

Mathématiques pour ingénieurs

Khalid Koufany

ENSEM 1A NRJ et ISN

2022/2023

Table des matières

Chapitre 1. Rappels d'analyse	3
1. Produit scalaire, norme et distance	3
2. Petit vocabulaire de topologie	6
3. Convexe, connexe, étoilé	7
4. Fonctions de \mathbb{R}^n dans \mathbb{R}^m	8
5. Limite et continuité de champs scalaires et vectoriels	8
6. Quelques notations	11
7. Suites de fonctions	11
Chapitre 2. Calcul différentiel	15
1. Notations	16
2. Différentiabilité et différentielle	16
3. Premières propriétés	18
4. Différentiabilité des fonctions composées	18
5. Dérivée par rapport à un vecteur	19
6. Dérivée partielle	21
7. Calcul des différentielles	23
8. Réécriture du Théorème des fonctions composées	26
9. Dérivées partielles d'ordre supérieur	29
10. Exercices	31
Chapitre 3. Généralités et étude théorique des problèmes d'optimisation	33
1. Maximum, minimum, point-selle	33
2. Rappels sur les matrices réelles symétriques	34
3. Hessienne, Formule de Taylor au second ordre	35
4. Nature des points stationnaires	35
5. Résultats d'existence	37
6. Convexité	38
7. Conditions nécessaires ou suffisantes d'extremum	40
8. Optimisation sous contraintes	42
9. Rôle des multiplicateurs de Lagrange	44
10. Exercices	45
Chapitre 4. Opérateurs différentiels de la physique	49
1. Définition et résultats	49

2. Exemples	50
Chapitre 5. Intégrales sur \mathbb{R}	53
1. Interverision Limite-Intégrale	53
2. Interverision Série-Limite	55
3. Interverision Dérivation-Limite	55
4. Intégrales à paramètres	55
Chapitre 6. Intégrales multiples I	59
1. Introduction	59
2. Intégrabilité des fonctions continues	59
3. Intégrales doubles étendues à des régions plus générales	60
4. Interprétation géométrique des intégrales doubles	62
Chapitre 7. Intégrales curvilignes	63
1. Introduction	63
2. Rappels sur les courbes	63
3. Intégrales curvilignes	65
4. Propriétés de base	66
5. Intégrales de ligne par rapport à la longueur d'arc	67
6. Indépendance par rapport au chemin dans les ensembles ouverts connexes	68
7. Circulation d'un champ de gradient	68
Chapitre 8. Intégrales multiples II	71
1. Introduction	71
2. Le théorème de Green dans le plan	71
3. Calcul des aires planes	73
4. Calcul d'aire et Théorème de Green en coordonnées polaires	74
5. Changements de variables dans une intégrale double	75
6. Extension à des dimensions supérieures	76
Chapitre 9. Intégrales de surface et théorème de Stokes	79
1. Représentation paramétrique d'une surface	79
2. Plan tangent et normale à une surface	81
3. Aire d'une surface paramétrique	83
4. Intégrales de surface	84
5. Flux de champ à travers une surface	85
6. Le théorème de Stokes	86
7. Reconstruction d'un champ de vecteurs à partir de son rotationnel	90
8. Le théorème de la divergence (théorème de Gauss ou d'Ostrogradski)	90
Chapitre 10. Introduction et position du problème	93
Chapitre 11. Etude générale	97
1. Existence de la solution : Le théorème de Peano	97

2.	Solution maximale	100
3.	Unicité de la solution : Le théorème de Cauchy-Lipschitz	101
4.	Stabilité de la solution	103
5.	Représentations graphiques	104
Chapitre 12. Rappels d'algèbre		107
1.	Espace vectoriel de dimension finie	107
2.	Matrice d'une transformation linéaire	107
3.	Calcul matriciel	108
4.	Vecteurs propres	110
5.	Diagonalisation	111
6.	Réduction de Jordan	112
7.	Trigonalisation	113
8.	Résolution de systèmes linéaires	114
9.	Exercices	114
Chapitre 13. Les équations différentielles linéaires		117
1.	Systèmes différentiels homogènes	118
2.	Résolvante d'un Système différentiel	119
3.	Matrice de Wronski	121
4.	Le cas des équations différentielles linéaires à coefficients constants	121
Chapitre 14. Espaces L^1, L^2 et convergences		129
1.	La notion de 'presque-partout'	129
2.	L'espace $L^1(I)$	129
3.	L'espace $L^2(I)$	130
4.	Convergence de de suites fonctions et propriétés	131
Chapitre 15. La convolution		133
1.	Introduction	133
2.	Définition	133
3.	Quelques exemples	134
4.	Interprétation graphique de la convolution	135
5.	Exercices	136
Chapitre 16. Compléments sur les espaces de Hilbert. Séries de Fourier		139
1.	Compléments sur les espaces de Hilbert	139
2.	L'espace ℓ^2	142
3.	L'espace $L^2(]0; a[)$	143
4.	L'espace $L^2(\mathbb{R})$	144
5.	Approximation par des séries de Fourier	144
6.	Exercices	151
Chapitre 17. La transformée de Laplace		155

1. Intégrale de Laplace	155
2. Hypothèses sur f	155
3. Premières propriétés de la transformée de Laplace	156
4. Transformées de fonctions usuelles	159
5. Dérivation et intégration de la transformée de Laplace	161
6. La transformée de Laplace inverse	163
7. Convolution et produit de transformées de Laplace	167
8. Convolution et équations différentielles	168
9. Table	171
Chapitre 18. La transformée de Fourier	173
1. Transformée de Fourier d'une fonction L^1	173
2. Propriétés de la transformée de Fourier	177
3. Transformée de Fourier d'une fonction L^2	182
4. Transformée de Fourier et convolution	186
5. Exercices	188
Bibliographie	191

Chapitre 1

Rappels d'analyse

Nous supposons connues les notions élémentaires concernant les espaces vectoriels (réels ou complexes) et les applications linéaires. Ainsi \mathbb{R}^n est un espace vectoriel réel de dimension n ; sa *base canonique* sera notée (e_1, \dots, e_n) où

$$e_1 = (1, 0, \dots, 0), \dots, e_n = (0, 0, \dots, 0, 1).$$

Les vecteurs $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ seront considérés comme des *matrices colonnes*; les nombres réels x_1, \dots, x_n sont les coordonnées de \mathbf{x} dans la base canonique. Une matrice (m, n) (ou de type (m, n)) sera notée

$$A = [a_{ij}]_{1 \leq i \leq m, 1 \leq j \leq n} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}.$$

Ainsi $\mathbf{x} \in \mathbb{R}^n$ sera considéré comme une matrice du type $(n, 1)$. On note A^\top la *transposée* de A ,

$$A^\top = [b_{ij}], \quad b_{ij} = a_{ji}$$

Avec ces notations, si $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$, alors

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{x}^\top = [x_1 \quad \dots \quad x_n],$$

donc selon les règles usuelles de la multiplication des matrices on a

$$\mathbf{x}^\top \mathbf{x} = x_1^2 + \dots + x_n^2, \quad \mathbf{x} \mathbf{x}^\top = \begin{bmatrix} x_1^2 & x_1 x_2 & \dots & x_1 x_n \\ x_2 x_1 & x_2^2 & \dots & x_2 x_n \\ \vdots & & \ddots & \vdots \\ x_n x_1 & x_n x_2 & \dots & x_n^2 \end{bmatrix}.$$

1. Produit scalaire, norme et distance

Afin d'introduire la structure topologique sur \mathbb{R}^n nous avons besoin de quelques notions fondamentales comme le produit scalaire, la norme et la distance.

DÉFINITION 1.1. Soit E un espace vectoriel sur \mathbb{R} . Un produit scalaire sur E est une application

$$\begin{aligned} E \times E &\mapsto \mathbb{R} \\ (\mathbf{x}, \mathbf{y}) &\mapsto \langle \mathbf{x}, \mathbf{y} \rangle \end{aligned}$$

avec les propriétés suivantes :

1. *symétrique* : $\forall \mathbf{x}, \mathbf{y} \in E, \quad \langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$;
2. *bilinéaire* : $\forall \alpha \in \mathbb{R}, \forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in E, \quad \langle \alpha \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \alpha \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$; (par symétrie on a aussi la linéarité par rapport à la deuxième place) ;
3. *définie positive* : $\forall \mathbf{x} \in E, \quad \langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ et si $\langle \mathbf{x}, \mathbf{x} \rangle = 0$, alors $\mathbf{x} = \mathbf{0}$.

Un espace muni d'un produit scalaire $\langle \cdot, \cdot \rangle$ s'appelle un espace pré-hilbertien. Un espace pré-hilbertien réel de dimension finie est appelé **espace euclidien**.

DÉFINITION 1.2. On dit que deux vecteurs \mathbf{x} et \mathbf{y} d'un espace pré-hilbertien sont **orthogonaux** (que l'on note souvent $\mathbf{x} \perp \mathbf{y}$) si $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.

Le produit scalaire usuel sur \mathbb{R}^n est défini par

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n x_i y_i$$

où $\mathbf{x} = (x_1, x_2, \dots, x_n)$, $\mathbf{y} = (y_1, y_2, \dots, y_n)$.

DÉFINITION 1.3. Une norme sur E est une application

$$\begin{aligned} E &\mapsto \mathbb{R}_+ \\ x &\mapsto \|\mathbf{x}\| \end{aligned}$$

telle que

1. $\|\mathbf{x}\| = 0$ seulement si $\mathbf{x} = \mathbf{0}$
2. $\|\lambda \mathbf{x}\| = |\lambda| \|\mathbf{x}\|$, $\mathbf{x} \in E, \lambda \in \mathbb{R}$
3. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$, $\mathbf{x} \in E, \mathbf{y} \in E$

Un espace muni d'une norme $\|\cdot\|$ s'appelle espace normé.

REMARQUE 1.1. Un espace pré-hilbertien est toujours normé avec la norme $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$. Une norme générée par un produit scalaire sur un espace de dimension finie est appelé norme euclidienne.

THÉORÈME 1.1. (**inégalité de Cauchy-Schwarz**) Soit E un espace pré-hilbertien, $(\mathbf{x}, \mathbf{y}) \in E \times E$, alors

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \|\mathbf{y}\|$$

avec égalité si et seulement si \mathbf{x} et \mathbf{y} sont colinéaires ($\exists \lambda \in \mathbb{R}$ tel que $\mathbf{x} = \lambda \mathbf{y}$).

THÉORÈME 1.2. (**inégalité de Minkowski**) Soit E un espace pré-hilbertien. Alors, $\forall (\mathbf{x}, \mathbf{y}) \in E$

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$$

avec égalité si et seulement si \mathbf{x} et \mathbf{y} sont positivement colinéaires ($\exists \lambda \in \mathbb{R}^+$ tel que $\mathbf{x} = \lambda \mathbf{y}$).

THÉORÈME 1.3. (égalité du parallélogramme) Si E est un espace pré-hilbertien, alors, $\forall (\mathbf{x}, \mathbf{y}) \in E \times E$, nous avons

$$\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2$$

Celle-ci est appelée égalité du parallélogramme car elle traduit le fait que, pour une norme euclidienne dans le plan, la somme des carrés des longueurs des diagonales d'un parallélogramme est égale à la somme des carrés des longueurs des quatre côtés.

Les normes les plus utilisés sur \mathbb{R}^n sont :

$$\begin{aligned} \|(x_1, x_2, \dots, x_n)\|_1 &= \sum_{i=1}^n |x_i| \\ \|(x_1, x_2, \dots, x_n)\|_2 &= \sqrt{\sum_{i=1}^n x_i^2} \\ \|(x_1, x_2, \dots, x_n)\|_\infty &= \max_{1 \leq i \leq n} |x_i| \end{aligned}$$

EXERCICE 1.1. Montrez que $(x_1, x_2) \mapsto \sup_{t \in \mathbb{R}} \frac{|x_1 + tx_2|}{1 + t + t^2}$ est une norme sur \mathbb{R}^2 .

EXERCICE 1.2. [Norme de Frobenius] Pour $A \in \mathcal{M}_{m,n}(\mathbb{R})$, on pose $\langle A, B \rangle = \text{Tr}(A^\top B)$ et $\|A\| = \sqrt{\text{Tr}(A^\top A)}$.

Montrer que l'on définit un produit scalaire et une norme.

DÉFINITION 1.4. On appelle distance sur un espace vectoriel E une application

$$\begin{aligned} E \times E &\mapsto \mathbb{R}_+ \\ (\mathbf{x}, \mathbf{y}) &\mapsto d(\mathbf{x}, \mathbf{y}) \end{aligned}$$

avec les propriétés suivantes :

1. $\forall \mathbf{x}, \mathbf{y} \in E$, $d(\mathbf{x}, \mathbf{y}) \geq 0$ avec égalité seulement si $\mathbf{x} = \mathbf{y}$;
2. $\forall \mathbf{x}, \mathbf{y} \in E$, $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$;
3. $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in E$, $d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z}) \geq d(\mathbf{x}, \mathbf{z})$.

Un espace vectoriel muni d'une distance est un **espace métrique**.

EXEMPLE 1.1. Chaque norme génère une distance canonique définie par

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$$

EXEMPLE 1.2. Sur \mathbb{R}^n les distances le plus utilisés sont :

$$\begin{aligned} d_1(\mathbf{x}, \mathbf{y}) &= \|\mathbf{x} - \mathbf{y}\|_1 = \sum_{i=1}^n |x_i - y_i| \\ d_2(\mathbf{x}, \mathbf{y}) &= \|\mathbf{x} - \mathbf{y}\|_2 = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \end{aligned}$$

$$d_\infty(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_\infty = \max_{1 \leq i \leq n} |x_i - y_i|$$

REMARQUE 1.2. La distance $d_2(\cdot, \cdot)$ et la norme $\|\cdot\|_2$ sont souvent appelées la distance euclidienne et la norme euclidienne respectivement.

DÉFINITION 1.5. Soit $(\mathbf{x}_n)_{n \in \mathbb{N}}$ une suite de termes d'un espace E , muni d'une distance $d(\cdot, \cdot)$. On dit que $(\mathbf{x}_n)_{n \in \mathbb{N}}$ est une **suite de Cauchy** si :

$$\forall \varepsilon > 0, \quad \exists N_0 \in \mathbb{N}, \text{ tel que } (\forall (p, q) \in \mathbb{N}^2, p \geq N_0, q \geq N_0), \quad \text{on a } d(\mathbf{x}_p, \mathbf{x}_q) < \varepsilon$$

Autrement dit, on peut toujours trouver un certain rang à partir duquel la distance entre deux termes de la suite peut être rendue aussi petite que l'on veut.

DÉFINITION 1.6. On dit qu'un espace métrique E est **complet** si toute suite de Cauchy converge, vers un élément de E .

EXEMPLE 1.3. \mathbb{R} est complet, mais \mathbb{Q} n'est pas complet : puisque \mathbb{Q} est dense dans \mathbb{R} , il existe une suite (r_n) de \mathbb{Q} telle que

$$\sqrt{2} < r_n < \sqrt{2} + \frac{1}{n}, \quad \forall n \in \mathbb{N}^*$$

Le réel $\sqrt{2}$ est limite de la suite (r_n) . Ainsi (r_n) est une suite de Cauchy dans \mathbb{Q} mais ne converge pas vers un élément de \mathbb{Q} .

DÉFINITION 1.7. Un espace normé complet est appelé **espace de Banach**. Un espace pré-Hilbertien **complet** est appelé **espace de Hilbert**.

Ces derniers types d'espaces jouent des rôles fondamentaux en physique et ingénierie car ils permettent de définir de manière naturelle la notion de convergence. Les espaces de Hilbert sont plus intéressants que les espaces pré-hilbertiens, puisque en plus d'être normés, donc métriques, dans lesquels nous pouvons parler de limite, continuité, convergence, on a que toute suite de Cauchy est nécessairement convergente ! Rappelons que, en pratique, la notion de convergence nécessite, *a priori*, de connaître cette limite, ce qui n'est pas toujours le cas. Les suites de Cauchy sont un passage intéressant pour démontrer ce résultat et consistent à regarder la convergence en tant que comportement asymptotique des points les uns par rapport aux autres, d'où l'importance de la notion de "complétude".

Nous reviendrons aux espaces de Hilbert lors de l'étude des séries de Fourier

2. Petit vocabulaire de topologie

Soit \mathbf{x}_0 un point de \mathbb{R}^n et $r > 0$ un nombre réel donné. L'ensemble des points \mathbf{x} de \mathbb{R}^n tels que : $\|\mathbf{x} - \mathbf{x}_0\| < r$, est appelé une **n -boule ouverte** de rayon r et de centre \mathbf{x}_0 . On la note $\mathcal{B}(\mathbf{x}_0; r)$. Un exemple est donné en dimension 1 par un intervalle ouvert de centre \mathbf{x}_0 . Dans \mathbb{R}^2 , nous retrouvons le disque circulaire ouvert de centre \mathbf{x}_0 et de rayon r . Dans \mathbb{R}^3 , c'est la boule usuelle ouverte de centre \mathbf{x}_0 et de rayon r .

EXERCICE 1.3. Dessiner les boules unité associées aux distances $d_1(\cdot, \cdot)$, $d_2(\cdot, \cdot)$, $d_3(\cdot, \cdot)$.

DÉFINITION 1.8. Soit \mathcal{S} un sous-ensemble de \mathbb{R}^n et soit $\mathbf{x}_0 \in \mathcal{S}$. Alors, \mathbf{x}_0 est appelé un **point intérieur** de \mathcal{S} si il existe une n -boule ouverte de centre \mathbf{x}_0 incluse dans \mathcal{S} . L'ensemble de tous les points intérieurs de \mathcal{S} est appelé l'**intérieur** de \mathcal{S} et est noté $\text{int}\mathcal{S}$.

DÉFINITION 1.9. Un ensemble \mathcal{S} de \mathbb{R}^n est appelé **ouvert** si tous ses points sont des points intérieurs. En d'autres termes, si et seulement si $\mathcal{S} = \text{int}\mathcal{S}$.

Un ouvert contenant un point \mathbf{x}_0 est appelé un **voisinage** de \mathbf{x}_0 .

EXEMPLE 1.4. Dans \mathbb{R} , $]a, b[$ est un ouvert de \mathbb{R} , mais ce même intervalle $]a, b[$, plongé dans \mathbb{R}^2 n'est pas un ouvert de \mathbb{R}^2 . $[a, b]$ n'est pas un ouvert de \mathbb{R} . Un disque ouvert, ou un rectangle du type $]a, b[\times]c, d[$, est un ouvert de \mathbb{R}^2 .

Introduisons maintenant la notion d'extérieur et de frontière.

DÉFINITION 1.10. Un point \mathbf{x} est **extérieur** à un ensemble \mathcal{S} dans \mathbb{R}^n si il existe une n -boule $\mathcal{B}(\mathbf{x})$ ne contenant aucun point de \mathcal{S} . L'ensemble de tous les points dans \mathbb{R}^n extérieurs à \mathcal{S} est appelé l'**extérieur** de \mathcal{S} et est noté $\text{ext}\mathcal{S}$. Un point qui n'est ni dans l'extérieur ou l'intérieur de \mathcal{S} est appelé un **point frontière** de \mathcal{S} . L'ensemble des points frontière de \mathcal{S} forme la **frontière** de \mathcal{S} , et est notée $\partial\mathcal{S}$. Un ensemble \mathcal{S} de \mathbb{R}^n est dit **fermé** si son complémentaire dans \mathbb{R}^n (noté $-\mathcal{S}$ ou encore \mathcal{S}^c) est ouvert.

DÉFINITION 1.11. Un ensemble \mathcal{S} de \mathbb{R}^n est **borné** s'il existe un $r \in \mathbb{R}_+^*$ tel que $\mathcal{S} \subset \mathcal{B}(\mathbf{0}; r)$. Autrement dit $\forall \mathbf{x} \in \mathcal{S}$, $\|\mathbf{x}\| < r$ ou encore $\sup_{\mathbf{x} \in \mathcal{S}} \|\mathbf{x}\| < \infty$.

Un ensemble borné et fermé de \mathbb{R}^n est **compacte**.

EXERCICE 1.4. Montrer que :

(1) Si A_i , $i \in \mathbb{N}$ sont ouverts, alors $\bigcup_{i \in \mathbb{N}} A_i$ est ouvert.

(2) Si A_i , $i = 1, 2, \dots, n$ sont ouverts, alors $\bigcap_{i=1}^n A_i$ est ouvert.

EXERCICE 1.5. Soient A, B deux parties non vides d'un espace métrique E . On note $A + B = \{\mathbf{x} + \mathbf{y} \text{ t.q. } \mathbf{x} \in A, \mathbf{y} \in B\}$. Montrer que :

(1) Si A ou B est ouvert, alors $A + B$ est ouvert.

(2) Si A et B sont fermés, alors $A + B$ n'est pas nécessairement fermé. (Prendre $A = \{(x, y) \in \mathbb{R}^2 \text{ t.q. } xy = 1\}$ et $B = \{(x, 0) \text{ t.q. } x \in \mathbb{R}\}$)

3. Convexe, connexe, étoilé

DÉFINITION 1.12. Un ensemble $\mathcal{S} \subset \mathbb{R}^n$ est **convexe** si, pour tout $\mathbf{a}, \mathbf{b} \in \mathcal{S}$, le segment $[\mathbf{a}, \mathbf{b}]$ est inclus dans \mathcal{S} , c-à-d

$$\forall \mathbf{a}, \mathbf{b} \in \mathcal{S}, \forall t \in [0, 1], \quad t\mathbf{a} + (1-t)\mathbf{b} \in \mathcal{S}.$$

DÉFINITION 1.13. Un ensemble \mathcal{S} de \mathbb{R}^n est **étoilé** par rapport à un point \mathbf{P} de \mathcal{S} si, pour tout $\mathbf{M} \in \mathcal{S}$, le segment $[\mathbf{P}, \mathbf{M}]$ est inclus dans \mathcal{S} .

Un ensemble convexe est donc étoilé par rapport à un point.

DÉFINITION 1.14. Soit \mathcal{S} un ensemble ouvert de \mathbb{R}^n . L'ensemble \mathcal{S} est dit **connexe** si chaque couple de points de \mathcal{S} peut être joint par un chemin régulier par morceaux dont le graphe se trouve dans \mathcal{S} . Plus précisément, pour tous points \mathbf{a} et \mathbf{b} dans \mathcal{S} , il existe un chemin régulier par morceaux α défini sur $[a; b]$ tel que $\alpha(t) \in \mathcal{S}$, $\forall t \in [a; b]$, avec $\alpha(a) = \mathbf{a}$, $\alpha(b) = \mathbf{b}$. Un ensemble est dit **non connexe** si \mathcal{S} est la réunion de deux ou plus ensembles ouverts non vides disjoints.

REMARQUE 1.3. On a donc : connexe \Rightarrow convexe \Rightarrow étoilé.

4. Fonctions de \mathbb{R}^n dans \mathbb{R}^m

Nous considérons ici des fonctions

$$\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m.$$

Lorsque $m = n = 1$, il s'agit d'une fonction d'une variable réelle à valeurs réelles.

Lorsque $n \geq 1$ et $m = 1$, la fonction est appelée **champ scalaire**.

Lorsque $n \geq 1$ et $m > 1$, elle est appelée **champ de vecteurs**.

Notation : Si f est un champ scalaire défini en un point $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$, les notations $f(\mathbf{x})$ et $f(x_1, \dots, x_n)$ seront utilisées pour désigner la valeur de f en ce point particulier. Si \mathbf{f} est un champ de vecteurs, nous écrivons également $\mathbf{f}(\mathbf{x})$ ou $\mathbf{f}(x_1, \dots, x_n)$.

Nous allons nous intéresser ici à étendre les concepts supposés connus, de limite, continuité, et dérivée à des champs scalaires et vectoriels (en ce qui concerne la dérivée, nous renvoyons au chapitre de Calcul Différentiel).

5. Limite et continuité de champs scalaires et vectoriels

Nous formulons les concepts de limite et continuité dans le cas des champs vectoriels. Le cas des champs scalaires s'en déduit directement (un champ scalaire est un cas particulier de champ vectoriel).

Mais avant, rappelons les définitions de limite et continuité d'une fonction dans \mathbb{R} .

DÉFINITION 1.15. Soit f une fonction de \mathbb{R} dans \mathbb{R} .

(a) La fonction f admet une **limite** L en un point x_0 si

$$\forall \varepsilon > 0, \exists \eta > 0, \text{ tel que } |x - x_0| < \eta \text{ implique } |L - f(x)| < \varepsilon.$$

Nous noterons

$$\lim_{x \rightarrow x_0} f(x) = L.$$

(b) La fonction f est **continue** en x_0 si

$$\forall \varepsilon > 0, \exists \eta > 0, \text{ tel que } |x - x_0| < \eta \text{ implique } |f(x_0) - f(x)| < \varepsilon.$$

Considérons maintenant un champ vectoriel $\mathbf{f} : \mathcal{S} \rightarrow \mathbb{R}^m$, où \mathcal{S} est un sous-ensemble de \mathbb{R}^n .

DÉFINITION 1.16. (a) Soient $\mathbf{x}_0 \in \mathbb{R}^n$ et $\mathbf{L} \in \mathbb{R}^m$. \mathbf{f} a une limite \mathbf{L} en \mathbf{x}_0 si

$$\lim_{\|\mathbf{x} - \mathbf{x}_0\| \rightarrow 0} \|\mathbf{f}(\mathbf{x}) - \mathbf{L}\| = 0. \quad (5.1)$$

Nous notons

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \mathbf{L}, \quad (5.2)$$

Le symbole \lim dans l'équation (5.1) est la limite au sens usuel du calcul élémentaire :

$$\forall \varepsilon > 0, \exists \eta > 0, \text{ tel que } \|\mathbf{x} - \mathbf{x}_0\| < \eta \text{ implique } \|\mathbf{f}(\mathbf{x}) - \mathbf{L}\| < \varepsilon.$$

Dans cette définition, il n'est pas nécessaire que \mathbf{f} soit définie en \mathbf{x}_0 . Soit $\mathbf{h} = \mathbf{x} - \mathbf{x}_0$; alors, l'équation (5.1) devient

$$\lim_{\|\mathbf{h}\| \rightarrow 0} \|\mathbf{f}(\mathbf{x}_0 + \mathbf{h}) - \mathbf{L}\| = 0.$$

(b) \mathbf{f} est **continue** en \mathbf{x}_0 si \mathbf{f} est définie en \mathbf{x}_0 et si

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0).$$

EXEMPLE 1.5. Pour des points dans \mathbb{R}^2 , $\mathbf{x} = (x, y)$ et $\mathbf{x}_0 = (x_0, y_0)$. La relation (5.2) s'écrit

$$\lim_{(x,y) \rightarrow (x_0,y_0)} \mathbf{f}(x, y) = \mathbf{L}.$$

Pour des points dans \mathbb{R}^3 , $\mathbf{x} = (x, y, z)$ et $\mathbf{x}_0 = (x_0, y_0, z_0)$. La relation (5.2) s'écrit

$$\lim_{(x,y,z) \rightarrow (x_0,y_0,z_0)} \mathbf{f}(x, y, z) = \mathbf{L}.$$

EXEMPLE 1.6. Si un champ vectoriel \mathbf{f} a ses valeurs dans \mathbb{R}^m , chaque valeur $\mathbf{f}(\mathbf{x})$ de la fonction a m composantes et nous pouvons écrire

$$\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x})).$$

Les m champs scalaires f_1, \dots, f_m sont appelés **composantes** du champ de vecteur \mathbf{f} . On peut montrer que \mathbf{f} est continue en un point si et seulement si chaque composante f_k est continue en ce point.

DÉFINITION 1.17. Nous dirons que \mathbf{f} est **continue** sur un ensemble \mathcal{S} si \mathbf{f} est continue en tout point de \mathcal{S} . On le note $\mathbf{f} \in \mathcal{C}^0(\mathcal{S})$.

EXERCICE 1.6. Soit A une partie de \mathbb{R}^n non vide. Pour $\mathbf{x} \in \mathbb{R}^n$ on note : $d_A(\mathbf{x}) = \inf\{\|\mathbf{x} - \mathbf{y}\| \text{ t.q. } \mathbf{y} \in A\}$.

(1) Montrez que d_A est continue.

- (2) Soient deux parties de \mathbb{R}^n non vides A, B . Donner une condition équivalente à $d_A = d_B$.

Puisque ces définitions sont des extensions directes de celles établies dans le cas unidimensionnel, il n'est pas surprenant d'apprendre que beaucoup de propriétés familières de la limite et de la continuité peuvent aussi être étendues. Pour les champs scalaires, les théorèmes basiques concernant les limites et continuités de sommes, produits et quotients de champs scalaires peuvent être étendus directement. Pour les champs vectoriels, les quotients ne sont pas définis mais nous avons le théorème suivant.

THÉORÈME 1.4. *Si $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \mathbf{L}$ et $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \mathbf{g}(\mathbf{x}) = \mathbf{M}$, nous avons :*

- a) $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} (\mathbf{f} + \mathbf{g})(\mathbf{x}) = \mathbf{L} + \mathbf{M}$,
- b) $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} (\lambda \mathbf{f})(\mathbf{x}) = \lambda \mathbf{L}$, $\forall \lambda \in \mathbb{R}$,
- c) $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \langle \mathbf{f}(\mathbf{x}), \mathbf{g}(\mathbf{x}) \rangle = \langle \mathbf{L}, \mathbf{M} \rangle$,
- d) $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \|\mathbf{f}(\mathbf{x})\| = \|\mathbf{L}\|$.

Le théorème suivant indique que la continuité est conservée par composition.

THÉORÈME 1.5. (*continuité des fonctions composées*) *Soient \mathbf{f} et \mathbf{g} des fonctions telles que la fonction composée $\mathbf{f} \circ \mathbf{g}$ soit définie en \mathbf{x}_0 , où*

$$(\mathbf{f} \circ \mathbf{g})(\mathbf{x}) = \mathbf{f}(\mathbf{g}(\mathbf{x})).$$

Si \mathbf{g} est continue en \mathbf{x}_0 et si \mathbf{f} est continue en $\mathbf{g}(\mathbf{x}_0)$, alors la composée $\mathbf{f} \circ \mathbf{g}$ est continue en \mathbf{x}_0 .

EXEMPLE 1.7. Le théorème précédent implique la continuité des champs scalaires h , où $h(x, y)$ est donnée par les formules telles que

- i) $\sin(x^2 y)$
- ii) $\log(x^2 + y^2)$
- iii) $\frac{\exp(x + y)}{x + y}$
- iv) $\log(\cos(x^2 + y^2))$.

La première fonction est continue en tous les points du plan, la seconde en tous les points en dehors de l'origine, la troisième en tous les points tels que $x + y \neq 0$, et, enfin, la quatrième en tous les points tels que $x^2 + y^2$ n'est pas un multiple impair de $\pi/2$. Plus précisément, ce dernier ensemble correspond aux points (x, y) tels que

$$x^2 + y^2 = \frac{\ell\pi}{2}, \ell = 1, 3, 5, \dots$$

C'est une famille de cercles centrés à l'origine. Ces exemples montrent notamment que l'ensemble des discontinuités d'une fonction de deux variables peut être un ou des points isolés, des courbes entières ou des familles de courbes.

EXEMPLE 1.8. Une fonction de deux variables peut être continue en chacune des variables séparément et être discontinue comme une fonction de deux variables. Vous

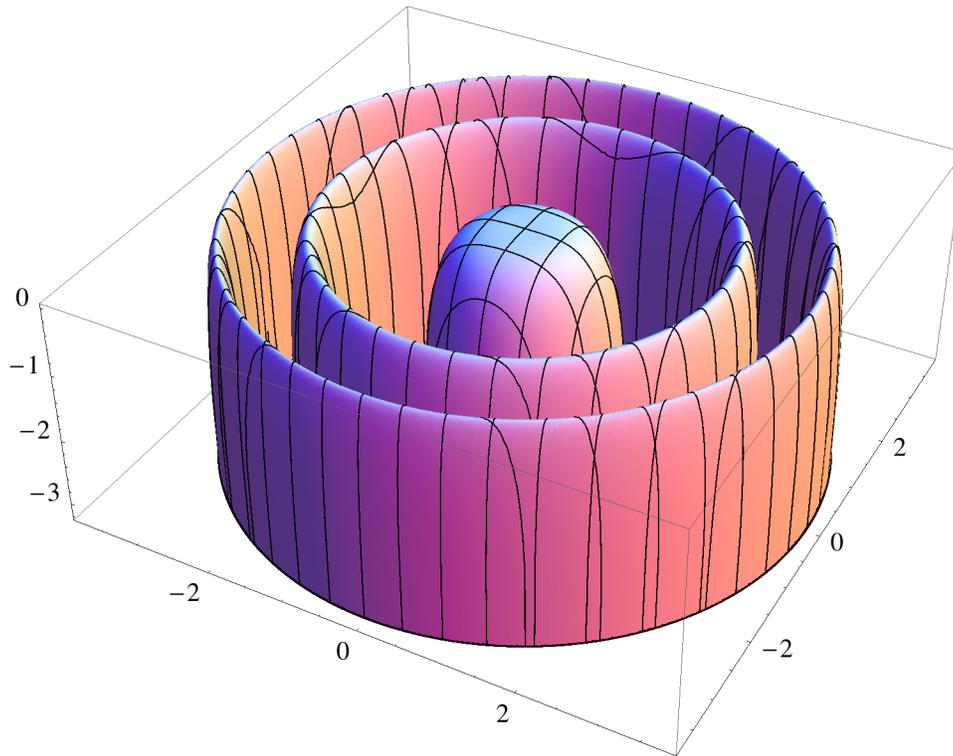


FIGURE 1 – Graphe de l'exemple iv)

pouvez considérer à titre d'exemple la fonction définie par

$$f(x, y) = \frac{xy}{x^2 + y^2}, \text{ si } (x, y) \neq (0, 0),$$

$$f(0, 0) = 0.$$

6. Quelques notations

Soit $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$, et $\mathbf{h} \in \mathbb{R}^n$.

Dire que $\mathbf{f}(\mathbf{h}) = o(\mathbf{h})$ signifie que $\lim_{\mathbf{h} \rightarrow \mathbf{0}} \mathbf{f}(\mathbf{h}) / \|\mathbf{h}\| = \mathbf{0}$.

Dire que $\mathbf{f}(\mathbf{h}) = O(\mathbf{h})$ signifie qu'il existe $C > 0$ et $\delta > 0$ tels que si $\|\mathbf{h}\| < \delta$, alors $\|\mathbf{f}(\mathbf{h})\| < C\|\mathbf{h}\|$, autrement dit, \mathbf{f} est au plus de l'ordre de $\|\mathbf{h}\|$ lorsque $\mathbf{h} \rightarrow \mathbf{0}$. Attention : un $o(\mathbf{h})$ est toujours un $O(\mathbf{h})$, mais la réciproque est fautive.

7. Suites de fonctions

DÉFINITION 1.18. Soient X un ensemble quelconque, E un espace métrique muni d'une distance d , et $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions définies sur X à valeurs dans E . On dit que la suite de fonctions $(f_n)_{n \in \mathbb{N}}$ converge simplement si : $\forall x \in X$, la suite $(f_n(x))_{n \in \mathbb{N}}$ converge dans E .

Autrement dit, une suite de fonctions $(f_n)_n$ converge simplement sur X vers une fonction f si et seulement si :

$$\forall x \in X, \forall \epsilon > 0, \exists N_{\epsilon, x}, \forall n \in \mathbb{N}, n \geq N_{\epsilon, x} \Rightarrow d(f_n(x), f(x)) < \epsilon$$

EXEMPLE 1.9. Considérons la suite de fonctions $(f_n)_n$ définie sur \mathbb{R}_+ par $f_n(x) = \frac{nx}{1+nx}$. On vérifie facilement que cette suite converge simplement vers la fonction définie sur \mathbb{R}_+ par

$$f(x) = \begin{cases} 1 & \text{si } x > 0 \\ 0 & \text{si } x = 0 \end{cases}$$

Pour $\epsilon \in]0; 1[$ donné et $x > 0$, on aura

$$|f_n(x) - f(x)| = \frac{1}{1+nx} < \epsilon$$

pour $nx > \frac{1-\epsilon}{\epsilon}$, soit pour $n \geq N_{\epsilon, x} = E(\frac{1-\epsilon}{\epsilon}) + 1$.

Supposons qu'il existe un entier N_ϵ indépendant de $x \in \mathbb{R}_+$ tel que $|f_n(x) - f(x)| < \epsilon$ pour tout $n \geq N_\epsilon$. On aura alors pour tout $x > 0$ et $n \geq N_\epsilon$, $\frac{1}{1+nx} < \epsilon$ et faisant tendre x vers 0 pour n fixé, on aboutit à $1 \leq \epsilon$, ce qui n'est pas pas.

Il est donc impossible de trouver un tel N_ϵ valable pour tout $x \in \mathbb{R}_+$ ou même pour tout $x > 0$. On dit dans ce cas que la convergence de la suite de fonction $(f_n)_n$ n'est pas uniforme sur \mathbb{R}_+ (ou \mathbb{R}_+^*).

L'exemple précédent nous conduit à la définition suivante.

DÉFINITION 1.19. Soient X un espace quelconque, (Y, d) , un espace métrique et $A \subset X$ un sous-ensemble de X . Soient $(f_n)_n$ une suite de fonctions définies sur X et à valeurs dans Y et f une fonction définie sur X à valeurs dans Y . On dit que la suite $(f_n)_n$ converge uniformément vers f sur A si :

$$\forall \epsilon > 0, \exists N_\epsilon \in \mathbb{N}, \forall n \in \mathbb{N}, [n \geq N_\epsilon \Rightarrow \forall x \in A, d(f_n(x), f(x)) \leq \epsilon] . \quad (7.1)$$

En introduisant la notation $d_{\infty, A}(f, g) = \sup_{x \in A} d(f(x), g(x))$, (dans laquelle le sup peut a priori être infini), la propriété (7.1) est équivalente à :

$$\forall \epsilon > 0, \exists N_\epsilon \in \mathbb{N}, \forall n \in \mathbb{N}, [n \geq N_\epsilon \Rightarrow d_{\infty, A}(f_n, f) \leq \epsilon] .$$

Autrement dit, $(f_n)_n$ converge uniformément vers f sur A si et seulement si

$$\lim_{n \rightarrow +\infty} d_{\infty, A}(f_n, f) = 0 .$$

REMARQUE 1.4. Avec les inégalités $d(f_n(x), f(x)) \leq \sup_{x \in A} d(f_n(x), f(x))$, on déduit que la convergence uniforme entraîne la convergence simple.

PROPOSITION 1.1. Si $(f_n)_n$ converge uniformément vers f sur A et $\forall n \in \mathbb{N}$ f_n continue sur A , alors f est continue sur A .

EXEMPLE 1.10. (a) Soit la suite de fonction $(f_n)_n$ définie sur $[0; 1]$ par $f_n(x) = \frac{nx}{1+nx}$.

On a pour tout $x \in]0; 1]$, $\lim_{n \rightarrow +\infty} f_n(x) = 1$ et pour $x = 0$, on a $f_n(x) = 0$. Donc la suite de fonctions $(f_n)_n$ converge simplement vers la fonction f donné par

$$f(x) = \begin{cases} 1 & \text{si } x \in]0; 1] \\ 0 & \text{si } x = 0 \end{cases}$$

Il ne peut y avoir convergence uniforme car la fonction limite f n'est pas continue sur $[0; 1]$, alors les fonctions f_n le sont.

(b) Soit la suite de fonction $(f_n)_n$ définie sur $[0; 1]$ par $f_n(x) = \frac{nx}{1+n^2x}$.

On a pour tout $x \in [0; 1]$, $\lim_{n \rightarrow +\infty} f_n(x) = \lim_{n \rightarrow +\infty} \frac{x}{\frac{1}{n} + nx^4} = 0$. Donc la suite de fonctions $(f_n)_n$ converge simplement vers la fonction nulle. Examinons maintenant la convergence uniforme et déterminons $\sup_{x \in [0; 1]} |f_n(x) - f(x)| = \sup_{x \in [0; 1]} f_n(x)$. Pour cela dérivons f_n (qui est une fonction dérivable sur $[0; 1]$). Pour tout $x \in [0; 1]$, on a $f'_n(x) = \frac{n-3n^3x^4}{(1+n^2x^4)^2}$. Donc

$$f'_n(x) = 0 \iff x = x_n = (3n^2)^{-\frac{1}{4}}.$$

La dérivée est positive sur $[0; x_n]$ et négative sur $[x_n; 1]$, donc f_n admet un maximum en x_n et par suite

$$\sup_{x \in [0; 1]} |f_n(x) - f(x)| = f_n(x_n) = \frac{3^{\frac{3}{4}} \sqrt{n}}{4}$$

qui ne tend pas vers 0 quand n tend vers $+\infty$. La suite de fonction $(f_n)_n$ ne converge donc pas uniformément vers la fonction nulle sur $[0; 1]$.

(c) Soit la suite de fonction $(f_n)_n$ définie sur \mathbb{R}_+^* par $f_n(x) = nxe^{-nx^2}$.

Pour $x > 0$, on a $\lim_{n \rightarrow +\infty} f_n(x) = \lim_{n \rightarrow +\infty} nxe^{-nx^2} = 0$. La suite de fonctions $(f_n) - N$ converge simplement vers la fonction nulle sur \mathbb{R}_+^* .

Pour la convergence uniforme, cherchons $\sup_{x \in]0; +\infty[} |f_n(x)|$. La fonction f_n étant dérivable sur \mathbb{R}_+ , nous avons

$$f'_n(x) = ne^{nx^2}(1 - 2nx^2)$$

$f'_n(x)$ s'annule pour $x_n = \frac{1}{\sqrt{2n}}$. La dérivée est positive sur $[0; x_n]$ et négative sur $[x_n; +\infty[$. Donc f_n est strictement croissante sur $[0; x_n]$ et strictement décroissante sur $[x_n; +\infty[$. La fonction f_n admet un maximum en x_n . Par conséquent

$$\sup_{x \in]0; +\infty[} |f_n(x)| = \sup_{x \in]0; +\infty[} f_n(x) = f_n(x_n) = \sqrt{\frac{n}{2}} e^{-\frac{1}{2}}$$

qui ne tend pas vers 0 quand n tend vers $+\infty$. La suite de fonction $(f_n)_n$ ne converge donc pas uniformément vers la fonction nulle sur \mathbb{R}_+^* .

(d) La suite de fonctions $(f_n)_n$ définie par $f_n(x) = \frac{\sin(nx)}{n}$ converge simplement vers la fonction nulle sur $[0; \pi]$. De plus

$$\sup_{x \in [0; \pi]} |f_n(x) - f(x)| = \sup_{x \in [0; \pi]} \left| \frac{\sin(nx)}{n} \right| = \frac{1}{n} \rightarrow 0.$$

Donc la suite de fonctions $(f_n)_n$ converge uniformément vers la fonction nulle sur $[0; \pi]$.

EXERCICE 1.7. Les suites $(f_n)_n$ suivantes convergent-elles simplement/uniformément ? vers quelles fonctions ?

a) $f_n : [0, 1] \mapsto \mathbb{R}$, $f_n(x) = nx(1-x)^n$;

b) $f_n : [0, 1] \mapsto \mathbb{R}$, $f_n(x) = \sqrt{x^2 + \frac{1}{n^2}}$;

EXERCICE 1.8. Montrer que la suite $f_n : \mathbb{R} \mapsto \mathbb{R}$, $f_n(x) = \frac{1}{\sqrt{n}} \cos 2nx$ converge uniformément mais la suite des dérivés ne converge pas simplement.

EXERCICE 1.9. Montrer que la suite $f_n : [0, \frac{\pi}{2}] \mapsto \mathbb{R}$, $f_1(x) = x$, $f_{n+1}(x) = \sin(f_n(x))$, $n \geq 1$ converge uniformément. Vers quelle fonction ?

Calcul différentiel

Les champs scalaires et vectoriels définis sur des sous-ensembles de \mathbb{R}^2 ou \mathbb{R}^3 (voire plus) apparaissent très souvent dans les sciences de l'ingénieur. En effet, dans de nombreux problèmes, on s'intéresse aux variations d'un champ \mathbf{f} . Dans le cas unidimensionnel, c'est la dérivée qui traduit cette idée :

$$f'(a) = \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h},$$

ou autrement dit

$$f(a+h) = f(a) + hf'(a) + o(h)$$

où $o(h)$ est un reste, qui tend vers 0 lorsque h tend vers 0. Au voisinage de a , la fonction f se comporte à peu près comme l'application affine : $x \mapsto f(a) + f'(a)(x-a)$ cf. [figure 1](#).

Nous généralisons cette idée dans le cas de \mathbb{R}^n . La variable est maintenant un **vecteur** de \mathbb{R}^n : plus question de diviser par h ! Et la dérivée s'appelle **différentielle**.

1. Notations

Soit U un ouvert de \mathbb{R}^n , on note

— un champ scalaire $f : U \rightarrow \mathbb{R}$,

— un champ vectoriel $\mathbf{f} = (f_1, \dots, f_m) : U \rightarrow \mathbb{R}^m$ où $m \geq 1$. Les $f_i : U \rightarrow \mathbb{R}$ sont des champs scalaires.

REMARQUE 2.1. Un champ scalaire est un cas particulier de champ vectoriel ($m = 1$). Dans la suite, nous donnerons souvent les définitions dans le cas d'un champ vectoriel \mathbf{f} .

2. Différentiabilité et différentielle

DÉFINITION 2.1. Soit \mathbf{f} une fonction définie sur U un ouvert de \mathbb{R}^n et à valeurs dans \mathbb{R}^m ($n, m \in \mathbb{N}$). La fonction \mathbf{f} est **différentiable** en $\mathbf{a} \in U$ s'il existe une **application linéaire** L de \mathbb{R}^n dans \mathbb{R}^m telle que

$$\mathbf{f}(\mathbf{a} + \mathbf{h}) = \mathbf{f}(\mathbf{a}) + L(\mathbf{h}) + o(\mathbf{h}). \quad (2.1)$$

THÉORÈME 2.1. (1) Si elle existe, l'**application linéaire** L définie plus haut est unique.

(2) Si \mathbf{f} est différentiable en \mathbf{a} , alors \mathbf{f} est continue en \mathbf{a} .

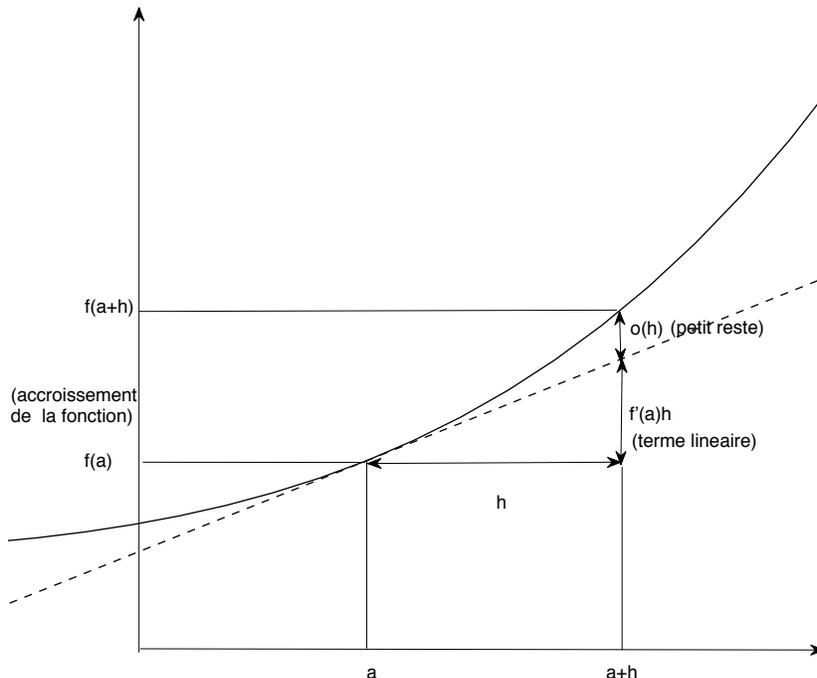


FIGURE 1 – Le b.a.-ba du calcul différentiel : $f(a+h) - f(a) = f'(a)h + o(h)$.

- (3) Un champ vectoriel $\mathbf{f} = (f_1, \dots, f_m)$ est différentiable si et seulement si ses composantes $f_i, i = 1, \dots, m$ sont des champs scalaires différentiables.

Notations. Quand \mathbf{f} est un champ vectoriel, L est parfois notée

$$L = D\mathbf{f}(\mathbf{a}), \quad \text{ou} \quad D_{\mathbf{a}}\mathbf{f}$$

Quand f est un champ scalaire, L est parfois notée

$$L = \nabla f(\mathbf{a})$$

On peut trouver d'autres notations : $f'(\mathbf{a})$, ou $df_{\mathbf{a}}$, ou $df(\mathbf{a})$

DÉFINITION 2.2. L'application linéaire $D\mathbf{f}(\mathbf{a})$ définie dans le théorème précédent est appelée **différentielle de \mathbf{f} en \mathbf{a}** .

DÉMONSTRATION. (du [Théorème 2.1](#))

- (1) Si L_1 et L_2 vérifient la définition, alors pour $r > 0$ assez petit, pour $\|\mathbf{h}\| < r$, on a :

$$\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) - L_1\mathbf{h} = o_1(\mathbf{h}) \text{ et } \mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) - L_2\mathbf{h} = o_2(\mathbf{h}), \text{ et donc}$$

$$L_1\mathbf{h} - L_2\mathbf{h} = (L_1 - L_2)\mathbf{h} = -o_1(\mathbf{h}) + o_2(\mathbf{h}) = o(\mathbf{h}).$$

Soit alors $\mathbf{x} \in \mathbb{R}^n$, quelconque. Et soit $t \in \mathbb{R}$ tel que $\|t\mathbf{x}\| < r$, alors $(L_1 - L_2)(t\mathbf{x}) = t(L_1 - L_2)\mathbf{x} = o(t\mathbf{x})$, et donc

$$(L_1 - L_2)\mathbf{x} = \frac{o(t\mathbf{x})}{t} = \frac{o(t\mathbf{x})\|\mathbf{x}\|}{t\|\mathbf{x}\|}.$$

Faisant ensuite tendre t vers 0, on a $(L_1 - L_2)\mathbf{x} = 0$.

Ceci étant valable pour tout $\mathbf{x} \in \mathbb{R}^n$, on a montré que $L_1 = L_2$ sur \mathbb{R}^n .

- (2) Si \mathbf{f} est différentiable en \mathbf{a} alors $\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) = L\mathbf{h} + o(\mathbf{h})$. Or, $\lim_{\|\mathbf{h}\| \rightarrow 0} L\mathbf{h} + o(\mathbf{h}) = 0$. En effet, L est une application linéaire de \mathbb{R}^n sur \mathbb{R}^m , elle est donc continue car on est en dimension finie¹. Donc $\lim_{\|\mathbf{h}\| \rightarrow 0} \mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) = 0$, ce qui prouve la continuité de \mathbf{f} en \mathbf{a} .
- (3) Il suffit d'écrire la formule de [Définition 2.1](#) et de regarder ses composantes. Les composantes de L sont bien sûr des applications linéaires.

□

En résumé :

la différentiabilité en \mathbf{a} d'une fonction \mathbf{f} de \mathbb{R}^n dans \mathbb{R}^m s'écrit :

$$\mathbf{f}(\mathbf{a} + \mathbf{h}) = \mathbf{f}(\mathbf{a}) + D\mathbf{f}(\mathbf{a})\mathbf{h} + o(\mathbf{h}) \quad (2.2)$$

où

- L'équation (2.2) est appelée **formule de Taylor** du premier ordre pour \mathbf{f} au point \mathbf{a} .
- $D\mathbf{f}(\mathbf{a})$ est une application linéaire de \mathbb{R}^n dans \mathbb{R}^m . C'est une approximation linéaire de $\mathbf{f}(\mathbf{a} + \cdot) - \mathbf{f}(\mathbf{a})$,
- et $o(\mathbf{h})$ est un reste qui tend vers $\mathbf{0}$ plus vite que $\|\mathbf{h}\| \rightarrow 0$ (i.e. $\frac{o(\mathbf{h})}{\|\mathbf{h}\|} \rightarrow \mathbf{0}$ si $\|\mathbf{h}\| \rightarrow 0$).

En pratique [Définition 2.1](#) et [Définition 2.2](#) sont théoriques, vous devez les connaître. Nous verrons plus loin une méthode calculatoire pour montrer la différentiabilité et calculer la différentielle d'une fonction.

REMARQUE 2.2. On retrouve encore l'approximation : au voisinage de \mathbf{a} , la fonction \mathbf{f} se comporte à peu près comme l'application affine $\mathbf{x} \mapsto \mathbf{f}(\mathbf{a}) + D\mathbf{f}(\mathbf{a})(\mathbf{x} - \mathbf{a})$, pour laquelle on pourra utiliser les outils de l'algèbre linéaire : calcul matriciel, rang, etc. Sous forme géométrique (voir [figure 1](#)), cela signifie qu'une courbe² est, au voisinage d'un point, à peu près confondue avec une droite (la *tangente*) ; ou encore qu'une surface est, au voisinage d'un point, à peu près confondue avec son *plan tangent*. D'ailleurs, $D\mathbf{f}(\mathbf{a})$ est parfois appelée *application linéaire tangente à \mathbf{f} en \mathbf{a}* .

1. Rappel : une application linéaire, en dimension finie, est continue
2. à condition que la fonction qui la représente soit différentiable

3. Premières propriétés

Linéarité : Si \mathbf{f} et \mathbf{g} sont différentiables en \mathbf{a} , et si le scalaire λ est constant, alors $\mathbf{f} + \mathbf{g}$, et $\lambda\mathbf{f}$ sont différentiables en \mathbf{a} et on a :

$$D(\mathbf{f} + \mathbf{g})(\mathbf{a}) = D\mathbf{f}(\mathbf{a}) + D\mathbf{g}(\mathbf{a}), \quad D(\lambda\mathbf{f})(\mathbf{a}) = \lambda D\mathbf{f}(\mathbf{a}).$$

Différentielle d'une constante : Une application $\mathbf{f} : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ constante est différentiable sur U , et pour tout $\mathbf{x} \in U$, on a :

$$D\mathbf{f}(\mathbf{x}) = \mathbf{0}.$$

En effet : $\mathbf{f}(\mathbf{x} + \mathbf{h}) = \mathbf{f}(\mathbf{x}) + \mathbf{0}\cdot\mathbf{h} + o(\mathbf{h})$, où $o(\mathbf{h}) = \mathbf{0}$.

Différentielle d'une application linéaire : Soit φ une application linéaire de \mathbb{R}^n dans \mathbb{R}^m , alors φ est différentiable sur U , et pour tout $\mathbf{x} \in \mathbb{R}^n$, on a :

$$D\varphi(\mathbf{x}) = \varphi.$$

En effet, $\varphi(\mathbf{x} + \mathbf{h}) = \varphi(\mathbf{x}) + \varphi(\mathbf{h}) = \varphi(\mathbf{x}) + D\varphi(\mathbf{x})(\mathbf{h}) + o(\mathbf{h})$, où $D\varphi(\mathbf{x})(\mathbf{h}) = \varphi(\mathbf{h})$ et $o(\mathbf{h}) = 0$.

REMARQUE 2.3. $D\varphi(\mathbf{x})$ est une application!! et $D\varphi(\mathbf{x})(\mathbf{h})$ est un vecteur de \mathbb{R}^m !!

La fonction Identité est une application linéaire. Sa différentielle est donc elle-même, c'est à dire la fonction Identité.

4. Différentiabilité des fonctions composées

Voici à présent un énoncé sur la différentielle des fonctions composées, sans lequel le calcul différentiel n'existerait pas.

Dans le cas unidimensionnel, on sait calculer la dérivée d'une fonction composée $\phi(t) = f(g(t))$ par la formule

$$\phi'(t) = f'(g(t))g'(t).$$

Nous allons étendre cette formule lorsque

- f est remplacée par un champ vectoriel $\mathbf{f} : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$
- g est remplacée par un champ vectoriel $\mathbf{g} : V \subset \mathbb{R}^p \rightarrow U \subset \mathbb{R}^n$.

THÉORÈME 2.2. Soient $\mathbf{f} : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ et $\mathbf{g} : V \subset \mathbb{R}^p \rightarrow \mathbb{R}^n$ des champs de vecteurs tels que la composition $\mathbf{f} \circ \mathbf{g}$ soit définie dans un voisinage d'un point $\mathbf{x}_0 \in V$. Supposons que \mathbf{g} soit différentiable en \mathbf{x}_0 , de différentielle $D\mathbf{g}(\mathbf{x}_0)$. Soit $\mathbf{y}_0 = \mathbf{g}(\mathbf{x}_0)$ et supposons que \mathbf{f} soit différentiable en \mathbf{y}_0 , de différentielle $D\mathbf{f}(\mathbf{y}_0)$. Alors, $\mathbf{f} \circ \mathbf{g}$ est différentiable en \mathbf{x}_0 , et la différentielle $D(\mathbf{f} \circ \mathbf{g})(\mathbf{x}_0)$ est donnée par

$$D(\mathbf{f} \circ \mathbf{g})(\mathbf{x}_0) = D\mathbf{f}(\mathbf{y}_0) \circ D\mathbf{g}(\mathbf{x}_0).$$

DÉMONSTRATION. On écrit

$$\begin{aligned} (\mathbf{f} \circ \mathbf{g})(\mathbf{x}_0 + \mathbf{h}) &= \mathbf{f}(\mathbf{g}(\mathbf{x}_0 + \mathbf{h})) = \mathbf{f}(\mathbf{g}(\mathbf{x}_0) + D\mathbf{g}(\mathbf{x}_0)\mathbf{h} + o(\mathbf{h})) \\ &= \mathbf{f}(\mathbf{y}_0 + \mathbf{H}) \text{ où } \mathbf{H} = D\mathbf{g}(\mathbf{x}_0)\mathbf{h} + o(\mathbf{h}) \\ &= \mathbf{f}(\mathbf{y}_0) + D\mathbf{f}(\mathbf{y}_0)\mathbf{H} + o(\mathbf{H}) \\ &= \mathbf{f}(\mathbf{y}_0) + D\mathbf{f}(\mathbf{y}_0) \circ D\mathbf{g}(\mathbf{x}_0)\mathbf{h} + D\mathbf{f}(\mathbf{y}_0)o(\mathbf{h}) + o(\mathbf{H}) \end{aligned}$$

Il reste à montrer que $D\mathbf{f}(\mathbf{y}_0)o(\mathbf{h}) + o(\mathbf{H}) = o(\mathbf{h})$. C'est en effet le cas. \square

5. Dérivée par rapport à un vecteur

Nous souhaitons étudier la manière dont varie une fonction (un champ scalaire ou un champ vectoriel) lorsque nous bougeons de \mathbf{a} vers un point proche. Par exemple, dans le cas d'un champ scalaire, supposons que $f(\mathbf{a})$ représente la température en un point \mathbf{a} donné dans une salle chauffée dont une fenêtre est ouverte. Si nous nous rapprochons de la fenêtre, la température tend à décroître, si nous nous rapprochons du chauffage, la température augmente. En général, la manière dont le champ change dépend de la direction selon laquelle nous nous dirigeons à partir de \mathbf{a} .

Supposons que nous spécifions cette direction par un second vecteur \mathbf{v} . Plus précisément, supposons que nous allions de \mathbf{a} vers $\mathbf{a} + \mathbf{v}$ le long de la ligne joignant \mathbf{a} et $\mathbf{a} + \mathbf{v}$. Chaque point de cette ligne est alors de la forme $\mathbf{a} + t\mathbf{v}$, où $t \in \mathbb{R}$. La distance de \mathbf{a} à $\mathbf{a} + t\mathbf{v}$ est $\|t\mathbf{v}\| = |t| \|\mathbf{v}\|$.

La courbe $t \mapsto f(\mathbf{a} + t\mathbf{v})$ représente la température sur la ligne joignant \mathbf{a} à $\mathbf{a} + \mathbf{v}$. Formons le quotient

$$\frac{f(\mathbf{a} + t\mathbf{v}) - f(\mathbf{a})}{t}. \quad (5.1)$$

Le numérateur de ce quotient nous dit comment varie la fonction lorsque nous bougeons de \mathbf{a} à $\mathbf{a} + t\mathbf{v}$. Ce quotient est appelé **taux de variation de f** sur la ligne joignant \mathbf{a} à $\mathbf{a} + t\mathbf{v}$. Intéressons nous au comportement de ce quotient lorsque $t \rightarrow 0$.

DÉFINITION 2.3. Soit $\mathbf{f} : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ une fonction donnée. Soit $\mathbf{a} \in U$ et soit \mathbf{v} un vecteur arbitraire dans \mathbb{R}^n . La **dérivée de \mathbf{f} en \mathbf{a} par rapport au vecteur \mathbf{v}** , ou la **dérivée directionnelle en \mathbf{a} dans la direction \mathbf{v}** , notée $\partial_{\mathbf{v}}\mathbf{f}(\mathbf{a})$ (il n'existe pas de notation standard), est définie par l'équation

$$\partial_{\mathbf{v}}\mathbf{f}(\mathbf{a}) = \lim_{t \rightarrow 0} \frac{\mathbf{f}(\mathbf{a} + t\mathbf{v}) - \mathbf{f}(\mathbf{a})}{t} \quad (5.2)$$

lorsque la limite définie dans le membre de droite de l'équation ci-dessus existe.

Autrement dit, $\partial_{\mathbf{v}}\mathbf{f}(\mathbf{a})$ est la vitesse en 0, ou la dérivée en 0, de la courbe de variable réelle $t \mapsto \mathbf{f}(\mathbf{a} + t\mathbf{v})$.

REMARQUE 2.4. Puisque \mathbf{a} est un point de U , il existe une n -boule $\mathcal{B}(\mathbf{a}; r)$ contenue entièrement dans U . Si t est choisit tel que $|t| \|\mathbf{v}\| < r$, la ligne reliant \mathbf{a} à $\mathbf{a} + t\mathbf{v}$ se trouve dans U . Il suffit de prendre t suffisamment petit pour garantir que $\mathbf{a} + t\mathbf{v} \in U$.

EXERCICE 2.1. Calculer $\partial_{\mathbf{v}}f(\mathbf{a})$ si la fonction f est définie par $f(\mathbf{x}) = \|\mathbf{x}\|^2$, pour tout $\mathbf{x} \in \mathbb{R}^n$.

Soit une application différentiable, regardons la relation entre sa différentielle et sa dérivée directionnelle.

THÉORÈME 2.3. *Si \mathbf{f} est différentiable en $\mathbf{a} \in \mathbb{R}^n$, alors, pour tout $\mathbf{v} \in \mathbb{R}^n$, $\partial_{\mathbf{v}}\mathbf{f}(\mathbf{a})$ existe et s'exprime suivant la relation :*

$$\partial_{\mathbf{v}}\mathbf{f}(\mathbf{a}) = D\mathbf{f}(\mathbf{a})(\mathbf{v})$$

DÉMONSTRATION. Comme \mathbf{f} est différentiable en \mathbf{a} , on a :

$$\frac{\mathbf{f}(\mathbf{a} + t\mathbf{v}) - \mathbf{f}(\mathbf{a})}{t} = \frac{D\mathbf{f}(\mathbf{a})(t\mathbf{v}) + o(t\mathbf{v})}{t} = D\mathbf{f}(\mathbf{a})(\mathbf{v}) + \frac{o(t\mathbf{v})}{t},$$

par linéarité de la différentielle. Ensuite, faisons tendre t vers 0. Le terme de droite converge vers $D\mathbf{f}(\mathbf{a})(\mathbf{v})$, qui est fini, car \mathbf{f} est différentiable. Donc la limite du terme de gauche existe, et

$$\partial_{\mathbf{v}}\mathbf{f}(\mathbf{a}) = D\mathbf{f}(\mathbf{a})(\mathbf{v}).$$

□

Attention : la réciproque est fautive. L'exemple suivant décrit un champ scalaire qui possède une dérivée directionnelle selon chaque direction partant de $\mathbf{0}$ mais qui n'est pas continu en ce point, il n'est donc pas différentiable.

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ telle que

$$f(x, y) = \frac{xy^2}{x^2 + y^4}, \quad \text{si } x \neq 0,$$

$$f(0, y) = 0, \quad \text{sinon.}$$

Soit $\mathbf{x}_0 = (0, 0)$ et $\mathbf{v} = (a, b)$ un vecteur. Si $a \neq 0$ et si $t \neq 0$, nous avons

$$\frac{f(\mathbf{0} + t\mathbf{v}) - f(\mathbf{0})}{t} = \frac{f(ta, tb)}{t} = \frac{ab^2}{a^2 + t^2b^4}.$$

Soit $t \rightarrow 0$. Nous trouvons $\partial_{\mathbf{v}}f(\mathbf{0}) = \frac{b^2}{a}$. Si $\mathbf{v} = (0, b)$, nous trouvons, de manière similaire que $\partial_{\mathbf{v}}f(\mathbf{0}) = 0$. Ainsi, $\partial_{\mathbf{v}}f(\mathbf{0})$ existe pour toute direction \mathbf{v} . De plus, $f(\mathbf{x}) \rightarrow 0$ lorsque $\mathbf{x} \rightarrow 0$ le long de toute ligne droite partant de l'origine. Toutefois, en chaque point de la parabole $x = y^2$ (excepté à l'origine) la fonction a comme valeur $1/2$. Puisque de tels points existent arbitrairement proche de l'origine et que $f(\mathbf{0}) = 0$, la fonction f n'est pas continue en $\mathbf{0}$. Cet exemple montre que l'existence de toutes les dérivées directionnelles n'implique pas la continuité en ce point, et donc *a fortiori* pas la différentiabilité.

EXERCICE 2.2. Soit un champ scalaire f , défini sur \mathbb{R}^n par l'équation $f(\mathbf{x}) = \mathbf{x}_0 \cdot \mathbf{x}$, où \mathbf{x}_0 est un vecteur constant.

- (1) Calculer $\partial_{\mathbf{v}}f(\mathbf{x})$ pour des vecteurs arbitraires \mathbf{x} et \mathbf{v} .

- (2) Même question lorsque $f(\mathbf{x}) = \|\mathbf{x}\|^4$.
- (3) Prendre $n = 2$ pour cette dernière fonction et trouver l'ensemble des points tels que

$$\partial_{x_i+y_j} f(2\mathbf{i} + 3\mathbf{j}) = 6.$$

EXERCICE 2.3. Evaluer la dérivée directionnelle du champ scalaire $f(x, y, z) = x^2 + y^2 + 3z^2$ au point $(1, 1, 0)$ selon la direction $\mathbf{y} = \mathbf{i} - \mathbf{j} + 2\mathbf{k}$.

6. Dérivée partielle

On munit \mathbb{R}^n de la base orthonormée $(\mathbf{e}_1, \dots, \mathbf{e}_n)$. Soit \mathbf{f} un champ vectoriel ($\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$).

DÉFINITION 2.4. Dans la [Définition 2.3](#), si on prend $\mathbf{v} = \mathbf{e}_k$ (le k -ième vecteur unitaire des coordonnées), la dérivée directionnelle $\partial_{\mathbf{e}_k} \mathbf{f}(\mathbf{a})$ est appelée la **dérivée partielle de \mathbf{f} par rapport à \mathbf{e}_k** et est également notée $\partial_k \mathbf{f}(\mathbf{a})$. Ainsi, nous notons

$$\partial_k \mathbf{f}(\mathbf{a}) = \partial_{\mathbf{e}_k} \mathbf{f}(\mathbf{a}).$$

Les notations suivantes sont également utilisées pour les dérivées partielles en un point \mathbf{a}

$$\partial_k \mathbf{f}(a_1, \dots, a_n), \quad \frac{\partial \mathbf{f}}{\partial x_k}(a_1, \dots, a_n), \quad \mathbf{f}'_{x_k}(a_1, \dots, a_n), \quad \partial_{x_k} \mathbf{f}(a_1, \dots, a_n), \quad \partial_k \mathbf{f}(a_1, \dots, a_n) = \mathbf{f}_{x_k}.$$

Dans \mathbb{R}^2 , les vecteurs unitaires des coordonnées s'écrivent \mathbf{i} et \mathbf{j} . Si $\mathbf{a} = (a_1, a_2)$, les dérivées partielles $\partial_1 \mathbf{f}(\mathbf{a})$ et $\partial_2 \mathbf{f}(\mathbf{a})$ s'écrivent aussi

$$\frac{\partial \mathbf{f}}{\partial x}(a_1, a_2) \quad \text{et} \quad \frac{\partial \mathbf{f}}{\partial y}(a_1, a_2),$$

Dans \mathbb{R}^3 , si $\mathbf{a} = (a_1, a_2, a_3)$, les dérivées partielles $\partial_1 \mathbf{f}(\mathbf{a})$, $\partial_2 \mathbf{f}(\mathbf{a})$ et $\partial_3 \mathbf{f}(\mathbf{a})$ sont aussi notées

$$\frac{\partial \mathbf{f}}{\partial x}(a_1, a_2, a_3), \quad \frac{\partial \mathbf{f}}{\partial y}(a_1, a_2, a_3) \quad \text{et} \quad \frac{\partial \mathbf{f}}{\partial z}(a_1, a_2, a_3).$$

Remarques sur la notation :

- (1) La notation $\frac{\partial \mathbf{f}}{\partial x_i}$ peut être trompeuse, car elle est sous la forme d'un quotient. C'est la dérivée de \mathbf{f} par rapport à sa i ème variable, et rien de plus ! Les notations $\partial_i \mathbf{f}$ et \mathbf{f}'_i n'ont pas ce défaut.
- (2) D'autre part, il faut prendre garde à la notation $\frac{\partial \mathbf{f}}{\partial x_i}$ en dépit de son aspect intuitif. Par exemple, si $f(x, y, z) = ax^2 + by^2 + cz^2$, écrire $\frac{\partial f}{\partial x} = 2ax$ est sans équivoque. Mais pour une fonction $g(x, y, z) = f(x, xy, x+z)$, écrire $\frac{\partial f(x, xy, x+z)}{\partial x}$ est source d'ambiguïté. Il faut savoir si vous voulez dériver f par rapport à x , puis mettre x , xy et $x+z$ respectivement en première, deuxième et troisième composantes, ou si vous voulez dériver g par rapport à sa première variable.

Dans ce dernier cas, il s'agit de dérivation de fonctions composées. Les deux points de vue sont totalement différents !

EXERCICE 2.4. Dans chacun des exemples suivants, calculer les dérivées partielles du premier ordre des champs scalaires :

(1) $f(x, y) = x^2 + y^2 \sin(xy)$,

(2) $f(x, y) = \frac{x+y}{x-y}$, pour $x \neq y$,

(3) $f(x, y, z) = x^2 - y^2 + 2z^2$.

(4) Dans \mathbb{R}^n , muni de la base orthonormée $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$, on note $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i$ la décomposition de $\mathbf{x} \in \mathbb{R}^n$ dans \mathcal{B} . Calculer les dérivées partielles du premier ordre des champs scalaires suivants

(a) $f(\mathbf{x}) = \mathbf{x}_0 \cdot \mathbf{x}$, le vecteur \mathbf{x}_0 étant fixé (forme linéaire),

(b) $f(\mathbf{x}) = \sum_{i,j=1}^n a_{ij} x_i x_j$, la matrice $A = (a_{ij})_{i,j}$ étant fixée (forme quadratique),

Examinons à présent la **relation entre la différentielle et les dérivées partielles d'une fonction différentiable**.

THÉORÈME 2.4. Si $\mathbf{f} : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ est différentiable en \mathbf{a} , alors pour tout $i \in \{1, \dots, n\}$, $\partial_i \mathbf{f}(\mathbf{a})$ existe et on a la relation :

$$D\mathbf{f}(\mathbf{a})(\mathbf{h}) = \sum_{i=1}^n h_i \partial_i \mathbf{f}(\mathbf{a}),$$

où $\mathbf{h} = (h_1, \dots, h_n)$ dans la base $(\mathbf{e}_1, \dots, \mathbf{e}_n)$.

DÉMONSTRATION. Si \mathbf{f} est différentiable en \mathbf{a} , alors on a vu précédemment que les dérivées directionnelles de \mathbf{f} existent. Les dérivées partielles, qui sont des dérivées directionnelles particulières, existent donc aussi $\partial_i \mathbf{f}(\mathbf{a}) = \partial_{\mathbf{e}_i} \mathbf{f}(\mathbf{a}) = D\mathbf{f}(\mathbf{a})(\mathbf{e}_i)$.

D'autre part, par linéarité de la différentielle $D\mathbf{f}(\mathbf{a})$, on a

$$D\mathbf{f}(\mathbf{a})(\mathbf{h}) = D\mathbf{f}(\mathbf{a})\left(\sum_{i=1}^n h_i \mathbf{e}_i\right) = \sum_{i=1}^n h_i D\mathbf{f}(\mathbf{a})(\mathbf{e}_i) = \sum_{i=1}^n h_i \partial_i \mathbf{f}(\mathbf{a}).$$

□

Attention : la réciproque est fautive. L'existence des dérivées partielles de f en \mathbf{a} ne suffit pas à entraîner sa différentiabilité. Par exemple, la fonction

$$f(x, y) = \frac{xy}{x^2 + y^2} \text{ si } (x, y) \neq (0, 0), \quad f(0, 0) = 0$$

admet des dérivées partielles nulles à l'origine, mais n'est pas différentiable, ni même continue en ce point.

On a cependant le théorème suivant, qui, sous des hypothèses plus fortes, donne une condition suffisante pour la différentiabilité d'une fonction. Il est démontré dans un supplément de cours disponible sur l'ENT.

THÉORÈME 2.5. *Si les différentielles partielles $\partial_i \mathbf{f}$ existent et sont continues sur U , alors l'application $\mathbf{f} : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ est différentiable sur U .*

DÉFINITION 2.5. *Une fonction \mathbf{f} satisfaisant les hypothèses du théorème précédent est dite **continûment différentiable** sur U .*

EXERCICE 2.5. *Dans \mathbb{R}^3 , soit $\mathbf{r}(x, y, z) = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, et $r = \|\mathbf{r}\|$.*

- (1) *Montrer que ∇r est un vecteur unitaire dans la direction de \mathbf{r} .*
- (2) *Montrer que $\nabla r^n = nr^{n-2}\mathbf{r}$, avec n un entier strictement positif.*
- (3) *Montrer que la formule reste vraie pour $n \leq 0$.*
- (4) *Trouver un champ scalaire f tel que $\nabla f = \mathbf{r}$.*

7. Calcul des différentielles

Supposons que les fonctions considérées ici sont différentiables. Le problème est donc de donner la différentielle $D\mathbf{f}(\mathbf{a})$ d'une fonction \mathbf{f} donnée.

7.1. Cas général. Soit $\mathbf{f}(\mathbf{x}) = (f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n))$ définie sur U ouvert de \mathbb{R}^n et à valeurs dans \mathbb{R}^m :

$$\mathbf{f} : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$$

Soit $\mathbf{a} \in U$.

La différentielle $D\mathbf{f}(\mathbf{a})$ est l'application linéaire définie, dans les bases canoniques³ de \mathbb{R}^n et \mathbb{R}^m , par la **matrice jacobienne** :

$$\begin{bmatrix} \partial_1 f_1 & \dots & \partial_n f_1 \\ \vdots & \vdots & \vdots \\ \partial_1 f_m & \dots & \partial_n f_m \end{bmatrix},$$

où les dérivées partielles sont calculées au point \mathbf{a} .

Ici, $D\mathbf{f}(\mathbf{a})(\mathbf{h})$ est le produit matriciel de la matrice jacobienne de \mathbf{f} en \mathbf{a} par le vecteur \mathbf{h} :

$$D\mathbf{f}(\mathbf{a})(\mathbf{h}) = D\mathbf{f}(\mathbf{a}) \cdot \mathbf{h}, \quad \text{pour tout } \mathbf{h} \in \mathbb{R}^n.$$

REMARQUE 2.5. Dans la matrice jacobienne $D\mathbf{f}(\mathbf{a})$, la première ligne est la différentielle de la première composante f_1 de \mathbf{f} , la première colonne est la première dérivée partielle de \mathbf{f} , etc.

Si on craint de confondre lignes et colonnes en écrivant cette matrice, on peut se souvenir qu'elle doit pouvoir être multipliée à droite par un vecteur *colonne* $\mathbf{h} = (h_1, \dots, h_n)$ de manière à donner $D\mathbf{f}(\mathbf{a}) \cdot \mathbf{h}$.

3. La base canonique de \mathbb{R}^p est formée des p vecteurs dont les composantes sont toutes nulles sauf une qui vaut 1. Par exemple, la base $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ est la base canonique de \mathbb{R}^3 .

Le déterminant de $D\mathbf{f}(\mathbf{a})$ est appelé **Jacobien** de \mathbf{f} en \mathbf{a} .

7.2. Cas particuliers.

7.2.1. *Cas d'une fonction de variable réelle.* Soit $\mathbf{f} : U \subset \mathbb{R} \rightarrow \mathbb{R}^m$ définie sur un ouvert U de \mathbb{R} .

Ici, $D\mathbf{f}(\mathbf{a})$ est la dérivée classique de $\mathbf{f} : \mathbf{f}'(a)$, définie par la limite suivante⁴

$$\mathbf{f}'(a) = \lim_{h \rightarrow 0} \frac{\mathbf{f}(a+h) - \mathbf{f}(a)}{h} \in \mathbb{R}^m,$$

La différentielle $D\mathbf{f}(a)$ est l'application linéaire $h \mapsto h\mathbf{f}'(a)$ de \mathbb{R} dans \mathbb{R}^m . Ici, $D\mathbf{f}(a)(h)$ est le produit de h par le vecteur $\mathbf{f}'(a)$:

$$D\mathbf{f}(a)(h) = h\mathbf{f}'(a)$$

pour tout $h \in \mathbb{R}$.

La formule de Taylor au premier ordre s'écrit

$$\mathbf{f}(a+h) = \mathbf{f}(a) + h\mathbf{f}'(a) + o(h).$$

Plus concrètement, l'application \mathbf{f} s'écrit

$$\begin{bmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{bmatrix}, \text{ d'où } \mathbf{f}'(a) = \begin{bmatrix} f'_1(a) \\ \vdots \\ f'_m(a) \end{bmatrix} \in \mathbb{R}^m.$$

La différentiabilité de \mathbf{f} équivaut ici à la dérivabilité des composantes f_1, \dots, f_m .

REMARQUE 2.6. $\mathbf{f}'(a)$ est appelé **vecteur dérivé** ou **vecteur vitesse** de \mathbf{f} en a .

Cas encore plus particulier. Si $f : U \subset \mathbb{R} \rightarrow \mathbb{R}$ est à valeurs réelles, $Df(a)$ est tout simplement la dérivée classique de $f : f'(a)$. Et $Df(a)(h)$ est le produit de $f'(a)$ et h :

$$Df(a)(h) = f'(a)h$$

pour tout $h \in \mathbb{R}$.

Ici, *différentiable* équivaut à *dérivable*, et la différentielle est la fonction linéaire $h \mapsto hf'(a)$, de \mathbb{R} dans \mathbb{R} .

7.2.2. *Cas d'un champ scalaire d'une variable vectorielle.* Soit f un champ scalaire :

$$f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$$

défini sur un ouvert U de \mathbb{R}^n , et soit $\mathbf{a} \in U$.

La différentielle $Df(\mathbf{a})$ est la forme linéaire sur \mathbb{R}^n , de composantes $(\partial_1 f(\mathbf{a}), \dots, \partial_n f(\mathbf{a}))$ dans la base canonique de \mathbb{R}^n . Dans ce contexte, la notation $df(\mathbf{a})$, ou $\nabla f(\mathbf{a})$ est souvent préférée à $Df(\mathbf{a})$. Ici, $Df(\mathbf{a})(\mathbf{h})$ est le produit scalaire entre $\nabla f(\mathbf{a})$ et \mathbf{h} :

$$Df(\mathbf{a})(\mathbf{h}) = \nabla f(\mathbf{a}) \cdot \mathbf{h} = \langle \nabla f(\mathbf{a}), \mathbf{h} \rangle$$

⁴ comme $h \in \mathbb{R}$, on peut diviser par h

pour tout $\mathbf{h} \in \mathbb{R}^n$.

La formule de Taylor au premier ordre s'écrit

$$f(\mathbf{a} + \mathbf{h}) = f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot \mathbf{h} + o(\mathbf{h}).$$

REMARQUE 2.7. $\nabla f(\mathbf{a})$ est appelé **gradient** de f en \mathbf{a} .

Dans le cas bidimensionnel, le vecteur gradient s'écrit

$$\nabla f(x, y) = \frac{\partial f}{\partial x}(x, y)\mathbf{i} + \frac{\partial f}{\partial y}(x, y)\mathbf{j}.$$

En dimension trois, nous avons

$$\nabla f(x, y, z) = \frac{\partial f}{\partial x}(x, y, z)\mathbf{i} + \frac{\partial f}{\partial y}(x, y, z)\mathbf{j} + \frac{\partial f}{\partial z}(x, y, z)\mathbf{k}.$$

Le gradient d'une champ scalaire a une interprétation géométrique : Le gradient $\nabla f(\mathbf{a})$ indique la direction dans laquelle f croît le plus vite. Par exemple, si $f(x, y)$ représente l'altitude au point (x, y) alors $\nabla f(x, y)$ indique la ligne de plus grande pente dans le sens croissant. Il est orthogonal aux lignes de niveaux, lignes où $f(x, y)$ est constante. Voir les figures ci-dessous dans le cas où $f(x, y) = x^2y$.

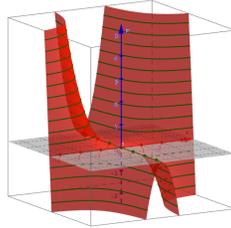


FIGURE 2 – Le graphe de la fonction $f(x, y) = x^2y$ et quelques lignes de niveaux tracées sur le graphe.

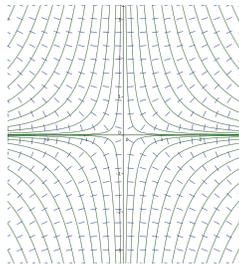


FIGURE 3 – Les lignes de niveaux dans le plan et le champ de gradient de $f(x, y) = x^2y$.

8. Réécriture du Théorème des fonctions composées

On reprend les notations du [Théorème 2.2](#) des fonctions composées.

On peut exprimer la règle de dérivation des fonctions composées grâce aux matrices jacobiniennes $D\mathbf{f}(\mathbf{y}_0)$ et $D\mathbf{g}(\mathbf{x}_0)$. Puisque la composition d'applications linéaires correspond à la multiplication des matrices correspondantes, nous obtenons

$$D(\mathbf{f} \circ \mathbf{g})(\mathbf{x}_0) = D\mathbf{f}(\mathbf{y}_0) \cdot D\mathbf{g}(\mathbf{x}_0) \quad (8.1)$$

où $\mathbf{y}_0 = \mathbf{g}(\mathbf{x}_0)$.

8.1. Détails et exemples associés à l'équation (8.1). On peut exprimer la règle (8.1) sous la forme d'un ensemble d'équations scalaires. Supposons que $\mathbf{x}_0 \in \mathbb{R}^p$, $\mathbf{y}_0 = \mathbf{g}(\mathbf{x}_0) \in \mathbb{R}^n$, et $\mathbf{f}(\mathbf{y}_0) \in \mathbb{R}^m$. Alors, $\mathbf{f} \circ \mathbf{g}(\mathbf{x}_0) \in \mathbb{R}^m$ et nous pouvons écrire

$$\begin{aligned} \mathbf{g} &= (g_1, \dots, g_n), \\ \mathbf{f} &= (f_1, \dots, f_m), \end{aligned}$$

Alors, $D(\mathbf{f} \circ \mathbf{g})(\mathbf{x}_0)$ est une matrice $m \times p$, $D\mathbf{f}(\mathbf{y}_0)$ est une matrice $m \times n$ et $D\mathbf{g}(\mathbf{x}_0)$ est une matrice $n \times p$, données respectivement par

$$\begin{aligned} D(\mathbf{f} \circ \mathbf{g})(\mathbf{x}_0) &= [\partial_j(\mathbf{f} \circ \mathbf{g})_i(\mathbf{x}_0)]_{i,j=1}^{m,p}, \\ D\mathbf{f}(\mathbf{y}_0) &= [\partial_j f_i(\mathbf{y}_0)]_{i,j=1}^{m,n}, \\ D\mathbf{g}(\mathbf{x}_0) &= [\partial_j g_i(\mathbf{x}_0)]_{i,j=1}^{n,p}. \end{aligned}$$

L'équation (8.1) est équivalente à mp équations scalaires

$$\partial_j(\mathbf{f} \circ \mathbf{g})_i(\mathbf{x}_0) = \sum_{k=1}^n \partial_k f_i(\mathbf{y}_0) \partial_j g_k(\mathbf{x}_0), \quad \forall 1 \leq i \leq m, \quad 1 \leq j \leq p.$$

EXEMPLE 2.1. (Règle de dérivation en chaîne développée pour les champs scalaires). Si $m = 1$, f est un champ scalaire, $f \circ \mathbf{g}$ aussi. Les p équations (une pour chaque dérivée partielle de $f \circ \mathbf{g}$) s'écrivent

$$\partial_j(f \circ \mathbf{g})(\mathbf{x}_0) = \sum_{k=1}^n \partial_k f(\mathbf{y}_0) \partial_j g_k(\mathbf{x}_0), \quad \forall 1 \leq j \leq p.$$

Le cas particulier $p = 1$ donne une équation

$$(f \circ \mathbf{g})'(\mathbf{x}_0) = \sum_{k=1}^n \partial_k f(\mathbf{y}_0) g'_k(\mathbf{x}_0).$$

Ce qui s'écrit plus simplement, si $m = p = 1$:

cas d'un champ scalaire composé avec un champ vectoriel :

$$(f \circ \mathbf{g})'(\mathbf{x}_0) = \nabla f(\mathbf{y}_0) \cdot \mathbf{g}'(\mathbf{x}_0)$$

où $\mathbf{y}_0 = \mathbf{g}(\mathbf{x}_0)$.

EXEMPLE 2.2. Considérons $p = 2$ et $n = 2$. Ecrivons $\mathbf{x}_0 = (s, t)$ et $\mathbf{b} = (x, y)$. Alors, les composantes de x et y sont reliées à s et t par les équations

$$x = g_1(s, t) \quad \text{et} \quad y = g_2(s, t).$$

La règle de dérivation en chaîne donne un couple d'équations pour les dérivées partielles de $\mathbf{f} \circ \mathbf{g}$

$$\partial_1(\mathbf{f} \circ \mathbf{g})(s, t) = \partial_1 \mathbf{f}(x, y) \partial_1 g_1(s, t) + \partial_2 \mathbf{f}(x, y) \partial_1 g_2(s, t).$$

De même, nous avons

$$\partial_2(\mathbf{f} \circ \mathbf{g})(s, t) = \partial_1 \mathbf{f}(x, y) \partial_2 g_1(s, t) + \partial_2 \mathbf{f}(x, y) \partial_2 g_2(s, t).$$

Ce couple d'équations s'écrit aussi (c'est la même chose, écrite autrement) :

$$\begin{aligned} \frac{\partial(\mathbf{f} \circ \mathbf{g})}{\partial s} &= \frac{\partial \mathbf{f}}{\partial x} \frac{\partial g_1}{\partial s} + \frac{\partial \mathbf{f}}{\partial y} \frac{\partial g_2}{\partial s}, \\ \frac{\partial(\mathbf{f} \circ \mathbf{g})}{\partial t} &= \frac{\partial \mathbf{f}}{\partial x} \frac{\partial g_1}{\partial t} + \frac{\partial \mathbf{f}}{\partial y} \frac{\partial g_2}{\partial t}. \end{aligned}$$

REMARQUE 2.8. Par abus, on écrit parfois $\frac{\partial \mathbf{f}}{\partial s}$ et $\frac{\partial \mathbf{f}}{\partial t}$ au lieu de $\frac{\partial(\mathbf{f} \circ \mathbf{g})}{\partial s}$ et $\frac{\partial(\mathbf{f} \circ \mathbf{g})}{\partial t}$.

REMARQUE 2.9. Supposons $m = 1$. Plutôt que de multiplier de volumineuses matrices, on pourra retenir que, si $\mathbf{y} = \mathbf{g}(\mathbf{x})$ et $z = f(\mathbf{y}) = f(\mathbf{g}(\mathbf{x}))$, les dérivées partielles de $z = f(g_1(x_1, \dots, x_p), \dots, g_n(x_1, \dots, x_p))$ sont les

$$\frac{\partial(f \circ \mathbf{g})}{\partial x_i} = \frac{\partial f}{\partial y_1} \cdot \frac{\partial g_1}{\partial x_i} + \dots + \frac{\partial f}{\partial y_n} \cdot \frac{\partial g_n}{\partial x_i}.$$

EXEMPLE 2.3. Soit $\mathbf{g}(t) = (P(t), Q(t))$ où P et Q sont deux fonctions dérivables sur \mathbb{R} , et Q ne s'annule pas sur \mathbb{R} . Soit $f(u, v) = u/v$. On souhaite dériver $f \circ \mathbf{g}(t) = f(\mathbf{g}(t)) = P(t)/Q(t)$. On a

$$\nabla f(\mathbf{g}(t_0)) \cdot \mathbf{g}'(t_0) = \begin{pmatrix} 1/Q(t_0) \\ -P(t_0)/Q^2(t_0) \end{pmatrix} \cdot \begin{pmatrix} P'(t_0) \\ Q'(t_0) \end{pmatrix} = P'(t_0)/Q(t_0) - P(t_0)Q'(t_0)/Q^2(t_0)$$

et on retrouve la formule bien connue de la dérivée de P/Q .

8.2. La différentielle de l'inverse. Soit $\mathbf{f} : E \rightarrow F$ bijective et différentiable, où E et F sont deux ouverts de \mathbb{R}^n et \mathbb{R}^m . Supposons que, pour $\mathbf{y} \in E$, la différentielle de \mathbf{f} en $\mathbf{f}^{-1}(\mathbf{y})$ soit inversible⁵. Alors,

$$D\mathbf{f}^{-1}(\mathbf{y}) = (D\mathbf{f}(\mathbf{f}^{-1}\mathbf{y}))^{-1}.$$

5. i.e. le Jacobien de \mathbf{f} en $\mathbf{f}^{-1}(\mathbf{y})$ est non nul

8.3. Notation différentielle. Elle donne une autre représentation de ces mêmes calculs. Si $z = f(\mathbf{y})$, et $\mathbf{y} = \mathbf{g}(\mathbf{x})$, on écrit

$$dz = \frac{\partial z}{\partial y_1} dy_1 + \dots + \frac{\partial z}{\partial y_n} dy_n, \text{ puis } dy_j = \frac{\partial y_j}{\partial x_1} dx_1 + \dots + \frac{\partial y_j}{\partial x_p} dx_p,$$

on reporte, d'où⁶

$$dz = \left(\sum_{j=1}^n \frac{\partial z}{\partial y_j} \cdot \frac{\partial y_j}{\partial x_1} \right) dx_1 + \dots + \left(\sum_{j=1}^n \frac{\partial z}{\partial y_j} \cdot \frac{\partial y_j}{\partial x_n} \right) dx_p.$$

Il n'y a plus qu'à *identifier* les coefficients avec ceux de

$$dz = \frac{\partial z}{\partial x_1} dx_1 + \dots + \frac{\partial z}{\partial x_p} dx_p$$

pour obtenir l'expression voulue des $\partial z / \partial x_i$!

EXERCICE 2.6. On suppose que toutes les dérivées des fonctions suivantes sont continues et existent. Les équations $u = f(x, y)$, $x = X(t)$, $y = Y(t)$, définissent u comme une fonction de t , que nous notons $u = F(t)$.

a) Montrer que

$$F'(t) = \frac{\partial f}{\partial x} X'(t) + \frac{\partial f}{\partial y} Y'(t),$$

où $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ sont évaluées en $(X(t), Y(t))$.

b) De la même manière, calculer $F''(t)$ en fonction de f , X et Y .

c) Appliquer ces résultats aux fonctions

$$f(x, y) = x^2 + y^2, \quad X(t) = t, \quad Y(t) = t^2.$$

EXERCICE 2.7. Soit $\mathbf{f}(x, y) = (f_1(x, y), f_2(x, y))$ un champ vectoriel à valeurs dans \mathbb{R}^2 , où f_1 et f_2 sont deux champs scalaires définis sur \mathbb{R}^2 .

On suppose que \mathbf{f} dérive d'un potentiel φ sur \mathbb{R}^2 , c'est-à-dire : il existe φ champ scalaire défini sur \mathbb{R}^2 tel que $\mathbf{f} = \nabla \varphi$ sur \mathbb{R}^2 .

Soit l'équation différentielle

$$f_2(t, u(t))u'(t) + f_1(t, u(t)) = 0, \quad t \in \mathbb{R}, \quad (8.2)$$

(1) Montrer que si $u = u(t)$ est une solution de l'équation différentielle (8.2), alors u vérifie

$$\varphi(t, u(t)) = \text{constante}, \quad \forall t \in \mathbb{R}.$$

6. On se gardera bien de 'simplifier' par ∂y_j ! Rappelons que la notation $\partial z / \partial y_j$ ne représente pas un quotient, mais seulement la dérivée partielle de $z = f(\mathbf{y})$ par rapport à sa j ème variable y_j .

(2) Application : En déduire une solution de

$$\begin{cases} u^2(t)u'(t) + \sin t = 0 \\ u(0) = 3 \end{cases}$$

Étapes suggérées : explicitez \mathbf{f} , vérifiez qu'il dérive d'un potentiel φ , et si c'est le cas, calculez φ .

9. Dérivées partielles d'ordre supérieur

Nous donnons les définitions dans le cas des champs scalaires. Les définitions sont les mêmes dans le cas de champs vectoriels. Les dérivées partielles produisent de nouveaux champs scalaires $\partial_1 f, \dots, \partial_n f$ pour un champ f donné. Les dérivées partielles de $\partial_1 f, \dots, \partial_n f$ sont appelées **dérivées secondes** de f . Pour les fonctions de deux variables, il y a quatre dérivées partielles secondes que l'on écrit

$$\partial_1(\partial_1 f) = \frac{\partial^2 f}{\partial x^2}, \quad \partial_1(\partial_2 f) = \frac{\partial^2 f}{\partial x \partial y}, \quad \partial_2(\partial_1 f) = \frac{\partial^2 f}{\partial y \partial x}, \quad \text{et} \quad \partial_2(\partial_2 f) = \frac{\partial^2 f}{\partial y^2}.$$

On utilise quelques fois la notation $\partial_{i,j} f$ pour la dérivée seconde $\partial_i(\partial_j f)$. Par exemple, $\partial_{1,2} f = \partial_1(\partial_2 f)$. Selon la notation ∂ , on précise l'ordre de dérivation en écrivant

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right).$$

Il est possible que cette quantité soit égale ou non à l'autre dérivée mixte $\frac{\partial^2 f}{\partial y \partial x}$. Si f est une fonction à valeurs réelles de deux variables, les deux dérivées mixtes $\partial_{1,2} f$ et $\partial_{2,1} f$ ne sont pas nécessairement égales. Un exemple est donné par l'exercice suivant.

EXERCICE 2.8. Soit f la fonction définie par

$$f(x, y) = xy \frac{x^2 - y^2}{x^2 + y^2}, \quad \text{pour } (x, y) \neq (0, 0),$$

et $f(0, 0) = 0$ sinon. Montrer que $\partial_{2,1} f(0, 0) = -1$ et $\partial_{1,2} f(0, 0) = 1$.

Dans l'exercice précédent, les deux dérivées partielles $\partial_{2,1} f$ et $\partial_{1,2} f$ ne sont pas continues à l'origine. On peut montrer que les deux dérivées mixtes sont égales en un point (x_0, y_0) si au moins l'une d'entre elles est continue en un voisinage de ce point. On peut montrer dans un premier temps qu'elles sont égales si elles sont toutes deux continues. Ceci fait l'objet du théorème suivant.

THÉORÈME 2.6. (Une condition suffisante pour l'égalité des dérivées partielles mixtes). Supposons que f soit un champ scalaire tel que les dérivées partielles $\partial_1 f$, $\partial_2 f$, $\partial_{1,2} f$ et $\partial_{2,1} f$ existent sur un ouvert \mathcal{S} . Si (x_0, y_0) est un point de \mathcal{S} où $\partial_{1,2} f$ et $\partial_{2,1} f$ sont continues, nous avons alors

$$\partial_{1,2} f(x_0, y_0) = \partial_{2,1} f(x_0, y_0).$$

On peut en fait démontrer une version plus forte de ce théorème.

THÉORÈME 2.7. *Soit f un champ scalaire tel que les dérivées partielles $\partial_1 f$, $\partial_2 f$ et $\partial_{2,1} f$ existent sur un ouvert \mathcal{S} contenant (x_0, y_0) . Supposons de plus que $\partial_{2,1} f$ est continue sur \mathcal{S} . Alors, la dérivée partielles $\partial_{1,2} f(x_0, y_0)$ existe et nous avons*

$$\partial_{1,2} f(x_0, y_0) = \partial_{2,1} f(x_0, y_0).$$

DÉFINITION 2.6. *Soit \mathbf{f} une fonction telle que ses dérivées partielles premières et secondes existent et sont continues sur un ouvert S de \mathbb{R}^n , alors \mathbf{f} est dite de **classe \mathcal{C}^2** sur S ou encore **deux fois continûment différentiable** sur S .*

EXERCICE 2.9. (Coordonnées polaires). Soit f un champ scalaire dépendant de (x, y) . En coordonnées polaires, nous avons $x = r \cos(\theta)$ et $y = r \sin(\theta)$. Posons : $\varphi(r, \theta) = f(r \cos(\theta), r \sin(\theta))$.

(a) Exprimer $\frac{\partial \varphi}{\partial r}$ et $\frac{\partial \varphi}{\partial \theta}$ en fonction de $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$.

(b) Exprimer la dérivée partielle du second ordre $\frac{\partial^2 \varphi}{\partial \theta^2}$ en fonction de celles de f .

EXERCICE 2.10. Soient $x = r \cos(\theta)$ et $y = r \sin(\theta)$, $\Omega = \{(x, y) \in \mathbb{R}^2 : x, y > 0\}$ et $f : \Omega \rightarrow \mathbb{R}$ une fonction $\mathcal{C}^2(\Omega)$.

(a) Montrer que le laplacien de f

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

s'écrit

$$\Delta f = \frac{\partial^2 f}{\partial r^2} + \frac{1}{r} \frac{\partial f}{\partial r} + \frac{1}{r^2} \frac{\partial^2 f}{\partial \theta^2} \quad (9.1)$$

(b) En utilisant (9.1), calculer Δf en coordonnées polaires pour la fonction f définie par

$$f(x, y) = \sqrt{x^2 + y^2} + \left(\arctg\left(\frac{y}{x}\right)\right)^2.$$

Pour cet exercice, on pourra commencer par montrer par exemple que

$$\frac{\partial r}{\partial x} = \frac{x}{r}, \quad \frac{\partial r}{\partial y} = \frac{y}{r}, \quad \frac{\partial \theta}{\partial x} = -\frac{y}{r^2}, \quad \frac{\partial \theta}{\partial y} = \frac{x}{r^2}.$$

EXERCICE 2.11. Soient $\mathbf{x} \in \mathbb{R}^n$, $r = \sqrt{x_1^2 + \dots + x_n^2}$, f et ψ des champs scalaires, respectivement sur \mathbb{R}^n et \mathbb{R}^+ .

(1) Montrer que $\forall i \in \{1, \dots, n\}$,

$$\frac{\partial r}{\partial x_i} = \frac{x_i}{r}.$$

(2) Montrer que si $f(\mathbf{x}) = \psi(r)$, alors pour $\mathbf{x} \neq \mathbf{0}$, on a :

$$\Delta f = \frac{d^2 \psi}{dr^2}(r) + \frac{n-1}{r} \frac{d\psi}{dr}(r).$$

(3) En déduire une solution de $\Delta f = 0$ dans $\mathbb{R}^n \setminus \{0\}$

10. Exercices

EXERCICE 2.12. Soit la fonction f définie sur \mathbb{R}^2 par

$$f(x, y) = \frac{x^2 y}{x^4 - 2x^2 y + 3y^2}, \quad \text{si } (x, y) \neq (0, 0),$$

et $f(0, 0) = 0$.

- Montrer que la restriction de f à toute droite passant par l'origine est continue.
- Montrer que f n'est pas continue à l'origine.

EXERCICE 2.13. (Calcul explicite de différentielles)

Calculer la différentielle à l'origine des applications suivantes :

- $f : \mathbb{R}^2 \rightarrow \mathbb{R}$
 $(x, y) \mapsto f(x, y) = x^3 + \sqrt{1 + x^2 + y^2}$
- $g : \mathbb{R}^3 \rightarrow \mathbb{R}$
 $(x, y, z) \mapsto g(x, y, z) = xyz \sin(xy) + 2x + 5$.

EXERCICE 2.14. (Linéarité des opérateurs différentiels).

On définit les fonctions Δ et Φ par :

$$\begin{aligned} \Delta : \mathcal{C}^\infty(\mathbb{R}^2) &\longrightarrow \mathcal{C}^\infty(\mathbb{R}^2) & \text{et } \Phi : \mathcal{C}^\infty(\mathbb{R}^2) &\longrightarrow \mathcal{C}^\infty(\mathbb{R}^2) \\ \varphi &\longmapsto \frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial y^2} & \varphi &\longmapsto \frac{\partial \varphi}{\partial x} \cdot \frac{\partial \varphi}{\partial y}. \end{aligned}$$

Δ et Φ sont-elles linéaires ?

EXERCICE 2.15. (Continuité et différentiabilité d'une fonction de plusieurs variables). Soit f , la fonction définie par :

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} \\ (x, y) \longmapsto \begin{cases} 0 & \text{si } (x, y) = (0, 0) \\ \frac{|x|^{\frac{3}{2}} y}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \end{cases}$$

- f est-elle continue ?
- Calculer les dérivées partielles de f . Sont-elles continues ?
- L'application f est-elle différentiable en $(0, 0)$?

EXERCICE 2.16. (Continuité et différentiabilité d'une fonction de plusieurs variables).

$$\begin{aligned} f : \mathbb{R}^2 \setminus \{(0, 0)\} &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \frac{xy}{\sqrt{x^2 + y^2}}. \end{aligned}$$

- (1) Montrer que l'on peut prolonger f par continuité. On appelle \tilde{f} , ce prolongement.
- (2) Étudier la différentiabilité de \tilde{f} .
- (3) \tilde{f} admet-elle des dérivées partielles ?
- (4) f est-elle \mathcal{C}^1 sur son ensemble de définition ?

EXERCICE 2.17. (Continuité et différentiabilité d'une fonction de plusieurs variables). Reprendre les questions de l'exercice précédent avec la fonction :

$$\begin{aligned} g : \mathbb{R}^2 \setminus \{(0, 0)\} &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto (x^4 + y^4) \sin \left(\frac{1}{\sqrt{x^4 + y^4}} \right). \end{aligned}$$

EXERCICE 2.18. (Différentiabilité d'une fonction de plusieurs variables). La fonction suivante est-elle différentiable ?

$$\begin{aligned} f : \mathbb{R}^2 \setminus \{(0, 0)\} &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto f(x, y) = \begin{cases} \frac{y^2 - x}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{sinon} \end{cases} \end{aligned}$$

EXERCICE 2.19. (Régularité d'une fonction de plusieurs variables). Soit f , la fonction de \mathbb{R}^2 dans \mathbb{R} définie par :

$$f(x, y) = x^2 y \sin \left(\frac{y}{x} \right).$$

- (1) Montrer que l'on peut définir un prolongement par continuité de la fonction f .
- (2) f admet-elle des dérivées partielles sur \mathbb{R}^2 ?
- (3) f est-elle de classe \mathcal{C}^1 sur \mathbb{R}^2 ?
- (4) Calculer $\frac{\partial^2 f}{\partial x \partial y}(0, 0)$ et $\frac{\partial^2 f}{\partial y \partial x}(0, 0)$.

Remarque : attention à donner un sens convenable aux deux expressions ci-dessus, avant de les calculer. Rien ne certifie que f est de classe \mathcal{C}^2 au voisinage de $(0, 0)$.

Généralités et étude théorique des problèmes d'optimisation

Au cours de ce chapitre, nous allons nous intéresser à résoudre des problèmes d'optimisation qui se résumeront à maximiser ou minimiser certaines fonctions. Nos méthodes seront celles du calcul différentiel et nous nous restreindrons à la dimension finie. Voici quelques exemples de problème que nous pourrons traiter :

- Après une série de mesures expérimentales, on souhaite faire une régression linéaire. Quelle est la droite qui approche le mieux les mesures. On verra que la méthode des moindres carrés apporte une réponse à ce problème.
- Fixons un volume V . Quel est l'emballage (parallélépipédique) le plus économique ? C'est-à-dire comment minimiser la surface sous la contrainte d'un volume donné ?
- Comment communiquer le plus d'information sous la contrainte d'une quantité de données fixée ? C'est l'entropie (dûe à Shannon en théorie de l'information) que l'on va maximiser.
- Les points d'équilibre en mécanique du point correspondent aux extrémums de l'énergie potentielle. Les minimums sont des points d'équilibre stables, les maximums des points d'équilibre instables.

1. Maximum, minimum, point-selle

DÉFINITION 3.1. Soit f un champ défini sur $C \subset \mathbb{R}^n$ à valeurs dans \mathbb{R} .

- Un point $\mathbf{x}_0 \in C$ est un **minimum global** (ou **absolu**) de f sur C si

$$\forall \mathbf{x} \in C, \quad f(\mathbf{x}_0) \leq f(\mathbf{x}).$$

- Un point $\mathbf{x}_0 \in C$ est un **minimum local** (ou **relatif**) de f sur C si il existe un voisinage de \mathbf{x}_0 dans C , noté $\mathcal{B}(\mathbf{x}_0)$, tel que

$$\forall \mathbf{x} \in \mathcal{B}(\mathbf{x}_0), \quad f(\mathbf{x}_0) \leq f(\mathbf{x}).$$

- Un point \mathbf{x}_0 est un point **stationnaire** de f si f est supposé différentiable en \mathbf{x}_0 , et si

$$\nabla f(\mathbf{x}_0) = \mathbf{0}.$$

- Un point \mathbf{x}_0 est un **point-selle** de f si c'est un point stationnaire tel que dans toute boule $\mathcal{B}(\mathbf{x}_0)$, il existe des points \mathbf{x} tels que $f(\mathbf{x}) < f(\mathbf{x}_0)$ et d'autres tels que $f(\mathbf{x}) > f(\mathbf{x}_0)$.

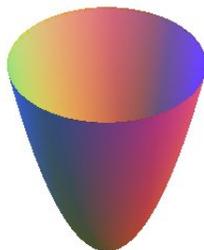


FIGURE 1 – Exemple 1 : surface d'équation $z = f(x, y) = x^2 + y^2$ (paraboloïde de révolution). L'origine est un minimum global.

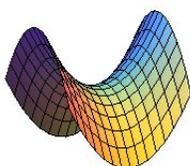


FIGURE 2 – Exemple 2 : surface d'équation $z = f(x, y) = x^2 - y^2$ (paraboloïde hyperbolique). Il y a un point-selle à l'origine.

2. Rappels sur les matrices réelles symétriques

On rappelle que $\mathcal{M}_n(\mathbb{R})$ désigne l'ensemble des matrices réelles de dimension $n \times n$.

THÉORÈME 3.1. *Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice **réelle symétrique**. Alors les valeurs propres de A sont réelles, et A est diagonalisable dans une base orthonormée de vecteurs propres de A .*

REMARQUE 3.1. — Les vecteurs propres de A sont donc orthogonaux. Il suffit de les normer pour en faire une base orthonormée.

— Dans le théorème, le fait que A soit réelle a son importance. Par exemple $B = \begin{bmatrix} 1 & i \\ i & -1 \end{bmatrix}$ a 0 comme seule valeur propre. Si elle était diagonalisable, elle serait nulle.

— Avec une matrice symétrique $A \in \mathcal{M}_n(\mathbb{R})$, on définit la **forme quadratique** :

$$\mathbf{y} \in \mathbb{R}^n \mapsto \mathbf{A}\mathbf{y} \cdot \mathbf{y},$$

ce qui s'écrit si $A = [a_{ij}]_{1 \leq i, j \leq n}$ et $\mathbf{y} = y_i \mathbf{e}_i$ dans la base orthonormée $(\mathbf{e}_i)_{1 \leq i \leq n}$ de \mathbb{R}^n :

$$\mathbf{A}\mathbf{y} \cdot \mathbf{y} = \sum_{i, j=1}^n a_{ij} y_i y_j.$$

DÉFINITION 3.2. — La matrice symétrique $A \in \mathcal{M}_n(\mathbb{R})$ est **définie positive** si

$$\forall \mathbf{y} \in \mathbb{R}^n, \quad \mathbf{y} \neq \mathbf{0}, \quad \mathbf{A}\mathbf{y} \cdot \mathbf{y} > 0.$$

— La matrice symétrique $A \in \mathcal{M}_n(\mathbb{R})$ est **définie négative** si

$$\forall \mathbf{y} \in \mathbb{R}^n, \quad \mathbf{y} \neq \mathbf{0}, \quad A\mathbf{y} \cdot \mathbf{y} < 0.$$

On peut relier le signe de la forme quadratique associée à A au signe du spectre de A :

THÉORÈME 3.2. *Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice réelle symétrique, alors*

— *A est définie positive \Leftrightarrow toutes les valeurs propres de A sont strictement positives.*

— *A est définie négative \Leftrightarrow toutes les valeurs propres de A sont strictement négatives.*

DÉMONSTRATION. Comme A est réelle symétrique, il existe une base orthonormée de vecteurs propres pour A , que l'on note $(\mathbf{v}_i)_{1 \leq i \leq n}$. Dans cette base, $A\mathbf{y} \cdot \mathbf{y}$ s'écrit $\sum_{i=1}^n \lambda_i y_i^2$, où les λ_i sont les valeurs propres de A et y_i les composantes de \mathbf{y} dans la base $(\mathbf{v}_i)_{1 \leq i \leq n}$. \square

3. Hessienne, Formule de Taylor au second ordre

DÉFINITION 3.3. *Soit f une fonction de classe \mathcal{C}^2 au voisinage de $\mathbf{x} \in \Omega \subset \mathbb{R}^n$, à valeurs dans \mathbb{R} . On note $H(\mathbf{x})$ ou $D^2 f(\mathbf{x})$ la **matrice hessienne** de f en $\mathbf{x} \in \Omega$, i.e.*

$$H(\mathbf{x}) = [\partial_{x_i x_j} f(\mathbf{x})]_{1 \leq i, j \leq n}.$$

La formule de Taylor au second ordre s'écrit au point \mathbf{x} :

$$f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + \nabla f(\mathbf{x}) \cdot \mathbf{y} + \frac{1}{2} H(\mathbf{x}) \mathbf{y} \cdot \mathbf{y} + o(\|\mathbf{y}\|^2).$$

Si f est \mathcal{C}^2 sur Ω , on en déduit que $H(\mathbf{x})$ est symétrique. C'est le lemme de Schwarz.

Si \mathbf{x} est un point critique alors la formule de Taylor s'écrit

$$f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + \frac{1}{2} H(\mathbf{x}) \mathbf{y} \cdot \mathbf{y} + o(\|\mathbf{y}\|^2).$$

Ainsi, le terme prépondérant au voisinage de \mathbf{x} (après le terme constant) correspond à une forme quadratique donné par la matrice symétrique $1/2 H(\mathbf{x})$. Pour comprendre les fonctions différentiables deux fois au voisinage de point critique, on a besoin de comprendre les matrices symétriques et les formes quadratiques associées.

4. Nature des points stationnaires

Le théorème suivant relie la nature des points stationnaires et le spectre de la matrice hessienne.

THÉORÈME 3.3. *Soit $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ et \mathbf{x}_0 un **point stationnaire** de f . On suppose que f est de classe \mathcal{C}^2 au voisinage $\mathcal{B}(\mathbf{x}_0)$ de \mathbf{x}_0 . Soit $H(\mathbf{x}_0)$ la **hessienne de f au point \mathbf{x}_0** . Alors*

- Si toutes les valeurs propres de $H(\mathbf{x}_0)$ sont positives strictement, f a un minimum relatif en \mathbf{x}_0 ,
- Si toutes les valeurs propres de $H(\mathbf{x}_0)$ sont négatives strictement, f a un maximum relatif en \mathbf{x}_0 ,
- Si $H(\mathbf{x}_0)$ possède au moins une valeur propre strictement positive et une valeur propre strictement négative, alors \mathbf{x}_0 est un point-selle pour f .

DÉMONSTRATION. D'après [Théorème 3.2](#), si les valeurs propres de $H(\mathbf{x}_0)$ sont toutes strictement positives, alors $\forall \mathbf{y} \in \mathbb{R}^n$, $H(\mathbf{x}_0)\mathbf{y} \cdot \mathbf{y} > 0$. De plus $\mathbf{y} \mapsto H(\mathbf{x}_0)\mathbf{y} \cdot \mathbf{y}$ est continue, donc atteint son minimum sur la sphère unité : $\alpha = \inf_{\|\mathbf{y}\|=1} H(\mathbf{x}_0)\mathbf{y} \cdot \mathbf{y} > 0$. Donc si $\mathbf{y} \rightarrow \mathbf{0}$, la formule de Taylor s'écrit

$$\begin{aligned} f(\mathbf{x}_0 + \mathbf{y}) - f(\mathbf{x}_0) &= \frac{1}{2} [H(\mathbf{x}_0)\mathbf{y} \cdot \mathbf{y} + o(\|\mathbf{y}\|^2)] \\ &= \frac{1}{2} \|\mathbf{y}\|^2 \left[H(\mathbf{x}_0) \frac{\mathbf{y}}{\|\mathbf{y}\|} \cdot \frac{\mathbf{y}}{\|\mathbf{y}\|} + o(1) \right] \\ &\geq \frac{1}{2} \|\mathbf{y}\|^2 [\alpha + o(1)] \geq 0 \end{aligned}$$

sur un voisinage de $\mathbf{y} = \mathbf{0}$. On en déduit que $f(\mathbf{x}_0 + \mathbf{y}) \geq f(\mathbf{x}_0)$ sur ce voisinage. \square

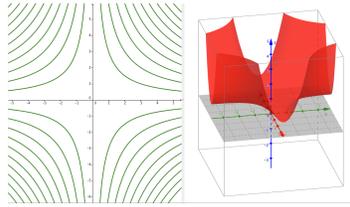
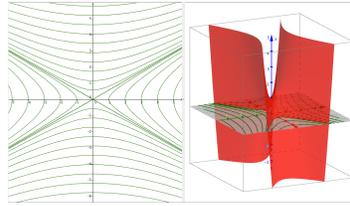
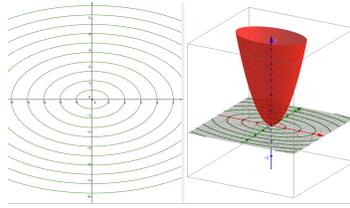
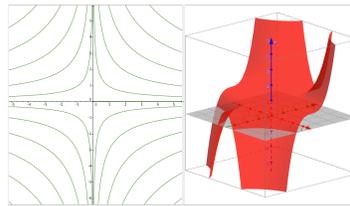
REMARQUE 3.2. Si on suppose seulement que toutes les valeurs propres sont positives, on ne peut pas conclure. Il faut approfondir l'étude.

EXEMPLE 3.1. (1) Si $f(x, y) = x^2 + 3y^2$ (paraboloïde de révolution). L'origine est un point stationnaire, f est de classe C^2 sur \mathbb{R}^2 , et $H(\mathbf{0})$ a pour valeurs propres 2 et 2. L'origine est donc un minimum global. Les lignes de niveaux sont des ellipses.

(2) Si $f(x, y) = x^2 - 3y^2$ (paraboloïde hyperbolique). L'origine est un point stationnaire, f est de classe C^2 sur \mathbb{R}^2 , et $H(\mathbf{0})$ a pour valeurs propres 2 et -2. L'origine est un point selle.

(3) Si $f(x, y) = 1 + x^2y^2$. L'origine est un point stationnaire, f est de classe C^2 sur \mathbb{R}^2 , et $H(\mathbf{0})$ est nulle. On ne peut donc pas conclure. Mais on voit clairement que $f(x, y) \geq 1 = f(\mathbf{0})$, et l'origine est donc un maximum global.

(4) Si $f(x, y) = 1 + x^2y^3$. Les points avec une coordonnée nulle sont des points stationnaires, f est de classe C^2 sur \mathbb{R}^2 , et la matrice hessienne a toujours (au moins) une valeur propre nulle sur les axes Ox et Oy . Selon le point, on obtient un extremum local ou un point selle.

FIGURE 3 – Lignes de niveaux et graphe de $f(x, y) = 1 + x^2y^2$ FIGURE 4 – Lignes de niveaux et graphe de $f(x, y) = x^2 - 3y^2$ FIGURE 5 – Lignes de niveaux et graphe de $f(x, y) = x^2 + 3y^2$ FIGURE 6 – Lignes de niveaux et graphe de $f(x, y) = 1 + x^2y^3$

5. Résultats d'existence

En suite on s'intéresse aux problèmes suivants : "trouver le minimum d'une fonction sans ou avec contrainte(s)". D'un point de vue mathématique, on considère un champ J défini sur \mathbb{R}^n ou sur un sous-ensemble Ω de \mathbb{R}^n , et on cherche

— le minimum de J sans contrainte

$$\min_{\mathbf{x} \in \mathbb{R}^n} J(\mathbf{x}),$$

— le minimum de J avec contrainte

$$\min_{\mathbf{x} \in \Omega} J(\mathbf{x}).$$

REMARQUE 3.3. — Maximiser J revient à minimiser $-J$.

— Il est important de distinguer extremum local et global.

— Les contraintes sont souvent de type inégalités :

$$C = \{\mathbf{x} \in \mathbb{R}^n \text{ tels que } \varphi_i(\mathbf{x}) \leq 0, \forall i \in I\},$$

ou de type égalités

$$C = \{\mathbf{x} \in \mathbb{R}^n \text{ tels que } \varphi_i(\mathbf{x}) = 0, \forall i \in I\},$$

les fonctions φ_i étant continues de \mathbb{R}^n dans \mathbb{R} .

Dans les deux théorèmes ci-dessous, J est supposée continue sur $\Omega \subset \mathbb{R}^n$. Dans le premier théorème Ω est compact (fermé, borné de \mathbb{R}^n) :

THÉORÈME 3.4. (*théorème de Heine*) Si J est une fonction continue, et si Ω est un compact de \mathbb{R}^n . Alors J admet un minimum et un maximum sur Ω : il existe $\mathbf{x}_m \in \Omega$ et $\mathbf{x}_M \in \Omega$ tels que

$$J(\mathbf{x}_m) = \inf_{\mathbf{x} \in C} J(\mathbf{x}) \quad \text{et} \quad J(\mathbf{x}_M) = \sup_{\mathbf{x} \in \Omega} J(\mathbf{x}).$$

REMARQUE 3.4. On rappelle q'un compact de \mathbb{R}^n est un ensemble fermé et borné.

Dans le second théorème, Ω est non borné mais J est coercive (i.e. $\lim_{\|\mathbf{x}\| \rightarrow +\infty} J(\mathbf{x}) = +\infty$) :

THÉORÈME 3.5. Si J est une fonction continue sur Ω un fermé non borné de \mathbb{R}^n , et si $\lim_{\|\mathbf{x}\| \rightarrow +\infty} J(\mathbf{x}) = +\infty$, alors J admet un minimum sur Ω : il existe $\mathbf{x}_m \in \Omega$ tel que

$$J(\mathbf{x}_m) = \inf_{\mathbf{x} \in \Omega} J(\mathbf{x}).$$

DÉMONSTRATION. Soit $\mathbf{0} \in \Omega$, comme $J(\mathbf{x}) \rightarrow +\infty$ si $\|\mathbf{x}\| \rightarrow +\infty$ on sait qu'il existe A tel que si $\|\mathbf{x}\| > A$ alors $J(\mathbf{x}) > J(\mathbf{0})$. Soit $\Gamma = \{\mathbf{x} \in C \text{ tel que } J(\mathbf{x}) \leq J(\mathbf{0})\}$. On a que Γ est borné (si $\mathbf{x} \in \Gamma$ alors $\|\mathbf{x}\| \leq A$) et Γ est fermé (car J est continue). Γ est donc un compact de \mathbb{R}^n , et J admet un minimum sur Γ . Par construction $\inf_{\mathbf{x} \in \Gamma} J(\mathbf{x}) = \inf_{\mathbf{x} \in \Omega} J(\mathbf{x})$ et donc J admet un minimum dans Ω . \square

6. Convexité

La convexité joue un rôle important en optimisation. Voici quelques définitions.

DÉFINITION 3.4. Soit C un sous-ensemble de \mathbb{R}^n . On rappelle que si $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, le segment $[x; y]$ est défini par

$$[\mathbf{x}; \mathbf{y}] = \{\mathbf{x} + t(\mathbf{y} - \mathbf{x}), t \in [0; 1]\} = \{t\mathbf{x} + (1 - t)\mathbf{y}, t \in [0; 1]\}.$$

— L'ensemble C est **convexe** si

$$\forall (\mathbf{x}, \mathbf{y}) \in C \times C, \text{ on a } [\mathbf{x}, \mathbf{y}] \subset C.$$

— La **fonction** J définie sur un ensemble convexe est **convexe** si

$$\forall (\mathbf{x}, \mathbf{y}) \in C \times C, \forall t \in [0, 1] \text{ on a } J(t\mathbf{x} + (1-t)\mathbf{y}) \leq tJ(\mathbf{x}) + (1-t)J(\mathbf{y}).$$

— La **fonction** J définie sur un ensemble convexe est **strictement convexe** si

$$\forall (\mathbf{x}, \mathbf{y}) \in C \times C, \mathbf{x} \neq \mathbf{y}, \forall t \in]0, 1[\text{ on a } J(t\mathbf{x} + (1-t)\mathbf{y}) < tJ(\mathbf{x}) + (1-t)J(\mathbf{y}).$$

Le résultat suivant montre l'impact de la convexité dans les problèmes d'optimisation.

THÉORÈME 3.6. *Soit J une fonction convexe définie sur $C \subset \mathbb{R}^n$ ensemble convexe. Alors*

— tout minimum local de J sur C est un minimum global.

— si J est strictement convexe, il y a au plus un minimum global.

DÉMONSTRATION. (a) Soit \mathbf{x}_0 un minimum local : il existe $r > 0$ tel que $B := B(\mathbf{x}_0, r) \subset C$ et

$$\forall \mathbf{x} \in B, \quad J(\mathbf{x}) \geq J(\mathbf{x}_0). \quad (6.1)$$

Soit $\mathbf{y} \in C$, et g définie par

$$g(t) = J(t\mathbf{y} + (1-t)\mathbf{x}_0).$$

Soit $\mathbf{z} \in \partial B \cap [\mathbf{x}_0, \mathbf{y}]$. Comme $\mathbf{z} \in [\mathbf{x}_0, \mathbf{y}]$, il existe $t_0 \in [0, 1]$ tel que $\mathbf{z} = t_0\mathbf{y} + (1-t_0)\mathbf{x}_0$. La relation (2.1) implique que

$$\forall t \in [0, t_0], \quad g(t) \geq J(\mathbf{x}_0).$$

En utilisant cette dernière égalité et le caractère convexe de J nous avons

$$\forall t \in [0, t_0], \quad J(\mathbf{x}_0) \leq g(t) \leq tJ(\mathbf{y}) + (1-t)J(\mathbf{x}_0),$$

d'où $t(J(\mathbf{y}) - J(\mathbf{x}_0)) \geq 0, \forall t \in [0, t_0]$, et donc $J(\mathbf{y}) \geq J(\mathbf{x}_0)$.

(b) Soit J strictement convexe. On suppose que J admet deux minima globaux distincts \mathbf{x} et \mathbf{y} . Bien sûr, $J(\mathbf{x}) = J(\mathbf{y}) = \inf_C J$. Soit alors $t \in]0, 1[$ et $\mathbf{z} = t\mathbf{x} + (1-t)\mathbf{y}$. On a $J(\mathbf{z}) < tJ(\mathbf{x}) + (1-t)J(\mathbf{y}) = \inf_C J$. Ce qui est absurde. \square

Pour vérifier qu'une fonction est convexe, les caractérisations suivantes peuvent être utiles :

PROPOSITION 3.1. *Soit C un ensemble convexe de \mathbb{R}^n , et J une fonction définie sur C .*

(1) Si J est une fonction différentiable définie sur un convexe C de \mathbb{R}^n . Alors

$$J \text{ est convexe} \Leftrightarrow \forall (\mathbf{x}, \mathbf{y}) \in C \times C, \quad \nabla J(\mathbf{x}) \cdot (\mathbf{y} - \mathbf{x}) \leq J(\mathbf{y}) - J(\mathbf{x}).$$

(2) Si J est une fonction de classe \mathcal{C}^2 sur C de \mathbb{R}^n . Alors

$$J \text{ est convexe} \Leftrightarrow \forall (\mathbf{x}, \mathbf{y}) \in C \times C, \quad H(\mathbf{x})\mathbf{y} \cdot \mathbf{y} \geq 0.$$

Autrement dit la forme quadratique $H(\mathbf{x}) = D^2J(\mathbf{x})$ est positive.

DÉMONSTRATION. (1) (a) Si J est convexe., soit $\theta \in [0, 1]$ et $\mathbf{x}, \mathbf{y} \in C$. Comme C est convexe, $\mathbf{x} + \theta(\mathbf{y} - \mathbf{x}) \in C$, et la convexité de J donne $J(\mathbf{x} + \theta(\mathbf{y} - \mathbf{x})) \leq (1 - \theta)J(\mathbf{x}) + \theta J(\mathbf{y})$. D'où

$$J(\mathbf{y}) - J(\mathbf{x}) \geq \frac{J(\mathbf{x} + \theta(\mathbf{y} - \mathbf{x})) - J(\mathbf{x})}{\theta}.$$

On conclut en faisant tendre θ vers 0.

(b) Réciproquement, si $J(\mathbf{b}) \geq J(\mathbf{a}) + \nabla J(\mathbf{a}) \cdot (\mathbf{b} - \mathbf{a})$ pour tout $\mathbf{a}, \mathbf{b} \in C$, alors avec $\mathbf{b} = \mathbf{y}$ et $\mathbf{a} = \mathbf{x} + \theta(\mathbf{y} - \mathbf{x})$, puis avec $\mathbf{b} = \mathbf{x}$ et $\mathbf{a} = \mathbf{x} + \theta(\mathbf{y} - \mathbf{x})$, on a $J(\mathbf{y}) \geq J(\mathbf{x} + \theta(\mathbf{y} - \mathbf{x})) + (1 - \theta)\nabla J(\mathbf{x} + \theta(\mathbf{y} - \mathbf{x})) \cdot (\mathbf{y} - \mathbf{x})$ et $J(\mathbf{x}) \geq J(\mathbf{x} + \theta(\mathbf{y} - \mathbf{x})) - \theta\nabla J(\mathbf{x} + \theta(\mathbf{y} - \mathbf{x})) \cdot (\mathbf{y} - \mathbf{x})$. On multiplie la première inégalité par θ , la deuxième par $1 - \theta$ et on fait la somme :

$$\theta J(\mathbf{y}) + (1 - \theta)J(\mathbf{x}) \geq J((1 - \theta)\mathbf{x} + \theta\mathbf{y}).$$

(2) (a) Si J est convexe, alors, pour tout $\mathbf{h} \in \mathbb{R}^n$ et pour tout t tel que $\mathbf{x} + t\mathbf{h} \in C$, $J(\mathbf{x} + t\mathbf{h}) \geq J(\mathbf{x}) + t\nabla J(\mathbf{x}) \cdot \mathbf{h}$. Or, $J(\mathbf{x} + t\mathbf{h}) = J(\mathbf{x}) + t\nabla J(\mathbf{x}) \cdot \mathbf{h} + \frac{1}{2}t^2 H(\mathbf{x})\mathbf{h} \cdot \mathbf{h} + t^2 \|\mathbf{h}\|^2 o(1)$. On a donc $H(\mathbf{x})\mathbf{h} \cdot \mathbf{h} + 2\|\mathbf{h}\|^2 o(1) \geq 0$. En faisant tendre $t \rightarrow 0$, on a $H(\mathbf{x})\mathbf{h} \cdot \mathbf{h} \geq 0$.

(b) Réciproquement, on applique la formule de Taylor-Mac-Laurin (Taylor-Lagrange) à l'ordre 2 en \mathbf{x} : il existe $\theta \in]0, 1[$ tel que :

$$J(\mathbf{y}) = J(\mathbf{x}) + \nabla J(\mathbf{x}) \cdot (\mathbf{y} - \mathbf{x}) + \frac{1}{2}H(\mathbf{x} + \theta(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x}) \cdot (\mathbf{y} - \mathbf{x}).$$

Or $\frac{1}{2}H(\mathbf{x} + \theta(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x}) \cdot (\mathbf{y} - \mathbf{x}) \geq 0$ donc $J(\mathbf{y}) \geq J(\mathbf{x}) + \nabla J(\mathbf{x}) \cdot (\mathbf{y} - \mathbf{x})$ et J est convexe. □

7. Conditions nécessaires ou suffisantes d'extremum

Nous supposons dans tout ce paragraphe que J est une ou deux fois différentiable. On note \mathbf{x}_m un minimum (local) de J . On rassemble et on reformule dans cette section les résultats vus au paravant rassemblant les liens entre extremums locaux et les conditions sur l'annulation du gradient et le signe de la matrice hessienne.

7.1. Conditions nécessaires.

THÉORÈME 3.7. Pour que \mathbf{x}_m soit un minimum (local) de J , les conditions suivantes sont nécessaires :

— Si J est différentiable en \mathbf{x}_m alors $\nabla J(\mathbf{x}_m) = \mathbf{0}$,

— Si J est deux fois différentiable en \mathbf{x}_m alors la forme quadratique $\mathbf{y} \mapsto D^2J(\mathbf{x}_m)\mathbf{y} \cdot \mathbf{y} = H(\mathbf{x}_m)\mathbf{y} \cdot \mathbf{y}$ est positive, i.e. $H(\mathbf{x}_m)\mathbf{y} \cdot \mathbf{y} \geq 0$ pour tout $\mathbf{y} \in \mathbb{R}^n$.

DÉMONSTRATION. — Soit le DL à l'ordre 1 en \mathbf{x}_m :

$$J(\mathbf{x}_m + \mathbf{h}) = J(\mathbf{x}_m) + \nabla J(\mathbf{x}_m) \cdot \mathbf{h} + \|\mathbf{h}\|o(1).$$

Soit $t \geq 0$. On choisit $\mathbf{h} = -t\nabla J(\mathbf{x}_m)$ et on pose $\alpha = \|\nabla J(\mathbf{x}_m)\|$. On a donc :

$$J(\mathbf{x}_m + \mathbf{h}) = J(\mathbf{x}_m) - t\alpha^2 + t\alpha o(1).$$

Supposons que $\alpha > 0$. Puisque $o(1) \rightarrow 0$ si $t \rightarrow 0$, il existe t^* tel que $\forall t \leq t^* : J(\mathbf{x}_m + \mathbf{h}) < J(\mathbf{x}_m)$. Ceci indique que \mathbf{x}_m n'est pas un minimum, ce qui est absurde. Donc $\alpha = 0$.

— Supposons qu'il existe $\mathbf{y} \in \mathbb{R}^n$, $\mathbf{y} \neq \mathbf{0}$ tel que $a = H(\mathbf{x}_m)\mathbf{y} \cdot \mathbf{y} < 0$. Alors le DL à l'ordre 2 donne :

$$J(\mathbf{x}_m + t\mathbf{y}) = J(\mathbf{x}_m) + \frac{1}{2}t^2a + t^2\|\mathbf{y}\|^2o(1).$$

Puisque $o(1) \rightarrow 0$ si $t \rightarrow 0$, il existe t^* tel que $\forall t \leq t^* : J(\mathbf{x}_m + t\mathbf{y}) < J(\mathbf{x}_m)$. Ceci indique que \mathbf{x}_m n'est pas un minimum, ce qui est absurde. Donc $a \geq 0$. \square

7.2. Conditions suffisantes.

THÉORÈME 3.8. *Soit J une fonction de classe \mathcal{C}^1 sur \mathbb{R}^n . On suppose que $\nabla J(\mathbf{x}_m) = \mathbf{0}$ et que J est deux fois différentiable en \mathbf{x}_m . Alors \mathbf{x}_m est un minimum (local) de J si l'une des deux conditions suivantes est vérifiée :*

- (i) $H(\mathbf{x}_m) = D^2J(\mathbf{x}_m)$ est définie positive,
- (ii) Il existe $r > 0$ tel que J est deux fois différentiable sur $B(\mathbf{x}_m, r)$ et la forme quadratique $H(\mathbf{x}) = D^2J(\mathbf{x})$ est positive pour tout $\mathbf{x} \in B(\mathbf{x}_m, r)$.

DÉMONSTRATION. Pour (i), voir [Théorème 3.3](#). Pour (ii), on considère la formule de Taylor-Mac-Laurin (Taylor-Lagrange) en \mathbf{x}_m à l'ordre 2. Pour tout $\mathbf{h} \in B(\mathbf{0}, r)$, il existe $\lambda \in [0, 1]$ (qui dépend de \mathbf{h}) tel que

$$J(\mathbf{x}_m + \mathbf{h}) = J(\mathbf{x}_m) + \frac{1}{2}D^2J(\mathbf{x}_m + \lambda\mathbf{h})\mathbf{h} \cdot \mathbf{h} \geq J(\mathbf{x}_m)$$

ce qui montre bien que \mathbf{x}_m est un minimum. \square

Dans le cas où J est **convexe**, la condition suffisante s'exprime beaucoup plus facilement :

PROPOSITION 3.2. *Soit J une fonction convexe de classe \mathcal{C}^1 sur \mathbb{R}^n et \mathbf{x}_m un point de \mathbb{R}^n . Alors \mathbf{x}_m est un minimum (global) de J si et seulement si $\nabla J(\mathbf{x}_m) = \mathbf{0}$.*

DÉMONSTRATION. Il suffit de montrer que la condition est suffisante. Soit \mathbf{y} quelconque dans \mathbb{R}^n . D'après Proposition 3.1, nous avons

$$J(\mathbf{y}) - J(\mathbf{x}_m) \geq \nabla J(\mathbf{x}_m) \cdot (\mathbf{y} - \mathbf{x}_m) = 0$$

ce qui montre bien que \mathbf{x}_m est minimum global. □

8. Optimisation sous contraintes

8.1. Cas d'une fonction convexe. Dans cette situation, nous avons le résultat suivant :

PROPOSITION 3.3. *Soit J une fonction convexe de classe \mathcal{C}^1 sur un ensemble convexe $C \subset \mathbb{R}^n$, et $\mathbf{x}_m \in C$. Alors \mathbf{x}_m est un minimum (global) de J sur C si et seulement si*

$$\forall \mathbf{y} \in C, \quad \nabla J(\mathbf{x}_m) \cdot (\mathbf{y} - \mathbf{x}_m) \geq 0.$$

DÉMONSTRATION. La condition suffisante se prouve en utilisant Proposition 3.1. La condition nécessaire se prouve en utilisant le DL à l'ordre 1 pour t petit :

$$J(\mathbf{x}_m + t(\mathbf{y} - \mathbf{x}_m)) = J(\mathbf{x}_m) + t\nabla J(\mathbf{x}_m) \cdot (\mathbf{y} - \mathbf{x}_m) + t\|\mathbf{y} - \mathbf{x}_m\|o(1) \geq J(\mathbf{x}_m).$$

Donc $\nabla J(\mathbf{x}_m) \cdot (\mathbf{y} - \mathbf{x}_m) + \|\mathbf{y} - \mathbf{x}_m\|o(1) \geq 0$, $\forall t > 0$ et assez petit. Faisant $t \rightarrow 0$, on trouve bien le résultat. □

8.2. Un exemple. On s'intéresse à l'exemple suivant.

EXEMPLE 3.2. Maximiser l'aire d'une parcelle rectangulaire de périmètre P donné.

Pour cela notons x et y les longueurs et largeur du rectangle. Le périmètre est alors $2(x + y)$ et l'aire xy . Posons $g(x, y) = 2(x + y) - P$ et $f(x, y) = xy$. On souhaite alors maximiser $f(x, y)$ sous la contrainte $g(x, y) = 0$.

D'une manière plus générale on voudra s'intéresser à résoudre le problème suivant : Maximiser ou minimiser $f(x_1, \dots, x_n)$ sous la contrainte $g_1(x_1, \dots, x_n) = 0, g_2(x_1, \dots, x_n) = 0, \dots, g_m(x_1, \dots, x_n) = 0$ où f et les g_i seront des fonctions différentiables de \mathbb{R}^n dans \mathbb{R} .

8.3. Idée de sous-variété. Soit g_1, \dots, g_m m fonctions différentiables sur \mathbb{R}^n . L'ensemble

$$S = \{(x_1, \dots, x_n); g_1(x_1, \dots, x_n) = \dots = g_m(x_1, \dots, x_n) = 0\}$$

est une sous-variété dès que le rang de $(\nabla g_1, \dots, \nabla g_m)$ est de rang m en tout point de S . L'entier $n - m$ est appelé dimension de la sous-variété. Si $n - m = 1$ on parle de courbe, si $n - m = 2$ on parle de surface.

REMARQUE 3.5. C'est un cas particulier d'une définition plus générale de sous-variété.

EXEMPLE 3.3. Le cercle donné par l'équation $x^2 + y^2 = 1$ est une courbe du plan \mathbb{R}^2 . Cela correspond au cas $n = 2, m = 1$ et $g_1(x, y) = x^2 + y^2 - 1$.

Une sous-variété peut aussi se *paramétrer*, c'est-à-dire que si S est la sous-variété ci-dessous alors au voisinage de n'importe quel point $\mathbf{x} \in S$, il existe $\Phi: U \subset \mathbb{R}^{n-m} \rightarrow \mathbb{R}^n$ différentiable telle que $\varphi(U)$ soit un voisinage de \mathbf{x} dans S . Un énoncé précis s'appelle le *théorème des fonctions implicites*.

EXEMPLE 3.4. La sphère unité dans \mathbb{R}^3 est donnée par l'équation $x^2 + y^2 + z^2 - 1 = 0$. On peut aussi la paramétrer par

$$\Phi(\theta, \varphi) = (\cos(\theta) \sin(\varphi), \sin(\theta) \sin(\varphi), \cos(\varphi))$$

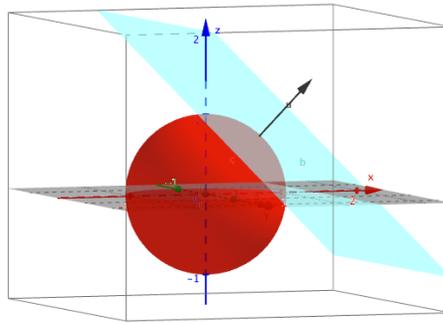
pour $\theta \in]-\pi, \pi[$ et $\varphi \in]0, \pi[$ autour du point $(1, 0, 0)$. On remarque que ce paramétrage ne donne pas tous les points de la sphère, il faut introduire un autre paramétrage autour du pôle nord par exemple.

DÉFINITION 3.5. L'espace tangent à une sous-variété S au point \mathbf{x} est le plan affine passant par \mathbf{x} et de direction vectorielle $\cap_{i=1}^n \ker Dg_i(\mathbf{x})$.

Par définition, cet espace affine est de dimension $n - m$ et c'est aussi l'ensemble des vecteurs tangents en \mathbf{x} à des courbes différentiables incluses dans S et passant par \mathbf{x} . La direction vectorielle $\cap_{i=1}^n \ker Dg_i(\mathbf{x})$ s'interprète aussi comme l'orthogonal à l'espace vectoriel engendré par les ∇g_i car

$$\cap_{i=1}^n \ker Dg_i(\mathbf{x}) = \cap_{i=1}^n (\nabla g_i)^\perp = (\text{Vect}(\nabla g_1, \dots, \nabla g_m))^\perp.$$

EXEMPLE 3.5. Soit S la sphère unité. Au point de coordonnées $(x_0, y_0, z_0) \in S$, le gradient de $g(x, y, z) = x^2 + y^2 + z^2 - 1$ est $2(x_0, y_0, z_0)$. Le plan tangent est le plan affine d'équation $x_0x + y_0y + z_0z - 1 = 0$.



8.4. Multiplicateurs de Lagrange.

THÉORÈME 3.9. Soient J, g_i avec $i \in \{1, \dots, m\}$ des fonctions de classe C^1 . On veut minimiser J sur l'ensemble

$$C = \{\mathbf{x} \in \mathbb{R}^n, \text{ tels que } g_i(\mathbf{x}) = 0, i \in \{1, \dots, p\}\}.$$

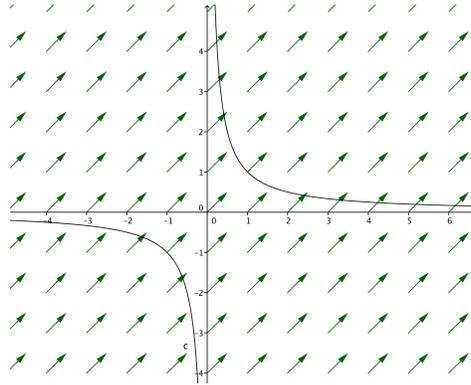
On suppose que les vecteurs $\nabla g_i(\mathbf{x}_m)$, $i \in \{1, \dots, m\}$ sont linéairement indépendants. Alors une condition nécessaire pour que \mathbf{x}_m soit un minimum de J est qu'il existe des réels $\lambda_1, \dots, \lambda_m$ tels que

$$\nabla J(\mathbf{x}_m) = \sum_{i=1}^m \lambda_i \nabla g_i(\mathbf{x}_m).$$

Les coefficients λ_i et μ_i s'appellent **multiplicateurs de Lagrange**.

REMARQUE 3.6. On a l'interprétation géométrique suivante : ∇f est orthogonal à l'espace tangent de la sous-variété définie par les g_i .

EXEMPLE 3.6. On veut minimiser le périmètre d'un rectangle sous la contrainte d'aire fixée égal à 1. On note x, y les côtés de ce rectangle. On pose $g(x, y) = xy - 1$ et $f(x, y) = 2(x + y)$. La courbe d'équation $g(x, y) = 0$ est une hyperbole. On trace cet hyperbole et la direction du gradient sur le graphique suivant



On voit sur le dessin que le champ de gradient est orthogonal à l'hyperbole si et seulement si $x = y$ et donc la solution est un carré unitaire.

9. Rôle des multiplicateurs de Lagrange

Soit $T(\mathbf{x})$ la température en un point \mathbf{x} de \mathbb{R}^3 . On veut déterminer les maxima et minima de température sur une courbe C de \mathbb{R}^3 .

On doit donc résoudre un problème d'optimisation sous contrainte

$$\min_{\mathbf{x} \in C} J(\mathbf{x})$$

avec $J(\mathbf{x}) = T(\mathbf{x})$. On décrit C comme l'intersection entre 2 surfaces d'équations $f_1(\mathbf{x}) = 0$ et $f_2(\mathbf{x}) = 0$. Nous avons donc un problème d'optimisation avec 2 contraintes d'égalité. Les 2 vecteurs gradients ∇f_1 et ∇f_2 sont normaux à ces surfaces, et donc à la courbe C (intersection des 2 surfaces).

On décrit la courbe C par un champ vectoriel $\alpha(t)$ où $t \in [a, b]$. Sur la courbe C la température devient donc une fonction de t : $g(t) = T(\alpha(t))$. Si g a un extremum relatif en un point $t^* \in]a, b[$, nous devons avoir $g'(t^*) = 0$. Or,

$$g'(t) = \nabla J(\alpha(t)) \cdot \alpha'(t).$$

Ce produit est nul si ∇J est perpendiculaire à $\alpha'(t)$. Or, $\alpha'(t)$ est tangent à C . Donc $\nabla J(\alpha(t))$ est normal à C , tout comme ∇f_1 et ∇f_2 . Si ∇f_1 et ∇f_2 sont linéairement indépendants alors ∇J est une combinaison linéaire de ∇f_1 et ∇f_2

$$\nabla J = \mu_1 \nabla f_1 + \mu_2 \nabla f_2.$$

EXEMPLE 3.7. Si $J(\mathbf{x}) = x^2 + y^2 + z^2 = \|\mathbf{x}\|^2$ sur la courbe C intersection entre le plan $S_1 = \{\mathbf{x} \mid f_1(\mathbf{x}) = z = 0\}$ et du cylindre $S_2 = \{\mathbf{x} \mid f_2(\mathbf{x}) = z^2 - (y - 1)^3 = 0\}$. La courbe C est alors la ligne $\{y = 1, z = 0\}$. Le problème $\inf_C J$ a comme solution $\mathbf{x}_m = (0, 1, 0) = \mathbf{j}$. Or $\nabla f_1(\mathbf{x}_m) = (0, 0, 1) = \mathbf{k}$ et $\nabla f_2(\mathbf{x}_m) = \mathbf{0}$. En \mathbf{x}_m , ∇f_1 et ∇f_2 ne sont donc pas linéairement indépendants. Et $\nabla J(\mathbf{x}_m) = 2\mathbf{j} \neq \mu_1 \nabla f_1(\mathbf{x}_m) + \mu_2 \nabla f_2(\mathbf{x}_m)$.

10. Exercices

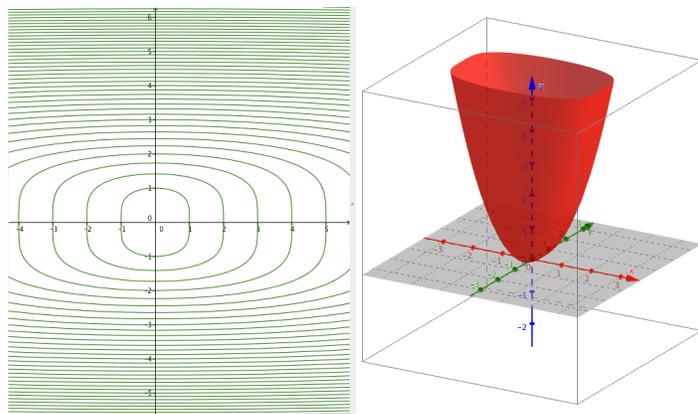
EXERCICE 3.1. Déterminer les points d'extremum local des fonctions suivantes :

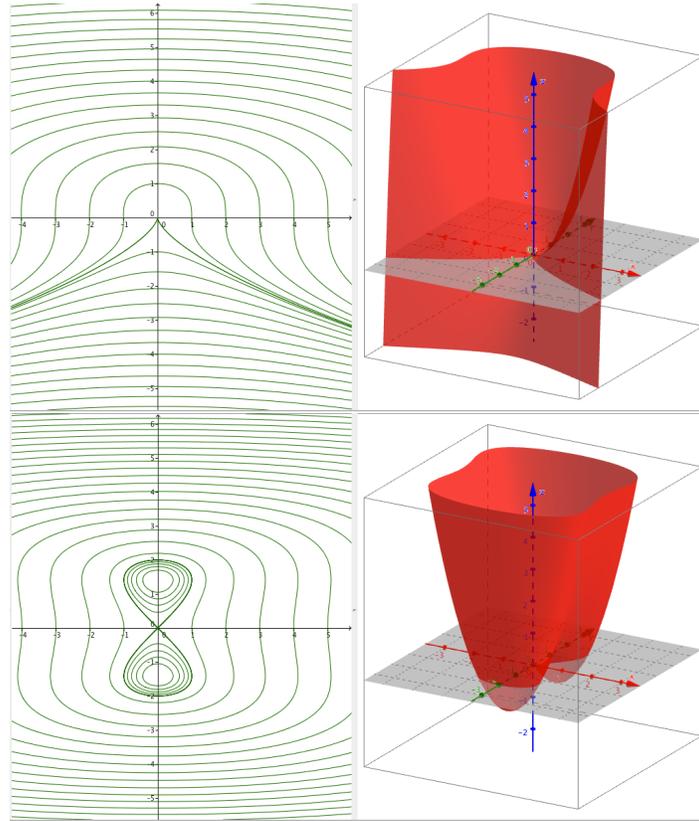
- a) $f_1 : \mathbb{R}^2 \mapsto \mathbb{R}, f_1(x, y) = x^3 + y^3 - 3xy$;
- b) $f_2 : \mathbb{R}^2 \mapsto \mathbb{R}, f_2(x, y) = x^2 y \exp(2x + 3y)$;
- c) $f_3 : \mathbb{R}^2 \mapsto \mathbb{R}, f_3(x, y) = x^4 + y^4 - 4y^3 - 3x^2 + \frac{9}{2}x$;
- d) $f_4 : D \subset \mathbb{R}^2 \mapsto \mathbb{R}, f_4(x, y) = 2xy + \frac{1}{x} + \frac{a}{y}$.

EXERCICE 3.2. Pour chacune des fonctions f suivantes, rechercher les extremums sur \mathbb{R}^2 et esquisser les courbes de niveau.

- (1) $f(x, y) = x^2 + y^4$,
- (2) $f(x, y) = x^2 + y^3$,
- (3) $f(x, y) = x^2 - y^2 + y^4/4$.

Voici les graphes et courbes de niveau de ces fonctions





EXERCICE 3.3. Montrer que $f : \mathbb{R}^2 \mapsto \mathbb{R}$, $f(x, y) = y \exp(y) - (1 + \exp(y)) \cos x$ a une infinité des points de minimums locaux mais aucun maximum.

EXERCICE 3.4. Etant donnés n points du plan \mathbb{R}^2 de coordonnées (x_i, y_i) , non alignés sur une même droite verticale, montrer qu'il existe λ et μ uniques tels que la somme $\sum_{i=1}^n (\lambda x_i + \mu - y_i)^2$ soit minimale.

La droite d'équation $y = \lambda x + \mu$ est appelée *droite des moindres carrés*.

EXERCICE 3.5. Soient $f : A \subset \mathbb{R}^2 \mapsto \mathbb{R}$ de classe \mathcal{C}^1 sur A et $\mathbf{x} \in A$. Trouver le vecteur $\mathbf{v} \in \mathbb{R}^2$ tel que $\|\mathbf{v}\| = 1$ et $Df(\mathbf{x})(\mathbf{v})$ est maximal/minimal.

EXERCICE 3.6. Déterminer les points d'extremum local :

- de la fonction $y : D \subset \mathbb{R} \mapsto \mathbb{R}$ définie par $x^3 + y^3 - 6xy = 0$ où $D = \{x \in \mathbb{R} \mid y^2 - 2x \neq 0\}$;
- de la fonction $z : D \subset \mathbb{R}^2 \mapsto \mathbb{R}$, $x^2 + y^2 + z^2 - 4z = 0$ où $D = \{(x, y) \in \mathbb{R}^2 \mid z - 2 \neq 0\}$.

EXERCICE 3.7. Maximiser le volume d'un parallélépipède rectangle de surface S donnée.

EXERCICE 3.8. Déterminer les points d'extremum local de $f(x, y, z) = x + y + z$ avec les contraintes : $x^2 + y^2 + z^2 = 1$ et $x + 2y + 2z = 1$.

EXERCICE 3.9. Soit $(t, x) \mapsto F(t, x) = f_t(x)$ une fonction de classe C^2 sur \mathbb{R}^2 . On suppose que f_0 admet un minimum local strict en un point $x = a$ avec $f_0''(a) > 0$.

- (1) Montrer que pour tout t suffisamment proche de 0, la fonction f_t admet un minimum local strict en un point $a(t)$ voisin de a et donner un développement limité au premier ordre de f_t en ce point.
- (2) Illustrer ce qui précède avec $F(t, x) = x^3/3 - (1 + t)x$.

Opérateurs différentiels de la physique

Dans la suite Ω désigne un ouvert de \mathbb{R}^n , avec $n \geq 2$ et on notera $\mathbf{x} = (x_1, \dots, x_n)$ un vecteur de \mathbb{R}^n . Pour $k \geq 1$ et $m \geq 1$, on désigne par $\mathcal{C}^k(\Omega)$ l'ensemble des fonction scalaires $f : \Omega \rightarrow \mathbb{R}$ de classe \mathcal{C}^k et par $\mathcal{C}^k(\Omega, \mathbb{R}^m)$ l'ensemble des fonctions vectorielles $F : \Omega \rightarrow \mathbb{R}^m$ de classe \mathcal{C}^k .

1. Définition et résultats

DÉFINITION 4.1. (1) Si $f \in \mathcal{C}^1(\Omega)$ on définit pour $\mathbf{x} \in \Omega$, le vecteur

$$\text{grad } f(\mathbf{x}) = \nabla f(\mathbf{x}) = \left(\frac{\partial f}{\partial x_1}(\mathbf{x}), \dots, \frac{\partial f}{\partial x_n}(\mathbf{x}) \right) \in \mathbb{R}^n \quad (1.1)$$

appelé le **gradient** de f .

(2) Si $f \in \mathcal{C}^2(\Omega)$ on définit pour $\mathbf{x} \in \Omega$, le scalaire

$$\Delta f(\mathbf{x}) = \sum_{i=1}^n \frac{\partial^2 f}{\partial^2 x_i}(\mathbf{x}) \in \mathbb{R} \quad (1.2)$$

appelé le **laplacien** de f .

(3) Si $F \in \mathcal{C}^1(\Omega, \mathbb{R}^m)$ et $\mathbf{x} \in \Omega$, avec $F(\mathbf{x}) = (F_1(x), \dots, F_m(x))$, on définit le scalaire

$$\text{div } F(\mathbf{x}) = \sum_{i=1}^m \frac{\partial F_i}{\partial x_i}(\mathbf{x}) \in \mathbb{R} \quad (1.3)$$

appelé la **divergence** de F . On peut écrire symboliquement, $\text{div } F(\mathbf{x}) = \nabla \cdot F(\mathbf{x})$, où \cdot désigne le produit scalaire usuel de \mathbb{R}^m .

(4) Si $F \in \mathcal{C}^1(\Omega, \mathbb{R}^m)$ et $\mathbf{x} \in \Omega$, avec $F(\mathbf{x}) = (F_1(\mathbf{x}), \dots, F_m(\mathbf{x}))$, on définit le vecteur

$$\text{rot } F(\mathbf{x}) = \left((-1)^{i+j} \left(\frac{\partial F_i}{\partial x_j}(\mathbf{x}) - \frac{\partial F_j}{\partial x_i}(\mathbf{x}) \right) \right)_{1 \leq i < j \leq m} \in \mathbb{R}^{\frac{m(m-1)}{2}} \quad (1.4)$$

appelé le **rotationnel** de F .

Si $N = 2$, alors

$$\text{rot } F(\mathbf{x}) = \frac{\partial F_2}{\partial x_1}(\mathbf{x}) - \frac{\partial F_1}{\partial x_2}(\mathbf{x}) \in \mathbb{R} \quad (1.5)$$

Si $N = 3$, alors

$$\operatorname{rot} F(\mathbf{x}) = \left(\frac{\partial F_3}{\partial x_2}(\mathbf{x}) - \frac{\partial F_2}{\partial x_3}(\mathbf{x}), \frac{\partial F_1}{\partial x_3}(\mathbf{x}) - \frac{\partial F_3}{\partial x_1}(\mathbf{x}), \frac{\partial F_2}{\partial x_1}(\mathbf{x}) - \frac{\partial F_1}{\partial x_2}(\mathbf{x}) \right) \in \mathbb{R}^3 \quad (1.6)$$

Dans ce cas, nous pouvons écrire symboliquement, $\operatorname{rot} F(\mathbf{x}) = \nabla \wedge F(\mathbf{x})$, où \wedge est le produit vectoriel dans \mathbb{R}^3 .

Les propriétés suivantes sont faciles à établir.

THÉORÈME 4.1. Soit $\Omega \subset \mathbb{R}^n$.

(1) Soit $f \in \mathcal{C}^2(\Omega)$, alors

$$\operatorname{div} \operatorname{grad} f = \Delta f. \quad (1.7)$$

(2) Soient $f \in \mathcal{C}^1(\Omega)$ et $F \in \mathcal{C}^1(\Omega, \mathbb{R}^3)$ alors

$$\operatorname{rot} \operatorname{grad} f = 0 \quad (1.8)$$

$$\operatorname{grad} \operatorname{rot} F = 0 \quad (1.9)$$

(3) Soient $f \in \mathcal{C}^1(\Omega)$ et $g \in \mathcal{C}^2(\Omega)$ alors

$$\operatorname{div}(f \operatorname{grad} g) = f \Delta g + \operatorname{grad} f \cdot \operatorname{grad} g. \quad (1.10)$$

(4) Soient $f, g \in \mathcal{C}^1(\Omega)$ alors

$$\operatorname{grad}(fg) = f \operatorname{grad} g + g \operatorname{grad} f. \quad (1.11)$$

(5) Soient $f \in \mathcal{C}^1(\Omega)$ et $F \in \mathcal{C}^1(\Omega, \mathbb{R}^N)$ alors

$$\operatorname{div}(fF) = f \operatorname{div} F + F \cdot \operatorname{grad} f. \quad (1.12)$$

(6) Soit $F \in \mathcal{C}^2(\Omega, \mathbb{R}^3)$ alors

$$\operatorname{rot} \operatorname{rot} F = -\Delta F + \operatorname{grad} \operatorname{div} F, \quad (1.13)$$

où $\Delta F = (\Delta F_1, \Delta F_2, \Delta F_3)$.

(7) Soient $f \in \mathcal{C}^1(\Omega)$ et $F \in \mathcal{C}^1(\Omega, \mathbb{R}^3)$ alors

$$\operatorname{rot}(fF) = \operatorname{grad} f \wedge F + f \operatorname{rot} F. \quad (1.14)$$

2. Exemples

EXEMPLE 4.1. Soient $\mathbf{x} = (x_1, \dots, x_n)$, $\mathbf{a} = (a_1, \dots, a_n)$ et r tels que

$$r = \sqrt{\sum_{i=1}^n (x_i - a_i)^2}$$

Soit $f(r) = \frac{1}{r}$. Calculons $F = \operatorname{grad} f$, Δf et $\operatorname{div} F$.

(1) Pour tout $i = 1, \dots, n$, on a

$$\frac{\partial f}{\partial x_i} = \frac{\partial}{\partial x_i}(r^{-1}) = -r^{-2} \frac{\partial}{\partial x_i}(r) = \frac{x_i - a_i}{r^3}$$

Donc

$$F(\mathbf{x}) = \text{grad } f(\mathbf{x}) = -\frac{1}{r^3}(x_1 - a_1, \dots, x_n - a_n) = -\frac{1}{r^3}(x - a).$$

(2) Pour tout $i = 1, \dots, n$, on a

$$\begin{aligned} \frac{\partial^2 f}{\partial x_i^2} &= \frac{\partial}{\partial x_i} \left(-\frac{1}{r^3}(x - a) \right) = -\frac{1}{r^6} \left(r^3 - 3(x_i - a_i)r^2 \frac{\partial r}{\partial x_i} \right) \\ &= -\frac{1}{r^4} \left(r - 3(x_i - a_i) \frac{1}{r}(x_i - a_i) \right) = \frac{1}{r^5} (3(x_i - a_i)^2 - r^2) \end{aligned}$$

Donc

$$\Delta f = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2} = \frac{1}{r^5} (3r^2 - nr^2) = \frac{3-n}{r^3}.$$

On remarque que si $n = 3$, alors $\Delta f = 0$.

(3) Pour tout $i = 1, \dots, n$, on a $F_i = \frac{\partial f}{\partial x_i}$, donc

$$\frac{\partial F_i}{\partial x_i} = \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_i} \right) = \frac{\partial^2 f}{\partial x_i^2}$$

et

$$\text{div } F = \Delta f = -\frac{1}{r^3}(x - a)$$

on retrouve donc $\text{div } F = \text{div grad } f = \Delta f$.

(4) On a

$$\begin{aligned} \text{rot } F = \nabla \wedge F &= \begin{bmatrix} \partial_{x_1} \\ \partial_{x_2} \\ \partial_{x_3} \end{bmatrix} \wedge \begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix} \\ &= \begin{bmatrix} \partial_{x_1} \\ \partial_{x_2} \\ \partial_{x_3} \end{bmatrix} \wedge \begin{bmatrix} f_{x_1} \\ f_{x_2} \\ f_{x_3} \end{bmatrix} \\ &= \begin{bmatrix} f_{x_3 x_2} - f_{x_2 x_3} \\ f_{x_1 x_3} - f_{x_3 x_1} \\ f_{x_2 x_1} - f_{x_1 x_2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \end{aligned}$$

On retrouve donc $\text{rot } F = \text{rot grad } f = 0$.

EXEMPLE 4.2. Soit le champs de vecteurs

$$F(x, y, z) = (x^2 - e^y, \sin z, y^2/z)$$

On a

$$\text{div } F = 2x + 0 + 1 = 2x + 1$$

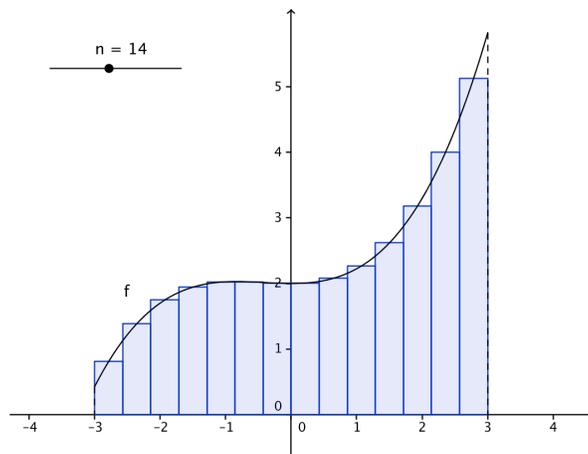
et

$$\text{rot } F = \begin{bmatrix} \partial_x \\ \partial_y \\ \partial_z \end{bmatrix} \wedge \begin{bmatrix} x^2 - e^y \\ \sin z \\ y^2 + z \end{bmatrix} = \begin{bmatrix} 2y - \cos z \\ 0 \\ e^y \end{bmatrix}$$

Intégrales sur \mathbb{R}

On veut définir une notion d'intégrale pour les fonctions continues par morceaux pour les intervalles compacts du type $[a, b]$ avec l'intuition que cela correspond à l'aire sous la courbe si la fonction est positive. Pour cela, on approche une telle fonction par des fonctions f_n en escalier (c'est-à-dire constantes par morceaux) et on calcule aisément l'aire sous la courbe pour de telles fonctions en escalier, c'est simplement une somme d'aires de rectangles notée S_n . C'est un théorème que cette suite de somme d'aires converge lorsque la suite de fonctions converge (uniformément) vers notre fonction cela permet de définir l'intégrale de f comme la limite des S_n :

$$\int_a^b f(x)dx = \lim_{n \rightarrow +\infty} S_n.$$



On donne ici quelques rappels en vrac, qui peuvent parfois vous être utiles.

Dans les hypothèses des théorèmes de cette section, on peut remplacer l'expression "pour tout x " par "pour presque tout x ". Pour "presque tout x " signifie ici pour tout x sauf éventuellement en un nombre fini ou dénombrable de points.

1. Intersion Limite-Intégrale

1.1. Théorème de convergence monotone. Les deux résultats suivants s'appliquent aux cas de fonctions positives et intégrables.

THÉORÈME 5.1. Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions **continues par morceaux, positives, intégrables** sur I un intervalle de \mathbb{R} , telles que $f = \lim_n f_n$ est une fonction continue par morceaux sur I . Alors

$$\int_I f(x) dx = \int_I \lim_n f_n(x) dx = \lim_n \int_I f_n(x) dx,$$

cette dernière égalité pouvant être égale à $+\infty$.

THÉORÈME 5.2. Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions continues par morceaux, positives, intégrables sur I un intervalle de \mathbb{R} . Alors

$$\int_I \sum_{n=0}^{\infty} f_n(x) dx = \sum_{n=0}^{\infty} \int_I f_n(x) dx,$$

et cette intégrale peut éventuellement être égale à $+\infty$.

1.2. Théorème de convergence dominée.

THÉORÈME 5.3. Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions définies sur I un intervalle de \mathbb{R} , **continues par morceaux, intégrables** sur I telles que

- (1) f_n converge vers une fonction f continue par morceaux sur I : $\lim_{n \rightarrow \infty} f_n(x) = f(x), \forall x \in I$.
- (2) Il existe une fonction g définie sur I , continue par morceaux et intégrable sur I telle que

$$\forall n \in \mathbb{N}, \forall x \in I, \quad |f_n(x)| \leq g(x).$$

Alors f est intégrable sur I et

$$\lim_{n \rightarrow \infty} \int_I f_n(x) dx = \int_I \lim_{n \rightarrow \infty} f_n(x) dx = \int_I f(x) dx.$$

THÉORÈME 5.4. Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions définies sur I un intervalle de \mathbb{R} , **continues par morceaux, et intégrables** sur I telles que $\sum_{n \in \mathbb{N}} \int_I |f_n(x)| dx < \infty$, alors la série $\sum_{n \in \mathbb{N}} f_n$ converge vers une fonction f intégrable et on a :

$$\sum_{n \in \mathbb{N}} f_n(x) = f(x), \text{ pour presque tout } x \in I$$

et

$$\int_I \sum_{n \in \mathbb{N}} f_n(x) dx = \int_I f(x) dx = \sum_{n \in \mathbb{N}} \int_I f_n(x) dx.$$

2. Interversiion Série-Limite

THÉORÈME 5.5. Si $\sum_n f_n$ est une série de fonctions définies sur I , un intervalle de \mathbb{R} et telle que $\sum_n \|f_n\| < \infty$, où $\|f_n\| = \sup_{x \in I} |f_n(x)|$, et si chaque fonction f_n admet une limite en a , a étant un point sur la frontière de I , alors on a la formule de sommation terme à terme des limites :

$$\lim_{x \rightarrow a} \sum_n f_n(x) = \sum_n \lim_{x \rightarrow a} f_n(x).$$

3. Interversiion Dérivation-Limite

THÉORÈME 5.6. Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions de classe \mathcal{C}^1 sur un intervalle I de \mathbb{R} . On suppose que

- (1) f_n converge vers f , où f est une fonction définie sur I ,
- (2) et qu'il existe une fonction g définie sur I telle que $\lim_{n \rightarrow \infty} \|f'_n - g\| = 0$, où $\|f'_n - g\| = \sup_{x \in I} |f'_n(x) - g(x)|$.

Alors f est de classe \mathcal{C}^1 sur I et $f' = g$, i.e. :

$$\frac{d \lim_n f_n}{dx} = \lim_n \frac{df_n}{dx}.$$

THÉORÈME 5.7. Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions de classe \mathcal{C}^1 sur un intervalle I de \mathbb{R} , on suppose que

- (1) $\sum_{n \in \mathbb{N}} f_n$ converge sur I , vers une fonction S ,
- (2) et que $\sum_{n \in \mathbb{N}} \|f'_n\| < \infty$ où $\|f'_n\| = \sup_{x \in I} |f'_n(x)|$.

Alors la somme $S = \sum_{n \in \mathbb{N}} f_n$ est de classe \mathcal{C}^1 sur I et $S' = \sum_{n \in \mathbb{N}} f'_n$, i.e. :

$$\frac{d \sum_n f_n}{dx} = \sum_n \frac{df_n}{dx}.$$

On a donc une formule de dérivation terme à terme des dérivées.

4. Intégrales à paramètres

4.1. Continuité sous le signe intégrale.

THÉORÈME 5.8. Soient I et J deux intervalles de \mathbb{R} . Soit f définie sur $I \times J$ à valeurs dans \mathbb{R} ou \mathbb{C} ,

$$f : \begin{cases} I \times J \rightarrow \mathbb{R} \text{ ou } \mathbb{C} \\ (x, t) \mapsto f(x, t) \end{cases}$$

telle que

- pour tout $x \in I$, la fonction $t \rightarrow f(x, t)$ est continue sur J ,
- il existe une fonction φ définie et intégrable sur I , telle que pour tout $t \in J$ et tout $x \in J$, on a $|f(x, t)| \leq \varphi(x)$.

Alors la fonction $F : t \rightarrow \int_I f(x, t) dx$ est définie et continue sur J .

4.2. Dérivabilité sous le signe intégrale.

THÉORÈME 5.9. *On reprend les notations du théorème précédent. Si*

- *pour tout $x \in I$, la fonction $t \rightarrow f(x, t)$ est dérivable sur J ,*
- *pour tout $t \in J$, la fonction $x \rightarrow f(x, t)$ est intégrable sur I ,*
- *il existe une fonction φ définie et intégrable sur I , telle que pour tout $t \in J$ et tout $x \in I$, on a $|\frac{\partial f}{\partial t}(x, t)| \leq \varphi(x)$.*

Alors la fonction $F : t \rightarrow \int_I f(x, t)dx$ est dérivable sur J et

$$F'(t) = \int_I \frac{\partial f}{\partial t}(x, t)dx.$$

REMARQUE 5.1. *La remarque du début de cette section est toujours valable, pour x . Mais pour le paramètre t , il faut bien garder l'hypothèse "pour tout t " dans ces deux derniers théorèmes.*

Si les limites d'intégration ne sont pas fixes mais elles dépendent aussi de t on obtient le résultat suivant.

THÉORÈME 5.10. *Soit f une fonction continue de \mathbb{R}^2 dans \mathbb{R} et possédant une dérivée partielle continue sur \mathbb{R} par rapport à la première variable. Soient a et b deux fonctions dérivables de \mathbb{R} dans \mathbb{R} , si F est l'intégrale paramétrique définie par :*

$$F(t) = \int_{a(t)}^{b(t)} f(t, x)dx$$

on montre que F est dérivable sur tout \mathbb{R} et que

$$F'(t) = f(t, b(t))b'(t) - f(t, a(t))a'(t) + \int_{a(t)}^{b(t)} \frac{\partial}{\partial t} f(t, x)dx$$

EXERCICE 5.1. Soient $f : [0, 1] \times (0, \infty) \mapsto \mathbb{R}$, $f(x, y) = \frac{x}{y^2} e^{-\left(\frac{x}{y}\right)^2}$ et l'intégrale à paramètre

$$F(y) = \int_0^1 f(x, y)dx.$$

a) Calculer $\lim_{y \rightarrow 0} \int_0^1 f(x, y)dx$

b) Calculer $\int_0^1 \lim_{y \rightarrow 0} f(x, y)dx$

Commenter le résultat.

EXERCICE 5.2. Soient $f(x, y) = \begin{cases} \ln \sqrt{x^2 + y^2}, & (x, y) \in [0, 1] \times (0, \infty) \\ \ln x, & (x, y) \in [0, 1] \times \{0\} \end{cases}$ et l'intégrale à paramètre

$$F(y) = \begin{cases} \int_0^1 f(x, y)dx, & y > 0 \\ -1, & y = 0 \end{cases}.$$

a) Montrer que F est continue. $\lim_{y \rightarrow 0} \int_0^1 f(x, y) dx$

b) Calculer $F'(0)$.

c) Calculer $\int_0^1 \frac{\partial f}{\partial y}(x, y) dx$

Commenter le résultat.

EXERCICE 5.3. Soit $F(t) = \int_0^t \frac{\ln(1+tx)}{1+x^2} dx$, $t > 0$. Calculer $F'(t)$.

Intégrales multiples I

1. Introduction

Nous nous intéressons dans cette partie à l'extension de la notion d'intégrale. Plutôt qu'un intervalle $[a; b]$, nous considérons un domaine bidimensionnel Q , appelé la région (ou domaine) d'intégration. L'intégrande est un champ scalaire f défini et borné sur Q . L'intégrale qui en résulte est appelée **intégrale double** et on la note

$$\iint_Q f(x, y) dx dy.$$

2. Intégrabilité des fonctions continues

Dans le cas des fonctions continues sur un rectangle, on a le droit d'**échanger l'ordre des deux intégrations** : c'est le **théorème de Fubini**.

THÉORÈME 6.1. *Soit f une fonction **continue** sur un rectangle $Q = [a; b] \times [c; d]$, alors f est **intégrable** sur Q . De plus, on a*

$$\iint_Q f(x, y) dx dy = \int_a^b \left(\int_c^d f(x, y) dy \right) dx = \int_c^d \left(\int_a^b f(x, y) dx \right) dy.$$

EXERCICE 6.1. Montrer que, si $Q = [-1; 1] \times [0; 2\pi]$, alors,

$$\iint_Q (x \sin y - ye^x) dx dy = 2 \left(\frac{1}{e} - e \right) \pi^2.$$

EXERCICE 6.2. Montrer que, si $Q = [-1; 1] \times [0; 2]$, alors nous avons

$$\iint_Q f(x, y) dx dy = \frac{4}{3} + \frac{\pi}{2},$$

avec $f(x, y) = \sqrt{|y - x^2|}$.

REMARQUE 6.1. Si f n'est pas continue sur Q , ni même bornée sur Q , on peut avoir $\int_c^d \left(\int_a^b f(x, y) dx \right) dy \neq \int_a^b \left(\int_c^d f(x, y) dy \right) dx$.

REMARQUE 6.2. Soit $Q = [0, 1] \times [0, 1]$, et $f(x, y) = \frac{x^2 - y^2}{(x^2 + y^2)^2}$. On a f non bornée au voisinage de $(0, 0)$, car $f(x, 0) = 1/x^2 \rightarrow +\infty$ lorsque $x \rightarrow 0$. Donc f n'est pas continue

sur Q , et le théorème de Fubini n'est pas vérifié. En effet : $\int_0^1 \left(\int_0^1 f(x, y) dx \right) dy = -\pi/4$ et $\int_0^1 \left(\int_0^1 f(x, y) dy \right) dx = \pi/4$.

3. Intégrales doubles étendues à des régions plus générales

Nous supposons maintenant que l'on veut intégrer une fonction f **continue** sur une région d'intégration S définie par

$$S = \{(x, y) / a \leq x \leq b, \varphi_1(x) \leq y \leq \varphi_2(x)\},$$

où φ_1 et φ_2 sont des fonctions **continues** sur un intervalle fermé $[a; b]$ satisfaisant $\varphi_1 \leq \varphi_2$.

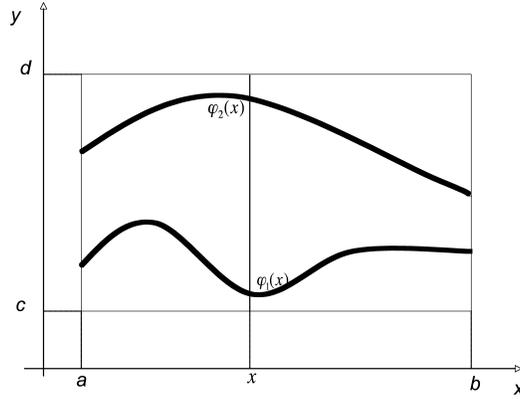


FIGURE 1 – Dans cet exemple de région, pour chaque point $x \in [a; b]$, la ligne verticale d'abscisse x intersecte S en un segment joignant la courbe $y = \varphi_1(x)$ à $y = \varphi_2(x)$. Une telle région est bornée puisque φ_1 et φ_2 sont continues et ainsi bornées sur $[a; b]$.

Comme φ_1 et φ_2 sont continues, S est une région bornée.

Enfermons S dans un rectangle Q . Par exemple, en prenant c un minorant de φ_1 et d un majorant de φ_2 , alors $S \subset Q = [a, b] \times [c, d]$.

Définissons \tilde{f} sur Q comme : $\tilde{f} = f$ sur S , $\tilde{f} = 0$ sur Q privé de S .

Supposons que la fonction f est intégrable sur S , alors \tilde{f} l'est également sur Q , et l'on a

$$\iint_Q \tilde{f}(x, y) dx dy = \iint_S f(x, y) dx dy.$$

Pour tout $x \in [a, b]$, l'application partielle $y \mapsto \tilde{f}(x, y)$ est intégrable sur $[c, d]$ d'après les hypothèses faites sur f et l'on a :

$$\int_c^d \tilde{f}(x, y) dy = \int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y) dy,$$

puisque \tilde{f} est nulle sur les intervalles $[c, \varphi_1(x)[,]\varphi_2(x), d]$. On peut donc calculer l'intégrale de \tilde{f} sur $[c, d]$:

$$\iint_Q \tilde{f}(x, y) dx dy = \int_a^b \left(\int_c^d \tilde{f}(x, y) dy \right) dx.$$

On obtient alors

$$\iint_S f(x, y) dx dy = \int_a^b \left(\int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y) dy \right) dx.$$

On a alors démontré le théorème suivant :

THÉORÈME 6.2. *Soit S une région, définie par*

$$S = \{(x, y) / a \leq x \leq b, \varphi_1(x) \leq y \leq \varphi_2(x)\},$$

située entre les graphes de φ_1 et φ_2 deux fonctions continues. Supposons que f est définie et continue sur S . Alors, l'intégrale double $\iint_S f(x, y) dx dy$ existe et peut être évaluée par

$$\iint_S f(x, y) dx dy = \int_a^b \left[\int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y) dy \right] dx.$$

Un résultat analogue peut être obtenu pour les régions de type :

$$T = \{(x, y) / c \leq y \leq d, \psi_1(y) \leq x \leq \psi_2(y)\},$$

où ψ_1 et ψ_2 sont continues sur un intervalle $[c, d]$ avec $\psi_1 \leq \psi_2$. Dans ce cas, les lignes horizontales intersectent T en segments, et l'on a :

$$\iint_S f(x, y) dx dy = \int_c^d \left[\int_{\psi_1(y)}^{\psi_2(y)} f(x, y) dx \right] dy.$$

Si nous avons une région qui est à la fois de type S et T , nous pouvons écrire

$$\int_a^b \left[\int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y) dy \right] dx = \int_c^d \left[\int_{\psi_1(y)}^{\psi_2(y)} f(x, y) dx \right] dy.$$

On voit donc, qu'avant d'intégrer, il est bénéfique d'analyser le domaine sur lequel on intègre.

EXERCICE 6.3. Calculer l'intégrale double

$$I = \iint_S (x^2 - y^2) dx dy,$$

S étant le disque elliptique

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 \leq 0.$$

EXERCICE 6.4. Calculer l'intégrale double : $I = \iint_D y dx dy$, avec $D = \{(x, y) \in \mathbb{R}^2; \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq r^2; y > c; a > 0, b > 0, r > 0, c \in \mathbb{R}\}$.

EXERCICE 6.5. Calculer l'intégrale double

$$I = \iint_D \frac{2y}{(1+x^2+y^2)^2} dx dy,$$

avec $D = \{(x, y) \in \mathbb{R}^2; 0 \leq x \leq y \leq 1\}$.

4. Interprétation géométrique des intégrales doubles

Soit $f : U \subset \mathbb{R}^2 \mapsto \mathbb{R}$, une fonction continue sur une partie U délimitée par une courbe \mathcal{C} . Le graphe de f est alors une surface S de \mathbb{R}^3 . Soit V le **volume** délimité par S , U , et les parallèles à l'axe $(0, z)$ menées par \mathcal{C} , alors $V = \iint_S f dS = \iint_U f(x, y) dx dy$.

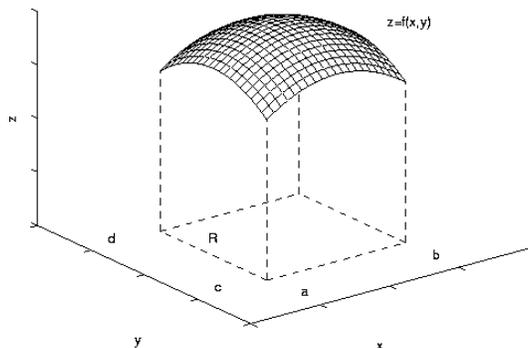


FIGURE 2 – Interprétation géométrique de $\iint_Q f(x, y) dx dy$ avec Q un rectangle.

EXERCICE 6.6. Calculer le volume V du solide défini par l'ellipsoïde d'équation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

et montrer que $V = \frac{4}{3}\pi abc$.

Intégrales curvilignes

1. Introduction

L'intégrale

$$\int_a^b f(x) dx,$$

est l'intégrale de f le long du segment rectiligne $[a; b]$. Ce segment est paramétré par x , x variant de a à b . Elle représente par exemple la masse du segment $[a; b]$, $f(x)$ étant la densité au point x .

Nous voulons étendre cette notion au cas où le segment $[a; b]$ est remplacé par une courbe dans un espace n -dimensionnel décrite par un champ vectoriel α , et l'intégrande est un champ de vecteurs \mathbf{f} défini et borné sur cette courbe.

2. Rappels sur les courbes

Rappelons quelques définitions.

DÉFINITION 7.1. — Soit $\alpha : J = [a; b] \rightarrow \mathbb{R}^n$ un champ vectoriel défini sur un intervalle J . Lorsque t parcourt J , $\alpha(t)$ décrit un ensemble de points appelé **graphe de la fonction**.

- Si α est continue sur J , le graphe est appelé une **courbe** (décrite par α) et α est appelée **chemin continu** dans \mathbb{R}^n . α , défini sur J , est appelé **paramétrage** ou **paramétrisation** de la courbe.
- La courbe est **régulière** si la dérivée α' existe, ne s'annule pas sur J , et est continue dans l'intervalle ouvert $]a; b[$.
- La courbe est **régulière par morceaux** si l'intervalle $[a; b]$ peut être partitionné en un nombre fini de sous-intervalles où la courbe est régulière.
- Une courbe est **simple** si pour tout $t_1 \neq t_2$ dans $]a; b[$, $\alpha(t_1) \neq \alpha(t_2)$.
- Une courbe simple est **fermée** si $\alpha(a) = \alpha(b)$.

Donnons aussi la définition d'orientation. On en donne une définition plus intuitive que rigoureuse...

DÉFINITION 7.2. Soit Γ une courbe fermée. On dira qu'elle est **orientée positivement** si, en se déplaçant sur Γ , on a l'intérieur de Γ à notre gauche. Si Γ est orientée positivement, on la note Γ^+ . Si elle est orientée négativement (i.e. dans le sens contraire), on la note Γ^- .

EXEMPLE 7.1. Si Γ est le cercle trigonométrique, Γ^+ est le cercle trigonométrique parcouru dans le sens trigonométrique (inverse des aiguilles d'une montre).

Intéressons nous maintenant à ce qui se passe si on change la paramétrisation d'une courbe.

DÉFINITION 7.3. Soit α un chemin continu défini sur $[a; b]$, $u : [c, d] \rightarrow [a, b]$ une fonction **bijective, différentiable** telle que u' **ne s'annule pas** sur $[c; d]$. Alors, la fonction β définie sur $[c; d]$ par l'équation $\beta(t) = \alpha[u(t)]$ est un chemin continu ayant le même graphe que α . Deux chemins α et β ainsi reliés sont dits **équivalents**.

Soit C la courbe commune de deux chemins équivalents α et β . Si la dérivée de u est toujours positive sur $[c; d]$, α et β tracent C dans le même sens (u préserve l'orientation). Si elle est toujours négative, u renverse l'orientation.

Enfin, définissons une mesure particulière : l'abscisse curviligne.

DÉFINITION 7.4. Soit α un chemin avec α' continue sur $[a; b]$. Alors, l'**abscisse curviligne** est donnée par

$$s(t) = \int_a^t \|\alpha'(x)\| dx$$

sa dérivée est : $s'(t) = \|\alpha'(t)\|$.

La courbe est dite **rectifiable** si elle est de longueur finie, i.e. si $\int_a^b \|\alpha'(u)\| du < \infty$.

La **longueur de l'arc** entre deux points $M_1 = \alpha(t_1)$ et $M_2 = \alpha(t_2)$, $t_1, t_2 \in [a, b]$ est donnée par $|s(t_2) - s(t_1)|$.

Le **vecteur tangent unitaire** à la courbe en $\alpha(t)$ est défini par

$$\mathbf{T}(t) = \frac{d\alpha}{ds} = \frac{\alpha'(t)}{\|\alpha'(t)\|}$$

REMARQUES 7.1. (1) L'abscisse curviligne, la longueur d'arc et le vecteur tangent unitaire à la courbe ne dépendent pas du paramétrage α choisi. Heureusement, entre deux points quelconques d'une courbe, la longueur d'arc est indépendante du paramétrage choisi.

En effet, soit β défini sur $[c, d]$ un autre paramétrage tel que α et β soient équivalents. Alors il existe $u : [c, d] \rightarrow [a, b]$ bijective, différentiable telle que $\beta(t) = \alpha[u(t)]$. Supposons que $u' > 0$, de telle sorte que la courbe conserve son orientation (si $u' < 0$, le raisonnement sera le même, il faudra juste échanger c et d). Soient alors $t \in [a, b]$, et $\tilde{t} = u(t)$, on a $\alpha(t) = \beta(\tilde{t})$, i.e. $t \in [a, b]$ et $\tilde{t} \in [c, d]$ correspondent au même point sur la courbe. Calculons la longueur entre $\alpha(a)$ et $\alpha(t)$: $s(t) = \int_a^t \|\alpha'(x)\| dx = \int_{u^{-1}(a)}^{u^{-1}(t)} \|\alpha'(u(y))\| u'(y) dy = \int_c^{\tilde{t}} \|\beta'(y)\| dy$. L'abscisse curviligne, i.e. la longueur, est la même entre $\alpha(a)$ et $\alpha(t)$, qu'entre $\beta(c)$ et $\beta(\tilde{t})$.

- (2) L'abscisse curviligne est un 'paramétrage'. Il s'agit d'un paramètre intrinsèque à la courbe¹.

EXERCICE 7.1. (1) Calculer la longueur de la courbe $\Gamma = \{(x, y) \in \mathbb{R}^2 : y = \cosh x, x \in [0, 1]\}$.

- (2) Soit $\Gamma' = \{(x, y) \in \mathbb{R}^2 : x(t) = r(t) \cos t, y(t) = r(t) \sin t, t \in [a, b]\}$. Calculer la longueur de Γ' en fonction de r . Calculer la longueur d'un cercle de rayon R .

- (3) Calculer la longueur de l'hélice donnée dans l'exercice 7.3.

3. Intégrales curvilignes

DÉFINITION 7.5. Soit Γ une courbe régulière par morceaux, paramétrée par $\alpha : [a, b] \rightarrow \mathbb{R}^n$. Soit \mathbf{f} un champ de vecteurs défini et borné sur Γ , à valeurs dans \mathbb{R}^n . L'intégrale curviligne de \mathbf{f} le long de Γ est définie par :

$$\int_{\Gamma} \mathbf{f} = \int_a^b \mathbf{f}(\alpha(t)) \cdot \alpha'(t) dt = \int_{\Gamma} \mathbf{f} \cdot d\alpha, \quad \text{notée aussi } \int_{\Gamma} \mathbf{f} \cdot d\Gamma \quad (3.1)$$

dès que l'intégrale de droite existe. Γ est appelée **chemin d'intégration**.

Lorsque le chemin est fermé, on utilise alors souvent le symbole \oint .

EXEMPLE 7.2. Dans le cas bidimensionnel, si la courbe Γ est paramétrée par le chemin $\alpha = (\alpha_1, \alpha_2)$ sur $[a, b]$, et si $\mathbf{f} = (P, Q)$ est un champ vectoriel à valeur dans \mathbb{R}^2 on a

$$\int_{\Gamma} \mathbf{f} \cdot d\alpha = \int_a^b \left(P(\alpha(t))\alpha_1'(t) + Q(\alpha(t))\alpha_2'(t) \right) dt.$$

Ce que l'on note aussi $\int_{\Gamma} Pdx + Qdy$, voir la fin de cette section.

REMARQUES 7.2. (1) L'intégrale curviligne est aussi appelée **travail** ou **circulation** du champ de vecteur \mathbf{f} sur la courbe Γ . Le mot *travail* évoque plutôt l'idée d'une force, et le mot *circulation* celle de la vitesse d'un élément fluide.

- (2) Pour calculer une intégrale curviligne, il faut définir un paramétrage du chemin d'intégration. Le calcul se ramène alors à celui d'une intégrale connue depuis le lycée...

- (3) Dans la notation $\int_{\Gamma} \mathbf{f}$, on ne spécifie pas de paramétrage. Effectivement, l'intégrale curviligne ne dépend pas du paramétrage choisi, pourvu que le paramétrage préserve l'orientation de la courbe. On a le théorème suivant :

1. utile essentiellement dans les exercices théoriques, dans la mesure où l'expression $s(t)$ s'exprime rarement à l'aide des fonctions usuelles de l'analyse.

THÉORÈME 7.1. Soient α et β deux chemins réguliers par morceaux et équivalents, représentant tous deux la courbe Γ . Alors, nous avons

$$\int_{\Gamma} \mathbf{f} \cdot d\alpha = \int_{\Gamma} \mathbf{f} \cdot d\beta$$

si α et β tracent Γ dans le même sens. On a

$$\int_{\Gamma} \mathbf{f} \cdot d\alpha = - \int_{\Gamma} \mathbf{f} \cdot d\beta$$

si il y a une inversion d'orientation.

DÉMONSTRATION. Soient α et β deux chemins réguliers par morceaux et équivalents, orientant Γ dans le même sens, définis respectivement sur $[a, b]$ et $[c, d]$. Alors, il existe u de classe \mathcal{C}^1 , bijection de $[a, b] \rightarrow [c, d]$ tel que $\beta(t) = \alpha[u(t)]$, pour tout $t \in [a, b]$. Effectuons, dans l'intégrale (3.1), le changement de variable $t = u(y)$. Alors $u(a) = c, u(b) = d$, d'après la croissance de u , et

$$\int_a^b \mathbf{f}(\alpha(t)) \cdot \alpha'(t) dt = \int_c^d \mathbf{f}(\alpha(u(y))) \cdot \alpha'(u(y)) u'(y) dy = \int_c^d \mathbf{f}(\beta(y)) \cdot \beta'(y) dy.$$

La première partie du théorème est démontrée. Si il y a changement d'orientation, on reproduit le même raisonnement, avec $u(a) = d, u(b) = c$, c'est pourquoi on change de signe. \square

EXERCICE 7.2. Soit \mathbf{f} le champ de vecteurs donné par

$$\mathbf{f}(x, y) = \left(\frac{\sqrt{y}}{x^3 + y} \right), \quad \forall (x, y), \quad y \geq 0.$$

- (1) Calculer l'intégrale curviligne de \mathbf{f} de $(0, 0)$ à $(1, 1)$ le long du chemin : $x = t^2, y = t^3, 0 \leq t \leq 1$. Montrer que l'intégrale vaut $59/42$.
- (2) Sur le chemin $x = t, y = t, 0 \leq t \leq 1$, montrer que l'intégrale vaut $17/12$. Conclure.
- (3) La même courbe que dans le point (1) est décrite par la paramétrisation : $\beta(t) = t\mathbf{i} + t^{3/2}\mathbf{j}$, pour $0 \leq t \leq 1$. Montrer que l'on retrouve la valeur de l'intégrale.

EXERCICE 7.3. Soit $\mathbf{f}(x, y, z) = (y, z, x)$, et l'hélice circulaire Γ paramétrée par $\alpha(t) = (r \cos t, r \sin t, ht)$ pour $t \in [0, \frac{\pi}{2}]$. Montrer que

$$\int_{\Gamma} \mathbf{f} \cdot d\alpha = \frac{\pi r}{4}(2h - r).$$

4. Propriétés de base

Nous avons les propriétés suivantes :

— **linéarité**

$$\int (\lambda \mathbf{f} + \mu \mathbf{g}) \cdot d\alpha = \lambda \int \mathbf{f} \cdot d\alpha + \mu \int \mathbf{g} \cdot d\alpha, \quad \forall \lambda, \mu \in \mathbb{R}, \text{ et } \mathbf{f}, \mathbf{g} \text{ des champs vectoriels}$$

— **additivité**

$$\int_{\Gamma} \mathbf{f} \cdot d\alpha = \int_{\Gamma_1} \mathbf{f} \cdot d\alpha + \int_{\Gamma_2} \mathbf{f} \cdot d\alpha,$$

où Γ est reconstruite par Γ_1 et Γ_2 , avec Γ_1 défini sur $[a; c]$ et Γ_2 sur $[c; b]$,

— et la propriété suivante

$$\int_{\Gamma^+} \mathbf{f} \cdot d\alpha = - \int_{\Gamma^-} \mathbf{f} \cdot d\alpha.$$

5. Intégrales de ligne par rapport à la longueur d'arc

Soit Γ une courbe régulière de longueur finie, paramétrée par α , définie sur $[a; b]$. Soit φ un **champ scalaire** défini et borné sur Γ . L'**intégrale curviligne de φ sur Γ par rapport à la longueur d'arc**, notée $\int_{\Gamma} \varphi ds$, est définie par

$$\int_{\Gamma} \varphi ds = \int_a^b \varphi(\alpha(t)) s'(t) dt$$

dès que l'intégrale de droite existe.

Considérons maintenant un champ scalaire φ donné par, $\varphi(\alpha(t)) = \mathbf{f}(\alpha(t)) \cdot \mathbf{T}(t)$, avec \mathbf{T} le vecteur tangent unitaire $\mathbf{T}(t) = \frac{d\alpha}{ds} = \frac{\alpha'(t)}{\|\alpha'(t)\|}$. Alors, nous avons

$$\int_{\Gamma} \varphi ds = \int_{\Gamma} \mathbf{f} \cdot d\alpha$$

REMARQUE 7.1. Lorsque \mathbf{f} désigne un champ de vitesse, le produit $\mathbf{f} \cdot \mathbf{T}$ désigne la **composante tangentielle** du champ et l'intégrale $\int_{\Gamma} \mathbf{f} \cdot \mathbf{T} ds$ est appelée le **flot intégral** de \mathbf{f} le long de Γ . Lorsque Γ est fermée, on parle de **circulation** de \mathbf{f} le long de Γ (mécanique des fluides par exemple).

EXERCICE 7.4. Un exemple d'intégrale curviligne est donné par les problèmes concernant la distribution d'une masse le long d'une courbe. Pensez par exemple à une courbe C en dimension 3 comme un fil mince fait d'un fin matériau de densité variable. En supposant que la densité est décrite par un champ scalaire φ , où $\varphi(x, y, z)$ est la masse par unité de longueur au point $(x, y, z) \in C$. La masse totale M du fil est définie par l'intégrale curviligne de φ par rapport à l'abscisse curviligne

$$M = \int_C \varphi(x, y, z) ds.$$

Le centre de masse du fil est défini par le point $(\bar{x}, \bar{y}, \bar{z})$ dont les coordonnées sont définies par les équations

$$\begin{aligned}\bar{x}M &= \int_C x\varphi(x, y, z)ds, \\ \bar{y}M &= \int_C y\varphi(x, y, z)ds, \\ \bar{z}M &= \int_C z\varphi(x, y, z)ds.\end{aligned}$$

Un fil de densité constante est dit uniforme. Dans ce cas, le centre de masse est appelé centroïde.

- (1) Calculer la masse M d'un enroulement d'une bobine ayant une forme hélicoïdale dont l'équation vectorielle est

$$\alpha(t) = a \cos t \mathbf{i} + a \sin t \mathbf{j} + bt \mathbf{k},$$

si la densité est $x^2 + y^2 + z^2$.

- (2) Calculer le centre de masse $\bar{z}M$.
- (3) Le centre d'inertie d'un fil par rapport à un axe peut également être calculé grâce aux intégrales de ligne. Si $\delta(x, y, z)$ représente la distance perpendiculaire d'un point $(x, y, z) \in C$ à un axe L , le moment d'inertie I_L est défini par

$$I_L = \int_C \delta^2(x, y, z)\varphi(x, y, z)ds.$$

Les moments d'inertie selon les axes des coordonnées sont notés I_x , I_y et I_z . Calculer le moment d'inertie I_z relatif à l'exercice.

6. Indépendance par rapport au chemin dans les ensembles ouverts connexes

Soit \mathbf{f} un champ de vecteurs continu sur un ouvert connexe S . Choisissons deux points \mathbf{a} et \mathbf{b} dans S et considérons l'intégrale curviligne de \mathbf{f} de \mathbf{a} vers \mathbf{b} le long d'un chemin régulier par morceaux S . La valeur de l'intégrale dépend en général du chemin joignant \mathbf{a} à \mathbf{b} . Pour certains champs de vecteurs, l'intégrale dépend seulement des extrémités \mathbf{a} et \mathbf{b} et non du chemin les joignant. Dans cette situation, nous dirons que l'intégrale est indépendante du chemin allant de \mathbf{a} vers \mathbf{b} . Nous dirons que l'intégrale curviligne de \mathbf{f} est **indépendante du chemin** de S si elle est indépendante du chemin joignant \mathbf{a} à \mathbf{b} .

7. Circulation d'un champ de gradient

Dans le cas de fonctions réelles, nous connaissons déjà le

THÉORÈME 7.2. Soit φ une fonction réelle qui est continue sur un intervalle fermé $[a; b]$ et supposons que l'intégrale

$$\int_a^b \varphi'(t) dt$$

existe. Si φ' est continue sur l'intervalle ouvert $]a; b[$, nous avons

$$\int_a^b \varphi'(t) dt = \varphi(b) - \varphi(a).$$

Dans le cas de **champs scalaires**, nous avons le théorème suivant :

THÉORÈME 7.3. Soit φ un champ scalaire différentiable de gradient $\nabla\varphi$ continu sur un ouvert **connexe** U de \mathbb{R}^n . Alors, pour tous les points \mathbf{a} et \mathbf{b} joints par un chemin régulier par morceaux α dans U , nous avons

$$\int_{\mathbf{a}}^{\mathbf{b}} \nabla\varphi \cdot d\alpha = \varphi(\mathbf{b}) - \varphi(\mathbf{a}).$$

La conséquence est que l'intégrale curviligne d'un gradient continu sur un ouvert connexe est indépendante du chemin. Pour un chemin fermé, puisque $\varphi(\mathbf{a}) = \varphi(\mathbf{b})$, la circulation d'un gradient continu est nulle le long de tout chemin fermé régulier par morceaux de U .

DÉMONSTRATION. Soit une courbe Γ de U , régulière par morceaux, paramétrée par α définie sur $[a, b]$, et telle que $\alpha(a) = \mathbf{a}$ et $\alpha(b) = \mathbf{b}$. Alors $\int_{\mathbf{a}}^{\mathbf{b}} \nabla\varphi \cdot d\alpha = \int_a^b \nabla\varphi(\alpha(t))\alpha'(t) dt = \int_a^b (\varphi \circ \alpha)'(t) dt = \varphi(\alpha(b)) - \varphi(\alpha(a)) = \varphi(\mathbf{b}) - \varphi(\mathbf{a})$. □

REMARQUES 7.3. (1) φ est parfois appelé **potentiel**.

- (2) Dans les conditions du théorème, l'intégrale curviligne dépend seulement des valeurs du *potentiel* φ en \mathbf{a} et \mathbf{b} , elle ne dépend pas du chemin Γ pris pour aller de \mathbf{a} à \mathbf{b} , du moment que Γ reste dans le domaine connexe U .
- (3) L'intégrale curviligne d'un gradient sur une courbe fermée d'un ouvert connexe est nulle !
- (4) On peut généraliser ces résultats de la façon suivante : on appelle *surfaces équipotentielle*s, ou *surfaces de niveau*, les surfaces d'équation

$$\varphi(x, y, z) = \text{constante}.$$

Soient S_1 et S_2 deux surfaces équipotentielle s incluses toutes deux dans un ensemble connexe U , \mathbf{a} un point quelconque de S_1 et \mathbf{b} un point quelconque de S_2 , Γ une courbe joignant \mathbf{a} à \mathbf{b} , dans U . Alors l'intégrale curviligne

$$\int_{\Gamma} \frac{\partial\varphi}{\partial x} dx + \frac{\partial\varphi}{\partial y} dy + \frac{\partial\varphi}{\partial z} dz$$

notée aussi $\int_{\Gamma} \nabla \varphi \cdot d\alpha$ a une valeur constante, indépendante des points \mathbf{a} et \mathbf{b} sur S_1 et S_2 respectivement, et qui ne dépend que de S_1 et S_2 . Elle représente la *différence de potentiel* entre S_1 et S_2 .

On peut montrer en fait que les gradients sont les seuls champs de vecteurs ayant cette propriété, c'est l'objet du théorème suivant, dont la preuve est donnée dans [4] p. 341-351, ou [5] p.261-264.

THÉORÈME 7.4. *Soit U un **ouvert connexe** de \mathbb{R}^n , et \mathbf{f} un champ de vecteur continu défini sur U à valeurs dans \mathbb{R}^n . On a les équivalences suivantes :*

- i) \mathbf{f} dérive d'un potentiel φ , champ scalaire défini sur U (i.e. $\mathbf{f} = \nabla \varphi$),*
- ii) Pour toute courbe fermée Γ de U , $\oint_{\Gamma} \mathbf{f} \cdot d\alpha = 0$,*
- iii) L'intégrale curviligne de \mathbf{f} entre deux points \mathbf{a} et \mathbf{b} de U est indépendante de la courbe Γ de U considérée pour aller de \mathbf{a} à \mathbf{b} .*

Si en plus, \mathbf{f} est de classe C^1 et si U est étoilé, on verra en Analyse Tensorielle que i), ii), iii) sont alors équivalents avec le fait que la matrice jacobienne de \mathbf{f} est symétrique : $\partial_i f_j = \partial_j f_i$, ou encore que le rotationnel de \mathbf{f} est nul.

EXERCICE 7.5. Soit le champ de vecteurs \mathbf{f} de coordonnées $P(x, y) = \frac{2x}{y}$ et $Q(x, y) = \frac{1-x^2}{y^2}$. Soit $U = \mathbb{R}^2 \setminus (y = 0)$ le plan \mathbb{R}^2 privé de la droite $y = 0$, et Γ une courbe d'extrémités $\mathbf{a} = (a_1, a_2)$, $\mathbf{b} = (b_1, b_2)$ incluse dans U . Calculer $\int_{\Gamma} \mathbf{f} \cdot d\alpha$.

EXERCICE 7.6. Calculer l'intégrale

$$\int_{\Gamma} \frac{xdy - ydx}{x^2 + y^2}$$

dans les cas suivants :

- (1) Si Γ est une courbe incluse dans un pavé inclus dans le demi-plan supérieur ouvert $\{(x, y) \in \mathbb{R}^2, y > 0\}$, d'extrémités \mathbf{a} et \mathbf{b} ,
- (2) si Γ est le cercle trigonométrique.

Intégrales multiples II

1. Introduction

Cette partie contient des éléments de calcul intégral destinés à compléter les notions d'intégrales multiples. Essentiellement, nous verrons : le théorème de Green dans le plan, une condition nécessaire et suffisante pour qu'un champ de vecteur bidimensionnel soit un gradient, les changements de variables dans une intégrale double ainsi que l'extension à des dimensions supérieures.

2. Le théorème de Green dans le plan

Il existe un analogue au [Théorème 7.3](#) pour la dimension deux. Ce théorème, appelé **théorème de Green**, relie l'intégrale double sur une région Ω à une intégrale de ligne prise le long de la **frontière** Γ du domaine. La forme de cette relation est

$$\iint_{\Omega} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy = \int_{\Gamma^+} P dx + Q dy. \quad (2.1)$$

La courbe Γ qui apparaît à droite est la frontière de Ω , la courbe est orientée dans le **sens direct**. Par sens direct, on entend que si l'on se promène dans le sens direct sur la courbe Γ , l'intérieur de la courbe (i.e. la région Ω) se trouve à notre gauche.

Deux types d'hypothèses doivent être prises en compte pour la validité de la relation (2.1). Les hypothèses les plus classiques sont que P et Q sont continûment différentiables sur un ouvert S contenant Ω . On a alors la continuité de $\partial_y P$ et $\partial_x Q$ sur Ω . On peut affaiblir également ces hypothèses. Il y a ensuite des hypothèses géométriques sur le domaine Ω et sa frontière Γ . La frontière Γ doit être **fermée rectifiable** (c'est-à-dire de longueur finie). Supposons que Γ est décrite par une fonction continue α à valeur vectorielle définie sur un intervalle $[a; b]$. Si $\alpha(a) = \alpha(b)$, Γ est fermée. Une courbe fermée telle que $\alpha(t_1) \neq \alpha(t_2)$, $\forall t_1 \neq t_2$ dans $]a; b[$ est appelée une **courbe fermée simple**. Celles d'entre elles qui se trouvent dans le plan sont appelées **courbes de Jordan**. On sépare donc le plan en deux parties disjointes ouvertes et connexes de frontière commune Γ : une est bornée, l'autre non. La première est appelée intérieure, la seconde extérieure. Le théorème de Green, qui n'est pas évident à démontrer, est vrai pour toute courbe de Jordan. Nous restreignons ici notre discussion à des courbes régulières par morceaux. Une autre difficulté est liée à la définition du

sens de circulation ("dans le sens direct"). Ceci peut s'exprimer grâce à α . Toutefois, ici, nous considérons une définition "intuitive" (l'intérieur de la courbe se trouve à gauche).

THÉORÈME 8.1 (Théorème de Green). *Soient P et Q des champs scalaires qui sont continûment différentiables sur un ouvert S dans le plan (x, y) . Soit Γ une courbe de Jordan \mathcal{C}^1 par morceaux, et soit Ω l'union de Γ et de son intérieur. Supposons que Ω est un sous-ensemble de S . Alors, nous avons l'identité*

$$\iint_{\Omega} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy = \int_{\Gamma^+} (P dx + Q dy),$$

où l'intégrale curviligne est prise le long de Γ dans le sens direct.

DÉMONSTRATION. On démontre le théorème dans le cas (cf. figure [figure 1](#)) où on peut écrire Ω indifféremment comme

$$\Omega = \{(x, y) / a \leq x \leq b, \varphi_1(x) \leq y \leq \varphi_2(x)\}, \quad (2.2)$$

φ_1, φ_2 étant deux fonctions à valeurs réelles, de classe \mathcal{C}^1 par morceaux sur $[a, b]$, et telles que $\varphi_1 \leq \varphi_2$;

ou comme

$$\Omega = \{(x, y) / c \leq y \leq d, \psi_1(y) \leq x \leq \psi_2(y)\}, \quad (2.3)$$

ψ_1, ψ_2 étant deux fonctions à valeurs réelles, de classe \mathcal{C}^1 par morceaux sur $[c, d]$, et telles que $\psi_1 \leq \psi_2$.

Considérons l'écriture (2.2) de Ω .

Soit Γ la frontière de Ω , orientée positivement, i.e. laissant Ω à sa gauche. Soit $(x, y) \mapsto P(x, y)$ une fonction continue sur Ω , admettant une dérivée $\frac{\partial P}{\partial y}$ continue sur Ω . On a

$$\iint_{\Omega} \frac{\partial P}{\partial y} dx dy = \int_a^b \left(\int_{\varphi_1(x)}^{\varphi_2(x)} \frac{\partial P}{\partial y} dy \right) dx = \int_a^b [P(x, \varphi_2(x)) - P(x, \varphi_1(x))] dx.$$

On reconnaît au membre de gauche l'intégrale curviligne de la forme différentielle $-P dx$ le long de Γ^+ . On a donc :

$$\iint_{\Omega} \frac{\partial P}{\partial y} dx dy = - \int_{\Gamma^+} P(x, y) dx.$$

Considérons maintenant l'écriture (2.3) de Ω . Soit ensuite $(x, y) \mapsto Q(x, y)$ une fonction continue sur Ω , admettant une dérivée $\frac{\partial Q}{\partial x}$ continue sur Ω . On a

$$\iint_{\Omega} \frac{\partial Q}{\partial x} dx dy = \int_c^d \left(\int_{\psi_1(y)}^{\psi_2(y)} \frac{\partial Q}{\partial x} dx \right) dy = \int_c^d [Q(\psi_2(y), y) - Q(\psi_1(y), y)] dy.$$

On reconnaît au second membre l'intégrale curviligne de la forme différentielle Qdy le long de Γ^+ . On a donc :

$$\iint_{\Omega} \frac{\partial Q}{\partial x} dx dy = \int_{\Gamma^+} Q(x, y) dy.$$

Finalement, on a

$$\iint_{\Omega} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy = \int_{\Gamma} (P dx + Q dy).$$

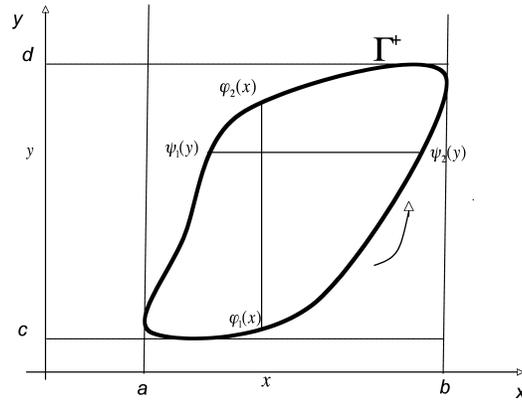


FIGURE 1 – Démonstration de la formule de Green.

□

EXERCICE 8.1. Utiliser le théorème de Green pour calculer le travail, défini par

$$\int_{\Gamma} P dx + Q dy,$$

résultant d'un champ de force $f(x, y) = (y + 3x)\mathbf{i} + (2y - x)\mathbf{j}$, en faisant bouger une particule dans le sens direct le long de l'ellipse : $4x^2 + y^2 = 4$. On prendra $P(x, y) = y + 3x$ et $Q(x, y) = 2y - x$.

EXERCICE 8.2. Evaluer l'intégrale curviligne

$$\int_{\Gamma} (5 - xy - y^2) dx - (2xy - x^2) dy,$$

où Γ est le carré de sommets $(0, 0)$, $(1, 0)$, $(1, 1)$, $(0, 1)$ dans le sens direct.

3. Calcul des aires planes

Considérons maintenant un domaine Ω pouvant avoir les deux écritures (2.2) et (2.3). L'aire de Ω est définie par $\iint_{\Omega} dx dy$.

Considérons dans un premier temps $P(x, y) = x$ et $Q(x, y) = 0$. Alors, la formule de Green donne :

$$\iint_{\Omega} dx dy = - \int_{\Gamma^+} y dx,$$

Γ^+ étant la frontière orientée de Ω .

De la même façon, considérons $P(x, y) = 0$ et $Q(x, y) = x$, alors

$$\iint_{\Omega} dx dy = \int_{\Gamma^+} x dy.$$

Par conséquent, l'aire de Ω se calcule de la façon suivante :

$$\iint_{\Omega} dx dy = \frac{1}{2} \int_{\Gamma^+} x dy - y dx.$$

REMARQUE 8.1. Cette formule est encore vraie pour un domaine Ω défini comme dans le [Théorème 8.1](#).

EXERCICE 8.3. (La surface exprimée comme une intégrale curviligne). Nous pouvons écrire

$$|\Omega| = \iint_{\Omega} dx dy.$$

Nous supposons que la frontière Γ est décrite par des équations paramétriques

$$x = X(t), \quad y = Y(t),$$

pour $t \in [a; b]$. Alors

- (i) Donner une formule donnant la surface en terme d'intégrale curviligne faisant intervenir a, b, X, Y, X', Y' .
- (ii) Calculer la surface de l'ellipse d'équation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

4. Calcul d'aire et Théorème de Green en coordonnées polaires

Soit un arc MN , orienté de M à N , comme illustré sur la [figure 2](#).

Le domaine A auquel on s'intéresse est l'intérieur du circuit formé des arcs OM , MN et NO . L'arc MN est paramétré par

$$\{(x = \rho \cos \theta, y = \rho \sin \theta), \quad \text{avec } \theta \in [\alpha, \beta], \text{ et } \rho = f(\theta)\},$$

où f est une fonction de classe \mathcal{C}^1 par morceaux sur $[\alpha, \beta]$. Remarquons que le paramétrage $x = \rho \cos \theta$ et $y = \rho \sin \theta$ entraîne

$$x dy - y dx = \rho^2 d\theta.$$

Par suite, pour l'aire de A , en supposant de plus que A est borné par une courbe de Jordan, on peut appliquer le théorème de Green, et on a (puisque $d\theta$ est nulle sur les rayons OM, ON) :

$$\text{Aire}(A) = \frac{1}{2} \int_{MN} \rho^2 d\theta$$

où MN est parcouru de M vers N .

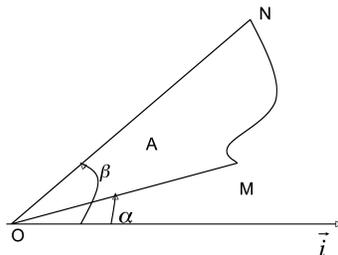


FIGURE 2 – Domaine A dont on veut calculer l'aire

EXERCICE 8.4. Calculer l'aire de l'astroïde d'équation $x(t) = a \cos^3(t)$; $y(t) = a \sin^3(t)$ pour $t \in [0, 2\pi]$.

5. Changements de variables dans une intégrale double

Lorsque nous travaillons sur l'intégration en dimension un, il est souvent commode de transformer une intégrale compliquée à calculer en une intégrale plus usuelle à travers un changement de variables en utilisant la formule

$$\int_a^b f(x) dx = \int_c^d f(g(t)) g'(t) dt,$$

où $a = g(c)$ et $b = g(d)$. Ceci est possible lorsque g a une dérivée continue sur un intervalle $[c; d]$ et que f est continue sur l'ensemble des valeurs de $g(t)$ lorsque t balaye l'intervalle $[c; d]$.

Il y a un analogue à cette formule en dimension supérieure pour les intégrales multiples. C'est toutefois plus compliqué puisque nous avons formellement deux substitutions, une pour x , une autre pour y . Supposons que nous ayons deux fonctions X et Y telles que

$$\begin{cases} x = X(u, v) \\ y = Y(u, v) \end{cases}$$

Ceci définit une application entre l'espace uv et le plan xy . On transporte un point (u, v) de T en un point (x, y) de S à travers une application \mathbf{r} définie par

$$\mathbf{r}(u, v) = X(u, v)\mathbf{i} + Y(u, v)\mathbf{j}, \quad \text{si } (u, v) \in T. \quad (5.1)$$

C'est ce que l'on appelle l'équation vectorielle de la transformation. Quelques fois, on peut retraduire l'équation de transformation en exprimant (u, v) en fonction de (x, y)

$$\begin{cases} u = U(x, y) \\ v = V(x, y) \end{cases}$$

C'est donc l'application inverse de celle précédente. Les applications **bijectives** sont de toute première importance. Elles transportent des points distincts de T en des points distincts de S . On établit alors une correspondance entre les points de T et S notamment, au moins théoriquement, à travers l'application inverse. Nous considérons ici des fonctions X et Y de classe \mathcal{C}^1 sur S . Nous supposons des hypothèses identiques sur U et V .

On peut écrire la formule de changement de variables sous la forme

$$\iint_S f(x, y) dS = \iint_T g(u, v) |J(u, v)| dT \quad (5.2)$$

où $dS = dx dy$, $dT = du dv$ et $g(u, v) = f(X(u, v), Y(u, v)) = (f \circ \mathbf{r})(u, v)$. Le facteur $J(u, v)$ qui apparaît dans l'intégrale de droite joue le rôle de $g'(t)$ dans la formule unidimensionnelle. Ce facteur est appelé **déterminant jacobien** de l'application (5.1) et est défini par

$$J(u, v) = \begin{vmatrix} \partial_u X & \partial_v X \\ \partial_u Y & \partial_v Y \end{vmatrix}$$

On écrit également quelques fois $\frac{\partial(X, Y)}{\partial(u, v)}$. Nous ne donnons pas ici les **hypothèses** optimales qui nous permettent de valider la formule (5.2) mais disons que, en plus des hypothèses de continuité sur X, Y, U et V données ci-dessus, nous supposons que l'application \mathbf{r} de T vers S est **bijective** et que le **jacobien** $J(u, v)$ **n'est jamais nul**.

EXERCICE 8.5. (1) Montrer que l'on a en coordonnées polaires

$$\iint_S f(x, y) dx dy = \iint_T f(r \cos \theta, r \sin \theta) r dr d\theta,$$

où S est le disque de centre l'origine et de rayon a , T est à déterminer.

(2) Calculer le volume de la sphère de rayon a située dans le premier octant par

$$\iint_S \sqrt{a^2 - x^2 - y^2} dS$$

où $S = \{(x, y) / x^2 + y^2 \leq a^2, x \geq 0, y \geq 0\}$.

EXERCICE 8.6. Refaire les exercices 8.8 et 6.3 avec un changement de variables. Est-ce plus facile ainsi ?

EXERCICE 8.7. On considère l'intégrale $I = \iint_D \arctan(x + y) dx dy$, D étant le domaine

$$D = \{(x, y) \in \mathbb{R}^2; \quad x \geq 0, y \geq 0, x + y \leq 1\}.$$

Effectuer le changement de variables donné par : $x = u, y = v - u$. Déterminer le nouveau domaine d'intégration Δ , et calculer le jacobien. En déduire la valeur de I .

6. Extension à des dimensions supérieures

On peut bien sûr étendre la notion d'intégrale multiple à des dimensions $n \geq 3$ de manière analogue au cas 2D. On notera alors l'intégrale sur un domaine n -dimensionnel sous la forme

$$\int \dots \int_S f dS, \quad \text{ou} \quad \int_S f dS.$$

Intéressons nous au changement de variables. Introduisons les variables $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ et $\mathbf{u} = (u_1, \dots, u_n) \in \mathbb{R}^n$ et supposons que l'on ait des relations du type

$$x_j = X_j(u_1, \dots, u_n), \quad 1 \leq j \leq n.$$

On peut alors définir une application vectorielle \mathbf{X}

$$\begin{aligned} \mathbf{X} : T &\rightarrow S, \\ \mathbf{u} &\mapsto \mathbf{x} = \mathbf{X}(\mathbf{u}) \end{aligned} \tag{6.1}$$

Supposons \mathbf{X} bijective et continûment différentiable sur T . La formule de changements de coordonnées s'écrit alors

$$\int_S f(\mathbf{x}) d\mathbf{x} = \int_T f(\mathbf{X}(\mathbf{u})) |\det D\mathbf{X}(\mathbf{u})| d\mathbf{u} \tag{6.2}$$

où $D\mathbf{X}(\mathbf{u})$ est la matrice jacobienne du champ de vecteurs \mathbf{X} . Nous avons la validité de la formule (6.2) si $J(\mathbf{u}) = \det D\mathbf{X}(\mathbf{u})$ n'est jamais nulle sur T .

Dans le cas tridimensionnel, posons $\mathbf{x} = (x, y, z)$, $\mathbf{X} = (X, Y, Z)$ et $\mathbf{u} = (u, v, w)$ comme notations. On a alors

$$\iiint_S f(x, y, z) dS = \iiint_T f(X(u, v, w), Y(u, v, w), Z(u, v, w)) |\det J(u, v, w)| du dv dw$$

où

$$J(u, v, w) = \begin{vmatrix} \partial_u X & \partial_v X & \partial_w X \\ \partial_u Y & \partial_v Y & \partial_w Y \\ \partial_u Z & \partial_v Z & \partial_w Z \end{vmatrix}$$

Soit $\mathbf{X}(u, v, w) = X\mathbf{i} + Y\mathbf{j} + Z\mathbf{k}$, $\mathbf{V}_1 = \partial_u \mathbf{X}$, $\mathbf{V}_2 = \partial_v \mathbf{X}$ et $\mathbf{V}_3 = \partial_w \mathbf{X}$. Alors, nous avons

$$|J(u, v, w)| = |\mathbf{V}_1 \cdot (\mathbf{V}_2 \wedge \mathbf{V}_3)|$$

EXERCICE 8.8. Calculer le volume V du solide défini par l'ellipsoïde d'équation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

et montrer que $V = \frac{4}{3}\pi abc$.

EXERCICE 8.9. Donner la formule de changement de coordonnées dans le cas

- d'une transformation en coordonnées cylindriques,
- d'une transformation en coordonnées sphériques.

EXERCICE 8.10. Calculer l'intégrale triple

$$I = \iiint_D \left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \right)^\alpha dx dy dz,$$

avec $a > 0$, $b > 0$, $c > 0$, $\alpha > 0$ et

$$D = \left\{ (x, y, z) \in \mathbb{R}^3 / \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leq 1 \right\}.$$

Intégrales de surface et théorème de Stokes

1. Représentation paramétrique d'une surface

Ce chapitre s'intéresse aux intégrales de surfaces et à leurs applications. De manière imagée, l'intégrale de surface est analogue à l'intégrale curviligne lorsque la région d'intégration est une surface plutôt qu'une courbe. Avant de commencer, mettons nous d'accord sur ce qu'est une **surface**.

DÉFINITION 9.1. Une **surface** est un ensemble de points de \mathbb{R}^3 , dont les coordonnées x, y, z sont fonctions continues de deux paramètres u, v . Ces fonctions sont définies sur un certain domaine de \mathbb{R}^2 . Il y a plusieurs façons de décrire une surface.

— Une **surface** peut être **paramétrée** par un couple (U, α) où U est un ouvert connexe de \mathbb{R}^2 , et $\alpha = (X, Y, Z)$ est un champ vectoriel de U dans \mathbb{R}^3 :

$$\begin{cases} x = X(u, v) \\ y = Y(u, v) \\ z = Z(u, v) \end{cases} \quad (1.1)$$

C'est ici l'analogie pour les surfaces de la représentation paramétrique à un paramètre pour les courbes de l'espace. On voit bien ici la présence de deux degrés de liberté u et v . Une autre manière de dire les choses est qu'une surface est une portion du plan déformée par la transformation (1.1).

Considérons le cas particulier où on représente la surface en exprimant une de ses coordonnées, z par exemple, en fonction de x et y : $f(x, y)$. C'est ce que l'on appelle la **représentation explicite**. Dans ce cas, le paramétrage α est égal à $(x, y, f(x, y))$.

— Une **surface** peut être représentée de manière **implicite**, par l'ensemble de points (x, y, z) satisfaisant une équation de la forme $F(x, y, z) = 0$.

EXEMPLE 9.1. Une sphère peut être représentée implicitement par une équation du type $x^2 + y^2 + z^2 = 1$ ou explicitement par $z = \pm\sqrt{1 - x^2 - y^2}$, le signe + correspondant à l'hémisphère nord et le signe - à l'hémisphère sud.

Dans beaucoup de situations, U est un rectangle, un disque ou d'autres ensembles connexes bornés par une courbe simple fermée.

REMARQUE 9.1. Soit une surface Σ , paramétrée par un champ vectoriel α défini sur l'ouvert connexe U . On peut voir $\alpha(u, v)$ comme le **point** de la surface Σ , image de (u, v) par α , ou bien, comme le vecteur partant de l'origine au point $\alpha(u, v)$.

Dans la théorie générale des surfaces, les fonctions X , Y et Z sont supposées **continues** sur U . La portion de surface $\alpha(U)$ est appelée **surface paramétrique**. Si la fonction α est **bijective** sur U , $\alpha(U)$ est appelée une **représentation paramétrique simple**. Dans une telle situation, les points différents de U sont transportés en points différents de la surface. Notamment, toute courbe fermée de U est transportée en une courbe fermée de la surface. Enfin, il est possible qu'une surface paramétrique dégénère en un point ou ligne de l'espace.

EXEMPLE 9.2. La sphère centrée en l'origine et de rayon a est décrite par la représentation paramétrique

$$\begin{cases} x = a \cos u \cos v \\ y = a \sin u \cos v \\ z = a \sin v \end{cases} \quad (1.2)$$

pour $(u, v) \in U = [0; 2\pi] \times [-\frac{\pi}{2}; \frac{\pi}{2}]$.

EXEMPLE 9.3. Le cône de hauteur $h \cos \gamma$ est décrit par

$$\begin{cases} x = v \sin \gamma \cos u \\ y = v \sin \gamma \sin u \\ z = v \sin \gamma \end{cases} \quad (1.3)$$

pour $(u, v) \in U = [0; 2\pi] \times [0; h]$.

Il y a bien sûr de nombreuses représentations paramétriques pour une surface donnée.

DÉFINITION 9.2. Soit Σ une surface paramétrée par une fonction α . Supposons que α transporte un ouvert connexe U du plan (u, v) en une surface paramétrique $\alpha(U)$. Supposons également que U est l'image d'un ouvert connexe V dans un plan (s, t) par une application \mathbf{G} **bijective, de classe C^1 , et de jacobien non nul**, donnée par

$$\begin{cases} \mathbf{G}(s, t) = u(s, t)\mathbf{i} + v(s, t)\mathbf{j} = (u, v) \\ si (s, t) \in V \end{cases} \quad (1.4)$$

Considérons la fonction β définie sur V par l'équation

$$\beta(s, t) = \alpha[\mathbf{G}(s, t)] \quad (1.5)$$

\mathbf{G} est appelé **changement de paramétrage**.

Deux fonctions α et β ainsi reliées sont dites **équivalentes**. Elles décrivent notamment la **même surface** Σ : $\alpha(U)$ et $\beta(V)$ sont identiques en tant qu'ensembles de points.

Notons $J_{\mathbf{G}} = \det D\mathbf{G}$ le jacobien de \mathbf{G} .

Si $J_{\mathbf{G}} > 0$, le changement de paramétrage est dit **direct**, on dit que l'**orientation est conservée**, si $J_{\mathbf{G}} < 0$, le changement de paramétrage est dit **indirect**, on dit que l'**orientation est inversée**.

2. Plan tangent et normale à une surface

Nous faisons l'hypothèse ici que le vecteur α est de classe \mathcal{C}^1 sur U , un ouvert connexe du plan :

$$\alpha(u, v) = X(u, v)\mathbf{i} + Y(u, v)\mathbf{j} + Z(u, v)\mathbf{k},$$

avec $(u, v) \in U$. On appelle Σ la surface paramétrée par α sur U .

Comme α est de classe \mathcal{C}^1 , on dit aussi que la surface Σ est de **classe \mathcal{C}^1** .

Puisque α est différentiable sur U , on peut alors considérer les deux vecteurs

$$\begin{cases} \frac{\partial \alpha}{\partial u} = \frac{\partial X}{\partial u}\mathbf{i} + \frac{\partial Y}{\partial u}\mathbf{j} + \frac{\partial Z}{\partial u}\mathbf{k} \\ \frac{\partial \alpha}{\partial v} = \frac{\partial X}{\partial v}\mathbf{i} + \frac{\partial Y}{\partial v}\mathbf{j} + \frac{\partial Z}{\partial v}\mathbf{k} \end{cases}$$

DÉFINITION 9.3. *On dira que le point $\alpha(u, v)$ de la surface paramétrée Σ est **régulier** si, en ce point,*

$$\frac{\partial \alpha}{\partial u} \wedge \frac{\partial \alpha}{\partial v} \neq \mathbf{0}$$

*Dans le cas contraire, on parle de **point singulier**. Si tous les points sont réguliers, la surface est dite **régulière**.*

En un point régulier, les vecteurs $\frac{\partial \alpha}{\partial u}$ et $\frac{\partial \alpha}{\partial v}$ sont linéairement indépendants, et engendrent donc un plan. Le plan affine contenant le point $M = \alpha(u, v)$ et les deux directions $\frac{\partial \alpha}{\partial u}$ et $\frac{\partial \alpha}{\partial v}$ est appelé **plan tangent** en M à la surface Σ .

Prouvons maintenant que ce plan est indépendant de la représentation paramétrique. En effet, si β est une paramétrisation définie sur un ouvert connexe V , équivalente à α , il existe une fonction \mathbf{G} bijective, de classe \mathcal{C}^1 , et de jacobien non nul, de V vers U telle que $\beta(s, t) = \alpha(\mathbf{G}(s, t))$. La fonction \mathbf{G} applique les nouveaux paramètres (s, t) sur les anciens (u, v) :

$$(u, v) = \mathbf{G}(s, t).$$

D'après la règle de différentiation des fonctions composées, on a :

$$\begin{cases} \frac{\partial \beta}{\partial s} = \frac{\partial \alpha}{\partial u} \frac{\partial u}{\partial s} + \frac{\partial \alpha}{\partial v} \frac{\partial v}{\partial s} \\ \frac{\partial \beta}{\partial t} = \frac{\partial \alpha}{\partial u} \frac{\partial u}{\partial t} + \frac{\partial \alpha}{\partial v} \frac{\partial v}{\partial t} \end{cases} \quad (2.1)$$

Comme le jacobien de \mathbf{G}

$$J_{\mathbf{G}} = \begin{vmatrix} \frac{\partial u}{\partial s} & \frac{\partial u}{\partial t} \\ \frac{\partial v}{\partial s} & \frac{\partial v}{\partial t} \end{vmatrix} \quad (2.2)$$

ne s'annule pas sur V , le sous-espace engendré par les vecteurs $\frac{\partial \beta}{\partial s}$ et $\frac{\partial \beta}{\partial t}$ coïncide avec le sous-espace engendré par les vecteurs $\frac{\partial \alpha}{\partial u}$ et $\frac{\partial \alpha}{\partial v}$.

Les notions de point régulier et de plan tangent sont donc indépendantes du paramétrage choisi pour Σ , pourvu que les paramétrisations soient équivalentes.

On peut donc énoncer le théorème-définition suivant :

THÉORÈME 9.1. *En tout point régulier d'une surface paramétrée de classe \mathcal{C}^1 , on peut définir un plan tangent, indépendant du paramétrage choisi pour Σ , pourvu qu'on considère des paramétrisations équivalentes. Si (α, U) est un tel paramétrage de Σ , le plan tangent en $M = \alpha(u, v) \in \Sigma$ est le plan passant par M et de directions $\frac{\partial \alpha}{\partial u}$ et $\frac{\partial \alpha}{\partial v}$.*

*En ce point, la droite orthogonale au plan tangent est appelée **normale**. Elle est dirigée par le vecteur*

$$\mathbf{N} = \frac{\partial \alpha}{\partial u} \wedge \frac{\partial \alpha}{\partial v}$$

Les formules (2.1) prouvent que, dans un changement de paramétrage \mathbf{G} ,

$$\mathbf{N}_1 = \frac{\partial \beta}{\partial s} \wedge \frac{\partial \beta}{\partial t} = \left(\frac{\partial \alpha}{\partial u} \wedge \frac{\partial \alpha}{\partial v} \right) J_{\mathbf{G}} = \mathbf{N} J_{\mathbf{G}},$$

$J_{\mathbf{G}}$ étant le jacobien de G .

On peut exprimer la normale, grâce à la matrice jacobienne comme :

$$\frac{\partial \alpha}{\partial u} \wedge \frac{\partial \alpha}{\partial v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial X}{\partial u} & \frac{\partial Y}{\partial u} & \frac{\partial Z}{\partial u} \\ \frac{\partial X}{\partial v} & \frac{\partial Y}{\partial v} & \frac{\partial Z}{\partial v} \end{vmatrix} = \frac{\partial(Y, Z)}{\partial(u, v)} \mathbf{i} + \frac{\partial(Z, X)}{\partial(u, v)} \mathbf{j} + \frac{\partial(X, Y)}{\partial(u, v)} \mathbf{k} \quad (2.3)$$

où

$$\begin{cases} \frac{\partial(Y, Z)}{\partial(u, v)} &= \frac{\partial Y}{\partial u} \frac{\partial Z}{\partial v} - \frac{\partial Y}{\partial v} \frac{\partial Z}{\partial u} \\ \frac{\partial(Z, X)}{\partial(u, v)} &= \frac{\partial Z}{\partial u} \frac{\partial X}{\partial v} - \frac{\partial Z}{\partial v} \frac{\partial X}{\partial u} \\ \frac{\partial(X, Y)}{\partial(u, v)} &= \frac{\partial X}{\partial u} \frac{\partial Y}{\partial v} - \frac{\partial X}{\partial v} \frac{\partial Y}{\partial u} \end{cases}$$

REMARQUE 9.2. On travaille avec des ouverts. Mais, dans le cas où l'on travaille avec un ensemble fermé, il se peut que la surface ne soit pas de classe \mathcal{C}^1 sur tout l'ensemble, et le fait qu'un point soit régulier ou singulier peut dépendre de la paramétrisation choisie. C'est ce qu'illustre l'exercice suivant.

EXERCICE 9.1. Soit une surface avec représentation explicite $z = f(x, y)$.

- (1) Donner la représentation paramétrique α de cette surface pour x et y les paramètres.
- (2) Calculer $\frac{\partial \alpha}{\partial x} \wedge \frac{\partial \alpha}{\partial y}$ en tout point où f est différentiable. Décrire les points singuliers (en donnant des conditions sur f ou plutôt ses dérivées partielles).
- (3) On considère l'hémisphère nord donné par $z = \sqrt{1 - x^2 - y^2}$, avec $x^2 + y^2 \leq 1$. Montrer que les seuls points singuliers se trouvent sur l'équateur.
- (4) Considérons maintenant l'hémisphère nord comme l'image du rectangle $U = [0; 2\pi] \times [0; \frac{\pi}{2}]$ par la transformation

$$\alpha(u, v) = \cos u \cos v \mathbf{i} + \sin u \cos v \mathbf{j} + \sin v \mathbf{k}.$$

Calculer $\frac{\partial \alpha}{\partial u} \wedge \frac{\partial \alpha}{\partial v}$ et montrer que le pôle nord est l'unique point singulier de cette représentation.

3. Aire d'une surface paramétrique

On considère une surface Σ paramétrée par (α, U) . Considérons un segment horizontal de U . Son image par α est une *courbe* qui se trouve sur $\alpha(U)$. Pour v fixé, pensez à u comme une variable de temps. Le vecteur $\frac{\partial \alpha}{\partial u}$ est alors le vecteur vitesse de cette courbe.

Lorsque u change de Δu , un point originellement en $\alpha(u, v)$ bouge le long de la courbe d'une distance approximativement égale à $\Delta u \left\| \frac{\partial \alpha}{\partial u} \right\|$ puisque $\left\| \frac{\partial \alpha}{\partial u} \right\|$ représente la vitesse le long de la courbe. On peut faire le même raisonnement selon v . Un rectangle dans U de surface $\Delta u \Delta v$ trace une *portion de surface* sur $\alpha(U)$ approchée par le parallélogramme porté par $\Delta u \frac{\partial \alpha}{\partial u}$ et $\Delta v \frac{\partial \alpha}{\partial v}$. La surface de ce parallélogramme est alors

$$\left\| \frac{\partial \alpha}{\partial u} \Delta u \wedge \frac{\partial \alpha}{\partial v} \Delta v \right\| = \Delta u \Delta v \left\| \frac{\partial \alpha}{\partial u} \wedge \frac{\partial \alpha}{\partial v} \right\| \quad (3.1)$$

On peut donc voir

$$\left\| \frac{\partial \alpha}{\partial u} \wedge \frac{\partial \alpha}{\partial v} \right\|$$

comme une sorte de facteur local de modification de la mesure de surface.

En un point singulier, on obtient 0 : le parallélogramme dégénère en un point ou courbe. En chaque point régulier, les vecteurs $\frac{\partial \alpha}{\partial u}$ et $\frac{\partial \alpha}{\partial v}$ déterminent le plan tangent dont le vecteur $\frac{\partial \alpha}{\partial u} \wedge \frac{\partial \alpha}{\partial v}$ est la normale. La continuité de $\frac{\partial \alpha}{\partial u}$ et $\frac{\partial \alpha}{\partial v}$ implique la continuité de la normale évitant ainsi la présence d'arêtes ou coins sur la surface. La non annulation de la normale évite la dégénérescence.

On introduit alors la définition suivante

DÉFINITION 9.4. *L'aire* de Σ , notée $a(\Sigma)$ ou $|\Sigma|$, est définie par l'intégrale double

$$a(\Sigma) = \iint_U \left\| \frac{\partial \alpha}{\partial u} \wedge \frac{\partial \alpha}{\partial v} \right\| du dv \quad (3.2)$$

REMARQUE 9.3. On voit que cette formule est assez similaire à celle du calcul de la longueur $\ell(C)$ d'une courbe C définie par un chemin α continu et différentiable sur $[a, b]$ et donnée par

$$\ell(C) = \int_a^b \left\| \frac{d\alpha}{du}(u) \right\| du \quad (3.3)$$

Puisque (3.2) fait intervenir *a priori* la paramétrisation, l'aire d'une surface devrait donc dépendre de la paramétrisation. Comme pour les intégrales de lignes, ce n'est pas vrai :

THÉORÈME 9.2. *L'aire d'une surface paramétrée est indépendante de la paramétrisation choisie, pourvu que les paramétrisations soient équivalentes.*

DÉMONSTRATION. Soit un changement de paramétrage \mathbf{G} , tel que $\alpha(u, v) = \beta(\mathbf{G}(s, t))$ et les deux paramétrisations sont équivalentes. On a, en faisant le changement de variable $(u, v) = \mathbf{G}(s, t)$ dans l'intégrale :

$$\iint_U \left\| \frac{\partial \alpha}{\partial u} \wedge \frac{\partial \alpha}{\partial v} \right\| du dv = \iint_V \left\| \frac{\partial \beta}{\partial s} \wedge \frac{\partial \beta}{\partial t} J_{\mathbf{G}^{-1}} \right\| J_{\mathbf{G}} ds dt = \iint_V \left\| \frac{\partial \beta}{\partial s} \wedge \frac{\partial \beta}{\partial t} \right\| ds dt.$$

□

EXERCICE 9.2. (Surface d'un hémisphère). En partant de la représentation paramétrique de l'hémisphère, calculer sa surface.

EXERCICE 9.3. (théorème de Pappus). Un des théorèmes de Pappus dit qu'une surface de révolution obtenue par rotation d'une courbe plane de longueur L autour d'un axe dans le plan de la courbe, a comme aire $2\pi Lh$, où h est la distance du centroïde de la courbe à l'axe de rotation (conseil : utiliser l'exercice 7.4 du chapitre sur les intégrales de lignes).

4. Intégrales de surface

A partir de la représentation paramétrique, les intégrales de surface sont assez analogues aux intégrales de lignes.

DÉFINITION 9.5. Soit $\Sigma = \alpha(U)$ une surface paramétrique décrite par une fonction différentiable α définie sur une région U du plan (u, v) et soit f un champ scalaire défini et borné sur Σ . L'**intégrale de surface** de f sur Σ est définie par l'équation :

$$\iint_{\Sigma} f dS = \iint_U f(\alpha(u, v)) \left\| \frac{\partial \alpha}{\partial u} \wedge \frac{\partial \alpha}{\partial v} \right\| du dv$$

dès que l'intégrale de droite existe. On appelle souvent dS l'élément de surface.

EXEMPLE 9.4. L'aire est une intégrale de surface :

$$a(\Sigma) = \int_{\Sigma} dS$$

On se demande maintenant si la valeur d'une intégrale de surface dépend de la représentation paramétrique choisie. La réponse est non pourvu que les paramétrisations soient équivalentes, et qu'elles aient même orientation.

THÉORÈME 9.3. Soient (α, U) et (β, V) des paramétrages équivalents. En supposant que les intégrales suivantes existent, on a

$$\iint_{\alpha(U)} f dS = \iint_{\beta(V)} f dS$$

DÉMONSTRATION. Soit \mathbf{G} tel que $\beta(s, t) = \alpha(\mathbf{G}(s, t))$ i.e. le changement de variables $(u, v) = \mathbf{G}(s, t)$. Alors :

$$\begin{aligned} \iint_{\alpha(U)} f dS &= \iint_U f(\alpha(u, v)) \left\| \frac{\partial \alpha}{\partial u} \wedge \frac{\partial \alpha}{\partial v} \right\| du dv \\ &= \iint_V f(\beta(s, t)) \left\| \frac{\partial \beta}{\partial s} \wedge \frac{\partial \beta}{\partial t} J_{\mathbf{G}^{-1}} \right\| J_{\mathbf{G}} ds dt \\ &= \iint_{\beta(V)} f dS \end{aligned}$$

□

5. Flux de champ à travers une surface

Si $\Sigma = \alpha(U)$ est une surface paramétrique, le vecteur

$$\mathbf{N} = \frac{\partial \alpha}{\partial u} \wedge \frac{\partial \alpha}{\partial v}$$

est normal à Σ en chaque point régulier de la surface. En chaque point, il y a deux normales unitaires, une \mathbf{n}^+ qui a la même direction que \mathbf{N} , et une autre \mathbf{n}^- de direction opposée. Ainsi, on a

$$\mathbf{n}^\pm = \pm \frac{\mathbf{N}}{\|\mathbf{N}\|}$$

Soit \mathbf{n} une de ces normales. Soit \mathbf{F} un champ de vecteurs défini sur Σ et supposons que l'intégrale de surface

$$\iint_{\Sigma} \mathbf{F} \cdot \mathbf{n} dS$$

existe, on a alors

$$\iint_{\Sigma} \mathbf{F} \cdot \mathbf{n} dS = \pm \iint_U \mathbf{F}[\alpha(u, v)] \bullet \frac{\partial \alpha}{\partial u} \wedge \frac{\partial \alpha}{\partial v} du dv$$

selon que $\mathbf{n} = \mathbf{n}^\pm$.

Cette intégrale de surface est appelée **flux du champ \mathbf{F} à travers la surface Σ** .

REMARQUE 9.4. Le flux d'un champ de vecteur à travers une surface dépend de l'orientation.

EXEMPLE 9.5. Le flot d'un fluide à travers une surface.

Nous considérons un fluide comme une collection de particules (ou de points). En chaque point, nous définissons un vecteur vitesse de la particule $\mathbf{V}(x, y, z)$. C'est le champ de vitesse du flot. Ce champ peut varier ou non selon le temps. Supposons ici qu'on soit à un état stationnaire, donc \mathbf{V} ne dépend que de la position de la particule et non du temps. Désignons par $\rho(x, y, z)$ la densité (masse par unité de volume) du fluide en un point (x, y, z) . Si le fluide est incompressible, la densité est constante dans le fluide. Pour un fluide compressible tel qu'un gaz, la densité peut varier selon le point

de l'espace. Dans tous les cas, la densité est un champ scalaire associé au mouvement. Le produit de la densité et de la vitesse est noté \mathbf{F}

$$\mathbf{F}(x, y, z) = \rho(x, y, z) \mathbf{V}(x, y, z)$$

ou encore $\mathbf{F} = \rho \mathbf{V}$. C'est un champ de vecteur appelé densité de flux du flot. Le vecteur \mathbf{F} a la même direction que la vitesse et ses dimensions sont

$$\frac{\text{masse}}{(\text{unité de surface})(\text{unité de temps})}$$

En d'autres termes, \mathbf{F} nous donne la masse de fluide par unité de surface et de temps en mouvement dans la direction \mathbf{V} au point (x, y, z) . Soit $\Sigma = \alpha(U)$ une surface paramétrique simple. En chaque point de Σ , soit \mathbf{n} la normale unitaire, c'est-à-dire :

$$\mathbf{n} = \frac{\frac{\partial \alpha}{\partial u} \wedge \frac{\partial \alpha}{\partial v}}{\left\| \frac{\partial \alpha}{\partial u} \wedge \frac{\partial \alpha}{\partial v} \right\|} \quad (5.1)$$

La quantité $\mathbf{F} \cdot \mathbf{n}$ représente alors la composante du vecteur densité de flux dans la direction \mathbf{n} . La masse de fluide en mouvement à travers Σ par unité de temps est définie par :

$$\iint_{\alpha(U)} \mathbf{F} \cdot \mathbf{n} dS = \int \int_U \mathbf{F} \cdot \mathbf{n} \left\| \frac{\partial \alpha}{\partial u} \wedge \frac{\partial \alpha}{\partial v} \right\| du dv$$

EXERCICE 9.4. Soit une gaine cylindrique de longueur L et de base circulaire de rayon ρ . La gaine est une conduite dans laquelle circule un fluide de température $T(r)$ ne dépendant que de la distance r à l'axe de la conduite. Sur l'axe de la conduite, la température est égale à T_1 , et à l'extérieur de la gaine, elle est égale à T_2 . La température du fluide est donnée par $T(r) = T_1 + (T_2 - T_1) \ln(\frac{r}{\rho})$. Calculer la quantité (i.e. le flux) de chaleur s'échappant de la gaine chaque seconde, sachant que le flux de chaleur passant à travers une surface Σ est donné par l'équation

$$\dot{Q} = -\lambda \iint_{\Sigma} \nabla T \cdot \mathbf{e}_r dS,$$

où λ est la conductivité thermique de la gaine et \mathbf{e}_r est le vecteur des coordonnées cylindriques, d'axe celui de la conduite.

EXERCICE 9.5. Soit Σ l'hémisphère $x^2 + y^2 + z^2 = 1, z \geq 0$, $\mathbf{F}(x, y, z) = x\mathbf{i} + y\mathbf{j}$. Soit \mathbf{n} la normale unitaire sortante à Σ . Calculer le flux du champ \mathbf{F} à travers la surface Σ , en utilisant la représentation

$$\alpha(u, v) = \sin u \cos v \mathbf{i} + \sin u \sin v \mathbf{j} + \cos u \mathbf{k}$$

6. Le théorème de Stokes

Rappelons le théorème de Green :

$$\iint_{\Sigma} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy = \int_C P dx + Q dy$$

où Σ est une région plane bornée par une courbe simple fermée C , traversée dans le sens direct.

Le théorème de Stokes est une extension directe du théorème de Green. Dans le théorème de Stokes, on ne demande plus que la surface et son contour soient dans le plan.

THÉORÈME 9.4. *On suppose que Σ est une surface paramétrique, de bord C fermé¹. On suppose que $\Sigma = \alpha(U)$ et $C = \alpha(\partial U)$, où*

- *U est une région dans le plan (u, v) bornée par une courbe de Jordan ∂U régulière par morceaux,*
- *α est une application bijective dont les composantes admettent des dérivées secondes continues sur un ouvert contenant $U \cup \partial U$ (i.e. Σ est une surface paramétrique simple, de classe \mathcal{C}^2).*

Soit ensuite un champ de vecteur \mathbf{f} de classe \mathcal{C}^1 sur Σ .

Alors nous avons

$$\iint_{\Sigma} \mathbf{rot} \mathbf{f} \cdot \mathbf{n} dS = \oint_{C^+} \mathbf{f} \cdot d\alpha \quad (6.1)$$

où \mathbf{n} est la normale \mathbf{n}^+ et α est une paramétrisation de C^+ , dont l'orientation est héritée de ∂U^+ par l'application α .

DÉMONSTRATION. Calculons $I = \int_{C^+} \mathbf{f} \cdot d\alpha$, et prenons la paramétrisation de C : $\alpha(u, v) = \alpha(u, v)$ définie sur ∂U . On a alors

$$I = \int_{\partial U^+} \mathbf{f}(\alpha(u, v)) \cdot \left(\frac{\partial \alpha}{\partial u} du + \frac{\partial \alpha}{\partial v} dv \right).$$

Posons alors

$$P = \mathbf{f}(\alpha(u, v)) \frac{\partial \alpha}{\partial u}, \quad Q = \mathbf{f}(\alpha(u, v)) \frac{\partial \alpha}{\partial v},$$

et appliquons le théorème de Green à $U \cup \partial U$:

$$I = \iint_U \left(\frac{\partial Q}{\partial u} - \frac{\partial P}{\partial v} \right) dudv.$$

Le calcul des dérivées partielles donne :

$$\frac{\partial Q}{\partial u} = \frac{\partial \mathbf{f}(\alpha(u, v))}{\partial u} \frac{\partial \alpha}{\partial v} + \mathbf{f}(\alpha(u, v)) \frac{\partial^2 \alpha}{\partial u \partial v}$$

$$\frac{\partial P}{\partial v} = \frac{\partial \mathbf{f}(\alpha(u, v))}{\partial v} \frac{\partial \alpha}{\partial u} + \mathbf{f}(\alpha(u, v)) \frac{\partial^2 \alpha}{\partial v \partial u}$$

et on obtient

$$I = \iint_U \left(\frac{\partial \mathbf{f}(\alpha(u, v))}{\partial u} \frac{\partial \alpha}{\partial v} - \frac{\partial \mathbf{f}(\alpha(u, v))}{\partial v} \frac{\partial \alpha}{\partial u} \right) dudv.$$

1. ici, fermé signifie fermé au sens bouclé, pas au sens mathématique du contraire de ouvert !!

On peut montrer que

$$\frac{\partial \mathbf{f}(\alpha(u, v))}{\partial u} \frac{\partial \alpha}{\partial v} - \frac{\partial \mathbf{f}(\alpha(u, v))}{\partial v} \frac{\partial \alpha}{\partial u} = \mathbf{rot} \mathbf{f} \cdot \mathbf{n} \left\| \frac{\partial \alpha}{\partial u} \wedge \frac{\partial \alpha}{\partial v} \right\|$$

et le théorème est ainsi démontré. □

REMARQUES 9.1. (1) On a supposé que ∂U est une courbe de Jordan car la preuve utilise le théorème de Green.

(2) Dans le cas où Σ est une région du plan (x, y) et $\mathbf{n} = \mathbf{k}$, on retrouve la formule de Green.

(3) Si Σ est une surface fermée, son bord est vide, l'intégrale curviligne et l'intégrale de surface du théorème de Stokes sont donc nulles.

(4) Sous certaines conditions, le **flux du rotationnel** de \mathbf{f} à travers la surface Σ est égal à la **circulation de \mathbf{f} le long du bord \mathcal{C}** de Σ . On parle parfois de **flux à travers une courbe fermée**, sans préciser la surface.

EXERCICE 9.6. Démontrer le théorème de Green, à partir du théorème de Stokes.

REMARQUE 9.5. Le rotationnel d'un gradient est nul. On en déduit que l'intégrale curviligne d'un gradient sur une courbe C de \mathbb{R}^3 , image d'une courbe de Jordan Γ par une application α bijective, régulière et de classe \mathcal{C}^2 , est nulle.

EXERCICE 9.7. Calculer à l'aide du théorème de Stokes l'intégrale

$$\iint_{\Sigma} \mathbf{rot} \mathbf{F} \cdot \mathbf{n} dS$$

avec $\mathbf{F}(x, y, z) = y^2 \mathbf{i} + xy \mathbf{j} + xz \mathbf{k}$, où Σ est l'hémisphère $x^2 + y^2 + z^2 = 1$, $z \geq 0$, et \mathbf{n} est la normale unitaire sortante (i.e. avec une composante positive selon z).

Vérifier votre résultat par le calcul direct de $\iint_{\Sigma} \mathbf{rot} \mathbf{F} \cdot \mathbf{n} dS$.

EXERCICE 9.8. Utiliser le théorème de Stokes pour montrer que l'intégrale curviligne est calculée par

$$\int_{C^+} y dx + z dy + x dz = -\pi a^2 \sqrt{3},$$

où C^+ est la courbe intersection de la sphère d'équation $x^2 + y^2 + z^2 = a^2$ et du plan $x + y + z = 0$, orientée positivement.

EXERCICE 9.9. (1) Soit la surface du cône $\Sigma = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = z^2, 0 \leq z \leq 1\}$. On donne la paramétrisation suivante pour Σ : $\mathbf{r}(u, v) = u \cos v \mathbf{i} + u \sin v \mathbf{j} + u \mathbf{k}$, $u \in [0, 1]$, $v \in [0, 2\pi]$. Donner l'expression de la normale à Σ . Quel est le bord $\partial \Sigma$ de la surface ? Donner en une paramétrisation.

(2) Vérifier le théorème de Stokes pour le champ vectoriel $\mathbf{f}(x, y, z) = (x^2 y, z^2, 0)$ et la surface Σ : i.e. calculer les deux intégrales de l'équation (6.1) pour vérifier qu'elles sont bien égales.

EXERCICE 9.10. (Dérivées normale). Nous avons vu, lors de l'étude des intégrales de lignes, que si $\mathbf{f} = P\mathbf{i} + Q\mathbf{j}$, l'intégrale curviligne par rapport à la longueur d'arc est telle que

$$\int_{\Gamma} Pdx + Qdy = \int_{\Gamma} \mathbf{f} \cdot \mathbf{T}ds,$$

où \mathbf{T} est le vecteur unitaire tangent à Γ .

- (1) Montrer que si Γ est une courbe de Jordan décrite par une fonction différentiable α , avec $\alpha(t) = X(t)\mathbf{i} + Y(t)\mathbf{j}$, alors une normale unitaire à Γ est donnée par

$$\mathbf{n}(t) = \frac{1}{\|\alpha'(t)\|} (Y'(t)\mathbf{i} - X'(t)\mathbf{j}),$$

dès que $\|\alpha'(t)\| \neq 0$.

Si φ est un champ scalaire de gradient $\nabla\varphi$ sur Γ , la dérivée normale extérieure à Γ , notée $\frac{\partial\varphi}{\partial\mathbf{n}}$ est définie par l'équation

$$\frac{\partial\varphi}{\partial\mathbf{n}} = \nabla\varphi \cdot \mathbf{n} = \partial_{\mathbf{n}}\varphi.$$

C'est en fait la dérivée directionnelle de φ selon la direction \mathbf{n} .

2. Si $\tilde{\mathbf{f}} = Q\mathbf{i} - P\mathbf{j}$, montrer que

$$\int_{\Gamma} Pdx + Qdy = \int_{\Gamma} \tilde{\mathbf{f}} \cdot \mathbf{n}ds.$$

$\tilde{\mathbf{f}} \cdot \mathbf{n}$ est appelé la composante normale du champ \mathbf{f} .

3. Soient f et g des champs scalaires ayant des dérivées partielles du premier et second ordre continues sur un ouvert S dans le plan (deux fois continûment différentiable). Soit Ω la région (dans S) dont la frontière est une courbe de Jordan Γ régulière par morceaux. Définissons le laplacien $\Delta u = \nabla^2 u = \partial_x^2 u + \partial_y^2 u$. Prouver que

(a)

$$\int_{\Gamma} \partial_{\mathbf{n}} g ds = \iint_{\Omega} \nabla^2 g dx dy.$$

(b)

$$\int_{\Gamma} f \partial_{\mathbf{n}} g ds = \iint_{\Omega} f \Delta g + \nabla f \cdot \nabla g dx dy.$$

(c)

$$\int_{\Gamma} (f \partial_{\mathbf{n}} g - g \partial_{\mathbf{n}} f) ds = \iint_{\Omega} (f \Delta g - g \Delta f) dx dy$$

(on appelle cette identité formule de Green)

(d) si f et g sont des fonctions harmoniques ($\Delta f = \Delta g = 0$), alors nous avons

$$\int_{\Gamma} f \partial_{\mathbf{n}} g ds = \int_{\Gamma} g \partial_{\mathbf{n}} f ds.$$

7. Reconstruction d'un champ de vecteurs à partir de son rotationnel

Nous nous sommes posé la question de savoir si un champ de vecteur donné dérive d'un gradient. La question que l'on se pose maintenant est la suivante : étant donné un champ de vecteurs \mathbf{f} , existe-t-il \mathbf{g} tel que $\mathbf{rot} \mathbf{g} = \mathbf{f}$?

Nous avons alors le théorème suivant.

THÉORÈME 9.5. *Soit \mathbf{f} , un champ de vecteurs continûment différentiable sur un ouvert Σ de \mathbb{R}^3 . Alors, il existe un champ de vecteurs \mathbf{g} tel que*

$$\mathbf{rot} \mathbf{g} = \mathbf{f}$$

si et seulement si

$$\operatorname{div} \mathbf{f} = 0$$

partout dans Σ .

*Un champ de vecteurs \mathbf{g} tel que $\mathbf{rot} \mathbf{g} = \mathbf{0}$ est appelé **irrotationnel** et un champ \mathbf{f} tel que $\operatorname{div} \mathbf{f} = 0$ est appelé **solénoïdal**.*

8. Le théorème de la divergence (théorème de Gauss ou d'Ostrogradski)

Le théorème de Stokes exprime une relation entre une intégrale double étendue à une surface et une intégrale de ligne prise le long d'une ou plusieurs courbes formant la frontière de cette surface. Le théorème de la divergence exprime une relation entre une intégrale triple étendue à un solide et une intégrale de surface prise sur la frontière de ce solide.

THÉORÈME 9.6. *Soit V , un solide de \mathbb{R}^3 borné par une surface fermée² orientable Σ et soit \mathbf{n} la normale unitaire sortante à Σ . Si \mathbf{f} est un champ de vecteurs continûment différentiable défini sur V , nous avons*

$$\iiint_V \operatorname{div} \mathbf{f} \, dx dy dz = \iint_{\Sigma} \mathbf{f} \cdot \mathbf{n} dS \quad (8.1)$$

DÉMONSTRATION. Elle peut être trouvée dans [4] par exemple. □

REMARQUE 9.6. Orientable, dans le cas d'une surface fermée, signifie que l'on peut distinguer la face intérieure et la face extérieure de la surface. (Il existe des surfaces fermées 'biscornues', où cela n'est pas possible).

2. fermée au sens bouclée, pas au sens mathématique contraire de ouvert

EXERCICE 9.11. Soit φ un champ scalaire de classe \mathcal{C}^2 sur \mathbb{R}^3 , qui ne s'annule jamais et vérifiant

$$\|\nabla\varphi\|^2 = \lambda\varphi$$

et

$$\operatorname{div}(\varphi\nabla\varphi) = \mu\varphi,$$

les paramètres λ et μ étant fixés et réels. Evaluer l'intégrale

$$\iint_{\Gamma} \frac{\partial\varphi}{\partial\mathbf{n}} dS = \iint_{\Gamma} \nabla\varphi \cdot \mathbf{n} dS,$$

où Σ est la surface d'un objet V compact de l'espace tridimensionnel, et \mathbf{n} est la normale unitaire sortante à Σ .

EXERCICE 9.12. Vérifier le théorème de la divergence (i.e. calculer les deux intégrales de l'équation (8.1) pour vérifier qu'elles ont bien égales) avec $\mathbf{F}(x, y, z) = (xz, y, y)$ et $V = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq 1\}$.

EXERCICE 9.13. Soit le champ de vecteurs

$$\mathbf{F}(x, y, z) = (0, 0, xyz)$$

défini sur \mathbb{R}^3 .

Soit la boule unité paramétrée par

$$x = r \cos \theta \sin \varphi, \quad y = r \sin \theta \sin \varphi, \quad z = r \cos \varphi$$

avec $\theta \in [0, 2\pi]$, $\varphi \in [0, \pi]$ et $r \in [0, 1]$.

- (1) Calculer $\iiint_V \operatorname{div} \mathbf{F} dx dy dz$ où V est la boule unité.
- (2) Calculer $\iint_{\Sigma} \mathbf{F} \cdot \mathbf{n} dS$ où \mathbf{n} est la normale unitaire sortante à la boule unité, et Σ est la sphère unité.
- (3) Commenter ces deux résultats.

Introduction et position du problème

Les tentatives de représentation et de modélisation de phénomènes évolutifs ont donné lieu à l'étude des systèmes dynamiques. Le phénomène évolutif est alors décrit au moyen d'une relation mathématique qui à partir de conditions initiales détermine le futur du phénomène. Parmi l'ensemble des systèmes dynamiques (cf cours d'automatique), ceux représentés par des équations différentielles sont historiquement à la base de la modélisation des phénomènes physiques.

Une équation différentielle ordinaire d'ordre n est une équation liant une fonction à ses dérivées jusqu'à l'ordre n par une relation du type :

$$F(t, x(t), \dot{x}(t), \dots, x^{(n)}(t)) = 0 \quad (0.1)$$

où F est une fonction définie sur une partie de \mathbb{R}^{n+2} et à valeurs dans \mathbb{R}^m . Une solution est alors la donnée d'une fonction de classe C^n à valeurs dans \mathbb{R} et vérifiant (0.1) sur un intervalle I de \mathbb{R} .¹

Commençons par observer qu'une équation différentielle d'ordre n (0.1) peut toujours se mettre sous la forme d'une équation différentielle d'ordre 1. En effet, si on note par $u = (u_1, u_2, \dots, u_n)$ le vecteur de \mathbb{R}^n et si on définit la fonction $\tilde{F} : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^{n-1} \times \mathbb{R}^m$ par :

$$\tilde{F}(t, u, v) = \begin{bmatrix} v_1 - u_2 \\ v_2 - u_3 \\ \dots \\ v_{k-1} - u_k \\ F(t, u_1, u_2, \dots, u_k, v_k) \end{bmatrix} \quad (0.2)$$

Alors, en posant $u(t) = (x(t), \dot{x}(t), \dots, x^{(n-1)}(t)) \in \mathbb{R}^n$, on voit que l'équation (0.1) se ramène à l'étude de l'équation différentielle d'ordre 1 :

$$\tilde{F}(t, u, \dot{u}) = 0 \quad (0.3)$$

1. De manière plus générale, on peut envisager de remplacer \mathbb{R}^n et \mathbb{R}^m ci-dessus par des espaces vectoriels normés complets (Espace de Banach). Si ces espaces sont de dimension finie, cela ne porte pas à conséquence. En revanche dans le cas inverse, les problèmes posés par la dimension infinie sort du cadre de ce cours.

Nous venons donc de réduire l'ordre de l'équation différentielle initiale en augmentant la dimension de l'espace de travail. C'est donc la forme générale (0.3) plus facile à manipuler que nous retiendrons par la suite.

Exemple : L'équation d'ordre 2 : $\ddot{x} - \dot{x} + x = f(t)$ peut se ramener en posant

$$u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} x \\ \dot{x} \end{bmatrix}$$

à l'équation d'ordre 1 :

$$\dot{u} = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix} u + \begin{bmatrix} 0 \\ f(t) \end{bmatrix}$$

Vocabulaire : Dans l'expression $F(t, x, \dot{x}) = 0$, la variable x ($x(t) \in \mathbb{R}^n$) est appelée état du système, \mathbb{R}^n espace d'état et t le temps. On comprend alors que l'état x permet de caractériser l'état (position, vitesse, température,...) d'un système physique et sa dimension (ici n) le nombre de variable nécessaire pour décrire le phénomène. L'évolution du système physique est donc vue comme celle d'un point dans un espace de dimension n .

Dans ce cours, nous nous limiterons à l'étude de la forme explicite (la dérivée \dot{x} est une fonction de t et x) :

$$\tilde{F}(t, x, \dot{x}) = \dot{x}(t) - f(t, x(t)) = 0 \iff \dot{x}(t) = f(t, x(t)) \quad (0.4)$$

En effet, la forme implicite (0.3) ne se prête pas à une étude générale pour la mise en évidence de l'existence d'une solution. Confronté à une forme implicite, on peut essayer de se ramener à une forme explicite ou si ce n'est pas possible consulter des ouvrages traitant du problème par exemple l'ouvrage [9].

Sous cette dernière forme (0.4), la solution n'est pas définie sans équivoque. Il nous faut donc ajouter une condition supplémentaire : la donnée d'une condition initiale portant sur la valeur de la solution à un instant t_0 , soit $x(t_0) = x_0$.

On aboutit alors au problème dit de Cauchy :

Problème de Cauchy :

Trouver une fonction x de classe C^1 définie sur un intervalle I contenant t_0 à valeurs dans \mathbb{R}^n et vérifiant :

$$\begin{cases} \dot{x}(t) = f(t, x(t)) \text{ sur } I \\ x(t_0) = x_0 \end{cases} \quad (0.5)$$

Avant de chercher à résoudre ce problème, une question essentielle est de savoir, sous quelles conditions simples peut-on garantir l'existence et l'unicité d'une solution ?

Exemples :

a. L'équation différentielle,

$$\begin{cases} \dot{x} = x^{2/3} \\ x(0) = 0 \end{cases}$$

admet les deux solutions :

$$\begin{aligned} x_1(t) &= 0, \quad \forall t \in \mathbb{R} \\ x_2(t) &= \left(\frac{t}{3}\right)^3, \quad \forall t \in \mathbb{R} \end{aligned}$$

b. L'équation différentielle,

$$\begin{cases} \dot{x} = x^2 \\ x(0) = 1 \end{cases}$$

a pour solution $x = \frac{1}{1-t}$ qui est définie uniquement sur $] -\infty, 1[$.

Ces deux exemples nous montrent que la solution lorsqu'elle existe, n'est pas nécessairement unique et définie pour tout instant t .

Un des objectifs de ce cours est donc de présenter deux résultats fondamentaux : l'un pour l'existence d'une solution avec le Théorème de Peano, l'autre pour l'existence et l'unicité de la solution avec le Théorème de Cauchy-Lipschitz. Le premier théorème présente l'avantage d'être constructif et est à mettre en parallèle avec la méthode de résolution numérique d'Euler (voir la fin du cours). Le second utilise le théorème des approximations successives et illustre bien les techniques employées en analyse numérique.

Nous aborderons également la notion de solution maximale qui permet de déterminer le plus grand intervalle de temps pour lequel la solution existe. Ceci nous amènera à nous interroger sur la stabilité de la solution et sa sensibilité aux variations de paramètres.

Une classe particulièrement importante pour laquelle un certain nombre de résultats généraux peuvent être énoncés, est la classe des équations différentielles linéaires. Une section entière lui est donc consacrée où les notions d'opérateur résolvant et d'exponentielle de matrice seront abordées. En complément de cette section, nous verrons comment l'étude d'un système différentiel non linéaire peut se ramener localement à l'étude d'un système linéaire par linéarisation.

Enfin, une brève introduction à la résolution numérique des équations différentielles donnera une idée au lecteur des techniques employées pour la résolution et la simulation des systèmes dynamiques.

Etude générale

1. Existence de la solution : Le théorème de Peano

Au préalable, on va avoir besoin de la définition et du théorème d'analyse fonctionnelle suivant :

DÉFINITION 11.1. *Une suite de fonctions (f_n) définies sur un espace métrique (E, d) et à valeurs dans \mathbb{R}^m (ou \mathbb{C}^m) est équicontinue en x_0 si*

$$\forall \varepsilon > 0 \quad \exists \delta \quad \forall n \quad \forall x, \quad d(x, x_0) < \delta \Rightarrow \|f_n(x) - f_n(x_0)\| < \varepsilon$$

C'est donc le même δ qui assure la continuité des fonctions f_n . Autrement dit, δ ne dépend pas de n mais uniquement de la famille (f_n) .

THÉORÈME 11.1. (*d'Ascoli*) *Soit (K, d) un espace métrique compact. Soit (f_n) une suite équicontinue de fonctions définies sur K à valeurs dans \mathbb{R}^m (ou \mathbb{C}^m) telle que $\forall x \in K, \sup_n \|f_n(x)\| < \infty$. Alors on peut extraire une sous suite (f_{n_k}) qui converge uniformément sur K .*

Ce théorème caractérise les parties compactes de l'espace des fonctions continues sur un compact. On rappelle qu'une caractéristique d'un ensemble compact K est que de toute suite définie sur K on peut extraire une sous suite qui converge sur K .

Le deuxième exemple introductif nous a permis de constater que l'existence d'une solution n'avait pas un caractère global mais plutôt local. Le bon cadre pour étudier l'existence d'une solution est donc de se placer dans un voisinage des conditions initiales et de supposer la fonction f continue dans ce voisinage. On montre alors qu'il existe une solution locale au problème (1.1) avec le théorème :

THÉORÈME 11.2. (*Peano*) *Si f est définie et continue dans un voisinage V ouvert de (t_0, x_0) , alors il existe une solution x de classe C^1 de l'équation*

$$\begin{cases} \dot{x}(t) &= f(t, x(t)) \\ x(t_0) &= x_0 \end{cases} \quad (1.1)$$

sur un voisinage de t_0 .

DÉMONSTRATION. La démonstration est constructive. Tout d'abord, on montre que le problème initial : Trouver x de classe C^1 sur un voisinage de t_0 , solution de (1.1) est équivalent au problème : déterminer x continue sur un voisinage de t_0 et vérifiant l'équation intégrale :

$$x(t) = x_0 + \int_{t_0}^t f(s, x(s)) ds \quad (1.2)$$

Par intégration, il est évident que (1.1) implique (1.2). Inversement, sur un voisinage de t_0 comme x est continue, $s \mapsto f(s, x(s))$ est continue et $t \mapsto \int_{t_0}^t f(s, x(s)) ds$ est de classe C^1 . On en conclut que x est C^1 sur un voisinage de t_0 et par dérivation on retrouve (1.1).

Nous allons montrer à présent qu'il existe des solutions continues sur un voisinage de t_0 à (1.2).

Puisque f est continue sur un voisinage ouvert V de (t_0, x_0) , il existe un compact K de la forme $[a, b] \times \overline{\mathbf{B}}(x_0, R)$ contenant le point (t_0, x_0) et inclus dans V (il y a toujours de la place dans un ouvert). Alors f est bornée sur K . Posons :

$$M = \sup_{(t,x) \in K} \|f(t, x)\|$$

$$\tau = \frac{R}{M}$$

et considérons l'intervalle $I = [\max(a, t_0 - \tau), \min(b, t_0 + \tau)] \subset [a, b]$ (voir figure 1).

Posons également $P_0 = (t_0, x_0)$ et construisons pour $\varepsilon > 0$ fixé, le segment $[P_0, P_1]$ de pente $f(t_0, x_0)$ où le point $P_1 = (t_1, x_1)$ est déterminé de telle sorte que $t_1 \in I$ et $\forall P = (t, x) \in [P_0, P_1]$, $\|f(t, x) - f(t_0, x_0)\| < \varepsilon$ (voir figure 2).

On construit alors par récurrence, une suite **finie** de points $P_k = (t_k, x_k)$ sur tout l'intervalle I , car f est uniformément continue sur K (on utilise le même procédé à gauche de t_0). Notons que la suite de points P_k appartient bien au compact K . C'est évidemment par construction le cas pour les t_k . Pour les x_k , comme :

$$x_{k+1} - x_k = f(t_k, x_k)(t_{k+1} - t_k)$$

il est aisé de montrer par récurrence que $\|x_k - x_0\| < M|t_k - t_0| < M\tau = R$ i.e. $x_k \in \overline{\mathbf{B}}(x_0, R)$.

Cette suite de segments définit une fonction affine par morceaux $x_\varepsilon : I \rightarrow \mathbb{R}^n$ telle que,

$$\|x_\varepsilon(t') - x_\varepsilon(t)\| < M|t' - t|$$

$$\|x_\varepsilon(t)\| < M|t - t_0| + \|\dots\|$$

Pour une suite $\varepsilon_n \rightarrow 0$, x_{ε_n} définit une suite uniformément bornée et équicontinue de fonctions sur I . Les conditions du théorème d'Ascoli étant vérifiées, il existe une sous suite extraite x_{ε_m} qui converge uniformément sur I vers une certaine fonction x .

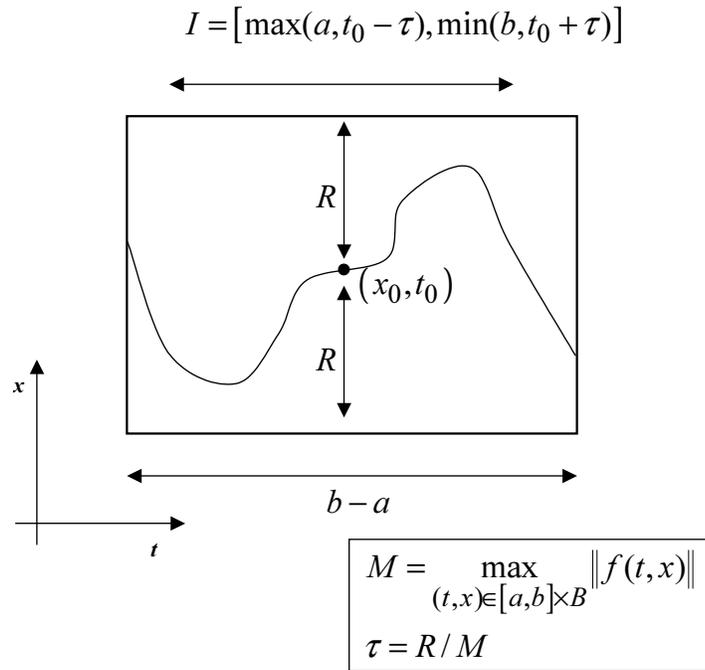


FIGURE 1 – Choix du compact

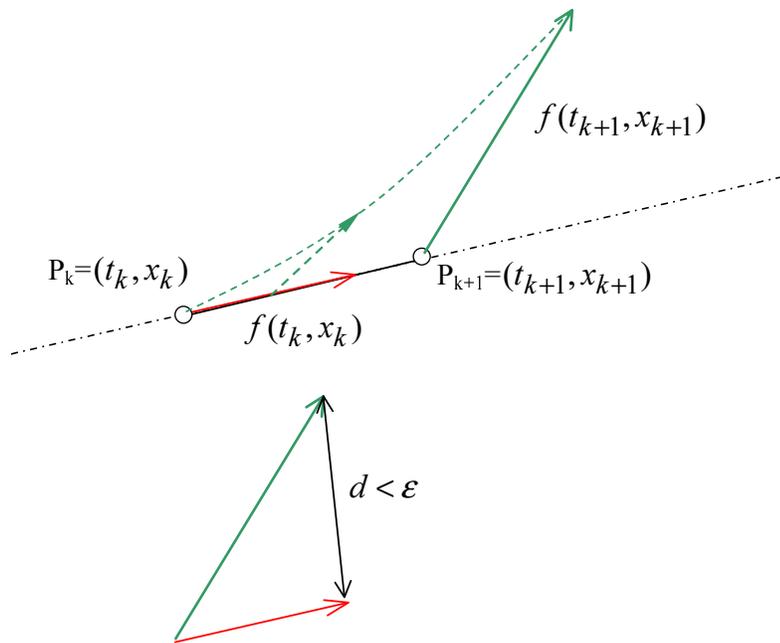


FIGURE 2 – Construction d'une trajectoire approchée

Montrons alors que x vérifie (1.2) pour tout $t \in I$. Evaluons la quantité $A(t)$ pour $t \in I$:

$$\begin{aligned} A(t) &= \|x(t) - x_0 - \int_{t_0}^t f(s, x(s)) ds\| \\ &\leq \|x(t) - x_{\varepsilon_m}(t)\| + \|x_{\varepsilon_m}(t) - x_0 - \int_{t_0}^t f(s, x(s)) ds\| \end{aligned}$$

Comme la fonction x_{ε_m} est définie sur une union finie d'intervalles $I_k(\varepsilon_m)$ dépendant de ε_m , pour $t \in I_k(\varepsilon_m) = [t_k, t_{k+1}[$, $x_{\varepsilon_m}(t)$ s'écrit :

$$x_{\varepsilon_m}(t) = x_0 + \sum_{i=0}^{k-1} f(t_i, x_i)(t_{i+1} - t_i) + f(t_k, x_k)(t - t_k)$$

La quantité $B(t) = \|x_{\varepsilon_m}(t) - x_0 - \int_{t_0}^t f(s, x(s)) ds\|$ correspondant au second membre de l'inégalité précédente, peut être majorée par :

$$\begin{aligned} B(t) &\leq \sum_{i=0}^k \int_{t_i}^{t_{i+1}} \|f(s, x(s)) - f(t_i, x_i)\| ds \\ B(t) &\leq \sum_{i=0}^k \int_{t_i}^{t_{i+1}} \|f(s, x(s)) - f(s, x_{\varepsilon_m}(s))\| + \|f(s, x_{\varepsilon_m}(s)) - f(t_i, x_i)\| ds \end{aligned}$$

Pour ε fixé, comme x_{ε_m} converge uniformément vers x et comme f est uniformément continue sur K , il existe un indice m_0 tel que $\forall m > m_0, \forall s \in I$:

$$\|f(s, x(s)) - f(s, x_{\varepsilon_m}(s))\| \leq \varepsilon$$

et comme par construction,

$$\|f(s, x_{\varepsilon_m}(s)) - f(t_i, x_i)\| ds \leq \varepsilon_m$$

on a la majoration dès que $\varepsilon_m \leq \varepsilon$:

$$B(t) \leq 2\varepsilon |t_{k+1} - t_0| \leq 2\varepsilon \text{mes}(I)$$

Enfin, puisque x_{ε_m} converge uniformément vers x , $A(t)$ peut être majorée par un nombre arbitrairement petit indépendant de t ce qui montre que x vérifie (1.2) pour tout $t \in I$. □

PROPOSITION 11.1. (Régularité de la solution) *Si x est solution de (1.1) sur un voisinage V de (t_0, x_0) et si f est de classe C^k sur ce voisinage alors $x \in C^{k+1}$.*

2. Solution maximale

Le procédé utilisé établit l'existence d'une solution sur un voisinage de t_0 . En se plaçant aux extrémités de cet intervalle de temps, on peut de proche en proche réitérer la procédure et rechercher le plus grand intervalle de temps sur lequel il existe une solution.

DÉFINITION 11.2. *Une solution (I, x) est dite maximale si toute solution (J, u) telle que $I \subset J$ et $u|_I = x$, est nécessairement identique à (I, x) . On dit alors que (I, x) n'est pas prolongeable. Si $I = \mathbb{R}$, on dit que la solution est globale dans le temps.*

On a alors la proposition :

PROPOSITION 11.2. *Si $f \in C([t_0, +\infty[\times \mathbb{R}^n, \mathbb{R}^n)$. Si x est une solution maximale de (1.1) à droite de t_0 (on peut faire la même chose à gauche) c'est à dire sur un intervalle de la forme $I = [t_0, T[$, on a l'alternative :*

- (1) ou bien $T = +\infty$
- (2) ou bien $T < +\infty$ et $\lim_{t \rightarrow T} \|x(t)\| = +\infty$. On dit alors que la solution explose en temps fini.

DÉMONSTRATION. Si $T < +\infty$ et $\lim_{t \rightarrow T} \|x(t)\| \leq R < +\infty$. Alors pour toute suite croissante (t_k) telle que $t_k \rightarrow T$, on a $\|x(t_k)\| \leq R$. La suite $x(t_k)$ de \mathbb{R}^n est donc bornée et on peut en extraire une sous suite qui converge vers x_1 ($\|x_1\| \leq R$). Comme f est continue il existe τ et R_1 tels que sur le compact $K = [T - \tau, T + \tau] \times \overline{\mathbf{B}}(x_1, R_1)$ on ait : $\tau < R_1/M$ où $M = \sup_{(t,x) \in K} \|f(t,x)\|$. Alors pour t_k suffisamment proche de T , c'est à dire $|T - t_k| < \tau/4$ et $\|x(t_k) - x_1\| < R_1/2$, il existe une solution sur $[t_k - \tau/2, t_k + \tau/2]$ à l'équation différentielle $\dot{x} = f(t, x)$ pour la condition initiale $x(t_k)$. On a donc prolongé la solution au delà de T ce qui contredit l'hypothèse de solution maximale. □

3. Unicité de la solution : Le théorème de Cauchy-Lipschitz

Le théorème précédent nous a permis de mettre en évidence l'existence d'une solution. Dans ce chapitre, nous allons utiliser une condition supplémentaire sur f qui va nous permettre de garantir l'unicité de la solution.

DÉFINITION 11.3. *La fonction f satisfait une condition de Lipschitz en x sur un voisinage V de (t_0, x_0) s'il existe une constante positive C telle que*

$$\|f(t, x_2) - f(t, x_1)\| < C\|x_2 - x_1\|$$

pour tous les couples (t, x_1) et (t, x_2) dans V

THÉORÈME 11.3. (Cauchy-Lipschitz) *On suppose que f est définie, continue dans un voisinage ouvert de (t_0, x_0) et satisfait une condition de Lipschitz en x sur ce même voisinage. Alors l'équation différentielle :*

$$\begin{cases} \dot{x}(t) &= f(t, x(t)) \\ x(t_0) &= x_0 \end{cases} \quad (3.1)$$

a une solution unique sur un voisinage de t_0

DÉMONSTRATION. Comme dans la démonstration du théorème de Peano, on choisit un compact K de la forme $I \times \overline{\mathbf{B}}(x_0, R)$ tel que toutes solutions définies sur I appartiennent à la boule $\overline{\mathbf{B}}(x_0, R)$.

On considère l'application $\Psi : C(I, \overline{\mathbf{B}}(x_0, R)) \rightarrow C(I, \overline{\mathbf{B}}(x_0, R))$ qui à tout élément $x \in C(I, \overline{\mathbf{B}}(x_0, R))$ associe la fonction u définie par $u(t) = x_0 + \int_{t_0}^t f(s, x(s)) ds$ pour tout $t \in I$. Notons que le choix de l'intervalle I garantit pour tout $t \in I$, $u(t) \in \overline{\mathbf{B}}(x_0, R)$. Il nous faut donc prouver qu'il existe un x tel que $\Psi(x) = x$.

Soient x_1 et x_2 deux fonctions de $C(I, \overline{\mathbf{B}}(x_0, R))$. On a alors :

$$\begin{aligned} \|\Psi(x_2(t)) - \Psi(x_1(t))\| &\leq \int_{t_0}^t \|f(s, x_2(s)) - f(s, x_1(s))\| ds \\ \|\Psi(x_2(t)) - \Psi(x_1(t))\| &\leq C \int_{t_0}^t \|x_2(s) - x_1(s)\| ds \quad \text{Lipschitz} \\ \|\Psi(x_2(t)) - \Psi(x_1(t))\| &\leq C|t - t_0| \|x_2 - x_1\|_\infty \end{aligned}$$

Montrons alors par récurrence que :

$$\forall t \in I, \|\Psi^k(x_2(t)) - \Psi^k(x_1(t))\| \leq \frac{C^k |t - t_0|^k}{k!} \|x_2 - x_1\|_\infty$$

C'est déjà fait pour $k = 1$. Supposons donc que la relation est vérifiée jusqu'à l'ordre k , alors on peut écrire $\forall t \in I$:

$$\begin{aligned} \|\Psi^{k+1}(x_2(t)) - \Psi^{k+1}(x_1(t))\| &\leq \int_{t_0}^t \|f(s, \Psi^k(x_2(s))) - f(s, \Psi^k(x_1(s)))\| ds \\ \|\Psi^{k+1}(x_2(t)) - \Psi^{k+1}(x_1(t))\| &\leq C \int_{t_0}^t \|\Psi^k(x_2(s)) - \Psi^k(x_1(s))\| ds \quad \text{Lipschitz} \\ \|\Psi^{k+1}(x_2(t)) - \Psi^{k+1}(x_1(t))\| &\leq \frac{C^{k+1}}{k!} \|x_2 - x_1\|_\infty \int_{t_0}^t |s - t_0|^k ds \\ \|\Psi^{k+1}(x_2(t)) - \Psi^{k+1}(x_1(t))\| &\leq \frac{C^{k+1} |t - t_0|^{k+1}}{(k+1)!} \|x_2 - x_1\|_\infty \end{aligned}$$

Posons $\delta = \max_{t \in I} |t - t_0|$.

$$\|\Psi^k(x_2) - \Psi^k(x_1)\|_\infty \leq \frac{C^k \delta^k}{(k)!} \|x_2 - x_1\|_\infty$$

On voit donc que l'application Ψ^k est lipschitzienne de rapport $\frac{C^k \delta^k}{(k)!}$ avec C et δ fixés. Comme $\frac{C^k \delta^k}{(k)!} \rightarrow 0$ lorsque $k \rightarrow \infty$, Ψ^k est contractante à partir d'un certain rang. Le théorème des approximations successives permet de conclure (car l'espace des fonctions continues sur un compact est complet pour la norme infini) :

- (1) Il existe $x \in C(I, \overline{\mathbf{B}}(x_0, R))$ unique tel que $\Psi(x) = x$.
- (2) $\forall u \in C(I, \overline{\mathbf{B}}(x_0, R))$, $\Psi^n(u)$ converge uniformément vers x .

□

4. Stabilité de la solution

4.1. Dépendance continue aux conditions initiales. Une question importante est de savoir comment une erreur sur la condition initiale x_0 se propage au cours du temps. Cette question est particulièrement cruciale lorsque l'on a recours à l'analyse numérique pour le calcul d'une solution. Le théorème suivant permet d'estimer cette erreur :

THÉORÈME 11.4. *On suppose que f est définie, continue et satisfait une condition de Lipschitz en x dans un voisinage ouvert V de (t_0, x_0) .*

Soient x et y les solutions de (3.1) pour les conditions initiales (t_0, x_0) et (t_0, y_0) respectivement ($\in V$). Alors il existe $T > t_0$ et $C > 0$ tel que pour tout $t \in [t_0, T]$,

$$\|x(t) - y(t)\| \leq \|x_0 - y_0\| e^{C(t-t_0)}$$

et l'application qui à $x_0 \mapsto x$ sur $[t_0, T]$ est continue. On dit alors que la solution dépend continûment des conditions initiales sur V .

DÉMONSTRATION. On sait que :

$$x(t) - y(t) = x_0 - y_0 + \int_{t_0}^t f(s, x(s)) - f(s, y(s)) ds.$$

En posant : $\Delta(t) = \|x(t) - y(t)\|$, comme f satisfait une condition de lipschitz, on obtient l'inégalité :

$$\Delta(t) \leq \Delta(t_0) + \int_{t_0}^t C \Delta(s) ds$$

Le **Lemme de Gronwall** implique le résultat :

LEMME 11.1. *Soient $a > 0$ une constante positive, b et c deux fonctions positives respectivement intégrable et continue sur $[t_0, T]$ telles que :*

$$\forall t \in [0, T], \quad c(t) \leq a + \int_{t_0}^t b(s)c(s) ds$$

alors

$$c(t) \leq a \exp\left(\int_{t_0}^t b(s) ds\right)$$

Posons $\varphi(t) = a + \int_{t_0}^t b(s)c(s)ds$. φ est une fonction continue et dérivable presque partout et on a :

$$\varphi'(t) = b(t)c(t) \leq b(t)\varphi(t)$$

ce qui signifie que :

$$\frac{d}{dt} \left(\varphi(t) \exp \left(- \int_{t_0}^t b(s)ds \right) \right) \leq 0$$

d'où par intégration,

$$\varphi(t) \leq \varphi(t_0) \exp \left(\int_{t_0}^t b(s)ds \right)$$

On conclut en notant que $\varphi(t_0) = a$.

□

4.2. Dépendance continue aux paramètres. On suppose dans cette partie que la fonction f dépend également d'un paramètre $\alpha \in \Omega \subset \mathbb{R}^d$ ouvert i.e.

$$\begin{cases} \dot{x}(t) &= f(t, x(t), \alpha) \\ x(t_0) &= x_0 \end{cases} \quad (4.1)$$

THÉORÈME 11.5. *Soit f définie, continue et satisfaisant une condition de Lipschitz en (x, α) dans un voisinage ouvert V de (t_0, x_0, α_0) .*

Soient x et y les solutions de (4.1) pour les conditions initiales (t_0, x_0, α_1) et (t_0, x_0, α_2) respectivement ($\in V$). Alors il existe $T > t_0$ et $C > 0$ tels que pour tout $t \in [t_0, T]$,

$$\|x(t) - y(t)\| \leq \|\alpha_1 - \alpha_2\| e^{C(t-t_0)}$$

et l'application qui a $\alpha_0 \mapsto x$ sur $[t_0, T]$ est continue. On dit alors que la solution dépend continûment du paramètre α .

DÉMONSTRATION. Il suffit d'ajouter l'équation $\dot{\alpha} = 0$ à (4.1) et d'appliquer le résultat de la section précédente. □

5. Représentations graphiques

5.1. Espaces des phases et portrait de phases. On considère une équation différentielle de la forme

$$x^{(n)}(t) = F(t, x(t), \dots, x^{(n-1)}(t))$$

avec F fonction continue sur $I \times \mathbb{R}^k \times \dots \times \mathbb{R}^k$ et à valeurs dans \mathbb{R}^k . On suppose de plus que $F(t, x_1, \dots, x_n)$ est localement Lipschitz par rapport à (x_1, \dots, x_n) . On a vu en exercice que le théorème de Cauchy-Lipschitz se reformule dans ce cadre et donne l'existence et l'unicité d'une solution maximale à condition initiale du type

$$\begin{aligned}x(t_0) &= x_1^0 \\x'(t_0) &= x_2^0 \\&\vdots \\x^{(n-1)}(t_0) &= x_n^0\end{aligned}$$

fixée. La donnée du vecteur $(x(t), \dots, x^{(n-1)}(t))$ détermine complètement l'évolution du système.

DÉFINITION 11.4. *La donnée de $(x(t), \dot{x}(t), \dots, x^{(n-1)}(t))$ s'appelle la phase du système au temps t . L'espace $(\mathbb{R}^k)^n$ où vit $(x(t), \dot{x}(t), \dots, x^{(n-1)}(t))$ s'appelle l'espace des phases ou encore espace des états.*

Une courbe $t \mapsto (x(t), \dot{x}(t), \dots, x^{(n-1)}(t))$ dans l'espace des phases s'appelle une trajectoire de phase.

La représentation graphique d'une famille de trajectoires de phase dans l'espace des phases, s'appelle un portrait de phase.

EXEMPLE 11.1. Considérons l'équation (sans les constantes physiques) d'un oscillateur harmonique

$$\ddot{x} + x = 0.$$

Le portrait de phase est alors constitué de cercles concentriques centrés en zéro. En effet les solutions de cette équation sont les fonctions de la forme $x : t \mapsto \lambda \cos(t) + \mu \sin(t)$ et on calcule $x(t)^2 + \dot{x}(t)^2 = \lambda^2 + \mu^2$, ce qui montre que la solution parcourt le cercle de rayon $\sqrt{\lambda^2 + \mu^2}$. Voir [figure 3](#).

5.2. Courbes intégrales. Il ne faut pas confondre les trajectoire de phases avec les *courbes intégrales* qui sont les graphes des fonctions y en fonction de x pour une équation du type

$$\dot{y} = f(x, y).$$

EXEMPLE 11.2. L'équation de l'oscillateur harmonique donne une équation différentielle d'ordre 1 qui s'écrit

$$\dot{Y} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} Y$$

avec Y qui joue le rôle de (x, \dot{x}) pour x solution de l'équation de l'oscillateur harmonique. On obtient alors la représentation des courbes intégrales sur [figure 4](#).

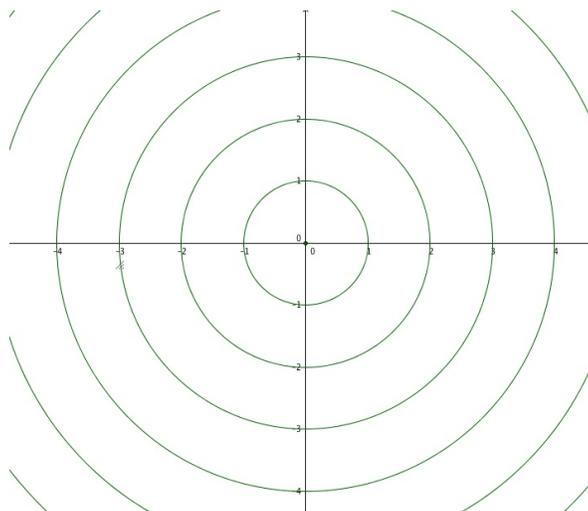


FIGURE 3 – Portrait de phases de l'oscillateur harmonique en coordonnées (x, \dot{x}) .

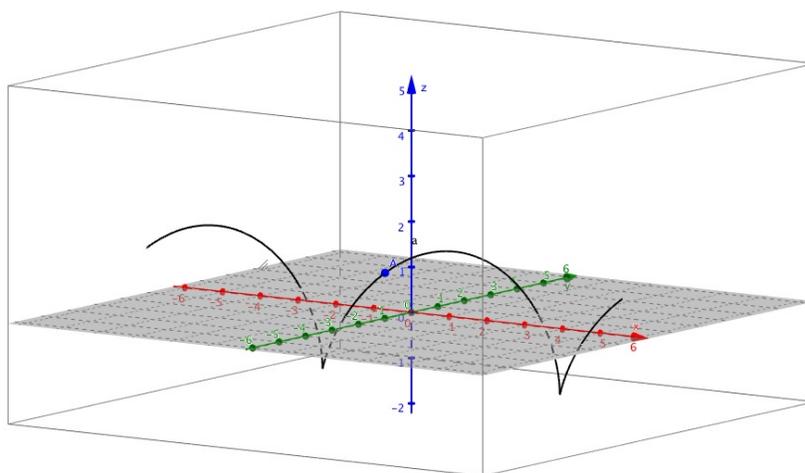


FIGURE 4 – Courbes intégrales de l'oscillateur harmonique

Rappels d'algèbre

1. Espace vectoriel de dimension finie

Soit $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} .

On supposera connues les notions d'espace vectoriel, sous-espaces vectoriel, base, matrice, *etc.*

2. Matrice d'une transformation linéaire

Soient E et F deux \mathbb{K} -espaces vectoriels de dimensions finies et $\mathbf{f} : E \rightarrow F$ une application linéaire.

On associe à une \mathbf{f} la notion de **matrice**, qui est une représentation pratique, pour deux bases \mathcal{B} et \mathcal{B}' respectivement de E et F , de cette application. En ce sens, cette représentation n'est pas intrinsèque (c'est-à-dire indépendante de la base ou de l'observateur). Nous reviendrons sur cette problématique dans le cadre plus général des tenseurs définis comme des applications multi-linéaires.

DÉFINITION 12.1. *Etant donnés E et F dotés de bases $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ et $\mathcal{B}' = (\mathbf{e}'_1, \dots, \mathbf{e}'_m)$ respectivement, $\mathbf{f} \in \mathcal{L}(E, F)$, la matrice*

$$A = (a_{i,j}) = \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{bmatrix}$$

est définie par rapport aux bases \mathcal{B} et \mathcal{B}' par

$$\mathbf{f}(\mathbf{e}_j) = \sum_{i=1}^m a_{i,j} \mathbf{e}'_i, \quad \text{pour } j = 1, \dots, n$$

On a donc une relation entre une quantité fonctionnelle et une représentation matricielle de la forme

$$\mathbf{y} = \mathbf{f}(\mathbf{x}) \iff \mathbf{Y} = \mathbf{A}\mathbf{X},$$

pour des vecteurs $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i$, $\mathbf{y} = \sum_{j=1}^m y_j \mathbf{e}'_j$, \mathbf{X} et \mathbf{Y} , où \mathbf{X} et \mathbf{Y} sont les vecteurs colonnes dont les composantes sont les x_i et y_j .

REMARQUE 12.1. *On note parfois $\mathcal{M}_{n,m}(\mathbb{K})$ l'ensemble des matrices à n lignes et m colonnes dont les termes appartiennent à \mathbb{K} .*

Un coup d'oeil rapide montre que le **rang d'une matrice** A (défini par le rang de \mathbf{f} , l'application linéaire associée) est le nombre maximal de vecteurs colonnes linéairement indépendants. En cela, c'est aussi la taille maximale d'une sous-matrice extraite de A dont le déterminant est non nul.

Changement de base : Supposons que $\mathbf{f} : E \rightarrow E$, donc A est une matrice carrée, définie par rapport à une base \mathcal{B} . Supposons que l'on passe de la base \mathcal{B} à une base \mathcal{B}' , nous obtenons une nouvelle représentation matricielle A_2 telle que

$$A_2 = P^{-1}AP$$

où P est la matrice des coordonnées des vecteurs de l'ancienne base \mathcal{B} dans la nouvelle base \mathcal{B}' . P est appelée **matrice de changement de base**, ou **matrice de passage** de la base \mathcal{B} à la base \mathcal{B}' .

Dans le cas de \mathbb{R}^3 , on note $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ et $\mathcal{B}' = (\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3)$ deux bases de $E = \mathbb{R}^3$.

$$P = \begin{array}{ccc} & \mathbf{e}'_1 & \mathbf{e}'_2 & \mathbf{e}'_3 & & \\ & \downarrow & \downarrow & \downarrow & & \\ \left[\begin{array}{ccc} \clubsuit_1 & \heartsuit_1 & \diamondsuit_1 \\ \clubsuit_2 & \heartsuit_2 & \diamondsuit_2 \\ \clubsuit_3 & \heartsuit_3 & \diamondsuit_3 \end{array} \right] & \leftarrow & \mathbf{e}_1 \\ & \leftarrow & \mathbf{e}_2 \\ & \leftarrow & \mathbf{e}_3 \end{array}$$

Plus précisément, les colonnes de P sont l'expression des vecteurs de la base \mathcal{B}' dans la base \mathcal{B} . Dans cet exemple, le vecteur \mathbf{e}'_1 s'écrit $\mathbf{e}'_1 = \clubsuit_1 \mathbf{e}_1 + \clubsuit_2 \mathbf{e}_2 + \clubsuit_3 \mathbf{e}_3$. De même, le vecteur \mathbf{e}_1 s'écrit $\mathbf{e}_1 = \clubsuit_1 \mathbf{e}'_1 + \heartsuit_1 \mathbf{e}'_2 + \diamondsuit_1 \mathbf{e}'_3$.

3. Calcul matriciel

On se donne deux matrices $A, B \in \mathcal{M}_{n,n}(\mathbb{K})$, $A = (a_{i,j})_{1 \leq i,j \leq n}$ et $B = (b_{i,j})_{1 \leq i,j \leq n}$. Soient deux scalaires λ et $\mu \in \mathbb{K}$, alors on définit

- (1) la **trace** de A comme

$$\text{tr}(A) = \sum_{i=1}^n a_{i,i}$$

- (2) la combinaison linéaire de relations matricielles par

$$\lambda A + \mu B = (\lambda a_{i,j} + \mu b_{i,j})_{1 \leq i,j \leq n}$$

- (3) le **produit matriciel** AB par

$$AB = \left(\sum_{k=1}^n a_{i,k} b_{k,j} \right)_{1 \leq i,j \leq n}$$

qui correspond à la composée de deux applications linéaires \mathbf{f} et \mathbf{g}

- (4) la **transposition** A^\top par

$$A^\top = (a_{j,i})_{1 \leq i,j \leq n}$$

(5) la **conjuguée** \bar{A} par

$$\bar{A} = (\bar{a}_{i,j})_{1 \leq i,j \leq n}$$

(6) la **matrice hermitienne** A^* par

$$A^* = (\bar{a}_{j,i})_{1 \leq i,j \leq n} = \overline{A^\top} = \bar{A}^\top$$

Si $A_{i,j}$ est une sous-matrice créée à partir de A , en enlevant à celle-ci la colonne et la ligne contenant $a_{i,j}$, alors le **déterminant** de A est donné par

$$\det(A) = |A| = \sum_{j=1}^n (-1)^{i+j} a_{i,j} \det(A_{i,j})$$

où $A_{i,j}$ est appelé un **mineur** et $(-1)^{i+j} \det(A_{i,j})$ le **cofacteur** de A . On verra une autre formule dans le cours sur les tenseurs.

On dit qu'une matrice est

- (1) **symétrique** si $A = A^\top$
- (2) **orthogonale** si $A^\top A = I$ (I la matrice identité)
- (3) **unitaire** si $A^* A = I$
- (4) **nilpotente** si $\exists k_0$ tel que $A^k = 0, \forall k \geq k_0$
- (5) **définie positive** si $\langle Ax, x \rangle > 0, \forall x \notin E \setminus \{0_E\}$
- (6) **singulière** si $\det(A) = 0, A$ est donc non inversible
- (7) une **projection** (notée P) si $P^2 = P$.

En général, deux matrices ne commutent pas

$$AB \neq BA$$

On a

$$\operatorname{tr}(BA) = \operatorname{tr}(AB),$$

$$\det(AB) = \det(A) \det(B) = \det(BA),$$

$$\det(A^\top) = \det(A),$$

$$\det(A^{-1}) = (\det(A))^{-1} \text{ dès que } A^{-1} \text{ existe,}$$

$$\det(\lambda A) = \lambda^n \det(A), \forall \lambda \in \mathbb{K},$$

$$\det(A^*) = \det(A),$$

$$\text{si } A, B \in \mathcal{M}_{n,n}(\mathbb{K}), \text{ on a : } \det(I - AB) = \det(I - BA).$$

EXERCICE 12.1. Soit \mathcal{B} et \mathcal{B}' deux bases orthonormées de E , et P la matrice de changement de base. Montrer que $P^{-1} = P^\top$ et que $\det(P) = \pm 1$.

THÉORÈME 12.1. (**Théorème de décomposition des noyaux**) Soit $P(X)$ un polynôme à coefficients dans \mathbb{K} , tel que $P(X) = P_1(X)P_2(X)$ avec P_1 et P_2 polynômes premiers entre eux. Soit $A \in \mathcal{M}_{n,n}(\mathbb{K})$. On note $\operatorname{Ker} P(A) = \{\mathbf{x} \in E, P(A)\mathbf{x} = \mathbf{0}_E\}$. Alors tout élément $\mathbf{x} \in \operatorname{Ker} P(A)$ se décompose de manière unique en la somme d'un

élément (unique) de $\text{Ker } P_1(A)$ et d'un élément (unique) de $\text{Ker } P_2(A)$. Ce que l'on écrit

$$\text{Ker } P(A) = \text{Ker } P_1(A) \oplus \text{Ker } P_2(A)$$

4. Vecteurs propres

DÉFINITION 12.2. Soit $f : E \rightarrow E$ une application linéaire. Un scalaire λ est appelé **valeur propre** de f si $\exists \mathbf{u} \neq \mathbf{0}_E$ tel que

$$f(\mathbf{u}) = \lambda \mathbf{u}$$

Dans ce cas, \mathbf{u} est appelé **vecteur propre** de f . Le sous-espace associé à λ , noté E_λ , est appelé **sous-espace propre** et est défini par

$$E_\lambda = \{\mathbf{x} \in E / f(\mathbf{x}) = \lambda \mathbf{x}\}$$

Introduisons le même type de définition pour A une matrice carrée de dimension $n \times n$.

DÉFINITION 12.3. Un scalaire λ est appelé **valeur propre** de A si $\exists \mathbf{x} \neq \mathbf{0}, \mathbf{x} \in \mathbb{R}^n$ tel que

$$A\mathbf{x} = \lambda \mathbf{x}$$

Dans ce cas, \mathbf{x} est appelé **vecteur propre** de A . Le sous-espace associé à λ , noté E_λ , est appelé **sous-espace propre** et est défini par

$$E_\lambda = \{\mathbf{x} \in \mathbb{R}^n / A\mathbf{x} = \lambda \mathbf{x}\}$$

On introduit alors le **polynôme caractéristique** P_A d'une matrice A le polynôme défini par :

$$P_A(\lambda) = \det(A - \lambda I)$$

C'est un polynôme en λ , de degré n , dont le développement est donné par :

$$P_A(\lambda) = (-1)^n \lambda^n + (-1)^{n-1} \text{tr}(A) \lambda^{n-1} + \dots + \det(A)$$

L'équation caractéristique d'une matrice A est :

$$P_A(\lambda) = 0$$

Les valeurs propres sont alors données comme racines de ce polynôme.

DÉFINITION 12.4. Pour toute valeur propre λ d'une matrice A , on définit :

(a) sa **multiplicité algébrique**, qui est l'ordre de multiplicité de la racine λ dans le polynôme caractéristique de A ;

(b) sa **multiplicité géométrique**, qui est la dimension du sous-espace propre associé.

On a également le théorème de **Cayley-Hamilton** qui dit que la matrice A satisfait l'équation caractéristique

$$P_A(A) = 0.$$

5. Diagonalisation

Une matrice $D \in \mathcal{M}_{n,n}(\mathbb{K}) = (d_{i,j})_{i,j}$ est **diagonale** si $d_{i,j} = 0, \forall i \neq j$.

On dit qu'une matrice $A \in \mathcal{M}_{n,n}(\mathbb{K})$ est **diagonalisable** si elle possède n vecteurs propres linéairement indépendants. Autrement dit, il existe une base constituée de ses n vecteurs propres. Dit encore autrement, A est diagonalisable si il existe une matrice P inversible telle que

$$D = P^{-1}AP$$

où D est diagonale.

REMARQUE 12.2. Chaque vecteur colonne \mathbf{c}_k de la matrice P est un vecteur propre pour la matrice A correspondant à la valeur propre $\lambda_k = D_{kk}$. Alors, les valeurs propres de A sont les termes de la diagonale de D .

La **diagonalisation** est la détermination effective d'une matrice de passage P transformant une matrice diagonalisable en une matrice diagonale. La remarque précédente nous dit que la diagonalisation consiste à trouver des bases pour chaque espace propre et construire une matrice par la juxtaposition de tous ces vecteurs. Si la matrice est diagonalisable alors la matrice P obtenue est inversible et $P^{-1}AP$ est une matrice diagonale.

REMARQUE 12.3. Les puissances d'une matrice diagonalisable s'expriment sous la forme

$$A^k = PD^kP^{-1}$$

où la puissance de la diagonale se calcule en élevant simplement chaque coefficient diagonal à la même puissance k .

PROPOSITION 12.1. Soit $A \in \mathcal{M}_n(\mathbb{K})$. On a équivalence entre

- (a) La matrice A est diagonalisable ;
- (b) $\mathcal{M}_n(\mathbb{K}) = \bigoplus_{\lambda \in Sp(A)} E_\lambda(A)$;
- (c) $n = \sum_{\lambda \in Sp(A)} \dim(E_\lambda(A))$;
- (d) le polynôme caractéristique de A est scindé dans $\mathbb{K}[X]$ et $\forall \lambda \in Sp(A), \dim(E_\lambda(A)) = m_\lambda(A)$.

On peut montrer les propriétés suivantes

- toute matrice A symétrique ($A = A^\top$) et réelle est diagonalisable et ses valeurs propres sont réelles,
- toute matrice A symétrique, réelle et définie positive est diagonalisable et ses valeurs propres réelles sont strictement positives.

PROPOSITION 12.2. Soit A une matrice symétrique. Si $\lambda \neq \mu$ sont deux valeurs propres de A alors

$$E_\lambda \perp E_\mu.$$

DÉMONSTRATION. Soient $\mathbf{x} \in E_\lambda$ et $\mathbf{y} \in E_\mu$.

$$\lambda \langle \mathbf{x}, \mathbf{y} \rangle = \langle \lambda \mathbf{x}, \mathbf{y} \rangle = \langle A\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, A^\top \mathbf{y} \rangle = \langle \mathbf{x}, A\mathbf{y} \rangle = \langle \mathbf{x}, \mu \mathbf{y} \rangle = \mu \langle \mathbf{x}, \mathbf{y} \rangle$$

Alors on obtient $(\lambda - \mu)\langle \mathbf{x}, \mathbf{y} \rangle = 0$ ce qui implique $\langle \mathbf{x}, \mathbf{y} \rangle = 0$. Donc $\mathbf{x} \perp \mathbf{y}$, $\forall \mathbf{x} \in E_\lambda, \mathbf{y} \in E_\mu$. \square

THÉORÈME 12.2. Soit E un sous-espace fermé de \mathbb{R}^n . Chaque $\mathbf{x} \in \mathbb{R}^n$ admet une décomposition unique $\mathbf{x} = \mathbf{y} + \mathbf{z}$ avec $\mathbf{y} \in E$ et $\mathbf{z} \in E^\perp$. L'application linéaire

$$P_E : \mathbb{R}^n \mapsto \mathbb{R}^n, \quad P_E(\mathbf{x}) = \mathbf{y}$$

s'appelle la **projection** sur E . Les projections satisfont les propriétés suivantes :

- (1) $P_E^2 = P_E$;
- (2) si $E \perp F$ sont deux sous-espaces fermés de \mathbb{R}^n alors $P_E P_F = 0 = P_F P_E$
- (3) si $E \perp F$ sont deux sous-espaces fermés de \mathbb{R}^n alors $P_E + P_F$ est une projection sur $E \oplus F$

Les deux derniers résultats conduisent au **théorème de décomposition spectrale d'une matrice symétrique** :

Soient $A \in \mathbb{R}^{n \times n}$ une matrice symétrique et $\lambda_1, \lambda_2, \dots, \lambda_m$ ses valeurs propres distinctes ($\lambda_i \neq \lambda_j$ si $i \neq j$). On note P_i la projection sur l'espace propre associé à λ_i , $i = 1, 2, \dots, m$. Alors :

- (1) $P_i P_j = O, \forall i \neq j$;
- (2) $P_1 + P_2 + \dots + P_m = Id$;
- (3) $A = \lambda_1 P_1 + \lambda_2 P_2 + \dots + \lambda_m P_m$. Cette formule s'appelle la décomposition spectrale de A .

6. Réduction de Jordan

En algèbre linéaire, la forme de Jordan d'un opérateur linéaire sur un espace vectoriel de dimension finie est une matrice triangulaire supérieure d'une forme particulière appelée matrice de Jordan, représentant l'opérateur sur une certaine base. Précisément, pour une matrice $A \in \mathbb{C}^{n \times n}$ il existe une matrice P inversible telle que :

$$A = PJP^{-1}, \quad J = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_m \end{pmatrix} \text{ où } J_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}, \quad \forall i = 1, \dots, m$$

J est appelé la forme de Jordan associée à la matrice A et J_i sont appelés les blocs de Jordan de A . Un bloc de Jordan nilpotent est un bloc de Jordan avec les éléments diagonaux égaux à 0. On observe que la forme de Jordan est une décomposition de la matrice A comme la somme d'une matrice diagonale et une nilpotente.

PROPOSITION 12.3. On peut prouver les propriétés suivantes :

- (1) les valeurs propres de A (et J) sont les éléments diagonaux de J ;
- (2) le nombre de blocs de Jordan correspondant à λ_i coincide avec la multiplicité géométrique de λ_i ;
- (3) le bloc de Jordan correspondant à λ_i est de la forme $\lambda_i Id + N$ où N est un bloc de Jordan nilpotent.

La "jordanisation" est la détermination effective d'une matrice de passage P . En utilisant le [Théorème 12.1](#) on déduit que $Ker(A - \lambda_i Id)^{r_i} \perp Ker(A - \lambda_j Id)^{r_j}$, $\forall i \neq j$ où r_i est la multiplicité algébrique de λ_i , $i = 1, \dots, m$. Comme $r_1 + \dots + r_m = n$, $\mathbb{C}^n = \bigoplus_{i=1}^m Ker(A - \lambda_i Id)^{r_i}$ et, l'union des bases pour les sous-espaces $Ker(A - \lambda_i Id)^{r_i}$ est une base de \mathbb{C}^n . Cette base s'appelle une base de Jordan et les colonnes de P sont les vecteurs qui appartient à la base de Jordan.

EXEMPLE 12.1. Considérons l'exemple suivant

$$A = \begin{pmatrix} 5 & 4 & 2 & 1 \\ 0 & 1 & -1 & -1 \\ -1 & -1 & 3 & 0 \\ 1 & 1 & -1 & 2 \end{pmatrix}.$$

Les valeurs propres de A sont 4, 4, 2 et 1. De plus, on remarque que :

$$\dim Ker(A - Id) = 1, \dim Ker(A - 2Id) = 1, \dim Ker(A - 4Id) = 1, \dim Ker(A - 4Id)^2 = 2$$

Une base pour $Ker(A - Id)$ est $(-1, 1, 0, 0)^\top$, une base pour $Ker(A - 2Id)$ est $(1, -1, 0, 1)^\top$ et une base pour $Ker(A - 4Id)^2$ est $\{v, (A - 4Id)v\} = \{(0, 0, -1, 1)^\top, (-1, 0, 1, -1)^\top\}$ où $v = (0, 0, -1, 1)^\top$ est une base pour $Ker(A - 4Id)$. Nous en déduisons la matrice de passage et la forme de Jordan :

$$P = \begin{pmatrix} -1 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \\ 1 & -1 & 0 & 0 \\ -1 & 1 & 1 & 0 \end{pmatrix} \text{ et } J = P^{-1}BP = \begin{pmatrix} 4 & 1 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

7. Trigonalisation

Une matrice $T \in \mathcal{M}_{n,n}(\mathbb{K}) = (t_{i,j})_{i,j}$ est **triangulaire supérieure** si $t_{i,j} = 0, \forall i > j$.

On dit qu'une matrice $A \in \mathcal{M}_{n,n}(\mathbb{K})$ est **trigonalisable** si il existe une base \mathcal{B} dans laquelle elle est triangulaire supérieure. Dit autrement, A est trigonalisable si il existe une matrice P inversible telle que

$$T = P^{-1}AP$$

où T est triangulaire supérieure. Dans ce cas, $P_A(\lambda) = P_T(\lambda)$ et donc A et T ont même valeurs propres. Ces valeurs propres sont les termes de la diagonale de T .

THÉORÈME 12.3. *Soit $A \in \mathcal{M}_{n,n}(\mathbb{K})$. A est trigonalisable si et seulement si son polynôme caractéristique est scindé dans \mathbb{K} .*

En particulier, toute matrice $A \in \mathcal{M}_{n,n}(\mathbb{C})$ est trigonalisable¹.

8. Résolution de systèmes linéaires

Il est classique de se poser le problème de la résolution d'un système linéaire

$$A\mathbf{X} = \mathbf{b}$$

où $A \in \mathcal{M}_{n,n}(\mathbb{K})$ est donné, $\mathbf{b} \in \mathbb{R}^n$ également et où \mathbf{X} est un vecteur que l'on cherche à calculer.

PROPOSITION 12.4. *Les transformations suivantes changent tout système en un système équivalent :*

- (1) échanger deux lignes,
- (2) multiplier une ligne par un réel non nul,
- (3) ajouter une ligne à une autre ligne.

Nous dirons qu'un système est échelonné s'il se présente sous la forme suivante.

$$(S_E) \left\{ \begin{array}{l} p_1 y_1 + c_{1,2} y_2 + \dots + c_{1,r} y_r + \dots + c_{1,n} y_n = d_1 \\ p_2 y_2 + \dots + c_{2,r} y_r + \dots + c_{2,n} y_n = d_2 \\ \vdots \\ p_r y_r + \dots + c_{r,n} y_n = d_r \\ 0 = d_{r+1} \\ \vdots \\ 0 = d_m \end{array} \right.$$

Mettre un système (S) sous forme échelonnée, c'est passer de (S) à (S_E) par les transformations de la Proposition 12.4, et une permutation éventuelle des coordonnées, de sorte que

- (1) les inconnues (y_1, \dots, y_n) de (S_E) sont celles de (S) , mais dans un ordre qui peut être différent,
- (2) les coefficients p_1, \dots, p_r sont tous non nuls.

9. Exercices

EXERCICE 12.2. Diagonaliser $A = \begin{bmatrix} -3 & -1 & 6 \\ 6 & 4 & -12 \\ 1 & 1 & -2 \end{bmatrix}$. Calculer A^n pour un entier n quelconque.

1. car \mathbb{C} est algébriquement clos, i.e. tout polynôme à coefficients dans \mathbb{C} et de degrés n a n racines dans \mathbb{C} , distinctes ou non.

EXERCICE 12.3. Diagonaliser $A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$. Calculer A^n pour un entier n quelconque.

EXERCICE 12.4. Diagonaliser $A = \begin{bmatrix} 2 & 1 & 2 \\ 1 & 2 & 2 \\ 2 & 2 & 1 \end{bmatrix}$. Calculer A^n pour un entier n quelconque. A est-elle inversible ? Si oui calculer l'inverse.

EXERCICE 12.5. Soit $A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 4 & 1 & -2 \\ 2 & 1 & 2 & -1 \\ 1 & 2 & 1 & 0 \end{bmatrix} \in \mathcal{M}_4(\mathbb{R})$. Calculer le polynôme caractéristique de A . Déterminer les sous-espaces propres. Pourquoi A est-il non diagonalisable ? Est-il triangularisable ? Calculer A^5 .

Les équations différentielles linéaires

Dans ce qui suit, E est un espace vectoriel (sur \mathbb{R} ou \mathbb{C}) de dimension finie n , I un intervalle de \mathbb{R} et t_0 un élément de I . On considère également :

- (1) \mathcal{A} une fonction définie et continue sur I à valeurs dans l'espace $\mathcal{L}(E)$ des applications linéaires de E dans lui même ($\mathcal{A} \in C(I, \mathcal{L}(E))$).
- (2) \mathcal{B} une fonction définie et continue sur I à valeurs dans E ($\mathcal{B} \in C(I, E)$).

Si E est muni d'une base naturelle $\{e_1, \dots, e_n\}$, alors pour tout $t \in I$, on peut associer à $\mathcal{A}(t)$ et $\mathcal{B}(t)$ la matrice $A(t)$ et le vecteur $B(t)$ dans cette base. On notera donc par $A \in C(I, \mathcal{M}_n)$ et $B \in C(I, \mathbb{R}^n)$ les fonctions du temps associées respectivement à \mathcal{A} et à \mathcal{B} (\mathcal{M}_n désigne l'espace des matrices carrées de dimension $n \times n$). On pourra donc remplacer à l'envie \mathcal{A} par A et \mathcal{B} par B .

On cherche à résoudre le problème : Trouver $x \in C^1(I, E)$ tel que

$$\begin{cases} \dot{x}(t) &= \mathcal{A}(t)x(t) + \mathcal{B}(t) \\ x(t_0) &= x_0 \end{cases} \quad (0.1)$$

THÉORÈME 13.1. *Il existe une solution unique sur I à (0.1).*

DÉMONSTRATION. Montrons que pour tout compact $K \subset I$, il existe une solution unique sur K .

Par prolongement on en déduira qu'il existe une unique solution sur I .

Soit $K \subset I$ un compact. Posons $f(t, x) = \mathcal{A}(t)x(t) + \mathcal{B}(t)$

(a) f vérifie une condition de lipchitz sur $K \times E$

$\forall (t, x_2), (t, x_1) \in K \times E$,

$$\begin{aligned} \|f(t, x_2) - f(t, x_1)\| &= \|\mathcal{A}(t)x_2 - \mathcal{A}(t)x_1\| \\ \|f(t, x_2) - f(t, x_1)\| &= \|\mathcal{A}(t)(x_2 - x_1)\| \\ \|f(t, x_2) - f(t, x_1)\| &\leq M\|x_2 - x_1\| \end{aligned}$$

avec $M = \sup_{t \in K} \|\mathcal{A}(t)\|$

(b) f est continue sur $K \times E$,

$$\|f(t_2, x_2) - f(t_1, x_1)\| \leq \|\mathcal{A}(t_2)x_2 - \mathcal{A}(t_1)x_1\| + \|\mathcal{B}(t_2) - \mathcal{B}(t_1)\|$$

or

$$\begin{aligned}
\|\mathcal{A}(t_2)x_2 - \mathcal{A}(t_1)x_1\| &\leq \|\mathcal{A}(t_2)x_2 - \mathcal{A}(t_2)x_1\| + \|\mathcal{A}(t_2)x_1 - \mathcal{A}(t_1)x_1\| \\
\|\mathcal{A}(t_2)x_2 - \mathcal{A}(t_1)x_1\| &\leq \|\mathcal{A}(t_2)(x_2 - x_1)\| + \|(\mathcal{A}(t_2) - \mathcal{A}(t_1))x_1\| \\
\|\mathcal{A}(t_2)x_2 - \mathcal{A}(t_1)x_1\| &\leq M\|x_2 - x_1\| + \|\mathcal{A}(t_2) - \mathcal{A}(t_1)\|\|x_1\| \\
&\text{avec } M = \sup_{t \in K} \|\mathcal{A}(t)\|
\end{aligned}$$

d'où l'on peut conclure :

$$\forall \varepsilon > 0, \exists \delta, \forall (t_2, x_2), (t_1, x_1) \in K \times E, \|(t_2, x_2) - (t_1, x_1)\| < \delta \Rightarrow \|f(t_2, x_2) - f(t_1, x_1)\| < \varepsilon$$

Comme f est continue et vérifie une condition de lipschitz sur $K \times E$, le théorème des approximations successives permet de conclure

Comme ce raisonnement est valable quel que soit le compact K , on peut prolonger la solution sur I en entier. □

PROPOSITION 13.1. *Toute solution de (0.1) est la somme d'une solution de l'équation homogène $\dot{x}(t) = \mathcal{A}(t)x(t)$ et d'une solution particulière de $\dot{x}(t) = \mathcal{A}(t)x(t) + \mathcal{B}(t)$.*

DÉMONSTRATION. Si x_1 et x_2 sont solutions de $\dot{x}(t) = \mathcal{A}(t)x(t) + \mathcal{B}(t)$ alors $y(t) = x_1(t) - x_2(t)$ vérifie l'équation

$$\begin{aligned}
\dot{y}(t) &= \dot{x}_1(t) - \dot{x}_2(t) = \mathcal{A}(t)x_1(t) + \mathcal{B}(t) - (\mathcal{A}(t)x_2(t) + \mathcal{B}(t)) \\
&= \mathcal{A}(t)(x_1(t) - x_2(t)) = \mathcal{A}(t)y(t)
\end{aligned}$$

Donc $x_1(t) = y(t) + x_2(t)$ où y est solution de l'équation homogène et x_2 une solution particulière. □

1. Systèmes différentiels homogènes

DÉFINITION 13.1. *Le système (0.1) est dit homogène lorsque $\mathcal{B} \equiv 0$*

Notons par S l'espace vectoriel engendré par les solutions de l'équation

$$\dot{x}(t) = \mathcal{A}(t)x(t) \text{ sur } I.$$

Vérifions qu'il s'agit bien d'un sous espace vectoriel de $C(I, E)$

- (1) $0 \in S$
- (2) $\forall x_1, x_2 \in S, x_1 + x_2 \in S$
- (3) $\forall \lambda \in \mathbb{R}, \forall x, \lambda x \in S$

THÉORÈME 13.2. *L'application de $\varphi : E \rightarrow S$ qui à toute condition initiale x_0 associe la solution $x \in S$ est linéaire, continue et bijective.*

DÉMONSTRATION. Linéarité : Vérifier que l'on a bien :

$$\forall \lambda, \mu \in \mathbb{R}, \forall x_1, x_2 \in E, \varphi(\lambda x_1 + \mu x_2) = \lambda \varphi(x_1) + \mu \varphi(x_2)$$

Injective : Il s'agit de montrer que $\text{Ker} \varphi = \{0\}$

Comme $\varphi(x_0)(t) = x_0 + \int_{t_0}^t \mathcal{A}(s)\varphi(x_0)(s)ds \Rightarrow x_0 = 0$ si $\varphi(x_0) \equiv 0$

Surjective : Si $x \in S$ alors $x \in C(I, E)$ et $x_0 = x(t) - \int_{t_0}^t \mathcal{A}(s)x(s)ds$

Continuité : on a (cf théorème 3.3) $\|x(t) - y(t)\| \leq \|x_0 - y_0\| \exp\left(\int_{t_0}^t \|\mathcal{A}(s)\| ds\right)$ \square

Ce théorème implique la proposition :

PROPOSITION 13.2. (1) *L'espace vectoriel des solutions S est de même dimension que E .*

(2) *Si on note $x_i = \varphi(e_i)$, $i = 1, \dots, n$, alors toute solution x se décompose sous la forme $x = \sum_{i=1}^n \lambda_i x_i$ si $x_0 = \sum_{i=1}^n \lambda_i e_i$.*

Soit $(x_1 \cdots x_n)$ n solutions de l'équation homogène $\dot{x}(t) = \mathcal{A}(t)x(t)$.

THÉORÈME 13.3. *Si la famille $\{x_1, \dots, x_n\}$ est linéairement indépendante à l'instant t_0 , alors elle le reste au cours du temps.*

DÉMONSTRATION. Supposons qu'il existe un instant t_1 tel que la famille $\{x_1(t_1) \cdots x_n(t_1)\}$ ne soit pas linéairement indépendante. Alors il existe des λ_i non tous nuls tels que $\sum_{i=1}^n \lambda_i x_i(t_1) = 0$. Comme l'unique solution de $\dot{x}(t) = \mathcal{A}(t)x(t)$ avec conditions initiales nulles à l'instant t_1 , est la fonction identiquement nulle sur I , on en déduit $\sum_{i=1}^n \lambda_i x_i(t_0) = 0$ d'où la contradiction. \square

2. Résolvante d'un Système différentiel

DÉFINITION 13.2. *On appelle opérateur résolvant du système (0.1), l'application $R : I \times I \rightarrow \mathcal{L}(E)$ définie par $\forall x_0 \in E, R(t, t_0)x_0 = x(t)$ où x est la solution de l'équation homogène :*

$$\begin{cases} \dot{x}(t) &= \mathcal{A}(t)x(t) \\ x(t_0) &= x_0 \end{cases} \quad (2.1)$$

PROPOSITION 13.3. *L'opérateur résolvant vérifie les propriétés suivantes :*

(1) $\forall t, t_0 \in I, R(t, t_0)$ est une application linéaire bijective de E

(2) $\forall t_0 \in I, R(t_0, t_0) = Id$

(3) $\forall t, s, t_0 \in I, R(t, s)R(s, t_0) = R(t, t_0)$

Par définition, on a $\forall x_0 \in E, \forall t, t_0 \in I$:

$$\begin{aligned} R(t, t_0)x_0 &= x_0 + \int_{t_0}^t \mathcal{A}(s)R(s, t_0)x_0 ds \\ &= \left(Id + \int_{t_0}^t \mathcal{A}(s)R(s, t_0) ds \right) x_0 \end{aligned}$$

d'où

$$R(t, t_0) = Id + \int_{t_0}^t \mathcal{A}(s)R(s, t_0) ds$$

PROPOSITION 13.4. $R(., t_0)$ est donc l'unique application continue de I dans $\mathcal{L}(E)$ qui vérifie :

$$\begin{cases} R'(t, t_0) &= \mathcal{A}(t)R(t, t_0) \\ R(t_0, t_0) &= Id \end{cases} \quad (2.2)$$

THÉORÈME 13.4. La solution de l'équation (0.1) est donnée par la formule :

$$x(t) = R(t, t_0)x_0 + \int_{t_0}^t R(t, s)\mathcal{B}(s)ds$$

Le terme $\int_{t_0}^t R(t, s)\mathcal{B}(s)ds$ correspond à la solution de (0.1) avec la condition initiale $x_0 = 0$. La solution du système avec second membre est donc la somme de la solution sans second membre prenant la valeur x_0 en t_0 et de la solution de (0.1) avec $x_0 = 0$.

DÉMONSTRATION. On pose $x(t) = R(t, t_0)y(t)$ et $x(t_0) = x_0$ ($\Rightarrow y(t_0) = x_0$) puis on dérive :

$$\begin{aligned} \dot{x}(t) &= R'(t, t_0)y(t) + R(t, t_0)y'(t) \\ &= A(t)R(t, t_0)y(t) + R(t, t_0)y'(t) \\ &= A(t)x(t) + R(t, t_0)y'(t) \end{aligned}$$

On identifie $R(t, t_0)y'(t)$ avec $B(t)$ de sorte que x est la solution de :

$$\begin{cases} \dot{x}(t) &= A(t)x(t) + B(t) \\ x(t_0) &= x_0 \end{cases}$$

Comme $R(t, t_0)$ est inversible, d'inverse $R(t_0, t)$, il vient :

$$\begin{cases} y'(t) &= R(t_0, t)B(t) \\ y(t_0) &= x_0 \end{cases} \quad (2.3)$$

Par intégration, on obtient :

$$y(t) = x_0 + \int_{t_0}^t R(t_0, s)\mathcal{B}(s)ds$$

et puisque $x(t) = R(t, t_0)y(t)$:

$$x(t) = R(t, t_0)x_0 + \int_{t_0}^t R(t, t_0)R(t_0, s)\mathcal{B}(s)ds$$

d'où le résultat annoncé. □

Il nous faut à présent déterminer l'opérateur résolvant.

3. Matrice de Wronski

On se place dans une base de E et on considère n solutions (x_1, \dots, x_n) de l'équation homogène $\dot{x}(t) = A(t)x(t)$.

DÉFINITION 13.3. On appelle matrice de Wronski la matrice $W(t) = [x_1 \cdots x_n]$.

THÉORÈME 13.5. Si les n solutions (x_1, \dots, x_n) sont linéairement indépendantes alors

$$R(t, t_0) = W(t)W(t_0)^{-1}$$

4. Le cas des équations différentielles linéaires à coefficients constants

On cherche donc à résoudre dans cette partie le cas où $\forall t, A(t) \equiv A$.

4.1. Préliminaire : Exponentielle de Matrice. Soit \mathcal{A} un endomorphisme de E , A sa matrice dans une base.

Montrons que la série $(\sum_{n=0}^{\infty} \frac{A^n}{n!})$ est normalement sommable pour la norme $\|A\| = \sup_{i,j} |a_{ij}|$.

Comme $\|AB\| = \sup_{i,j} \left| \sum_{k=1}^n a_{ik}b_{kj} \right| \leq \sup_{i,j} \sum_{k=1}^n |a_{ik}| |b_{kj}| \leq n \sup_{i,j} |a_{ij}| \sup_{i,j} |b_{kj}| = n\|A\|\|B\|$,

il vient $\|A^k\| \leq n^{k-1}\|A\|^k$.

Par suite $\sum_{k=0}^{\infty} \frac{\|A^k\|}{k!} \leq \frac{1}{n} \sum_{k=0}^{\infty} \frac{(n\|A\|)^k}{k!} = \frac{1}{n} e^{n\|A\|} < +\infty$

Puisque E est complet ($\dim E = n$ fini), on en déduit que la série $(\sum_{n=0}^{\infty} \frac{A^n}{n!})$ converge.

On peut alors définir pour toute matrice A :

DÉFINITION 13.4. On appelle exponentielle de A et on note $\exp(A) = e^A$ la somme de la série convergente $\sum_{n=0}^{\infty} \frac{A^n}{n!}$.

Propriétés : Soit M et N deux matrices qui **commutent** entre elles. Alors $\exp(M + N) = \exp(M)\exp(N)$. En particulier, $\exp(-M) = (\exp(M))^{-1}$.

4.2. Calcul pratique de $\exp(A)$.

- (1) Cas où A est diagonalisable. Si $A = P\Delta P^{-1}$ avec Δ diagonale alors $\forall n, A^n = P\Delta^n P^{-1}$. On en déduit que $\sum_{n=0}^{\infty} \frac{A^n}{n!} = P \sum_{n=0}^{\infty} \frac{\Delta^n}{n!} P^{-1}$ et par passage à la limite $\exp(A) = P\exp(\Delta)P^{-1}$ avec

$$\exp(\Delta) = \begin{bmatrix} \exp(\delta_1) & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \exp(\delta_i) & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & \exp(\delta_n) \end{bmatrix}$$

où les $\delta_i, i = 1, \dots, n$ sont les valeurs propres de A .

- (2) Cas où A est nilpotente. Si $A^k = 0$, le terme général de la série est nul à partir du rang k et alors $\exp(A) = Id + A + \frac{A^2}{2} + \cdots + \frac{A^{k-1}}{(k-1)!}$
- (3) On sait que toute matrice carré A se décompose sous la forme $M + N$ où M est diagonalisable et N nilpotente et où $MN = NM$. Il en résulte que $\exp(A) = \exp(M)\exp(N)$.

4.3. Rappel sur la décomposition sous forme de Jordan. Mise sous forme d'une matrice de Jordan (**Cas des matrices non diagonalisables**)

Soit A la matrice associée à un endomorphisme de E .

Le théorème spectral permet de décomposer E comme la somme directe des sous espaces spectraux, soit $E = \bigoplus_{i=1}^p E_i$ où $E_i = \text{Ker}(A - \delta_i Id)^{r_i}$ avec $r_i \leq$ à la multiplicité de δ_i . (δ_i valeur propre de A).

On a les propriétés suivantes :

- L'image de chacun des E_i par A est stable ($A(E_i) \subset E_i$).
- Par définition des E_i , il existe r_i tel que $\Phi_i^{(r_i)} = 0$ où $\Phi_i = A|_{E_i} - \delta_i Id_{E_i} : E_i \rightarrow E_i$.

Ainsi $\text{Ker}\Phi_i^{(k)}, k = 0, 1, 2, \dots$ forment une suite croissante et on a :

$$\{0\} = \text{Ker}\Phi_i^{(0)} \subset \text{Ker}\Phi_i^{(1)} \subset \cdots \subset \text{Ker}\Phi_i^{(r_i)} = E_i$$

La suite est strictement croissante. (c'est facile à montrer : si $\text{Ker}\Phi_i^{(p)} = \text{Ker}\Phi_i^{(p+1)} \Rightarrow \text{Ker}\Phi_i^{(p+1)} = \text{Ker}\Phi_i^{(p+2)}$)

On construit une base des E_i de la manière suivante :

- (1) On cherche tous les v tels que $v \in Ker\Phi_i$
- (2) Puis pour chacun des vecteurs v trouvés, on cherche les w tels que $v = \Phi_i(w)$.

Et on voit alors que : $w \in Ker\Phi_i^{(2)}$ et $w \notin Ker\Phi_i$

- (3) On itère le procédé jusqu'à épuisement.

Au final dans la base de E_i ainsi construite, $A|_{E_i}$ prend la forme de Jordan

$$\begin{bmatrix} \delta_i & * & 0 & 0 \\ 0 & \delta_i & \ddots & 0 \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & \delta_i \end{bmatrix} \text{ avec } * = 0 \text{ ou } 1$$

Intérêt de la décomposition : On a directement la décomposition de A sous la forme $M + N$ recherchée au point (3) de la section précédente.

4.4. Exemple : Soit la matrice $A = \begin{bmatrix} 4 & 2 & 0 \\ 1 & 5 & -1 \\ -1 & 1 & 5 \end{bmatrix}$. Calculer $\exp(At)$.

Le polynôme caractéristique de A est $P(\lambda) = (\lambda - 4)^2(\lambda - 6)$

On résout $(A - 6Id)x = 0$ et on trouve $v_6 = [1 \ 1 \ 0]^T$. Le sous espace spectral $E_1 = Ker(A - 6Id)$ est donc engendré par v_6 .

Puis on résout $(A - 4Id)x = 0$, on trouve un seul vecteur $v_4 = [1 \ 0 \ 1]^T$ d'où l'on déduit que $E_2 = Ker(A - 4Id)^2$. On cherche alors w_4 tel que $v_4 = (A - 4Id)w_4$ et on trouve $w_4 = [-0.5 \ 0.5 \ 0]^T$

Ainsi $\mathbb{R}^3 = E_1 \oplus E_2$ avec $E_1 = Ker(A - 6Id)$ et $E_2 = Ker(A - 4Id)^2$ et dans la base $\{v_6, v_4, w_4\}$, A s'écrit :

$$\begin{bmatrix} 6 & 0 & 0 \\ 0 & 4 & 1 \\ 0 & 0 & 4 \end{bmatrix} = P^{-1}AP$$

avec $P = [v_6 \ v_4 \ w_4]$.

On peut maintenant calculer l'exponentielle de tA .

On a :

$$\exp\left(\begin{bmatrix} 6 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} t\right) = \begin{bmatrix} e^{6t} & 0 & 0 \\ 0 & e^{4t} & 0 \\ 0 & 0 & e^{4t} \end{bmatrix}$$

$$\exp\left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} t\right) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{bmatrix}$$

d'où

$$\exp\left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} t\right) \exp\left(\begin{bmatrix} 6 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} t\right) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} e^{6t} & 0 & 0 \\ 0 & e^{4t} & 0 \\ 0 & 0 & e^{4t} \end{bmatrix}$$

$$= \begin{bmatrix} e^{6t} & 0 & 0 \\ 0 & e^{4t} & te^{4t} \\ 0 & 0 & e^{4t} \end{bmatrix}$$

Au final,

$$\exp(At) = P^{-1} \begin{bmatrix} e^{6t} & 0 & 0 \\ 0 & e^{4t} & te^{4t} \\ 0 & 0 & e^{4t} \end{bmatrix} P$$

$$= \begin{bmatrix} \frac{1}{2}e^{4t} - te^{4t} + \frac{1}{2}e^{6t} & te^{4t} - \frac{1}{2}e^{4t} + \frac{1}{2}e^{6t} & te^{4t} + \frac{1}{2}e^{4t} - \frac{1}{2}e^{6t} \\ -\frac{1}{2}e^{4t} + \frac{1}{2}e^{6t} & \frac{1}{2}e^{6t} + \frac{1}{2}e^{4t} & \frac{1}{2}e^{4t} - \frac{1}{2}e^{6t} \\ -te^{4t} & te^{4t} & te^{4t} + e^{4t} \end{bmatrix}$$

4.5. Résolvante. On peut maintenant énoncer le théorème :

THÉORÈME 13.6. *Lorsque $\forall t \in I, A(t) \equiv A$, la solution de l'équation (0.1) est donnée par la formule :*

$$x(t) = \exp(A(t - t_0))x_0 + \int_{t_0}^t \exp(A(t - s))\mathcal{B}(s)ds$$

REMARQUE 13.1. On observera que le terme intégral $\int_{t_0}^t \exp(A(t - s))\mathcal{B}(s)ds$ n'est autre que le produit de convolution : $(f * g)(t) = \int_{t_0}^t f(t - s)g(s)ds$ avec $f(t) = \exp(At)$ et $g(t) = \mathcal{B}(t)$.

DÉMONSTRATION. Il suffit de constater que :

$$\begin{cases} \frac{d}{dt} \exp(A(t - t_0)) &= A(t) \exp(A(t - t_0)) \\ \exp(A(t_0 - t_0)) &= Id \end{cases} \quad (4.1)$$

Posons $S_k : \mathbb{R} \rightarrow \mathcal{M}_n$ la suite de fonctions définies par $t \mapsto S_k(t) = \sum_{i=0}^k \frac{A^i t^i}{i!}$.

$\forall k \forall t$, S_k est continûment dérivable de dérivée égale à : $S'_k(t) = A S_{k-1}(t)$. Sur tout compact K de \mathbb{R} , S'_k converge uniformément vers $A \exp(At)$ ce qui entraîne $(\exp(At))' = A \exp(At)$ sur tout compact de \mathbb{R} donc sur \mathbb{R} tout entier. \square

4.6. Point d'équilibre et stabilité. Commençons par quelques définitions :

DÉFINITION 13.5. On appelle état (ou point) d'équilibre un point x_e de \mathbb{R}^n tel que

$$f(t, x_e) \equiv 0, \forall t \geq t_e.$$

Autrement dit si la solution x passe par un tel point elle y reste par la suite.

DÉFINITION 13.6. Un point d'équilibre x_e est stable depuis t_0 si

$$\forall \varepsilon > 0, \exists \delta(\varepsilon, t_0) > 0, \|x_0 - x_e\| < \delta \Rightarrow \|x(t) - x_e\| < \varepsilon \quad \forall t > t_0$$

Il est dit instable dans le cas contraire. Autrement dit si la solution x part suffisamment près du point d'équilibre, elle y reste d'aussi près que l'on veut.

DÉFINITION 13.7. Un point d'équilibre x_e est localement asymptotiquement stable depuis t_0 si

$$\exists \delta(t_0) > 0, \|x_0 - x_e\| < \delta \Rightarrow \lim_{\substack{t \rightarrow +\infty \\ t \geq t_0}} \|x(t) - x_e\| = 0$$

DÉFINITION 13.8. Un point d'équilibre x_e est globalement asymptotiquement stable depuis t_0 si

$$\forall x_0, \lim_{\substack{t \rightarrow +\infty \\ t \geq t_0}} \|x(t) - x_e\| = 0$$

Ces définitions sur la stabilité ont pour but de répondre aux questions suivantes :

Que ce passe-t-il lorsque l'on écarte l'état x du système étudié d'une position d'équilibre x_e ?

- l'état x reste-t-il dans un voisinage x_e ? (stabilité)
- l'état x revient-il naturellement vers l'équilibre x_e ? (stabilité asymptotique)
- ou bien s'écarte-t-il irrémédiablement de l'équilibre x_e

EXEMPLE 13.1. Tout le monde connaît l'exemple du pendule qui possède deux points d'équilibres, l'un stable l'autre pas.

PROPOSITION 13.5. Un système linéaire homogène $\dot{x}(t) = A(t)x(t)$ ne possède qu'un seul point d'équilibre **à l'origine** si $A(t)$ n'est pas identiquement singulière.

Deux propriétés remarquables pour les systèmes linéaires homogènes à coefficients constants sont :

PROPOSITION 13.6. Le système $\dot{x}(t) = Ax(t)$ est stable si aucune des valeurs propres de la matrice A n'est strictement supérieure à 0.

PROPOSITION 13.7. *Le système $\dot{x}(t) = Ax(t)$ est globalement asymptotiquement stable i.e. :*

$$\forall x_0, \quad x(t) \rightarrow 0 \text{ lorsque } t \rightarrow +\infty$$

si et seulement si toutes les valeurs propres de la matrice A sont à parties réelles strictement négatives.

PROPOSITION 13.8. *Pour que le système $\dot{x}(t) = Ax(t)$ soit instable, il faut et il suffit qu'il existe une valeur propre de la matrice A strictement supérieure à 0.*

4.7. Linéarisation et stabilité d'un point d'équilibre. A contrario des systèmes linéaires, les systèmes non linéaires ne possèdent pas en générale de propriétés globales, c'est à dire vrai sur tout l'espace de travail. L'étude est alors menée de manière locale, dans un voisinage i.e. autour d'un point de cet espace.

Une technique très utile est la linéarisation. Elle consiste à faire le développement limité à l'ordre 1 autour d'un point d'un système non linéaire et de négliger les termes d'ordres supérieurs. Le résultat de cette opération est un *système linéaire* permettant d'approcher le comportement du système non linéaire dans un voisinage du point où le développement est effectué. L'étude du système linéarisé permet alors dans certains cas de tirer des conclusions sur le système non linéaire **mais attention, ces conclusions ne sont valables que localement.**

Considérons un système non linéaire

$$\dot{x} = f(t, x)$$

le système linéarisé autour du point (t_0, x_0) est alors défini comme

$$\begin{aligned} \dot{x} &= f(t_0, x_0) + \left. \frac{\partial f}{\partial x} \right|_{(t_0, x_0)} (x - x_0) + \left. \frac{\partial f}{\partial t} \right|_{(t_0, x_0)} (t - t_0) \\ \dot{x} &= f(t_0, x_0) + A(x - x_0) + B(t - t_0) \end{aligned}$$

avec

$$A = \left. \frac{\partial f}{\partial x} \right|_{(t_0, x_0)} \quad \text{et} \quad B = \left. \frac{\partial f}{\partial t} \right|_{(t_0, x_0)}$$

REMARQUE 13.2. On conservera à l'esprit que le modèle linéarisé change avec le point (t_0, x_0) où s'effectue cette linéarisation.

REMARQUE 13.3. Le système linéarisé est également appelé système linéarisé tangent pour bien marquer l'approximation faite et le caractère locale de cette opération.

Linéarisation et stabilité d'un point d'équilibre :

On s'intéresse à la stabilité d'un point d'équilibre x_e d'un système non linéaire autonome (sans dépendance explicite au temps),

$$\dot{x} = f(x)$$

Puisqu'il s'agit d'un point d'équilibre, $f(x_e) = 0$ et le système linéarisé autour de ce point x_e s'écrit $\dot{x} = A(x - x_e)$.

PROPOSITION 13.9. *Si A a toutes ces valeurs propres à partie réelle **strictement** négative alors le système non linéaire est localement asymptotiquement stable en x_e .*

*Si A a toutes ces valeurs propres à partie réelle négative mais **non strictement** alors on ne peut pas se prononcer sur la stabilité local en x_e du système non linéaire.*

*Si A a **au moins** une valeur propre à partie réelle **strictement** positive alors le système non linéaire est instable en x_e .*

EXEMPLE 13.2. En automatique, on est amené très souvent à considérer des équations différentielles non linéaires qui s'écrivent sous la forme :

$$\dot{x} = f(x, u)$$

où la variable u est une fonction du temps et représente une entrée (commande) sur le système. On agit sur l'état x du système par le biais de cette commande u .

Il est alors fréquent que l'on cherche à amener le système dans un état particulier x_e , et de chercher à l'y maintenir. On détermine alors la commande u_e telle que $f(x_e, u_e) = 0$.

Il est alors important de savoir si ce point d'équilibre (x_e, u_e) est stable ou non. On effectue une linéarisation autour de ce point :

$$\dot{x} = \left. \frac{\partial f}{\partial x} \right|_{(x_e, u_e)} (x - x_e) + \left. \frac{\partial f}{\partial u} \right|_{(x_e, u_e)} (u - u_e)$$

d'où le système linéaire

$$\dot{x} = A(x - x_e) + B(u - u_e)$$

Si on écarte x de sa position d'équilibre x_e , on peut se prononcer sur la stabilité de ce point en calculant les valeurs propres la matrice A et en utilisant la proposition ci-dessus.

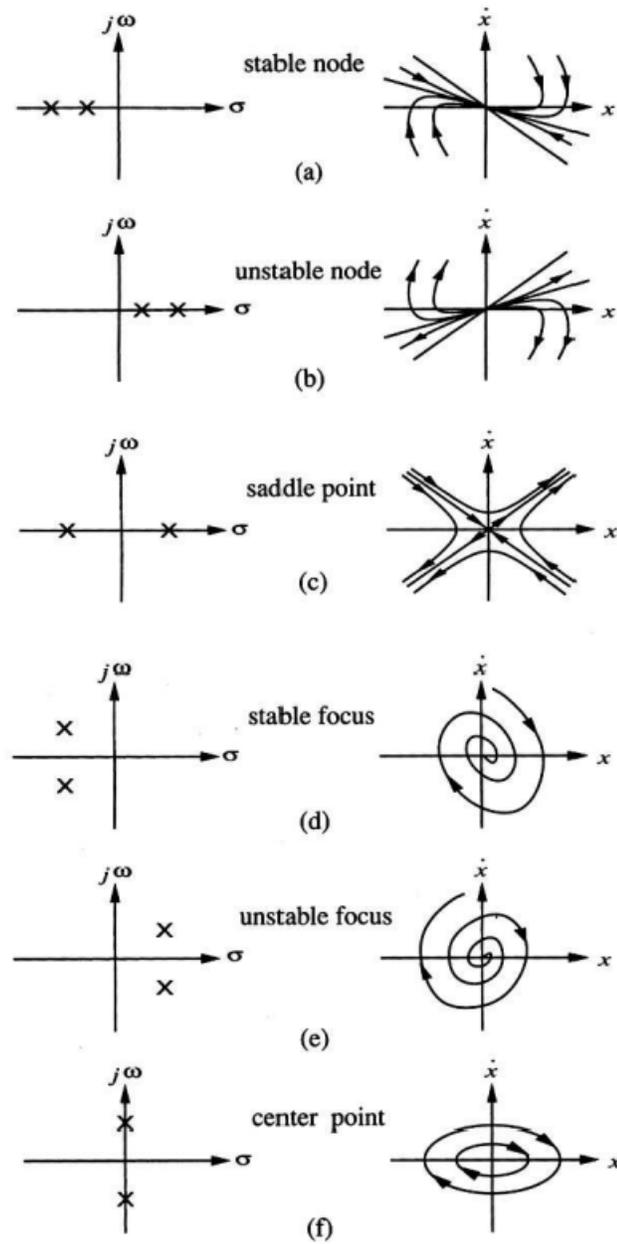


Figure 2.9 : Phase-portraits of linear systems

FIGURE 1 – Trajectoire dans le plan de phase d'un système linéaire homogène d'ordre 2 à coefficients constants en fonction de ses valeurs propres

Espaces L^1, L^2 et convergences

Ce très court chapitre donne quelques rappels sur la théorie de l'intégration. Pour être plus clair, ce chapitre nécessiterait un cours de théorie de la mesure de Lebesgue... Nous nous contentons de donner quelques définitions et résultats. Dans tout ce chapitre, I est un intervalle de \mathbb{R} (ce peut être \mathbb{R} tout entier)

1. La notion de 'presque-partout'

Presque partout signifie partout sauf en un nombre fini ou dénombrable de points. En abrégé, on écrit *p.p.*.

Par exemple la fonction $\mathbf{1}_{[0,1]}$ est continue *p.p.*.

2. L'espace $L^1(I)$

Soit l'ensemble $\mathcal{L}^1(I) = \{f : I \rightarrow \mathbb{C}, \text{ telles que } \int_I |f(t)| dt < \infty.\}$

On définit, pour $f \in \mathcal{L}^1(I)$:

$$\|f\|_1 = \int_I |f(t)| dt.$$

PROPOSITION 14.1. *Soit $f \in \mathcal{L}^1(I)$. Si $\|f\|_1 = 0$, nous n'avons pas $f = 0$, mais $f = 0$ p.p.*

De même, si $\|f - g\|_1 = 0$, alors $f = g$ p.p.

Donc $\|\cdot\|_1$ n'est pas une norme sur $\mathcal{L}^1(I)$.

On définit alors un autre espace, sur lequel $\|\cdot\|_1$ est une norme. Pour cela, on décide que si $f = g$ p.p., alors, elles représentent le même objet. Ainsi, lorsque l'on parlera d'une fonction, on parlera de la classe de cette fonction modulo une égalité presque partout (p.p.).

DÉFINITION 14.1. *On appelle espace $L^1(I)$ l'espace vectoriel des fonctions définies sur I , à valeurs dans \mathbb{C} et de module intégrable (définies à une égalité p.p.) muni de la norme*

$$\|f\|_1 = \int_I |f(x)| dx$$

On admet que

$$\|f\|_1 = 0 \Leftrightarrow f = 0 \text{ p.p.} \Leftrightarrow f = 0_{L^1(I)}$$

$\|\cdot\|_1$ est donc une norme pour $L^1(I)$, i.e. $L^1(I)$ est donc un espace vectoriel **normé**, de norme $\|\cdot\|_1$.

3. L'espace $L^2(I)$

De même, on définit $\mathcal{L}^2(I) = \{f : I \rightarrow \mathbb{C}, \text{ telles que } \int_I |f(t)|^2 dt < \infty.\}$ et pour $f \in \mathcal{L}^2(I)$:

$$\|f\|_2 = \sqrt{\int_I |f(t)|^2 dt} = \sqrt{\langle f, f \rangle}.$$

Pour les mêmes raisons que précédemment, on définit un nouvel espace, pour lequel $\|\cdot\|_2$ est une norme.

DÉFINITION 14.2. *On appelle espace $L^2(I)$ l'espace vectoriel des fonctions définies sur I , à valeurs dans \mathbb{C} et de carré intégrable (définies à une égalité p.p.).*

On peut munir $L^2(I)$ du produit scalaire

$$\langle f, g \rangle = \int_I f(t)g(\bar{t})dt$$

La norme $\|\cdot\|_2$ est la norme associée à ce produit scalaire.

On admet que

$$\|f\|_2 = 0 \Leftrightarrow f = 0 \text{ p.p.} \Leftrightarrow f = 0_{L^2(I)}$$

On a l'**inégalité de Schwartz** : si $f, g \in L^2(I)$, alors

$$|\langle f, g \rangle| \leq \|f\|_2 \|g\|_2$$

i.e.

$$\int_I f(t)g(\bar{t})dt \leq \sqrt{\int_I |f(t)|^2 dt} \sqrt{\int_I |g(t)|^2 dt}.$$

De l'inégalité de Schwartz, on déduit que, **si I est borné** (par exemple $I = [a, b]$ avec a, b des réels), **alors $L^2(I) \subset L^1(I)$** et

$$\|f\|_1 \leq \sqrt{b-a} \|f\|_2 \tag{3.1}$$

DÉMONSTRATION. Soit $f \in L^2(I)$. On sait que la fonction constante égale à 1 est aussi dans $L^2(I)$ car I est borné.

Donc, $\langle |f|, 1 \rangle = \int_I |f(t)| dt = \|f\|_1 \leq \sqrt{\int_I |f(t)|^2 dt} \sqrt{\int_I |1|^2 dt} = \sqrt{b-a} \|f\|_2 < \infty$. Attention, on n'a pas de telle relation entre $L^2(\mathbb{R})$ et $L^1(\mathbb{R})$. D'ailleurs la preuve utilisée n'est plus applicable si $I = \mathbb{R}!!!$ \square

4. Convergence de de suites fonctions et propriétés

Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions, et f une fonction. Toutes ces fonctions étant définies sur I et à valeurs dans \mathbb{C} . On a différents types de convergence de f_n vers f .

On parle de convergence

- **ponctuelle** ou **simple** si $f_n(t) \xrightarrow{n \rightarrow \infty} f(t)$ pour chaque t .
- **presque partout** si $f_n(t) \xrightarrow{n \rightarrow \infty} f(t)$ pour presque tout t .
- **uniforme sur** I si $\sup_{t \in I} |f_n(t) - f(t)| \xrightarrow{n \rightarrow \infty} 0$.

On parle aussi de la convergence dans $L^1(I)$ ou $L^2(I)$ lorsque $\|f_n - f\|_1 \xrightarrow{n \rightarrow \infty} 0$ ou $\|f_n - f\|_2 \xrightarrow{n \rightarrow \infty} 0$.

Pour $L^2(I)$, on dit aussi **convergence en moyenne quadratique** ou **au sens de l'énergie**.

On a les propriétés suivantes :

- Convergence uniforme \Rightarrow Convergence ponctuelle \Rightarrow Convergence presque partout.
- Si une suite f_n converge uniformément vers f sur I :
 - si les f_n sont continues, alors f est continue.
 - $\int_I f_n$ tend vers $\int_I f$.
- Si une suite f_n converge ponctuellement vers f , si la suite des f'_n converge uniformément vers g alors $g = f'$.
- Pour $I = [a, b]$ **borné**, on a d'après (3.1), Convergence uniforme \Rightarrow Convergence dans $L^2(I) \Rightarrow$ Convergence dans $L^1(I)$.

EXERCICE 14.1. Soit la suite de fonctions $f_n(t) = \frac{1}{n^p} \sqrt{n^2 - t^2} \mathbf{1}_{[-n, n]}(t)$. Si $p = 2$, montrer que f_n converge uniformément et au sens de l'énergie sur \mathbb{R} , vers 0. Si $p = 1$, montrer que f_n converge ponctuellement vers 1 sur \mathbb{R} , mais ne converge pas uniformément ni au sens de l'énergie sur \mathbb{R} .

EXERCICE 14.2. Soit la suite de signaux $f_n(t) = te^{-nt^2}$. Examiner les convergences ponctuelle et uniforme sur \mathbb{R} . A-t-on $\lim_{n \rightarrow \infty} f'_n(t) = \frac{d}{dt} [\lim_{n \rightarrow \infty} f_n(t)]$?

La convolution

1. Introduction

La convolution est un opérateur qui, à partir de deux fonctions, en crée une troisième. Le produit de convolution est d'une grande importance en physique (pour modéliser la transmission d'un signal par un appareil, l'impulsion électrique en fonction du temps, en diffraction, en acoustique linéaire, etc.)

2. Définition

Considérons deux fonctions f et g intégrables sur \mathbb{R} . On appelle **produit de convolution** de deux fonctions f et g et on le note $f * g$, l'intégrale définie par

$$(f * g)(x) = \int_{\mathbb{R}} f(t)g(x - t)dt$$

lorsque cette intégrale existe.

On peut aussi définir le produit de convolution de deux fonctions discrètes

$$(f * g)(n) = \sum_{m=-\infty}^{+\infty} f(m)g(n - m)$$

lorsque cette somme existe.

Dans la suite, pour ne pas alourdir les notations, on ne travaillera plus qu'avec le produit de convolution dans le cas continu. Mais les résultats s'étendent au cas discret.

REMARQUE 15.1. On peut montrer que si f et g sont intégrables, alors $f * g$ est intégrable. Plus particulièrement :

$$\|f * g\|_1 \leq \|f\|_1 \|g\|_1$$

En effet, par Fubini, $\int_{\mathbb{R}} |(f * g)(x)|dx \leq \int_{\mathbb{R}} \int_{\mathbb{R}} |f(t)g(x - t)|dtdx = \int_{\mathbb{R}} |f(t)|(\int_{\mathbb{R}} |g(x - t)|dx)dt = \int_{\mathbb{R}} |f(t)|\|g\|_1 dt = \|f\|_1 \|g\|_1$ où $\|f\|_1$ et $\|g\|_1$ sont les intégrales de $|f|$ et $|g|$ respectivement. \square

Pour toutes fonctions f_1, f_2, g_1, g_2, f et g intégrables, et tous scalaires λ_1 et λ_2 dans \mathbb{R} , nous avons (preuve à faire en exercice)

$$\begin{aligned} (\lambda_1 f_1 + \lambda_2 f_2) * g &= \lambda_1 (f_1 * g) + \lambda_2 (f_2 * g), \\ f * (\lambda_1 g_1 + \lambda_2 g_2) &= \lambda_1 (f * g_1) + \lambda_2 (f * g_2), \\ f * g &= g * f \\ &\text{p.p.} \end{aligned}$$

Par ailleurs, sous des hypothèses d'existence des intégrales correspondantes, nous avons les relations

$$\text{Si } f, g, h \in L^1(\mathbb{R}), \quad (f * g) * h \underset{p.p.}{=} f * (g * h) \quad (2.1)$$

$$\text{Si } f, g \in L^1(\mathbb{R}) \cap \mathcal{C}_b^1(\mathbb{R}), \quad \frac{d}{dx}(f * g) = \frac{df}{dx} * g = f * \frac{dg}{dx} \quad (2.2)$$

où on a noté $\mathcal{C}_b^1(\mathbb{R})$ l'ensemble des fonctions de classe \mathcal{C}^1 sur \mathbb{R} , à dérivées bornées sur \mathbb{R} .

3. Quelques exemples

Dans le cas où les fonctions f et g sont **causales**, i.e. nulles sur \mathbb{R}^{-*} (par exemple si on considère que f et g sont des variables du temps $t \geq 0$), on a $f(t) = 0$ si $t \leq 0$ et $g(x - t) = 0$ si $x \leq t$. La convolution se réduit alors à une intégrale à support fini

$$(f * g)(x) = \int_0^x f(t)g(x - t)dt$$

Considérons l'**échelon unité** ou encore ce que l'on appelle la **fonction de Heaviside**. Celle-ci est définie par :

$$\begin{cases} H(x) = 0, & x < 0 \\ H(x) = 1, & x \geq 0 \end{cases}$$

Par conséquent, $H(x - t) = 0$ si $x < t$ et 1 si $t \leq x$. Nous obtenons ainsi, pour $f \in L^1(\mathbb{R})$

$$(f * H)(x) = \int_{-\infty}^x f(t)dt$$

Considérons maintenant ce que l'on appelle la **porte de Dirac**. Celle-ci est définie par

$$\Pi_\varepsilon(x) = \begin{cases} \frac{1}{\varepsilon} & \text{si } |x| \leq \frac{\varepsilon}{2} \\ 0 & \text{si } |x| > \frac{\varepsilon}{2} \end{cases}$$

En fait, puisque $\Pi_\varepsilon(x) = 0$, si $|x| > \frac{\varepsilon}{2}$, et 1 sinon, nous avons

$$f_\varepsilon(x) = (f * \Pi_\varepsilon)(x) = \frac{1}{\varepsilon} \int_{-\frac{\varepsilon}{2}}^{\frac{\varepsilon}{2}} f(x - t)dt$$

qui se récrit, par changement de variable $t \leftarrow x - t$, sous la forme

$$(f * \Pi_\varepsilon)(x) = \frac{1}{\varepsilon} \int_{x - \frac{\varepsilon}{2}}^{x + \frac{\varepsilon}{2}} f(t)dt$$

Lorsque ε tend vers zéro, nous avons formellement ce que l'on appelle l'**impulsion** ou le **pic de Dirac** δ_0 qui vaut 0 si $x \neq 0$ et $+\infty$ (par convention) sinon. Le calcul que nous

allons effectuer maintenant est très formel¹ et demanderait plus de développements mathématiques à travers notamment l'introduction de la notion de distribution (vue en seconde année). Ainsi, formellement, nous avons

$$(f * \delta_0)(x) = \lim_{\varepsilon \rightarrow 0} f_\varepsilon(x) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \int_{x-\frac{\varepsilon}{2}}^{x+\frac{\varepsilon}{2}} f(t) dt = \lim_{\varepsilon \rightarrow 0} f(c)$$

avec $c \in [x - \frac{\varepsilon}{2}; x + \frac{\varepsilon}{2}]$. Par conséquent, nous en déduisons que

$$(f * \delta_0)(x) = f(x),$$

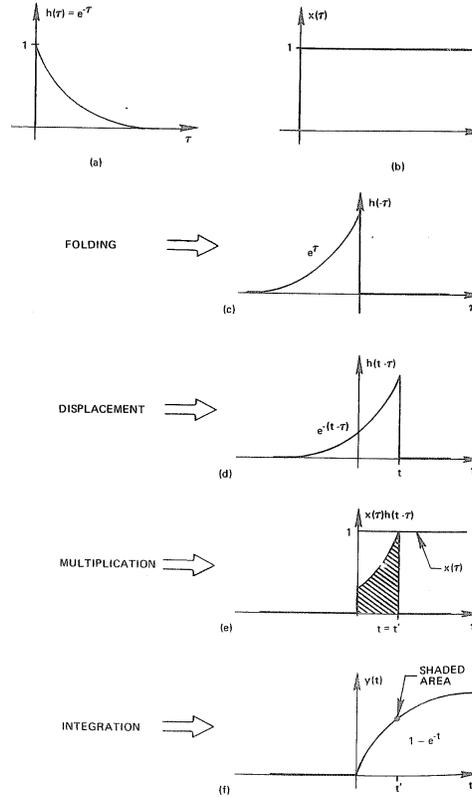
si f est continue. Ceci indique que δ_0 est l'**élément neutre** pour la convolution.

4. Interprétation graphique de la convolution

Considérons deux fonctions causales x et h . Pour évaluer $(x * h)(t)$, il faut :

1. Prendre le symétrique de $h(\tau)$ par rapport à l'axe des ordonnées
2. Translater la fonction $h(-\tau)$ de t
3. Multiplier la fonction translatée par $x(\tau)$
4. Intégrer la produit $x(\tau)h(t - \tau)$ sur \mathbb{R} .

1. En fait, δ_0 est une distribution (que vous verrez en 2ème année), pas une fonction, il est défini par sa transformée de Fourier $\widehat{\delta_0}(\xi) = 1 = \int_{\mathbb{R}} \delta_0(t) e^{-2i\pi ft} dt$, et $\delta_0(t) = 0$ si $t \neq 0$. Il s'agit d'un pic infiniment fin, dont l'intégrale vaut 1.



5. Exercices

EXERCICE 15.1. *Le problème du calcul de la charge u d'un condensateur à travers une résistance est régi par l'équation différentielle*

$$\tau \frac{du}{dt} + u = E, \quad u(0) = u_0, \quad (5.1)$$

où E est l'entrée du système. On appelle également u la sortie ou encore la réponse.

- (1) Calculer la solution de l'équation différentielle dans le cas où E est constant.
- (2) Dans le cas général, montrer que la solution s'écrit sous la forme d'une convolution.
- (3) Montrer que le régime transitoire devient rapidement négligeable. Quel est le régime permanent? Dépend-t-il de la donnée initiale?
- (4) Retrouver le résultat de la première question à partir de la formule générale.

EXERCICE 15.2. *On définit l'opérateur de dérivation fractionnaire $\partial_t^{1/2} f$ d'ordre un demi pour une fonction f par*

$$\partial_t^{1/2} f(t) = \partial_t(G * f)(t)$$

où G est le noyau défini par $G(t) = \frac{1}{\sqrt{\pi}\sqrt{t}}$. On considère la fonction $f(t) = t$. Calculer sa dérivée fractionnaire. Dériver fractionnairement le résultat obtenu. Que constatez vous ?

Compléments sur les espaces de Hilbert. Séries de Fourier

Les séries de Fourier permettent de décomposer une fonction périodique en une série de fonctions trigonométriques élémentaires. Elles permettent de résoudre de nombreux problèmes, comme l'équation des cordes vibrantes (cf. exercice 16.5), l'équation de la chaleur (cf. exercice 16.9), etc.

Avant de montrer comment une telle décomposition est possible, rappelons quelques points fondamentaux.

1. Compléments sur les espaces de Hilbert

1.1. Rappels - Définitions.

1.1.1. *Espace préhilbertien.* On rappelle ici qu'un **espace pré-hilbertien** H est un espace vectoriel muni d'un produit scalaire $\langle \cdot, \cdot \rangle$. On définit la norme sur H , $\|\mathbf{x}\|_H = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$.

1.1.2. *Convergence dans un espace préhilbertien.* On dira que la suite \mathbf{x}_n d'éléments de H converge vers $\mathbf{x} \in H$ si $\|\mathbf{x}_n - \mathbf{x}\|_H \xrightarrow{n \rightarrow \infty} 0$.

1.1.3. *Espace de Hilbert.* Un **espace de Hilbert** est un espace pré-hilbertien **complet**, c'est-à-dire tel que toute suite de Cauchy dans H est convergente. Les espaces considérés ici sont de dimension infinie. De ce fait, l'analyse est plus compliquée qu'en dimension finie. Néanmoins, les espaces de Hilbert généraux permettent d'avoir encore accès à des propriétés importantes que nous avons en dimension finie. Il est, par exemple, possible de parler de **base** comme nous allons le voir un peu plus loin.

1.1.4. *Familles...* Comme dans le cas des espaces vectoriels de dimension finie, nous avons les définitions suivantes :

DÉFINITION 16.1. Soit $(\mathbf{e}_i)_{i \in I}$ une famille de vecteurs de H indexée par un ensemble I , dénombrable ou non. La famille $(\mathbf{e}_i)_{i \in I}$ est dite

- **orthogonale** si $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = 0$, $\forall i \neq j$,
- **orthonormée** si $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{sinon} \end{cases}$,
- **totale** si $\{\mathbf{e}_i, i \in I\}^\perp = \{\mathbf{0}\}$, c'est-à-dire si étant donné $\mathbf{x} \in H$, $\langle \mathbf{x}, \mathbf{e}_i \rangle = 0, \forall i \in I \implies \mathbf{x} = \mathbf{0}$,
- **libre** si étant donnée une famille finie de complexes $(\lambda_i)_{i \in I}$, $\sum_{i \in I} \lambda_i \mathbf{e}_i = \mathbf{0} \implies \lambda_i = 0, \forall i \in I$.

— **génératrice de l'espace** H si pour tout élément $\mathbf{x} \in H$, il existe une famille de réels (ou complexes) $(\lambda_i)_{i \in I}$, telle que $\mathbf{x} = \sum_{i \in I} \lambda_i \mathbf{e}_i$.

On définit les bases d'un espace Hilbertien dans la section [sous-section 1.3](#). Avant cela, donnons quelques propriétés des familles orthonormées.

1.2. Quelques résultats sur les familles orthonormées dénombrables. Dans toute cette section, H est un espace pré-hilbertien.

Tout d'abord, une famille dénombrable orthonormée est libre (le montrer en exercice).

DÉFINITION 16.2. Soit $\mathbf{x} \in H$, un vecteur et $(\mathbf{e}_n)_{n \in \mathbb{N}}$ une famille orthonormée. On pose

$$\mathbf{S}_p(\mathbf{x}) = \sum_{n=0}^p \langle \mathbf{x}, \mathbf{e}_n \rangle \mathbf{e}_n, \quad \forall p \in \mathbb{N}$$

que l'on appelle **somme partielle de Fourier** de \mathbf{x} relativement à la famille orthonormée $(\mathbf{e}_n)_{n \in \mathbb{N}}$. Le coefficient

$$c_n(\mathbf{x}) = \langle \mathbf{x}, \mathbf{e}_n \rangle$$

est appelé **n -ième coefficient de Fourier** de \mathbf{x} . On a alors

$$\mathbf{S}_p(\mathbf{x}) = \sum_{n=0}^p c_n(\mathbf{x}) \mathbf{e}_n$$

On dit aussi que $\mathbf{S}_p(\mathbf{x})$ est la **projection orthogonale** de \mathbf{x} sur le sous-espace vectoriel $\text{Vect}\{\mathbf{e}_0, \dots, \mathbf{e}_p\}$ de dimension finie. Il est clair que $\mathbf{x} - \mathbf{S}_p(\mathbf{x}) \perp \text{Vect}\{\mathbf{e}_0, \dots, \mathbf{e}_p\}$ et que l'on a la formule de Pythagore

$$\|\mathbf{x}\|_H^2 = \|\mathbf{x} - \mathbf{S}_p(\mathbf{x})\|_H^2 + \|\mathbf{S}_p(\mathbf{x})\|_H^2$$

donc

$$\|\mathbf{S}_p(\mathbf{x})\|_H^2 = \sum_{n=0}^p |\langle \mathbf{x}, \mathbf{e}_n \rangle|^2 \leq \|\mathbf{x}\|_H^2$$

Cette dernière inégalité étant vraie pour tout indice p , nous avons le théorème suivant, appelé **inégalité de Bessel**.

THÉORÈME 16.1. Soit $(\mathbf{e}_n)_{n \in \mathbb{N}}$ une famille orthonormale dénombrable et $\mathbf{x} \in H$. Alors, la série $\sum_n |\langle \mathbf{x}, \mathbf{e}_n \rangle|^2$ converge et

$$\sum_{n \in \mathbb{N}} |\langle \mathbf{x}, \mathbf{e}_n \rangle|^2 \leq \|\mathbf{x}\|_H^2$$

THÉORÈME 16.2. Pour tout vecteur $\mathbf{x} \in H$, la suite des projections orthogonales $(\mathbf{S}_n(\mathbf{x}))_{n \in \mathbb{N}}$ converge dans H .

DÉMONSTRATION. On a $\lim_{n \rightarrow \infty} \|S_{n+p} - S_n\|_H = 0$, car c'est le reste d'une série convergente. S_n est donc une suite de Cauchy de H , qui est complet. S_n converge donc vers un élément de H . \square

On voit ici que notre suite pourrait aussi bien converger vers \mathbf{x} ou un autre vecteur de H . La question est de savoir si $\mathbf{S}_p(\mathbf{x}) \rightarrow_{p \rightarrow \infty} \mathbf{x}$? En général, la réponse est négative. Le [Théorème 16.3](#) indiquera dans quel cas on a $\mathbf{S}_p(\mathbf{x}) \rightarrow_{p \rightarrow \infty} \mathbf{x}$.

1.3. Bases hilbertiennes. La question que l'on se pose maintenant est de savoir si on peut étendre la notion de base d'un espace vectoriel de dimension finie à un espace de Hilbert.

La réponse est oui :

DÉFINITION 16.3. Soit H un espace de Hilbert. On dit que la famille dénombrable $(\mathbf{e}_n)_{n \in \mathbb{Z}}$ est une **base hilbertienne** si c'est une famille libre et génératrice. Une **base orthonormée hilbertienne** est une base hilbertienne de vecteurs orthonormés.

Nous aurons aussi besoin de la définition suivante.

DÉFINITION 16.4. Soit H un espace de Hilbert. Si H possède une base hilbertienne, on dit que H est **séparable**.

Ce sera dorénavant notre hypothèse de travail.

La propriété d'avoir accès à une base hilbertienne est très forte car elle permet de décomposer tout vecteur de H en une combinaison (infinie) d'éléments de la base, et cette décomposition est unique. C'est ce que l'on précise maintenant.

THÉORÈME 16.3. Soit H un espace de Hilbert et $(\mathbf{e}_n)_{n \in \mathbb{N}}$ une famille orthonormée (donc en particulier libre). On a alors équivalence entre les différents points suivants

- i) $(\mathbf{e}_n)_{n \in \mathbb{N}}$ est génératrice de H
- ii) $\forall \mathbf{x} \in H$, si l'on pose

$$\mathbf{S}_p(\mathbf{x}) = \sum_{n=0}^p \langle \mathbf{x}, \mathbf{e}_n \rangle \mathbf{e}_n$$

alors

$$\|\mathbf{S}_p(\mathbf{x}) - \mathbf{x}\|_H \xrightarrow{p \rightarrow \infty} 0.$$

- iii) $\forall (\mathbf{x}, \mathbf{y}) \in H \times H$, on a

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=0}^{+\infty} \langle \mathbf{x}, \mathbf{e}_i \rangle \langle \mathbf{e}_i, \mathbf{y} \rangle$$

- iv) $\forall \mathbf{x} \in H$, on a l'égalité de Parseval

$$\|\mathbf{x}\|_H^2 = \sum_{n=0}^{\infty} |\langle \mathbf{x}, \mathbf{e}_n \rangle|^2$$

- v) la famille $(\mathbf{e}_n)_{n \in \mathbb{N}}$ est une famille totale.

Donc, dans un espace de Hilbert H , si $(\mathbf{e}_n)_{n \in \mathbb{N}}$ est une **base orthonormée hilbertienne** de H , les projections orthogonales $(\mathbf{S}_n(\mathbf{x}))_{n \in \mathbb{N}}$ de tout $\mathbf{x} \in H$ convergent vers \mathbf{x} dans H . On peut donc décomposer \mathbf{x} , par

$$\mathbf{x} = \sum_{n \in \mathbb{N}} \langle \mathbf{x}, \mathbf{e}_n \rangle \mathbf{e}_n = \sum_{n \in \mathbb{N}} c_n(\mathbf{x}) \mathbf{e}_n \quad (1.1)$$

De plus,

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{n \in \mathbb{N}} c_n(\mathbf{x}) \overline{c_n(\mathbf{y})}$$

et on a la formule de Parseval :

$$\|\mathbf{x}\|_H^2 = \sum_{n \in \mathbb{N}} |c_n(\mathbf{x})|^2$$

Le corollaire suivant dit que la décomposition (1.1) est **unique**.

COROLLAIRE 16.1. *Deux éléments d'un espace de Hilbert ayant mêmes coefficients de Fourier sont égaux.*

DÉMONSTRATION. Soit $(\mathbf{e}_n)_{n \in \mathbb{N}}$ une base hilbertienne de H . Soient $\mathbf{x}, \mathbf{y} \in H$ tels que $c_n(\mathbf{x}) = c_n(\mathbf{y})$, pour tout $n \in \mathbb{N}$.

Alors $\|\mathbf{x} - \mathbf{y}\|_H \leq \|\mathbf{x} - \mathbf{x}_n\|_H + \|\mathbf{x}_n - \mathbf{y}_n\|_H + \|\mathbf{y}_n - \mathbf{y}\|_H = \|\mathbf{x} - \mathbf{x}_n\|_H + \|\mathbf{y} - \mathbf{y}_n\|_H \rightarrow 0$ quand n tend vers $+\infty$. On a donc $\mathbf{x} = \mathbf{y}$. □

Nous avons également la proposition ci-dessous.

PROPOSITION 16.1. *Soit H un espace de Hilbert séparable. Alors toute famille orthonormée est dénombrable.*

REMARQUE 16.1. *Que se passe-t-il si la famille $(\mathbf{e}_n)_{n \in \mathbb{N}}$ est orthonormée, mais pas génératrice de H ? L'inégalité de Bessel est encore vraie, mais on n'aura pas que $\mathbf{x} = \sum_{n \in \mathbb{N}} \langle \mathbf{x}, \mathbf{e}_n \rangle \mathbf{e}_n$. La famille n'est pas une base.*

Deux espaces de Hilbert sont très importants dans la suite de notre étude : les espaces l^2 et L^2 . Ils sont très utilisés en traitement du signal, mécanique quantique, ... Nous verrons que nous pourrons passer de l'un à l'autre de ces espaces, c'est ce que l'on appelle l'analyse de Fourier.

2. L'espace l^2

DÉFINITION 16.5. *On appelle espace l^2 l'ensemble des suites $(a_n)_{n \in \mathbb{N}}$ à valeurs dans \mathbb{C} telles que*

$$\sum_{n \in \mathbb{N}} |a_n|^2 < +\infty$$

THÉORÈME 16.4. *L'espace ℓ^2 , muni du produit scalaire hermitien*

$$\langle a, b \rangle = \sum_{n \in \mathbb{N}} a_n \overline{b_n}$$

est un espace de Hilbert.

Une base hilbertienne orthonormée de ℓ^2 est donnée par :

$$\mathbf{e}_n = (\underbrace{0, \dots, 0}_{(n-1) \text{ fois}}, 1, 0, \dots),$$

$n \in \mathbb{N}$. ℓ^2 est donc un espace séparable.

3. L'espace $L^2(]0; a[)$

Ici, on pose le fondement de l'analyse de Fourier pour les fonctions périodiques de carré intégrable.

REMARQUE 16.2. Nous étudions les fonctions périodiques sur une période, c'est pourquoi nous nous contentons d'une étude sur un intervalle borné du type $]0, a[$. On travaillera donc avec l'espace $L^2(]0, a[)$.

THÉORÈME 16.5. *L'espace $L^2(]0, a[)$ est un espace de Hilbert. La famille $(e_n)_{n \in \mathbb{Z}}$ définie par*

$$e_n(x) = \frac{1}{\sqrt{a}} e^{2in\pi x/a}, \quad \forall n \in \mathbb{Z}, \quad \forall x \in]0; a[$$

est une base orthonormée hilbertienne de $L^2(]0, a[)$.

DÉMONSTRATION. Le fait que la famille orthonormée $(e_n)_{n \in \mathbb{Z}}$ soit une base hilbertienne résulte du fait que l'ensemble des fonctions continues sur $[0, a]$ est dense dans $L^2(]0, a[)$, et du théorème de Stone-Weierstrass (nous admettons tout cela). \square

EXERCICE 16.1. *Vérifier que la famille $(e_n)_{n \in \mathbb{Z}}$ est bien orthonormée.*

Une fonction de $L^2(]0; a[)$ peut donc être représentée par une somme infinie de fonctions trigonométriques

$$f = \sum_{n \in \mathbb{Z}} c_n(f) e_n$$

On a égalité dans $L^2(]0, a[)$, i.e. presque partout, et $c_n(f) = \langle f, e_n \rangle$. On dit qu'on a **convergence quadratique** (i.e. en norme L^2) de la série partielle $\sum_{n=-p}^p c_n(f) e_n$ vers la fonction de départ f .

De plus, l'égalité de Parseval s'écrit

$$\sum_n |c_n(f)|^2 = \|f\|_2^2 < +\infty.$$

L'égalité de Parseval dit que l'énergie d'un signal périodique est égale à la somme des énergies associées aux différents harmoniques.

Ce résultat est le fondement de l'analyse en séries de Fourier des fonctions périodiques.

4. L'espace $L^2(\mathbb{R})$

Un théorème fondamental dit que $L^2(\mathbb{R})$ est complet. C'est donc un espace de Hilbert. On peut démontrer qu'il admet une base hilbertienne.

5. Approximation par des séries de Fourier

Une manière commode de considérer $L^2(]0; a[)$ est de dire que c'est l'espace des fonctions $f : \mathbb{R} \rightarrow \mathbb{C}$ qui sont a -périodiques et de carré intégrable sur $]0; a[$.

Nous donnons maintenant la théorie des séries de Fourier, qui n'est autre que la traduction des résultats précédents sur $L^2(]0, a[)$.

5.1. Coefficients de Fourier d'une fonction.

DÉFINITION 16.6. Soit $f \in L^2(]0; a[)$. Pour tout $n \in \mathbb{Z}$, on appelle n -ième **coefficient de Fourier** le nombre

$$c_n = \langle f, e_n \rangle = \frac{1}{\sqrt{a}} \int_0^a f(t) e^{-2i\pi nt/a} dt.$$

On appelle suite des **sommes partielles de Fourier** la suite (f_n) définie par

$$f_n(t) = \frac{1}{\sqrt{a}} \sum_{k=-n}^n c_k e^{2i\pi kt/a} = \sum_{k=-n}^n c_k e_k(t)$$

REMARQUE 16.3. — La somme partielle f_n n'est autre que la **projection orthogonale** de f sur le sous-espace $\text{Vect}(e_i)_{i=-n}^n$. En tant que projection orthogonale de f , f_n est la '**meilleure approximation trigonométrique**' de f à partir de fonctions trigonométriques de périodes a/k , $k = 1, \dots, n$.

— D'après les sections précédentes, on sait que la **décomposition d'une fonction** $f \in L^2(]0, a[)$ est **unique**. Les caractérisations de f par ses valeurs temporelles $f(t)$, ou par ses coefficients de Fourier $c_n(f)$, $n \in \mathbb{Z}$ sont donc équivalentes.

PROPOSITION 16.2. On considère deux fonctions $f, g \in L^2(]0, a[)$ et $c_n(f), c_n(g)$ leurs coefficients de Fourier respectifs. On a les propriétés suivantes :

	Signal temporel	Coefficients de Fourier
<i>Linéarité</i>	$\alpha f(t) + \beta g(t)$	$\alpha c_n(f) + \beta c_n(g)$
<i>Décalage temporel</i>	$f(t - t_0)$	$e^{-2i\pi n t_0/a} c_n(f)$
<i>Symétrie temporelle</i>	$f(-t)$	$c_{-n}(f)$
<i>Conjuguée</i>	$\overline{f(t)}$	$\overline{c_{-n}(f)}$
<i>Fonction réelle paire</i>	$f(t) = f(-t) \in \mathbb{R}$	$c_n(f) = c_{-n}(f) \in \mathbb{R}$
<i>Fonction imaginaire paire</i>	$f(t) = f(-t) \in i\mathbb{R}$	$c_n(f) = c_{-n}(f) \in i\mathbb{R}$
<i>Fonction réelle impaire</i>	$f(t) = -f(-t) \in \mathbb{R}$	$c_n(f) = -c_{-n}(f) \in i\mathbb{R}$

Fonction imaginaire impaire $f(t) = -f(-t) \in i\mathbb{R}$ $c_n(f) = -c_{-n}(f) \in \mathbb{R}$

De plus, si f est de classe \mathcal{C}^1 par morceaux sur $]0, a[$, alors la série de Fourier de f' s'obtient en dérivant terme à terme celle de f et elle converge quadratiquement vers f' . De plus,

$$c_n(f') = \frac{2i\pi n}{a} c_n(f), \forall n \in \mathbb{Z}.$$

DÉMONSTRATION. En exercice. □

5.2. Convergence quadratique. Soit une fonction $f \in L^2(]0; a[)$. Puisque $L^2(]0; a[)$ est un espace de Hilbert et $(e_n)_{n \in \mathbb{N}}$, une base hilbertienne orthonormée, on peut utiliser toutes les propriétés des bases hilbertiennes.

On a **convergence quadratique de la série partielle de Fourier f_n vers f** :

$$\|f - f_n\|_{L^2} \rightarrow_{n \rightarrow +\infty} 0$$

ce que l'on écrit

$$f(t) = \frac{1}{\sqrt{a}} \sum_{n \in \mathbb{Z}} c_n e^{2i\pi n t/a} = \sum_{n \in \mathbb{Z}} c_n e_n(t) \text{ p.p.} \quad (5.1)$$

pour presque tout $t \in]0, a[$, et on appelle **développement en série de Fourier** cette écriture.

Donc, en plus d'être une projection orthogonale de f , f_n est aussi une bonne approximation, au sens de la norme $\|\cdot\|_2$, de f .

REMARQUE 16.4. Pour vous convaincre qu'il n'y a pas égalité partout, faites la première question de l'exercice 16.3. La série de Fourier de f n'est pas égale à f au point $T/2$. Vous remarquerez que ce point est une discontinuité de la fonction f ... Est-ce un hasard? Nous verrons cela plus tard...

La famille $(|c_n|^2)_{n \in \mathbb{Z}}$ est sommable et on a l'égalité de Parseval :

$$\sum_{n \in \mathbb{Z}} |c_n|^2 = \int_0^a |f(t)|^2 dt = \|f\|_2^2$$

De plus, si $f, g \in L^2(]0; a[)$, alors $f\bar{g} \in L^1(]0; a[)$, c'est-à-dire

$$\sum_{n \in \mathbb{Z}} c_n(f) \overline{c_n(g)} = \int_0^a f(t) \overline{g(t)} dt = \langle f, g \rangle$$

Ceci permet d'affirmer que la représentation en série de Fourier est une isométrie¹ entre les espaces $L^2(]0; a[)$ et ℓ^2 .

REMARQUES 16.1. On trouve également cette notation pour la série de Fourier :

$$\frac{1}{a} \sum_{n \in \mathbb{Z}} \int_{-a/2}^{a/2} f(x) e^{-2i\pi nx/a} dx e^{2i\pi nt/a} = \sum_{n \in \mathbb{Z}} C_n e^{2i\pi nt/a}$$

avec $C_n = \frac{1}{a} \int_{-a/2}^{a/2} f(x) e^{-2i\pi nx/a} dx$.

Vous choisirez la notation qui vous convient le mieux, mais faites très attention de garder la même du début à la fin de vos calculs, afin de ne pas faire d'erreurs!! Dans ce polycopié, la notation utilisée sera celle des c_n .

On utilise parfois les notations réelles suivantes (avec égalité abusive) : pour presque tout $t \in]0, a[$,

$$f(t) = \frac{a_0}{2\sqrt{a}} + \frac{1}{\sqrt{a}} \sum_{n=1}^{+\infty} a_n \cos\left(n \frac{2\pi t}{a}\right) + \frac{1}{\sqrt{a}} \sum_{n=1}^{+\infty} b_n \sin\left(n \frac{2\pi t}{a}\right)$$

avec

$$a_n = \frac{2}{\sqrt{a}} \int_0^a f(t) \cos\left(n \frac{2\pi t}{a}\right) dt = c_n + c_{-n}$$

$$b_n = \frac{2}{\sqrt{a}} \int_0^a f(t) \sin\left(n \frac{2\pi t}{a}\right) dt = i(c_n - c_{-n})$$

et

$$\int_0^a |f(t)|^2 dt = \frac{|a_0|^2}{4} + \frac{1}{2} \sum_{n=1}^{+\infty} (|a_n|^2 + |b_n|^2).$$

REMARQUE 16.5. Si f est paire, $b_n = 0$, pour tout $n \in \mathbb{N}^*$. Si f est impaire, $a_n = 0$, pour tout $n \in \mathbb{N}$.

1. C'est-à-dire une application qui préserve la norme

5.3. Série de Fourier d'une fonction L^1 $]0; a[$. Ici, on s'intéresse aux fonctions intégrables sur $]0, a[$.

Donnons deux remarques sur $L^1(]0, a[)$:

— $L^1(]0, a[)$ n'est pas un espace de Hilbert.

— Nous avons vu au [chapitre 14](#) que $L^2(]0, a[) \subset L^1(]0, a[)$. Donc $L^1(]0, a[)$ est un espace plus grand que $L^2(]0, a[)$.

Si $f \in L^1(]0; a[)$, il est possible de définir

$$c_n = \frac{1}{\sqrt{a}} \int_0^a f(t) e^{-2i\pi nt/a} dt$$

Ainsi, nous avons la proposition suivante.

PROPOSITION 16.3. *Toute fonction $f \in L^1(]0; a[)$ admet des coefficients de Fourier.*

Que peut-on dire alors des sommes partielles de Fourier $(f_n)_{n \in \mathbb{N}}$ de f ? Convergent-elles vers f ? Nous ne sommes plus dans un cadre hilbertien et il faut donc analyser les choses de manière détaillée.

5.4. Convergence ponctuelle de la série de Fourier. Nous avons vu que pour toute fonction f de $L^2(]0, a[)$, la suite f_n des sommes partielles de Fourier converge en norme quadratique vers f , et que l'on écrit abusivement $f(t) = \frac{1}{\sqrt{a}} \sum_{n \in \mathbb{Z}} c_n e^{2i\pi nt/a} = \sum_{n \in \mathbb{Z}} c_n e_n(t)$, l'égalité ayant lieu presque partout. On n'a donc pas la convergence simple des f_n vers f (et encore moins la convergence uniforme!!!). Est-il quand même possible d'écrire (5.1) de manière rigoureuse ?

Pour répondre à cette question, travaillons avec l'espace $L^1(]0, a[)$. Comme $L^2(]0, a[) \subset L^1(]0, a[)$, ce qui est fait ici est aussi valable pour $L^2(]0, a[)$.

5.4.1. *La première question est la suivante : $\lim_{n \rightarrow \pm\infty} c_n = 0$?* C'est une question centrale en pratique car elle permet lorsque l'on tronque la série infinie de savoir si l'approximation est correcte. De plus, pour que $\sum_n c_n e_n$ converge, il faut que $|c_n e_n| \rightarrow 0$ lorsque $n \rightarrow \infty$. Une condition nécessaire (mais non suffisante) est donc que $c_n \xrightarrow{n \rightarrow +\infty} 0$.

La réponse à cette question est donnée par le théorème de Riemann-Lebesgue.

THÉORÈME 16.6. *Soit $f \in L^1(]0; a[)$ une fonction intégrable. Alors, nous avons*

$$\int_0^a f(t) e^{-2i\pi nt/a} dt \xrightarrow{n \rightarrow \pm\infty} 0$$

5.4.2. *La seconde question est la suivante : a-t-on convergence ponctuelle de la série partielle $(f_n)_{n \in \mathbb{N}}$ vers f ?* C'est en général faux (!), comme le montrent la première question de l'exercice [16.3](#), ou l'exemple suivant.

EXEMPLE 16.1. On considère la fonction f impaire, de période 1, définie par

$$f(t) = \begin{cases} 1, & \text{pour } 0 \leq t < 1/2, \\ -1, & \text{pour } -1/2 \leq t < 0. \end{cases}$$

Cette fonction est continue **sauf** aux points de la forme $p/2$, $p \in \mathbb{Z}$, points pour lesquels, elle possède cependant une limite à gauche et une limite à droite (on note que la demi-somme de ces limites est nulle).

La fonction est impaire. On a donc $a_0(f) = a_n(f) = 0$, $\forall n \geq 1$, et après calculs : $b_n(f) = \frac{4}{\pi n}$ si n impair, $=0$ sinon. Sa série de Fourier est donnée par

$$\sum_{k=0}^{\infty} \frac{4 \sin 2\pi(2k+1)t}{\pi(2k+1)}.$$

Prenons $t = 0$. La série se réduit à $0 \neq f(0)$; il n'y a donc pas convergence au point 0 vers $f(0)$. Notons cependant que l'on trouve la demi-somme des limites à gauche et à droite de f en 0 : $\frac{1}{2}[f(0^-) + f(0^+)]$, qui est une façon assez naturelle de définir la valeur de f en 0.

Par contre, si nous prenons $t = 1/4$, la série de Fourier converge en ce point (série alternée) et sa somme est bien égale à $f(1/4)$. En effet, on a

$$\frac{4}{\pi} \sum_{k=0}^{\infty} \frac{1}{2k+1} \sin(2k+1)\frac{\pi}{2} = \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} = \frac{4}{\pi} \arctan(1) = 1.$$

Pour avoir convergence ponctuelle de la série de Fourier, il faut, en pratique, plus d'hypothèses de régularité sur f . D'un point de vue pratique le résultat suivant est le plus important. Avant cela, introduisons la définition suivante.

DÉFINITION 16.7. *Soit f , une fonction admettant en tout point, une limite à gauche et une limite à droite. On appelle **régularisée de la fonction** f la fonction*

$$\check{f} : x \rightarrow \frac{1}{2} [f(x^+) + f(x^-)]$$

On a alors le théorème suivant.

THÉORÈME 16.7. (de **Dirichlet**). *Soient $f \in L^1(]0; a[)$, $t_0 \in [0; a]$. Si*

- les limites à gauche et à droite $f(t_0^-)$ et $f(t_0^+)$
- et les limites à gauche et à droite $f'(t_0^-)$ et $f'(t_0^+)$

existent, alors, la suite des sommes partielles $(f_n(t_0))_{n \in \mathbb{N}}$ converge vers la régularisée de f en t_0 , c'est-à-dire :

$$\lim_{n \rightarrow +\infty} f_n(t_0) = \frac{1}{2} [f(t_0^-) + f(t_0^+)] = \check{f}(t_0)$$

REMARQUE 16.6. On remarque que la régularité est sur f mais également sur f' .

COROLLAIRE 16.2. *Si f est \mathcal{C}^1 par morceaux, alors la série de Fourier de f converge simplement vers la régularisée de f .*

5.4.3. *La troisième question est : a-t-on convergence uniforme de la série partielle $(f_n)_{n \in \mathbb{N}}$ vers f ?* Avant de commencer, rappelons ce qu'est la convergence uniforme d'une série de fonctions. La série $\sum_{n \in \mathbb{Z}} c_n(f) e_n$ converge uniformément vers f sur \mathbb{R} si :

$$\forall \varepsilon > 0, \quad \exists n_0 \in \mathbb{N}, \quad \forall n > n_0, \quad \text{on a} \quad \sup_{t \in \mathbb{R}} \left| \sum_{k=-n}^n c_n(f) e_n(t) - f(t) \right| < \varepsilon.$$

La convergence uniforme est une notion bien plus forte que la convergence simple. D'ailleurs elle implique la convergence simple.

Nous avons le

THÉORÈME 16.8. *Soit $f : \mathbb{R} \rightarrow \mathbb{C}$, une fonction continue, périodique et \mathcal{C}^1 par morceaux, alors la série de Fourier de f converge uniformément vers f sur \mathbb{R} et $\sum_{n \in \mathbb{Z}} |c_n(f)| < \infty$.*

DÉMONSTRATION. cf [13] p. 42. □

Nous avons également cet autre théorème :

THÉORÈME 16.9. *Soit $(c_n)_{n \in \mathbb{Z}}$ une suite telle que $\sum_{n \in \mathbb{Z}} |c_n| < \infty$, alors la série $\sum_{n \in \mathbb{Z}} c_n e_n(t)$ converge uniformément vers une fonction continue et périodique, de période a , dont le coefficient de Fourier d'ordre n est égal à c_n . C'est-à-dire, il existe f continue et a -périodique telle que pour tout $t \in \mathbb{R}$*

$$f(t) = \sum_{n \in \mathbb{Z}} c_n e_n(t),$$

et $c_n(f) = c_n$ pour tout $n \in \mathbb{Z}$.

5.4.4. *Que se passe-t-il si f n'est pas continue ?* Si f n'est pas continue, on n'a pas convergence uniforme. C'est ce que l'on appelle le phénomène de Gibbs (cf. figure 1).

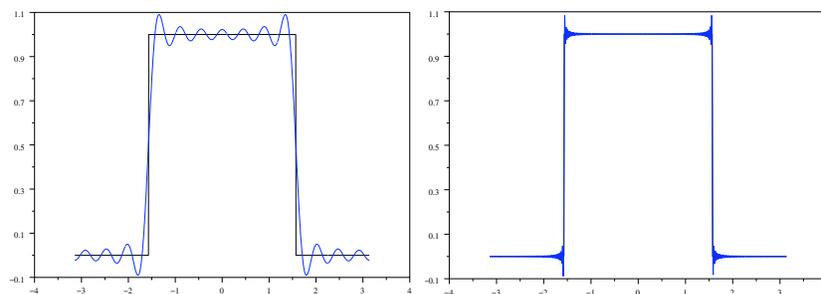


FIGURE 1 – Phénomène de Gibbs : représentation graphique de f_{13} et f_{199} pour $f(t) = \mathbf{1}_{[-\pi/2, \pi/2]}(t)$.

5.4.5. *Enfin, quel est le lien entre la régularité de f et la vitesse de décroissance des coefficients de Fourier ? On peut donner les différents points suivants.*

(a) On sait que si $f \in L^1(]0; a[)$, alors

$$c_n \xrightarrow[n \rightarrow +\infty]{} 0$$

d'après Riemann-Lebesgue.

(b) Si $f \in L^2(]0; a[)$, alors

$$\sum_{n \in \mathbb{Z}} |c_n|^2 < +\infty$$

d'après Parseval.

(c) Si $f \in \mathcal{C}^1(]0; a[)$, alors

$$\sum_{n \in \mathbb{Z}} |c_n| < +\infty$$

d'après le [Théorème 16.8](#).

REMARQUE 16.7. Si on a $\sum_{n \in \mathbb{Z}} |c_n| < \infty$, alors on a $\sum_{n \in \mathbb{Z}} |c_n|^2 < \infty$.

En effet, si $\sum_{n \in \mathbb{Z}} |c_n| < \infty$, alors à partir d'un certain rang, $|c_n| < 1$ (puisque les $|c_n|$ tendent vers 0 quand n tend vers l'infini), et donc $|c_n|^2 < |c_n|$, et donc la série des carrés converge.

(d) On peut montrer que si f est de classe \mathcal{C}^k sur $[0, a]$ et telle que $f(a) = f(0)$, alors

$$c_n(f^{(k)}) = \left(\frac{2i\pi n}{a} \right)^k c_n(f)$$

et donc

$$c_n(f) = \mathcal{O}\left(\frac{1}{n^k}\right).$$

(e) De plus, on a

$$f \in \mathcal{C}^\infty([0, a]) \text{ et telle que } f(a) = f(0) \implies c_n(f) = \mathcal{O}\left(\frac{1}{n^k}\right), \forall k \in \mathbb{N}.$$

DÉMONSTRATION. Preuve du point (d). Pour $f \in \mathcal{C}^1([0, a])$, a -périodique, on a

$$c_n(f) = \int_0^a f(t) e^{-2i\pi n t/a} dt = \left[\frac{-a}{2i\pi n} f(t) e^{-2i\pi n t/a} \right]_0^a + \frac{a}{2i\pi n} \int_0^a f'(t) e^{-2i\pi n t/a} dt = \frac{a}{2i\pi n} c_n(f').$$

Pour $f \in \mathcal{C}^k([0, a])$, on obtient le résultat en intégrant k fois par parties.

De plus, $|c_n(f)| = \left| \frac{1}{n^k} \left(\frac{a}{2i\pi} \right)^k c_n(f^{(k)}) \right| \leq M \frac{1}{n^k}$, puisque $|c_n(f^{(k)})|$ converge vers 0 lorsque n tend vers l'infini, d'où $c_n(f) = \mathcal{O}\left(\frac{1}{n^k}\right)$. \square

5.5. Convolution et série de Fourier. Soient f et g définies de \mathbb{R} dans \mathbb{C} , deux fonctions continues et a -périodiques. On note h le produit de convolution de f par g défini par

$$h(x) = (f * g)(x) = \frac{1}{\sqrt{a}} \int_0^a f(x-t) g(t) dt \quad (5.2)$$

On a alors le résultat suivant.

PROPOSITION 16.4. *Une fonction h définie par (5.2) est a -périodique. De plus, on a la relation au niveau des coefficients*

$$c_n(h) = c_n(f) c_n(g)$$

REMARQUE 16.8. La (5.2) de la convolution est un peu différente de la formule de convolution vue au chapitre 15. C'est une adaptation de la convolution dans le cadre des séries de Fourier.

6. Exercices

EXERCICE 16.2. (1) Donner le développement en série de Fourier de la fonction f , de période $T = 2$, définie sur $[0; 2[$ par

$$\begin{aligned} f(x) &= 1, & x \in [0; 1[\\ f(x) &= 0, & x \in [1; 2[\end{aligned}$$

(2) Que vous donne l'égalité de Parseval ?

EXERCICE 16.3. (1) Donner le développement en série de Fourier du signal carré

$$\begin{aligned} f(x) &= 1, & x \in [0; T/2[\\ f(x) &= -1, & x \in [T/2; T[\end{aligned}$$

(2) Qu'obtenez-vous en intégrant ?

EXERCICE 16.4. Soit une fonction u de \mathbb{R} dans \mathbb{C} de classe \mathcal{C}^1 et périodique de période L . Montrer que l'on a l'inégalité

$$\int_0^L |u(x)|^2 dx \leq \frac{L^2}{4\pi^2} \int_0^L |u'(x)|^2 dx + \frac{1}{L} \left| \int_0^L u(x) dx \right|^2$$

Supposons maintenant que $L = 2\pi$, et $\int_0^L u(x) dx = 0$. L'inégalité précédente peut-elle être une égalité ?

EXERCICE 16.5. (équation des cordes vibrantes.) On considère une corde dont les extrémités sont fixées. Le déplacement transversal, $u = u(x, t)$ est une fonction de l'abscisse x du point et du temps t .

Cette fonction u vérifie l'équation d'évolution aux dérivées partielles

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad (6.1)$$

pour a dans \mathbb{R} .

Le problème consiste à déterminer la forme de la corde à l'instant t connaissant les conditions initiales. Autrement dit, on veut résoudre l'équation (6.1) pour obtenir la fonction u . L'exemple traité ici est celui d'une corde fixée à ses extrémités et dont la forme initiale est une parabole. On suppose de plus que la vitesse initiale de ses points est nulle. Nous avons donc

$$u(0, t) = 0, \quad u(\ell, t) = 0, \quad (6.2)$$

et

$$u(x, 0) = bx(\ell - x), \quad \ell \in \mathbb{R}, \quad \frac{\partial u}{\partial t}(x, 0) = 0. \quad (6.3)$$

Les conditions (6.2) sont appelées conditions aux limites et les conditions (6.3) conditions initiales.

On utilise la méthode de séparation de variables ou encore de Fourier.

- (1) Déterminer les solutions non nulles de l'équation (6.1) qui s'écrivent sous la forme particulière $u(x, t) = X(x)T(t)$.
- (2) Donner les solutions qui vérifient (6.2).
- (3) Former la série obtenue en faisant la somme des solutions particulières trouvées dans la question précédente.
- (4) Déterminer celle qui vérifie (6.3) et que l'on peut dériver terme à terme par rapport à t .

EXERCICE 16.6. On cherche à calculer $u(x, y)$ solution de l'équation de Laplace dans le disque Ω centré à l'origine et de rayon R , avec condition de Dirichlet

$$\begin{cases} \Delta u = 0, & (x, y) \in \Omega, \\ u(x, y) = \varphi(x, y), & (x, y) \in \partial\Omega, \end{cases}$$

où $\Omega = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < R^2\}$, φ est une donnée au bord que nous supposons C^1 . Calculer la solution u à l'aide des séries de Fourier. On traitera en particulier et explicitement le cas

$$\varphi(x, y) = x^2 + y.$$

EXERCICE 16.7. Développer en série de Fourier la fonction f définie par

$$f(x) = \begin{cases} -\sin x, & \text{pour } -\pi \leq x < 0, \\ \sin x, & \text{pour } 0 \leq x < \pi, \end{cases}$$

EXERCICE 16.8. (1) Écrire la série de Fourier pour la fonction 2π -périodique impaire qui coïncide avec $x(\pi - x)$ sur $[0, \pi]$.

- (2) En utilisant la question précédente ainsi que l'identité de Parseval, en déduire la somme de la série

$$\sum_{k=1}^{\infty} \frac{1}{(2k-1)^6}.$$

EXERCICE 16.9. (*Equation de la chaleur*) Considérons une barre métallique. Connaissant, à l'instant initial, la température en chaque point de la barre et, à tout instant, la température aux deux extrémités, on peut déterminer, à tout moment et en tout point, la température de la barre. Ce problème a en effet été modélisé et la température u qui est une fonction du temps t et du point x est solution de l'équation de la chaleur :

$$\begin{cases} \frac{\partial u}{\partial t}(t, x) - \frac{\partial^2 u}{\partial x^2}(t, x) = 0, & t > 0, x \in]0, L[, \\ u(t, 0) = u(t, L) = 0, & t > 0, \\ u(0, x) = u_0(x), & x \in [0, L] \end{cases},$$

où u_0 est une fonction \mathcal{C}^1 sur $[0, L]$ telle que $u_0(0) = u_0(L) = 0$.

Trouver une solution de cette équation en utilisant les séries de Fourier. Vous vous aiderez des étapes de l'exercice 16.5.

EXERCICE 16.10. (*Le phénomène de Gibbs*) Soit $f(t) = t^2$ si $t \in [0, \pi]$, f paire et de période 2π . Soit $K(t) = t$ si $t \in]0, 2\pi]$, de période 2π .

On note $S_n(t) = h_0(t) + h_1(t) + \dots + h_n(t)$ la somme des $n+1$ premiers harmoniques.

- (1) Tracer les graphes de S_1, S_3, S_{15} et S_{40} pour les fonctions f et K , à l'aide d'une calculatrice ou d'un ordinateur. Que constate-t-on ?
- (2) On s'intéresse à la fonction K et à $S_n(t) = \pi - 2 \sum_{k=1}^n \frac{\sin kt}{k}$.

(a) Calculer $S'_n(t)$ en utilisant la relation :

$$\cos \theta + \cos 2\theta + \dots + \cos n\theta = \cos \left(\frac{n+1}{2} \theta \right) \frac{\sin n\frac{\theta}{2}}{\sin \frac{\theta}{2}} = \frac{\sin \left(\left(n + \frac{1}{2} \right) \theta \right)}{2 \sin \frac{\theta}{2}} - \frac{1}{2}$$

- (b) Quelle est la plus petite valeur positive de t correspondant à un extremum de S_n ? Quelle est la valeur m_n de cette extremum ?
- (c) Quelle est la limite de m_n quand $n \rightarrow \infty$?

La transformée de Laplace

La transformée de Laplace est un outil qui peut simplifier la résolution d'équations différentielles, notamment lorsque l'on travaille sur une équation différentielle dont le second membre est un champ de force donné. La transformée de Laplace est une des nombreuses transformées que l'on peut rencontrer en physique et ingénierie (transformée de Fourier, de Radon, etc...). Sa compréhension est donc un premier pas vers ces outils de calcul.

1. Intégrale de Laplace

La variable t , dans tout ce qui suit, sera réelle, et on la supposera positive (elle représentera souvent le temps).

On fait correspondre à une fonction f de la variable t , une fonction $\mathcal{L}f$ d'une nouvelle variable p par la formule

$$\mathcal{L}f(p) = \int_0^{\infty} e^{-pt} f(t) dt \quad (1.1)$$

La variable p sera généralement complexe, on posera $p = x + iy$. Si l'intégrale précédente est définie, $\mathcal{L}f$ s'appelle la **transformée de Laplace** de f .

2. Hypothèses sur f

On ne peut pas calculer la transformée de Laplace de n'importe quelle fonction f . On supposera que

- (1) f est nulle sur \mathbb{R}^{-*} (c'est une convention, dont l'utilité apparaîtra plus tard), une telle fonction est appelée **causale**,
- (2) f est définie et continue par morceaux sur \mathbb{R}^+ ,
- (3) il existe une constante M positive telle que le produit

$$e^{-Mt}|f(t)|$$

reste borné pour toutes les valeurs de t assez grandes (i.e. pour $t > t_0$ avec $t_0 \in \mathbb{R}^+$, t_0 dépendant de M et de f).

La dernière condition signifie que f est d'**ordre exponentiel** M :

DÉFINITION 17.1. Une fonction f telle que

$$|f(t)| \leq C e^{Mt}, \quad \text{pour tout } t \geq t_0$$

avec $C \in \mathbb{R}^+$, est dite d'**ordre exponentiel** M .

Cette condition est suffisante pour que l'intégrale (1.1) soit définie pour toutes les valeurs de $p = x + iy$ telles que $x > M$.

THÉORÈME 17.1. Si f est une fonction continue par morceaux sur \mathbb{R}^+ , et d'ordre exponentiel M , alors, $\forall p \in \mathbb{C}$ tel que $\Re(p) > M$, l'intégrale (1.1) est bien définie.

DÉMONSTRATION. Dire que $e^{-Mt}|f(t)|$ est borné pour $t > t_0$ équivaut à dire qu'il existe une constante C telle que

$$t > t_0 \text{ entraîne } e^{-Mt}|f(t)| < C.$$

Alors on a évidemment

$$|e^{-pt}f(t)| < C e^{(M-x)t}$$

et comme l'intégrale $\int_{t_0}^{\infty} e^{(M-x)t} dt$ est convergente car $M - x < 0$, l'intégrale (1.1) est bien définie. □

Citons également un autre critère, plus général :

THÉORÈME 17.2. Soit f une fonction continue par morceaux sur \mathbb{R}^+ . Supposons qu'il existe un nombre positif p_0 et un nombre positif C tels que, pour toute valeur $A > 0$, on ait

$$\left| \int_0^A e^{-p_0 t} f(t) dt \right| < \infty$$

alors l'intégrale (1.1) est définie pour tout p tel que $\Re(p) > p_0$.

3. Premières propriétés de la transformée de Laplace

Linéarité. Considérons deux fonctions f et g telles que l'on puisse définir leur transformée de Laplace, pour $\Re(p) > p_f$ pour f et $\Re(p) > p_g$ pour g . Alors, nous avons

$$\forall (a, b) \in \mathbb{R}, \quad \mathcal{L}\{af + bg\}(p) = a\mathcal{L}f(p) + b\mathcal{L}g(p)$$

pour tout p tel que $\Re(p) > \max(p_f, p_g)$.

Dans la suite, on considère une fonction f continue par morceaux sur \mathbb{R}^+ , et dont la transformée de Laplace existe pour $\Re(p) > p_f$, avec $p_f \in \mathbb{R}^+$.

Changement d'échelle. Soit f une fonction admettant une transformée de Laplace, et k une constante positive. Si on change l'échelle dans f en créant la nouvelle fonction $\tilde{f}(t) = f(kt)$, alors

$$\mathcal{L}\tilde{f}(p) = \frac{1}{k}\mathcal{L}f\left(\frac{p}{k}\right),$$

pour $\Re(p) > kp_f$.

La démonstration est immédiate en utilisant le changement de variables $u = kt$.

Translation sur la variable symbolique p . Si $\mathcal{L}f(p)$ est la transformée de Laplace de $f(t)$, alors

$$\mathcal{L}f(p+a) = \mathcal{L}\{e^{-at}f(t)\}(p),$$

pour $\Re(p+a) > p_f$.

En effet : $\mathcal{L}\{e^{-at}f(t)\}(p) = \int_0^\infty e^{-(a+p)t}f(t)dt = \mathcal{L}f(p+a)$, à condition que $\mathcal{L}f(p+a)$ soit définie.

EXERCICE 17.1. *Calculer*

$$\mathcal{L}\{e^{3t}\sin(2t)\}(p).$$

Translation sur la variable temporelle t . Soit $h > 0$, alors :

$$\mathcal{L}\{f(t-h)\} = e^{-ph}\mathcal{L}f(p),$$

pour $\Re(p) > p_f$.

En effet, ayant supposé que f est nulle sur \mathbb{R}^{-*} , on a $f(t-h) = 0$ pour $t < h$:

$$\mathcal{L}\{f(t-h)\}(p) = \int_0^\infty e^{-pt}f(t-h)dt = e^{-ph}\int_h^\infty e^{-p(t-h)}f(t-h)dt = e^{-ph}\int_0^\infty e^{-pu}f(u)du.$$

Considérons une fonction f nulle pour $t < 0$, et périodique, de période T . Soit la fonction f_0 égale à f sur le segment $[0, T[$, nulle à l'extérieur. Elle a pour transformée de Laplace

$$\varphi_0(p) = \mathcal{L}f_0(p).$$

On peut alors représenter $f(t)$ entre 0 et $+\infty$ comme la somme de la série

$$f_0(t) + f_0(t-T) + f_0(t-2T) + \dots + f_0(t-nT) + \dots$$

et par conséquent, sa transformée de Laplace est

$$\mathcal{L}f(p) = \varphi_0(p)[1 + e^{-pt} + \dots + e^{-npt} + \dots] = \frac{\varphi_0(p)}{1 - e^{-pT}},$$

à condition qu'on ait le droit d'invertir l'intégrale et la série et que $|e^{-pT}| < 1$.

EXERCICE 17.2. *Prenons $f(t) = \sin t$ de période 2π . Retrouver alors que*

$$\mathcal{L}(\sin)(p) = \frac{1}{1 + p^2}.$$

Transformée d'une dérivée. Le théorème suivant est fondamental dans la théorie du calcul symbolique.

THÉORÈME 17.3. *Si f est une fonction de classe \mathcal{C}^1 sur \mathbb{R}^+ , si elle est d'ordre exponentiel M et si f' admet une transformée de Laplace, alors*

$$\mathcal{L}\{f'\}(p) = p\mathcal{L}f(p) - f(0),$$

pour $\Re(p) > M$.

DÉMONSTRATION. Soit p tel que $\Re(p) > M$, alors par intégration par parties, on a :

$$\int_0^{\infty} e^{-pt} f'(t) dt = [e^{-pt} f(t)]_0^{\infty} + p \int_0^{\infty} e^{-pt} f(t) dt.$$

Or, $|e^{-pt} f(t)| = e^{-(\Re(p)-M)t} e^{-Mt} |f(t)| \leq C e^{-(\Re(p)-M)t} \rightarrow 0$ lorsque $t \rightarrow \infty$ car $\Re(p) > M$.

On en déduit que la $\mathcal{L}\{f'\}(p)$ existe pour $\Re(p) > M$ et la formule annoncée. \square

Si l'on applique plusieurs fois la règle précédente, on trouve, si f est de classe \mathcal{C}^2 , et f, f' d'ordre exponentiel M :

$$\mathcal{L}\{f''\}(p) = p^2 \mathcal{L}f(p) - pf(0) - f'(0),$$

et plus généralement, si f est de classe \mathcal{C}^n et f, f', \dots, f^{n-1} d'ordre exponentiel M :

$$\mathcal{L}\{f^n\}(p) = p^n \mathcal{L}f(p) - p^{n-1} f(0) - p^{n-2} f'(0) - \dots - f^{(n-1)}(0).$$

REMARQUE 17.1. (1) Une fonction f peut avoir une transformée de Laplace, sans que sa dérivée f' en ait une.

La fonction $f(t) = t^{-1/2}$ a une transformée de Laplace valant $\sqrt{\pi/p}$, mais $f'(t)$ n'a pas de transformée de Laplace, car pour cette fonction, l'intégrale de Laplace diverge en 0.

(2) La transformation de Laplace fait correspondre à la dérivation par rapport à t la multiplication de la transformée de Laplace par p , avec l'addition d'une constante. C'est la simplicité de cette règle qui explique l'emploi du calcul symbolique.

EXERCICE 17.3. Trouver la transformée de Laplace de y lorsque y est solution du problème initial

$$\begin{cases} \frac{d^2 y}{dt^2} - 7 \frac{dy}{dt} + 12y = 16e^{2t} \\ y(0) = 6 \\ \frac{dy}{dt}(0) = 4 \end{cases}$$

Transformée d'une primitive. Soit f une fonction continue sur \mathbb{R}^+ admettant une transformée de Laplace. Si $F(t)$ est une primitive de $f(t)$, nulle pour $t = 0$, admettant une transformée de Laplace, on a :

$$\mathcal{L}F(p) = \frac{1}{p}\mathcal{L}f(p).$$

C'est une conséquence immédiate du théorème précédent. On peut écrire

$$\mathcal{L}\left\{\int_0^t f(u)du\right\}(p) = \frac{1}{p}\mathcal{L}f(p)$$

et en appliquant n fois cette règle :

$$\mathcal{L}\left\{\int_0^t \int_0^{t_1} \dots \int_0^{t_{n-1}} f(u)du dt_1 \dots dt_{n-1}\right\}(p) = \frac{1}{p^n}\mathcal{L}f(p).$$

Comportement de $\mathcal{L}f(p)$ pour $p = 0$ ou $+\infty$.

(1) Si $\int_0^\infty f(t)dt$ converge, alors $\int_0^\infty f(t)dt = \mathcal{L}f(0)$.

Démonstration évidente.

(2) Si f est de classe \mathcal{C}^1 et vérifie les hypothèses du [Théorème 17.3](#), on peut écrire :

$$\int_0^\infty f'(t)dt = \mathcal{L}\{f'\}(0) = \lim_{p \rightarrow 0} p\mathcal{L}f(p) - f(0)$$

et d'autre part,

$$\int_0^\infty f'(t)dt = \lim_{t \rightarrow \infty} f(t) - f(0),$$

on en déduit

$$\lim_{t \rightarrow \infty} f(t) = \lim_{p \rightarrow 0} p\mathcal{L}f(p).$$

(3) On peut montrer, et nous l'admettons, que si f est continue à droite en 0 et si l'argument de p reste compris entre des valeurs $-a$ et $+a$ avec $0 < a < \pi/2$, alors

$$\lim_{p \rightarrow \infty} p\mathcal{L}f(p) = f(0^+).$$

4. Transformées de fonctions usuelles

On considère ici des fonctions causales.

EXERCICE 17.4. Soit $f(t) = t^n$, avec $n \in \mathbb{N}$. Montrer que

$$\mathcal{L}f(p) = \frac{n!}{p^{n+1}}, \quad \text{si } \Re(p) > 0.$$

EXERCICE 17.5. Soit $f(t) = e^{at}$, avec $a \in \mathbb{R}$. Montrer que

$$\mathcal{L}f(p) = \frac{1}{p-a}, \quad \text{si } \Re(p) > a.$$

EXERCICE 17.6. Montrer que, si $f(t) = e^{iat}$, avec $a \in \mathbb{R}$, nous avons

$$\mathcal{L}f(p) = \frac{1}{p - ia}, \quad \text{si } \Re(p) > 0.$$

EXERCICE 17.7. Montrer que les transformées de Laplace des fonctions $\cos(\omega t)$ et $\sin(\omega t)$, pour tout ω réel, sont données par

$$\mathcal{L}\{\cos(\omega t)\}(p) = \frac{p}{p^2 + \omega^2}, \quad \text{si } \Re(p) > 0.$$

et

$$\mathcal{L}\{\sin(\omega t)\}(p) = \frac{\omega}{p^2 + \omega^2}, \quad \text{si } \Re(p) > 0.$$

EXERCICE 17.8. Montrer que les transformées de Laplace des fonctions $\cosh(bt)$ et $\sinh(bt)$, pour tout b réel, sont données par

$$\mathcal{L}\{\cosh(bt)\}(p) = \frac{p}{p^2 - b^2}, \quad \text{si } \Re(p) > |b|.$$

et

$$\mathcal{L}\{\sinh(bt)\}(p) = \frac{b}{p^2 - b^2}, \quad \text{si } \Re(p) > |b|.$$

Maintenant, on désire calculer

$$\mathcal{L}\{t^\alpha\}(p) = \int_0^\infty e^{-pt} t^\alpha dt, \quad p > 0$$

avec $\alpha \in]-1; +\infty[$ (pour assurer la convergence de l'intégrale).

Posons $u = tp$. Alors, par changement de variables, nous avons

$$\mathcal{L}\{t^\alpha\}(p) = \frac{1}{p^{\alpha+1}} \int_0^\infty e^{-u} u^\alpha du$$

pour $p > 0$ et $\alpha > -1$. La dernière intégrale est indépendante de p , on peut donc poser

$$\mathcal{L}\{t^\alpha\}(p) = \frac{c_\alpha}{p^{\alpha+1}}$$

où

$$c_\alpha = \int_0^\infty e^{-u} u^\alpha du$$

Introduisons la fonction Γ définie par

$$\Gamma(x) = \int_0^\infty e^{-u} u^{x-1} du.$$

Alors, $c_\alpha = \Gamma(\alpha + 1)$, ce qui implique que

$$\mathcal{L}\{t^\alpha\}(p) = \frac{\Gamma(\alpha + 1)}{p^{\alpha+1}}, \quad p > 0.$$

On peut montrer que la fonction Γ , bien connue dans le domaine des mathématiques, est telle que

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

et

$$\Gamma(x+1) = x\Gamma(x), \quad x > 0.$$

Considérons enfin la fonction de Heaviside H et prenons la translatée de cette fonction par a : $H(t-a)$, telle que

$$u_a(t) = \begin{cases} 0, & \text{si } t < a \\ 1, & \text{si } t \geq a \end{cases}$$

Alors, nous en déduisons que

$$\mathcal{L}\{u_a(t)\}(p) = \int_0^{\infty} e^{-pt} u_a(t) dt$$

En découpant le calcul entre $[0; a]$ et $[a; \infty]$, on obtient, si $\Re(p) > 0$,

$$\mathcal{L}\{u_a(t)\}(p) = \frac{1}{p} e^{-ap}, \quad a \geq 0.$$

Il existe des tables et on peut s'y référer pour connaître des transformées de Laplace connues. Une autre possibilité consiste également à utiliser des logiciels de calcul formel (Mathematica par exemple) pour déterminer ces transformées.

5. Dérivation et intégration de la transformée de Laplace

Dérivation de $\mathcal{L}f$.

THÉORÈME 17.4. *Soit $\varphi = \mathcal{L}\{f\}$, où f est une fonction continue par morceaux d'ordre exponentiel M . Alors $\varphi(p)$ est différentiable en tout p tel que $\Re(p) > M$ et*

$$\mathcal{L}\{tf(t)\}(p) = -\frac{d\varphi}{dp}$$

DÉMONSTRATION. On applique le théorème de dérivation sous l'intégrale : la fonction $t \mapsto e^{-pt}f(t)$ est continue par morceaux et intégrable sur \mathbb{R}^+ , la fonction $t \mapsto -te^{-pt}f(t)$ est continue par morceaux, dérivable et intégrable sur \mathbb{R}^+ , et pour tout p, t , $|-te^{-pt}f(t)| \leq Cte^{-Mt}$, qui est intégrable sur \mathbb{R}^+ dès que $\Re(p) > M$. On a donc le droit de dériver sous l'intégrale et on obtient la formule du théorème. \square

REMARQUE 17.2. On a démontré le théorème dans le cas $p \in \mathbb{R}$. Le cas $p \in \mathbb{C}$ demande d'avoir eu le cours sur la dérivation des fonctions complexes, au second semestre.

On a également la proposition.

PROPOSITION 17.1. Soit $\varphi = \mathcal{L}\{f\}$, où f est une fonction continue par morceaux d'ordre exponentiel M . Alors $\varphi(p)$ est infiniment différentiable pour p tel que $\Re(p) > M$ et

$$\mathcal{L}\{t^n f(t)\}(p) = (-1)^n \frac{d^n \varphi}{dp^n}, \quad n = 1, 2, \dots$$

EXERCICE 17.9. Soit la fonction de Heaviside $H(t) = 1$ pour tout $t \in \mathbb{R}^+$ et 0 sur \mathbb{R}^{-*} . Montrer que

$$\mathcal{L}H(p) = \begin{cases} \frac{1}{p}, & \text{si } \Re(p) > 0 \\ \infty, & \text{sinon} \end{cases}$$

En déduire

(1) la transformée de Laplace de la fonction $f(t) = a \cdot H(t)$, où a est une constante complexe.

(2) la transformée de Laplace de la fonction $f(t) = t^n H(t)$, $n \in \mathbb{N}$.

EXERCICE 17.10. Calculer la transformée de Laplace de $t \sin(3t)$.

Intégration de $\mathcal{L}f$. On a, pour $p \in \mathbb{R}^{+*}$,

$$\int_p^\infty \mathcal{L}f(s) ds = \mathcal{L}\left\{\frac{f(t)}{t}\right\}(p).$$

En effet, formellement, en intégrant par rapport à p l'égalité (1.1) de définition de $\mathcal{L}f$, on a :

$$\int_p^\infty \mathcal{L}f(s) ds = \int_0^\infty \int_p^\infty e^{-st} f(t) ds dt = \int_0^\infty \frac{e^{-pt}}{t} f(t) dt$$

en admettant que l'interversion des intégrations est légitime.

Dérivation par rapport à un paramètre. Soit $f(t, \alpha)$ une fonction dépendant d'un paramètre α , dérivable par rapport à α , qui admette une transformée de Laplace $\mathcal{L}f(p, \alpha)$, et telle que la dérivée $\frac{\partial f}{\partial \alpha}$ admette une transformée de Laplace $\mathcal{L}_1 f(p, \alpha)$.

Si les hypothèses du théorème de dérivation sous l'intégrale sont vérifiées, alors la fonction $\mathcal{L}_1 f(p, \alpha)$ est la dérivée de $\mathcal{L}f(p, \alpha)$ par rapport à α .

Exercice : Trouver la transformée de Laplace $Y(s) = \mathcal{L}\{y(t)\}|_s$, lorsque y est solution du problème initial

$$\begin{aligned} y'' - 7y' + 12y &= 16e^{2t} \\ y(0) &= 6 \\ y'(0) &= 4 \end{aligned}$$

Solution : Prenons la transformée de Laplace de l'équation :

$$\mathcal{L}\{y'' - 7y' + 12y\}|_s = \mathcal{L}\{16e^{2t}\}|_s$$

Alors, ceci se signifie :

$$[s^2 Y(s) - s y(0) - y'(0)] - 7[s Y(s) - y(0)] + 12Y(s) = 16 \frac{1}{s-2}$$

ou encore, ??? les calculs et en utilisant : $y'(0) = 4$
 $y(0) = 6$

$$Y(s) = \frac{6s + 38}{s^2 - 7s + 12} + \frac{16}{(s-2)(s^2 - 7s + 12)}$$

Exercice : Calculer la transformée de Laplace de $t \sin(3t)$

Solution : Nous avons :

$$\begin{aligned} \mathcal{L}\{f(t)\} &= \mathcal{L}\{t \sin(3t)\} \\ &= -\frac{d}{ds} \mathcal{L}\{\sin(3t)\} \Big|_s \\ &= -\frac{d}{ds} \frac{3}{s^2 + 9} = \frac{6s}{(s^2 + 9)^2} \end{aligned}$$

6. La transformée de Laplace inverse

Injectivité de la transformée de Laplace. Comme nous l'avons vu, la transformation de Laplace est très intéressante lorsque l'on veut simplifier des calculs, par exemple sur une

THÉORÈME 17.5. (*unicité ou injectivité de la transformée*). *Supposons que f et g soient deux fonctions continues par morceaux d'ordre exponentiel et ayant les mêmes transformées de Laplace*

$$\mathcal{L}\{f\} = \mathcal{L}\{g\}.$$

Alors, nous avons l'égalité

$$f(t) = g(t), \quad \forall t \geq 0.$$

REMARQUE 17.3. En fait, c'est une égalité presque partout (i.e. partout sauf en des points isolés), et c'est une égalité partout si f et g sont continues.

Dans cette dernière proposition, nous voyons que nous pouvons conclure pour $t \geq 0$. Que se passe-t-il alors pour $t < 0$? En fait, nous ne pouvons pas reconstruire d'informations sur f pour $t < 0$ puisque la formule donnant la transformée de Laplace a lieu pour $t \geq 0$. En pratique, nous supposons que nous résolvons un problème temporel et donc que f est une fonction causale : $f(t) = 0$, pour $t < 0$.

Formule de la transformée de Laplace inverse. Définissons maintenant la transformée de Laplace inverse : étant donnée une fonction $\varphi(p)$, la transformée de Laplace inverse de φ , notée $\mathcal{L}^{-1}\{\varphi\}$, est la fonction f dont la transformée de Laplace est φ .

On n'a, grâce à l'unicité de la représentation, aucune ambiguïté sur ce qu'est réellement $\mathcal{L}^{-1}\{\varphi\}$. Ceci a pour conséquence qu'une table de transformée de Laplace peut être lue "de gauche à droite" ou "de droite à gauche". Nous avons donc :

$$f(t) = \mathcal{L}^{-1}\{\varphi(p)\}(t) \iff \mathcal{L}\{f(t)\}(p) = \varphi(p)$$

Avant de passer à la suite, notons qu'il existe une autre formule (de Mellin-Fourier) pour calculer une transformée de Laplace inverse. Nous pouvons, en passant dans le plan complexe, montrer que

$$\mathcal{L}^{-1}\{\varphi(p)\}(t) = \frac{1}{2i\pi} \lim_{\beta \rightarrow \infty} \int_{\gamma-i\beta}^{\gamma+i\beta} e^{tz} \varphi(z) dz$$

L'intégrale est prise le long d'une ligne se trouvant dans le plan complexe. Toutefois, nous n'aborderons pas l'utilisation de cette relation ici qui demande à ce que vous ayez déjà eu un cours d'analyse complexe plus poussé (faisant l'objet du cours du deuxième semestre).

Linéarité de la transformée de Laplace inverse. De manière similaire à la transformée de Laplace, la transformée de Laplace inverse est une transformation linéaire.

THÉORÈME 17.6. *La transformée de Laplace est une transformation linéaire, c'est-à-dire :*

$$\mathcal{L}^{-1}\{a\varphi_1(p) + b\varphi_2(p)\} = a\mathcal{L}^{-1}\{\varphi_1(p)\} + b\mathcal{L}^{-1}\{\varphi_2(p)\}$$

où $a, b \in \mathbb{C}$ et φ_1, φ_2 sont des fonctions admettant une transformée de Laplace inverse.

La linéarité permet en particulier de trouver la transformée de Laplace inverse d'une fraction rationnelle $\varphi(p)$ pour laquelle le degré du dénominateur est supérieur au degré du numérateur. Il suffit de décomposer $\varphi(p)$ en éléments simples de première et seconde espèces, et d'utiliser les propriétés de la transformée de Laplace (translation, changement d'échelle, dérivation, intégration).

Nous illustrons ci-dessous l'intérêt de la propriété

$$\mathcal{L}^{-1}\{\varphi(p-a)\}(t) = e^{at} f(t)$$

où $\varphi = \mathcal{L}f$.

Cette formule peut s'avérer très utile comme le montre l'exemple suivant. On se propose de calculer la transformée de Laplace inverse

$$\frac{1}{p^2 - 8p + 25}$$

On a l'identité

$$\frac{1}{p^2 - 8p + 25} = \frac{1}{(p - 4)^2 + 9}$$

Ainsi, une application de la transformée de Laplace inverse, suivie d'une translation donne

$$\mathcal{L}^{-1} \left\{ \frac{1}{p^2 - 8p + 25} \right\} (t) = \mathcal{L}^{-1} \{ \varphi(p - 4) \}$$

où

$$\varphi(p) = \frac{1}{p^2 + 9}$$

Ceci implique que :

$$\mathcal{L}^{-1} \left\{ \frac{1}{p^2 - 8p + 25} \right\} (t) = e^{4t} \mathcal{L}^{-1} \left\{ \frac{1}{p^2 + 9} \right\}$$

Mais, nous avons également

$$\mathcal{L}^{-1} \left\{ \frac{1}{p^2 + 9} \right\} = \frac{1}{3} \mathcal{L}^{-1} \left\{ \frac{3}{p^2 + 9} \right\} = \frac{\sin 3t}{3}$$

d'où la formule finale

$$\mathcal{L}^{-1} \left\{ \frac{1}{p^2 - 8p + 25} \right\} (t) = \frac{e^{4t} \sin(3t)}{3}$$

EXERCICE 17.11. Calculer la transformée de Laplace inverse de

$$\frac{1}{p^2 + p - 6}$$

EXERCICE 17.12. Trouver la transformée de Laplace inverse de

$$\frac{3p + 1}{p^4 + p^2}.$$

EXERCICE 17.13. Trouver la transformée de Laplace inverse de

$$\frac{p^2 + 2}{p(p + 1)(p + 2)}.$$

Exercice : trouver la transformée de Laplace inverse de

$$F(s) = \frac{30}{s^7} + \frac{8}{s - 4}$$

Solution : Nous savons que :

$$\mathcal{L}^{-1} \left\{ \frac{6!}{s^7} \right\} \Big|_{t=0} = t^6$$

et

$$\mathcal{L}^{-1} \left\{ \frac{1}{s - 4} \right\} \Big|_{t=0} = e^{4t}$$

ce qui donne :

$$\begin{aligned}\mathcal{L}^{-1}\left\{\frac{30}{s^7} + \frac{8}{s-4}\right\} &= \frac{30}{6!}t^6 + 8e^{4t} \\ &= \frac{1}{24}t^6 + 8e^{4t}\end{aligned}$$

Exercice : Calculer la transformée de Laplace inverse :

$$\mathcal{L}^{-1}\left\{\frac{1}{s^4 + s - 6}\right\} |t$$

Solution : Une décomposition en fractions rationnelles montre que :

$$\frac{1}{s^4 + s - 6} = \frac{1}{5(s-2)} - \frac{1}{5(s+3)}$$

Par conséquent,

$$\begin{aligned}\mathcal{L}^{-1}\left\{\frac{1}{s^4 + s - 6}\right\} |t &= \frac{1}{5}\mathcal{L}^{-1}\left\{\frac{1}{s-2}\right\} - \frac{1}{5}\mathcal{L}^{-1}\left\{\frac{1}{s+3}\right\} \\ &= \frac{1}{5}e^{2t} - \frac{1}{5}e^{-3t}\end{aligned}$$

Exercice : Trouver la transformée de Laplace inverse de $\frac{3s+1}{s^4+s^2}$

Solution : On montre qu'on a la décomposition :

$$\frac{3s+1}{s^4+s^2} = \frac{As+B}{s^2} + \frac{Cs+D}{S^2+1}$$

Des calculs simples montrent que :

$$A = 3, B = 1, C = -3, D = -1$$

Ainsi :

$$\frac{3s+1}{s^4+s^2} = \frac{3}{S} + \frac{-3s}{s^2+1} - \frac{1}{s^2+1}$$

En considérant la transformée de Laplace inverse, nous avons :

$$\mathcal{L}^{-1}\left\{\frac{3s+1}{s^4+s^2}\right\} |t = 3\mathcal{L}^{-1}\left\{\frac{1}{s}\right\} |t + \mathcal{L}^{-1}\left\{\frac{1}{s^2}\right\} |t - 3\mathcal{L}^{-1}\left\{\frac{s}{s^2+1}\right\} |t - \mathcal{L}^{-1}\left\{\frac{1}{s^2+1}\right\} |t$$

ou encore :

$$\mathcal{L}^{-1}\left\{\frac{3s+1}{s^4+s^2}\right\} |t = 3 + \frac{1}{t} - 3\cos(t) - \sin(t)$$

7. Convolution et produit de transformées de Laplace

Afin de voir l'intérêt de la transformation de Laplace lors du calcul de convolution, donnons le théorème suivant :

THÉORÈME 17.7. *Soient deux fonctions f et g , d'ordre exponentiel M . On note leurs transformées de Laplace*

$$\mathcal{L}f(p) = \int_0^{\infty} e^{-pt} f(t) dt$$

et

$$\mathcal{L}g(p) = \int_0^{\infty} e^{-pt} g(t) dt.$$

Alors $f * g$ est aussi d'ordre exponentiel M , sa transformée de Laplace existe, et elle est égale au produit des transformées de Laplace de f et g :

$$\mathcal{L}(f * g)(p) = \mathcal{L}f(p) \cdot \mathcal{L}g(p)$$

De plus

$$\mathcal{L}^{-1}\{\mathcal{L}f(p) \cdot \mathcal{L}g(p)\}(t) = f * g(t)$$

DÉMONSTRATION. Soit p tel que $\Re(p) > M$, montrons que $f * g$ admet une transformée de Laplace.

Montrons que l'intégrale

$$\int_0^{\infty} |(f * g)(t) e^{-pt}| dt = \int_0^{\infty} \int_0^t e^{-\Re(p)t} |f(x) g(t-x)| dx dt$$

est finie.

Si nous posons $\mathcal{R} = \{(t, x) / 0 < t < \infty, 0 < x < t\}$, nous avons

$$\int_0^{\infty} \int_0^t e^{-\Re(p)t} |f(x) g(t-x)| dx dt = \iint_{\mathcal{R}} e^{-\Re(p)t} |f(x) g(t-x)| dx dt \quad (7.1)$$

La région \mathcal{R} désigne donc le deuxième octant du plan (t, x) que l'on peut récrire

$$\mathcal{R} \{(t, x) / 0 < x < +\infty ; x < t < +\infty\}$$

Il s'ensuit que (7.1) peut se récrire :

$$\int_0^{+\infty} \int_x^{+\infty} e^{-\Re(p)t} |f(x) g(t-x)| dt dx$$

ou encore :

$$\int_0^{+\infty} |f(x)| \left\{ \int_x^{+\infty} e^{-\Re(p)t} |g(t-x)| dt \right\} dx$$

Effectuons le changement de variables : $y = t - x$. Nous avons alors :

$$\int_0^{+\infty} |f(x)| \left(\int_0^{+\infty} e^{-\Re(p)(y+x)} |g(y)| dy \right) dx$$

c'est-à-dire :

$$\mathcal{L}|f|(\Re(p)) \cdot \mathcal{L}|g|(\Re(p)) < \infty$$

car $|f|$ et $|g|$ sont d'ordre exponentiel M et $\Re(p) > M$.

La transformée de Laplace de $f * g$ existe donc, pour $\Re(p) > M$. Les mêmes calculs montrent alors que $\mathcal{L}(f * g)(p) = \mathcal{L}f(p) \cdot \mathcal{L}g(p)$. \square

EXERCICE 17.14. Calculer la transformée de Laplace inverse de la fonction

$$\frac{1}{\sqrt{p}(p^2 + 9)},$$

$p > 0$.

8. Convolution et équations différentielles

La transformée de Laplace joue un rôle essentiel notamment dans la résolution des équations différentielles et, en particulier, par la convolution. Considérons une équation différentielle du second ordre

$$\begin{cases} ay'' + by' + cy = f \\ y(0) = y'(0) = 0 \end{cases}$$

où a , b et c sont des constantes.

Si nous posons : $Y(p) = \mathcal{L}\{y(t)\}(p)$ et $F(p) = \mathcal{L}\{f(t)\}(p)$, alors des calculs immédiats donnent

$$Y(p) = W(p)F(p)$$

avec $W(p) = \frac{1}{ap^2 + bp + c}$. Par conséquent, nous avons

$$y(t) = (w * f)(t) = \int_0^t w(x) f(t-x) dx$$

où

$$w(x) = \mathcal{L}^{-1} \left\{ \frac{1}{ap^2 + bp + c} \right\} (t)$$

Cette dernière forme de la solution est souvent appelée **principe de Duhamel**, la fonction W est appelée **fonction de transfert** et son inverse, **fonction poids** ou (fonction) **réponse impulsionnelle**. La fonction f est le **signal d'entrée** et y le **signal de sortie**. C'est une formule intéressante si, par exemple, on veut calculer y pour différents temps. On peut généraliser ce résultat pour des systèmes d'ordres plus élevés.

THÉORÈME 17.8. (*Principe de Duhamel*). Soit $n \in \mathbb{N}^*$, $a_0, \dots, a_n \in \mathbb{R}$, f une fonction admettant une transformée de Laplace. Alors la solution de

$$\begin{aligned} \sum_{j=0}^n a_j \frac{d^j}{dt^j} y(t) &= f(t) \\ \frac{d^j}{dt^j} y(0) &= 0, \quad 0 \leq j \leq n-1 \end{aligned} \quad (8.1)$$

est donnée par :

$$y(t) = (w * f)(t) = \int_0^t w(x) f(t-x) dx$$

où

$$w(x) = \mathcal{L}^{-1} \{W(p)\}(x) = \mathcal{L}^{-1} \left\{ \frac{1}{\sum_{j=0}^n a_j p^j} \right\} (x)$$

Il est facile de montrer que si l'on veut calculer la solution de :

$$\sum_{j=0}^n a_j \frac{d^j}{dt^j} y(t) = f(t)$$

avec des conditions initiales non nulles, il suffit d'ajouter la solution obtenue par le principe de Duhamel à la solution de l'équation différentielle homogène

$$\sum_{j=0}^n a_j \frac{d^j}{dt^j} y(t) = 0$$

avec les conditions initiales choisies.

EXERCICE 17.15. On s'intéresse à la résolution explicite du problème de l'oscillateur harmonique de type masse-ressort. On considère une masse m et un ressort de raideur k . On peut alors montrer à l'aide de considérations simples physiques que le problème est régi par le système différentiel suivant : trouver $x(t)$ solution de

$$m \frac{d^2 x}{dt^2}(t) + kx(t) = 0, \quad t > 0,$$

assujetti aux conditions initiales $x(0) = x_0$ et $\frac{dx}{dt}(0) = v_0$.

(1) Donner la transformée de Laplace de $e^{-at}H(t)$, où H est la fonction de Heaviside, a un réel (éventuellement à préciser).

(2) On pose $X(p) = \mathcal{L}(x(t))(p)$ la transformée de Laplace de x . Montrer que

$$X(p) = \frac{r}{p + i\omega_0} + \frac{\bar{r}}{p - i\omega_0}$$

où $\omega_0 = \sqrt{\frac{k}{m}}$, et r s'écrit $r = R_r e^{i\theta_r}$. Donner les expressions de R_r et θ_r en fonction de x_0 , v_0 et ω_0 .

(3) En déduire, par transformée de Laplace inverse que

$$x(t) = \frac{\sqrt{v_0^2 + \omega_0^2 x_0^2}}{\omega_0} \cos(\omega_0 t - \tan^{-1}(\frac{v_0}{\omega_0 x_0})).$$

EXERCICE 17.16. Soit u la solution de l'équation dite "des télégraphistes"

$$\partial_t^2 u = a \partial_x^2 u + ku,$$

qui satisfait également aux conditions aux limites

$$\begin{aligned} u(0, t) &= f(t) \\ u(+\infty, t) &= 0, \end{aligned}$$

où f est une fonction donnée, et aux conditions initiales

$$u(x, 0) = \partial_t u(x, 0) = 0.$$

Exprimer la solution à partir de la transformée de Laplace.

EXERCICE 17.17. On s'intéresse à l'équation différentielle suivante :

$$\begin{cases} y''(t) + y(t) = \sin t, t > 0 \\ y(0) = y'(0) = 1 \end{cases}$$

Résoudre cette équation. On pourra introduire $Y(p)$ la transformée de Laplace de y .

EXERCICE 17.18. Soit l'équation suivante :

$$t f'(t) + 2 \int_0^t f(u) \sin(t-u) du = 0, \text{ pour } t > 0$$

avec $f(0) = 1$. $f(t)$ sera supposée nulle pour $t < 0$. Résoudre cette équation. On pourra introduire $F(p) = Lf(p)$ la transformée de Laplace de f , pour $p \in \mathbb{R}$.

EXERCICE 17.19. Soit f une fonction de classe \mathcal{C}^1 sur \mathbb{R}^+ . On suppose qu'il existe γ_0 tel que $\int_0^\infty |f(t)| e^{-\gamma_0 t} dt < \infty$. On suppose de plus que $\int_0^\infty |f'(t)| dt < \infty$. On notera F la transformée de Laplace de f .

Pour cet exercice, on rappelle la formule d'interversion de la limite et d'une intégrale : Soient U un ouvert de \mathbb{R} , et k une fonction définie sur $U \times \mathbb{R} : (s, x) \rightarrow k(s, x)$, telle que pour tout $s \in U$, $K(s) = \int_{-\infty}^\infty k(s, x) dx$ existe. Soit s_0 un point de l'adhérence de U .

- Si pour tout $x \in \mathbb{R}$, la limite suivante existe : $\lim_{s \rightarrow s_0} k(s, x) = l(x)$
- et si il existe une fonction m définie sur \mathbb{R} , positive, intégrable sur \mathbb{R} et telle que

$$\forall s \in U, \forall x \in \mathbb{R}, |k(s, x)| \leq m(x),$$

alors on a

$$\lim_{s \rightarrow s_0} K(s) = \int_{-\infty}^{\infty} \lim_{s \rightarrow s_0} k(s, x) dx = \int_{-\infty}^{\infty} l(x) dx.$$

(1) Exprimer la transformée de Laplace de $f' : L(f')(p) = \int_0^{\infty} f'(t)e^{-pt} dt$ en fonction de F et f . Montrer que

$$\lim_{p \rightarrow \infty} pF(p) = f(0)$$

(2) On suppose que $\lim_{t \rightarrow \infty} f(t)$ existe. Montrer que

$$\lim_{p \rightarrow 0} pF(p) = \lim_{t \rightarrow \infty} f(t)$$

9. Table

Toutes les fonctions temporelles considérées ci-dessous sont multipliées par l'échelon unité de Heaviside (elles sont causales).

Fonction	$f(t)$	$F(p)$	Région de convergence
impulsion unité ¹	$\delta(t)$	1	$\forall p \in \mathbb{C}$
décalé idéal ²	$\delta(t - \tau)$	$e^{-\tau p}$	$\forall p \in \mathbb{C}$
échelon unité	$H(t)$	$\frac{1}{p}$	$\Re(p) > 0$
échelon retardé	$H(t - \tau)$	$\frac{e^{-\tau p}}{p}$	$\Re(p) > 0$
rampe	t	$\frac{1}{p^2}$	$\Re(p) > 0$
puissance n ème	$\frac{t^n}{n!}$	$\frac{1}{p^{n+1}}$	$\Re(p) > 0$
idem avec décalage fréquentiel	$\frac{t^n}{n!} e^{-\alpha t}$	$\frac{1}{(p+\alpha)^{n+1}}$	$\Re(p + \alpha) > 0$
idem avec retard temporel	$\frac{(t-\tau)^n}{n!} e^{-\alpha(t-\tau)}$	$\frac{e^{-\tau p}}{(p+\alpha)^{n+1}}$	$\Re(p + \alpha) > 0$
puissance q ème ($q > -1$)	$\frac{t^q}{\Gamma(q+1)}$	$\frac{1}{p^{q+1}}$	$\Re(p) > 0$
décroissance exponentielle	$e^{-\alpha t}$	$\frac{1}{p+\alpha}$	$\Re(p + \alpha) > 0$
approche exponentielle	$1 - e^{-\alpha t}$	$\frac{\alpha}{p(p+\alpha)}$	$\Re(p + \alpha) > 0$
sinus	$\sin(\omega t)$	$\frac{\omega}{p^2 + \omega^2}$	$\Re(p) > 0$
cosinus	$\cos(\omega t)$	$\frac{p}{p^2 + \omega^2}$	$\Re(p) > 0$
sinus hyperbolique	$\sinh(\alpha t)$	$\frac{\alpha}{p^2 - \alpha^2}$	$\Re(p) > \alpha $
cosinus hyperbolique	$\cosh(\alpha t)$	$\frac{p}{p^2 - \alpha^2}$	$\Re(p) > \alpha $
décr. expo. d'une onde sinusoïdale	$e^{-\alpha t} \sin(\omega t)$	$\frac{\omega}{(p+\alpha)^2 + \omega^2}$	$\Re(p + \alpha) > 0$
décr. expo. d'une onde cosinusoïdale	$e^{-\alpha t} \cos(\omega t)$	$\frac{p+\alpha}{(p+\alpha)^2 + \omega^2}$	$\Re(p + \alpha) > 0$

La transformée de Fourier

L'analyse de Fourier est une analyse de type fréquentiel, étendue à des signaux qui ne sont pas forcément sinusoïdaux. Elle permet de représenter certaines fonctions d'une variable réelle comme 'superposition d'harmoniques', i.e. de fonctions de la forme $t \mapsto e^{2i\pi ut}$, pour $u \in \mathbb{R}$ fixé. Cette représentation est obtenue sous la forme de séries dans le cas de fonctions périodiques (séries de Fourier), ou sous la forme d'une intégrale pour des fonctions qui ne sont pas périodiques.

L'analyse de Fourier est un moyen d'analyser les grandeurs physiques variant dans le temps (par exemple des signaux accoustiques), c'est aussi un outil pour synthétiser ces grandeurs. Les signaux apparaissent ainsi sous deux formes de représentations suivant qu'on les considère dans l'espace temporel, ou dans l'espace des fréquences. La transformée de Fourier permet de passer de l'une à l'autre de ces représentations.

Avant de commencer, précisons qu'il existe différentes conventions de notations sur les transformées de Fourier qui peuvent modifier la forme des résultats que nous énonçons ici.

1. Transformée de Fourier d'une fonction L^1

La transformée de Fourier est en quelque sorte une généralisation de la notion de série de Fourier à des fonctions définies sur \mathbb{R} et non périodiques. Nous allons commencer par regarder les fonctions de L^1 , puis celles de L^2 .

1.1. Définition. Nous avons la définition suivante.

DÉFINITION 18.1. Soit h , une fonction d'une variable réelle. On appelle **transformée de Fourier** de h , si elle existe, la fonction complexe de la variable f (réelle)

$$\mathcal{F}(h)[\nu] = \widehat{h}(\nu) = \int_{\mathbb{R}} h(x) e^{-2i\pi\nu x} dx, \quad \forall \nu \in \mathbb{R}.$$

REMARQUE 18.1. On notera indifféremment la transformée de Fourier de h par $\mathcal{F}(h)$ ou \widehat{h} .

ν est souvent appelé '**fréquence**'. La transformée de Fourier d'une fonction d'une variable temporelle est une fonction d'une variable fréquentielle.

Cette transformation est linéaire, et $\widehat{h}(0) = \int_{\mathbb{R}} h(t) dt$.

1.2. Existence de la transformée de Fourier. L'intégrale n'existe pas toujours : \widehat{h} peut ne pas exister ou $\widehat{h}(\nu)$ peut ne pas être définie pour certaines valeurs de ν . Si nous cherchons des conditions d'existence de la transformée de Fourier, on peut s'apercevoir que cela est difficile de manière générale. Toutefois, il est clair que \widehat{h} existe si $h \in L^1(\mathbb{R})$. En effet, si $h : \mathbb{R} \rightarrow \mathbb{R}$, on a l'équivalence

$$h \text{ est intégrable} \iff |h| \text{ est intégrable.} \quad (1.1)$$

Ainsi, si h est intégrable, cela implique que la fonction :

$$x \rightarrow h(x) e^{-2i\pi x\nu}$$

est aussi intégrable, $\forall f$. On a alors le théorème suivant.

THÉORÈME 18.1. *Soit $h \in L^1(\mathbb{R})$. Alors h possède une transformée de Fourier : $\nu \rightarrow \widehat{h}(\nu)$.*

EXEMPLE 18.1. Montrer que si on introduit la fonction "porte"

$$\Pi(x) = \begin{cases} 0 & \text{si } |x| > 1/2 \\ 1 & \text{si } |x| < 1/2 \end{cases}$$

alors

$$\widehat{\Pi}(\nu) = \begin{cases} \frac{\sin(\pi\nu)}{\pi\nu} & \text{si } \nu \neq 0 \\ 1 & \text{si } \nu = 0 \end{cases}$$

Remarquez que $\Pi \in L^1(\mathbb{R})$ mais $\widehat{\Pi} \notin L^1(\mathbb{R})$.

REMARQUE 18.2. — L'appartenance à $L^1(\mathbb{R})$ est une condition suffisante pour l'existence de la transformée de Fourier, mais ce n'est pas une condition nécessaire.

Par exemple, la fonction $\frac{\sin x}{x}$ n'est pas dans $L^1(\mathbb{R})$ mais on peut calculer $\int_{-\infty}^{\infty} \frac{\sin x}{x} e^{-2i\pi x\nu} dx$.

— L'exemple précédent montre que la transformée de Fourier d'une fonction de $L^1(\mathbb{R})$ n'est, elle-même pas obligatoirement dans $L^1(\mathbb{R})$. Il s'ensuit que la transformée de Fourier ne laisse pas stable l'espace $L^1(\mathbb{R})$. Nous verrons plus loin que ce n'est pas le cas de l'espace L^2 qui est stable par la transformée de Fourier, propriété essentielle.

1.3. Propriétés générales.

THÉORÈME 18.2. *Soit $h \in L^1(\mathbb{R})$ et \widehat{h} sa transformée de Fourier. Alors, \widehat{h} est une fonction continue sur \mathbb{R} et bornée :*

$$\|\widehat{h}\|_{\infty} = \sup_{x \in \mathbb{R}} |\widehat{h}(x)| \leq \|h\|_1,$$

DÉMONSTRATION. Si $h \in L^1(\mathbb{R})$, \widehat{h} est clairement bien définie, et $|\widehat{h}(\nu)| \leq \int_{\mathbb{R}} |h(t)| dt = \|h\|_1$. La continuité de la fonction $\nu \rightarrow \widehat{h}(\nu)$ résulte du théorème de continuité des intégrales à paramètres. \square

On a aussi le lemme de **Riemann-Lebesgue** :

THÉORÈME 18.3. *Soit $h \in L^1(\mathbb{R})$ et \widehat{h} sa transformée de Fourier. Alors,*

$$\lim_{\nu \rightarrow \pm\infty} |\widehat{h}(\nu)| = 0.$$

DÉMONSTRATION. On peut trouver une démonstration de ce théorème à la [Remarque 18.7](#). \square

Le résultat suivant précise la linéarité et la continuité de la transformation $\widehat{\cdot}$

THÉORÈME 18.4. *La transformée de Fourier est une opération linéaire et continue de $L^1(\mathbb{R})$ dans $L^\infty(\mathbb{R})$ (ensemble des fonctions bornées sur \mathbb{R}), c'est-à-dire*

i) $\forall h, g \in L^1(\mathbb{R}), \forall \lambda, \mu \in \mathbb{R}$, on a

$$\widehat{\lambda h + \mu g} = \lambda \widehat{h} + \mu \widehat{g} \quad (1.2)$$

ii) si la suite $(h_n)_{n \in \mathbb{N}}$ vérifie

$$h_n \xrightarrow[n \rightarrow \infty]{L^1} h \in L^1(\mathbb{R}) \quad (1.3)$$

alors

$$\widehat{h_n} \xrightarrow[n \rightarrow \infty]{} \widehat{h} \text{ uniformément} \quad (1.4)$$

La propriété de linéarité implique que la transformée de Fourier est bien adaptée à la résolution d'équations linéaires. Mais attention, on ne pourra trouver par ce moyen que des solutions admettant une transformée de Fourier... Si l'équation à résoudre a une solution n'admettant pas de transformée de Fourier, on ne la trouvera pas... à moins d'effectuer des calculs formels et de vérifier que le résultat trouvé par cette méthode est bien solution de l'équation.

Par exemple, l'équation différentielle linéaire suivante $y' - ky = 0$ a une solution triviale $y(t) = y(0)e^{kt}$. Or, cette solution n'admet pas de transformée de Fourier! Et pourtant, si on cherche la solution par le biais des transformées de Fourier, on écrirait $\widehat{y}'(\nu) - k\widehat{y}(\nu) = 0$, ce qui équivaut formellement à (on le verra plus loin) $(2i\pi f - k)\widehat{y}(\nu) = 0$, ce qui implique que $\widehat{y}(\nu) = 0$ partout sauf en $\nu = k/2i\pi$. On pourrait donc écrire, formellement, $\widehat{y}(\nu) = c\delta_{k/2i\pi}(\nu)$ où δ est le pic de Dirac, et c une constante. Il suffirait ensuite d'inverser formellement \widehat{y} pour en déduire y , et vérifier, finalement, que la solution trouvée par ce moyen, formel et pas du tout rigoureux, est bien correcte!

1.4. Inversion de la transformée de Fourier. La question est maintenant la suivante : peut-on inverser la transformée de Fourier? Nous venons de voir que $h \in L^1$ n'implique pas que $\widehat{h} \in L^1$. Dans le cadre actuel de nos espaces, nous ne pouvons donc pas prendre une "transformée" de Fourier de \widehat{h} . Toutefois, nous avons le théorème suivant.

THÉORÈME 18.5. (*théorème d'inversion*) Soit $h \in L^1(\mathbb{R})$, une fonction telle que $\widehat{h} \in L^1$. Alors on peut calculer $\widehat{\widehat{h}}$, et on a, pour presque tout t (pour tout t si h est continue) :

$$h(t) = \int_{\mathbb{R}} \widehat{h}(\nu) e^{2i\pi\nu t} d\nu = \widehat{[\widehat{h}(\nu)]}(-t).$$

REMARQUE 18.3. (1) On exprime ici le 'signal' $h(t)$ comme une superposition de signaux élémentaires $e_{\nu}(t) = e^{2i\pi\nu t}$ de fréquence $|\nu|$, de période $\frac{1}{|\nu|}$. $\widehat{h}(\nu)$ est appelée **densité spectrale**.

(2) $\int_{\mathbb{R}} \widehat{h}(\nu) e^{2i\pi\nu t} d\nu$ est une fonction continue (preuve par le théorème de continuité sous le signe intégral). On obtient ainsi une **version continue** de h .

(3) Le théorème d'inversion est valide quand $\widehat{h} \in L^1$. Faisons une petite analogie avec les séries de Fourier : si $\sum_{n \in \mathbb{Z}} |c_n| < \infty$ alors on sait que $\sum_{n \in \mathbb{Z}} c_n e^{2i\pi n t}$ converge vers une fonction continue, périodique dont le coefficient de Fourier $c_n(\nu) = c_n$. Autrement dit, lorsque la suite c_n est sommable, on peut retrouver f à partir de c_n .

DÉMONSTRATION. On démontre le théorème dans le cas simple où $h \in L^1(\mathbb{R})$ et h est bornée. \square

Un résultat pratique est le suivant.

PROPOSITION 18.1. Si h est de classe \mathcal{C}^2 , et si h , h' et h'' sont toutes intégrables, alors \widehat{h} l'est également.

DÉMONSTRATION. La démonstration de ce théorème est donnée à la [Remarque 18.7](#). \square

Enfin, la formule d'inversion nous donne une propriété importante.

THÉORÈME 18.6. — Soit $h \in L^1$ telle que $\widehat{h} = 0$. Alors $h = 0$ presque partout.
— Soient $g, h \in L^1$ telles que $\widehat{g} = \widehat{h}$. Alors $g = h$ presque partout.

DÉMONSTRATION. D'après la formule d'inversion, on a pour presque tout t : $h(t) = \widehat{\widehat{h}}(-t) = 0$. De même, $g - h$ vérifie les hypothèses de la formule d'inversion, et donc pour presque tout t on a : $(g - h)(t) = \int_{\mathbb{R}} (\widehat{g - h})(\nu) e^{2i\pi\nu t} d\nu = 0$. \square

REMARQUE 18.4. La formule d'inversion permet de déduire d'autres transformées de Fourier. On va pouvoir lire des tables de transformées de Fourier de gauche à droite et de droite à gauche, en faisant toutefois bien attention au signe qui

1.5. Extension de la formule d'inversion. Discutons le cas maintenant où h admet une transformée de Fourier \widehat{h} qui n'est pas intégrable. Dans certains cas

$$\lim_{R \rightarrow +\infty} \int_{-R}^{+R} \widehat{h}(\nu) e^{2i\pi\nu t} d\nu$$

peut être définie (ce n'est pas à proprement parler une intégrale impropre car les bornes sont fixées de manière symétrique). On a alors les théorèmes suivants.

THÉORÈME 18.7. *Si h est continue sur \mathbb{R} et $h \in L^1(\mathbb{R})$, on a*

$$\lim_{R \rightarrow +\infty} \int_{-R}^{+R} \widehat{h}(\nu) e^{2i\pi\nu t} d\nu = h(t), \quad \forall t \in \mathbb{R}$$

et

THÉORÈME 18.8. *Soit h une fonction bornée et $h \in L^1(\mathbb{R})$ et de classe \mathcal{C}^1 par morceaux sur \mathbb{R} , alors*

$$\lim_{R \rightarrow +\infty} \int_{-R}^{+R} \widehat{h}(\nu) e^{2i\pi\nu t} d\nu = \frac{1}{2} [h(t^+) + h(t^-)], \quad \forall t \in \mathbb{R}$$

Par abus, on notera encore $\widehat{[\widehat{h}(\nu)]}(-t)$ cette limite. En d'autres termes, la formule d'inversion, prise au sens d'une limite d'intégrale (appelée valeur principale), nous donne la régularisée de h .

DÉMONSTRATION. cf [13] p. 139. □

2. Propriétés de la transformée de Fourier

Intéressons nous maintenant à des propriétés calculatoires de la transformée de Fourier.

Si h est une fonction telle que sa transformée de Fourier existe, alors nous avons les propriétés suivantes :

2.0.1. *Conjuguée.*

$$\widehat{\widehat{h}}(\nu) = \overline{\widehat{h}(-\nu)}$$

2.0.2. *Symétrie temporelle.*

$$\widehat{\widehat{h(-t)}}(\nu) = \widehat{\widehat{h(t)}}(-\nu)$$

Des propriétés de symétrie et de conjugaison, on déduit :

2.0.3. *Parité.* Soit $h \in L^1(\mathbb{R})$ à valeurs réelles. h et \widehat{h} ont même parité

2.0.4. *Parties réelle et imaginaire.*

$$\begin{aligned} h \text{ réelle} &\iff \widehat{h} \text{ hermitienne} \\ h \text{ imaginaire} &\iff \widehat{h} \text{ anti-hermitienne} \end{aligned} \quad (2.1)$$

où une fonction hermitienne est telle que $g(-x) = \overline{g(x)}$ et anti-hermitienne $g(-x) = -\overline{g(x)}$

2.0.5. *Translation temporelle.* Soit $a \in \mathbb{R}$, alors on a la relation

$$\widehat{h(t-a)}(\nu) = e^{-2i\pi\nu a} \widehat{h}(t)(\nu)$$

A la translation de $h(t)$ correspond un déphasage de $\widehat{h}(\nu)$ proportionnel à ν .

REMARQUE 18.5. Si h est périodique, $h(t) = h(t-a)$, alors $(e^{-2i\pi\nu a} - 1)\widehat{h}(\nu) = 0$ ce qui implique que \widehat{h} est nulle partout sauf aux points $\nu_n = \frac{n}{a}$. L'intégrale de Fourier dégénère en série de Fourier.

2.0.6. *Translation fréquentielle / Modulation de fréquence.* Soit $a \in \mathbb{R}$, alors on a la relation

$$\widehat{h}(\nu - \nu_0) = \widehat{h(t)e^{2i\pi\nu_0 t}}(\nu)$$

A la modulation de $h(t)$ (i.e. la multiplication de $h(t)$ par un facteur oscillatoire $e^{2i\pi\nu_0 t}$) correspond la translation de $\widehat{h}(\nu)$.

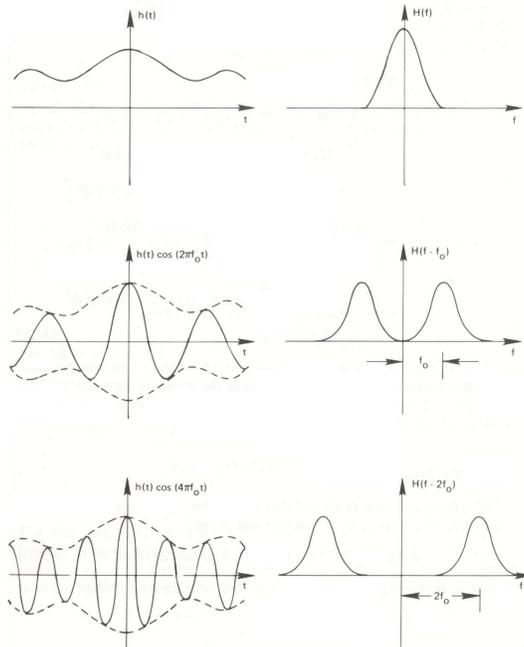


FIGURE 1 – Propriété de translation fréquentielle avec \widehat{h} une fonction réelle. Ce phénomène est appelé **modulation de fréquence**.

2.0.7. *Changement d'échelle.* Soit $a \in \mathbb{R}^*$, alors on peut montrer que l'on a la relation

$$\widehat{h(at)}(\nu) = \frac{1}{|a|} \widehat{h(t)}\left(\frac{\nu}{a}\right).$$

De même,

$$\widehat{h(t)}(a\nu) = \frac{1}{|a|} \widehat{h\left(\frac{t}{a}\right)}(\nu).$$

A la dilatation de $h(t)$ correspond une compression et une augmentation des amplitudes de $\widehat{h}(\nu)$. De même, à une dilatation de $\widehat{h}(\nu)$ correspond une compression et une augmentation de l'amplitude de $h(t)$.

Remarquons que dans [figure 3](#), l'aire du rectangle reste constante. De même, dans [figure 2](#), si $\widehat{h} \in L^1$, l'intégrale de la fonction fréquentielle reste constante. Ce phénomène est bien connu dans la théorie des radars et des antennes.

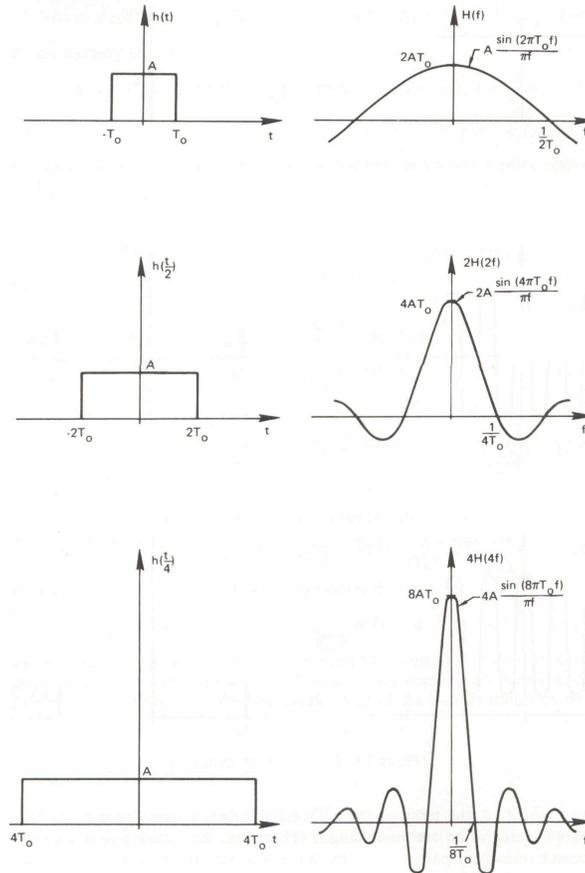


FIGURE 2 – Propriété de changement d'échelle. \widehat{h} est noté H .

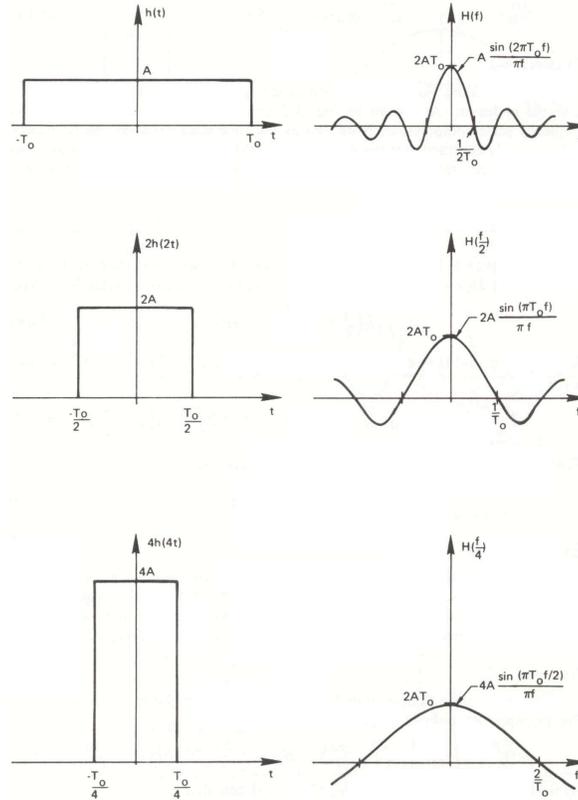


FIGURE 3 – Propriété de changement d'échelle (suite). \widehat{h} est noté H .

2.0.8. *Dérivation.* Plus $h(t)$ décroît rapidement, plus $\widehat{h}(\nu)$ est dérivable et à dérivées bornées.

THÉORÈME 18.9. Soit $h \in L^1$, une fonction décroissant suffisamment vite pour que $t \mapsto t^p h(t)$

soit également dans L^1 . Alors, \widehat{h} est de classe \mathcal{C}^p , et les dérivées de \widehat{h} jusqu'à l'ordre p sont bornées. De plus,

$$\widehat{h}^{(p)}(\nu) = (-2i\pi)^p \int_{\mathbb{R}} t^p h(t) e^{-2i\pi\nu t} dt = [(-2i\pi t)^p \widehat{h}(t)](\nu). \quad (2.2)$$

REMARQUE 18.6. (1) On a donc les dérivées successives de \widehat{h} en 'dérivant' sous le signe intégral.

(2) Sous les hypothèses du [Théorème 18.9](#), les dérivées de \widehat{h} sont bornées :

$$\left| \widehat{h}^{(p)}(\nu) \right| \leq \int_{\mathbb{R}} |2\pi t|^p |h(t)| dt = (2\pi)^p \|t^p h(t)\|_1.$$

- (3) On démontre le [Théorème 18.9](#) par récurrence en utilisant le théorème de dérivation sous le signe intégral.

THÉORÈME 18.10. *Inversement, si $h \in L^1$, si h est de classe \mathcal{C}^p , et si de plus, les dérivées successives $h^{(k)}$ sont intégrables pour $k = 1, \dots, p$, alors on a*

$$\widehat{h^{(k)}}(\nu) = (2i\pi\nu)^k \widehat{h}(\nu), k = 1, \dots, p. \quad (2.3)$$

REMARQUE 18.7. (1) A la dérivation de $h(t)$ par rapport à t correspond la multiplication de $\widehat{h}(\nu)$ par $2i\pi\nu$.

- (2) Plus une fonction est dérivable et à dérivées intégrables, plus sa transformée de Fourier décroît vite vers 0 :

$$|\widehat{h}(\nu)| \leq \|\widehat{h^{(p)}}\|_1 |2\pi\nu|^{-p} = \|h^{(p)}\|_\infty |2\pi\nu|^{-p}, \quad \forall \nu \neq 0,$$

et nous pouvons écrire $\widehat{h}(\nu) = \mathcal{O}(|\nu|^{-p})$. (On retrouve ici, le même résultat que Riemann-Lebesgue : $\lim_{|f| \rightarrow \infty} |\widehat{h}(\nu)| = 0$.)

- (3) Donnons ici une démonstration de la [Proposition 18.1](#). Vérifions que \widehat{h} est intégrable. D'après le [Théorème 18.2](#), la fonction \widehat{h} est continue sur \mathbb{R} . D'après le [Théorème 18.10](#), nous avons $\widehat{h}(\nu) = \mathcal{O}(\frac{1}{\nu^2})$ pour ν grand. Elle est donc intégrable sur \mathbb{R} . \square

Sous les hypothèses de [Théorème 18.9](#) et [Théorème 18.10](#), on a

$$\widehat{h'}(\nu) = 2i\pi\nu \widehat{h}(\nu) \quad (2.4)$$

et

$$[-2i\pi t \widehat{h}(t)](\nu) = \frac{d\widehat{h}(\nu)}{d\nu} \quad (2.5)$$

Introduisons maintenant le support d'une fonction h défini par $\text{Supp } h = \{t \in \mathbb{R} / h(t) \neq 0\}$. On a alors la proposition suivante.

PROPOSITION 18.2. *Si h est L^1 à support borné, alors \widehat{h} est une fonction de classe \mathcal{C}^∞ .*

Ces propriétés de dérivation jouent un rôle essentiel pour la résolution des équations différentielles linéaires. Une équation différentielle linéaire se transforme par Fourier en une équation algébrique, ce qui simplifie la résolution. Voir par exemple les exercices [18.5](#), [18.6](#) et [18.7](#).

Pour toute fonction h définie sur \mathbb{R} , notons Mh la fonction $t \rightarrow (Mh)(t) = th(t)$. Si h est dérivable, notons Dh sa dérivée. Si h est intégrable, notons $\mathcal{F}h$ sa transformée de Fourier. D est donc l'opérateur de dérivation, M celui de la multiplication par t , et \mathcal{F} l'opérateur qui à h fait correspondre sa transformée de Fourier.

Les résultats précédents montrent que la transformée de Fourier échange la dérivation et la multiplication par t et conduisent ainsi aux relations formelles entre les opérateurs \mathcal{F} , D et M :

$$D\mathcal{F} = -2i\pi\mathcal{F}M, \quad (2.6)$$

$$\mathcal{F}D = 2i\pi M\mathcal{F}. \quad (2.7)$$

Plus généralement, on a formellement :

$$D^k\mathcal{F} = (-2i\pi)^k\mathcal{F}M^k, \quad (2.8)$$

$$\mathcal{F}D^k = (2i\pi)^k M^k\mathcal{F}. \quad (2.9)$$

Pour pouvoir appliquer ces formules à une fonction h , nous devons bien sûr vérifier que les hypothèses des propositions précédentes sont satisfaites.

3. Transformée de Fourier d'une fonction L^2

Nous avons défini la transformée de Fourier pour des fonctions appartenant à $L^1(\mathbb{R})$. Nous avons vu qu'alors la transformée de Fourier d'une fonction $h \in L^1$ n'est pas nécessairement dans L^1 , et donc que la formule d'inversion n'est pas nécessairement valide.

Nous travaillons maintenant sur $L^2(\mathbb{R}) = \{f : \mathbb{R} \mapsto \mathbb{R}, \text{ telle que } \int_{\mathbb{R}} |f(t)|^2 dt < \infty\}$.

Faisons quelques remarques sur $L^2(\mathbb{R})$.

- Il n'y a pas de relation d'inclusion entre $L^1(\mathbb{R})$ et $L^2(\mathbb{R})$. Il existe des fonctions $h \in L^2(\mathbb{R})$ mais telles que $h \notin L^1(\mathbb{R})$: par exemple, $h(t) = \frac{1}{t}\mathbf{1}_{|t|>1}(t) \in L^2(\mathbb{R})$ mais $h \notin L^1(\mathbb{R})$; ou $h(t) = \frac{1}{\sqrt{t}}\mathbf{1}_{]0,1]}(t) \in L^1(\mathbb{R})$ mais $h \notin L^2(\mathbb{R})$.
- L'espace L^2 est très important en physique, car on calcule souvent l'énergie comme l'intégrale du carré d'une fonction. Il est donc important que l'outil des transformées de Fourier fonctionne aussi pour les fonctions de L^2 .

3.1. La transformée de Fourier et l'espace de Schwartz. Avant d'avoir des résultats sur L^2 , définissons l'espace de Schwartz $\mathcal{S} \subset L^2(\mathbb{R})$, sur lequel on aura

$$\widehat{\cdot} : \mathcal{S} \rightarrow \mathcal{S} \quad (3.1)$$

$$h \mapsto \widehat{h} \quad (3.2)$$

qui sera une bijection, facile à inverser. Pour cela, il faudra que la formule d'inversion soit encore valide, et donc que les fonctions de cet espace soient à la fois intégrables et de transformée de Fourier intégrables.

La bonne classe de fonctions cherchée est celle des fonctions qui sont d'une part très régulières, d'autre part qui décroissent très vite vers 0. Cette dernière notion est précisée dans la définition suivante.

DÉFINITION 18.2. On appelle **fonction à décroissance rapide** une fonction h vérifiant

$$\lim_{|t| \rightarrow +\infty} |t^k h(t)| = 0, \quad \forall k \in \mathbb{N} \quad (3.3)$$

Ainsi, h décroît plus vite que toute puissance de t . On a alors, le théorème suivant.

THÉORÈME 18.11. Si h est une fonction intégrable à décroissance rapide, alors \widehat{h} est de classe \mathcal{C}^∞ , si h est une fonction intégrable \mathcal{C}^∞ , et si $\forall k \in \mathbb{N}$, $h^{(k)}$ est intégrable, alors \widehat{h} est à décroissance rapide.

DÉMONSTRATION. (a) Soit $M \in \mathbb{R}^+$, et $k \in \mathbb{N}$. Pour $t \in [-M, M]$, on a $|t^k h(t)| \leq M^k |h(t)|$, pour $|t| > M$, on a $|t^k h(t)| \leq \frac{\tilde{M}}{|t|^2}$ car $|t^{k+2} h(t)| \rightarrow 0$, donc est borné, par $\tilde{M} \in \mathbb{R}^+$. Ainsi, la fonction $t^k h(t)$ est intégrable, pour tout $k \in \mathbb{N}$, et donc par le [Théorème 18.9](#), \widehat{h} est de classe \mathcal{C}^∞ .

(b) Soit $k \in \mathbb{N}$, grâce au [Théorème 18.10](#), on a $|(2i\pi\nu)^k \widehat{h}(\nu)| = |\widehat{h^{(k)}}(\nu)| \rightarrow 0$ pour $|\nu| \rightarrow \infty$ par le théorème de Riemann-Lebesgue. \square

Nous pouvons maintenant définir l'espace que nous cherchions.

DÉFINITION 18.3. On appelle **espace de Schwartz**, et on le note \mathcal{S} , l'espace des fonctions qui sont \mathcal{C}^∞ à décroissance rapide, ainsi que toutes leurs dérivées.

REMARQUE 18.8. (a) Toute fonction \mathcal{C}^∞ à support borné est dans \mathcal{S} .

(b) La fonction $h(t) = e^{-t^2}$ est dans \mathcal{S} .

(c) $\mathcal{S} \subset L^1(\mathbb{R}) \cap L^2(\mathbb{R})$

On a alors le résultat de stabilité important suivant.

THÉORÈME 18.12. (de **Parseval-Plancherel**). La transformation de Fourier $\widehat{\cdot}$ définit une isométrie bijective de \mathcal{S} sur lui-même : si h et g sont deux fonctions de \mathcal{S} , on a

$$\int_{\mathbb{R}} h(t) \overline{g(t)} dt = \int_{\mathbb{R}} \widehat{h}(\nu) \overline{\widehat{g}(\nu)} d\nu \quad (3.4)$$

En particulier, en prenant $h = g$, nous avons

$$\|h\|_2 = \|\widehat{h}\|_2 \quad \textbf{Formule de Plancherel-Parseval} \quad (3.5)$$

De plus, toute fonction $h \in \mathcal{S}$ vérifie la formule d'inversion de la transformée de Fourier :

$$h(t) = \int_{\mathbb{R}} \widehat{h}(\nu) e^{2i\pi\nu t} d\nu, \quad \forall t \in \mathbb{R} \quad (3.6)$$

Pour toute fonction $h \in \mathcal{S}$, on a :

$$\frac{d^p}{df^p} \widehat{h}(\nu) = (-2i\pi)^p \int_{\mathbb{R}} t^p h(t) e^{-2i\pi f t} dt, \quad \forall p \geq 0$$

$$\widehat{h^{(p)}}(\nu) = (2i\pi\nu)^p \widehat{h}(\nu), \quad \forall p \geq 0$$

DÉMONSTRATION. Puisque $h \in \mathcal{S}$, les hypothèses de [Théorème 18.9](#) et [Théorème 18.10](#) sont vérifiées, ce qui justifie les formules sur la dérivation à l'ordre p de la transformée de Fourier. Ces formules montrent aussi la relation

$$\nu^n \frac{d^p}{d\nu^p} \widehat{h}(\nu) = (-2i\pi)^p \nu^n \widehat{g}_p(\nu) = (-2i\pi)^p \frac{1}{(2i\pi)^n} \widehat{g_p^{(n)}}(\nu)$$

avec $g_p(t) = t^p h(t)$. Pour tout $n \geq 1, p \geq 1$, la fonction $u^n \frac{d^p}{du^p} \widehat{h}(\nu)$ est donc la transformée de Fourier d'une fonction intégrable, ce qui implique qu'elle est bornée (cf [Théorème 18.2](#)). Nous en déduisons $\widehat{h} \in \mathcal{S}$. De plus, on a $\widehat{h} \in L^1(\mathbb{R})$, et la formule d'inversion est valide. On en déduit que la transformée de Fourier est une bijection de \mathcal{S} sur \mathcal{S} . Enfin, il reste à montrer la formule de Plancherel-Parseval. Soient $g, h \in \mathcal{S}$. Comme g et h sont de carré intégrables, et que la formule d'inversion est valide, on utilise la formule d'inversion et le théorème de Fubini :

$$\begin{aligned} \int_{\mathbb{R}} h(t) \overline{g(t)} dt &= \int_{\mathbb{R}} h(t) \left(\int_{\mathbb{R}} \overline{\widehat{g}(\nu)} e^{-2i\pi\nu t} d\nu \right) dt \\ &= \int_{\mathbb{R}} \overline{\widehat{g}(f)} \left(\int_{\mathbb{R}} h(t) e^{-2i\pi f t} dt \right) df \\ &= \int_{\mathbb{R}} \widehat{h}(\nu) \overline{\widehat{g}(\nu)} d\nu \end{aligned}$$

□

REMARQUE 18.9. En physique, si $h(t)$ représente une onde ou une vibration quelconque, et si $\widehat{h}(\nu)$ est sa transformée de Fourier, la formule de Plancherel-Parseval dit que l'énergie totale de l'onde correspond à la somme des énergies de toutes les vibrations harmoniques.

3.2. Transformée de Fourier dans L^2 . Nous prolongeons à $L^2(\mathbb{R})$ la transformée de Fourier définie sur \mathcal{S} . On peut car $\mathcal{S} \subset L^2(\mathbb{R})$, et surtout car :

LEMME 18.1. *L'espace \mathcal{S} est un sous-espace vectoriel dense de $L^2(\mathbb{R})$.*

Ce qui veut dire qu'on peut approcher toute fonction $h \in L^2(\mathbb{R})$, en norme $\|\cdot\|_2$, par une fonction $\varphi \in \mathcal{S}$.

On utilise ensuite le fait que

- \mathcal{S} est dense dans L^2 ,
- L^2 est complet,

pour montrer que la transformée de Fourier sur \mathcal{S} se **prolonge** en un opérateur de L^2 dans L^2 qui reste linéaire et continu. (Plus précisément, soit $h \in L^2(\mathbb{R})$, dont on veut sa transformée de Fourier. On sait que $h = \lim_n h_n$ où (h_n) est une suite de fonctions de \mathcal{S} . On définit alors \widehat{h} par $\lim_n \widehat{h}_n$. On sait que cette limite existe et qu'elle est dans

L^2 (car L^2 est complet et on peut montrer que \widehat{h}_n est une suite de Cauchy). De plus, comme $\|h_n\|_2 = \|\widehat{h}_n\|_2$ pour tout n , on a aussi $\|h\|_2 = \|\widehat{h}\|_2$.)

En conclusion, on se retrouve après ces étapes, avec un opérateur "transformation de Fourier" défini pour toute fonction de $L^2(\mathbb{R})$, à valeur dans $L^2(\mathbb{R})$.

- (1) Les formules (3.4) et (3.5) de **Plancherel-Parseval** sont encore vérifiées. En terme d'énergie, la formule (3.5) s'interprète comme l'égalité entre l'énergie d'un signal calculée dans l'espace temporel et son énergie calculée dans l'espace des fréquences.
- (2) La formule d'inversion (3.6) est encore valable, mais seulement 'presque-partout', i.e. pour "presque tout $t \in \mathbb{R}$ ".
- (3) Pour ce qui concerne les formules de dérivations :

THÉORÈME 18.13. *Soit $h \in L^2(\mathbb{R})$.*

Si h est continue, de classe C^1 par morceaux et telle que $h' \in L^2(\mathbb{R})$, alors pour presque tout $\nu \in \mathbb{R}$:

$$\widehat{h'}(\nu) = 2i\pi\nu\widehat{h}(\nu).$$

Si de plus, h est de classe C^m par morceaux, ou $m \in \mathbb{N}^$ et telle que les dérivées $h^{(k)}$ jusqu'à l'ordre m inclus sont de carré intégrables, alors pour presque tout $\nu \in \mathbb{R}$ et pour tout $1 \leq k \leq m$, on a*

$$\widehat{h^{(k)}}(\nu) = (2i\pi\nu)^k \widehat{h}(\nu).$$

REMARQUE 18.10. Le **Théorème 18.9** n'a pas d'analogue dans le cas $L^2(\mathbb{R})$. En effet, si les fonctions h et $x \mapsto th(t)$ appartiennent à $L^2(\mathbb{R})$, il n'est pas toujours vrai que \widehat{h} soit de classe C^1 .

La **Définition 18.1** de la transformation de Fourier sur L^1 , et la définition de la transformation de Fourier sur L^2 par prolongement sont *a priori* différentes. Toutefois, nous avons le résultat réconfortant suivant.

PROPOSITION 18.3. *La transformée de Fourier sur L^1 et celle sur L^2 coïncident sur $L^1 \cap L^2$.*

DÉMONSTRATION. cf [13] p. 160. □

Par conséquent, si h est une fonction intégrable et de carré intégrable, alors sa transformée de Fourier est

$$\widehat{h} : f \mapsto \int_{\mathbb{R}} h(t) e^{-2i\pi ft} dt \quad (3.7)$$

Mais, si ce n'est pas le cas, si h est de carré intégrable mais non intégrable, que se passe-t-il? C'est la cas par exemple pour $h(t) = t^{-1}\mathbf{1}_{[1,+\infty[}(t)$ ou $h(t) = t^{-1} \sin t$.

On ne peut pas définir la transformée de Fourier directement à l'aide de la formule de [Définition 18.1](#).

On peut montrer que, si $h \in L^2$, on peut obtenir la transformée de Fourier dans L^2 en prenant la limite suivante dans L^2 .

$$\widehat{h}(f) = \lim_{R \rightarrow +\infty} \int_{-R}^R h(t) e^{-2i\pi\nu t} dt, \quad \forall \nu \in \mathbb{R} \quad (3.8)$$

4. Transformée de Fourier et convolution

4.1. Formule de convolution. Rappelons la définition du produit de convolution de deux fonctions.

DÉFINITION 18.4. *Si f et g sont deux fonctions intégrables, leur produit de convolution, lorsqu'il existe est $h = f * g$ avec*

$$h(x) = \int_{\mathbb{R}} f(t) g(x-t) dt \quad (4.1)$$

La transformée de Fourier échange la convolution et la multiplication :

THÉORÈME 18.14. *Soient h et g deux fonctions admettant des transformées de Fourier et telles que leur produit de convolution existe et est intégrable. Alors, nous avons :*

$$\widehat{[h * g]} = \widehat{h} \widehat{g} \quad (4.2)$$

et lorsque ces expressions sont définies

$$\widehat{[hg]} = \widehat{h} * \widehat{g} \quad (4.3)$$

REMARQUE 18.11. — La preuve formelle de ces deux formules est facile et laissée en exercice. Elle utilise le théorème de Fubini. Cependant, pour être valide, l'interversion des intégrales demande quelques hypothèses. On en donne quelques unes dans la section suivante.

- Ce théorème est très important en analyse de Fourier. L'un des intérêts du calcul du produit de convolution par transformées de Fourier est que ces opérations sont moins coûteuses en temps pour un ordinateur que le calcul direct de l'intégrale.
- Le théorème de convolution est très utile pour résoudre des équations linéaires où figure un noyau ω , par exemple $h(x) = \phi(x) + \int_{\mathbb{R}} \omega(x-y)h(y)dy'$. Une telle équation dit que la valeur de h au point x est conditionnée par les valeurs de h en d'autres points y dont l'importance est pondérée par $\omega(x-y)$. Si x est le temps, on parle de système à mémoire. Si x est l'espace, on parle de système avec interactions. Voir exemple l'exercice [18.7](#).
- En traitement du signal, on peut décrire un signal de sortie $S(f)$ d'une boîte noire, via une fonction A convolée avec le signal d'entrée $S(\nu) = A * E(\nu) = \int_{\mathbb{R}} A(\nu - \nu')E(\nu')d\nu'$. La sortie $S(\nu)$ dépend donc des entrées $E(\nu')$ pondérées

par l'aptitude de la boîte noire, quantifiée par $A(\nu - \nu')$. Inversement, dans le domaine temporel, la sortie $s(t)$ peut s'exprimer comme le produit $a(t)e(t)$ où $a(t)$ est la déformation du signal d'entrée $e(t)$. Par exemple, si $A(\nu)$ est une fonction porte jouant le rôle de filtre passe bande, $a(t)$ est un sinus cardinal qui déforme le signal d'entrée.

— La convolution de deux gaussiennes reste une gaussienne (cf. exercice 18.2).

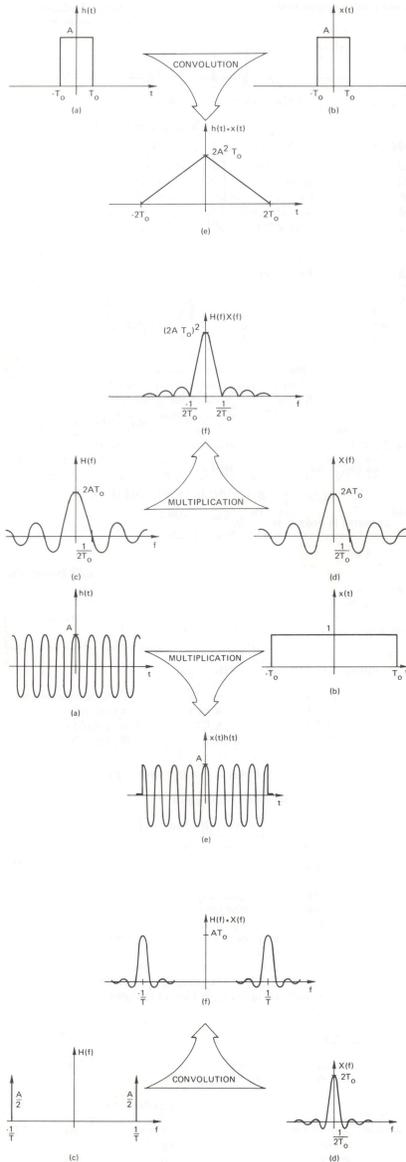


FIGURE 4 – Illustration graphique du théorème de convolution.

4.2. Validité de la formule de convolution. On peut montrer que la formule (4.2) est vraie lorsque :

- i) $h \in L^1, g \in L^1 : [\widehat{h * g}](\nu) = \widehat{h}\widehat{g}(\nu), \quad \forall \nu \in \mathbb{R}.$
- ii) $h \in L^2(\mathbb{R}), g \in L^2(\mathbb{R})$ et $h * g \in L^1(\mathbb{R})$ ou $L^2(\mathbb{R}) : [\widehat{h * g}](\nu) = \widehat{h}\widehat{g}(\nu), \quad$ pour presque tout $\nu \in \mathbb{R}.$
- iii) $h \in L^1(\mathbb{R}), g \in L^2(\mathbb{R}) : [\widehat{h * g}](\nu) = \widehat{h}\widehat{g}(\nu), \quad$ pour presque tout $\nu \in \mathbb{R}.$

On peut montrer que la formule (4.3) est vraie lorsque :

- iv) h et g sont dans L^1 et que leurs transformées de Fourier sont également dans $L^1,$
- v) h et g sont dans $L^2.$

Et dans l'un ou l'autre de ces deux cas, on a :

$$[\widehat{hg}](\nu) = (\widehat{h} * \widehat{g})(\nu), \quad \forall \nu \in \mathbb{R}.$$

5. Exercices

EXERCICE 18.1. On définit la fonction porte Π par

$$\Pi(x) = \begin{cases} 1 & \text{si } x \in [-\frac{1}{2}; \frac{1}{2}], \\ 0 & \text{sinon} \end{cases}$$

- (1) Calculer la transformée de Fourier de $\Pi.$
- (2) En utilisant le résultat précédent, calculer les transformées de Fourier des fonctions
 - i) f définie par

$$f(x) = \Pi\left(\frac{x - 1/2}{a}\right), \quad \text{pour } a > 0,$$

- ii) g définie par

$$g(x) = x\Pi(x),$$

- iii) h définie par

$$h(x) = (\Pi * \Pi)(x).$$

EXERCICE 18.2. (1) On veut calculer la transformée de Fourier de la fonction g telle que : $g(x) = e^{-a^2x^2},$ pour $a \neq 0.$ En utilisant le fait que g est solution d'une équation différentielle, en déduire $\widehat{g}.$

- (2) Soit maintenant $g_a,$ pour $a > 0,$ la gaussienne définie par

$$g_a(x) = \frac{1}{a\sqrt{2\pi}} e^{-\frac{x^2}{2a^2}}$$

- i) Calculer sa transformée de Fourier.
- ii) Calculer le produit de convolution $g_a * g_b.$

EXERCICE 18.3. (1) Calculer la transformée de Fourier de la fonction f donnée par

$$f : x \mapsto f(x) = e^{-|x|}$$

(2) En déduire que

$$\forall \nu \in \mathbb{R}, \quad \mathcal{F}\left(\frac{1}{1+x^2}\right)(\nu) = \pi e^{-2\pi|\nu|}$$

(3) Résoudre l'équation fonctionnelle dans L^1

$$\int_{\mathbb{R}} e^{-a|x-t|} f(t) dt = e^{-x^2}$$

où $a \in \mathbb{R}^{*+}$.

EXERCICE 18.4. On se donne une fonction $\varphi \in L^1$. Résoudre alors l'équation fonctionnelle suivante : trouver $f \in L^1$ satisfaisant l'équation discrète

$$f(x) - k(f(x+1) + f(x-1)) = \varphi(x),$$

où $k \in \mathbb{R}$ tel que $|k| < \frac{1}{2}$. On supposera $\widehat{\varphi} \in L^1$.

EXERCICE 18.5. Trouver, en utilisant la transformée de Fourier, une fonction u de (x, y) solution du système

$$\begin{cases} \Delta u = 0, & x \in \mathbb{R}, y > 0 \\ u(x, 0) = e^{-x^2} \\ \lim_{y \rightarrow +\infty} u(x, y) = 0. \end{cases} \quad (5.1)$$

EXERCICE 18.6. Soient $a \neq 0$ et f une fonction de \mathbb{R} dans \mathbb{R} . On suppose que $f \in \mathcal{C}^1$ et qu'elle vérifie les inégalités d'intégrabilité

$$\|f\|_{L^1} < \infty \quad \|\widehat{f}\|_{L^1} < \infty \quad (5.2)$$

On cherche $u = u(x, t)$ solution de l'équation d'évolution

$$\begin{cases} \partial_t u = a^2 \partial_x^2 u, & x \in \mathbb{R}, t > 0 \\ u(x, 0) = f(x), & x \in \mathbb{R} \end{cases} \quad (5.3)$$

Utiliser pour cela la transformation de Fourier. On traitera ensuite explicitement le cas classique où $f(x) = e^{-x^2}$.

EXERCICE 18.7. Soit $f(t) = e^{-|t|}$ et $g(t) = te^{-t^2}$ deux fonctions définies sur \mathbb{R} .

(1) Calculer la transformée de Fourier de f .

(2) En déduire les valeurs de $\int_{-\infty}^{\infty} \frac{\cos \alpha x}{1+x^2} dx$ et $\int_{-\infty}^{\infty} \frac{\sin \alpha x}{1+x^2} dx$ pour tout $\alpha \in \mathbb{R}$.

(3) On considère l'équation

$$3y(t) + \int_{-\infty}^{\infty} (y''(s) - y(s))f(t-s)ds = g(t).$$

Calculer la transformée de Fourier de y en fonction de celles de f et g (on pourra s'aider du produit de convolution, et il n'est pas demandé de calculer la transformée de Fourier de g). En déduire y .

EXERCICE 18.8. Soit $\Pi(t) = \mathbf{1}_{[-\frac{1}{2}, \frac{1}{2}]}(t)$. Sa transformée de Fourier est donnée par :

$$\widehat{\Pi}(\nu) = \begin{cases} \frac{\sin(\pi\nu)}{\pi\nu} & \text{si } \nu \neq 0 \\ 1 & \text{si } \nu = 0 \end{cases}$$

- (1) Soit $\Pi_a(t) = \mathbf{1}_{[-\frac{a}{2}, \frac{a}{2}]}(t)$, où $a > 0$. Donner la transformée de Fourier de Π_a . On la note $\widehat{\Pi}_a$.
- (2) Calculer l'intégrale $\int_{-\infty}^{\infty} \frac{(\sin u)^2}{u^2} du$.
- (3) Calculer l'intégrale $\int_{-\infty}^{\infty} \frac{\sin u}{u} du$.
- (4) Calculer le produit de convolution $h = \widehat{\Pi}_a * \widehat{\Pi}_b$ (avec $a > 0$ et $b > 0$, et $\Pi_b(t) = \mathbf{1}_{[-\frac{b}{2}, \frac{b}{2}]}(t)$).

Bibliographie

- [1] Appel, W. Mathématiques pour la physique et les physiciens! *H et K* (2002)
- [2] Arfken G.B., Weber H.J. : Mathematical methods for physicists, *New York London, Academic Press* (1966)
- [3] Cartan H. : Cours de calcul différentiel, *Hermann* (1977).
- [4] Chatterji S.D. : Cours d'Analyse 1, Analyse vectorielle, *PPUR* (1997).
- [5] Fleming W. : Functions of several variables, second edition, *Springer Verlag* (1977).
- [6] Niglio , L. : Fonctions de plusieurs variables : rappels de cours, questions de réflexion, exercices d'entraînement, *Paris , Dunod* (1998).
- [7] K. E. Brenan, S. L. Campbell, and L. R. Petzold, *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*, SIAM 1996.
- [8] Michel Crouzeix, Alain L. Mignot, *Analyse numérique des équations différentielles* Edition Masson, Paris 1992
- [9] L. Dai, *Singular Control Systems*, Springer Verlag, 1989.
- [10] Demailly, Jean Pierre, *Analyse numérique et équations différentielles*, Presses universitaires de Grenoble, 1996.
- [11] J. Dieudonné, *Calcul infinitésimal*, Hermann, Paris 1980.
- [12] John H. Hubbard and Beverly Henderson, *Equations différentielles et systèmes dynamiques*, transl. by Véronique Gautheron Edition Cassini, 1999
- [13] G. Gasquet ; P. Witomski *Analyse de Fourier et Applications* Masson, 1990.
- [14] L. R. Petzold, Uri M. Ascher, *Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations*, SIAM 1998.
- [15] J. Vauthier, M. Maumy, *Calcul différentiel et intégral*, Ellipses 1998.