=t)\Leftrightarrow(x=\arctan(t))\ .$$

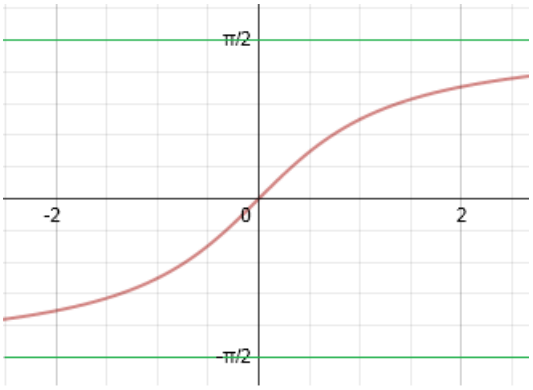
À partir de cette relation et de l'imparité de tan, on déduit que arctan est impaire ; et elle est strictement croissante puisque tan l'est. On peut calculer sa dérivée sans difficulté : pour tout $t\in]-\frac{\pi}{2},\frac{\pi}{2}[$, on a

$$\arctan'(t)=\frac{1}{\tan'(\arctan(t))}=\frac{1}{1+(\tan(\arctan(t)))^2}=\frac{1}{1+t^2}\ .$$

Notons que la dérivée de arctan est une fraction rationnelle, en fait c'est même un élément simple (sur $\mathbb{R}$) : plus tard, quand on essaiera de calculer des primitives de fractions rationnelles, la fonction arctan jouera un rôle important...

Ci-dessous, le graphe et le tableau de variations de arctan sur $]-\frac{\pi}{2},\frac{\pi}{2}[$.

x	$-\infty$		0		$+\infty$
$\frac{1}{1+x^2}$		$+$	1	$+$	
$\arctan(x)$	$-\frac{\pi}{2}$	↗	0	↗	$\frac{\pi}{2}$



6.5 Dérivées d'ordre supérieur

Définition 6.31. Soit I un intervalle de $\mathbb{R}$, et $f\colon I\to\mathbb{R}$ une fonction. Si f est k fois dérivable sur I, et sa dérivée k-ième $f^{(k)}$ est continue sur I , on dit que f est *de classe* $\mathcal{C}^k$ sur I. Par convention, si f est continue sur I on dit que f est de classe $\mathcal{C}^0$ sur I ; si jamais f est de classe $\mathcal{C}^k$ sur I pour tout entier k alors on dit que f est de classe $\mathcal{C}^\infty$.

Par convention, on note $f^{(0)}=f$; notons que $(f^{(k)})'=(f')^{(k)}=f^{(k+1)}$. Attention, le fait que f soit k fois dérivable sur I n'est pas suffisant pour conclure que f est de classe $\mathcal{C}^k$: il existe des fonctions dérivables dont la dérivée n'est pas continue...

Proposition 6.32. *Soit I un intervalle de $\mathbb{R}$, $k\geqslant 1$ un entier, $\lambda\in\mathbb{R}$ et f,g deux fonctions de classe $\mathcal{C}^k$ sur I. Alors fg et $f+g$ sont de classe $\mathcal{C}^k$; de plus on a les formules*

- $(\lambda f+g)^{(k)}=\lambda f^{(k)}+g^{(k)}$.
- $(fg)^{(k)}=\sum_{j=0}^{k}\binom{k}{j}f^{(j)}g^{(k-j)}$.

En particulier, les sommes et les produits de fonctions de classe $\mathcal{C}^\infty$ sont encore $\mathcal{C}^\infty$; on retrouve ainsi le fait que toute fonction polynôme est de classe $\mathcal{C}^\infty$.

Démonstration. Ces deux formules se démontrent par récurrence ; la première est très facile à montrer et on laisse la preuve en exercice. Montrons la deuxième propriété ; elle est bien vraie pour $k = 1$: si f, g sont de classe $\mathcal{C}^1$ sur I alors fg est dérivable sur I, et sa dérivée $f'g + g'f$ est continue. Supposons la propriété vraie au rang k et considérons f, g deux fonctions de classe $\mathcal{C}^{k+1}$ sur I. Alors f et g sont en particulier de classe $\mathcal{C}^k$, et on a par hypothèse de récurrence que

$$(fg)^{(k)} = \sum_{j=0}^{k} \binom{k}{j} f^{(j)} g^{(k-j)} \ .$$

Chaque fonction $f^{(j)}g^{(k-j)}$ est un produit de fonctions de classe (au moins) $\mathcal{C}^1$, et est donc de classe $\mathcal{C}^1$; en dérivant terme à terme on obtient que $(fg)^{(k)}$ est dérivable, de dérivée

$$\begin{aligned}
(fg)^{(k+1)} &= \sum_{j=0}^{k} \binom{k}{j} \left(f^{(j+1)} g^{(k-j)} + f^{(j)} g^{(k+1-j)} \right) \\
&= \sum_{j=1}^{k+1} \binom{k}{j-1} f^{(j)} g^{(k+1-j)} + \sum_{j=0}^{k} \binom{k}{j} f^{(j)} g^{(k+1-j)} \\
&= f^{(k+1)} g + \sum_{j=1}^{k} \left(\binom{k}{j-1} + \binom{k}{j} \right) f^{(j)} g^{(k+1-j)} + f g^{(k+1)} \\
&= \sum_{j=0}^{k+1} \binom{k+1}{j} f^{(j)} g^{(k-j)} \ .
\end{aligned}$$

□

Proposition 6.33. *Soit $k \geqslant 1$ un entier, I, J deux intervalles de $\mathbb{R}$, $f \colon J \to \mathbb{R}$ et $g \colon I \to \mathbb{R}$ telles que $g(I) \subseteq J$ et f, g sont de classe $\mathcal{C}^k$. Alors la fonction composée $f \circ g$ est de classe $\mathcal{C}^k$ sur I.*

Ce résultat entraîne qu'une composée de fonctions C^∞ est encore C^∞. La preuve se fait à nouveau par récurrence, sur le modèle de ce qu'on a fait pour les produits ; on a besoin de traiter avec attention le cas des fonctions de classe $\mathcal{C}^1$, et dans ce cas on peut utiliser la formule pour la dérivée d'une composée pour conclure. Ensuite, on remarque que si f, g sont de classe $\mathcal{C}^{n+1}$, la dérivée de $f \circ g$ est obtenue comme produit et composée de fonctions de classe $\mathcal{C}^n$, et est donc elle-même $\mathcal{C}^n$. On conclut alors par récurrence. On va s'épargner les détails de cette preuve, et aussi éviter d'essayer d'écrire une formule explicite pour la dérivée n-ième d'une fonction composée (qui est compliquée !)

Corollaire 6.34.

1. Soit I un intervalle de $\mathbb{R}$, $k \geqslant 1$ un entier et $f \colon I \to \mathbb{R}$ une fonction de classe $\mathcal{C}^k$ ne s'annulant pas sur I. Alors $\dfrac{1}{f}$ est de classe $\mathcal{C}^k$ sur I.
2. Soit I un intervalle de $\mathbb{R}$, $k \geqslant 1$ et f une fonction de classe $\mathcal{C}^k$ telle que f' ne s'annule pas sur I. Alors f est une bijection de I sur l'intervalle $J = f(I)$, et sa bijection réciproque est de classe $\mathcal{C}^k$.

Par conséquent, si f est de classe $\mathcal{C}^\infty$ et ne s'annule pas sur I, alors $\frac{1}{f}$ est de classe $\mathcal{C}^\infty$ sur I. Le deuxième point ci-dessus nous permet de retrouver le fait que arcsin et arccos sont de classe $\mathcal{C}^\infty$ sur $]-1, 1[$ et arctan est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ (ce qui se voyait directement sur la formule donnant leurs dérivées)

Démonstration. Pour démontrer le premier point, on écrit simplement que $\frac{1}{f}$ est obtenue en composant la fonction $x \mapsto \frac{1}{x}$ et f. Pour le deuxième point, on sait sous les hypothèses ci-dessus que f doit être strictement monotone (le théorème de Darboux nous assure que f' ne change pas de signe), et que $f(I) = J$ est un intervalle par le théorème des valeurs intermédiaires. On sait alors que f^{-1} est dérivable sur J et que

$$\forall x \in J, \quad (f^{-1})'(x) = \frac{1}{f'(f^{-1}(x))} \ .$$

Ceci nous montre que $(f^{-1})'$ est $\mathcal{C}^1$ si f est $\mathcal{C}^1$; ensuite on raisonne encore par récurrence sur k, on suppose que la propriété qu'on souhaite démontrer est vraie au rang k et que f est une fonction de classe $\mathcal{C}^{k+1}$ satisfaisant les conditions ci-dessus. Par la proposition précédente, $f' \circ f^{-1}$ est de classe $\mathcal{C}^k$ puisque c'est la composée de deux fonctions de classe $\mathcal{C}^k$. Donc $(f^{-1})'$ est de classe $\mathcal{C}^k$ en tant qu'inverse d'une fonction de classe $\mathcal{C}^k$ qui ne s'annule pas sur J ; autrement dit f^{-1} est de classe $\mathcal{C}^{k+1}$. □

6.6 Dériver des fonctions à valeurs dans $\mathbb{C}$

Dans le chapitre suivant, on va calculer des intégrales ; pour cela, il est parfois pratique d'intégrer des fonctions à valeurs complexes (le plus souvent, une fonction comme $t \mapsto e^{it}$) ; il est utile d'avoir pris le temps de parler de la dérivée d'une fonction à valeurs dans $\mathbb{C}$.

Définition 6.35. Soit I un intervalle de $\mathbb{R}$, et $f\colon I \to \mathbb{C}$ une fonction. Si sa partie réelle $\mathrm{Re}(f)$ et sa partie imaginaire $\mathrm{Im}(f)$ sont toutes deux dérivables en $x \in I$, alors on dit que f est dérivable en $x \in I$ et on pose

$$f'(x) = \mathrm{Re}(f)'(x) + i\mathrm{Im}(f)'(x) \ .$$

Notons que l'on pourrait définir la notion de limite d'une fonction à valeurs complexes (il suffit de remplacer la valeur absolue par un module dans la définition d'une limite) ; et alors $f'(x)$ est simplement la limite du taux d'accroissement $\dfrac{f(y)-f(x)}{y-x}$ quand y tend vers x. Dans ce cours on se contente de mentionner ce dont on a besoin pour simplifier certains calculs d'intégrales.

Proposition 6.36. *Si I est un intervalle de $\mathbb{R}$ et $f, g\colon I \to \mathbb{C}$ sont des fonctions dérivables sur I alors fg est dérivable sur I et on a*

$$\forall x \in I, \quad (fg)'(x) = f(x)g'(x) + f'(x)g(x) \ .$$

Démonstration. Notons f_1, g_1 les parties réelles de f, g et f_2, g_2 leurs parties imaginaires. Alors on a

$$\forall x \in I, \quad f(x)g(x) = (f_1(x)g_1(x) - f_2(x)g_2(x)) + i(f_1(x)g_2(x) + f_2(x)g_1(x)) \ .$$

On déduit des théorèmes sur les fonctions dérivables à valeurs dans $\mathbb{R}$ que la partie réelle et la partie imaginaire de fg sont toutes deux dérivables sur I, donc fg est dérivable sur I, et de plus on a pour tout $x \in I$:

$$\begin{aligned}(fg)'(x) &= (f_1'(x)g_1(x) + f_1(x)g_1'(x) - f_2'(x)g_2(x) - f_2(x)g_2'(x)) + i(f_1'(x)g_2(x) + f_1(x)g_2'(x) + f_2'(x)g_1(x) + f_2(x)g_2'(x)) \\ &= (f_1'(x) + if_2'(x))(g_1(x) + ig_2(x)) + (f_1(x) + if_2(x))(g_1'(x) + ig_2'(x)) \\ &= f'(x)g(x) + f(x)g'(x) \ .\end{aligned}$$

□

Sur le même modèle, on vérifie que les théorèmes sur les dérivations de quotients et de fonctions composées se généralisent mot pour mot aux fonctions à valeurs complexes. Notons en particulier le fait suivant : si I est un intervalle de $\mathbb{R}$, et $f\colon I \to \mathbb{C}$ est dérivable, alors $x \mapsto e^{f(x)}$ est une fonction dérivable à valeurs dans $\mathbb{C}$, et sa dérivée vaut $f'(x)e^{f(x)}$. Souvent, on utilisera simplement que si $\alpha \in \mathbb{C}$ alors $x \mapsto e^{\alpha x}$ est dérivable sur $\mathbb{R}$, de dérivée $\alpha e^{\alpha x}$.

Finissons ce chapitre en mentionnant de nouveau que l'égalité des accroissements finis ne se généralise pas aux fonctions à valeurs dans $\mathbb{C}$: par exemple, considérons la fonction $f\colon x \mapsto e^{ix}$. Alors f est dérivable sur $\mathbb{R}$, de dérivée ie^{ix} ; en particulier $|f'(x)| = 1$ pour tout x et donc f' ne s'annule pas. Pourtant, $f(0) = f(2\pi) = 1$; il ne peut pas exister de c tel que $f(2\pi) - f(0) = 0 = 2\pi f'(c)$.

Chapitre 7

Intégration

7.1 Intégrale des fonctions en escalier

Définition 7.1. Soient $a < b$ deux réels. Une *subdivision* de $[a, b]$ est une suite finie $(a_0, a_1, \ldots, a_n)$ telle que $a_0 = a$, $a_n = b$ et $a_i < a_{i+1}$ pour tout $i \in \{0, \ldots, n-1\}$. On définit le *pas* d'une subdivision $(a_0, a_1, \ldots, a_n)$ comme étant égal à la quantité $\max\{a_{i+1} - a_i : i \in \{0, \ldots, n-1\}\}$.

Intuitivement, considérer une subdivision $(a_0, \ldots, a_n)$ revient à considérer un découpage de $[a, b]$ en n intervalles $[a_0, a_1], \ldots, [a_{n-1}, b]$; dire que le pas de la subdivision est petit signifie que tous les intervalles créés lors du découpage sont petits.

Définition 7.2. Soient $a < b$ deux réels, et $(a_0, \ldots, a_n)$, $(b_0, \ldots, b_m)$ deux subdivisions de $[a, b]$. On dit que $(b_0, \ldots, b_m)$ *raffine* $(a_0, \ldots, a_n)$ si chaque intervalle $[b_j, b_{j+1}]$ est contenu dans un invervalle de la forme $[a_k, a_{k+1}]$.

Cela signifie que la subdivision $(b_0, \ldots, b_m)$ a été obtenue en découpant les intervalles de la subdivision $(a_0, \ldots, a_n)$.

Exercice 7.3. Soient $a < b$ deux réels, et $(a_0, \ldots, a_n)$ et $(b_0, \ldots, b_m)$ deux subdivisions de $[a, b]$. Alors il existe une subdivision $(c_0, \ldots, c_p)$ qui raffine *à la fois* $(a_0, \ldots, a_n)$ et $(b_0, \ldots, b_m)$.

(Indication : $c_0, \ldots, c_p$ peuvent par exemple être obtenus en écrivant dans l'ordre croissant l'ensemble $\{a_0, \ldots, a_n; b_0, \ldots, b_m\}$)

Définition 7.4. Soient $a < b$ deux réels ; $f\colon [a, b] \to \mathbb{R}$ est une *fonction en escalier* s'il existe une subdivision $(a_0, \ldots, a_n)$ de $[a, b]$ telle que f soit constante sur chaque intervalle $]a_i, a_{i+1}[$. On dit que $(a_0, \ldots, a_n)$ *témoigne* du fait que f est en escalier, ou encore est une *subdivision adaptée* à f.

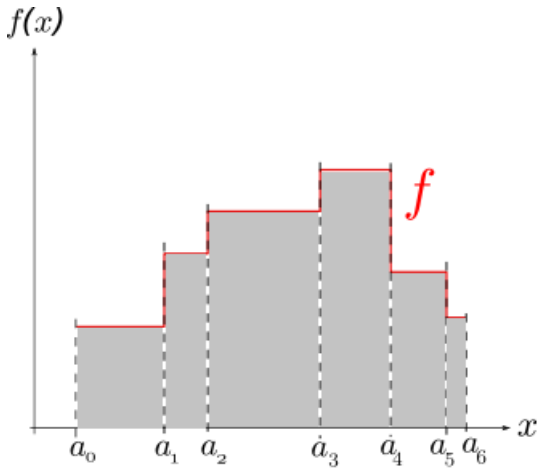
Proposition 7.5.
1. *Une fonction en escalier ne prend qu'un nombre fini de valeurs.*
2. *L'ensemble des fonctions en escaliers sur $[a, b]$ est un sous-espace vectoriel des fonctions réelles sur $[a, b]$.*
3. *Un produit de fonctions en escalier sur $[a, b]$ est une fonction en escalier sur $[a, b]$.*

Démonstration. La première propriété découle immédiatement de la définition. Les preuves des deuxième et troisième propriétés sont très similaires. Tout d'abord la fonction constante 0 est bien une fonction en escalier. Soient f, g deux fonctions en escalier, $\lambda \in \mathbb{R}$, $(a_0, \ldots, a_n)$ une subdivision qui témoigne du fait que f est en escalier et $(b_0, \ldots, b_m)$ une subdivision qui témoigne du fait que g est en escalier, $a_0 = b_0 = a$, $a_n = b_m = b$. Par l'exercice précédent, on peut trouver une subdivision $(c_0, \ldots, c_p)$ qui raffine ces deux subdivisions. Etant donné i entre 0 et $p-1$, il existe j, k tel que $[c_i, c_{i+1}]$ soit contenu dans $[a_j, a_{j+1}]$ et dans $[b_k, b_{k+1}]$. En particulier, les deux fonctions f et g sont constantes sur $]c_i, c_{i+1}[$, donc $\lambda f + g$ et fg y sont constantes aussi. Ainsi, la subdivision $(c_0, \ldots, c_p)$ témoigne du fait que $\lambda f + g$ et fg sont des fonctions en escalier. □

Définition 7.6. Soient $a < b$ deux réels, $f\colon [a, b] \to \mathbb{R}$ une fonction en escalier, et $\sigma = (a_0, \ldots, a_n)$ une subdivision adaptée à f. On pose

$$I(f, \sigma) = \sum_{k=0}^{n-1} (a_{k+1} - a_k) f\left(\frac{a_k + a_{k+1}}{2}\right) .$$

Remarque 7.7. Dans la définition de $I(f,\sigma)$, on aurait pu remplacer $\frac{a_k+a_{k+1}}{2}$ par n'importe quel point de $]a_k, a_{k+1}[$ sans changer la valeur de $I(f,\sigma)$.



On voit que $I(f,\sigma)$ correspond à l'"aire sous le graphe de f" quand f est à valeurs positives.

Lemme 7.8. *Soient $a<b$ deux réels, $f\colon [a,b]\to\mathbb{R}$ une fonction en escalier, et σ,τ deux subdivisions adaptées à f. Alors $I(f,\sigma)=I(f,\tau)$.*

Démonstration. Commençons par le cas où $\tau=(b_0,\dots,b_m)$ raffine $\sigma=(a_0,\dots,a_n)$. Alors il existe $j_0,\dots,j_n$ tels que pour tout $k\in\{0,\dots,n\}$ on ait $b_{j_k}=a_k$ (en particulier $j_0=0$, $j_n=m$). Alors on a

$$\begin{aligned}
I(f,\tau) &= \sum_{j=0}^{m-1}(b_{j+1}-b_j)f\left(\frac{b_j+b_{j+1}}{2}\right)\\
&= \sum_{k=0}^{n-1}\sum_{j=i_k}^{i_{k+1}-1}(b_{j+1}-b_j)f\left(\frac{b_j+b_{j+1}}{2}\right)\\
&= \sum_{k=0}^{n-1}\sum_{j=i_k}^{i_{k+1}-1}(b_{j+1}-b_j)f\left(\frac{a_k+a_{k+1}}{2}\right)\\
&= \sum_{k=0}^{n-1}f\left(\frac{a_k+a_{k+1}}{2}\right)\left(\sum_{j=i_k}^{i_{k+1}-1}b_{j+1}-b_j\right)\\
&= \sum_{k=0}^{n-1}f\left(\frac{a_k+a_{k+1}}{2}\right)(a_{k+1}-a_k)\\
&= I(f,\sigma)\ .
\end{aligned}$$

On a donc démontré le résultat désiré dans le cas où τ raffine σ. Si maintenant τ et σ sont deux subdivisions quelconques adaptées à f, il existe une subdivision γ qui raffine à la fois τ et σ. Cette subdivison est encore adaptée à f, donc par le cas précédent on a $I(f,\sigma)=I(f,\gamma)$ et $I(f,\gamma)=I(f,\tau)$. On en déduit bien que $I(f,\sigma)=I(f,\tau)$. □

Ce lemme nous permet finalement de définir l'intégrale d'une fonction en escalier.

Définition 7.9. Soient $a<b$ deux réels, $f\colon[a,b]\to\mathbb{R}$ une fonction en escalier, et σ une subdivision adaptée à f. On pose

$$\int_a^b f(x)dx = I(f,\sigma)\ .$$

Le lemme précédent nous dit que cette définition est légitime : quelle que soit la subdivision σ adaptée à f que l'on choisisse, $I(f,\sigma)$ a toujours la même valeur.

Répétons quelle est l'intuition derrière cette définition : pour une fonction f à valeurs positives, l'intégrale est censée représenter « l'aire sous la courbe de f ». Dans le cas où f est en escalier et à valeurs positives, le domaine sous la courbe de f est une union finie de rectangles, et la formule que l'on a donnée pour l'intégrale de f correspond à la somme des aires de ces rectangles.

Evidemment, on ne veut pas intégrer que des fonctions en escalier ! Si jamais la partie délimitée par l'axe des abscisses et le graphe de f peut être bien approchée par des unions de rectangles, on peut définir l'aire sous la courbe comme la limite de la somme des aires de ces rectangles ; c'est comme ça qu'on va définir les fonctions intégrables dans la section suivante.

pour étendre la définition de l'intégrale à des fonctions plus générales (et particulièrement pour pouvoir intégrer des fonctions continues !) on a besoin des propriétés suivantes, qui suivent facilement (même si les preuves peuvent être un peu pénibles à écrire...) de la définition de l'intégrale des fonctions en escalier.

Proposition 7.10. *1. Etant donnés deux réels $a<b$, on a $\int_a^b 1dx = b-a$.*

2. *Etant donnés deux réels $a<b$ et une fonction f en escalier sur $[a,b]$ et à valeurs positives, $\int_a^b f(x)dx \geqslant 0$* **(positivité de l'intégrale)**.
3. *Etant donnés deux réels $a<b$, deux fonctions f,g en escalier sur $[a,b]$ et deux constantes α,β, on a* **(linéarité de l'intégrale)** *:*

$$\int_a^b (\alpha f(x)+\beta g(x))\,dx = \alpha\int_a^b f(x)\,dx + \beta\int_a^b g(x)\,dx\ .$$

4. *Etant donnés deux réels $a<b$, et f une fonction en escalier sur $[a,b]$, on a* **(inégalité triangulaire)**

$$\left|\int_a^b f(x)\,dx\right| \leqslant \int_a^b |f(x)|\,dx\ .$$

7.2 Fonctions intégrables

Définition 7.11. Soient $a<b$ deux réels. Une *subdivision pointée* $\sigma=(a_0,\dots,a_n;x_1,\dots,x_n)$ de $[a,b]$ est la donnée :

— d'une subdivision $(a_0,\dots,a_n)$ de $[a,b]$;

— de points $x_1,\dots,x_n$ de $[a,b]$ tels que pour tout $k\in\{1,\dots,n\}$ on ait $x_k\in[a_{k-1},a_k]$.

Un cas particulier très important en pratique est celui ou l'on divise l'intervalle $[a,b]$ en n intervalles de même longueur $\frac{b-a}{n}$, et où l'on choisit x_k à gauche de l'intervalle $[a_{k-1},a_k]$ (c'est-à-dire $x_k=a+(k-1)\frac{b-a}{n}$) ou à droite de cet intervalle ($x_k=a+k\frac{b-a}{n}$).

Définition 7.12. Soit $a<b$ deux réels, $f\colon[a,b]\to\mathbb{R}$ et $\sigma=(a_0,\dots,a_n;x_1,\dots,x_n)$ une subdivision pointée de $[a,b]$. On appelle *somme de* Riemann *associée à f et à σ* le nombre

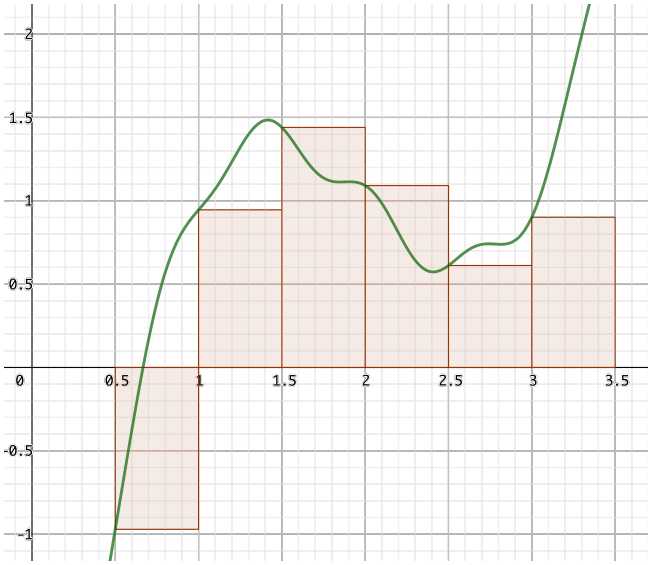
$$S(f,\sigma)=\sum_{k=1}^{n}(a_k-a_{k-1})f(x_k)\ .$$

Notons que, par définition, $S(f,\sigma)$ est l'intégrale d'une fonction en escalier qui est égale à $f(x_i)$ sur chaque $]a_{i-1},a_i[$. En particulier, si σ est la subdivision obtenue en découpant $[a,b]$ en n intervalles de même longueur et que x_k est choisi à gauche du k-ième intervalle de la subdivision, on obtient la formule

$$S(f,\sigma)=\sum_{k=1}^{n}\frac{b-a}{n}f\left(a+(k-1)\frac{b-a}{n}\right)=\frac{b-a}{n}\sum_{k=0}^{n-1}f\left(a+k\frac{b-a}{n}\right)\ .$$

De même, quand τ est obtenue en découpant $[a,b]$ en n intervalles de même longueur et que x_k est l'extrémité droite du k-ième intervalle, on a la somme de Riemann

$$S(f,\tau)=\frac{b-a}{n}\sum_{k=0}^{n}f\left(a+k\frac{b-a}{n}\right)\ .$$



Ci-dessus on a $n = 5$, c'est-à-dire qu'on a partitionné $[a, b]$ en 6 intervalles $[a, a_1], \ldots, [a_4, a_5]$. Dans chacun de ces intervalles on a choisi le point $x_k = a_{k-1}$ et formé le rectangle dont deux côtés sont de longueur $a_k - a_{k-1}$ et les deux autres de longueur *algébrique*[i] $f(x_k)$; et la somme de Riemann correspondante est la somme des aires (algébriques) de ces rectangles. La différence entre ce qu'on voudrait appeler « l'aire sous la courbe » de f et la somme de Riemann, si les intervalles de la subdivision sont de plus en plus fins, va tendre vers 0, et l'intégrale de f sur $[a, b]$ sera ainsi obtenue comme la limite des aires des rectangles rouges.

L'explication intutive qu'on vient de donner nous amène à la définition suivante.

Définition 7.13. Soit $a < b$ deux réels et $f\colon [a, b] \to \mathbb{R}$. On dit que f est *intégrable sur* $[a, b]$ s'il existe un réel l tel que pour tout $\varepsilon > 0$ on peut trouver $\delta > 0$ satisfaisant

$$|S(f, \sigma) - l| \leqslant \varepsilon$$

pour toute subdivision pointée σ de $[a, b]$ de pas $< \delta$. On note alors $l = \displaystyle\int_a^b f(x)dx$.

Autrement dit : quand le pas de la subdivision σ tend vers 0, la somme de Riemann $S(f, \sigma)$ tend vers l ; une autre formulation équivalente : pour toute suite σ_i de subdivisions pointées dont le pas tend vers 0, la suite de sommes de Riemann $S(f, \sigma_i)$ est convergente.

En particulier, si f est intégrable sur $[a, b]$ alors les sommes

$$\frac{b-a}{n} \sum_{k=0}^{n-1} f\left(a + k\frac{b-a}{n}\right) \quad \text{et} \quad \frac{b-a}{n} \sum_{k=1}^{n} f\left(a + k\frac{b-a}{n}\right)$$

convergent toutes deux vers $S(f, \sigma)$ (de même que leur consœur $\displaystyle\frac{b-a}{n} \sum_{k=0}^{n} f\left(a + k\frac{b-a}{n}\right)$).

Définition 7.14. Soient $a < b$ deux réels. On dit qu'une fonction $f\colon [a, b] \to \mathbb{R}$ est *continue par morceaux* sur $[a, b]$ s'il existe une subdivision $(a_0, \ldots, a_n)$ de $[a, b]$ telle que la restriction de f à chaque intervalle $]a_i, a_{i+1}[$ soit continue et admette un prolongement par continuité à $[a_i, a_{i+1}]$.

C'est-à-dire que f ait une limite finie à droite en a_i, respectivement à gauche en a_{i+1}. Comme premiers exemples, notons qu'une fonction continue sur $[a, b]$ est bien sûr continue par morceaux, et qu'une fonction en escalier est également continue par morceaux.

Exercice 7.15. Soit $f\colon [a, b] \to \mathbb{R}$ une fonction continue par morceaux. Montrer que f est bornée.

Dans la suite, on va se contenter d'étudier les propriétés de l'intégrale des fonctions continues par morceaux, qui est bien définie comme le montre le théorème suivant.

i. longueur algébrique : elle est négative si $f(x_k) < 0$...

Théorème 7.16. *Soit $a < b$ deux réels et $f\colon [a,b] \to \mathbb{R}$ une fonction continue par morceaux. Alors f est intégrable sur $[a,b]$.*

Démonstration. On ne va rédiger la preuve que dans le cas où f est continue.

Fixons $\varepsilon > 0$. D'après le théorème 6.12 de HEINE, f est uniformément continue sur $[a,b]$ et on peut trouver $\delta > 0$ tel que

$$\forall x, y \in [a,b],\ |x-y| \leqslant \delta \Rightarrow |f(x) - f(y)| \leqslant \varepsilon\ .$$

Soit maintenant $\sigma = (a_0, \dots, a_n; x_1, \dots, x_n)$, $\tau = (a'_0, \dots, a'_m; x'_1, \dots, x'_m)$ deux subdivisions pointées de $[a,b]$ de pas $\leqslant \frac{\delta}{2}$, et f_σ une fonction en escalier égale à $f(x_i)$ sur chaque $[a_{i-1}, a_i[$, f_τ une fonction en escalier égale à $f(x'_i)$ sur chaque $[a'_{i-1}, a'_i[$. On pose aussi $f_\sigma(b) = f_\tau(b) = f(b)$. Rappelons qu'on a par définition de l'intégrale d'une fonction en escalier les égalités

$$S(f,\sigma) = \int_a^b f_\sigma(x)dx \quad \text{et} \quad S(f,\tau) = \int_a^b f_\tau(x)dx\ .$$

Pour $x \in [a,b[$, il existe un unique i tel que $x \in [a_{i-1}, a_i[$, et $f_\sigma(x) = f(x_i)$; et un unique j tel que $x \in [a'_{j-1}, a'_j[$ tel que $f_\tau(x) = f(x'_j)$. Puisque les pas de σ et τ sont tous deux inférieurs à $\frac{\delta}{2}$, on a $|x - x_i| \leqslant \frac{\delta}{2}$ et $|x - x'_j| \leqslant \frac{\delta}{2}$, d'où $|x_i - x'_j| \leqslant \delta$. Ceci impose que

$$|f_\sigma(x) - f_\tau(x)| = |f(x_i) - f(x'_j)| \leqslant \varepsilon.$$

On en déduit que

$$\left| \int_a^b f_\sigma(x) - f_\tau(x) \right| \leqslant \int_a^b |f_\sigma(x) - f_\tau(x)|dx \leqslant \int_a^b \varepsilon dx \leqslant \varepsilon(b-a)\ .$$

(Ci-dessus, on a appliqué l'inégalité triangulaire à l'intégrale de la fonction en escalier $f_\sigma - f_\tau$). Autrement dit,

$$|S(f,\sigma) - S(f,\tau)| \leqslant \varepsilon(b-a)\ .$$

On vient de montrer que, si σ et τ sont des subdivisions de pas suffisamment petit, $S(f,\sigma)$ et $S(f,\tau)$ deviennent arbitrairement proches. Montrons que ceci entraîne l'intégrabilité de f : si f n'est pas intégrable, alors il existe une suite de subdivsions pointées (σ_i) dont le pas tend vers 0 et qui ne converge pas. Mais f est bornée, disons $|f(x)| \leqslant M$ pour tout $x \in [a,b]$, et on en déduit

$$|S(f,\sigma_i)| \leqslant M(b-a)\ .$$

Par conséquent, on peut utiliser le théorème 5.16 de BOLZANO-WEIERSTRASS pour trouver une première suite strictement croissante (i_k) telle que $S(f,\sigma_{i_k})$ converge vers $l \in \mathbb{R}$; et comme on a supposé que $S(f,\sigma_i)$ ne converge pas, on peut trouver une autre suite (j_k) strictement croissante et telle que $S(f,\sigma_{j_k})$ converge vers $l' \neq l$. Mais alors, σ_{i_k} et σ_{j_k} sont deux suites de subdivisions pointées, dont le pas tend vers 0, et telles que $S(f,\sigma_{i_k}) - S(f,\sigma_{j_k})$ tend vers $l - l' \neq 0$, une contradiction avec ce qu'on a démontré au début de la preuve. □

On a maintenant une méthode pour calculer des intégrales, mais elle n'est pas très pratique ! Essayons de calculer quelques intégrales de fonction continues en passant par des sommes de RIEMANN :

$$\int_0^1 x\,dx = \lim_{n\to+\infty} \frac{1}{n} \sum_{k=0}^{n-1} \frac{k}{n} = \lim_{n\to+\infty} \frac{1}{n^2} \frac{n(n+1)}{2} = \frac{1}{2}\ .$$

Essayons un autre calcul : comment calculer $\int_0^\pi \sin(x)dx$? On doit calculer la limite de

$$S_n = \frac{\pi}{n} \sum_{k=0}^{n-1} \sin\left(\frac{k\pi}{n}\right)\ .$$

A priori, ce n'est pas facile ! Mais on ne se laisse pas décourager ; il est parfois plus facile de passer par les nombres complexes, et on peut reconnaître que S_n est la partie imaginaire de

$$U_n = \frac{\pi}{n} \sum_{k=0}^{n-1} e^{ik\frac{\pi}{n}}\ .$$

Ici on peut reconnaître la somme des termes d'une suite géométrique de raison $e^{i\frac{\pi}{n}}$, et on obtient

$$\begin{aligned} U_n &= \frac{\pi}{n}\frac{1-e^{\frac{in\pi}{n}}}{1-e^{i\frac{\pi}{n}}} \\ &= \frac{2\pi}{n(1-e^{i\frac{\pi}{n}})} \end{aligned}$$

Si on forme la partie imaginaire de U_n, on obtient la formule

$$S_n = \frac{2\pi \sin\left(\frac{\pi}{n}\right)}{(1-\cos\left(\frac{\pi}{n}\right))^2 + (\sin\left(\frac{\pi}{n}\right))^2} \ .$$

La limite de S_n n'est pas facile à calculer avec les outils dont on dispose pour le moment ! En trichant un peu, on peut reconnaître que

$$n\frac{e^{i\frac{\pi}{n}}-1}{\pi} = -\frac{e^{i\frac{\pi}{n}}-e^{i.0}}{\frac{\pi}{n}-0}$$

est un taux d'accroissement entre 0 et $\frac{\pi}{n}$ de la fonction $x \mapsto e^{ix}$ et doit donc tendre vers $(e^{ix})'(0) = i$ quand x tend vers 0 ; on en déduit que (U_n) tend vers $\frac{-2}{i} = 2i$, et donc que sa partie imaginaire S_n tend vers 2. On en conclut finalement que

$$\int_0^\pi \sin(x)dx = 2 \ .$$

On vient de souffrir un peu, et de passer par des limites de suites de nombres complexes, pour calculer l'intégrale de $\sin(x)$ entre 0 et π : on a besoin d'une méthode plus efficace ! Bien sûr, vous en connaissez une : passer par le calcul de primitives ! Mais on n'est pas encore tout à fait équipé pour faire le lien entre intégration et primitives, qu'on va établir dans la prochaine section.

Avant de passer à cette section, notons que souvent on utilise les sens de Riemann en sens inverse de ce qu'on a fait ci-dessus : on les utilise pour calculer la limite de certaines suites ; par exemple, si l'on sait calculer $\int_0^\pi \sin(x)dx$, le lien avec les sommes de Riemann rend immédiat le fait que

$$\lim_{n\to+\infty} \frac{\pi}{n}\sum_{k=0}^{n-1} \sin\left(\frac{k\pi}{n}\right) = 2 \ .$$

7.3 Propriétés fondamentales de l'intégrale

Notation. Soient $a > b$ deux réels, et $f\colon [b,a] \to \mathbb{R}$ une fonction continue par morceaux. On pose $\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx$. On pose aussi $\int_a^a f(x)\,dx = 0$.

Proposition 7.17. — *Soient a, b, c trois réels et f une fonction continue par morceaux sur un intervalle qui contient a,b et c. Alors on a* **(Relation de Chasles)**

$$\int_a^b f(x)\,dx + \int_b^c f(x)\,dx = \int_a^c f(x)\,dx \ .$$

— *Etant donnés deux réels $a < b$ et une fonction f continue par morceaux sur $[a,b]$ et à valeurs positives, $\int_a^b f(x)\,dx \geqslant 0$* **(positivité de l'intégrale)**.

— *Etant donnés deux réels $a < b$, deux fonctions f, g continue par morceauxs sur $[a,b]$ et deux constantes α, β, on a* **(linéarité de l'intégrale)** *:*

$$\int_a^b (\alpha f(x) + \beta g(x))\,dx = \alpha\int_a^b f(x)\,dx + \beta\int_a^b g(x)\,dx \ .$$

— *Etant donnés deux réels $a < b$, et f une fonction continue par morceaux sur $[a,b]$, on a* **(inégalité triangulaire)**

$$\left|\int_a^b f(x)\,dx\right| \leqslant \int_a^b |f(x)|\,dx \ .$$

Toutes ces propriétés se déduisent facilement à partir de la définition de l'intégrale d'une fonction continue par morceaux et des propriétés analogues pour l'intégrale d'une fonction en escalier ; leur vérification est laissée en exercice.

Théorème 7.18. *Soient $a < b$ deux réels et $f\colon [a,b] \to \mathbb{R}$ une fonction continue par morceaux. Alors on a*

$$(b-a) \inf_{[a,b]} f \leqslant \int_a^b f(x)\,dx \leqslant (b-a) \sup_{[a,b]} f(x)\,dx \ .$$

Si f est de plus supposée continue, alors il existe $c \in [a,b]$ tel que $f(x) = \frac{1}{b-a}\int_a^b f(x)\,dx$.

Notons que l'inf et le sup dans le théorème ci-dessus sont bien définis puisqu'une fonction continue par morceaux sur $[a,b]$ est nécessairement bornée.

Démonstration. Par définition d'un inf et d'un sup, on a, pour tout $x \in [a,b]$, $\inf_{[a,b]} f \leqslant f(x) \leqslant \sup_{[a,b]} f$. Par positivité de l'intégrale, on en déduit

$$\int_a^b \inf_{[a,b]} f\,dx \leqslant \int_a^b f(x)\,dx \leqslant \int_a^b \sup_{[a,b]} f\,dx \ .$$

Ceci prouve l'inégalité désirée. Si maintenant f est continue sur $[a,b]$, alors l'inf et le sup sont un min et un max puisqu'une fonction continue sur un intervalle fermé borné est bornée et atteint ses bornes sur cet intervalle. Appelons m le minimum de f sur $[a,b]$ et M le maximum de f sur $[a,b]$. On a alors $f([a,b]) = [m,M]$ par le théorème des valeurs intermédiaires, et l'inégalité ci-dessus donne

$$m \leqslant \frac{1}{b-a}\int_a^b f(x)\,dx \leqslant M \ .$$

Ceci achève la démonstration. □

Théorème 7.19. *Soient $a < b$ deux réels, et f une fonction continue par morceaux sur $[a,b]$ et à valeurs positives. Alors $\displaystyle\int_a^b f(x)\,dx = 0$ si, et seulement si, f est nulle partout sauf peut-être en un nombre fini de points de $[a,b]$.*

Démonstration. Si f est continue par morceaux et nulle partout sauf peut-être en a et en b, alors il suit de la définition de l'intégrale que $\int_a^b f(x)dx = 0$; et si f est nulle sauf peut-être en $a_1 < a_2 < \ldots, < a_n$ alors on peut utiliser la relation de Chasles pour écrire

$$\int_a^b f(x)dx = \int_a^{a_1} f(x)dx + \int_{a_1}^{a_2} f(x)dx + \ldots + \int_{a_n}^b f(x)dx = 0 + 0 + \ldots + 0 = 0 \ .$$

Réciproquement, supposons f d'intégrale nulle. Commençons par le cas où f est continue ; si f prend une valeur strictement positive en $x_0 \in [a,b]$, alors il existe un intervalle $[c,d] \subseteq [a,b]$ avec $c < d$ sur lequel f est à valeurs strictement positives, donc le minimum de f sur $[c,d]$ est strictement positif, et le théorème précédent nous permet de conclure que $\displaystyle\int_c^d f(x)\,dx > 0$; la relation de Chasles et la positivité de l'intégrale nous donnent alors :

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^d f(x)\,dx + \int_d^b f(x)\,dx > 0 \ .$$

On vient de montrer que, si f est continue, à valeurs positives et prend une valeur non nulle, alors son intégrale est strictement positive ; autrement dit, si $\displaystyle\int_a^b f(x)\,dx = 0$ et f est continue à valeurs positives sur $[a,b]$ alors f est la fonction nulle.

Reste le cas où f n'est que supposée continue par morceaux : alors il existe $a_1 < \ldots < a_n$ avec $a_1 = a$, $a_n = b$ tels que f coïncide sur chaque intervalle $]a_i, a_{i+1}[$ avec une fonction f_i qui se prolonge continûment à $[a_i, a_{i+1}]$, et on a

$$\int_a^b f(x)\,dx = \sum_{i=1}^{n-1} \int_{a_i}^{a_{i+1}} f_i(x)\,dx \ .$$

Par positivité de l'intégrale, si $\int_a^b f(x)\,dx = 0$ alors on doit avoir $\int_{a_i}^{a_{i+1}} f_i(x)\,dx = 0$ pour tout $i \in \{1, \dots, n-1\}$; comme les f_i sont continues on en déduit que chaque f_i est nulle sur $[a_i, a_{i+1}]$. Par conséquent, f est nulle partout sauf peut-être en $a_1, \dots, a_n$. □

Le théorème suivant peut s'avérer très utile pour comprendre le comportement des intégrales de produits de fonctions, en particulier lorsqu'on étudie des *intégrales généralisées* (ce qu'on ne fera pas ce semestre !).

Théorème 7.20 (Première formule de la moyenne)**.** *Soient $a < b$ deux réels, $f\colon [a,b] \to \mathbb{R}$ une fonction continue et $g\colon [a,b] \to \mathbb{R}$ une fonction continue par morceaux et à valeurs positives. Alors il existe $c \in [a,b]$ tel que*

$$\int_a^b f(x)g(x)\,dx = f(c)\int_a^b g(x)\,dx\ .$$

Démonstration. Comme g est continue par morceaux et à valeurs positives, l'intégrale de g ne peut valoir 0 que si g est nulle partout sauf peut-être en un nombre fini de points, auquel cas il en va de même de fg, dont l'intégrale est donc nulle elle aussi. Par conséquent, tout $c \in [a,b]$ est tel que $\int_a^b f(x)g(x)\,dx = f(c)\int_a^b g(x)\,dx$ dans ce cas (très) particulier.

On peut donc supposer que $\int_a^b g(x)\,dx \neq 0$. Pour tout $x \in [a,b]$, $m \leqslant f(x) \leqslant M$ avec $[m,M] = [\inf_{[a,b]} f, \sup_{[a,b]} f]$.

Par positivité et linéarité de l'intégrale,

$$\int_a^b mg(x)\,dx \leqslant \int_a^b f(x)g(x)\,dx \leqslant \int_a^b Mg(x)\,dx.$$

Or f est continue et le théorème des valeurs intermédiaires nous assure que toute valeur entre m et M est atteinte. Par conséquent il existe $c \in [a,b]$ tel que $m \leqslant f(c) = \frac{\int_a^b f(x)g(x)\,dx}{\int_a^b g(x)\,dx} \leqslant M$.

□

Théorème 7.21. *Soient $a < b$ deux réels, et f une fonction continue par morceaux sur $[a,b]$. Alors la fonction F définie sur $[a,b]$ par $F(t) = \int_a^t f(x)\,dx$ est continue sur $[a,b]$.*

Démonstration. Une fonction continue par morceaux est nécessairement bornée, autrement dit il existe M tel que $|f(x)| \leqslant M$ pour tout $x \in [a,b]$. On a alors, pour tout $s,t \in [a,b]$:

$$|F(t) - F(s)| = \left|\int_s^t f(x)\,dx\right| \leqslant \left|\int_s^t |f(x)|\,dx\right| \leqslant M|t-s|\ .$$

En particulier, $F(s)$ tend vers $F(t)$ quand s tend vers t, donc F est continue sur $[a,b]$ [ii] □

Il serait tentant de penser que la fonction F définie ci-dessus est toujours dérivable, et que $F' = f$. C'est faux en général : le théorème de Darboux nous dit qu'une fonction dérivée vérifie la conclusion du théorème des valeurs intermédiaires (*i.e.* l'image d'un intervalle par une fonction dérivée est un intervalle) donc une fonction en escalier non constante ne peut jamais être une dérivée. Il existe néanmoins un cas essentiel où ce résultat est vrai.

Théorème 7.22 (Théorème fondamental de l'analyse)**.** *Soient $a < b$ deux réels, f une fonction continue sur $[a,b]$, et F la fonction définie sur $[a,b]$ par $F(t) = \int_a^t f(x)\,dx$. Alors F est dérivable sur $[a,b]$, $F' = f$, et F est l'unique primitive de f sur $[a,b]$ qui s'annule en a.*

Démonstration. Commençons par montrer que $F' = f$; pour cela, fixons $t \in [a,b]$. Pour tout $s \in [a,b]$, on a :

$$F(t) - F(s) - (t-s)f(t) = \int_s^t f(x)\,dx - \int_s^t f(t)\,dx = \int_s^t (f(x) - f(t))\,dx\ .$$

ii. On vient en fait de montrer que F est *lipschitzienne* sur $[a,b]$.

Fixons $\varepsilon > 0$. Comme f est continue en t, il existe $\delta > 0$ tel que, pour tout $s \in [a,b]$, on ait

$$|t-s| \leqslant \delta \Rightarrow |f(t) - f(s)| \leqslant \varepsilon \ .$$

Notons que si $|t-s| \leqslant \delta$ alors $|t-x| \leqslant \delta$ pour tout x appartenant au segment d'extrémités t et s. Par conséquent, $|f(t) - f(x)| \leqslant \varepsilon$ pour tout x appartenant à ce segment, et l'inégalité triangulaire nous donne, pour tout s tel que $|t-s| \leqslant \delta$:

$$\left| \int_s^t (f(x) - f(t))\, dx \right| \leqslant \varepsilon |t-s| \ .$$

On obtient donc finalement, pour tout $s \in [a,b]$ tel que $|t-s| \leqslant \delta$, que

$$\left| \frac{F(t) - F(s)}{t-s} - f(t) \right| \leqslant \varepsilon \ .$$

Ceci prouve que $\lim_{s \to t} \frac{F(s)-F(t)}{s-t} = f(t)$, autrement dit que F est dérivable en t et $F'(t) = f(t)$.

Si maintenant G est une autre primitive de f sur $[a,b]$ qui s'annule en a, alors $(G-F)' = 0$, donc l'égalité des accroissements finis appliquée à $G-F$ (qui est continue et dérivable sur $[a,b]$) entraîne que $G-F$ est constante sur $[a,b]$. Comme $G(a) = F(a) = 0$ par hypothèse, on obtient bien que $G(x) = F(x)$ pour tout x de $[a,b]$. □

Exercice 7.23. En utilisant le théorème fondamental de l'analyse, donner une nouvelle démonstration du théorème 7.19.

On peut maintenant se rappeler de notre technique habituelle pour calculer des intégrales : utiliser les primitives (technique qui ne marche, hélas, pas toujours : parfois on ne connaît pas de primitive de la fonction qu'on souhaite intégrer !).

Corollaire 7.24. Soient $a < b$ deux réels, $f\colon [a,b] \to \mathbb{R}$ une fonction continue sur $[a,b]$ et F une primitive de f sur $[a,b]$. Alors on a $\int_a^b f(x)\, dx = F(b) - F(a)$.

On note souvent $[F]_a^b$ la quantité $F(b) - F(a)$.

Démonstration. Si F est une primitive de f sur $[a,b]$, alors $F - F(a)$ est encore une primitive de f sur $[a,b]$, par conséquent le théorème précédent nous donne que, pour tout $t \in [a,b]$, on a $F(t) - F(a) = \int_a^t f(x)\, dx$. On obtient le résultat désiré en appliquant cette formule pour $t = b$. □

Corollaire 7.25 (Formule d'intégration par parties)**.** Soient $a < b$ deux réels et $f, g\colon [a,b] \to \mathbb{R}$ deux fonctions de classe C^1. Alors on a

$$\int_a^b f(x)g'(x)\, dx = [fg]_a^b - \int_a^b f'(x)g(x)\, dx \ .$$

Démonstration. On utilise la formule $(fg)' = f'g + g'f$. Comme $f'g + g'f$ est continue, et a pour primitive fg, on peut lui appliquer le résultat précédent et obtenir (par linéarité de l'intégrale)

$$[fg]_a^b = \int_a^b f'(x)g(x)\, dx + \int_a^b f(x)g'(x)\, dx \ .$$

C'est la formule qu'on souhaitait démontrer. □

A titre d'exemple, utilisons une intégration par parties pour déterminer une primitive de la fonction ln sur $]0, +\infty[$, par exemple la primitive F qui s'annule en 1 : pour $x > 0$ on a

$$\begin{aligned} F(x) &= \int_1^x \ln(t) dt \\ &= [t \ln(t)]_1^x - \int_1^x 1 dt \\ &= x \ln(x) - x + 1. \end{aligned}$$

On voit ainsi que $x \mapsto x\ln(x) - x$ est une primitive de ln, ce qu'on vérifie facilement en dérivant :

$$\forall x > 0, \quad (x\ln(x) - x)' = \ln(x) + x\frac{1}{x} - 1 = \ln(x) \ .$$

Un autre exemple : essayons de trouver une primitive de $x \arctan(x)$. En dérivant arctan et en intégrant x (toutes les fonctions en jeu sont $\mathcal{C}^\infty$ et on peut donc bien intégrer par parties) on obtient

$$\begin{aligned}\int_0^x t \arctan(t) dt &= \left[\arctan(t)\frac{t^2}{2}\right]_0^x - \int_0^x \frac{t^2}{2(1+t^2)} dt \\ &= \frac{x^2}{2}\arctan(x) - \frac{1}{2}\int_0^x \left(1 - \frac{1}{1+t^2}\right) dt \\ &= \frac{1+x^2}{2}\arctan(x) - \frac{x}{2} \ .\end{aligned}$$

Le fait que, pour intégrer $\frac{t^2}{1+t^2}$, on ait eu besoin de décomposer cette fraction rationnelle en éléments simples, n'est pas un hasard : c'est toujours comme cela qu'on intègrera les fractions rationnelles.

Corollaire 7.26 (Formule de changement de variables)**.** Soient $a < b, c < d$ quatre réels, f une fonction continue sur $[a, b]$ à valeurs dans $\mathbb{R}$ et φ une fonction de classe C^1 sur $[c, d]$ et telle que $\varphi([c, d]) \subseteq [a, b]$. Alors on a

$$\int_c^d f(\varphi(t))\varphi'(t)\, dt = \int_{\varphi(c)}^{\varphi(d)} f(x)\, dx \ .$$

Démonstration. Soit F une primitive de f sur $[a, b]$. Alors, par la formule de dérivation des fonctions composées, $F \circ \varphi$ est une primitive de $(f \circ \varphi)\varphi'$ sur $[c, d]$, et le théorème fondamental de l'analyse nous permet d'écrire :

$$\int_c^d f(\varphi(t))\varphi'(t)\, dt = [F \circ \varphi]_c^d = F(\varphi(d)) - F(\varphi(c)) = \int_{\varphi(c)}^{\varphi(d)} f(x)\, dx \ .$$

□

Il est **très important**, pour la formule de changement de variables écrite ci-dessus, que f soit continue et que φ soit de classe C^1. Il n'est pas nécessaire, par contre, que φ soit une bijection de $[c, d]$ sur son image. Souvent, on applique la formule ci-dessus en « partant de la droite vers la gauche », i.e. on veut poser $x = \varphi(t)$. Il est alors un peu délicat de trouver les bonnes bornes pour l'intégrale de gauche, sauf dans le cas où φ est une bijection. Ce cas particulier est particulièrement utile en pratique.

Corollaire 7.27 (Cas particulier de la formule de changement de variables)**.** Soient $a < b, c < d$ quatre réels, f une fonction continue sur $[a, b]$ à valeurs dans $\mathbb{R}$, et φ une bijection de classe C^1 de $[c, d]$ sur $[a, b]$. Alors on a

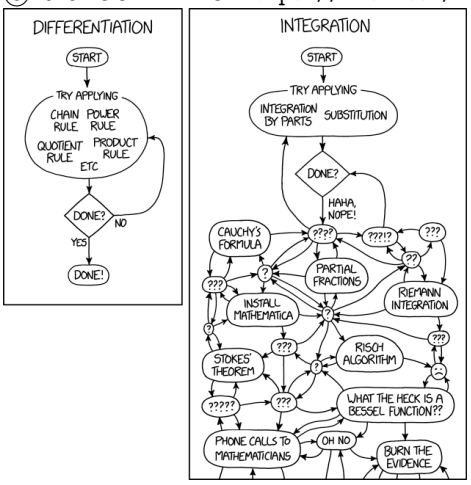
$$\int_a^b f(t)\, dt = \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} f(\varphi(x))\varphi'(x)\, dx \ .$$

Un autre intérêt de la formule ci-dessus est qu'elle se généralise au cas où f est continue par morceaux et aux intégrales de fonctions de plusieurs variables.

On verra quelques exemples dans la dernière section de ce chapitre ; pour l'instant, utilisons un changement de variables pour calculer l'aire d'un disque de centre 0 et de rayon $R > 0$. Ce disque est délimité par les courbes $y = \sqrt{R^2 - x^2}$ et $y = -\sqrt{R^2 - x^2}$ pour $x \in [-R, R]$, donc son aire A est égale à $2\displaystyle\int_{-R}^{R} \sqrt{R^2 - x^2} dx$. On utilise le changement de variables de classe $\mathcal{C}^1$ $x = R\cos(\theta)$, où $\theta \in [0, \pi]$; $\varphi\colon [0, \pi] \to [-R, R]$ est une bijection, on a $\varphi^{-1}(-R) = \pi$, $\varphi^{-1}(R) = 0$,et $\varphi'(\theta) = -R\sin(\theta)$.

La formule du changement de variable nous donne donc

$$\begin{aligned} A &= 2\int_{\pi}^{0} \sqrt{R^2 - R^2\cos^2(\theta)}(-R\sin(\theta))d\theta \\ &= 2\int_{0}^{\pi} (R\sin(\theta))(R\sin(\theta))d\theta \\ &= 2R^2\int_{0}^{\pi} \sin^2(\theta)d\theta \\ &= 2R^2\int_{0}^{\pi} \frac{1-\cos(2\theta)}{2}d\theta \\ &= R^2\left[\theta - \frac{1}{2}\sin(2\theta)\right]_{0}^{\pi} \\ &= \pi R^2 \ . \end{aligned}$$



7.4 Intégration de fonctions à valeurs complexes

Définition 7.28. Soit $a < b$ deux réels et $f\colon [a,b] \to \mathbb{C}$. On dit que f est *intégrable* sur $[a,b]$ si $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$ sont toutes deux intégrables, et dans ce cas on pose

$$\int_a^b f(x)dx = \int_a^b \mathrm{Re}(f)(x)dx + i\int_a^b \mathrm{Im}(f)(x)dx \ .$$

En particulier, si $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$ sont toutes deux continues par morceaux, on peut intégrer f sur $[a,b]$. La raison pour laquelle on souhaite parfois intégrer des fonctions à valeurs complexes dans ce cours est qu'elles peuvent nous permettre de simplifier des calculs d'intégrales de fonctions de $[a,b]$ dans $\mathbb{R}$.

Proposition 7.29 (Théorème fondamental de l'analyse pour des fonctions à valeurs dans $\mathbb{C}$)**.** *Si $f\colon [a,b] \to \mathbb{C}$ est telle que $Im(f)$ et $Re(f)$ sont toutes deux continues, et qu'on définit $F(x) = \int_a^x f(t)dt$, alors $F\colon [a,b] \to \mathbb{C}$ est dérivable et $F'(x) = f(x)$ pour tout $x \in [a,b]$.*

Pour démontrer ce résultat, il suffit de l'appliquer à la fois à la partie réelle et à la partie imaginaire de f. Il nous est utile pour les calculs : grâce à lui, les théorèmes que nous avons démontré à partir du théorème fondamental de l'analyse, en particulier la formule d'intégration par parties et la formule du changement de variables, restent vrais pour des fonctions à valeurs dans $\mathbb{C}$.

7.5 Quelques calculs de primitives et d'intégrales

Quelques exemples

On utilise souvent des intégrations par parties pour établir des relations de récurrence entre suites d'intégrales, par exemple considérons $I_n = \int_0^1 x^n e^x dx$. En intégrant par partie (toutes les fonctions impliquées sont de classe $\mathcal{C}^\infty$) on a pour tout $n \geqslant 1$ que

$$I_n = [x^n e^x]_0^1 - n\int_0^1 x^{n-1}e^x = e - nI_{n-1} \ .$$

il est clair que $I_0 = [e^x]_0^1 = e - 1$; on en déduit de proche en proche (écrire une formule pour I_n serait un bon exercice !) que

$$I_1 = e - (e-1) = 1; \quad I_2 = e - 2 \quad \text{etc.}$$

De la même façon, calculons $I = \int_0^\pi t^2 \sin(t)dt$:

$$\begin{aligned} I &= \int_0^\pi t^2 \sin(t) \\ &= \left[-t^2\cos(t)\right]_0^\pi + 2\int_0^\pi t\cos(t)dt \\ &= 2\pi^2 + 2\left(\left[t\sin(t)\right]_0^\pi - \int_0^\pi \sin(t)dt\right) \\ &= 2\pi^2 - 2. \end{aligned}$$

Passer par les complexes est parfois utile, en particulier pour calculer une intégrale d'une fonction de la forme $e^{ax}\cos(bx)$ ou $e^{ax}\sin(bx)$; par exemple, calculons $\int_0^t e^x\cos(x)$: c'est la partie réelle de

$$\begin{aligned} \int_0^t e^x e^{ix}dx &= \int_0^t e^{(1+i)x}dx \\ &= \frac{1}{1+i}\left[e^{(1+i)x}\right]_0^t \\ &= \frac{e^{(1+i)t}-1}{1+i} \\ &= \frac{(e^{(1+i)t}-1)(1-i)}{2} \\ &= \frac{e^t\cos(t)-1+e^t\sin(t)+i(e^t\sin(t)-e^t\cos(t)+1)}{2}. \end{aligned}$$

On vient d'obtenir

$$\int_0^t \cos(x)e^x dx = \frac{e^t\cos(t)-1+e^t\sin(t)}{2} \text{ et } \int_0^t e^x\sin(x)dx = \frac{e^t\sin(t)-e^t\cos(t)+1}{2}.$$

Exercice 7.30. A l'aide de deux intégrations par parties bien choisies, trouver la valeur de $\int_0^t e^x\cos(x)\,dx$ pour $t \in \mathbb{R}$.

Soit $I(t) = \int_0^t e^x\cos(x)\,dx$ et $J(t) = \int_0^t e^x\sin(x)\,dx$, par intégration par partie, on a $I(t) = [e^x\cos(x)]_0^t + J(t)$ et $J(t) = [e^x\sin(x)]_0^t - I(t)$ et on retrouve les résultats précédents.

Quand une intégrale paraît compliquée à calculer, on peut parfois essayer de la simplifier à l'aide d'un changement de variables ; par exemple, calculons

$$I = \int_1^e \frac{\cos(\ln(x))}{x}dx \ .$$

La fonction intégrée est continue, et on essaye le changement de variables $\mathcal{C}^1$ $\varphi(x) = \ln(x)$; autrement dit on pose $u = \ln(x)$, $du = \frac{dx}{x}$, et on obtient

$$I = \int_0^1 \cos(u)du = \sin(1) \ .$$

(On aurait aussi pu reconnaître tout de suite qu'on intégrait la dérivée de $\sin(\ln(t))$!)

Polynômes trigonométriques

Quand on doit intégrer un polynôme en $\sin, \cos$, on se retrouve face à une somme d'intégrales de la forme $\int \cos^n(x)\sin^m(x)\,dx$. Le cas le plus sympathique est quand m ou n est impair ; si par exemple $m = 2k+1$, on écrit $\sin^m(x) = \sin^{2k}(x)\sin(x) = (1-\cos^2(x))^k\sin(x)$, et on est en situation pour utiliser le changement de

variables de classe $\mathcal{C}^1$, $u = \cos(x)$, ce qui donne $P\big(\cos(x)\big)\sin(x)\,dx = P(u)\,du$. De même, si n est impair on peut utiliser le changement de variables $u = \sin(x)$. Par exemple, calculons

$$\begin{aligned}
\int_0^t \sin^7(x)\cos^{12}(x)dx &= \int_0^t (1-\cos^2(x))^3\cos^{12}(x)(\sin(x)dx)\\
&= \int_1^{\cos(t)} (1-u^2)^3u^{12}(-du)\\
&= \int_{\cos(t)}^1 (1-3u^2+3u^4-u^6)u^{12}du\\
&= \left[\frac{u^{13}}{13}-\frac{3u^{15}}{15}+\frac{3u^{17}}{17}-\frac{u^{19}}{19}\right]_{\cos(t)}^1\\
&= \frac{1}{13}-\frac{1}{5}+\frac{3}{17}-\frac{1}{19}-\frac{\cos^{13}(t)}{13}+\frac{\cos^{15}(t)}{5}-\frac{3\cos^{17}(t)}{17}+\frac{\cos^{19}(t)}{19}\\
&= \frac{16}{20995}-\frac{\cos^{13}(t)}{13}+\frac{\cos^{15}(t)}{5}-\frac{3\cos^{17}(t)}{17}+\frac{\cos^{19}(t)}{19}.
\end{aligned}$$

Notons que, si l'on est intéressé par le calcul d'une primitive, le calcul de la constante $\frac{16}{20995}$ ci-dessus est parfaitement inutile.

Le cas le plus pénible est celui où n et m sont pairs ; il faut alors arriver à linéariser $\cos^n(x)\sin^m(x)$, par exemple en passant par les nombres complexes. A titre d'exemple, calculons

$$\begin{aligned}
\int_0^t \cos^4(x)dx &= \int_0^t \left(\frac{e^{ix}+e^{-ix}}{2}\right)^4 dx\\
&= \int_0^t \frac{e^{4ix}+4e^{3ix}e^{-ix}+6e^{2ix}e^{-2ix}+4e^{ix}e^{-3ix}+e^{-4ix}}{16}dx\\
&= \int_0^t \frac{e^{4ix}+e^{-4ix}+4(e^{2ix}+e^{-2ix})+6}{16}dx\\
&= \int_0^t \frac{2\cos(4x)+8\cos(2x)+6}{16}dx\\
&= \frac{\sin(4t)}{32}+\frac{\sin(2t)}{4}+\frac{3t}{8}\ .
\end{aligned}$$

Fractions rationnelles

Pour trouver des primitives de fractions rationnelles, on commence par décomposer en éléments simples (4.5), puis on intègre séparément chacun des éléments simples.

- Les éléments simples de la forme $\frac{1}{X-\alpha}$ ont comme primitive $x \mapsto \ln|x-\alpha|+K$, $K \in \mathbb{R}$.
- Les éléments simples de la forme $\frac{1}{(X-\alpha)^k}$ avec $k>1$ ont comme primitive $-\frac{1}{k-1}\frac{1}{(X-\alpha)^{k-1}}+K$, $K \in \mathbb{R}$.
- Un seul type d'élément simple est difficile à intégrer : $\dfrac{\lambda X+\mu}{P(X)^n}$, où $\lambda,\mu \in \mathbb{R}$ et $P(X)=aX^2+bX+c$ est un polynôme de degré 2 sans racine réelles, $\Delta = b^2-4ac<0$.

La première chose à faire est de s'occuper du terme linéaire au numérateur en faisant apparaitre au numérateur la dérivée $2aX+b$ du dénominateur :

$$\frac{\lambda X+\mu}{(aX^2+bX+c)^k} = \frac{\lambda}{2a}\frac{2aX+b}{(aX^2+bX+c)^k}+\frac{\mu-\frac{\lambda}{2a}b}{(aX^2+bX+c)^k}.$$

La première partie a comme primitive la fraction rationnelle $-\frac{\lambda}{2a}\frac{1}{(k-1)(aX^2+bX+c)^{k-1}}$ si $k>1$ et la fonction $x \mapsto \frac{\lambda}{2a}\ln|ax^2+bx+c|$ pour $k=1$.

Le changement affine $X=\alpha Y+\beta$ avec $\alpha=\frac{\sqrt{-\Delta}}{2a}$ et $\beta=-\frac{b}{2a}$, permet d'obtenir $P(X)=-\frac{\Delta}{4a}(Y^2+1)$. N'apprenez pas cette formule par cœur ! En pratique, faites une translation de $-\frac{b}{2a}$ pour centrer la parabole, puis faire un autre changement de variable multiplicatif pour se retrouver sur une forme $\nu(Y^2+1)$. On peut

résumer cette opération par la forme canonique d'un trinome qui donne son sens à la formule classique des solutions $\frac{-b\pm\sqrt{\Delta}}{2a}$:

$$aX^2+bX+c = a\left(\big(X+\frac{b}{2a}\big)^2-\frac{\Delta}{4a^2}\right).$$

On a donc simplement besoin d'intégrer $x\mapsto\dfrac{1}{(1+x^2)^k}$.

Pour $k=1$ une primitive est la fonction arctan. Pour $k>1$, c'est plus difficile ! Une façon de faire est d'utiliser le changement de variable (de classe $\mathcal{C}^1$ et bijectif, voir 7.26) $x=\tan(u)$, pour obtenir

$$\begin{aligned}\int_0^t \frac{1}{(1+x^2)^k}dx &= \int_0^{\arctan(t)} \frac{1}{(1+\tan^2(u))^k}(1+\tan^2(u))du\\ &= \int_0^{\arctan(t)} \frac{1}{(1+\tan^2(u))^{k-1}}du\\ &= \int_0^{\arctan(t)} \cos^{2k-2}(u)du\ .\end{aligned}$$

On est ainsi ramené au calcul d'une primitive de $\cos^{2k-2}$, qu'on peut mener à bien en linéarisant comme on a vu à la section précédente.

Par exemple, on obtient ainsi

$$\begin{aligned}\int_0^t \frac{1}{(1+x^2)^2}dx &= \int_0^{\arctan(t)} \cos^2(u)du\\ &= \int_0^{\arctan(t)} \frac{\cos(2u)+1}{2}du\\ &= \frac{\sin\big(2\arctan(t)\big)}{4}+\frac{\arctan(t)}{2}\\ &= \frac{\sin\big(\arctan(t)\big)\cos\big(\arctan(t)\big)+\arctan(t)}{2}\ .\end{aligned}$$

On n'a pas fini : il nous reste à simplifier les expressions $\sin(\arctan(t))$ et $\cos(\arctan(t))$. Pour se faire, on peut (on doit ?) se rappeler que $\cos^2(u)=\dfrac{1}{1+\tan^2(u)}$, qui nous donne

$$\cos^2(\arctan(t))=\frac{1}{1+t^2}.$$

Puisque cos est positif sur $]-\frac{\pi}{2},\frac{\pi}{2}[$, $\cos(\arctan(t))$ est positif et $\cos(\arctan(t))=\dfrac{1}{\sqrt{1+t^2}}$. On a donc

$$\sin^2(\arctan(t))=\frac{t^2}{1+t^2}\quad\text{d'où}\quad\sin(\arctan(t))=\frac{t}{\sqrt{1+t^2}}$$

en effet sin et arctan étant toutes les deux des fonctions impaires du même signe que t sur l'intervalle considéré, leur composée $\sin(\arctan(t))$ est du même signe que t. Finalement, on arrive à

$$\int_0^t \frac{1}{(1+x^2)^2}dx=\frac{t}{2(1+t^2)}+\frac{\arctan(t)}{2}\ .$$

Exercice 7.31. En utlisant la relation

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

et une intégration par parties bien choisie, retrouver la formule donnant une primitive de $x\mapsto\dfrac{1}{(1+x^2)^2}$.

Un autre exemple : déterminons une primitive sur $]-\infty,1[$ ou sur $]1,+\infty[$ de $x \mapsto \dfrac{1}{x^3-1}$. On commence par écrire

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

Pour calculer a on évalue en 1 :

$$a = \frac{1}{1+1+1} = \frac{1}{3}.$$

Pour calculer b et c on pourrait multiplier par $(X-j)$ (où $j = e^{2i\pi/3}$) et évaluer en j ; ou évaluer en 0 pour obtenir $-1 = -a+c$ et donc $c = -\dfrac{2}{3}$, puis utiliser que la limite en $+\infty$ de $\frac{x}{x^3-1}$ vaut 0 pour arriver à $0 = a+b$ donc $b = -\dfrac{1}{3}$. Finalement, on doit donc calculer une primitive de

$$x \mapsto \frac{1}{3}\left(\frac{1}{x-1} - \frac{x+2}{x^2+x+1}\right) .$$

Sur $]1,+\infty[$, une primitive de $x \mapsto \frac{1}{x-1}$ est $x \mapsto \ln(x-1)$, et une primitive de $x \mapsto \frac{x+\frac{1}{2}}{x^2+x+1}$ est $x \mapsto \frac{1}{2}\ln(x^2+x+1)$. Modulo les constantes multiplicatives, il nous reste à intégrer

$$\begin{aligned}\frac{1}{(x^2+x+1)} &= \frac{1}{(x+\frac{1}{2})^2+\frac{3}{4}} \\ &= \frac{4}{3}\frac{1}{1+\frac{4}{3}(x+\frac{1}{2})^2} \\ &= \frac{4}{3}\frac{1}{1+\left(\frac{2}{\sqrt{3}}(x+\frac{1}{2})\right)^2}.\end{aligned}$$

On reconnaît (?) la dérivée de $\dfrac{2}{\sqrt{3}}\arctan(\dfrac{2}{\sqrt{3}}(x+\dfrac{1}{2}))$.

Finalement, une primitive de notre fonction sur $]1,+\infty[$ est

$$x \mapsto \frac{1}{3}\left(\ln(x-1) - \ln(\sqrt{x^2+x+1}) - \sqrt{3}\arctan\left(\frac{2x+1}{\sqrt{3}}\right)\right) .$$

Fractions rationnelles trigonométriques

Cette fois-ci, on essaie d'intégrer des fonctions de la forme $F(t) = \dfrac{P(\sin(t),\cos(t))}{Q(\sin(t),\cos(t))}$ où P et Q sont des polynômes. On peut toujours se ramener à une intégrale de fraction rationnelle, via l'application des *règles de* Bioche :

- Si $F(t)dt$ est laissé inchangé par le changement de variables $u = -t$ (autrement dit si F est impaire) alors on pose $u = \cos(t)$.
- Si $F(t)dt$ est laissé inchangé par le changement de variables $u = \pi - t$, alors on pose $u = \sin(t)$.
- Si $F(t)dt$ est laissé inchangé par le changement de variables $u = \pi + t$, alors on pose $u = \tan(t)$ (en se restreignant à un intervalle où ce changement de variables a un sens : il faut que tan soit définie sur tout l'intervalle d'intégration !).
- Si aucun des points précédents n'est vérifié, alors on pose $u = \tan(\frac{t}{2})$ (sur un intervalle où le changement de variables a un sens, et en se rendant compte que le calcul risque d'être long...)

Exercice 7.32. Exprimer $\sin(t)$ et $\cos(t)$ en fonction de $\tan\left(\frac{t}{2}\right)$ (on pourra commencer par démontrer que $\tan(2a) = \frac{2\tan(a)}{1-\tan^2(a)}$)

Contentons-nous de traiter un exemple : déterminons une primitive de $f\colon x \mapsto \frac{1}{\cos^4(x)+\sin^4(x)}$, qui est bien définie et de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ puisque le dénominateur ne s'annule jamais. Les règles de Bioche et les formules $\cos(\pi+t) = -\cos(t)$, $\sin(\pi+t) = -\sin(t)$ nous disent que, puisque $f(t)dt$ est laissé invariant par le changement de variables $u = \pi + t$, on peut essayer de poser $u = \tan(t)$; ceci sur un intervalle où tan est bien définie, par

exemple $I =]-\frac{\pi}{2}, \frac{\pi}{2}[$ (on utilisera ensuite la périodicité de la fonction pour trouver une primitive sur $\mathbb{R}$ tout entier à partir d'une primitive sur I). Ensuite, on écrit que

$$\begin{aligned}\frac{dt}{\cos^4(t)+\sin^4(t)} &= \frac{dt}{\cos^4(t)(1+\tan^4(t))}\\ &= \frac{dt}{\cos^2(t)}\frac{1}{\cos^2(t)}\frac{1}{1+\tan^4(t)}\\ &= \frac{dt}{\cos^2(t)}\frac{1+\tan^2(t)}{1+\tan^4(t)}\\ &= \frac{1+u^2}{1+u^4}du.\end{aligned}$$

Le changement de variables $u = \tan(t)$ nous amène donc à la formule, pour $x \in I$:

$$\int_0^x \frac{dt}{\cos^4(t)+\sin^4(t)} = \int_0^{\tan(x)} \frac{1+u^2}{1+u^4}du\ .$$

Là on est un peu désespéré de devoir décomposer en éléments simples $\frac{1+X^2}{1+X^4}$, mais on ne se laisse pas abattre et on écrit

$$\frac{X^2+1}{X^4+1} = \frac{X^2+1}{(X^2+\sqrt{2}X+1)(X^2-\sqrt{2}X+1)} = \frac{aX+b}{X^2+\sqrt{2}X+1} + \frac{cX+d}{X^2-\sqrt{2}X+1}\ .$$

Pour calculer a, b, c, d on peut passer par les complexes en évaluant en $e^{i\pi/4}$ (recommandé !) et utiliser la parité de la fonction pour voir que $a = c$, évaluer en 0 pour déduire que $b + d = 1$; ou simplement remarquer que

$$\begin{aligned}\frac{1+X^2}{(X^2+\sqrt{2}X+1)(X^2-\sqrt{2}X+1)} &= \frac{1}{2}\frac{(X^2+\sqrt{2}X+1)+(X^2-\sqrt{2}X+1)}{(X^2+\sqrt{2}X+1)(X^2-\sqrt{2}X+1)}\\ &= \frac{1}{2}\left(\frac{1}{X^2+\sqrt{2}X+1}+\frac{1}{X^2-\sqrt{2}X+1}\right).\end{aligned}$$

Une fois qu'on a décomposé en éléments simples, il nous reste à trouver les primitives, qui s'obtiennent à partir des formes canoniques :

$$\begin{aligned}\frac{1}{X^2+\sqrt{2}X+1} &= \frac{1}{(X+\frac{\sqrt{2}}{2})^2+\frac{1}{2}}\\ &= \frac{2}{1+(\sqrt{2}X+1)^2}.\end{aligned}$$

Donc une primitive de $u \mapsto \dfrac{1}{u^2+\sqrt{2}u+1}$ est $\sqrt{2}\arctan(\sqrt{2}u+1)$; de même une primitive de $u \mapsto \dfrac{1}{u^2-\sqrt{2}u+1}$ est $-\sqrt{2}\arctan(-\sqrt{2}u+1)$. Finalement, on arrive à la formule

$$\begin{aligned}\int_0^t \frac{dt}{\cos^4(t)+\sin^4(t)} &= \int_0^{\tan(x)} \frac{1+u^2}{1+u^4}du\\ &= \int_0^{\tan(x)} \frac{1}{2}\left(\frac{1}{u^2+\sqrt{2}u+1}+\frac{1}{u^2-\sqrt{2}u+1}\right)\\ &= \frac{1}{2}\left[\sqrt{2}\arctan(\sqrt{2}u+1)-\sqrt{2}\arctan(-\sqrt{2}u+1)\right]_0^{\tan(x)}\\ &= \frac{\sqrt{2}}{2}\left(\arctan(1+\sqrt{2}\tan(x))-\arctan(1-\sqrt{2}\tan(x))\right).\end{aligned}$$

Après ces quelques menus efforts, nous avons établi qu'une primitive de $f\colon x \mapsto \frac{1}{\cos^4(x)+\sin^4(x)}$ sur l'intervalle $I =]-\frac{\pi}{2}, \frac{\pi}{2}[$ est $F\colon x \mapsto \frac{\sqrt{2}}{2}\left(\arctan(1+\sqrt{2}\tan(x))-\arctan(1-\sqrt{2}\tan(x))\right)$. Notons qu'une primitive de fonction continue et π-périodique sur $\mathbb{R}$ doit être continue et π-périodique sur $\mathbb{R}$; par conséquent F doit avoir une limite en $\pm\frac{\pi}{2}$ (savez-vous la calculer ?) et on a en fait trouvé une formule valable sur $\mathbb{R}$ tout entier (modulo les prolongements par continuité en $k\pi + \frac{\pi}{2}$ pour $k \in \mathbb{Z}$).

Chapitre 8

Comparaison locale de fonctions et formules de Taylor

8.1 o, O et équivalents

8.1.1 Le cas des fonctions

Définition 8.1. Soit I un intervalle de $\mathbb{R}$, a un point adhérent à I et $f, g\colon I \to \mathbb{R}$. On dit que f est *négligeable devant g en a* si :

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x \in I,\ |x-a| \leqslant \delta \Rightarrow |f(x)| \leqslant \varepsilon |g(x)|\ .$$

On note alors $f \underset{a}{=} o(g)$ ou $f(x) = \underset{x\to a}{o}\big(g(x)\big)$.

Quand a est clairement précisé, on note simplement $f = o(g)$. Une autre façon d'exprimer cette comparaison est de dire qu'il existe une *fonction* $\varepsilon : U_a \to \mathbb{R}$ d'un voisinage U_a de a qui tend vers 0 en a telle que $\forall x \in U_a, f(x) = \varepsilon(x) \times g(x)$.

On étend cette notion (et les suivantes) « à gauche » (respectivement « à droite ») quand on se restreint à un voisinage à gauche, $x \in I \cap]a-\delta, a]$ (respectivement à droite, $x \in I \cap [a, a+\delta[$) avec $\delta > 0$, et on note $f \underset{a^-}{=} o(g)$, resp. $f \underset{a^+}{=} o(g)$.

Il faut bien comprendre que cette notation rompt avec le comportement habituel de l'égalité, elle n'est plus symétrique. Il aurait été plus adéquat de noter $f \in o(g)$ pour signifier que f appartient à un certain ensemble de fonctions, mais l'histoire[i] en a décidé autrement. Il faut donc être vigilant avec cette notation et ne pas succomber aux tentations de raccourcis opératoires habituels, mais il y a certaines égalités qui restent vraies :

Proposition 8.2. *Au voisinage de a, pour des fonctions réelles $f, g, h, f_1, f_2, g_1, g_2$ définies sur ce voisinage,*

1. *Si $f = o(h)$ et $g = o(h)$, alors une combinaison linéaire l'est encore, $\forall \lambda, \mu \in \mathbb{R}, \lambda f + \mu g = o(h)$ (linéarité).*
2. *Si $f = o(g)$ et $g = o(h)$, alors $f = o(h)$ (transitivité).*
3. *Si $f_1 = o(g_1)$ et $f_2 = o(g_2)$, alors $f_1 \times f_2 = o(g_1 \times g_2)$.*
4. *Si $f = o(g)$ alors, $\frac{1}{g} = o(\frac{1}{f})$.*

Preuve: Avec les ε et δ : On écrit les définitions de chacune des comparaisons en indiçant les variables ε et δ.

1. Si $\lambda\mu \neq 0$, on choisit $\varepsilon_f = \frac{\varepsilon}{2|\lambda|}$ pour définir δ_f, $\varepsilon_g = \frac{\varepsilon}{2|\mu|}$ pour définir δ_g, et en prenant $\delta = \min(\delta_f, \delta_g)$, on obtient, pour tout $|x-a| \leqslant \delta$, $|\lambda f(x) + \mu g(x)| \leqslant \varepsilon |h(x)|$.
2. En prenant $\varepsilon_f = \varepsilon_g = \sqrt{\varepsilon}$ et $\delta = \min(\delta_f, \delta_g)$, on obtient, pour tout $|x-a| \leqslant \delta$, $|f(x| \leqslant \sqrt{\varepsilon}|g(x)| \leqslant \varepsilon |h(x)|$.
3. On obtient le résultat en prenant $\delta = \min(\delta_{f_1}, \delta_{f_2})$.
4. On se restreint au cas où f, g ne s'annulent pas au voisinage de a et $0 < |f(x)| \leqslant \varepsilon |g(x)| \iff 0 < |\frac{1}{g(x)}| \leqslant \varepsilon |\frac{1}{f(x)}|$.

Avec les fonctions ε_f tendant vers 0 en a sur un voisinage U_f de a (resp. pour les autres fonctions) :

i. Voir Notation de Landau

1. La fonction $\lambda\varepsilon_f + \mu\varepsilon_g$ tend vers 0 sur le voisinage $U_f \cap U_g$ et remplit son office.
2. Le produit $\varepsilon_f \times \varepsilon_g$ remplit son office.
3. Le produit $\varepsilon_{f_1} \times \varepsilon_{f_2}$ remplit son office.
4. $f(x) = \varepsilon(x)g(x) \iff \frac{1}{g(x)} = \varepsilon(x)\frac{1}{f(x)}$ partout où ces fonctions sont toutes définies. On voit d'ailleurs que la condition que f et g ne s'annulent pas au voisinage de a, utile dans la première démonstration n'est pas nécessaire.

□

Remarquons que, si g ne s'annule pas, alors dire que f est négligeable devant g revient à dire que le quotient $\frac{f(x)}{g(x)}$ tend vers 0 quand x tend vers x_0. En particulier, la notation $f \underset{x_0}{=} o(1)$ signifie simplement que $f(x)$ tend vers 0 quand x tend vers x_0.

La notation $f = o(g)$ est délicate à manipuler et il faut faire très attention au début ; par exemple, en 0, on a par exemple $x^2 = x^3 + o(1)$, mais aussi $x^2 = o(1)$, et de manière générale pour un polynôme $P \in \mathbb{R}[X]$, $P(x) + o(x^n) = P_n(x) + o(x^n)$ où $P_n \in \mathbb{R}_n[X]$ est la version tronquée à l'ordre n de P.

Exercice 8.3. Soit I un intervalle de $\mathbb{R}$, $x_0 \in I$ et $f\colon I \to \mathbb{R}$. Montrer que f est dérivable en x_0, et $f'(x_0) = l$, si et seulement si on peut écrire $f(x) \underset{x_0}{=} f(x_0) + (x - x_0)l + o(x - x_0)$.

Définition 8.4. Soit I un intervalle de $\mathbb{R}$, x_0 un point adhérent à I et $f, g\colon I \to \mathbb{R}$. On dit que f est *dominée par* g en x_0, et on note $f \underset{x_0}{=} O(g)$, si

$$\exists M > 0,\ \exists \delta > 0,\ \forall x \in I, |x - x_0| \leqslant \delta \Rightarrow |f(x)| \leqslant M|g(x)|\ .$$

Quand g ne s'annule pas sur I, $f = O(g)$ revient à dire que f/g est bornée au voisinage de x_0.

La notion la plus importante en pratique est celle qui est présentée dans la définition suivante.

Définition 8.5. Soit I un intervalle de $\mathbb{R}$, x_0 un point adhérent à I et $f, g\colon I \to \mathbb{R}$. On dit que f est *équivalente à* g en x_0, et on note $f \underset{x_0}{\sim} g$, s'il existe $\delta > 0$ et une fonction $h\colon I \cap [x_0 - \delta, x_0 + \delta] \to \mathbb{R}$ telle que $h(x)$ tend vers 1 lorsque x tend vers x_0, et $f(x) = h(x)g(x)$ pour tout $x \in I \cap [x_0 - \delta, x_0 + \delta]$.

Quand g ne s'annule pas, dire que $f \underset{x_0}{\sim} g$ revient à dire que $\frac{f}{g}(x)$ tend vers 1 quand x tend vers x_0 (et c'est comme ça qu'il faut souvent y penser).

Exercice 8.6. Montrer que $f \underset{x_0}{\sim} g$ si et seulement si $g \underset{x_0}{\sim} f$. Montrer que $f \underset{x_0}{\sim} g$ si et seulement si $g \underset{x_0}{=} f + o(f)$.

Les équivalents sont particulièrement utiles pour calculer des limites : si l est un réel **non nul**, alors dire que $f \underset{x_0}{\sim} l$ revient à dire que $f(x)$ tend vers l quand x tend vers x_0. Pourquoi distinguer le cas où l est nul ? Parce que, si $f \underset{x_0}{\sim} 0$, alors la définition d'un équivalent impose qu'il existe $\delta > 0$ tel que $f = 0$ sur $I \cap [x_0 - \delta, x_0 + \delta]$, c'est-à-dire que f restreinte à cet ensemble est la fonction nulle, ce qui est bien sur différent de dire que f tend vers 0 !

Ce problème se retrouve lorsqu'on essaye d'appliquer des opérations algébriques aux équivalents : si on peut sans danger mutliplier des équivalents, on ne peut en général pas les ajouter (ni les soustraire). Voyons un exemple : soit $f(x) = x$ et $g(x) = -x + x^2$, toutes deux définies sur $\mathbb{R}$. En 0, on a $f(x) \sim x$ et $g(x) \sim -x \sim -f(x)$; mais puisque $f(x) + g(x) = x^2$, on n'a pas $f(x) + g(x) \sim 0$!

De la même façon, on ne peut pas composer des équivalents sans faire un minimum attention. Dans le même esprit que ci-dessus, considérons les fonctions $f, g\colon]0, +\infty[$ to $\mathbb{R}$ définies par $f(x) = \frac{1}{x^2}$ et $g(x) = \frac{1}{x^2} + \frac{1}{x}$. Alors en 0 on a $f(x) \sim g(x)$; pourtant, $e^{g(x)}$ et $e^{f(x)}$ ne sont pas équivalentes, puisque $e^{g(x)} = e^{1/x}e^{f(x)}$ et $e^{1/x}$ ne tend pas vers 1 quand x tend vers 0^+.

On pourrait énoncer un théorème selon lequel on peut composer des équivalents quand les fonctions équivalentes sont "à gauche" de la composition ; mais ce théorème n'est pas très utile en pratique et on préférera composer des développements limités.

Notons avant de décrire brièvement le cas des suites que toutes ces définitions (o, O, $\sim$) auraient aussi un sens en $\pm\infty$; c'est un bon exercice d'écrire les définitions correspondantes (vous pouvez vous inspirer si nécessaire des définitions données ci-dessous pour les suites).

8.1.2 Le cas des suites

Définition 8.7. Soit u, v deux suites de nombres réels.

1. On dit que u est *négligeable* devant v, et on note $u = o(v)$, si
$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall n \geqslant N, |u(n)| \leqslant \varepsilon |v(n)|.$$
Si v ne s'annule pas, cela revient à dire que $u(n)/v(n)$ tend vers 0 quand n tend vers $+\infty$.
2. On dit que u et v sont *équivalentes*, et on note $u \sim v$, s'il existe une suite w telle que w_n tend vers 1 quand n tend vers $+\infty$, et N tel que pour tout $n \geqslant N$ on ait $u_n = w_n v_n$. Si v ne s'annule pas, cela revient à dire que $u(n)/v(n)$ tend vers 1 quand n tend vers $+\infty$.

Comme pour les fonctions, on peut mutliplier ou diviser des équivalents (en évitant de diviser par 0) mais on ne peut en général pas les ajouter ou les soustraire.

Exemple. Soit $P \in \mathbb{R}[X]$ un polynôme de degré $k \geqslant 0$, et a_k son coefficient dominant. Alors on a $P(n) \sim a_k n^k$.
En effet, on peut écrire
$$P(n) = a_k n^k \left(1 + \sum_{j=0}^{k-1} \frac{a_j}{a_k} n^{j-k}\right) .$$
Puisque $\displaystyle\sum_{j=0}^{k-1} \frac{a_j}{a_k} n^{j-k}$ tend vers 0 quand n tend vers $+\infty$ comme toute $(n^\alpha)_{n\in\mathbb{N}}$ pour $\alpha < 0$, on vient de montrer que $P(n) \sim a_k n^k$.

Exercice 8.8. Soit $F = P/Q$ une fraction rationnelle. Si on suppose que P est de degré k et de coefficient dominant a, et Q est de degré l et de coefficient dominant b, montrer que $P(n)/Q(n) \sim \frac{a}{b} n^{k-l}$.

8.2 Les formules de Taylor

On a vu que, si f est dérivable en x_0, alors on peut écrire $f(x) = f(x_0) + (x - x_0)f'(x) + o(x - x_0)$. Autrement dit, la fonction affine $x \mapsto f(x_0) + (x - x_0)f'(x_0)$ est une approximation de f en x_0 avec une erreur au plus de l'ordre de $x - x_0$. Il est très utile de pouvoir approcher f en x_0 avec une meilleure précision par des fonctions polynomiales, et c'est ce que permettent les formules de Taylor.

La première formule de Taylor est celle qui a les hypothèses les plus faibles (et la conclusion aussi la plus faible).

Théorème 8.9 (Formule de Taylor–Young)**.** *Soit I un intervalle de $\mathbb{R}$, $n \in \mathbb{N}^*$, et f une fonction $n-1$ fois dérivable sur I et telle que $f^{(n)}(x_0)$ existe. Alors on a*
$$f(x) = f(x_0) + (x - x_0)f'(x_0) + \ldots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + o\big((x - x_0)^n\big) .$$
En notation plus condensée :
$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + o((x - x_0)^n) .$$

En s'avançant un peu, on dira plus tard que $\displaystyle\sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + o((x - x_0)^n)$ est le *développement limité* de f en x_0 à l'ordre n (dans ce cas particulier, le développement limité est obtenu en calculant le *polynôme de* Taylor de f à l'ordre n), et $\displaystyle f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$ est le reste du développement limité ; la formule de Taylor-Young nous permet simplement de dire que le reste du développement limité à l'ordre n est un $o((x - x_0)^n)$.

Démonstration. Pour $n = 0$ la formule de Taylor–Young a encore du sens, c'est une reformulation de la continuité de f en x_0 et pour $n = 1$, elle signifie sa dérivabilité, comme on a déjà eu plusieurs fois l'occasion de le remarquer. Ceci constitue l'initialisation du raisonnement par récurrence que nous allons mener :

Supposons donc le théorème de Taylor–Young démontré jusqu'à un certain rang n, et f une fonction n fois dérivable sur I telle que $f^{(n+1)}(x_0)$ existe, c'est-à-dire vérifiant les hypothèses nécessaires à l'application du théorème au rang $n+1$; on va faire l'hypothèse supplémentaire que f' est continue pour simplifier l'argument (cela change la démonstration seulement pour le cas $n = 2$).

Par l'hypothèse de récurrence appliquée à f', qui vérifie les hypothèses au rang n, on sait qu'on peut trouver une fonction ε qui tend vers 0 en x_0 et telle que

$$\forall x \in I,\ f'(x) = \sum_{k=0}^{n} \frac{(f')^{(k)}(x_0)}{k!}(x-x_0)^k + (x-x_0)^n \varepsilon(x).$$

Autrement dit, pour tout $x \in I$ on a

$$f'(x) = \sum_{k=0}^{n} \frac{f^{(k+1)}(x_0)}{k!}(x-x_0)^k + (x-x_0)^n \varepsilon(x).$$

Notons que le reste $R(x) = (x-x_0)^n\varepsilon(x)$ est une fonction continue puisque c'est la différence de f' et d'une fonction polynôme. Comme ε tend vers 0 en x_0, pour tout $\epsilon > 0$ il existe $\delta > 0$ tel que $|x-x_0| \leqslant \delta \Rightarrow |R(x)| \leqslant \epsilon|(x-x_0)^n|$. Comme R est continue, on peut l'intégrer entre x_0 et x

$$f'(x) - \sum_{k=0}^{n} \frac{f^{(k+1)}(x_0)}{k!}(x-x_0)^k = R(x)\ ,$$

et on obtient

$$f(x) - f(x_0) - \sum_{k=0}^{n} \frac{f^{(k+1)}(x_0)}{(k+1)!}(x-x_0)^{k+1} = \int_{x_0}^{x} R(t)dt.$$

Cependant, pour x proche de x_0, on peut majorer cette intégrale :

$$\forall x \in I,\ |x-x_0| \leqslant \delta \Rightarrow \left|\int_{x_0}^{x} R(t)dt\right| \leqslant \int_{x_0}^{x} |R(t)|\,dt \leqslant \int_{x_0}^{x} |\epsilon(t-x_0)^n|\,dt = \epsilon\frac{|x-x_0|^{n+1}}{n+1}\ .$$

Ceci implique que, pour tout $\epsilon > 0$ et x dans I tel que $|x-x_0| \leqslant \delta$, on a

$$\left|f(x) - \left(f(x_0) + \sum_{k=0}^{n} \frac{f^{(k+1)}(x_0)}{(k+1)!}(x-x_0)^{k+1}\right)\right| \leqslant \frac{\epsilon}{n+1}|x-x_0|^{n+1}.$$

On vient donc de montrer que

$$f(x) = f(x_0) + \sum_{k=0}^{n} \frac{f^{(k+1)}(x_0)}{(k+1)!}(x-x_0)^{k+1} + o((x-x_0)^{n+1})\ .$$

Ceci établit la formule de Taylor–Young à l'ordre $n+1$, ce qui nous permet de conclure par récurrence qu'elle est vraie à tous les ordres. □

Avec des hypothèses plus fortes sur f, on peut avoir une estimation plus précise du reste.

Théorème 8.10 (Formule de Taylor–Lagrange)**.** *Soit $a < b$ deux réels, $n \geqslant 0$, et f une fonction de classe $\mathcal{C}^n$ sur $[a,b]$ et telle que $f^{(n)}$ soit dérivable sur $]a,b[$. Alors pour tout $x, x_0 \in [a,b]$ il existe $c \in]x_0, x[$ tel que l'on ait*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \frac{(x-x_0)^{n+1}}{(n+1)!}f^{(n+1)}(c)\ .$$

Démonstration. Notons que pour $n = 0$ on obtient exactement l'égalité des accroissements finis, et on sait donc déjà que la formule de Taylor–Lagrange est vraie à l'ordre 0. Pour montrer qu'elle est vraie à un ordre

$n > 0$, on va utiliser le théorème de Rolle ; soit donc f, $n > 0$, x_0 et x comme dans l'énoncé et considérons la fonction auxiliaire φ_λ définie sur le segment $[x_0, x]$ par

$$\varphi_\lambda(t) = f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(t)}{k!}(x-t)^k - \lambda \frac{(x-t)^{n+1}}{(n+1)!}.$$

On choisit $\lambda = (n+1)! \frac{f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k}{(x-x_0)^{n+1}}$ de telle façon que $\varphi_\lambda(x_0) = 0$. On a aussi $\varphi_\lambda(x) = 0$. De plus les hypothèses sur f assurent que φ_λ est dérivable sur $]x_0, x[$ et continue sur $[x_0, x]$. On peut donc lui appliquer le théorème de Rolle pour conclure qu'il existe $c \in]x_0, x[$ tel que $\varphi'_\lambda(c) = 0$. Cette égalité est équivalente à

$$-\sum_{k=0}^{n} \frac{f^{(k+1)}(c)}{k!}(x-c)^k + \sum_{\ell=1}^{n} \frac{f^{(\ell)}(c)}{(\ell-1)!}(x-c)^{\ell-1} + \lambda \frac{(x-c)^n}{n!} = 0 \ .$$

En décalant les indices $\ell - 1 = k$ dans la deuxième somme, ceci est équivalent à

$$-\sum_{k=0}^{n} \frac{f^{(k+1)}(c)}{k!}(x-c)^k + \sum_{k=0}^{n-1} \frac{f^{(k+1)}(c)}{k!}(x-c)^k + \lambda \frac{(x-c)^n}{n!} = 0 \ .$$

Après simplification, on est donc arrivé à $\frac{f^{(n+1)}(c)}{n!}(x-c)^n = \lambda \frac{(x-c)^n}{n!}$, autrement dit $\lambda = f^{(n+1)}(c)$ (notons qu'ici on utilise le fait que $c < x$). En revenant à la définition de λ et au fait que $\varphi_\lambda(x_0) = 0$ on a obtenu

$$f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k - f^{(n+1)}(c)\frac{(x-x_0)^{n+1}}{(n+1)!} = 0 \ ,$$

et on a donc bien démontré la formule de Taylor–Lagrange. □

La formule de Taylor–Lagrange est une généralisation de l'égalité des accroissements finis et, comme elle, elle peut être utilisée pour établir des inégalités, ce que vous ferez en TD.

Et si on renforce encore les hypothèses sur f, on obtient même une formule explicite pour le reste.

Théorème 8.11 (Formule de Taylor avec reste intégral)**.** *Soit I un intervalle de $\mathbb{R}$, et f une fonction de classe $\mathcal{C}^{n+1}$ sur I. Alors pour tout $x, x_0 \in I$ on a*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \int_{x_0}^{x} \frac{(x-t)^n}{n!} f^{(n+1)}(t)dt \ .$$

Démonstration. On va raisonner par récurrence. Pour $n = 0$, on doit démontrer l'égalité

$$f(x) = f(x_0) + \int_{x_0}^{x} f'(t)dt$$

qui est vraie, d'après le théorème fondamental de l'analyse, dès que f' est continue. Supposons maintenant le résultat démontré jusqu'au rang n, et supposons f de classe $\mathcal{C}^{n+2}$. On peut en particulier lui appliquer la formule de Taylor avec reste intégral au rang n pour écrire que

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \int_{x_0}^{x} \frac{(x-t)^n}{n!} f^{(n+1)}(t)dt \ .$$

On applique la formule d'intégration par parties avec $u'(t) = \frac{(x-t)^n}{n!}$ et $v = f^{(n+1)}$: $\int u'v = [uv] - \int uv'$. Celà est possible puisque $f^{(n+1)}$ et $t \mapsto (x-t)^{n+1}$ sont de classe $\mathcal{C}^1$. On a alors

$$\begin{aligned}\int_{x_0}^{x} \frac{(x-t)^n}{n!} f^{(n+1)}(t)dt &= \left[-f^{(n+1)}(t)\frac{(x-t)^{n+1}}{(n+1)!}\right]_{x_0}^{x} + \int_{x_0}^{x} f^{(n+2)}(t)\frac{(x-t)^{n+1}}{(n+1)!} \\ &= f^{(n+1)}(x_0)\frac{(x-x_0)^{n+1}}{(n+1)!} + \int_{x_0}^{x} f^{(n+2)}(t)\frac{(x-t)^{n+1}}{(n+1)!}.\end{aligned}$$

En injectant cette égalité dans la formule pour $f(x)$ obtenue ci-dessus, on obtient exactement la formule de Taylor avec reste intégral à l'ordre $n+1$.

□

Exercice 8.12. Dans le cas où f est de classe $\mathcal{C}^{n+1}$, donner une démonstration de la formule de Taylor–Lagrange à l'aide de la formule de Taylor avec reste intégral et de la première formule de la moyenne.

Chapitre 9

Développements limités et applications

Les développements limités sont particulièrement utiles pour calculer des limites, ou déterminer des équivalents. Leur avantage par rapport aux équivalents est qu'on peut ajouter et composer des développements limités (ce qui nécessite toutefois un peu d'attention...)

9.1 Développements limités

Définition 9.1. Soit $n \in \mathbb{N}$, I un intervalle de $\mathbb{R}$, $x_0 \in I$ et $f\colon I \to \mathbb{R}$. On dit que f admet un *développement limité* à l'ordre n en x_0 s'il existe $P \in \mathbb{R}_n[X]$ tel que

$$f(x) \underset{x_0}{=} P(x - x_0) + o\big((x - x_0)^n\big).$$

On dit alors que P est la *partie régulière* du développement limité, tandis que $f - P$ est appelé le *reste*.

La formule de Taylor–Young nous permet d'assurer que, si f est n fois dérivable en x_0, alors f admet un développement limité à l'ordre n en x_0, et que la partie régulière de ce développement limité est

$$f(x_0) + f'(x_0)X + \frac{f''(x_0)}{2}X^2 + \frac{f^{(3)}(x_0)}{6}X^3 + \cdots + \frac{f^{(n)}(x_0)}{n!}X^n \ .$$

En particulier, quand f est de classe C^∞ sur $\mathbb{R}$ et qu'on sait calculer la suite des dérivées successives $f^{(n)}(0)$, alors on peut donner les développements limités de f à tout ordre en 0. On reviendra plus tard sur les calculs pratiques de développements limités et leurs applications, pour l'instant on va développer un peu leurs propriétés théoriques (qui nous permettront de simplifier certains calculs).

Proposition 9.2. *La partie régulière d'un développement limité est unique.*

Lemme 9.3. *Si un polynôme $R \in \mathbb{R}_n[X]$, en tant que fonction réelle est négligeable en 0 devant un monôme X^k, $k \geqslant n$, alors il est nul.*

Démonstration. Pour le lemme : Le premier terme non nul de R ne serait pas négligeable devant X^n, ce qui impose que tous ses termes sont nuls.

Pour la proposition : Si $P, Q \in \mathbb{R}_n[X]$ sont tels que $P(x - x_0) + o((x - x_0)^n) \underset{x_0}{=} Q(x - x_0) + o((x - x_0)^n)$ alors leur différence est un polynôme de degré au plus n, négligeable : $P(x - x_0) - Q(x - x_0) \underset{x_0}{=} o((x - x_0)^n)$, donc $P = Q$. □

En particulier, on voit que le développement limité à l'ordre k d'une fontion f qui admet un développement limité à l'ordre $n \geqslant k$ en x_0 est obtenu, à partir de sa partie régulière d'ordre n, en négligeant ses termes de degré $\geqslant k + 1$.

Proposition 9.4. *Soit I un intervalle de $\mathbb{R}$, n un entier, $x_0 \in I$ et $f, g\colon I \to \mathbb{R}$ des fonctions admettant des développements limités à l'ordre n en x_0 de parties régulières P, Q respectivement. Alors :*

1. *La combinaison linéaire $\lambda f + \mu g$, pour tout $\lambda, \mu \in \mathbb{R}$, admet un développement limité à l'ordre n en x_0, de partie régulière $\lambda P + \mu Q$.*

2. *Le produit $f \times g$ admet un développement limité à l'ordre n en x_0, de partie régulière obtenue en ne retenant que les termes de degré $\leqslant n$ du polynôme $P \times Q$.*

Il est important de noter que, dans l'énoncé ci-dessus, les développements limités de f et de g sont calculés **au même ordre** ! Pour s'éviter des calculs inutiles, il faut veiller à *tronquer* les parties régulières à partir du même degré.

Démonstration. Pour alléger la notation, on ne va traiter que le cas $x_0 = 0$, qui est le cas le plus important en pratique et auquel on se ramène par une simple translation. Pour montrer la première propriété, on écrit simplement

$$\begin{aligned}\lambda f(x) + \mu g(x) &= \lambda\big(P(x) + o(x^n)\big) + \mu\big(Q(x) + o(x^n)\big)\\ &= \big(\lambda P(x) + \mu Q(x)\big) + \big(\lambda o(x^n) + \mu o(x^n)\big)\\ &= \big(\lambda P(x) + \mu Q(x)\big) + o(x^n)\ .\end{aligned}$$

La deuxième propriété se montre de façon similaire, mais il faut bien comprendre que, pour tout $m > n$, on a $x^m \underset{0}{=} o(x^n)$; en particulier, si on note R le polynôme obtenu en ne conservant que les termes de degré $\leqslant n$ de $P \times Q$, alors on a $R(x) = P(x) \times Q(x) + o(x^n)$. On a aussi $o(x^n) \times o(x^n) \underset{0}{=} o(x^{2n}) = o(x^n)$, il faut bien faire attention : on n'a pas écrit que les fonctions négligeables devant x^n et x^{2n} étaient les mêmes ; on a écrit que toute fonction négligeable en 0 devant x^{2n} était en particulier négligeable en 0 devant x^n. Une fois qu'on a compris tout cela, on peut écrire

$$\begin{aligned}f(x) \times g(x) &= \big(P(x) + o(x^n)) \times \big(Q(x) + o(x^n)\big)\\ &= P(x) \times Q(x) + P(x) \times o(x^n) + Q(x) \times o(x^n) + o(x^n) \times o(x^n)\\ &= \big(R(x) + o(x^n)\big) + o(x^n) + o(x^n) + o(x^n)\\ &= R(x) + o(x^n)\ .\end{aligned}$$

□

Notons que, dans les preuves ci-dessus, on a été très lourd dans la rédaction et la manipulation des $o(x^n)$; une fois qu'on y sera un peu mieux habitué, on n'écrira plus $o(x^n)$ qu'une fois par ligne de calcul...

Exemple. $\sin(x) = x + o(x^2)$, $\cos(x) = 1 - \frac{x^2}{2} + o(x^2)$, donc $\sin(x) \times \cos(x) = x - \frac{x^3}{2} + o(x^2) = x + o(x^2)$ et c'est tout ! Pourtant, on a bien que $(\cos(x) - 1 + \frac{x^2}{2}) \times (\sin(x) - x) = -\frac{x^7}{144} + o(x^8)$ mais ça ne signifie pas que le développement limité de leur produit est le produit de leurs développements limités à un ordre extravagant. Car on a bien-sûr $\sin(x) = x - \frac{x^3}{6} + \frac{x^5}{5!} + o(x^6)$, $\cos(x) = 1 - \frac{x^2}{2} + \frac{x^4}{24} - \frac{x^6}{6!} + o(x^6)$ donc $\sin(x) \times \cos(x) = x - \frac{2x^3}{3} + \frac{2x^5}{15} + o(x^6)$, qu'on attrape aussi par $\sin(x) \times \cos(x) = \frac{1}{2}\sin(2x)$.

On peut aussi composer des développements limités, mais bien sûr il faut faire attention aux points où on compose !

Proposition 9.5. *Soit I, J deux intervalles de $\mathbb{R}$, $n \in \mathbb{N}$, $x_0 \in I$, $g\colon I \to J$ et $f\colon J \to \mathbb{R}$ deux fonctions telles que g ait un développement limité à l'ordre n en x_0 (de partie régulière Q) et f ait un développement limité à l'ordre n en $g(x_0)$ (de partie régulière P). Alors $f \circ g$ a un développement limité à l'ordre n en x_0, dont la partie régulière est le polynôme obtenu en tronquant $P \circ (Q - Q(0))$ à partir du degré $n + 1$.*

Encore une fois, on ne compose les développements limités que s'ils ont été calculés au même ordre pour f et pour g.

Démonstration. Si $n = 0$ il s'agit de montrer que $f \circ g$ est continue en x_0, ce qu'on sait déjà. Sinon, puisque g admet un développement limité à l'ordre $n \geqslant 1$ en x_0 on a en particulier $g(x) - g(x_0) \underset{x_0}{=} a(x - x_0) + o(x - x_0)$, ce dont on tire que $o((g(x) - g(x_0))^n) \underset{x_0}{=} o((x - x_0)^n)$. De plus, on voit en développant

$$\big(Q(x-x_0)-Q(0)+o((x-x_0)^n)\big)^k \times \big(Q(x-x_0)-Q(0)+o((x-x_0)^n)\big)^\ell = \big(Q(x-x_0)-Q(0)\big)^{k+\ell}+o((x-x_0)^n)$$ que

$$P\big(Q(x - x_0) - Q(0) + o((x - x_0)^n)\big) \underset{x_0}{=} P\big(Q(x - x_0) - Q(0)) + o((x - x_0)^n\big)\ .$$

Enfin, on a par définition d'un développement limité que $Q(0) = g(x_0)$. Une fois ces informations rassemblées, il nous suffit d'écrire

$$\begin{aligned} f(g(x)) &\underset{x_0}{=} P\big(g(x) - g(x_0)\big) + o\big((g(x) - g(x_0))^n\big) \\ &\underset{x_0}{=} P\big(Q(x - x_0) - Q(0) + o((x - x_0)^n)\big) + o((x - x_0)^n) \\ &\underset{x_0}{=} P\big(Q(x - x_0) - Q(0)\big) + o((x - x_0)^n) \end{aligned}$$

En notant R le polynôme obtenu en ne retenant que les termes de degré $\leqslant n$ de $P \circ \big(Q - Q(0)\big)$, on vient d'arriver à

$$f(g(x)) \underset{x_0}{=} R(x - x_0) + o((x - x_0)^n)$$

□

En pratique, on composera surtout des développements limités avec $Q(0) = 0$ et il s'agira simplement de calculer une composée.

Les deux résultats précédents nous disent que les développements limités sont plus faciles à manipuler que les équivalents : comme les équivalents, on peut les multiplier, mais contrairement aux équivalents on peut aussi les ajouter et les composer. La proposition suivante nous dit qu'on peut aussi les intégrer !

Proposition 9.6. *Soit I un intervalle de $\mathbb{R}$, $n \in \mathbb{N}$, $x_0 \in I$ et f une fonction dérivable sur I telle que f' soit continue et ait un développement limité à l'ordre n en x_0, de partie régulière P. Alors f a un développement limité à l'ordre $n+1$ en x_0, dont la partie régulière est la primitive Q de P telle que $Q(0) = f(x_0)$.*

Démonstration. Par hypothèse, on peut écrire

$$f'(x) \underset{x_0}{=} P(x - x_0) + o((x - x_0)^n) \ .$$

Exactement comme dans la preuve de la formule de Taylor–Young, on utilise le fait que $\displaystyle\int_{x_0}^{x} o((t - x_0)^n)dt \underset{x_0}{=} o\big((x - x_0)^{n+1}\big)$, et on obtient alors en intégrant (et en utilisant le fait que f' est continue pour pouvoir appliquer le théorème fondamental de l'analyse) que

$$\begin{aligned} f(x) - f(x_0) &\underset{x_0}{=} \int_{x_0}^{x} P(t - x_0)dt + o\big((x - x_0)^{n+1}\big) \\ &\underset{x_0}{=} Q(x - x_0) - Q(0) + o\big((x - x_0)^{n+1}\big) \ . \end{aligned}$$

Puisque $Q(0) = f(x_0)$, on obtient bien comme attendu

$$f(x) \underset{x_0}{=} Q(x - x_0) + o((x - x_0)^{n+1} \ .$$

□

Par contre, on ne peut pas en général dériver un développement limité : il peut arriver que f ait un développement limité à l'ordre n en 0, mais que f' n'ait pas de développement limité à l'ordre $n-1$ en 0...

9.2 Les développements limités classiques

Ci-dessous on ne va écrire que des développements limités en 0 (et on écrira $= o(x^n)$ au lieu de $\underset{x_0}{=} o(x^n)$) ; en général si on calcule un développement limité de f en $x_0 \neq 0$ il est souvent plus pratique de poser $x = x_0 + h$ et de se ramener à un développement limité de $h \mapsto f(x_0 + h)$ en 0.

Puisque on a $\exp'(x) = \exp(x)$, on a immédiatement $\exp^{(n)}(0) = 1$, et la formule de Taylor-Young nous donne

$$\forall n \in \mathbb{N}, \quad e^x = \sum_{k=0}^{n} \frac{x^n}{n!} + o(x^n) \ .$$

Pour les fonctions trigonométriques, c'est à peine plus difficile : on a $\sin'(x) = \cos(x)$, et $\cos'(x) = -\sin(x)$, ce dont on déduit les formules

$$\forall n \in \mathbb{N}, \sin^{(2n)}(0) = 0 \text{ et } \sin^{(2n+1)}(0) = (-1)^n \ .$$

$$\forall n \in \mathbb{N}, \cos^{(2n)}(0) = (-1)^n \text{ et } \cos^{(2n+1)}(0) = 0 \ .$$

La formule de Taylor–Young nous donne alors

$$\forall n \in \mathbb{N}, \ \sin(x) = \sum_{k=0}^{n} \frac{(-1)^k}{(2k+1)!} x^{2k+1} + o(x^{2n+1}) \ ;$$

$$\forall n \in \mathbb{N}, \ \cos(x) = \sum_{k=0}^{n} \frac{(-1)^k}{(2k)!} x^{2k} + o(x^{2n}) \ ;$$

On voit ci-dessus que, dans le développement limité de cos, n'apparaissent que des termes pairs, et dans celui de sin que des termes impairs. La raison en est explicitée dans la proposition suivante.

Proposition 9.7. *Soit I un intervalle de $\mathbb{R}$ centré en 0, et $f : I \to \mathbb{R}$ une fonction admettant un développement limité à l'ordre n en 0, de partie régulière P. Si f est paire, alors tous les coefficients de degré impair de P sont nuls ; et si f est impaire alors tous les coefficients de degré pair de P sont nuls.*

Démonstration. C'est une conséquence de l'unicité de la partie régulière d'un développement limité. De l'égalité $f(x) = P(x) + o(x^n)$ on déduit que $f(-x) = P(-x) + o((-x)^n) = P(-x) + o(x^n)$. Si $f(x) = f(-x)$, alors $P(X)$ et $P(-X)$ sont donc tous les deux des parties régulières du développement limité de f à l'ordre n en 0, d'où $P(X) = P(-X)$, et P n'a que de coefficients de degré pair. De même si $f(x) = -f(-x)$ on déduit que $P(X) = -P(-X)$ et P n'a que des coefficients d'ordre impair. □

Continuons notre liste de développements limités ; par la formule pour la somme des termes d'une suite géométrique, on sait que

$$\forall n \in \mathbb{N}, \ \sum_{k=0}^{n} x^k = \frac{1 - x^{n+1}}{1 - x} = \frac{1}{1 - x} + o(x^n) \ .$$

Autrement dit, on a

$$\forall n \in \mathbb{N}, \ \frac{1}{1 - x} = \sum_{k=0}^{n} x^k + o(x^n) \ .$$

On en déduit que

$$\forall n \in \mathbb{N}, \ \frac{1}{1 + x} = \sum_{k=0}^{n} (-1)^k x^k + o(x^n) \ .$$

En intégrant, on obtient

$$-\ln(1 - x) = \ln(1) + \sum_{k=0}^{n} \frac{x^{k+1}}{k+1} + o(x^{n+1}) \ , \text{ et donc}$$

$$\forall n \in \mathbb{N}, \ \ln(1 - x) = -\sum_{k=1}^{n} \frac{x^k}{k} + o(x^n) \ .$$

De même,

$$\forall n \in \mathbb{N}, \ \ln(1 + x) = \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k} x^k + o(x^n).$$

Ceci nous permet également de calculer le développement limité en 0 de arctan : à partir de $\frac{1}{1+x^2} = \sum_{k=0}^{n} (-1)^k x^{2k} + o(x^{2n})$, on obtient en intégrant que

$$\arctan(x) = \sum_{k=0}^{n} \frac{(-1)^k x^{2k+1}}{2k+1} + o(x^{2n+1}) \ .$$

Voyons une dernière formule à connaître : pour $\alpha \in \mathbb{R}$, on vérifie par récurrence que, pour $n \geqslant 1$, la dérivée n-ième de $f_\alpha \colon x \mapsto (1+x)^\alpha$ est égale à $\alpha(\alpha-1)\dots(\alpha-n+1)(1+x)^{\alpha-n}$. En particulier, pour $n \geqslant 1$ on a $f_\alpha^{(n)}(0) = \alpha(\alpha-1)\dots(\alpha-n+1)$, et la formule de Taylor–Young nous donne

$$\forall n \in \mathbb{N},\ (1+x)^\alpha = 1 + \alpha x + \dots + \alpha(\alpha-1)\dots(\alpha-n+1)\frac{x^n}{n!} + o(x^n)\ .$$

Récapitulons tous les développements que nous venons d'obtenir :

- $\dfrac{1}{1+x} = 1 - x + x^2 - \dots + (-1)^n x^n + o(x^n)$.
- $\ln(1+x) = x - \dfrac{x^2}{2} + \dots + (-1)^{n+1}\dfrac{x^n}{n} + o(x^n)$.
- $e^x = 1 + x + \dfrac{x^2}{2} + \dfrac{x^3}{6} + \dots + \dfrac{x^n}{n!} + o(x^n)$.
- $\cos(x) = 1 - \dfrac{x^2}{2} + \dots + (-1)^n \dfrac{x^{2n}}{(2n)!} + o\left(x^{2n+1}\right)$.
- $\sin(x) = x - \dfrac{x^3}{6} + \dots + (-1)^n \dfrac{x^{2n+1}}{(2n+1)!} + o\left(x^{2n+2}\right)$.
- $\arctan(x) = x - \dfrac{x^3}{3} + \dots + (-1)^n \dfrac{x^{2n+1}}{2n+1} + o\left(x^{2n+2}\right)$.
- $(1+x)^\alpha = 1 + \alpha x + \dots + \alpha(\alpha-1)\cdots(\alpha-n+1)\dfrac{x^n}{n!} + o(x^n)$ (valable pour tout $\alpha \in \mathbb{R}$).

Les résultats ci-dessus ont été obtenus à partir de la formule de Taylor–Young ; en général, c'est une mauvaise idée de calculer un développement limité à l'ordre n en calculant n dérivées successives de f, car c'est beaucoup trop lourd en calcul. On préférera si possible déduire les développements limités à partir des développements classiques ci-dessus (à connaître ou à savoir retrouver très vite) et les méthodes de calcul : produit, composition, intégration... Pour les développements limités des quotients, on peut soit raisonner par composition comme on va le voir ci-dessous, soit appliquer la méthode des *divisions par puissances croissantes* qu'on va aussi se contenter d'illustrer sur des exemples.

9.3 Premiers exemples de calculs de développements limités

Un exemple : trois méthodes pour calculer le développement limité de tan **en** 0 **à l'ordre** 5.

Première méthode : la division selon les puissances croissantes, qui ressemble à une division euclidienne, sauf qu'on fait en sorte d'éliminer successivement les termes de plus bas degré.

On commence par écrire

$$\tan(x) = \frac{\sin(x)}{\cos(x)} = \frac{x - \frac{x^3}{6} + \frac{x^5}{120} + o(x^5)}{1 - \frac{x^2}{2} + \frac{x^4}{24} + o(x^5)}\ .$$

Puis on utilise la méthode des puissances croissantes :

$$\begin{array}{rcccccc|l}
x & - & \frac{x^3}{6} & + & \frac{x^5}{120} & + & o(x^5) & 1 - \frac{x^2}{2} + \frac{x^4}{24} + o(x^5) \\
\cline{8-8}
- \ x & + & \frac{x^3}{2} & - & \frac{x^5}{24} & + & o(x^5) & x + \frac{x^3}{3} + \frac{2}{15}x^5 \\
\cline{1-7}
 & & \frac{x^3}{3} & - & \frac{x^5}{30} & + & o(x^5) & \\
 & - & \frac{x^3}{3} & + & \frac{x^5}{6} & + & o(x^5) & \\
\cline{2-7}
 & & & & \frac{2}{15}x^5 & + & o(x^5) &
\end{array}$$

On a donc obtenu la relation

$$x - \frac{x^3}{6} + \frac{x^5}{120} + o(x^5) = (1 - \frac{x^2}{2} + \frac{x^4}{24})(x + \frac{x^3}{3} + \frac{2}{15}x^5) + o(x^5)\ ,$$

qui nous donne finalement $\tan(x) = x + \frac{x^3}{3} + \frac{2}{15}x^5 + o(x^5)$.

Deuxième méthode : par composition, en utilisant le développement limité de $\frac{1}{1-u}$ (ci-dessous, les termes en gris clair sont ceux qu'on aurait pu se passer d'écrire, puisqu'ils font apparaître des termes de degré > 5).

$$\begin{aligned}
\tan(x) &= \frac{x - \frac{x^3}{6} + \frac{x^5}{120} + o(x^5)}{1 - \frac{x^2}{2} + \frac{x^4}{24} + o(x^5)} \\
&= \frac{x - \frac{x^3}{6} + \frac{x^5}{120}}{1 - (\frac{x^2}{2} - \frac{x^4}{24})} + o(x^5) \\
&= (x - \frac{x^3}{6} + \frac{x^5}{120})(1 + (\frac{x^2}{2} - \frac{x^4}{24}) + (\frac{x^2}{2} - \frac{x^4}{24})^2 + (\frac{x^2}{2} - \frac{x^4}{24})^3 + (\frac{x^2}{2} - \frac{x^4}{24})^4 + (\frac{x^2}{2} - \frac{x^4}{24})^5) + o(x^5) \\
&= (x - \frac{x^3}{6} + \frac{x^5}{120})(1 + \frac{x^2}{2} + \frac{5x^4}{24}) + o(x^5) \\
&= x + \frac{x^3}{3} + \frac{2x^5}{15} + o(x^5) \ .
\end{aligned}$$

(Note : dans la dernière ligne, on n'a, de nouveau, pas fait apparaître les termes de degré > 5, puisque ce sont tous des $o(x^5)$)

Troisième méthode : en utilisant une équation différentielle. On a $\tan'(x) = 1 + \tan^2(x)$. En écrivant le développement de $\tan'$ en 0 à l'ordre 4 sous la forme $a + bx^2 + cx^4 + o(x^4)$ (il n'y pas de termes d'ordre impair : tan est impaire, donc $\tan'$ est paire), le fait que $\tan(0) = 0$ et le théorème d'intégration des développements limités nous donne que le développement limité de tan en 0 à l'ordre 5 est égal à $ax + b\frac{x^3}{3} + c\frac{x^5}{5} + o(x^5)$. La formule $\tan'(x) = 1 + \tan^2(x)$ nous donne alors (par composition) :

$$a + bx^2 + cx^4 + o(x^4) = 1 + (ax + b\frac{x^3}{3} + c\frac{x^5}{5})^2 + o(x^4).$$

Autrement dit, on a

$$a + bx^2 + cx^4 + o(x^4) = 1 + a^2x^2 + \frac{2ab}{3}x^4 + \frac{b^2}{9}x^6 + o(x^4).$$

Le théorème d'unicité des développements limités nous permet d'identifier les deux développements terme à terme : ceci donne $a = 1$, $b = a^2 = 1$, et $c = \frac{2ab}{3} = \frac{2}{3}$. En reportant cela dans la formule donnant le développement de tan à l'ordre 5 en 0 en fonction de a, b, c, on obtient de nouveau $\tan(x) = x - \frac{x^3}{3} + \frac{2}{15}x^5 + o(x^5)$.

Voyons brièvement deux autres exemples. Si l'on veut calculer le développement limité de $x \mapsto \frac{1}{1-\sin(x)}$ à l'ordre 3 en 0, on peut utiliser la méthode de composition (en évitant de faire intervenir dans le calcul des termes de trop haut degré, comme on l'a fait pour la tangente) :

$$\begin{aligned}
\frac{1}{1 - \sin(x)} &= \frac{1}{1 - x + \frac{x^3}{6} + o(x^3)} \\
&= 1 + (x - \frac{x^3}{6}) + x^2 + x^3 + o(x^3) \\
&= 1 + x + x^2 + \frac{5}{6}x^3 + o(x^3) \ .
\end{aligned}$$

A la deuxième ligne ci-dessus, on n'a pas pris la peine d'écrire les « $\frac{x^3}{6}$ »quand on a composé après l'ordre 2 : dans le développement, ils auraient contribué des termes négligeables devant x^3. C'est un bon exercice d'essayer de retrouver ce résultat en utilisant la méthode de division par puissances croissantes...

De même, calculons le développement limité à l'ordre 2 en 0 de $\sqrt{\cos(x)}$:

$$\begin{aligned}
\sqrt{\cos(x)} &= \sqrt{1 - \frac{x^2}{2} + o(x^2)} \\
&= 1 + \frac{1}{2}(-\frac{x^2}{2}) + o(x^2) \\
&= 1 - \frac{x^2}{4} + o(x^2) \ .
\end{aligned}$$

Attention, on ne calcule pas des développements limités seulement en 0 ! Par exemple, essayons de développer $x \mapsto e^x$ à l'ordre 3 en 1. On écrit $x = 1 + h$, et on se ramène à développer $h \mapsto 1 + h$ à l'ordre 3 en 0, ce qui nous amène au calcul suivant :

$$\begin{aligned} e^{1+h} &= e.e^h \\ &= e(1 + h + \frac{h^2}{2} + \frac{h^3}{6} + o(h^3)) \ . \end{aligned}$$

En revenant à la relation $h = x - 1$, on a obtenu le développement limité $e^x = e(1 + (x-1) + \frac{(x-1)^2}{2} + \frac{(x-1)^3}{6}) + o((x-1)^3)$.

Exercice 9.8. Donner le développement limité à l'ordre 2 en $\frac{\pi}{4}$ de $x \mapsto \sin(x)$.

9.4 Exemples d'applications

Le fait de savoir calculer des développements limités nous rend accessible le calcul de limites qui étaient jusqu'à présent difficiles. Une difficulté est qu'on ne sait pas a priori à quel ordre calculer le développement limité pour trouver la limite : si on va à un ordre trop bas, les calculs ne permettent peut-être pas de répondre ; à un ordre trop élevé, on fait des calculs inutiles et on perd son temps. Il faut s'entraîner... Par exemple, essayons de déterminer la limite en 0 de $\dfrac{2}{\sin^2(x)} + \dfrac{1}{\ln(\cos(x))}$. En mettant au même dénominateur, on est confronté à $\dfrac{2\ln(\cos(x)) + \sin^2(x)}{\sin^2(x)\ln(\cos(x))}$. On sait que $\sin^2(x) \sim x^2$, et $\ln(\cos(x)) = \ln(1 - \frac{x^2}{2} + o(x^2)) = -\frac{x^2}{2} + o(x^2) \sim -\frac{x^2}{2}$; le dénominateur est donc équivalent en 0 à $-\frac{x^4}{2}$. Pour comprendre la limite (si elle existe), on doit donc développer le numérateur à l'ordre 4 en 0. On calcule :

$$\begin{aligned} 2\ln(\cos(x)) + \sin^2(x) &= 2\ln\left(1 - \frac{x^2}{2} + \frac{x^4}{24}\right) + (x - \frac{x^3}{6})^2 + + o(x^4) \\ &= 2\left((-\frac{x^2}{2} + \frac{x^4}{24}) - \frac{1}{2}(\frac{x^2}{2})^2\right) + x^2 - \frac{x^4}{3} + o(x^4) \\ &= -x^2 - \frac{x^4}{6} + x^2 - \frac{x^4}{3} + o(x^4) \\ &= \frac{-x^4}{2} + o(x^4) \\ &\sim -\frac{x^4}{2}. \end{aligned}$$

Il en ressort que en 0 on a

$$\frac{2\ln(\cos(x)) + \sin^2(x)}{\sin^2(x)\ln(\cos(x))} \sim \frac{-\frac{x^4}{2}}{-\frac{x^4}{2}} = 1 \ .$$

La limite recherchée est donc égale à 1.

On peut aussi utiliser les développements limités pour mener à bien des études fines de graphes de fonction : détermination de la position du graphe par rapport à sa tangente en un point, recherche d'asymptotes et positions par rapport à celles-ci...

Si f est dérivable en x_0, déterminer la position du graphe de f par rapport à sa tangente revient à étudier le signe, au voisinage de x_0, de $f(x) - (f(x_0) + (x - x_0)f'(x_0))$: on voit que, si on peut trouver un développement limité de f d'ordre $n \geqslant 2$ avec un terme d'ordre $\geqslant 2$ non nul, cela va nous permettre de déterminer ce signe (qui sera constant si le premier terme non nul est d'ordre pair, et changera si le premier terme non nul est d'ordre impair). Voyons deux exemples. Comme premier exemple, étudions l'application $f : x \mapsto \ln(1 + x + x^2)$ et la position de son graphe par rapport à sa tangente en 0 et en 1. En 0, un développement limité à l'ordre 2 nous donne

$$\begin{aligned} \ln(1 + x + x^2) &\underset{0}{=} (x + x^2) - \frac{(x + x^2)^2}{2} + o(x^2) \\ &\underset{0}{=} x + \frac{x^2}{2} + o(x^2). \end{aligned}$$

Ceci nous donne deux informations : d'une part, la tangente au graphe de f en 0 est d'équation $y = x$; d'autre part, puisque $f(x) - x = \frac{x^2}{2} + o(x^2) \sim \frac{x^2}{2}$, on voit que $f(x) - x \geqslant 0$ pour x proche de 0, et donc le graphe de f est (au voisinage de 0) au-dessus de sa tangente en 0.

En 1, on fait un calcul similaire, en posant $x = 1 + h$ pour se ramener à un développement limité en 0 :

$$\begin{aligned}\ln(1 + x + x^2) &= \ln(1 + (1 + h) + (1 + h)^2) \\ &= \ln(3 + 3h + h^2) \\ &= \ln(3) + \ln(1 + h + \frac{h^2}{3}) \\ &\underset{0}{=} \ln(3) + h + \frac{h^2}{3} - \frac{1}{2}(h + \frac{h^2}{3})^2 + o(h^2) \\ &\underset{0}{=} \ln(3) + h - \frac{1}{6}h^2 + o(h^2)\end{aligned}$$

Ceci nous permet de conclure que, en 1, on a

$$f(x) \underset{1}{=} \ln(3) + (x - 1) - \frac{1}{6}(x - 1)^2 + o((x - 1)^2) \ .$$

On en déduit que la tangente au graphe de f en 1 est d'équation $y = \ln(3) + (x - 1)$; et puisque $f(x) - (\ln(3) + (x - 1)) = -\frac{1}{6}(x - 1)^2 + o((x - 1)^2) \sim -\frac{1}{6}(x - 1)^2$, on voit que $f(x) - (\ln(3) + (x - 1)) \leqslant 0$ pour x proche de 1, et donc le graphe de f est (au voisinage de 1) en-dessous de sa tangente en 1.

Pour étudier des limites, des équivalents ou des asymptotes en $\pm\infty$, on peut aussi souvent utiliser des développements limités. Par exemple, étudions le comportement de $\left(x + \sqrt{x + 1}\right)^{\frac{1}{3}} - \left(x + \sqrt{x - 1}\right)^{\frac{1}{3}}$ en $+\infty$. On commence par faire sortir le terme dominant :

$$\left(x + \sqrt{x + 1}\right)^{\frac{1}{3}} = x^{\frac{1}{3}}(1 + \sqrt{\frac{1}{x} + \frac{1}{x^2}})$$

On a donc un premier équivalent, $x^{\frac{1}{3}}$, qui ne va pas nous donner assez d'information puisque l'autre terme de la différence est aussi équivalent à $x^{\frac{1}{3}}$ et qu'on ne peut pas soustraire des équivalents ; on écrit

$$\begin{aligned}\sqrt{\frac{1}{x} + \frac{1}{x^2}} &= \frac{1}{\sqrt{x}}\sqrt{1 + \frac{1}{x}} \\ &= \frac{1}{\sqrt{x}}\left(1 + \frac{2}{x} + o(\frac{1}{x})\right) \\ &= \frac{1}{\sqrt{x}} + \frac{1}{2x^{\frac{3}{2}}} + o(\frac{1}{x^{\frac{3}{2}}})\end{aligned}$$

On a ainsi obtenu que, en $+\infty$,

$$\begin{aligned}\left(x + \sqrt{x + 1}\right)^{\frac{1}{3}} &= x^{\frac{1}{3}}\left(1 + \frac{1}{\sqrt{x}} + \frac{1}{2x^{\frac{3}{2}}} + o(\frac{1}{x^{\frac{3}{2}}})\right)^{\frac{1}{3}} \\ &= x^{\frac{1}{3}}\left(1 + \frac{1}{3\sqrt{x}} - \frac{1}{9x} + \frac{1}{x\sqrt{x}}\frac{37}{162} + o(\frac{1}{x\sqrt{x}})\right)\end{aligned}$$

(Après calcul, en composant avec le développement limité de $(1 + u)^{\frac{1}{3}}$ à l'ordre 3 en 0).

Ce qu'on a obtenu ci-dessus s'appelle un *développement asymptotique* de $x \mapsto \left(x + \sqrt{x + 1}\right)^{\frac{1}{3}}$ en $+\infty$. Un calcul similaire amène à

$$\left(x + \sqrt{x - 1}\right)^{\frac{1}{3}} = x^{\frac{1}{3}}\left(1 + \frac{1}{3\sqrt{x}} - \frac{1}{9x} - \frac{1}{x\sqrt{x}}\frac{17}{162} + o(\frac{1}{x\sqrt{x}})\right)$$

Finalement,

$$\left(x + \sqrt{x + 1}\right)^{\frac{1}{3}} - \left(x + \sqrt{x - 1}\right)^{\frac{1}{3}} = x^{\frac{1}{3}}\left(\frac{1}{3x\sqrt{x}} + o(\frac{1}{x\sqrt{x}})\right) \sim \frac{1}{3}x^{-\frac{1}{6}} \ .$$

Ainsi, notre fonction tend vers 0 en $+\infty$, et on en a déterminé un équivalent précis.

Étudions maintenant le comportement en $+\infty$ de $f\colon x \mapsto x^2 \arctan(\frac{1}{1+x^2})$: a-t-elle une limite en $+\infty$ (ce qui revient à dire que le graphe a une asymptote horizontale en $+\infty$) ? Quelle est la position du graphe de f par rapport à une asymptote éventuelle ? Pour faire apparaître un développement limité de arctan en 0, on écrit

$$x^2 \arctan(\frac{1}{1+x^2}) = x^2 \arctan\left(\frac{1}{x^2}\left(\frac{1}{1+\frac{1}{x^2}}\right)\right).$$

En posant $u = \frac{1}{x}$, on a besoin d'un développement limité de $u \mapsto \arctan(\frac{u^2}{1+u^2})$ en 0 ; pour déterminer la limite de f en $+\infty$, un développement limité à l'ordre 2 (ou un simple équivalent, facile à obtenir ici) nous suffit, mais pour avoir une information plus précise, on va chercher le terme suivant non nul du développement limité, qui apparaîtra à l'ordre 4 (la fonction étant paire, le terme d'ordre 3 sera nul). On calcule :

$$\begin{aligned}
\arctan(\frac{u^2}{1+u^2}) &= \arctan(u^2(1-u^2+o(u^2))) \\
&= \arctan(u^2-u^4+o(u^4)) \\
&= (u^2-u^4) - \frac{(u^2-u^4)^3}{3} + o(u^4) \\
&= u^2 - u^4 + o(u^4)
\end{aligned}$$

En revenant à la définition de f, on a obtenu (en se rappelant que $u = 1/x$) :

$$\begin{aligned}
x^2 \arctan(\frac{1}{1+x^2}) &= x^2\left(\frac{1}{x^2} - \frac{1}{x^4} + o(\frac{1}{x^4})\right) \\
&= 1 - \frac{1}{x^2} + o(\frac{1}{x^2})
\end{aligned}$$

Ce développement asymptotique nous permet de conclure que $f(x)$ tend vers 1 quand x tend vers $+\infty$, et que le graphe de f est en-dessous de son asymptote horizontale d'équation $y = 1$; on aurait pu obtenir les mêmes informations en utilisant simplement le fait que $\arctan(u) \sim u$ en 0, et $\arctan(u) \leqslant u$ pour tout u positif (qui découle de l'inégalité des accroissements finis ; savez-vous le démontrer ?) mais ici, avec un développement asymptotique, on a une information plus précise (puisqu'on connaît précisément l'ordre de grandeur de $f(x)-1$ quand x tend vers $+\infty$, et pas seulement son signe).

Un dernier exemple : étudions la fonction $g\colon x \mapsto x^2 e^{\frac{2x}{x^2-1}}$, que nous allons étudier en $+\infty$. On commence par remarquer que $\dfrac{2x}{x^2-1}$ tend vers 0 quand x tend vers $+\infty$, et on va donc pouvoir étudier un développement limité de exp en 0 ; on commence comme toujours par le terme le plus à l'intérieur des parenthèses et on écrit

$$\begin{aligned}
\frac{2x}{x^2-1} &= \frac{2}{x}\left(\frac{1}{1-\frac{1}{x^2}}\right) \\
&= \frac{2}{x}(1+\frac{1}{x^2}+o(\frac{1}{x^2})) \\
&= \frac{2}{x} + \frac{2}{x^3} + o(\frac{1}{x^3})
\end{aligned}$$

En composant avec le développement limité de exp en 0, on obtient

$$\begin{aligned}
x^2 e^{\frac{2x}{x^2-1}} &= x^2(\frac{2}{x} + \frac{2}{x^3} + o(\frac{1}{x^3})) \\
&= 2x + \frac{2}{x} + o(\frac{1}{x})
\end{aligned}$$

On en déduit que le graphe de g admet la droite d'équation $y = 2x$ comme asymptote en $+\infty$; et puisque $g(x) - 2x \sim \frac{2}{x} \geqslant 0$, le graphe de g est au-dessus de cette asymptote en $+\infty$.

Exercice 9.9. Étudier la fonction g définie ci-dessus en ± 1 et en $-\infty$. Donner l'allure du graphe de g.

Chapitre 10

Équations différentielles

Une équation différentielle est une équation fonctionnelle faisant intervenir les valeurs de la fonction et de ses dérivées. Son inconnue est une fonction de régularité suffisante et répondant à des conditions particulières. Bien que les équations différentielles provenant de la modélisation de phénomènes réels soient le plus souvent complexes, on s'intéressera surtout aux équations différentielles linéaires, provenant de la linéarisation de ces phénomènes par développements limités, faisant intervenir une combinaison linéaire des valeurs et dérivées de la fonction cherchée, au même point.

10.1 Équations linéaires réelles d'ordre 1

Définition 10.1. Une *équation différentielle* est une équation dont l'inconnue est une fonction y ; une équation différentielle linéaire réelle d'ordre 1, à coefficients non constants, se présente sous la forme

$$\forall x \in I, \quad y'(x) + a(x)y(x) = b(x) \ , \tag{E}$$

où I est un intervalle de $\mathbb{R}$ et a, b sont des fonctions continues sur I. On appelle l'équation sans second membre *l'équation différentielle homogène* associée :

$$\forall x \in I, \quad y'(x) + a(x)y(x) = 0. \tag{E_0}$$

Une solution de l'équation est une fonction dérivable $y\colon I \to \mathbb{R}$ telle que $y'(x) + a(x)y(x) = b(x)$ pour tout $x \in I$. Notons qu'alors $y' = b - ay$ sera une combinaison linéaire de fonctions continues, donc y' sera continue : toute solution sera nécessairement de classe $\mathcal{C}^1$.

Observation. Une première observation est que, si y_1, y_2 sont deux solutions de (E), alors $y_1 - y_2$ est une solution de l'*équation homogène* (E_0). En effet, si on a à la fois $y_1' + ay_1 = b$ et $y_2' + ay_2 = b$ alors on déduit $(y_1 - y_2)' + a(y_1 - y_2) = 0$.

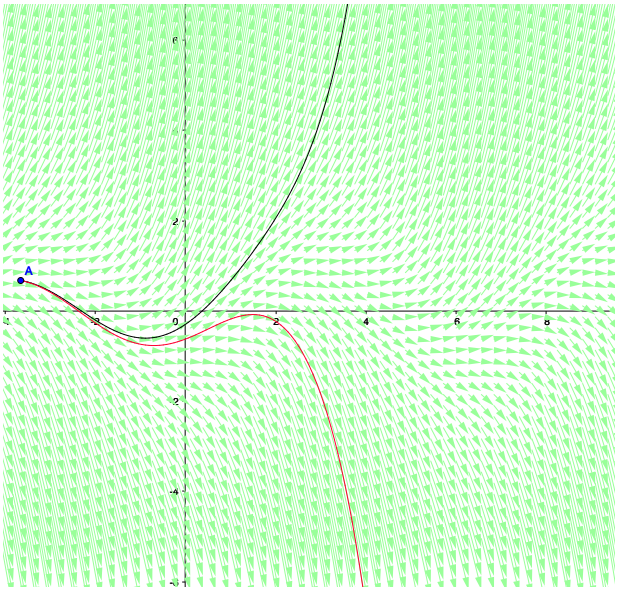
Par conséquent, si on sait déterminer :

— S_H l'ensemble des solutions de (E_0), et

— y_p une solution de (E), dite *solution particulière*,

alors l'ensemble S_E des solutions de E sont les fonctions qui s'écrivent comme somme de la solution particulière y_p de (E) et d'une solution quelconque de (E_0). On dit que S_E est l'espace affine dirigé par S_H passant par y_p. Ce schéma de raisonnement marche d'ailleurs dès qu'on travaille avec une équation différentielle linéaire (d'ordre quelconque) ; on le reverra plus loin dans le cas des équations différentielles linéaires d'ordre 2.

Interprétation géométrique Trouver une solution à une équation différentielle du premier ordre, c'est trouver une fonction dont la dérivée y' est prescrite par la valeur de la fonction y au point x. On peut traduire (E) sous la forme d'un champ de vecteurs, en tout point $(x, y) \in I \times \mathbb{R}$, prescrivant la tangente de la solution passant en ce point. En n'importe quelle condition initiale définie par un point $A(x_0, y_0)$, on imagine une plume suivre le vent dont la direction est donnée par V.



Deux conditions initiales très proches du point A peuvent conduire à des solutions qui se séparent rapidement comme dans l'exemple ci-dessus donné par $y' = y + \cos(x)$. On voit que les vecteurs « tirent » vers le haut d'un côté et vers le bas de l'autre, modulés par le cosinus de x. On conçoit qu'une plume peut hésiter à partir vers le haut ou vers le bas... Ce qu'on observe également c'est que jamais deux solutions à une équation différentielle d'ordre 1 ne se croiseront car en chaque point leur tangente est fixée. Une préoccupation importante dans le cadre des équations différentielles est le comportement asymptotique de la solution, en $\pm\infty$.

En prenant la métaphore de la plume dans le vent au pied de la lettre, c'est-à-dire en comprenant le vecteur $V(x, y) = \big(1, b(x) - y \times a(x)\big)$ comme un vecteur vitesse avec x le temps, on obtient un schéma numérique, appelé *la méthode d'*Euler, qui nous permet d'avoir une approximation du graphe d'une fonction solution, c'est-à-dire d'une courbe intégrale, passant par un point initial (x_0, y_0). On prend un pas de temps δ petit et on définit une suite de points $A_0 = (x_0, y_0)$ et $A_{n+1} = A_n + \delta V(A_n)$.

Dans ce cas particulier très simple d'équation différentielle, linéaire (pas de termes en y^k avec $k \neq 1$ ou de fonctions plus compliquées) du premier ordre (pas de y''), on sait trouver une solution de manière systématique à l'aide de deux intégrations, et on va maintenant expliquer comment.

Proposition 10.2 (Formule de Duhamel)**.** *L'équation (E) a une unique solution valant y_0 en un point $x_0 \in I$ et elle est de la forme $y(x) = K(x)e^{-A(x)}$ où A est une primitive de a et K une primitive de $x \mapsto b(x)e^{A(x)}$.*

Preuve: Comme l'équation est linéaire, si y_1 est une solution de (E) et y_2 est une solution de (E_0) alors $y_1 + y_2$ est une solution de (E).

Choisissons une primitive A de a sur I, qui existe puisque a est continue ; la fonction $x \mapsto e^{-A(x)}$ a comme dérivée $x \mapsto -a(x)e^{-A(x)}$, elle est donc solution de l'équation homogène (E_0). En pratique il faudra calculer A explicitement.

Étant donnée une fonction y dérivable sur I, on introduit $f : x \mapsto y(x) \times e^{A(x)}$. On a alors, pour tout $x \in I$, l'égalité

$$f'(x) = y'(x)e^{A(x)} + y(x)A'(x)e^{A(x)} = (y'(x) + ay(x))e^{A(x)}.$$

On en déduit que y est solution de l'équation homogène (E_0) si, et seulement si, $f'(x) = 0$ pour tout $x \in I$, c'est-à-dire si et seulement si f est constante : par conséquent, les solutions de (E_0) sont exactement les fonctions de la forme $x \mapsto Ke^{-A(x)}$, où K est une constante et A est une primitive de a sur I.

Ensuite, il nous reste à trouver une solution de (E) ; s'il n'y a pas de solution évidente, on peut appliquer la *méthode de variation de la constante*, en cherchant y sous la forme $x \mapsto K(x)e^{-A(x)}$ pour une fonction dérivable K. En dérivant, on obtient pour tout $x \in I$

$$\begin{aligned} y'(x) + a(x)y(x) &= K'(x)e^{-A(x)} - K(x)A'(x)e^{-A(x)} + a(x)K(x)e^{-A(x)}, \\ &= K'(x)e^{-A(x)}. \end{aligned}$$

Par conséquent, y est solution de (E) si, et seulement si, on a $K'(x)e^{-A(x)} = b(x)$ pour tout $x \in I$, autrement dit si, et seulement si, K est une primitive sur I de la fonction continue $x \mapsto e^{A(x)}b(x)$.

On vient de démontrer que les solutions de (E) sont exactement les fonctions de la forme $x \mapsto K(x)e^{-A(x)}$, où K est une primitive de $x \mapsto e^{A(x)}b(x)$.

On pourrait penser avoir deux degrés de liberté puisque la primitive A est déterminée à une constante additive près, ainsi que la primitive K. Cependant, ces deux constantes jouent exactement le même rôle, l'ensemble des fonctions solutions est, pour A et K des primitives données,

$$S_E = \{ \begin{cases} I & \to \mathbb{R} \\ x & \mapsto K(x)e^{-A(x)} + \lambda \end{cases} : \lambda \in \mathbb{R} \}.$$

En prenant $\lambda = y_0 - K(x_0)e^{-A(x_0)}$, on trouve ainsi l'unique solution de (E) respectant la *condition initiale* de valoir y_0 en x_0. On vient d'établir là un cas particulier du théorème de Cauchy–Lipschitz selon lequel les solutions d'une équation différentielle (satisfaisant certaines hypothèses !) sont uniquement déterminées par un choix de conditions initiales.

La forme intégrale de cette solution, due à Duhamel, avec $A(x) = \int_{x_0}^{x} a(t)dt$ la primitive s'annulant en x_0, est

$$y(x) = e^{-A(x)} \left(y_0 + \int_{x_0}^{x} e^{A(t)} b(t) dt \right).$$

□

En résumé, résoudre une équation différentielle linéaire réelle du premier ordre se réduit à calculer deux primitives ; c'est un cas très particulier, il est en général difficile de trouver les solutions d'équations différentielles, même dans le cas linéaire ! Pour les équations non homogènes, il peut être efficace de chercher des solutions évidentes. Étant donné la linéarité des équations, si le second membre est une somme de termes, ils peuvent être traités séparément et recombinés. On peut même faire apparaître des termes qui simplifient la recherche, en particulier pour les solutions polynomiales.

Exemples

1. Réécrivons $y' + x^2 y = x^5 = x^5 + 3x^2 - 3x^2$ car $y(x) = x^3$ est solution évidente de l'équation $y' + x^2 y = x^5 + 3x^2$ tandis que la constante -3 est solution de $y' + x^2 y = -3x^2$, ce qui donne $y = x^3 - 3$.

   De manière générale, la forme du second membre permet de chercher des solutions particulières dans la même classe, par exemple ceux de la forme $b(x) = P(x)e^{\lambda x}$ avec $P \in \mathbb{R}[X]$, ou bien les fonctions circulaires ou exponentielles. Il ne faut pas hésiter à *bricoler* un peu !

2. Cherchons à résoudre sur $\mathbb{R}$ l'équation $y' + 2xy = x$. L'équation homogène associée admet pour solutions les fonctions de la forme $x \mapsto Ke^{-\frac{x^2}{2}}$ où K est une constante ; et il y a une solution particulière évidente : la fonction constante $y = \frac{1}{2}$. Les solutions sont donc toutes les fonctions de la forme $y = \frac{1}{2} + Ke^{-\frac{x^2}{2}}$, où $K \in \mathbb{R}$ est une constante.

3. Cherchons toutes les solutions sur $\mathbb{R}$ de l'équation différentielle $y' - y = \cos(x)$. Les solutions de l'équation homogène $y' - y = 0$ sont les fonctions de la forme $x \mapsto Ke^x$, où $K \in \mathbb{R}$ est une constante. Pour trouver une solution de l'équation avec second membre, soit on trouve une solution particulière, soit on applique la méthode ci-dessus et on a besoin d'intégrer $x \mapsto e^{-x}\cos(x)$.

   On a déjà eu l'occasion de calculer une primitive de ce type ; on peut utiliser deux intégrations par parties ou passer par une intégrale complexe. Le calcul donne

$$\begin{aligned} \int_0^t e^{-x}\cos(x)dx &= \mathrm{Re}\left(\int_0^t e^{(-1+i)x} dx \right) \\ &= \mathrm{Re}\left(\left[\frac{e^{(-1+i)x}}{-1+i} \right]_0^t \right) \\ &= \mathrm{Re}\left(\frac{e^{(-1+i)t}(-1-i)+1+i}{2} \right) \\ &= \frac{-e^{-t}\cos(t) + e^{-t}\sin(t) + 1}{2}. \end{aligned}$$

Finalement, les solutions de notre équation sur $\mathbb{R}$ sont les fonctions de la forme

$$x \mapsto \left(\frac{-e^{-x}\cos(x) + e^{-x}\sin(x)}{2} + K\right)e^{x} = \frac{\sin(x) - \cos(x)}{2} + Ke^{x}$$

où K est une constante réelle. L'unique solution avec la condition initiale de passer par le point (x_0, y_0) est donc $x \mapsto \frac{\sin(x)-\cos(x)}{2} + (y_0 - \frac{\sin(x_0)-\cos(x_0)}{2})e^{x-x_0}$. Si on avait remarqué dès le départ que $x \mapsto \dfrac{\sin(x) - \cos(x)}{2}$ est une solution particulière (ce qui est confirmé par un calcul immédiat), on aurait obtenu ce résultat en s'épargnant le calcul d'une primitive de $x \mapsto e^{-x}\cos(x)$.

4. Trouvons la solution y de l'équation $y'(x) - y(x)\tan(x) = 0$ sur l'intervalle $I =]-\frac{\pi}{2}, \frac{\pi}{2}[$ qui vérifie $y(0) = \pi$. Une primitive de $-\tan$ sur I est $x \mapsto \ln\big(\cos(x)\big)$ (bien définie puisque $\cos(x) > 0$ pour $x \in I$) et les solutions de notre équation sont donc les fonctions de la forme $x \mapsto Ke^{-\ln\left(\cos(x)\right)} = \frac{K}{\cos(x)}$. On aura $y(0) = \pi$ si et seulement si $K = \pi$, et la fonction recherchée est donc $x \mapsto \frac{\pi}{\cos(x)}$.

Raccordements

L'équation différentielle réelle d'ordre 1 linéaire $y'(x) + a(x)y(x) = b(x)$ peut se concevoir sur des intervalles plus grands que sur l'intersection des domaines de définition des fonctions a et b. On rencontre ce cas plus fréquemment dans la formulation générale $\alpha(x)y' + \beta(x)y = \gamma(x)$ où α, β et γ sont continues sur un intervalle, mais où la fonction α peut s'annuler sans être identiquement nulle. La forme plus simple $y' + ay = b$ avec $a = \frac{\beta}{\alpha}$ et $b = \frac{\gamma}{\alpha}$ est appelée la forme *résolue*.

On détermine alors des solutions sur chacun des intervalles de définition conjointe des coefficients a et b et on cherche une solution maximale de continuité au bord des intervalles. Par exemple si a ou b n'est pas défini en α, on cherche y_- solution sur $]x_{-1}, \alpha[$ et y_+ solution sur $]\alpha, x_1[$, telles que

$$\begin{cases} \lim_{\alpha^-} y_- = \lim_{\alpha^+} y_+, \\ \lim_{\alpha^-} y'_- = \lim_{\alpha^+} y'_+. \end{cases}$$

On appelle ce prolongement une solution de l'équation différentielle sur $]x_{-1}, x_1[$, elle y est alors nécessairement $\mathcal{C}^1$. Cependant, le théorème de Cauchy-Lipschitz d'existence et d'unicité des solutions pour une condition $y(x_0) = y_0$ donnée ne tient plus, on peut ne plus avoir l'existence ou existence sans unicité.

Exemples

10.1 Considérons par exemple $y' + \tan(x)y = \cos(x)$ définie sur chaque intervalle $]k\pi - \frac{\pi}{2}, k\pi + \frac{\pi}{2}[$, $k \in \mathbb{Z}$, où elle a $\{x \mapsto (x + K)\cos(x) : K \in \mathbb{R}\}$ comme solutions. Il est clair que chacune de ces fonctions est en fait *définie sur* $\mathbb{R}$ tout entier, unique solution de continuité aux points $\frac{\pi}{2} + k\pi$. Il y a donc existence et unicité sur $\mathbb{R}$ tout entier d'une solution passant en un point (x_0, y_0) donné avec $x_0 \notin \mathbb{Z}\pi + \frac{\pi}{2}$, mais il n'y a pas existence de solution passant par (x_0, y_0) avec $x_0 \in \mathbb{Z}\pi + \frac{\pi}{2}$ et $y_0 \neq 0$ tandis qu'il n'y a pas unicité des solutions passant par $(x_0, 0)$ avec $x_0 \in \mathbb{Z}\pi + \frac{\pi}{2}$, car toutes les solutions définies sur $\mathbb{R}$ y passent. À part ce *pincement* en une *singularité*, on retrouve une unique droite affine de solutions définies sur tout $\mathbb{R}$.

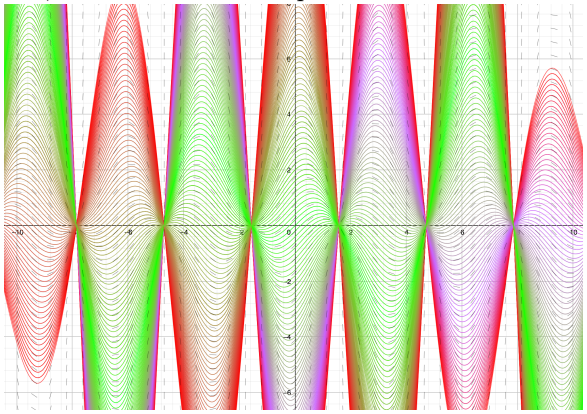


Figure 10.1 – Raccordement unique sur $\mathbb{R}$

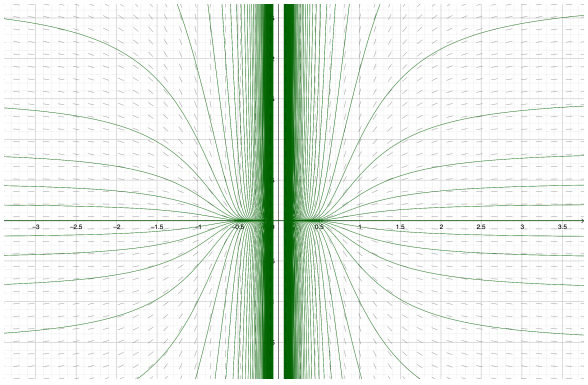


Figure 10.2 – Raccordements indépendants sur chaque intervalle

10.2 L'équation $y' - 2\frac{y}{x^3} = 0$ a comme solutions les fonctions de la forme $x \mapsto Ke^{-\frac{1}{x^2}}$ pour $K \in \mathbb{R}$, définies sur $]-\infty, 0[$ et définies sur $]0, +\infty[$. Elles sont toutes prolongeables en 0 par 0, et toutes leurs dérivées y

sont nulles. Donc la fonction raccordant $x \mapsto K_- e^{-\frac{1}{x^2}}$ sur $\mathbb{R}^-$ à $x \mapsto K_+ e^{-\frac{1}{x^2}}$ sur $\mathbb{R}^+$ est $\mathcal{C}^\infty(\mathbb{R})$, même avec $K_- \neq \mathbb{K}_+$. Les solutions sur $\mathbb{R}^-$ et $\mathbb{R}^+$ sont ainsi indépendantes, *il n'y a pas unicité des solutions*, il y a ici deux droites affines indépendantes de solutions qui peuvent se raccorder en une fonction définie sur tout $\mathbb{R}$.

10.3 En considérant $y' + (\frac{1}{x^2} - 1)y = \frac{e^x}{x^2}$, là encore les fonctions solutions ont exactement la même forme, $\mathbb{R}^{\pm *} \ni x \mapsto Ke^{x+1/x} + e^x$ avec $K \in \mathbb{R}$, et pas les mêmes ensembles de définitions, $\mathbb{R}^{-*}$ et $\mathbb{R}^{+*}$. Cependant, si toutes les solutions sur $\mathbb{R}^{-*}$ sont prolongeables en 0^- par $\begin{cases} y(0) = 1, \\ y'(0) = 1, \end{cases}$ la seule solution sur $\mathbb{R}^{+*}$ prolongeable en 0^+ est la solution particulière $x \mapsto e^x$, définie sur tout $\mathbb{R}$. Ceci implique

- pour les conditions initiales (x_0, y_0) avec $x_0 < 0$, l'existence et l'unicité d'une solution *définie sur tout* $\mathbb{R}$: Elle est de la forme $x \mapsto Ke^{x+1/x} + e^x$ pour $x \leqslant 0$ et $x \mapsto e^x$ pour $x \geqslant 0$;
- pour les conditions initiales (x_0, y_0) avec $x_0 > 0$ et $y_0 \neq e^{x_0}$, l'existence et l'unicité de la solution maximale $x \mapsto Ke^{x+1/x} + e^x$ seulement *définie sur* $\mathbb{R}^{+*}$, non prolongeable en 0 ;
- pour les données initiales du type (x_0, e^{x_0}) avec $x_0 \geqslant 0$, $\mathbb{R}^+ \ni x \mapsto e^x$ est prolongeable sur $\mathbb{R}^-$ par n'importe quelle fonction du type $\mathbb{R}^- \ni x \mapsto Ke^{x+1/x} + e^x$ avec $K \in \mathbb{R}$, et il n'y a donc *pas unicité* de la solution !

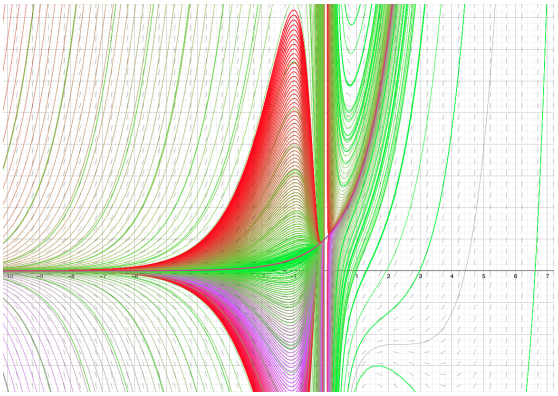
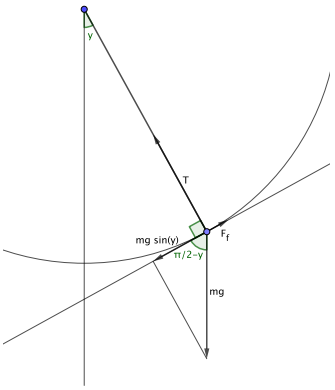


FIGURE 10.3 – Ensemble de définition dépendant de la condition

10.2 Équations linéaires réelles d'ordre 2, à coefficients constants

Tout d'abord, une équation différentielle d'ordre 2 donne la dérivée seconde $y'' = f(t, y, y')$ en fonction de la dérivée première y', de la position y et du temps t. Si y décrit la position d'un mobile au cours du temps, une telle équation va décrire *l'accélération* de ce mobile en fonction de sa position et sa vitesse, ce qui est physiquement la manière dont on décrit les choses : on agit sur un mobile en lui exerçant une force, ce qui conduit à lui donner une accélération, c'est le principe fondamental de la dynamique ou seconde loi de NEWTON. On les rencontre donc dans presque toutes les modélisations de phénomènes naturels et on note souvent la dérivée par rapport au temps avec un point plutôt qu'un prime : $\dot{y} = \frac{\mathrm{d}y}{\mathrm{d}t}$, $\ddot{y} = f(t, y, \dot{y})$ pour la différencier de la dérivée par rapport à la position physique x quand on a deux variables.

Exemple. *Le pendule pesant* Une masse est attachée à une tige légère de longueur fixe ℓ tournant autour d'un axe. Le système a un unique degré de liberté, l'angle y qu'il fait avec la verticale. Les forces exercées sur le mobile sont : la tension dans la tige (le long de la tige), la pesanteur (g constante et vers le bas) et le frottement (proportionnel d'un facteur μ et contraire à la vitesse y'). Le fait que la tige est rigide nous indique que la tension annule la composante radiale de la force et que seule compte sa composante angulaire, projection du poids perpendiculairement à la tige, donnant finalement l'équation :



$$y'' + \mu y' + \frac{g}{\ell} \sin(y) = 0. \tag{10.1}$$

Cette équation n'est pas linéaire, notez la fonction sin, ce qui fait qu'elle est difficile à résoudre. On peut cependant, en prenant un développement limité pour de petits angles, *linéariser* cette équation en une équation différentielle linéaire d'ordre 2 à coefficients constants dont la solution sont des *oscillations harmoniques amorties*, qui apparaissent dans énormément de modélisations de phénomènes réels.

On se concentrera sur un type particulier mais commençons par le cas général :
Une équation différentielle *linéaire* d'ordre 2 est de la forme

$$\forall t \in I, \quad y''(t) + a(t)y'(t) + b(t)y(t) = c(t) \tag{10.2}$$

où a, b, c sont des fonctions continues de I dans $\mathbb{R}$.

Une telle équation se ramène à une équation du premier ordre mais sur les matrices : la fonction inconnue y et sa dérivée y' composent la fonction à valeur vectorielle $Y = \begin{pmatrix} y \\ y' \end{pmatrix}$ qui vérifie

$$\forall t \in I, \quad Y'(t) + \mathcal{A}(t)Y(t) = \mathcal{B}(t) \tag{10.3}$$

où $\mathcal{A}(t) = \begin{pmatrix} 0 & -1 \\ b(t) & a(t) \end{pmatrix}$ et $\mathcal{B}(t) = \begin{pmatrix} 0 \\ c(t) \end{pmatrix}$. Et la méthode expliquée précédemment pourrait s'appliquer en trouvant la primitive matricielle A de la fonction à valeur matricielle $\mathcal{A}$, et la primitive vectorielle K de la fonction à valeur vectorielle $t \mapsto e^{A(t)}\mathcal{B}(t)$, quand nous saurons faire ces opérations. En tout cas c'est une méthode adaptée aux simulations numériques par un schéma d'Euler.

Dans ce chapitre nous nous contenterons de considérer des équations *à coefficients constants*, de la forme

$$\forall t \in I, \quad y''(t) + ay'(t) + by(t) = c(t) \ , \tag{E}$$

où a et b sont des constantes et seulement $c\colon I \to \mathbb{R}$ est une fonction continue (le forçage, une force extérieure comme l'action d'un moteur par exemple). La solution y recherchée doit être deux fois dérivable sur I, et sera automatiquement de classe $\mathcal{C}^2$ puisque $y'' = c - ay' - by$ sera une combinaison linéaire de fonctions continues.

On voit que si y_1, y_2 sont solutions de (E) alors $y_1 - y_2$ est une solution de l'équation homogène

$$\forall t \in I, \quad y''(t) + ay'(t) + by(t) = 0 \ . \tag{E_0}$$

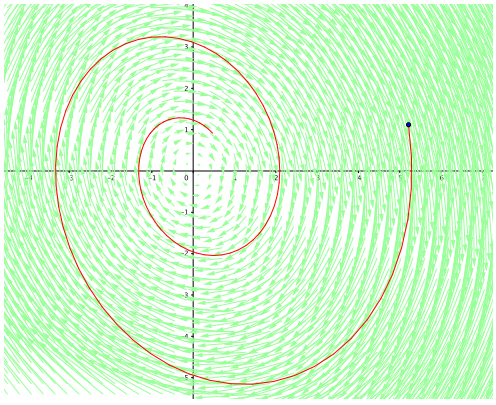
Réciproquement, si y_1 est solution de (E), et y_2 est solution de (E_0), alors $y_1 + \lambda y_2$ est solution de (E) pour tout $\lambda \in \mathbb{R}$: on voit donc que les solutions de (E) sont obtenues comme somme d'une solution particulière de (E) et d'une solution de l'équation homogène (E_0), exactement comme dans le cas des équations linéaires réelles d'ordre 1. Dans le cas des coefficients constants, qui est celui qui nous occupe ici, on va voir qu'on peut trouver les solutions de l'équation homogène. Trouver une solution particulière nécessite par contre plus de travail...

Observation. *Interprétation géométrique* La situation est compréhensible non plus dans le plan (x, y) mais dans le plan (y, y') de la position et de la vitesse, qu'on appelle *l'espace des phases*. On préfère utiliser la variable du temps plutôt que l'abscisse x pour décrire l'aspect dynamique du système, en tout point décrit par position et vitesse mais où le temps n'est pas représenté. En tout point $(y, y') = (q, p)$, le vecteur (constant par rapport au temps) $\begin{pmatrix} p \\ -ap - bq \end{pmatrix}$ décrit les modifications $Y' = \begin{pmatrix} y' \\ y'' \end{pmatrix}$ de la position et de la vitesse $Y = \begin{pmatrix} y \\ y' \end{pmatrix}$ selon la dynamique homogène (E_0).

Notez qu'une condition initiale n'est plus seulement donnée par une position $y(x_0) = y_0$ mais par une position $y(t_0) = q$ et une vitesse $y'(t_0) = p$.

Attention : Quand, dans les équations différentielles du premier ordre, dans le plan (x, y), nous n'avions jamais deux solutions différentes qui occupaient le même point, ici, au contraire pour une équation du second ordre, deux solutions décalées dans le temps (donc différentes en tant que fonctions dépendant du temps), sont représentées superposées dans l'espace des phases !

En rajoutant le second terme, le terme non homogène $c(t)$, la dynamique dépendante du temps est alors plus compliquée et le vecteur Y' en un point (q, p) n'est plus constant par rapport au temps, le vent change ! L'interprétation géométrique est donc plus délicate.



Définition 10.3. Étant donnée une équation différentielle linéaire homogène d'ordre 2, à coefficients constants, $y'' + ay' + by = 0$, on introduit son *équation caractéristique*

$$P(X) = X^2 + aX + b = 0.$$

Théorème 10.4. *Soit (E_0) : $y'' + ay' + by = 0$ une équation différentielle linéaire homogène d'ordre 2, à coefficients constants.*

1. *Si l'équation caractéristique $P(X) = X^2 + aX + b = 0$ de (E_0) admet deux solutions réelles $\alpha \neq \beta$, alors les solutions de (E_0)sont toutes les fonctions de la forme*

$$t \mapsto K_1 e^{\alpha t} + K_2 e^{\beta t}$$

où K_1, K_2 sont des constantes réelles.

2. *Si l'équation caractéristique admet une racine réelle double α, alors les solutions de (E_0) sont toutes les fonctions de la forme*

$$t \mapsto K_1 e^{\alpha t} + K_2 t e^{\alpha t} \ ,$$

où K_1, K_2 sont des constantes réelles.

3. *Si l'équation caractéristique admet deux racines complexes conjuguées $\alpha \pm i\beta$, alors les solutions de (E_0) sont toutes les fonctions de la forme*

$$t \mapsto K_1 e^{\alpha t}\cos(\beta t) + K_2 e^{\alpha t}\sin(\beta t) \ ,$$

où K_1, K_2 sont des constantes réelles, solutions qu'on peut aussi écrire sous la forme

$$t \mapsto A e^{\alpha t}\cos(\beta t - \phi)$$

où A, appelée l'amplitude *et ϕ, la* phase *sont des réels quelconques.*

Comme l'équation caractéristique est une équation polynomiale réelle de degré 2, on est forcément dans un des trois cas ci-dessus.

Preuve que les fonctions décrites ci-dessus sont bien des solutions. Pour le moment, on va se contenter de vérifier que, dans chacun des trois cas ci-dessus, les fonctions qui sont décrites ci-dessus sont bien des solutions ; on justifiera plus tard le fait que toutes les solutions sont de cette forme. Pour le premier cas, on doit vérifier que si α est solution de l'équation caractéristique alors $y\colon t \mapsto e^{\alpha t}$ est solution de (E_0). En effet, on a

$$y''(t) + ay'(t) + by(t) = (\alpha^2 + a\alpha + b)e^{\alpha t} = 0.$$

Pour le deuxième cas, il nous faut voir que, si α est racine double, alors $y\colon t \mapsto te^{\alpha t}$ est également solution de (E_0). Pour le démontrer, on commence par calculer que $y'(t) = e^{\alpha t} + \alpha t e^{\alpha t}$, et $y''(t) = 2\alpha e^{\alpha t} + \alpha^2 t e^{\alpha t}$, donc

$$y''(t) + ay'(t) + by(t) = (\alpha^2 + a\alpha + b)te^{\alpha t} + (2\alpha + a)e^{\alpha t} \ .$$

Comme α est racine double de P, on a $P(\alpha) = \alpha^2 + a\alpha + b = 0$ et $P'(\alpha) = 2\alpha + a = 0$.

Pour traiter le troisième cas, comme on s'autorise à dériver des fonctions à valeurs complexes, on voit que $t \mapsto e^{(\alpha+i\beta)t}$ et $t \mapsto e^{(\alpha-i\beta)t}$ sont toutes deux solutions de (E_0) (puisque $\alpha \pm i\beta$ sont racines de l'équation caractéristique ; c'est le même calcul que celui du premier point), donc leurs combinaisons linéaires aussi, ce qui nous permet de conclure que $t \mapsto e^{\alpha t}\cos(\beta t)$ et $t \mapsto e^{\alpha t}\sin(\beta t)$ sont toutes deux solutions de (E_0). Si on ne veut pas dériver des fonctions à valeurs complexes, on fait simplement le calcul, par exemple en posant $y(t) = e^{\alpha t}\cos(\beta t)$ on a

$$y'(t) = \alpha e^{\alpha t}\cos(\beta t) - \beta e^{\alpha t}\sin(\beta t) \quad \text{et}$$
$$y''(t) = \alpha^2 e^{\alpha t}\cos(\beta t) - 2\alpha\beta e^{\alpha t}\sin(\beta t) - \beta^2 e^{\alpha t}\cos(\beta t).$$

Par conséquent, on arrive à

$$y''(t) + ay'(t) + by(t) = \cos(\beta t)e^{\alpha t}(\alpha^2 - \beta^2 + a\alpha + b) + \sin(\beta t)e^{\alpha t}(-2\alpha\beta - a\beta) \ .$$

Comme par définition de α, β on a $(\alpha + i\beta)^2 + a(\alpha + i\beta) + b = 0$, on obtient

$$(\alpha^2 - \beta^2 + a\alpha + b) + i(2\alpha\beta + a\beta) = 0.$$

Comme α, β sont réels on en déduit $\alpha^2 - \beta^2 + a\alpha + b = 0 = 2\alpha\beta + a\beta$ et on trouve bien comme attendu que y est solution.

En remarquant que $\cos(\phi)\cos(\beta t) + \sin(\phi)\sin(\beta t) = \cos(\beta t - \phi)$ on a $K_1 e^{\alpha t}\cos(\beta t) + K_2 e^{\alpha t}\sin(\beta t) = Ae^{\alpha t}\cos(\beta t - \phi)$ pour $A = \sqrt{K_1^2 + K_2^2}$ et $\cos\phi = \frac{K_1}{A}$, $\sin\phi = \frac{K_2}{A}$. Suivant les conditions, une forme peut être plus adaptée que l'autre. □

Dans le cas complexe, la partie réelle $\alpha = -a/2$ pilote le comportement :

— $\alpha = 0$, c'est-à-dire $a = 0, b > 0$ dans (E_0), on a à faire à un *oscillateur harmonique*, périodique, purement circulaire conservatif, de pulsation propre $\beta = \sqrt{b}$.

— Sinon, on appelle le régime *pseudo-périodique*,

  — $\alpha < 0$, c'est un *oscillateur harmonique amorti*, dissipatif ($a > 0$ est un coefficient de frottement), dont l'amplitude tend vers 0 et

  — $\alpha > 0$, c'est un *oscillateur harmonique amplifié* dont l'amplitude diverge.

Il faut faire attention dans le cas complexe, l'espace vectoriel solution est

$$\mathrm{Re}\big(\mathrm{Vect}_{\mathbb{C}}(t \mapsto e^{(\alpha+i\beta)t}, t \mapsto e^{(\alpha-i\beta)t})\big) = \mathrm{Vect}_{\mathbb{R}}\big(t \mapsto e^{\alpha t}\cos(\beta t), t \mapsto e^{\alpha t}\sin(\beta t)\big)$$

qui contient strictement $\mathrm{Vect}_{\mathbb{R}}\big(t \mapsto \mathrm{Re}(e^{(\alpha+i\beta)t}), t \mapsto \mathrm{Re}(e^{(\alpha-i\beta)t})\big) = \mathrm{Vect}_{\mathbb{R}}\big(t \mapsto e^{\alpha t}\cos(\beta t)\big)$, mais aussi $\mathrm{Vect}_{\mathbb{R}}\big(t \mapsto \mathrm{Im}(e^{(\alpha+i\beta)t}), t \mapsto \mathrm{Im}(e^{(\alpha-i\beta)t})\big) = \mathrm{Vect}_{\mathbb{R}}\big(t \mapsto e^{\alpha t}\sin(\beta t)\big)$.

Remarquons que, dans chacun des cas, on voit que l'espace formé par toutes les fonctions solutions de (E_0) est un espace vectoriel de dimension *au moins* 2 car l'énoncé ci-dessus nous donne, dans chacun des cas, deux fonctions non proportionnelles qui appartiennent à l'espace des solutions ; ce qui manque encore dans la preuve est donc la justification que l'espace des solutions est exactement de dimension 2, ce qu'on verra plus en détail ci-dessous.

Pour l'instant, notons qu'il nous reste à résoudre l'équation avec second membre ! Ce second membre correspond physiquement à une excitation extérieure, un couplage avec l'environnement, une source d'énergie comme une houle, le vent, un moteur... Quand le second membre $c(t)$ est particulièrement simple, on peut rechercher une solution particulière sous une forme déterminée à l'avance, comme par exemple dans le cas ci-dessous.

Proposition 10.5. *Soit I un intervalle de $\mathbb{R}$, et (E) une équation différentielle de la forme*

$$y''(t) + ay'(t) + by(t) = Q(t)e^{\lambda t}, \tag{E}$$

où $a, b, \lambda \in \mathbb{R}$ et $Q \in \mathbb{R}[X]$. Alors il existe une solution de (E) de la forme $t^k R(t)e^{\lambda t}$, où $R \in \mathbb{R}[X]$ est du même degré que Q et k est égal à la multiplicité de λ comme racine de l'équation caractéristique de la forme homogène de (E) (donc $k = 0$, 1 ou 2).

Démonstration. Notons n le degré de Q. Étant donnée y une fonction deux fois dérivable sur I, introduisons $z : t \mapsto y(t)e^{-\lambda t}$. Alors y est deux fois dérivable si, et seulement si, z est deux fois dérivable ; de plus, on a

$$\forall t \in \mathbb{R},\ y'(t) = z'(t)e^{\lambda t} + \lambda z(t)e^{\lambda t} \text{ et } y''(t) = z''(t)e^{\lambda t} + 2\lambda z'(t)e^{\lambda t} + \lambda^2 z(t)e^{\lambda t}\ .$$

On en déduit que

$$\forall t \in \mathbb{R},\ y''(t) + ay'(t) + by(t) = \big(z''(t) + (2\lambda + a)z'(t) + (\lambda^2 + a\lambda + b)z(t)\big)\, e^{\lambda t}\ .$$

En particulier, y est solution de (E) si, et seulement si,

$$\forall t \in \mathbb{R},\ z''(t) + (2\lambda + a)z'(t) + (\lambda^2 + a\lambda + b)z(t) = Q(t)\ .$$

Considérons l'endomorphisme des polynômes $\Phi : \begin{cases} \mathbb{R}[X] & \to \mathbb{R}[X], \\ R & \mapsto R'' + (2\lambda + a)R' + (\lambda^2 + a\lambda + b)R. \end{cases}$

Maintenant, on distingue trois cas :

1. Supposons $P(\lambda) = \lambda^2 + a\lambda + b \neq 0$, c'est-à-dire que λ n'est pas racine de l'équation caractéristique de (E). Alors Φ vérifie que que pour tout polynôme R, le degré de $\Phi(R)$ est égal au degré de R. Donc $\Phi(R) = 0$ seulement si $R = 0$ et Φ est injective de $\mathbb{R}_n[X]$ dans $\mathbb{R}_n[X]$. Alors elle est également surjective, car tout endomorphisme injectif d'un espace vectoriel de dimension finie est surjectif. Il existe donc R du même degré que Q tel que $\Phi(R) = Q$, et $y : t \mapsto R(t)e^{\lambda t}$ est alors solution de (E).
2. Si $P(\lambda) = \lambda^2 + a\lambda + b = 0$ et $P'(\lambda) = 2\lambda + a \neq 0$, c'est-à-dire si λ est racine simple de P, l'application $\Phi : R \mapsto R'' + (2\lambda + a)R'$ est maintenant de $\mathbb{R}_{n+1}[X]$ dans $\mathbb{R}_n[X]$. Le noyau de Φ est formé des polynômes constants, donc est de dimension 1 ; par conséquent l'image est de dimension $\dim(\mathbb{R}_{n+1}[X]) - 1 = n + 1$, et est contenue dans $\mathbb{R}_n[X]$, qui est aussi de dimension $n + 1$. Donc Φ est surjective : il existe R tel que $\Phi(R) = Q$, et puisque $\Phi(R - R(0)) = \Phi(R)$ on peut supposer que le terme constant de R est nul. De nouveau, l'application $y : t \mapsto R(t)e^{\lambda t}$ est solution de (E).

3. Si $P(\lambda) = \lambda^2 + a\lambda + b = 0$ et $P'(\lambda) = 2\lambda + a = 0$, c'est-à-dire si λ est racine double de P, on considère sur le même modèle l'application $\Phi\colon R \mapsto R''$, qui est une surjection de $\mathbb{R}_{n+2}[X]$ sur $\mathbb{R}_n[X]$, et on trouve R tel que $\Phi(R) = Q$. Alors $y\colon t \mapsto R(t)e^{\lambda t}$ est solution de (E) ; de plus, puisque les polynômes de degré $\leqslant 1$ sont dans le noyau de Φ, on peut supposer que les termes de degré 0 et 1 de R sont nuls.

□

Revenons maintenant sur la dimension de l'espace vectoriel formé par les solutions d'une équation différentielle linéaire réelle homogène d'ordre 2.

Lemme 10.6. *Soit (E_0) une équation différentielle linéaire réelle homogène d'ordre 2, à coefficients constants. Etant données y_1, y_2 deux solutions de (E_0), on introduit leur* wronskien *w défini par*

$$\forall t \in I, \quad w(t) = y_1(t)y_2'(t) - y_2(t)y_1'(t) \ .$$

S'il existe $t_0 \in I$ tel que $w(t_0) = 0$ alors w est nul sur I. Sur un intervalle où y_2 ne s'annule pas, on a alors $y_1 = \lambda y_2$ pour une constante $\lambda \in \mathbb{R}$.

Le wronskien s'annule si, dans l'espace des phases, les vecteurs $\begin{pmatrix} y_1 \\ y_1' \end{pmatrix}$ et $\begin{pmatrix} y_2 \\ y_2' \end{pmatrix}$ sont colinéaires. Alors la linéarité du système implique que les deux solutions partant de conditions initiales colinéaires, restent colinéaires.

Démonstration. On note que w est dérivable, et que pour $t \in I$ on a

$$\begin{aligned} w'(t) &= y_1'(t)y_2'(t) + y_1(t)y_2''(t) - y_2'(t)y_1'(t) - y_2(t)y_1''(t) \\ &= y_1(t)y_2''(t) - y_2(t)y_1''(t) \\ &= y_1(t)(-ay_2'(t) - by_2(t)) - y_2(t)(-ay_1'(t) - by_1(t)) \\ &= -ay_1(t)y_2'(t) + ay_2(t)y_1'(t) \\ &= -aw(t). \end{aligned}$$

On vient de montrer que w est solution d'une équation différentielle homogène de degré 1, à coefficients constants : il existe une constante K telle que $w(t) = Ke^{-at}$ pour tout $t \in I$, en particulier si w s'annule en un point alors $w(t) = 0$ pour tout $t \in I$. Sur un intervalle où y_2 ne s'annule pas, on peut dériver la fonction $t \mapsto \frac{y_1}{y_2}$, dont la dérivée est égale à $\frac{w(t)}{y_2^2(t)} = 0$: cette fonction est constante, et on a bien $y_1 = \lambda y_2$ pour un certain $\lambda \in \mathbb{R}$. □

Pour finir la preuve du théorème 10.4, il nous reste à démontrer le résultat suivant.

Proposition 10.7. *Soit (E_0) une équation différentielle linéaire réelle homogène d'ordre 2, à coefficients constants. Alors l'ensemble $\mathcal{E}$ formé par les solutions de (E_0) est un sous-espace vectoriel de dimension 2 de l'espace des fonctions de I dans $\mathbb{R}$.*

Voyons une première preuve, basée sur les propriétés du wronskien.

Démonstration. On a déjà mentionné le fait que $\mathcal{E}$ est un sous-espace vectoriel : il est non vide puisqu'il contient la fonction nulle, et dès que y_1, y_2 sont solutions et λ, μ sont des réels, $\lambda y_1 + \mu y_2$ est encore un élément de $\mathcal{E}$. De plus, on a vu dans la preuve du théorème 10.4 que $\mathcal{E}$ est de dimension au moins 2 (selon la forme des solutions de l'équation caractéristique, on a à chaque fois fourni deux solutions non proportionnelles y_1, y_2). Fixons deux solutions non proportionnelles y_1, y_2. Pour tout $t_0 \in I$, comme le wronskien de (y_1, y_2) ne s'annule pas sur I, on sait que $(y_1(t_0), y_1'(t_0))$ et $(y_2(t_0), y_2'(t_0))$ sont non colinéaires, ils forment donc une base de l'espace des phases $\mathbb{R}^2$ et pour tout vecteur position/vitesse $(\lambda, \mu) \in \mathbb{R}^2$ il existe $(A_1, A_2) \in \mathbb{R}^2$ tels que $(\lambda, \mu) = A_1\big(y_1(t_0), y_1'(t_0)\big) + A_2\big(y_2(t_0), y_2'(t_0)\big)$.

Ainsi, si y est une solution quelconque de (E_0), les valeurs $y(t_0)$ et $y'(t_0)$ définissent une combinaison linéaire $y(t_0) = A_1y_1(t_0) + A_2y_2(t_0)$ et $y'(t_0) = A_1y_1'(t_0) + A_2y_2'(t_0)$. Mais cette combinaison est solution de (E_0) sur tout I. De plus, le wronskien de y et $A_1y_1 + A_2y_2$ s'annule en t_0, et s'annule donc sur tout un segment autour de t_0 où y ne s'annule pas, pour peu que $y(t_0) \neq 0$, ce qu'on supposera.

Pour en déduire que cette combinaison linéaire est égale à y sur tout I, il reste à traiter les valeurs où la fonction y s'annule. Nous allons voir que dans tous les cas, ces zéros sont isolés et la combinaison linéaire s'étend par continuité d'un intervalle sans zéro à un autre.

Dans le cas où l'équation caractéristique de (E_0) est à racines réelles, on vérifie sur la forme des solutions $y_1(t) = e^{\alpha t}$, $y_2(t) = e^{\beta t}$ que si $z = A_1y_1 + A_2y_2$ n'est pas la fonction nulle alors z s'annule au plus une fois sur I, disons en t_1 et qu'alors $z'(t_1) \neq 0$. Alors, $I \setminus \{t_1\}$ s'écrit comme réunion de deux intervalles I_1, I_2 sur lesquels z ne s'annule pas, et puisque le wronskien de (y, z) est nul sur I_1 et I_2 on en déduit l'existence de constantes λ_1, λ_2 telles que $y = \lambda_1 z$ sur I_1, $y = \lambda_2 z$ sur I_2. Par continuité de y' sur I, on obtient alors $\lambda_1 z'(t_1) = \lambda_2 z'(t_1)$, donc $\lambda_1 = \lambda_2$ puisque $z'(t_1) \neq 0$. Donc finalement $\lambda_1 = \lambda_2$ et $y = \lambda z$ pour tout $z \in I \setminus \{t_1\}$; cette égalité doit être vraie aussi en t_1 par continuité, donc finalement il existe λ tel que $y = \lambda z$ sur I. Comme on a choisi z de façon que $y(t_0) = z(t_0)$ on obtient $y = z$.

Dans le cas où l'équation caractéristique est à racines double ou complexes, $z = A_1y_1 + A_2y_2$ peut s'annuler plus d'une fois sur I (et même une infinité de fois si I est non borné). Néanmoins, on peut écrire I comme réunion d'une suite d'intervalles dans l'intérieur desquels z ne s'annule pas, et adapter le raisonnement ci-dessus pour conclure de nouveau que $z = y$ (on ne va pas détailler cet argument ici).

Dans tout les cas, on obtient que, étant donné $y \in \mathcal{E}$ on peut trouver A_1, A_2 tels que $y = A_1y_1 + A_2y_2$: tout élément de $\mathcal{E}$ s'écrit de manière unique comme combinaison linéaire de y_1 et y_2, ce qui montre que $\mathcal{E}$ est de dimension 2. □

Le raisonnement précédent contient en filigrane le résultat suivant, qui est un cas (très) particulier du théorème de Cauchy–Lipschitz .

Proposition 10.8. *Soit (E_0) une équation différentielle linéaire réelle homogène d'ordre 2, à coefficients constants. Etant donné $t_0 \in I$ et $\lambda, \mu \in \mathbb{R}$, il existe une unique solution de E_0 telle que $y_1(t_0) = \lambda$ et $y_1'(t_0) = \mu$.*

Démonstration. Notons $\mathcal{E}$ l'espace des solutions de (E_0), dont on a vu qu'il était de dimension 2. Considérons l'application $\Phi\colon \mathcal{E} \to \mathbb{R}^2$, définie par $\Phi(y) = (y(t_0), y'(t_0))$. On doit montrer que Φ est bijective ; comme $\mathcal{E}$ et $\mathbb{R}^2$ ont la même dimension, et Φ est linéaire, il nous suffit de montrer que Φ est surjective. Etant donnée une base (y_1, y_2) de $\mathcal{E}$, on a vu plus haut que $y_1(t_0)y_2'(t_0) - y_2(t_0)y_1'(t_0) \neq 0$, et c'est exactement ce dont on a besoin pour assurer que, pour tout $\lambda, \mu \in \mathbb{R}$, le système $\begin{cases} Ay_1(t_0) + By_2(t_0) &= \lambda, \\ Ay_1'(t_0) + By_2'(t_0) &= \mu, \end{cases}$ admet une solution A, B. Alors $\Phi(Ay_1 + By_2) = (\lambda, \mu)$, prouvant que Φ est surjective. □

On peut aussi donner une preuve directe de la proposition 10.8 et en déduire la proposition 10.7 :

Une autre preuve de 10.8. Considérons l'équation $y'' + ay' + by = 0$, où a, b sont des constantes ; on va de nouveau utiliser l'application Φ définie dans la preuve précédente. Cette fois on va montrer que Φ est injective, ce qui nous montrera que $\mathcal{E}$ est de dimension au plus 2, et donc exactement 2. Soit donc y dans le noyau de Φ ; on introduit la fonction $u\colon t \mapsto (y(t))^2 + (y'(t))^2$. On va utiliser l'inégalité $2xy \leqslant x^2 + y^2$ valable pour tout $x, y \in \mathbb{R}$.

En dérivant, on obtient

$$\begin{aligned} u'(t) &= 2y'(t)y(t) + 2y'(t)y''(t) \\ &= 2y'(t)\big(y''(t) + y(t)\big) \\ &= 2y'(t)\Big((1-b)y(t) - a\big(y'(t)\big)\Big) \\ &= (1-b)2y(t)y'(t) - 2a\big(y'(t)\big)^2 \\ &\leqslant |1-b|((y(t))^2 + (y'(t))^2) + 2|a|(y'(t))^2 \\ &\leqslant Cu(t) \end{aligned}$$

où $C = |1-b| + 2|a|$. Au lieu d'avoir une équation différentielle linéaire du premier ordre, on est face à une inéquation différentielle ! En repensant au cas des équations d'ordre 1, on introduit $\varphi\colon t \mapsto e^{-Ct}u(t)$ et on voit que $\varphi'(t) = (-Cu(t) + u'(t))e^{-Ct} \leqslant 0$ pour tout t. Comme φ est à valeurs positives et $\varphi(0) = 0$, on en déduit $\varphi(t) = 0$ et donc $u(t) = 0$ pour tout $t \geqslant 0$. Le même raisonnement appliqué à l'équation différentielle satisfaite par $t \mapsto y(-t)$ nous permet de conclure que $u(-t) = 0$ pour tout $t \geqslant 0$, donc u est la fonction nulle. On vient de montrer que $y(t) = 0$ pour tout t, ce qui prouve que le seul élément du noyau de Φ est la fonction nulle. □

Ci-dessous, on va montrer que ce résultat reste vrai quand l'équation n'est pas homogène ; il nous suffit en fait de justifier qu'une équation différentielle réelle d'ordre 2, à coefficients constants, a toujours une solution. Pour cela, on va décrire la

Proposition 10.9 (Méthode de Laplace)**.** *L'équation* (E) $y''+ay'+by=c(t)$ *définie sur un intervalle* I *par les constantes* $a,b\in\mathbb{R}$ *et* $c:I\to\mathbb{R}$ *une fonction continue, a comme solutions*

$$y(t)=\lambda_1(t)y_1(t)+\lambda_2(t)y_2(t)$$

où (y_1,y_2) *forment une base de l'espace des solutions de l'équation homogène et* λ_1*, resp.* λ_2 *sont des primitives des fonctions*

$$\forall t\in I,\quad \lambda_1'(t)=-\frac{c(t)y_2(t)}{w(t)},\ \textit{resp. }\lambda_2'(t)=+\frac{c(t)y_1(t)}{w(t)}$$

où $w(t)=y_1(t)y_2'(t)-y_1'(t)y_2(t)$ *est le wronskien de* y_1 *et* y_2*.*

Démonstration. Trouvons des conditions sur λ_1,λ_2 qui suffisent à ce que y soit solution (on a seulement besoin de trouver une solution particulière !) : d'abord, on dérive

$$y'(t)=\lambda_1'(t)y_1(t)+\lambda_1(t)y_1'(t)+\lambda_2'(t)y_2(t)+\lambda_2(t)y_2'(t)\ .$$

Pour simplifier la forme de $y'(t)$, on demande que $\lambda_1'(t)y_1(t)+\lambda_2'(t)y_2(t)=0$ (ce qui impose une équation reliant λ_1' et λ_2') ; si cette condition est vérifiée, on a simplement $y'(t)=\lambda_1(t)y_1'(t)+\lambda_2(t)y_2'(t)$ et

$$y''(t)=\lambda_1'(t)y_1'(t)+\lambda_1(t)y_1''(t)+\lambda_2'(t)y_2'(t)+\lambda_2(t)y_2''(t)\ .$$

Alors, on obtient en regroupant que $y''(t)+ay'(t)+by(t)$ est égal à

$$\begin{aligned}&(\lambda_1(t)y_1''(t)+\lambda_2(t)y_2''(t)+a\lambda_1(t)y_1'(t)+a\lambda_2(t)y_2'(t)+b\lambda_1(t)y_1(t)+b\lambda_2(t)y_2(t))+\lambda_1'(t)y_1'(t)+\lambda_2'(t)y_2'(t)\\&=\lambda_1'(t)y_1'(t)+\lambda_2'(t)y_2'(t).\end{aligned}$$

Finalement, on obtient que, dès lors que l'on a

$$\forall t\in I,\quad \begin{cases}\lambda_1'(t)y_1(t)+\lambda_2'(t)y_2(t) &=0\\ \lambda_1'(t)y_1'(t)+\lambda_2'(t)y_2'(t) &=c(t)\end{cases}$$

la fonction $t\mapsto\lambda_1(t)y_1(t)+\lambda_2(t)y_2(t)$ est une solution de (E). Le système en $(\lambda_1'(t),\lambda_2'(t))$ obtenu ci-dessus a une solution exactement parce que le wronskien de y_1,y_2 ne s'annule pas, autrement dit $y_1(t)y_2'(t)-y_2(t)y_1'(t)\neq 0$ pour tout $t\in I$! En effet, en résolvant le système ci-dessus on obtient

$$\forall t\in I,\quad \lambda_1'(t)=\frac{-c(t)y_2(t)}{y_1(t)y_2'(t)-y_1'(t)y_2(t)}\ \text{et}\ \lambda_2'(t)=\frac{-c(t)y_1(t)}{y_2(t)y_1'(t)-y_2'(t)y_1(t)}\ .$$

En choisissant pour λ_1,λ_2 des primitives pour les fonctions (continues, puisque le dénominateur ne s'annule pas) qu'on vient d'obtenir ci-dessus, nos calculs nous montrent que $\lambda_1y_1+\lambda_2y_2$ est une solution de E. □

Théorème 10.10 (Théorème de Cauchy–Lipschitz)**.** *Soit* (E) *une équation différentielle linéaire réelle d'ordre 2, à coefficients constants, et* $t_0\in I$*. Pour tout* $(\lambda,\mu)\in\mathbb{R}^2$ *il existe une unique solution* y *de* (E) *telle que* $y(t_0)=\lambda$ *et* $y'(t_0)=\mu$*.*

Démonstration. Commençons par trouver une solution z de (E), ce qui est possible d'après ce qu'on a fait ci-dessus ; on sait par 10.8 qu'il existe une solution y_0 de l'équation homogène associée à (E) telle que $y_0(t_0)=\lambda-z(t_0)$ et $y_0'(t_0)=\mu-z'(t_0)$. Alors $y=z+y_0$ est une solution de (E), et on a bien $y(t_0)=\lambda$, $y'(t_0)=\mu$. Ceci montre l'existence d'une solution satisfaisant nos conditions initiales ; pour montrer qu'elle est unique, supposons que y_1,y_2 sont deux solutions de (E) satisfaisant les mêmes conditions initiales. Alors $z=y_1-y_2$ est une solution de l'équation homogène (E_0) satisfaisant $z(0)=z'(0)=0$. On déduit alors de la proposition 10.8 que $z=0$, c'est-à-dire $y_1=y_2$. □

10.3 Équations différentielles linéaires en dimension supérieure

On fixe un entier n, et I un intervalle de $\mathbb{R}$. Étant donnée $Y\colon I\to\mathbb{R}^n$, on dit que Y est continue (resp. dérivable) si chacune des coordonnées de Y est une fonction continue (resp. dérivable). Une équation différentielle linéaire d'ordre 1 est alors une équation d'inconnue Y, où $Y\colon I\to\mathbb{R}^n$ est dérivable, de la forme

$$\forall t\in I,\ Y'(t)+A(t)Y(t)=B(t)\ ,$$

où $A\colon I \to M_n(\mathbb{R})$ et $B\colon I \to \mathbb{R}^n$ sont continues. Un cas particulier important est celui où A est une fonction constante, c'est-à-dire qu'il existe une matrice A telle que $A(t) = A$ pour tout $t \in I$.

Notons que toute équation différentielle d'ordre 2 à coefficients constants peut être vue sous cette forme : considérons l'équation

$$\forall t \in I, \quad y''(t) + ay'(t) + by(t) = c(t)\ , \tag{E}$$

et introduisons la matrice $A = \begin{pmatrix} 0 & -1 \\ b & a \end{pmatrix}$, $C\colon t \mapsto \begin{pmatrix} 0 \\ c(t) \end{pmatrix}$, et l'équation

$$\forall t \in I, \quad Y'(t) + A(t)Y(t) = C(t)\ . \tag{E'}$$

Proposition 10.11. *Si y est solution de (E) alors, $Y = \begin{pmatrix} y \\ y' \end{pmatrix}$ est solution de (E') ; si $Y = \begin{pmatrix} y \\ z \end{pmatrix}$ est solution de (E') alors $z = y'$ et y est solution de (E).*

Démonstration. Supposons y solution de (E), et calculons

$$\begin{aligned} Y'(t) + AY(t) &= \begin{pmatrix} y'(t) \\ y''(t) \end{pmatrix} + \begin{pmatrix} -y'(t) \\ by(t) + ay'(t) \end{pmatrix} \\ &= \begin{pmatrix} y'(t) - y'(t) \\ y''(t) + ay'(t) + by(t) \end{pmatrix} = \begin{pmatrix} 0 \\ c(t) \end{pmatrix} = C(t). \end{aligned}$$

Réciproquement, soit $Y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ une solution de (E'). Pour tout $t \in I$ on a

$$\begin{aligned} Y'(t) + AY(t) &= \begin{pmatrix} y'(t) \\ z'(t) \end{pmatrix} + \begin{pmatrix} -z(t) \\ by(t) + ay'(t) \end{pmatrix} \\ &= \begin{pmatrix} y'(t) - z(t) \\ z'(t) + by(t) + ay'(t) \end{pmatrix}. \end{aligned}$$

Par conséquent, Y est solution de (E') si, et seulement si, on a pour tout $t \in I$ à la fois $z(t) = y'(t)$ et $z'(t) + ay'(t) + by(t) = c(t)$, autrement dit si et seulement si $z = y'$ et y est solution de (E). □

En généralisant la méthode qu'on vient de voir, les équations différentielles réelles linéaires d'ordre quelconque peuvent se ramener à des équations différentielles linéaires d'ordre 1 dans $\mathbb{R}^n$; concluons en mentionnant le théorème suivant, qui généralise les théorèmes qu'on a vus pour les équations différentielles réelles linéaires d'ordre 1 et 2.

Théorème 10.12 (Théorème de Cauchy–Lipschitz linéaire)**.** *Soit I un intervalle de $\mathbb{R}$, et (E) l'équation différentielle linéaire d'ordre* 1

$$\forall t \in I, \quad Y'(t) + A(t)Y(t) = B(t)\ ,$$

où $A\colon I \to M_n(\mathbb{R})$ et $B\colon I \to \mathbb{R}^n$ sont des fonctions continues. Alors, pour tout $Y_0 \in \mathbb{R}^n$ et tout $t_0 \in I$ il existe une unique solution Y de (E) telle que que $Y(t_0) = Y_0$.

Exercice 10.13. En utilisant le résultat ci-dessus, montrer que l'espace formé par les solutions de l'équation homogène $Y' + AY = 0$ est de dimension n.

10.4 Quelques équations différentielles non linéaires

On a déjà vu l'équation du pendule pesant (10.1) $y'' + \mu y' + \frac{g}{\ell}\sin(y) = 0$ qui n'est pas linéaire mais linéarisable pour de petits angles, donnant lieu aux oscillations harmoniques amorties (ou amplifiées) qui se retrouvent partout en physique, dans la propagation des ondes, la modélisation des oscillateurs électriques RLC, les ressorts, les puits de potentiel... D'autres équations sont historiquement et pratiquement importantes, en particulier l'équation *Proies-prédateurs* de Lotka-Volterra et les équations épidémiologiques comme le modèle SIR. Comprendre le processus de modélisation qui leur donne sens est important pour un scientifique d'aujourd'hui.

Modélisation d'une population : la suite géométrique

Imaginons une population de lapins. Son accroissement est *a priori* proportionnel à sa population : plus il y a de lapins, plus il y a de petits. La première modélisation d'une telle population est ainsi discrète, on appellera u_n le nombre de lapins à la n-ème portée et on a la définition récurrente simple : l'accroissement de la population à la $n+1$-ème portée est proportionnelle d'un facteur q à celle de la n-ème portée, soit $u_{n+1} = u_n + qu_n$, ce qui nous amène à la solution $u_n = (1+q)^n u_0$, la suite géométrique de premier terme u_0 et de raison $1+q$.

C'est (malheureusement) également le cas des débuts d'épidémie. Nous avons énormément entendu parlé de croissance exponentielle de l'épidémie de COVID-19 car, représentés en fonction du temps le nombre de cas des premiers jours d'une épidémie donnent lieu, sur une échelle logarithmique, à une courbe presque droite, c'est-à-dire de la forme $t \mapsto N_0 e^{\lambda(t-t_0)}$ où le coefficient λ dépend du R_0, le nombre de reproduction de base. Diverses mesures des autorités tendent à modifier le comportement des personnes pour diminuer ce nombre de reproduction et *applatir la courbe*.

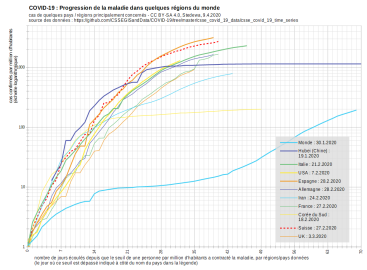


FIGURE 10.4 – Croissance exponentielle de l'épidémie Covid-19 Wikipedia.

Modélisation d'une population : la suite logistique

La première critique de cette modélisation est que la contrainte liée à la pression écologique d'une population à la croissance exponentielle n'est pas prise en compte. De fait, l'introduction de lapins dans un nouvel environnement avec peu de prédateurs, comme une île, est suivie d'une explosion exponentielle de leur nombre, puis d'une stabilisation, voire même d'un effondrement si le pâturage n'est pas soutenable. Comment prendre en compte cet effet ? En modulant l'accroissement de la population, la rendant nulle à une population d'équilibre u_e, soit

$$u_{n+1} = u_n + qu_n(u_e - u_n).$$

Ainsi, si la population est inférieure à la population d'équilibre, elle va croître, mais si elle est supérieure, elle va décroître. Sa version continue $y' = qy(p-y)$, où p est la population d'équilibre, est beaucoup plus régulière que la version discrète : elle ralentit progressivement vers l'équilibre là où la version discrète peut dépasser l'équilibre et introduire un comportement *chaotique*.

On résume le comportement de cette équation sous la forme plus simple à étudier $u_{n+1} = f_a(u_n)$ avec $f_a(x) = ax(1-x)$ un polynôme de degré 2, s'annulant en 0 et en 1. On appelle une telle suite *la suite logistique*. Remarquons que son sommet est en $(\frac{1}{2}, \frac{a}{4})$, donc pour $0 \leqslant a \leqslant 4$, f_a est une application de $[0,1]$ dans lui-même, on peut donc l'itérer sans explosion. Si la suite converge, vu que f_a est continue, sa limite ℓ est forcément un point fixe de f_a et $x = f_a(x)$ a deux solutions, la solution nulle et $\ell = 1 - \frac{1}{a}$. Pour $0 \leqslant a \leqslant 1$, $\ell < 0$ donc ce point fixe n'est pas attractif et la suite converge vers 0, l'autre point fixe où $f_a(x) = ax - ax^2 = ax + O(x^2)$ donc la suite $u_{n+1} = f_a(u_n)$ est comparable à une suite géométrique de raison $a < 1$ et la suite converge vers 0.

Pour $1 < a < 4$, étudions le comportement de la fonction autour du point fixe ℓ : $f_a(x) = f_a(\ell) + (x-\ell)f'_a(\ell) + \frac{1}{2}(x-\ell)^2 f''_a(\ell)$, avec $f_a(\ell) = \ell$, $f'_a(\ell) = a(1-2\ell) = 2-a$ et $f''_a(\ell) = -2a$, soit $f_a(x) - \ell = (x-\ell)(2-a) - a(x-\ell)^2$. Ainsi $u_{n+1} - \ell = (2-a)(u_n - \ell) + O(x-\ell)^2$ et la différence au point fixe est comparable à une suite géométrique de raison $(2-a)$. Donc pour $1 < a < 3$, si la suite est assez proche du point fixe ℓ, elle va converger vers ℓ car $-1 < 2-a < 1$.
Par contre, à partir de $a = 3$, le point fixe est répulsif car un petit écart au point fixe s'accroît d'un facteur proche de $|2-a| > 1$. La suite ne converge plus mais ses sous-suites paires et impaires, vérifiant $v_n = u_{2n} = f_a \circ f_a(v_{n-1})$ et $w_n = u_{2n+1} = f_a \circ f_a(w_{n-1})$. Quand vont-elles converger ? tout d'abord, elles ont, outre les points fixes 0 et $\ell = 1 - \frac{1}{a}$ de f_a, deux autres points fixes, $\frac{a+1 \pm \sqrt{(a-3)(a+1)}}{2a}$ et la dérivée de $f_a \circ f_a$ y vaut $1-(a-3)(a+1)$. Cette valeur est comprise entre -1 et 1 pour $3 < a < 1+\sqrt{6}$ à partir duquel les suites paires et impaires se dédoublent encore, et se dédoubleront encore et encore. Une applet permettant de modifier les paramètres.
Cette histoire est complexe mais pointe du doigt l'apparition du chaos dans des itérations pourtant très simples, ce qui se traduit, pour le modèle continu, dans la sensibilité aux conditions initiales.

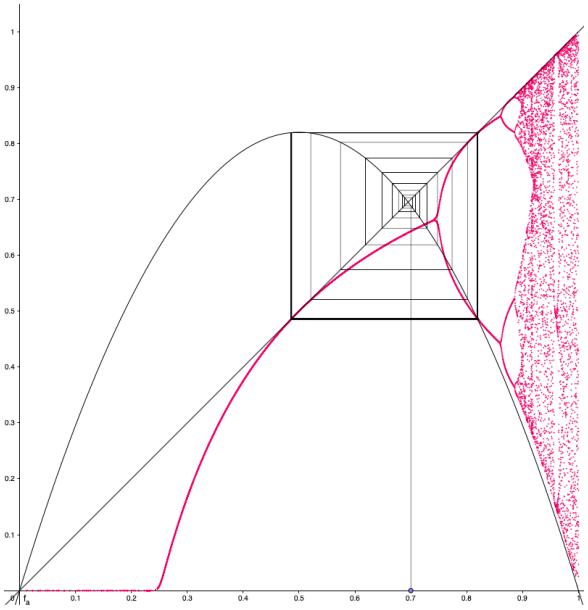


FIGURE 10.5 – Une suite $u_{n+1} = f_a(u_n)$ pour $a \approx 3,36$, et (en rouge) les points d'accumulation $(\frac{a}{4}, u_n)$ pour n grands et différentes valeurs de a.

Équation de Lotka-Volterra

Cette équation non linéaire d'ordre 1 en dimension 2 modélise des systèmes biologiques de deux populations (N_1, N_2) en interaction. Chaque population suit une équation différentielle d'ordre 1 comprenant différents termes. Un terme $\epsilon_i N_i$ de reproduction propre linéaire, qui conduirait à une explosion exponentielle ou au contraire à une extinction suivant le signe de ϵ_i. L'impact de l'environnement, comme précédemment dans le modèle logistique, est un terme non linéaire $-\lambda_i N_i^2$ proportionnel au carré de la population, qui mènerait à un point d'équilibre $N_i = \frac{\epsilon_i}{\lambda_i}$ sans présence de l'autre espèce ($N_i = 0$ étant toujours solution). Les choses se compliquent avec un terme croisé, $\gamma_j N_j N_i$ et suivant le signe du coefficient γ_i, la présence de l'autre population N_j est favorable (si $\gamma_j > 0$) ou défavorable.

$$\begin{cases} \frac{dN_1}{dt} &= (\epsilon_1 - \lambda_1 N_1 + \gamma_2 N_2) \times N_1, \\ \frac{dN_2}{dt} &= (\epsilon_2 - \lambda_2 N_2 + \gamma_1 N_1) \times N_2. \end{cases}$$

Suivant les différents signes des divers coefficients, on a affaire à des relations écologiques variées :

- $\epsilon_1 > 0$, $\gamma_2 < 0$, $\gamma_1 > 0$ signifie que N_1 est une *proie* pour le *prédateur* N_2 qui s'en nourrit. Quand en plus $\epsilon_2 < 0$, c'est une relation de *parasitisme* car N_2 ne peut pas survivre tout seul.
- $\epsilon_i > 0$, $\gamma_i < 0$, c'est une relation de *compétition* simple, la présence d'une espèce est au détriment de l'autre.
- Si $\gamma_1 = 0, \gamma_2 > 0$, c'est une relation de *commensalisme*, la présence d'une espèce étant indifférente pour l'une et bénéfique pour l'autre. Pour $\gamma_2 < 0$, on parle d'*amensalisme*.
- $\gamma_i > 0$ modélise une *symbiose* ou un *mutualisme*, la présence d'une espèce est bénéfique à l'autre.

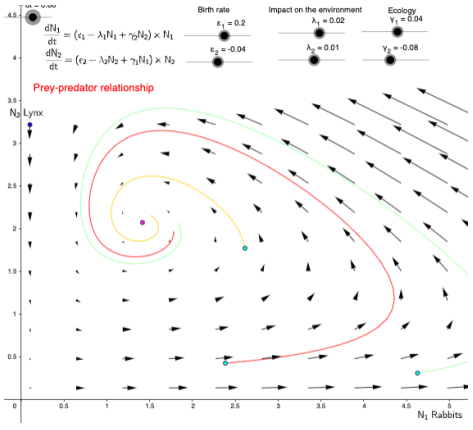


FIGURE 10.6 – Relation proie-prédateur.

Un tel système a potentiellement un point d'équilibre, quand le vecteur $\begin{pmatrix} N_1 \\ N_2 \end{pmatrix}$ vérifie

$$\begin{pmatrix} \lambda_1 & -\gamma_2 \\ -\gamma_1 & \lambda_2 \end{pmatrix} \begin{pmatrix} N_1 \\ N_2 \end{pmatrix} = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \end{pmatrix}.$$

Ceci nécessite que $\lambda_1\lambda_2 \neq \gamma_1\gamma_2$ et qu'ensuite les coordonnées du vecteur $\begin{pmatrix} \lambda_1 & -\gamma_2 \\ -\gamma_1 & \lambda_2 \end{pmatrix}^{-1} \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \end{pmatrix} = \begin{pmatrix} N_1 \\ N_2 \end{pmatrix}$ soient positifs pour avoir un sens dans le modèle.

On peut intégrer numériquement ce modèle à l'aide d'une applet permettant de modifier les paramètres. Plus sur le sujet proie-prédateur dans la thèse d'Hugues Santin Janin UCBL 2010.

Modèle SIR

Lors d'une épidémie, on divise la population en trois classes (ou compartiments), S les susceptibles d'être infectés, I les infectés et R les personnes guéries. On fait les hypothèses que le taux d'infection est proportionnel au nombre d'infectés et aux nombre de susceptible. Les nouveaux infectés sont retranchés des susceptibles S et ajoutés aux infectés I. On suppose aussi que la guérison se fait avec un taux fixe, proportionnel seulement au nombre d'infectés. De même, les nouveaux guéris sont retranchés des infectés I et ajoutés à R. Le taux d'infection est relié à R_0, le nombre de reproduction de base, qu'on supposera constant (ce n'est pas vrai dans une vraie épidémie, l'application de mesures d'hygiène et de confinement fait baisser ce taux). Le taux de guérison est également dépendant des décisions de santé publique. On aboutit ainsi à l'équation différentielle suivante dans $\mathbb{R}^3$:

$$\begin{cases} S' &= -p \times S \times I \\ I' &= +p \times S \times I - \alpha I \\ R' &= \alpha I \end{cases}$$

C'est une équation différentielle d'ordre 1, vectorielle, non linéaire. Remarquez la conservation $S + I + R$.

On peut l'intégrer numériquement à l'aide d'une applet permettant de modifier les paramètres. On peut complexifier le modèle en introduisant la vaccination, qui fait passer des cas Susceptibles dans le compartiment Remis, ou au contraire, une immunité imparfaite, des cas Remis devenant de nouveau Susceptibles. On a alors une endémie qui se met en place, avec parfois des bouffées épidémiques comme la grippe saisonnière.

Un article sur le site Images des Mathématiques.

Index