

# Cours de Mathématiques PCP S4

Xavier Buff (d'après le cours de Pierre Bousquet)

2021



# Table des matières

<b>1</b>	<b>Séries de fonctions</b>	<b>5</b>
1.1	Découvrir . . . . .	5
1.1.1	Rappels sur les suites de fonctions . . . . .	5
1.1.2	Les séries de fonctions . . . . .	6
1.2	Assimiler . . . . .	8
1.2.1	Convergence simple, absolue, uniforme, normale . . . . .	8
1.2.2	Propriétés de la somme d'une série de fonctions . . . . .	10
1.2.3	Cas des séries alternées . . . . .	11
1.2.4	Méthodes pour montrer la convergence uniforme d'une série de fonctions . . . . .	12
1.3	Approfondir . . . . .	13
1.3.1	Convergence normale implique convergence absolue et convergence uniforme . . . . .	13
1.3.2	Convergence uniforme des séries alternées de fonctions . . . . .	14
1.3.3	Critère de Cauchy . . . . .	14
<b>2</b>	<b>Séries entières</b>	<b>15</b>
2.1	Découvrir . . . . .	15
2.2	Assimiler . . . . .	18
2.2.1	Comparaisons de rayon de convergence . . . . .	18
2.2.2	Propriétés de la somme . . . . .	19
2.2.3	Fonctions développables en séries entières . . . . .	21
2.3	Approfondir . . . . .	23
<b>3</b>	<b>Séries de Fourier</b>	<b>27</b>
3.1	Découvrir . . . . .	27
3.1.1	L'analyse de Fourier et la compression d'images . . . . .	27
3.1.2	Intégrales de fonctions à valeurs complexes . . . . .	28
3.1.3	Polynômes trigonométriques à coefficients complexes . . . . .	28
3.1.4	Série de Fourier d'une fonction $T$ -périodique . . . . .	29
3.1.5	Propriétés des coefficients de Fourier . . . . .	30
3.1.6	Taille des coefficients de Fourier . . . . .	31
3.1.7	Problème inverse . . . . .	32
3.1.8	Objectif du chapitre . . . . .	32
3.2	Assimiler . . . . .	33
3.2.1	Convergence uniforme et convergence ponctuelle . . . . .	33
3.2.2	Produit de convolution . . . . .	33
3.2.3	Unité approchée . . . . .	34
3.2.4	Formule de Parseval . . . . .	35
3.2.5	Théorème de Dirichlet . . . . .	37
3.2.6	Théorème de Féjer . . . . .	38

3.2.7	Fonctions continues par morceaux . . . . .	40
3.2.8	Formule de Parseval . . . . .	41
3.3	Approfondir . . . . .	43
3.3.1	Notion de norme . . . . .	43
<b>4</b>	<b>Normes</b> . . . . .	<b>45</b>
4.1	Découvrir . . . . .	45
4.1.1	Normes et suites convergentes . . . . .	45
4.1.2	Normes équivalentes . . . . .	48
4.2	Assimiler . . . . .	48
4.2.1	Applications de l'équivalence des normes . . . . .	48
4.2.2	Boules et sphères . . . . .	49
4.2.3	Equivalence des normes et inclusions des boules . . . . .	50
4.3	Approfondir . . . . .	51
4.3.1	Quelques normes en dimension finie . . . . .	51
4.3.2	Des exemples de normes en dimension infinie . . . . .	54
4.3.3	Ouverts, fermés, bornés . . . . .	59
4.4	Rappels sur les espaces vectoriels . . . . .	62
4.4.1	Espaces vectoriels . . . . .	62
4.4.2	Sous-espaces vectoriels . . . . .	63
4.4.3	Dimension . . . . .	64
<b>5</b>	<b>Continuité des fonctions de plusieurs variables</b> . . . . .	<b>67</b>
5.1	Découvrir . . . . .	67
5.1.1	Limite et continuité pour les fonctions d'une variable . . . . .	67
5.1.2	Limite et continuité pour les fonctions de plusieurs variables à valeurs réelles . . . . .	68
5.2	Assimiler . . . . .	70
5.2.1	Continuité et critère séquentiel . . . . .	70
5.2.2	Opérations sur les fonctions continues . . . . .	70
5.2.3	Continuité des fonctions à valeurs dans $\mathbb{R}^p$ . . . . .	72
5.3	Approfondir . . . . .	73
5.3.1	Retour sur les limites et la continuité des fonctions à valeurs dans $\mathbb{R}^p$ . . . . .	73
5.3.2	Fonctions continues sur un fermé borné . . . . .	74
<b>6</b>	<b>Différentiabilité des fonctions de plusieurs variables</b> . . . . .	<b>75</b>
6.1	Découvrir . . . . .	75
6.1.1	Fonctions dérivables de $\mathbb{R}$ dans $\mathbb{R}$ . . . . .	75
6.1.2	Les dérivées partielles . . . . .	76
6.1.3	Gradient et différentielle d'une fonction . . . . .	77
6.2	Assimiler . . . . .	78
6.2.1	Fonctions de classe $C^1$ . . . . .	78
6.2.2	Dérivées partielles et différentielles des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}^p$ . . . . .	79
6.3	Approfondir . . . . .	82
6.3.1	Dérivées directionnelles . . . . .	82
6.3.2	Différentiabilité . . . . .	83
6.3.3	Différentiabilité des fonctions vectorielles . . . . .	87
6.3.4	Composition de fonctions différentiables . . . . .	88
6.3.5	Théorème des accroissements finis . . . . .	91

<b>7</b>	<b>Quelques compléments sur les fonctions de plusieurs variables</b>	<b>93</b>
7.1	Découvrir . . . . .	93
7.1.1	Changement de variables . . . . .	93
7.1.2	Dérivées partielles d'ordre supérieur . . . . .	94
7.2	Assimiler . . . . .	94
7.2.1	Quelques exemples de changements de variables . . . . .	94
7.2.2	Formule de Taylor-Young et extrema . . . . .	96
7.3	Approfondir . . . . .	99
7.3.1	Preuve de la formule de Taylor-Young : Théorème 7.14 . . . . .	99
7.3.2	Preuve du Théorème de Monge : Théorème 7.21 . . . . .	100
<b>8</b>	<b>Probabilités discrètes</b>	<b>103</b>
8.1	Événements aléatoires . . . . .	103
8.1.1	Issues et événements . . . . .	103
8.1.2	Probabilités . . . . .	104
8.2	Variables aléatoires discrètes . . . . .	107
8.3	Moments des variables aléatoires discrètes réelles . . . . .	108
8.3.1	Espérance . . . . .	108
8.3.2	Composition et linéarité . . . . .	109
8.3.3	Variance . . . . .	110
8.4	Indépendance . . . . .	113
8.4.1	Indépendance d'événements . . . . .	113
8.4.2	Indépendance de variables aléatoires . . . . .	113
8.4.3	Espérance et indépendance . . . . .	114
8.5	Processus de Bernoulli . . . . .	116
8.6	Statistiques . . . . .	119
8.6.1	Loi faible des grands nombres . . . . .	120
8.6.2	Théorème de Moivre-Laplace . . . . .	120
<b>9</b>	<b>Intégrales à paramètres et intégrales multiples</b>	<b>123</b>
9.1	Régularité des intégrales à paramètres . . . . .	123
9.1.1	Rappels sur l'intégrale sur un intervalle quelconque . . . . .	123
9.1.2	Continuité sous le signe intégral . . . . .	127
9.1.3	Dérivabilité sous le signe intégral . . . . .	128
9.1.4	Différentielle totale exacte . . . . .	129
9.1.5	Théorème de Fubini . . . . .	130



# Chapitre 1

## Séries de fonctions

### 1.1 Découvrir

#### 1.1.1 Rappels sur les suites de fonctions

Au semestre précédent, vous avez déjà étudié la convergence des suites de fonctions. On propose ici quelques rappels à ce sujet. Dans toute la suite, on notera  $I$  un intervalle (non vide) de  $\mathbb{R}$ . Toutes les fonctions seront définies sur  $I$  et à valeurs dans  $\mathbb{R}$ . Occasionnellement, il arrivera que les fonctions soient à valeurs dans  $\mathbb{C}$ . Les définitions se généralisent facilement à ce cas. Vous connaissez deux types de convergence :

1. la convergence simple ;
2. la convergence uniforme.

**Définition 1.1.** On dit qu'une suite de fonctions  $(f_n : I \rightarrow \mathbb{R})$  converge simplement sur  $I$  vers une fonction  $f : I \rightarrow \mathbb{R}$  si pour tout  $x \in I$ , la suite numérique  $(f_n(x))$  converge vers  $f(x)$ ; autrement dit,

$$\forall x \in I, \quad \forall \varepsilon > 0, \quad \exists N_{x,\varepsilon} \in \mathbb{N}, \quad \forall n \geq N_{x,\varepsilon}, \quad |f_n(x) - f(x)| \leq \varepsilon.$$

Pour montrer la convergence simple, il suffit donc d'étudier les suites numériques  $(f_n(x))_{n \in \mathbb{N}}$ .

Rappelons que  $\sup_{x \in I} |f(x)|$  désigne la borne supérieure (i.e. le plus petit majorant) de l'ensemble  $\{|f(x)|; x \in I\}$ . Cette borne supérieure est égale à  $+\infty$  si la fonction n'est pas bornée. On utilisera la notation

$$\|f\|_{L^\infty(I)} := \sup_{x \in I} |f(x)| \in [0, +\infty].$$

Quand il n'y a pas d'ambiguïté sur  $I$ , on note également  $\|f\|_\infty$  cette quantité. On appelle cette quantité la *norme uniforme* (ou encore norme infinie) de  $f$ . Le mot *norme* est un concept général en mathématique qu'on étudiera dans un prochain chapitre. Disons simplement ici que c'est un moyen de mesurer les distances entre deux objets.

**Définition 1.2.** On dit qu'une suite de fonctions  $(f_n : I \rightarrow \mathbb{R})$  converge uniformément vers une fonction  $f : I \rightarrow \mathbb{R}$  sur  $I$  si la suite numérique  $(\|f_n - f\|_{L^\infty(I)})$  converge vers 0. Autrement dit,

$$\forall \varepsilon > 0, \quad \exists N_\varepsilon \in \mathbb{N}, \quad \forall n \geq N_\varepsilon, \quad \forall x \in I, \quad |f_n(x) - f(x)| \leq \varepsilon.$$

Dans la formulation avec les quantificateurs, on constate que pour la convergence uniforme, l'entier  $N_\varepsilon$  doit pouvoir être choisi indépendamment de  $x$ , contrairement au cas de la convergence simple.

*Méthode 1.3.* Pour montrer qu'une suite de fonctions  $(f_n : I \rightarrow \mathbb{R})$  converge uniformément :

1. on commence par identifier une fonction  $f$  telle que la suite de fonctions  $(f_n)$  converge simplement vers  $f$ ,
2. on cherche à majorer la suite  $(\|f_n - f\|_{L^\infty(I)})$  par une suite qui tend vers 0.

Parfois, on peut calculer explicitement  $\|f_n - f\|_{L^\infty(I)}$  mais ce n'est pas toujours le cas.

La convergence uniforme implique la convergence simple mais la réciproque n'est généralement pas vraie.

*Exemple 1.4.* Pour  $n \in \mathbb{N}$ , soit  $f_n : [0, 1[ \rightarrow \mathbb{R}$  la fonction définie par  $f_n(x) = x^n$ . La suite de fonctions  $(f_n)_{n \in \mathbb{N}}$  converge simplement vers 0 sur  $[0, 1[$  mais ne converge pas uniformément vers 0 sur cet intervalle. En revanche, pour tout  $a \in ]0, 1[$ , la convergence est uniforme sur  $[0, a]$ .

*Exercice 1.5.*

1. Montrer les affirmations de l'exemple 1.4.
2. Soit  $(f_n)$  une suite de fonctions qui converge simplement vers une fonction  $f$ . On suppose que chaque  $f_n$  est croissante. Est-ce que  $f$  est croissante ?
3. Même question en remplaçant *croissante* par *strictement croissante*.

On dit également qu'une suite  $(f_n : I \rightarrow \mathbb{R})$  converge uniformément vers  $f$  sur tout segment de  $I$  si pour tout segment  $[a, b] \subseteq I$ , la suite des restrictions  $(f_n : [a, b] \rightarrow \mathbb{R})$  converge uniformément vers la restriction  $f : [a, b] \rightarrow \mathbb{R}$  sur  $[a, b]$ .

*Exercice 1.6.* Pour  $n \in \mathbb{N}$ , soit  $f_n : \mathbb{R} \rightarrow \mathbb{R}$  la fonction définie par  $f_n(x) = nxe^{-nx}$ . Etudier la convergence simple, puis uniforme de la suite de fonctions  $(f_n)_{n \in \mathbb{N}}$ .

L'intérêt de la convergence uniforme est mis en évidence dans les deux résultats suivants.

**Théorème 1.7.** Soit  $(f_n : I \rightarrow \mathbb{R})$  une suite de fonctions qui converge uniformément vers une fonction  $f$  sur tout segment de  $I$ .

1. Si chaque  $f_n$  est continue sur  $I$ , alors  $f$  est continue sur  $I$ .
2. Dans ce cas, pour tout  $a, b \in I$ ,

$$\lim_{n \rightarrow +\infty} \int_a^b f_n(t) dt = \int_a^b f(t) dt.$$

*Méthode 1.8.* Pour montrer qu'une suite de fonctions continues  $(f_n)$  ne converge pas uniformément, il suffit de montrer qu'elle converge simplement vers une fonction qui n'est pas continue.

**Théorème 1.9.** Soit  $(f_n : I \rightarrow \mathbb{R})$  une suite de fonctions de classe  $\mathcal{C}^1$ . On fait les hypothèses suivantes :

1. il existe  $a \in I$  tel que la suite numérique  $(f_n(a))$  converge ;
2. la suite de fonctions  $(f'_n : I \rightarrow \mathbb{R})$  converge uniformément vers une fonction  $g$  sur tout segment de  $I$ .

Alors la suite de fonctions  $(f_n : I \rightarrow \mathbb{R})$  converge uniformément sur tout segment  $[a, b] \subseteq I$  vers une fonction  $f : I \rightarrow \mathbb{R}$ . De plus,  $f$  est de classe  $\mathcal{C}^1$  sur  $I$  et  $f' = g$ .

## 1.1.2 Les séries de fonctions

La nouveauté dans ce chapitre est la traduction des définitions et propriétés liées à la convergence des *suites* de fonctions pour les *séries* de fonctions.

**Définition 1.10.** Soit  $(f_n : I \rightarrow \mathbb{R})_{n \geq 0}$  une suite de fonctions. La série  $\sum_{n \geq 0} f_n$  est la suite de fonctions  $(S_n)_{n \geq 0}$  définie par

$$S_n(x) := \sum_{k=0}^n f_k(x).$$

La fonction  $S_n$  s'appelle la  $n$ -ème somme partielle de la série.

**Définition 1.11.** Si la série de fonctions  $\sum_{n \geq 0} f_n$  converge sur  $I$ , c'est-à-dire si la suite  $(S_n)_{n \geq 0}$  des sommes partielles converge, la limite  $S : I \rightarrow \mathbb{R}$  s'appelle la somme de la série. On la note

$$S := \sum_{k=0}^{+\infty} f_k.$$

Dans ce cours, on distinguera la notation d'une série  $\sum_{n \geq 0} f_n$  de la notation pour désigner sa somme

$\sum_{n=0}^{+\infty} f_n$ . On peut toujours parler d'une série, mais on ne peut parler de sa somme que lorsqu'il y a convergence.

**Définition 1.12.** Si la série de fonctions  $\sum_{n \geq 0} f_n$  converge sur  $I$ , la fonction  $R_n := S - S_n : I \rightarrow \mathbb{R}$  s'appelle le  $n$ -ème reste de la série.

*Remarque 1.13.* On a

$$S = S_n + R_n \quad \text{et} \quad R_n = \sum_{k=n+1}^{+\infty} f_k.$$

Une grande famille de séries de fonctions est constituée des *séries entières*. Une série entière est une série de fonctions de la forme  $\sum_{n \geq 0} f_n$  où

$$f_n(x) = a_n x^n \quad \text{avec} \quad a_n \in \mathbb{C}.$$

La somme est définie sur l'ensemble des  $x \in \mathbb{R}$  tels que la série numérique  $\sum_{n \geq 0} a_n x^n$  converge. Avec un abus de notation, on identifie  $x^n$  avec la fonction  $x \mapsto x^n$  et on note la série entière

$$\sum_{n \geq 0} a_n x^n.$$

*Exemple 1.14.* Considérons la série entière  $\sum_{n \geq 0} x^n$ . La série numérique  $\left( \sum_{k=0}^n x^k \right)$  converge si  $|x| < 1$  et diverge grossièrement si  $|x| \geq 1$ . La série entière converge donc sur  $] -1, 1[$ . La somme  $S$  de la série entière est donc définie sur  $] -1, 1[$  et

$$\forall x \in ] -1, 1[, \quad S(x) = \sum_{n=0}^{+\infty} x^n = \frac{1}{1-x}.$$

Il y a bien sûr d'autres séries de fonctions que les séries entières. Par exemple, pour tout  $n \geq 1$ , considérons la fonction  $f_n : \mathbb{R} \rightarrow \mathbb{R}$  définie par

$$f_n(x) := \frac{1}{n^x}.$$

Considérons la série de fonctions  $\sum_{n \geq 1} f_n$  et notons  $S$  sa somme. Cette série de fonctions ne converge pas

sur  $\mathbb{R}$  et  $S$  n'est donc pas définie sur  $\mathbb{R}$  tout entier. Comme la série numérique  $\sum_{n \geq 1} \frac{1}{n^2}$  converge, on peut

dire que  $S(2)$  est bien défini. En revanche, comme la série numérique  $\sum_{n \geq 1} \frac{1}{n}$  diverge,  $x = 1$  n'appartient

pas au domaine de définition de  $S$ . En fait, d'après le critère de Riemann, la série converge précisément sur  $]1, +\infty[$ . Et pourtant, les fonctions  $f_n$  elles-mêmes sont définies sur  $\mathbb{R}$  ! Il faut donc distinguer le domaine de définition des fonctions  $f_n$  et le domaine de convergence de la série  $\sum_{n \geq 1} f_n$ .

L'une des difficultés de ce chapitre est qu'il faut solliciter à la fois la théorie des séries numériques et la théorie des suites de fonctions. Pour surmonter cette difficulté, il faut sans cesse se demander si l'objet qu'on considère est un nombre ou une fonction.

## 1.2 Assimiler

### 1.2.1 Convergence simple, absolue, uniforme, normale

Dans toute la suite, on se donne une suite de fonctions  $(f_n : I \rightarrow \mathbb{R})$  définies sur un intervalle  $I$  de  $\mathbb{R}$  et on considère la série de fonctions  $\sum_n f_n$ .

**Définition 1.15** (Convergence simple d'une série de fonctions). *La série de fonctions  $\sum_n f_n$  converge simplement sur  $I$  si  $\forall x \in I$ , la série numérique  $\sum_n f_n(x)$  converge.*

**Définition 1.16** (Convergence absolue d'une série de fonctions). *La série de fonctions  $\sum_n f_n$  converge absolument sur  $I$  si  $\forall x \in I$ , la série numérique  $\sum_n f_n(x)$  converge absolument, c'est-à-dire si  $\forall x \in I$ , la série numérique  $\sum_n |f_n(x)|$  converge.*

*Exemple 1.17.* Si  $f_n(x) = \frac{x^n}{n!}$ , la série de fonctions  $\sum_{n \geq 0} f_n$  converge absolument sur  $\mathbb{R}$ .

En effet, soit  $x \in \mathbb{R}$ . Si  $x = 0$ , alors  $f_n(0) = 0$  pour tout  $n \geq 0$  et la série  $\sum_n f_n(0)$  converge absolument. Si  $|x| > 0$ , on peut appliquer le critère de D'Alembert à la suite  $(u_n := f_n(x))$ . On a

$$\left| \frac{u_{n+1}}{u_n} \right| = \left| \frac{x^{n+1}/(n+1)!}{x^n/n!} \right| = \frac{|x|}{n+1} \xrightarrow{n \rightarrow +\infty} 0 < 1.$$

Donc la série  $\sum_n u_n$  converge absolument. Donc la série  $\sum_n f_n(x)$  converge absolument.

*Exemple 1.18.* Si  $f_n(x) = \frac{x^n}{n}$ , la série de fonctions  $\sum_{n \geq 0} f_n$  converge absolument sur  $] -1, 1[$  (critère de D'Alembert). Elle converge simplement sur  $[-1, 1[$  (en  $-1$ , on utilise le critère des séries alternées).

La convergence absolue d'une série numérique implique la convergence de la série numérique. On en déduit le résultat suivant.

**Proposition 1.19.** *Si une série de fonctions converge absolument sur  $I$ , alors elle converge simplement sur  $I$ .*

*Démonstration.* Supposons que la série de fonctions  $\sum_n f_n$  converge absolument sur  $I$ . Alors pour tout  $x \in I$ , la série numérique  $\sum_n f_n(x)$  est absolument convergente. Donc pour tout  $x \in I$ , la série numérique  $\sum_n f_n(x)$  est convergente. Par conséquent, la série de fonctions  $\sum_n f_n$  converge simplement sur  $I$ .  $\square$

**Définition 1.20** (Convergence uniforme d'une série de fonctions). *La série de fonctions  $\sum_{n \geq 0} f_n$  converge*

*uniformément sur  $I$  si la suite des sommes partielles  $\left( S_n := \sum_{k=0}^n f_k \right)$  converge uniformément sur  $I$ .*

*Remarque 1.21.* Étant donné que par définition, la série de fonctions  $\sum_{n \geq 0} f_n$  est la suite de fonction  $(S_n)$ , cette définition est en fait une tautologie.

Si une suite de fonctions converge uniformément vers une fonction, alors elle converge simplement vers cette fonction. Il en va de même pour les séries : si la suite des sommes partielles  $(S_n)$  converge uniformément, alors elle converge simplement. Dans ce cas, la limite de la suite  $(S_n)$  est la somme

$$S := \sum_{k=0}^{+\infty} f_k.$$

**Proposition 1.22.** *Pour qu'une série de fonctions  $\sum_n f_n$  converge uniformément sur  $I$ , il faut, et il suffit, qu'elle converge simplement sur  $I$  et que la suite  $(R_n)$  de ses restes converge uniformément vers 0 sur  $I$ .*

*Démonstration.* Supposons que la suite des sommes partielles  $(S_n)$  converge uniformément sur  $I$ . Alors elle converge simplement sur  $I$ . Notons  $S$  sa somme et  $R_n := S - S_n$  le  $n$ -ème reste, de sorte que

$$\|R_n\|_{L^\infty(I)} = \|S_n - S\|_{L^\infty(I)} \xrightarrow{n \rightarrow +\infty} 0.$$

La suite  $(R_n)$  converge donc uniformément vers 0 sur  $I$ .

Réciproquement, supposons que la suite  $(S_n)$  converge simplement vers  $S$  sur  $I$  et que la suite des restes  $(R_n = S - S_n)$  converge uniformément vers 0 sur  $I$ . Alors

$$\|S_n - S\|_{L^\infty(I)} = \|R_n\|_{L^\infty(I)} \xrightarrow{n \rightarrow +\infty} 0.$$

La série de fonctions  $\sum_n f_n$  converge donc uniformément vers  $S$  sur  $I$ .  $\square$

**Définition 1.23** (Convergence normale d'une série de fonctions). La série de fonctions  $\sum_n f_n$  converge normalement sur  $I$  si la série numérique  $\sum_n \|f_n\|_{L^\infty(I)}$  est convergente.

*Exemple 1.24.* Soit  $(f_n : \mathbb{R} \rightarrow \mathbb{R})$  la suite de fonctions définie par  $f_n(x) := \frac{x^n}{n!}$ . Alors pour tout  $M \in \mathbb{R}^+$ , la série de fonctions  $\sum_n f_n$  converge normalement sur  $[-M, M]$ . En revanche, elle ne converge pas normalement sur  $\mathbb{R}$ .

En effet, soit  $M \in \mathbb{R}^+$ . Alors, pour tout entier  $n \geq 0$

$$\|f_n\|_{L^\infty([-M, M])} = \frac{M^n}{n!}.$$

Donc la série numérique  $\sum_n \|f_n\|_{L^\infty([-M, M])}$  est convergente. Par conséquent, la série de fonctions  $\sum_n f_n$  converge normalement sur  $[-M, M]$ . En revanche,

$$\|f_n\|_{L^\infty(\mathbb{R})} = +\infty.$$

Donc la série  $\sum_n \|f_n\|_{L^\infty(\mathbb{R})}$  diverge et la série de fonctions  $\sum_n f_n$  ne converge pas normalement sur  $\mathbb{R}$ .

*Méthode 1.25.* Pour montrer la convergence normale d'une série de fonctions, on peut majorer  $\|f_n\|_{L^\infty(I)}$  par une quantité qui est le terme général d'une série convergente.

*Exemple 1.26.* La série  $\sum_{n \geq 1} \frac{1 + \cos(nx)}{n^2 + x^2}$  converge normalement sur  $\mathbb{R}$ . En effet, on peut utiliser la majoration

$$\|f_n\|_{L^\infty(I)} := \sup_{x \in \mathbb{R}} \frac{1 + \cos(nx)}{n^2 + x^2} \leq \frac{2}{n^2}.$$

Or la série  $\sum_{n \geq 1} \frac{1}{n^2}$  converge. Donc la série  $\sum_{n \geq 1} \|f_n\|_{L^\infty(I)}$  converge par critère de majoration des séries à termes positifs. Donc la série de fonctions  $\sum_{n \geq 1} \frac{1 + \cos(nx)}{n^2 + x^2}$  converge normalement.

**Proposition 1.27.** La convergence normale implique la convergence absolue et la convergence uniforme.

La preuve de la proposition est proposée dans la section *Approfondir*.

*Méthode 1.28.* Pour montrer la convergence uniforme d'une série de fonctions, il est souvent plus simple de chercher à montrer la convergence normale (il existe cependant des séries de fonctions qui convergent uniformément mais pas normalement).

## 1.2.2 Propriétés de la somme d'une série de fonctions

**Proposition 1.29.** Si chaque fonction  $f_n : I \rightarrow \mathbb{R}$  est continue et si la série  $\sum_n f_n$  converge uniformément sur tout segment de  $I$  vers une fonction  $S : I \rightarrow \mathbb{R}$ , alors  $S$  est continue sur  $I$ .

*Démonstration.* La  $n$ -ème somme partielle  $S_n$  est continue comme somme finie de fonctions continues. De plus, la suite de fonctions  $(S_n)$  converge uniformément vers  $S$  sur tout segment de  $I$ . D'après le point 1 du Théorème 1.7, la limite  $S$  est une fonction continue sur  $I$ .  $\square$

**Proposition 1.30.** Si chaque fonction  $f_n : I \rightarrow \mathbb{R}$  est continue et si la série  $\sum_n f_n$  converge uniformément sur  $I$ , alors pour tout  $a, b \in I$ , la série numérique  $\sum_n \int_a^b f_n(t) dt$  converge et

$$\int_a^b \sum_{n=0}^{+\infty} f_n(t) dt = \sum_{n=0}^{+\infty} \int_a^b f_n(t) dt.$$

*Démonstration.* En considérant toujours la  $n$ -ème somme partielle  $S_n := \sum_{k=0}^n f_k$ , on a par linéarité de l'intégrale (cas d'une somme finie) :

$$\sum_{k=0}^n \int_a^b f_k(t) dt = \int_a^b \sum_{k=0}^n f_k(t) dt = \int_a^b S_n(t) dt.$$

La suite  $(S_n)$  converge uniformément vers  $S$  sur tout segment de  $I$ . D'après le point 2 du Théorème 1.7,

$$\int_a^b S_n(t) dt \xrightarrow{n \rightarrow +\infty} \int_a^b S(t) dt = \int_a^b \sum_{k=0}^{+\infty} f_k(t) dt.$$

Donc la série du membre de gauche a aussi une limite, et on peut écrire

$$\sum_{k=0}^{+\infty} \int_a^b f_k(t) dt = \lim_{n \rightarrow +\infty} \sum_{k=0}^n \int_a^b f_k(t) dt = \lim_{n \rightarrow +\infty} \int_a^b S_n(t) dt = \int_a^b \sum_{k=0}^{+\infty} f_k(t) dt. \quad \square$$

**Proposition 1.31.** On suppose que

1. chaque  $f_n : I \rightarrow \mathbb{R}$  est de classe  $\mathcal{C}^1$ ,
2. il existe  $a \in I$  tel que la série numérique  $\sum_n f_n(a)$  converge simplement et
3. la série de fonctions  $\sum_n f'_n$  converge uniformément sur tout segment de  $I$ .

Alors, la série de fonctions  $\sum_n f_n$  converge uniformément sur tout segment de  $I$  vers une fonction  $S : I \rightarrow \mathbb{R}$  de classe  $\mathcal{C}^1$  sur  $I$  et

$$\left( \sum_{n=0}^{+\infty} f_n \right)' = S' = \sum_{n=0}^{+\infty} f'_n.$$

*Exercice 1.32.* Ecrire la preuve de la Proposition 1.31. Il est recommandé de s'appuyer sur le Théorème 1.7.

### 1.2.3 Cas des séries alternées

Il existe un cas important de séries pour lequel la convergence normale est souvent mise en défaut, mais où il est facile de montrer la convergence uniforme : les séries alternées de fonctions.

**Définition 1.33.** Soit  $(f_n : I \rightarrow \mathbb{R})$  une suite de fonctions. On dit que  $\sum_n f_n$  est une série alternée (de fonctions) si pour tout  $x \in I$ , la série numérique  $\sum_n f_n(x)$  est une série alternée, c'est-à-dire :

1.  $\forall x \in I, \forall n \in \mathbb{N}, (-1)^n f_n(x) \geq 0,$
2.  $\forall x \in I, \forall n \in \mathbb{N}, |f_{n+1}(x)| \leq |f_n(x)|,$
3.  $\forall x \in I, \lim_{n \rightarrow +\infty} f_n(x) = 0.$

D'après le théorème des séries alternées, la série numérique  $\sum_n f_n(x)$  converge pour tout  $x \in I$ .

Autrement dit, la série de fonctions  $\sum_n f_n$  converge simplement sur  $I$ .

**Théorème 1.34.** Soit  $\sum_n f_n$  une série alternée (de fonctions). On suppose que la suite  $(f_n)$  converge

uniformément vers 0 sur  $I$ . Alors la série de fonctions  $\sum_{n=0}^{+\infty} f_n$  converge uniformément sur  $I$ .

La preuve de ce théorème qui est proposée dans la section suivante (Exercice 1.40) repose sur l'estimation des restes  $R_n := \sum_{k=n+1}^{+\infty} f_k$ . Dans le cas d'une série alternée de fonctions, on a

$$\forall x \in I, \quad |R_n(x)| \leq |f_{n+1}(x)|.$$

**Corollaire 1.35.** Soit  $\sum_n f_n$  une série alternée (de fonctions). Si chaque  $f_n$  est continue sur  $I$ , alors

$\sum_{n=0}^{+\infty} f_n$  est continue sur  $I$ .

*Exemple 1.36.* La série de fonctions  $\sum_{n \geq 1} \frac{(-1)^n}{n+x^2}$  converge uniformément sur  $\mathbb{R}$ . Sa somme est donc continue sur  $\mathbb{R}$ .

## 1.2.4 Méthodes pour montrer la convergence uniforme d'une série de fonctions

On a vu dans cette section plusieurs méthodes pour montrer la convergence uniforme d'une série de fonctions.

*Méthode 1.37.* Pour montrer qu'une série de fonctions  $\sum_n f_n$  converge uniformément, on peut

1. essayer de montrer qu'elle converge normalement, c'est-à-dire que la série numérique

$$\sum_n \|f_n\|_{L^\infty(I)}$$

converge; pour cela,

- (a) on peut trouver un majorant  $\alpha_n$  de  $\|f_n\|_{L^\infty(I)}$ ,
  - (b) tel que la série  $\sum_n \alpha_n$  converge;
2. vérifier si  $\sum_n f_n$  est une série entière (voir chapitre suivant) : on sait alors qu'elle converge uniformément sur tout segment de son intervalle de convergence;
  3. vérifier si  $\sum_n f_n$  est une série alternée de fonctions; on sait alors que la série  $\sum_n f_n$  converge simplement; on pourra conclure que la série converge uniformément si la suite  $(f_n)$  converge uniformément vers 0.

*Exercice 1.38.* Etudier la convergence de la série de fonctions  $\sum_{n \geq 1} f_n$  avec  $f_n(x) := \frac{(-1)^n e^{-nx}}{n}$ .

*Exercice 1.39.* Montrer que la série de fonctions  $\sum_{n \geq 1} f_n$  avec  $f_n(x) := \frac{(-1)^n (2n+1)}{(2n+1)^2 + x^2}$  converge simplement sur  $\mathbb{R}$ . La somme est-elle continue ?

## 1.3 Approfondir

### 1.3.1 Convergence normale implique convergence absolue et convergence uniforme

Dans cette partie, nous montrons la Proposition 1.27. Supposons que la série de fonctions  $\sum_n f_n$  converge normalement sur un intervalle  $I$ . Posons

$$\mu_n := \|f_n\|_{L^\infty(I)}.$$

Alors, la série numérique  $\sum_n \mu_n$  converge.

On montre d'abord que cela implique la convergence absolue de la série de fonctions  $\sum_n f_n$ . Soit  $x \in I$ . Comme pour tout  $n \in \mathbb{N}$ , on a  $0 \leq |f_n(x)| \leq \mu_n$  et comme la série numérique  $\sum_n \mu_n$  converge, on en déduit que la série numérique  $\sum_n |f_n(x)|$  converge aussi (critère de majoration pour les séries à termes positifs). Cela montre la convergence absolue de la série de fonctions  $\sum_n f_n$ , et donc aussi sa convergence simple.

Montrons maintenant la convergence uniforme de la série de fonctions  $\sum_n f_n$ . D'après la Proposition 1.22, il suffit de montrer que la série  $(R_n)$  des restes  $R_n := \sum_{k=n+1}^{+\infty} f_k$  converge uniformément vers 0. D'après l'inégalité triangulaire, pour tout  $x \in I$ ,

$$0 \leq |R_n(x)| = \left| \sum_{k=n+1}^{+\infty} f_k(x) \right| \leq \sum_{k=n+1}^{+\infty} |f_k(x)| \leq \sum_{k=n+1}^{+\infty} \mu_k.$$

Comme l'encadrement précédent est valide pour tout  $x \in I$ , on a donc

$$0 \leq \|R_n\|_{L^\infty(I)} \leq \sum_{k=n+1}^{+\infty} \mu_k.$$

Comme la série numérique  $\sum_n \mu_n$  converge,

$$\sum_{k=n+1}^{+\infty} \mu_k = \sum_{k=0}^{+\infty} \mu_k - \sum_{k=0}^n \mu_k \xrightarrow{n \rightarrow +\infty} 0.$$

On a donc

$$\|R_n\|_{L^\infty(I)} \xrightarrow{n \rightarrow +\infty} 0,$$

ce qui montre que la série  $(R_n)$  des restes converge uniformément vers 0.

### 1.3.2 Convergence uniforme des séries alternées de fonctions

Nous proposons d'aborder la démonstration du Théorème 1.34 sous forme d'exercice.

*Exercice 1.40.*

Soit  $(a_n)_{n \in \mathbb{N}}$  une suite décroissante de nombres positifs.

1. Montrer que pour tout  $p, q \in \mathbb{N}$ , avec  $p < q$ , on a

$$\left| \sum_{k=p}^q (-1)^k a_k \right| \leq a_p.$$

2. On suppose de plus que la suite  $(a_n)_{n \in \mathbb{N}}$  tend vers 0. Montrer que pour tout  $p \in \mathbb{N}$ ,

$$\left| \sum_{k=p}^{+\infty} (-1)^k a_k \right| \leq a_p.$$

Soit  $\sum_n f_n$  une série alternée de fonctions définies sur  $I$ .

3. Montrer que pour tout  $x \in I$ , pour tout  $n \in \mathbb{N}$ ,

$$\left| \sum_{k=n+1}^{+\infty} f_k(x) \right| \leq |f_{n+1}(x)|.$$

4. En déduire une preuve du Théorème 1.34.

### 1.3.3 Critère de Cauchy

Il arrive qu'on ait à montrer la convergence uniforme d'une suite ou d'une série de fonctions, sans connaître la limite. Le critère de Cauchy est alors bien utile.

**Proposition 1.41** (Critère de Cauchy pour la convergence uniforme).

1. Une suite de fonctions  $(f_n)_{n \in \mathbb{N}}$  converge uniformément vers une fonction  $f$  sur un intervalle  $I$  si et seulement si

$$\forall \varepsilon > 0, \quad \exists n \in \mathbb{N}, \quad \forall p \geq 0, \quad \|f_{n+p} - f_n\|_{L^\infty(I)} \leq \varepsilon.$$

2. Une série de fonctions  $\sum_n f_n$  converge uniformément vers une fonction  $f$  sur un intervalle  $I$  si et seulement si

$$\forall \varepsilon > 0, \quad \exists n \in \mathbb{N}, \quad \forall p \geq 0, \quad \left\| \sum_{k=n}^{n+p} f_k \right\|_{L^\infty(I)} \leq \varepsilon.$$

La preuve de cette proposition repose sur le fait que  $\mathbb{R}$  est complet : cela signifie que toutes les suites de Cauchy dans  $\mathbb{R}$  convergent. Une suite de Cauchy dans  $\mathbb{R}$  est une suite numérique  $(x_n)_{n \in \mathbb{N}}$  telle que

$$\forall \varepsilon > 0, \quad \exists n \in \mathbb{N}, \quad \forall p \geq 0, \quad |x_{n+p} - x_n| \leq \varepsilon.$$

La complétude de  $\mathbb{R}$  est étroitement liée à la définition même de  $\mathbb{R}$ . Nous l'admettrons, ainsi que la preuve de la proposition.

# Chapitre 2

## Séries entières

### 2.1 Découvrir

On peut voir les séries entières comme une généralisation des polynômes, où il y aurait une infinité de termes. Comme chaque fois qu'on *manipule l'infini*, il faut en préciser le sens.

**Définition 2.1.** Une série entière est une série de fonctions  $\sum_n f_n$  où  $f_n(x) = a_n x^n$  avec  $a_n \in \mathbb{R}$ .

*Remarque 2.2.* On peut également considérer le cas de séries entières à coefficients complexes  $a_n \in \mathbb{C}$ .

Comme signalé dans le premier chapitre, avec un abus de notation, on identifie  $x^n$  et la fonction  $x \mapsto x^n$  et on note les séries entières  $\sum_{n \geq 0} a_n x^n$ .

Les sommes partielles  $S_n := \sum_{k=0}^n a_k x^k$  associées à une série entière, sont des polynômes. C'est en ce sens que les séries entières peuvent être considérées comme des limites de polynômes.

Notons que toutes les séries entières convergent en au moins un point :  $x = 0$ .

*Exemple 2.3.* La série entière  $\sum_{n \geq 0} x^n$  converge sur  $] -1, 1[$  et diverge grossièrement en dehors de cet intervalle. En fait, on peut calculer la somme partielle

$$S_n(x) = \sum_{k=0}^n x^k = \frac{1 - x^{n+1}}{1 - x}.$$

On en déduit que

$$\forall x \in ] -1, 1[, \quad \sum_{n=0}^{+\infty} x^n = \frac{1}{1 - x}.$$

*Exemple 2.4.*

1.  $\sum_n n x^n$  converge sur  $] -1, 1[$  (critère de D'Alembert),
2.  $\sum_n \frac{x^n}{n+1}$  converge sur  $[ -1, 1[$  (en  $-1$ , on utilise le critère des séries alternées),
3.  $\sum_n \frac{x^n}{n!}$  converge sur  $\mathbb{R}$  (critère de D'Alembert).

**Théorème 2.5.** Soit  $\sum_n a_n x^n$  une série entière. Il existe  $R \in [0, +\infty]$  tel que

1. pour tout  $x \notin [-R, R]$ , la série diverge,
2. pour tout  $x \in ]-R, R[$ , la série converge absolument.

Le théorème ne dit rien sur ce qui se passe pour  $x = R$  ou  $x = -R$ . Pour l'instant, on peut se contenter de retenir que la somme  $\sum_{n=0}^{+\infty} a_n x^n$  a pour domaine de définition soit  $] -R, R[$ , soit  $[-R, R]$ , soit  $[-R, R[$ , soit  $] -R, R]$ .

La preuve du théorème repose sur un lemme qui joue un rôle très important dans tout ce chapitre, le lemme d'Abel.

**Lemme 2.6.** Soit  $r \in \mathbb{R}^+$ . Si la suite  $(a_n r^n)_{n \in \mathbb{N}}$  est bornée, alors pour tout  $x \in \mathbb{R}$  tel que  $|x| < r$ , la série  $\sum_n a_n x^n$  converge absolument.

*Démonstration.* On écrit

$$|a_n x^n| = |a_n r^n| \left| \frac{x}{r} \right|^n.$$

Par hypothèse, il existe  $M > 0$  tel que pour tout  $n \geq 0$ ,  $|a_n r^n| \leq M$ . Donc

$$|a_n x^n| \leq M q^n,$$

avec  $q = \left| \frac{x}{r} \right|$ . Comme  $0 < q < 1$ , la série  $\sum_n q^n$  converge (voir l'Exemple 2.3). Par comparaison entre séries à termes positifs, on en déduit que la série  $\sum_n |a_n x^n|$  converge, ce qui conclut la preuve du lemme.  $\square$

*Démonstration du Théorème 2.5.* Posons

$$R := \sup \{ r \in \mathbb{R}^+ ; (a_n r^n) \text{ est bornée} \} \in [0, +\infty].$$

Cette borne supérieure est bien définie car l'ensemble est non vide : il contient  $r = 0$ .

On commence par montrer le premier point. Si  $R = +\infty$ , il n'y a rien à montrer. Si  $R < +\infty$  et si  $|x| > R$ , alors la suite  $(a_n x^n)$  n'est pas bornée. La série  $\sum_n a_n x^n$  diverge donc grossièrement.

On montre à présent le deuxième point. Si  $R = 0$ , il n'y a rien à démontrer. On suppose désormais  $R > 0$ . On va établir une conclusion plus forte que celle de l'énoncé : la série  $\sum_n a_n x^n$  converge normalement sur tout segment  $[-r, r] \subset ]-R, R[$ . Soit  $r \in ]0, R[$ . D'une part,

$$\sup_{x \in [-r, r]} |a_n x^n| = |a_n r^n|.$$

D'autre part,  $r$  n'est pas un majorant de  $\{ r \in \mathbb{R}^+ ; (a_n r^n) \text{ est bornée} \}$  (car  $R$  est le plus petit majorant). Il existe donc  $r_1 > r$  tel que la suite  $(a_n r_1^n)$  est bornée. D'après le lemme d'Abel, on en déduit que la série  $\sum_n |a_n r_1^n|$  est convergente. La série  $\sum_n a_n x^n$  converge donc normalement sur  $[-r, r]$ .  $\square$

**Définition 2.7.** Le rayon de convergence d'une série entière  $\sum_n a_n x^n$  est

$$R := \sup \{ r \in \mathbb{R}^+ ; (a_n r^n) \text{ est bornée} \} \in [0, +\infty].$$

La preuve du Théorème 2.5 a montré davantage que l'énoncé.

**Proposition 2.8.** Une série entière de rayon de convergence  $R$  converge normalement (et donc uniformément) sur tout segment de  $] -R, R[$ .

- Remarque 2.9.*
1. Si une série entière  $\sum_n a_n x^n$  diverge en un point  $x_0$ , alors son rayon de convergence est inférieur ou égal à  $|x_0|$ .
  2. Si  $\sum_n |a_n| R^n$  converge, alors la série entière converge normalement et donc uniformément sur  $[-R, R]$ .

La proposition suivante donne un critère pratique pour calculer le rayon de convergence d'une série entière  $\sum_n a_n x^n$  (c'est une conséquence du critère de D'Alembert).

**Proposition 2.10.** Si la suite  $\left( \left| \frac{a_{n+1}}{a_n} \right| \right)_n$  converge vers  $\ell \in [0, +\infty]$ , alors  $R = \frac{1}{\ell}$ .

Dans cet énoncé (et dans sa preuve ci-dessous), on utilise la convention  $\frac{1}{0} = +\infty$  et  $\frac{1}{+\infty} = 0$ . Il se peut que la limite de  $\left( \left| \frac{a_{n+1}}{a_n} \right| \right)_n$  n'existe pas. Mais le rayon de convergence, lui, est toujours défini.

*Démonstration.* Supposons que  $0 < r < \frac{1}{\ell}$ .<sup>1</sup> Alors

$$\lim_{n \rightarrow +\infty} \frac{|a_{n+1} r^{n+1}|}{|a_n r^n|} = \ell r < 1. \quad (2.1)$$

Donc par le critère de D'Alembert, la série  $\sum_n a_n r^n$  converge absolument, et donc la suite  $(a_n r^n)$  est bornée. Donc  $r \leq R$ . Comme cette inégalité est valide pour tout  $r < \frac{1}{\ell}$ , on en déduit que  $\frac{1}{\ell} \leq R$  en faisant tendre  $r$  vers  $\frac{1}{\ell}$ .

Supposons ensuite que  $r > \frac{1}{\ell}$ .<sup>2</sup> Alors  $\ell r > 1$  et donc

$$\lim_{n \rightarrow +\infty} \frac{|a_{n+1} x^{n+1}|}{|a_n x^n|} = \ell r > 1.$$

Le critère de D'Alembert implique alors que la série  $\sum_n a_n r^n$  diverge. Par le Théorème 2.5,  $r \geq R$ .

Comme cette inégalité est valide pour tout  $r > \frac{1}{\ell}$ , on en déduit que  $\frac{1}{\ell} \geq R$  en faisant tendre  $r$  vers  $\frac{1}{\ell}$ .

En conclusion,  $\frac{1}{\ell} \leq R \leq \frac{1}{\ell}$ , c'est-à-dire  $R = \frac{1}{\ell}$ .  $\square$

*Exercice 2.11.* Soient  $\sum_n a_n x^n$  de rayon de convergence  $R_a$  et  $\sum_n b_n x^n$  de rayon de convergence  $R_b$ .

1. Montrer que  $\sum_n (a_n + b_n) x^n$  a un rayon de convergence  $R \geq \min(R_a, R_b)$  avec égalité si  $R_a \neq R_b$ .
2. Montrer que pour tout  $x \in ] -R, R[$ ,

$$\sum_{n=0}^{+\infty} (a_n + b_n) x^n = \sum_{n=0}^{+\infty} a_n x^n + \sum_{n=0}^{+\infty} b_n x^n.$$

---

1. Ce cas n'advient que si  $\ell \neq +\infty$ .  
2. Ce cas n'advient que si  $\ell \neq 0$ .

*Exercice 2.12.* On suppose que la série  $\sum_n a_n x_0^n$  converge simplement mais pas absolument. Justifier que le rayon de convergence de la série entière  $\sum_n a_n x^n$  est égal à  $|x_0|$ .

*Exercice 2.13.* Montrer que si  $\lim_{n \rightarrow +\infty} \frac{|a_{n+2}|}{|a_n|} = \lambda$ , alors le rayon de convergence de  $\sum_n a_{2n} x^{2n}$  est  $\frac{1}{\sqrt{\lambda}}$ .

## 2.2 Assimiler

### 2.2.1 Comparaisons de rayon de convergence

Comme pour les séries numérique, la comparaison de séries entières est un outil puissant pour déterminer les rayons de convergence.

**Proposition 2.14.** Soient  $\sum_n a_n x^n$  et  $\sum_n b_n x^n$  deux séries entières de rayon de convergence respectifs  $R_a$  et  $R_b$ . S'il existe  $M \in \mathbb{R}^+$  et  $n_0 \in \mathbb{N}$  tels que  $|a_n| \leq M|b_n|$  pour  $n \geq n_0$ , alors  $R_a \geq R_b$ .

*Démonstration.* Si  $R_b = 0$ , il n'y a rien à démontrer. Supposons donc que  $R_b > 0$ . Supposons que  $0 < r < R_b$ . Alors, la série  $\sum_n b_n r^n$  converge, et donc la suite  $(b_n r^n)$  est bornée. Comme pour tout  $n \geq n_0$ , on a  $|a_n r^n| \leq M|b_n r^n|$ , on en déduit que la suite  $(a_n r^n)$  est bornée. On a donc  $R_a \geq r$ . Comme cette inégalité est valide pour tout  $r < R_b$ , on a  $R_a \geq R_b$  en faisant tendre  $r$  vers  $R_b$ .  $\square$

**Proposition 2.15.** Soient  $\sum_n a_n x^n$  et  $\sum_n b_n x^n$  deux séries entières. On suppose que  $a_n \sim b_n$  quand  $n \rightarrow +\infty$ . Alors les deux séries ont même rayon de convergence.

*Démonstration.* Notons  $R_a$  et  $R_b$  les rayons de convergence respectifs de  $\sum_n a_n x^n$  et  $\sum_n b_n x^n$ . Par hypothèse, on sait que pour  $n$  assez grand,  $a_n = b_n u_n$  avec  $u_n \rightarrow 1$  quand  $n \rightarrow +\infty$ . La suite  $(u_n)$  converge, elle est donc bornée. Par conséquent, il existe  $M \in \mathbb{R}^+$  et  $n_0 \in \mathbb{N}$  tels que  $|a_n| \leq M|b_n|$  pour  $n \geq n_0$ . D'après la proposition 2.14, on a  $R_a \geq R_b$ .

De même, pour  $n$  assez grand,  $b_n = a_n v_n$  avec  $v_n \rightarrow 1$  quand  $n \rightarrow +\infty$ . La suite  $(v_n)$  converge, elle est donc bornée. Par conséquent, il existe  $M' \in \mathbb{R}^+$  et  $n_1 \in \mathbb{N}$  tels que  $|b_n| \leq M'|a_n|$  pour tout  $n \geq n_1$ . D'après la proposition 2.14, on a  $R_b \geq R_a$ .

Par conséquent,  $R_a \geq R_b \geq R_a$ , et donc  $R_b = R_a$ .  $\square$

**Proposition 2.16.** Les séries  $\sum_{n \geq 0} a_n x^n$ ,  $\sum_{n \geq 1} n a_n x^{n-1}$  et  $\sum_{n \geq 0} \frac{a_n}{n+1} x^{n+1}$  ont le même rayon de convergence.

*Remarque 2.17.* La deuxième série est la série dérivée, et la troisième série est la série primitive.

*Démonstration.* Observons que le rayon de convergence de la série  $\sum_{n \geq 1} n a_n x^{n-1}$  est le même que celui de la série  $\sum_{n \geq 1} n a_n x^n$  (on passe de l'une à l'autre en multipliant ou en divisant par  $x$ ). Notons  $R$  le rayon de convergence et la série  $\sum_{n \geq 1} a_n x^n$  et  $R'$  celui de la série  $\sum_{n \geq 1} n a_n x^n$ . Comme pour tout  $n \geq 1$ , on a  $|a_n| \leq |n a_n|$ , on déduit de la proposition 2.14 que  $R \geq R'$ .

Il nous faut donc montrer que  $R' \geq R$ . Il suffit de montrer que  $\sum na_n x^n$  converge sur  $] -R, R[$ . Si  $R = 0$ , il n'y a rien à démontrer. Supposons donc que  $R > 0$ . Soient  $x \in ] -R, R[$  et  $r$  tels que  $|x| < r < R$ . Notons que

$$|na_n x^n| = |a_n r^n| \cdot |nq^n| \quad \text{avec} \quad 0 \leq q := \frac{|x|}{r} < 1.$$

La suite  $(nq^n)$  tend vers 0 par croissances comparées. Elle est donc bornée. Autrement dit, il existe  $M \in \mathbb{R}$  tel que  $|na_n x^n| \leq M \cdot |a_n r^n|$  pour tout  $n \geq 1$ . La série  $\sum_{n \geq 1} |a_n r^n|$  converge car  $r < R$ . Par comparaison des séries à termes positifs, la série  $\sum_{n \geq 1} na_n x^n$  converge.  $\square$

## 2.2.2 Propriétés de la somme

**Théorème 2.18.** *La somme d'une série entière de rayon de convergence  $R$  est de classe  $\mathcal{C}^\infty$  sur l'intervalle  $] -R, R[$ .*

*Démonstration.* Soit  $\sum_{n \geq 0} a_n x^n$  une série entière de rayon de convergence  $R$ . La fonction  $f_n : \mathbb{R} \rightarrow \mathbb{R}$  définie par  $f_n(x) := a_n x^n$  est de classe  $\mathcal{C}^1$ . La série numérique  $\sum_{n \geq 0} f_n(0)$  est constante égale à  $a_0$ ; elle est donc convergente. D'après la proposition 2.16, le rayon de convergence de la série des dérivées  $\sum_{n \geq 0} f'_n = \sum_{n \geq 1} na_n x^{n-1}$  est égal à  $R$ . La série  $\sum_{n \geq 0} f'_n$  est donc normalement (et donc uniformément) convergente sur tout segment de  $] -R, R[$ . D'après la Proposition 1.31, la somme  $S$  de la série entière  $\sum_{n \geq 0} a_n x^n$  est de classe  $\mathcal{C}^1$  sur  $] -R, R[$  et

$$S'(x) = \sum_{n=1}^{+\infty} na_n x^{n-1}.$$

En répétant cet argument pour  $S'$  au lieu de  $S$ , on conclut que  $S'$  est de classe  $\mathcal{C}^1$  sur  $] -R, R[$ , de dérivée  $x \mapsto \sum_{n=2}^{+\infty} n(n-1)a_n x^{n-2}$ . Par récurrence sur l'ordre de dérivée, on montre que  $S$  est de classe  $\mathcal{C}^k$  sur  $] -R, R[$ , pour tout  $k \in \mathbb{N}$ . Autrement dit,  $S$  est de classe  $\mathcal{C}^\infty$  sur  $] -R, R[$ .  $\square$

*Exemple 2.19.* La série entière  $\sum_n \frac{x^n}{n!}$  est de rayon de convergence  $R = +\infty$ . Sa somme  $f$  est donc  $\mathcal{C}^\infty$  sur  $\mathbb{R}$ . De plus,

$$f'(x) = \sum_{n=1}^{+\infty} n \frac{x^{n-1}}{n!} = \sum_{n=1}^{+\infty} \frac{x^{n-1}}{(n-1)!}.$$

Or, pour tout  $N \in \mathbb{N}^*$ ,

$$\sum_{n=1}^N \frac{x^{n-1}}{(n-1)!} = 1 + \frac{x}{1!} + \cdots + \frac{x^{N-1}}{(N-1)!} = \sum_{n=0}^{N-1} \frac{x^n}{n!}.$$

Comme les deux séries convergent, on peut donc passer à la limite  $N \rightarrow +\infty$  pour obtenir

$$\sum_{n=1}^{+\infty} \frac{x^{n-1}}{(n-1)!} = \sum_{n=0}^{+\infty} \frac{x^n}{n!}. \quad (2.2)$$

On a donc montré que  $f$  est solution de l'équation différentielle  $f' = f$ . De plus  $f(0) = 1$ . On vient de montrer l'existence d'une fonction  $f : \mathbb{R} \rightarrow \mathbb{R}$  telle que

$$f' = f \quad \text{et} \quad f(0) = 1.$$

Autrement dit, on vient de montrer l'existence de la fonction exponentielle.

Dans cet exemple, vous pouvez observer que pour établir l'égalité (2.2), on l'a d'abord justifiée pour les sommes partielles, avant de passer à la limite. C'est un fait général : pour toutes séries entières, on peut faire le changement d'indices<sup>3</sup> :

$$\sum_{n=1}^{+\infty} a_n x^{n-1} = \sum_{n=0}^{+\infty} a_{n+1} x^n.$$

*Exemple 2.20.* Si  $f(x) = \sum_{n=0}^{+\infty} x^n = \frac{1}{1-x}$ , alors

$$\sum_{n=1}^{+\infty} n x^{n-1} = f'(x) = \frac{1}{(1-x)^2}.$$

Une conséquence du Théorème 2.18 est que la somme d'une série entière est une fonction continue, et on peut donc considérer ses primitives.

**Corollaire 2.21.** Soit  $f : x \mapsto \sum_{n=0}^{\infty} a_n x^n$  la somme d'une série entière de rayon de convergence  $R > 0$ .

Alors la primitive de  $f$  sur  $] -R, R[$  qui s'annule en 0 est

$$F : x \mapsto \sum_{n=0}^{+\infty} a_n \frac{x^{n+1}}{n+1}.$$

*Démonstration.* On sait déjà (Proposition 2.16) que la série entière  $\sum_{n \geq 0} a_n \frac{x^{n+1}}{n+1}$  a pour rayon de convergence  $R$ . Par changement d'indices,

$$F(x) = \sum_{n=0}^{+\infty} a_n \frac{x^{n+1}}{n+1} = \sum_{n=1}^{+\infty} a_{n-1} \frac{x^n}{n}.$$

D'après le Théorème 2.18, la fonction  $F$  est de classe  $\mathcal{C}^\infty$  sur  $] -R, R[$ , et

$$F'(x) = \sum_{n=1}^{+\infty} a_{n-1} x^{n-1} = \sum_{n=0}^{+\infty} a_n x^n.$$

Ainsi, la fonction  $F$  est une primitive de  $f : x \mapsto \sum_{n=0}^{+\infty} a_n x^n$ . De plus,  $F(0) = 0$ . □

---

3. Et même plus généralement, pour tout  $n_1, k \in \mathbb{N}$  tel que  $k \leq n_1$ ,

$$\sum_{n=n_1}^{+\infty} a_n x^n = \sum_{n=n_1-k}^{+\infty} a_{n+k} x^{n+k}.$$

*Exemple 2.22.* Une primitive de  $f : x \mapsto \sum_{n=0}^{+\infty} x^n$  est la fonction  $F : x \mapsto \sum_{n=0}^{+\infty} \frac{x^{n+1}}{n+1} = \sum_{n=1}^{+\infty} \frac{x^n}{n}$ . Comme  $F'(x) = f(x) = \frac{1}{1-x}$  et comme  $F(0) = 0$ , on a

$$F(x) = -\ln(1-x).$$

On a donc montré que pour tout  $x \in ]-1, 1[$ , on a  $-\ln(1-x) = \sum_{n=1}^{+\infty} \frac{x^n}{n}$ ; soit encore, en faisant le changement de variables  $y = -x$ ,

$$\forall y \in \mathbb{R}, \quad \ln(1+y) = \sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{n} y^n.$$

### 2.2.3 Fonctions développables en séries entières

La formule de Taylor assure que si  $f$  est  $\mathcal{C}^\infty$  sur un intervalle ouvert  $I$  contenant 0, alors pour tout  $n \geq 0$ , quand  $x \rightarrow 0$ ,

$$f(x) = a_0 + a_1x + \cdots + a_nx^n + o(|x|^n) \quad \text{avec} \quad a_n = \frac{f^{(n)}(0)}{n!}.$$

Une question est alors de savoir pour quelles fonctions  $f$  et pour quels  $x$  on peut écrire

$$f(x) = \sum_{n=0}^{+\infty} a_nx^n.$$

Autrement dit, quand peut-on sommer jusqu'à l'infini ?

**Définition 2.23.** Soit  $f$  une fonction définie sur un intervalle ouvert  $I$  contenant un réel  $x_0$ . On dit que  $f$  est développable en série entière (DSE) en  $x_0$  sur  $]x_0 - R, x_0 + R[ \subseteq I$  s'il existe une série entière  $\sum_n a_ny^n$  de rayon de convergence  $\geq R$  tel que

$$\forall x \in ]x_0 - R, x_0 + R[, \quad f(x) = \sum_{n=0}^{+\infty} a_n(x - x_0)^n.$$

**Proposition 2.24.** Si  $f$  est développable en série entière (DSE) en  $x_0$  sur  $]x_0 - R, x_0 + R[ \subseteq I$ , alors  $f$  est de classe  $\mathcal{C}^\infty$  sur  $]x_0 - R, x_0 + R[$  et pour tout  $n \in \mathbb{N}$ ,

$$a_n = \frac{f^{(n)}(x_0)}{n!}.$$

*Démonstration.* Par hypothèse, on sait que

$$\forall x \in ]x_0 - R, x_0 + R[, \quad f(x) = \sum_{n=0}^{+\infty} a_n(x - x_0)^n.$$

Posons

$$g(y) = f(y + x_0).$$

Alors  $g$  est DSE en 0 sur  $] -R, R[$  et

$$\forall y \in ] -R, R[, \quad g(y) = \sum_{n=0}^{+\infty} a_n y^n.$$

Par le Théorème 2.18,  $g$  est de classe  $\mathcal{C}^\infty$  sur  $] -R, R[$ . Par composition, on en déduit que  $f$  est  $\mathcal{C}^\infty$  sur  $]x_0 - R, x_0 + R[$ . De plus, on sait qu'on peut dériver terme à terme :

$$\forall k \in \mathbb{N}^*, \quad \forall y \in ] -R, R[, \quad g^{(k)}(y) = \sum_{n=k}^{+\infty} a_n n(n-1) \cdots (n-k+1) y^{n-k}.$$

En particulier,  $g^{(k)}(0) = a_k k!$ . On en déduit par composition

$$a_k = \frac{f^{(k)}(x_0)}{k!}. \quad \square$$

Ainsi, les coefficients du développement en série entière sont entièrement déterminés par  $f$ , ils sont donc uniques :

$$f(x) = \sum_{n=0}^{+\infty} a_n (x - x_0)^n = \sum_{n=0}^{+\infty} b_n (x - x_0)^n \implies \forall n \in \mathbb{N}, \quad a_n = b_n.$$

Il existe des fonctions  $\mathcal{C}^\infty$  qui ne sont pas DSE. On en présente un exemple dans la troisième partie.

*Remarque 2.25.* Si  $f$  est DSE en 0 sur  $] -R, R[$  avec  $f(x) = \sum_{n=0}^{\infty} a_n x^n$ , alors  $f'$  est DSE en 0 sur  $] -R, R[$  avec

$$f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}.$$

De plus, la primitive  $F(x) = \int_0^x f(t) dt$  est DSE en 0 sur  $] -R, R[$  avec

$$F(x) = \sum_{n=0}^{+\infty} \frac{a_n}{n+1} x^{n+1}.$$

*Exemple 2.26.*

1.  $\exp x = \sum_{n=0}^{+\infty} \frac{x^n}{n!}, R = +\infty,$
2.  $\sin x = \sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}, R = +\infty,$
3.  $\cos x = \sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n}}{(2n)!}, R = +\infty,$
4.  $\frac{1}{1-x} = \sum_{n=0}^{+\infty} x^n, R = 1,$
5.  $(1+x)^\alpha = \sum_{n=0}^{+\infty} \frac{\alpha(\alpha-1) \cdots (\alpha-n+1)}{n!} x^n, R = 1,$

$$6. \ln(1+x) = \sum_{n=1}^{+\infty} (-1)^{n-1} \frac{x^n}{n}, R=1.$$

*Exercice 2.27.* Montrer que la fonction  $f(x) = (1+x)^\alpha$ , définie sur  $] -1, 1[$  est solution de l'équation différentielle

$$(1+x)y' - \alpha y = 0.$$

En déduire un DSE de  $f$ .

## 2.3 Approfondir

*Exercice 2.28* (Un exemple de fonction de classe  $C^\infty$  qui n'est pas DSE). Considérons la fonction  $f : \mathbb{R}^* \rightarrow \mathbb{R}$  définie par

$$f(x) := \exp\left(-\frac{1}{x^2}\right).$$

1. Justifier que  $f$  est de classe  $C^\infty$  sur  $\mathbb{R}^*$  puis montrer que pour tout  $n \in \mathbb{N}$ , il existe un polynôme  $P_n : \mathbb{R} \rightarrow \mathbb{R}$  tel que pour tout  $x \in \mathbb{R}^*$ ,

$$f^{(n)}(x) = P_n\left(\frac{1}{x}\right) \exp\left(-\frac{1}{x^2}\right).$$

On ne cherchera pas à expliciter ce polynôme  $P_n$ .

2. Montrer que  $f$  se prolonge par continuité en 0. On note encore  $f$  le prolongement. Montrer que la fonction  $f$  ainsi obtenue est dérivable en 0, et en fait  $C^1$  sur  $\mathbb{R}$ .
3. Montrer que la fonction  $f$  est même  $C^\infty$  sur  $\mathbb{R}$  et calculer toutes ses dérivées en 0.
4. Conclure que  $f$  n'est pas DSE en 0.

*Exercice 2.29* (Produit de Cauchy de séries). Soient  $\sum_k u_k$  et  $\sum_\ell v_\ell$  deux séries numériques complexes qu'on suppose absolument convergentes. Le but de l'exercice est de montrer que

$$\left(\sum_{k=0}^{+\infty} u_k\right) \left(\sum_{\ell=0}^{+\infty} v_\ell\right) = \sum_{m=0}^{+\infty} w_m \quad \text{avec} \quad w_m := \sum_{k=0}^m u_k v_{m-k}. \quad (2.3)$$

1. Montrer que  $\sum_{m=0}^n w_m = \sum_{\substack{(k,\ell) \in \mathbb{N}^2 \\ k+\ell \leq n}} u_k v_\ell$ .

2. Justifier les inégalités suivantes

$$\left| \sum_{\substack{(k,\ell) \in \mathbb{N}^2 \\ k+\ell \leq n}} u_k v_\ell - \left(\sum_{k=0}^n u_k\right) \left(\sum_{\ell=0}^n v_\ell\right) \right| \leq \sum_{\substack{(k,\ell) \in \mathbb{N}^2 \\ k \leq n, \ell \leq n, k+\ell > n}} |u_k v_\ell|$$

$$\leq \sup_{k \in \mathbb{N}} |u_k| \sum_{\ell=\lfloor \frac{n}{2} \rfloor}^{+\infty} |v_\ell| + \sup_{\ell \in \mathbb{N}} |v_\ell| \sum_{k=\lfloor \frac{n}{2} \rfloor}^{+\infty} |u_k|.$$

On a noté  $\lfloor \frac{n}{2} \rfloor$  la partie entière de  $\frac{n}{2}$ .

3. En déduire la convergence de la série  $\sum_m w_m$  et établir l'égalité (2.3).

A partir de l'exercice 2.29, on peut déduire :

**Proposition 2.30.** Soient  $\sum_n a_n x^n$  de rayon de convergence  $R_a$  et  $\sum_n b_n x^n$  de rayon de convergence  $R_b$ . Alors pour tout  $x \in ]-\min(R_a, R_b), \min(R_a, R_b)[$ ,

$$\left(\sum_{n=0}^{+\infty} a_n x^n\right) \left(\sum_{n=0}^{+\infty} b_n x^n\right) = \sum_{n=0}^{+\infty} c_n x^n \quad \text{avec} \quad c_n := \sum_{k=0}^n a_k b_{n-k}.$$

*Exercice 2.31.* Soit  $f : \mathbb{R} \rightarrow \mathbb{R}$  une fonction de classe  $\mathcal{C}^\infty$ . On suppose qu'il existe  $a > 0$  et  $M > 0$  tel que

$$\forall t \in [-a, a], \quad \forall n \in \mathbb{N}, \quad |f^{(n)}(t)| \leq \frac{Mn!}{a^n}.$$

Montrer que  $f$  est développable en série entière sur  $]-a, a[$  (penser à la formule de Taylor avec reste intégral).

La partie II du programme du concours Passingénieurs (pour l'option Mathématiques-Informatique) considère les séries entières de la variable complexe  $z$ . Etant donnée une suite de nombres complexes  $(a_n)_{n \in \mathbb{N}}$ , on considère la série de nombres complexes

$$\sum_{n \in \mathbb{N}} a_n z^n.$$

Dans ce nouveau contexte, on parle encore de rayon de convergence : il s'agit de l'unique  $R \in [0, +\infty[$  tel que pour tout  $z \in \mathbb{C}$ ,

$$|z| < R \implies \sum_n a_n z^n \text{ converge absolument,}$$

$$|z| > R \implies \sum_n a_n z^n \text{ diverge.}$$

La convergence absolue est définie maintenant avec le module, au lieu de la valeur absolue. On peut étendre à ce cadre une bonne partie des résultats précédents (grosso modo, dès qu'ils ont un sens, ils sont vrais !). En particulier, pour tout  $r < R$ , la série converge normalement sur le disque fermé centré en 0 de rayon  $r$ , c'est-à-dire sur  $\overline{D}_r := \{z \in \mathbb{C} : |z| \leq r\}$  :

$$\sum_n \sup_{|z| \leq r} |a_n z^n| = \sum_n |a_n| r^n \text{ converge.}$$

De plus, la fonction  $f : z \mapsto \sum_{n=0}^{+\infty} a_n z^n$  est continue sur le disque ouvert  $D_R = \{z \in \mathbb{C} : |z| < R\}$ .

Cela signifie que pour tout  $z \in D_R$ , pour toute suite  $(z_k)_{k \in \mathbb{N}}$  contenue dans  $D_R$  et convergeant vers  $z$ , la suite  $(f(z_k))_{k \in \mathbb{N}}$  converge vers  $f(z)$ . On peut aussi l'écrire avec des quantificateurs :

$$\forall \varepsilon > 0, \quad \exists \eta > 0, \quad \forall w \in \mathbb{C}, \quad |w - z| < \eta \implies |f(w) - f(z)| \leq \varepsilon.$$

Si  $\sum_n a_n z^n$  et  $\sum_n b_n z^n$  ont pour rayons de convergence respectifs  $R_a$  et  $R_b$ , alors la série entière  $\sum_n (a_n + b_n) z^n$  aura un rayon de convergence  $R \geq \min(R_a, R_b)$ , avec égalité si  $R_a \neq R_b$ .

L'exponentielle complexe  $z = x + iy \mapsto \exp(z) = e^x e^{iy}$  est DSE sur  $\mathbb{C}$  avec

$$\exp(z) = \sum_{n=0}^{+\infty} \frac{z^n}{n!}.$$

La fonction  $z \mapsto \frac{1}{1-z}$  est aussi DSE en 0, de rayon de convergence 1 avec pour tout  $z \in \mathbb{C}, |z| < 1$  :

$$\frac{1}{1-z} = \sum_{n=0}^{+\infty} z^n.$$

Les fonctions  $\cos z$  et  $\sin z$  sont également DSE en 0, et ont des développements identiques à ceux du cas réel.



# Chapitre 3

## Séries de Fourier

### 3.1 Découvrir

#### 3.1.1 L'analyse de Fourier et la compression d'images

En 1809, dans son *Mémoire sur la propagation de la chaleur*, Fourier s'intéresse au problème suivant : l'évolution de la température d'un anneau circulaire métallique dont les deux moitiés sont initialement à des températures différentes. Il énonce alors un principe tout à fait surprenant : toute fonction périodique peut être représentée par une série de sinus et de cosinus. C'est Dirichlet qui en 1829 donnera des bases solides à la théorie de Fourier. Cette théorie est un outil fondamental dans l'étude des équations aux dérivées partielles comme l'équation de la chaleur ou l'équation des ondes.

La première série qu'exhibe Fourier est :

$$\cos x - \frac{1}{3} \cos(3x) + \frac{1}{5} \cos(5x) - \frac{1}{7} \cos(7x) + \dots = \begin{cases} +\pi/4 & \text{pour } |x| < \pi/2 \\ 0 & \text{pour } |x| = \pi/2 \\ -\pi/4 & \text{pour } \pi/2 < |x| < \pi \end{cases}$$

(c'est en tout cas ce qu'il prétend). Un des objectifs de ce cours est de donner les outils qui permettent de justifier cette affirmation.

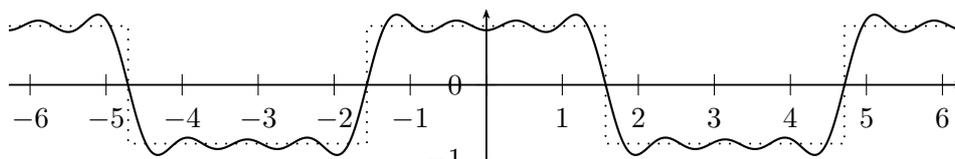


FIGURE 3.1 – La somme des quatre premiers termes de la série  $\sum_{n \geq 0} \frac{(-1)^n}{2n+1} \cos((2n+1)x)$  (en trait continu) et de la fonction  $2\pi$ -périodique qui vaut  $\pi/4$  pour  $|x| < \pi/2$  et  $-\pi/4$  pour  $\pi/2 < |x| < \pi$  (en pointillés).

Pour stocker une image numérique, produite par un appareil photo de  $3 \times 3$  mégapixels, on aurait théoriquement besoin de 9 Mo sur un ordinateur, ce qui limiterait considérablement le nombre de photos qu'on pourrait effectivement stocker. Heureusement, on peut compresser les images numériques (ou les sons, les films) sans trop altérer leur qualité visuelle. Le JPEG, format de compression mis au point dans les années 80, permet de réduire jusqu'à 20 fois la place occupée par une image avec une perte de qualité presque imperceptible.

Lorsqu'un signal est représenté comme une somme de sinusoides, chaque coefficient de cette somme représente le poids d'une certaine fréquence dans le signal. Pour stocker le signal, il suffit donc de conserver ces coefficients. Plus une fréquence dans la somme est associée à un coefficient petit, moins nous la

percevons. On peut donc réduire considérablement la quantité de données à stocker ou à transmettre en éliminant les coefficients trop faibles.

### 3.1.2 Intégrales de fonctions à valeurs complexes

Dans ce chapitre, toutes les fonctions considérées sont à valeurs dans  $\mathbb{C}$  (ce qui n'exclut pas les fonctions qui ne prennent que des valeurs réelles). Une telle fonction est continue si et seulement si sa partie réelle et sa partie imaginaire le sont :

$$\operatorname{Re} f : x \mapsto \operatorname{Re}(f(x)) \quad , \quad \operatorname{Im} f : x \mapsto \operatorname{Im}(f(x)).$$

On peut définir l'intégrale d'une telle fonction en décomposant  $f$  en partie réelle et partie imaginaire :  $f = \operatorname{Re} f + i \operatorname{Im} f$  :

$$\int_a^b f(x) dx = \int_a^b \operatorname{Re} f(x) dx + i \int_a^b \operatorname{Im} f(x) dx.$$

L'intégrale des fonctions complexes hérite de toutes les propriétés de l'intégrale des fonctions réelles (linéarité, relation de Chasles). On vérifie facilement que

$$\int_a^b \overline{f(t)} dt = \overline{\int_a^b f(t) dt}.$$

### 3.1.3 Polynômes trigonométriques à coefficients complexes

Dans tout ce chapitre, on se donne un réel  $T > 0$  et on pose

$$\omega := \frac{2\pi}{T}.$$

On note  $\mathcal{C}_T$  l'espace des fonctions continues  $f : \mathbb{R} \rightarrow \mathbb{C}$  qui sont  $T$ -périodiques :

$$\forall x \in \mathbb{R}, \quad f(x + T) = f(x).$$

On peut voir  $f \in \mathcal{C}_T$  comme un signal périodique qu'on cherche à transmettre après compression. On va chercher à l'écrire comme une somme de sinusôides, c'est-à-dire de fonctions de la forme  $t \mapsto a \cos(n\omega t)$  et  $t \mapsto b \sin(n\omega t)$ . Cette somme sera généralement *infinie*. Ainsi,  $f$  sera la somme d'une série de fonctions dont chaque terme est une sinusôide. On va commencer par considérer des sommes *finies* de fonctions. C'est ce qu'on appelle des *polynômes trigonométriques*.

**Définition 3.1.** Les fonctions exponentielles  $e_n \in \mathcal{C}_T$  sont définies par

$$e_n(t) := e^{i\omega n t}.$$

**Définition 3.2.** On appelle polynôme trigonométrique de degré  $N$  et de période  $T$  toute fonction de la forme :

$$P : t \mapsto \sum_{n=-N}^N c_n e^{i\omega n t} \quad \text{avec} \quad c_n \in \mathbb{C}.$$

On appelle série trigonométrique toute série de la forme :

$$c_0 + \sum_{n \geq 1} (c_n e^{i\omega n t} + c_{-n} e^{-i\omega n t}) \quad \text{avec} \quad c_n \in \mathbb{C}.$$

Pour simplifier l'écriture, nous utiliserons la notation  $\sum_{n \in \mathbb{Z}} c_n e_n$  pour désigner la série trigonométrique  $c_0 + \sum_{n \geq 1} (c_n e^{i\omega n t} + c_{-n} e^{-i\omega n t})$ .

Si on veut transmettre la fonction (le signal)  $P$ , il est évidemment impossible de communiquer les nombres  $P(t)$  pour toutes les valeurs possibles de  $t$  (puisque'il y a une infinité de réels dans  $[0, 2\pi]$ )! En revanche, on peut se contenter d'envoyer les valeurs  $c_n$  pour tout  $|n| \leq N$  et le destinataire pourra alors reconstruire le polynôme  $P$ . Réciproquement, étant donné un signal périodique  $P$ , on peut calculer les coefficients  $c_n$  grâce à des intégrales faisant intervenir la fonction  $P$ .

Le lemme suivant sera fréquemment utilisé.

**Lemme 3.3.** *On a*

$$\frac{1}{T} \int_0^T e_0(t) dt = 1 \quad \text{et} \quad \forall n \in \mathbb{Z} \setminus \{0\}, \quad \frac{1}{T} \int_0^T e_n(t) dt = 0.$$

*Démonstration.* D'une part

$$\frac{1}{T} \int_0^T e_0(t) dt = \frac{1}{T} \int_0^T dt = 1.$$

D'autre part, si  $n \neq 0$ ,

$$\frac{1}{T} \int_0^T e_n(t) dt = \frac{1}{T} \left[ \frac{e^{i\omega n t}}{i\omega n} \right]_0^T = \frac{e^{i2\pi n} - 1}{i2\pi n} = 0. \quad \square$$

**Proposition 3.4.** *Si  $P(t) = \sum_{n=-N}^N c_n e^{i\omega n t}$ , alors*

$$c_n = \frac{1}{T} \int_0^T P(t) e^{-i\omega n t} dt.$$

*Démonstration.* Par linéarité de l'intégrale, et d'après le lemme précédent, on a

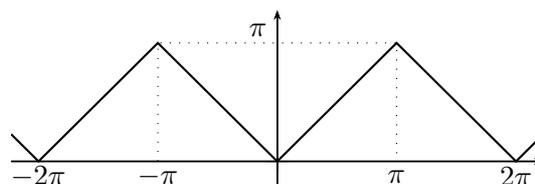
$$\frac{1}{T} \int_0^T P(t) e^{-i\omega n t} dt = \sum_{k=-N}^N \frac{c_k}{T} \int_0^T e^{i\omega(k-n)t} dt = \sum_{k=-N}^N \frac{c_k}{T} \int_0^T e_{k-n}(t) dt = c_n. \quad \square$$

### 3.1.4 Série de Fourier d'une fonction $T$ -périodique

**Définition 3.5.** *Si  $f \in \mathcal{C}_T$ , les coefficients de Fourier de  $f$  sont :*

$$c_n(f) := \frac{1}{T} \int_0^T f(t) e^{-i\omega n t} dt.$$

*Exercice 3.6.* Considérons la fonction  $f \in \mathcal{C}_{2\pi}$  définie par  $f(x) = |x|$  si  $x \in [-\pi, \pi]$  et prolongée à  $\mathbb{R}$  par périodicité. Déterminer les coefficients de Fourier de  $f$ .



**Définition 3.7.** Si  $f \in \mathcal{C}_T$ , la série de Fourier de  $f$  est la série trigonométrique  $\sum_{n \in \mathbb{Z}} c_n(f) e_n$ . La  $N$ -ème somme partielle de la série de Fourier de  $f$  est :

$$S_N(f) : t \mapsto \sum_{n=-N}^N c_n(f) e^{i\omega n t}.$$

Lorsqu'une fonction  $f \in \mathcal{C}_T$  est à valeurs réelles, il est fréquent d'utiliser une autre écriture des sommes partielles de Fourier. En effet, on peut toujours écrire :

$$e_n(t) = \cos(\omega n t) + i \sin(\omega n t).$$

Par conséquent :

$$\sum_{n=-N}^N c_n e^{i\omega n t} = \frac{a_0}{2} + \sum_{n=1}^N a_n \cos(\omega n t) + \sum_{n=1}^N b_n \sin(\omega n t)$$

avec, pour  $n \geq 0$  :

$$a_n = c_n + c_{-n} \quad \text{et} \quad b_n = i(c_n - c_{-n}).$$

Cela conduit à la définition suivante.

**Définition 3.8.** Si  $f \in \mathcal{C}_T$ , on pose, pour  $n \geq 0$  :

$$a_n(f) = \frac{2}{T} \int_0^T f(t) \cos(\omega n t) dt \quad \text{et} \quad b_n(f) = \frac{2}{T} \int_0^T f(t) \sin(\omega n t) dt.$$

La  $N$ -ème somme partielle de la série de Fourier d'une fonction  $f \in \mathcal{C}_T$  est donc :

$$S_N(f) : t \mapsto \frac{a_0(f)}{2} + \sum_{n=1}^N a_n(f) \cos(\omega n t) + \sum_{n=1}^N b_n(f) \sin(\omega n t).$$

### 3.1.5 Propriétés des coefficients de Fourier

**Proposition 3.9.** On a les relations suivantes entre coefficients de Fourier :

1. Coefficients de  $\bar{f}$  :  $c_n(\bar{f}) = \overline{c_{-n}(f)}$  et si  $f$  est une fonction à valeurs réelles,  $c_n(f) = \overline{c_{-n}(f)}$ .
2. Coefficients de  $g : t \mapsto f(-t)$  :  $c_n(g) = c_{-n}(f)$ . Si  $f$  est paire, alors  $c_{-n}(f) = c_n(f)$  et si  $f$  est impaire  $c_{-n}(f) = -c_n(f)$ .
3. Coefficients de  $h : t \mapsto f(t+a)$  :  $c_n(h) = e^{i\omega n a} c_n(f)$ .

*Démonstration.* Commençons par le premier point :

$$c_n(\bar{f}) = \frac{1}{T} \int_0^T \bar{f}(t) e^{-i\omega n t} dt = \overline{\frac{1}{T} \int_0^T f(t) e^{i\omega n t} dt} = \overline{c_{-n}(f)}.$$

Pour le deuxième point, il suffit de faire le changement de variables  $u = T - t$ . Observons que l'on a  $g(t) = f(-t) = f(T - t) = f(u)$ , et donc

$$c_n(g) = \frac{1}{T} \int_0^T g(t) e^{-i\omega n t} dt = -\frac{1}{T} \int_T^0 f(u) e^{i\omega n u} du = \frac{1}{T} \int_0^T f(u) e^{i\omega n u} du = c_{-n}(f).$$

Si  $f$  est paire, alors  $g = f$  et si  $f$  est impaire, alors  $g = -f$ . Le résultat en découle.

Pour le troisième point, on utilise le changement de variable  $u = t + a$ . On a  $h(t) = f(t+a) = f(u)$ , et donc :

$$c_n(h) = \frac{1}{T} \int_0^T g(t) e^{-i\omega n t} dt = \frac{1}{T} \int_a^{T+a} f(u) e^{-i\omega n(u-a)} du = \frac{e^{i\omega n a}}{T} \int_a^{T+a} f(u) e^{-i\omega n u} du.$$

La fonction  $u \mapsto f(u) e^{-i\omega n u}$  étant  $T$ -périodique, on a

$$\frac{e^{i\omega n a}}{T} \int_a^{T+a} f(u) e^{-i\omega n(u-a)} du = \frac{e^{i\omega n a}}{T} \int_0^T f(u) e^{-i\omega n u} du = e^{i\omega n a} c_n(f). \quad \square$$

**Corollaire 3.10.** Si  $f \in \mathcal{C}_T$  est une fonction à valeurs réelles, les coefficients  $a_n(f)$  et  $b_n(f)$  sont réels (ce qui n'est pas nécessairement le cas des coefficients  $c_n(f)$ ).

**Corollaire 3.11.** La fonction  $f$  est une fonction paire si, et seulement si, les coefficients  $b_n(f)$  sont tous nuls. La fonction  $f$  est une fonction impaire si, et seulement si, les coefficients  $a_n(f)$  sont tous nuls.

*Exemple 3.12.* Si l'on reprend l'exemple de la fonction  $f \in \mathcal{C}_{2\pi}$  qui coïncide avec  $x \mapsto |x|$  sur  $[-\pi, \pi]$ , on peut écrire les sommes partielles de la série de Fourier sous la forme :

$$S_N(f)(t) = \frac{\pi}{2} + \frac{4}{\pi} \sum_{p=0}^{\lfloor \frac{N-1}{2} \rfloor} \frac{1}{(2p+1)^2} \cos((2p+1)t).$$

La fonction  $f$  est paire. La série de Fourier ne fait intervenir que des cosinus.

**Proposition 3.13.** Si  $f \in \mathcal{C}_T$  est une fonction de classe  $\mathcal{C}^1$ , alors :

$$c_n(f') = i\omega n c_n(f).$$

*Démonstration.* Il suffit de faire une intégration par parties :

$$\begin{aligned} c_n(f') &= \frac{1}{T} \int_0^T f'(t) e^{-i\omega n t} dt = \frac{1}{T} [f(t) e^{-i\omega n t}]_0^T - \frac{1}{T} \int_0^T f(t) (-i\omega n) e^{-i\omega n t} dt \\ &= 0 + i\omega n c_n(f). \quad \square \end{aligned}$$

Par récurrence, on obtient le corollaire suivant.

**Corollaire 3.14.** Si  $f \in \mathcal{C}_T$  est une fonction de classe  $\mathcal{C}^\infty$ , alors pour tout  $k \geq 1$  :

$$c_n(f^{(k)}) = (i\omega n)^k c_n(f).$$

### 3.1.6 Taille des coefficients de Fourier

**Proposition 3.15.** Si  $f \in \mathcal{C}_T$ , alors pour tout  $n \in \mathbb{Z}$ , on a

$$|c_n(f)| \leq \frac{1}{T} \int_0^T |f(t)| dt.$$

La suite  $(c_n(f))$  est donc bornée. En fait, nous établirons ultérieurement le résultat suivant.

**Proposition 3.16.** Si  $f \in \mathcal{C}_T$ , alors  $c_n(f)$  tend vers 0 quand  $n \rightarrow \pm\infty$ .

**Corollaire 3.17.** Si  $f \in \mathcal{C}_T$  est de classe  $\mathcal{C}^k$ , alors quand  $n \rightarrow \pm\infty$ ,

$$c_n(f) = o\left(\frac{1}{n^k}\right).$$

**Corollaire 3.18.** Si  $f \in \mathcal{C}_T$  est de classe  $\mathcal{C}^\infty$ , alors quand  $n \rightarrow \pm\infty$ , pour tout  $k \geq 0$ , on a

$$c_n(f) = o\left(\frac{1}{n^k}\right).$$

### 3.1.7 Problème inverse

Supposons maintenant que  $\{c_n\}_{n \in \mathbb{Z}}$  soit une famille de nombres complexes. La suite de fonctions  $S_N := \sum_{n=-N}^N c_n \cdot e_n$  est-elle convergente ? Si oui, les  $c_n$  sont-ils les coefficients de Fourier de la limite ?

**Proposition 3.19.** Si  $(c_n)_{n \in \mathbb{Z}}$  est une suite de nombres complexes tels que  $\sum_{n \in \mathbb{Z}} |c_n| < +\infty$ , alors la suite

$\left( S_N := \sum_{n=-N}^N c_n e_n \right)$  converge normalement sur  $\mathbb{R}$ .

*Démonstration.* Pour tout  $n \in \mathbb{Z}$ , on a

$$\|c_n e_n\|_\infty = \|c_n e^{i\omega n t}\|_\infty = |c_n|.$$

Par hypothèse, la série  $\sum_{n \in \mathbb{Z}} |c_n|$  est convergente. Par conséquent, la série de fonctions  $\sum_{n \in \mathbb{Z}} c_n e_n$  converge normalement sur  $\mathbb{R}$ . En particulier la suite des sommes partielles  $(S_N)$  converge uniformément sur  $\mathbb{R}$ .  $\square$

**Proposition 3.20.** Si la suite  $\left( S_N := \sum_{n=-N}^N c_n e_n \right)$  converge normalement vers  $f$ , alors  $f \in \mathcal{C}_T$  et pour tout  $n \in \mathbb{Z}$ ,  $c_n(f) = c_n$ .

*Démonstration.* Comme chaque fonction  $c_n e_n$  est continue, on en déduit que  $f$  est également continue sur  $\mathbb{R}$ . Par convergence simple, la  $T$ -périodicité de  $S_N$  entraîne celle de  $f$ . Ainsi  $f \in \mathcal{C}_T$ .

Calculons les coefficients de Fourier de  $f$ . Par convergence uniforme de  $S_N$  vers  $f$  sur  $[0, T]$ , on a d'après la Proposition 3.4 :

$$\begin{aligned} c_n(f) &= \frac{1}{T} \int_0^T \left( \lim_{N \rightarrow +\infty} S_N(t) \right) e^{-i\omega n t} dt \\ &= \lim_{N \rightarrow +\infty} \frac{1}{T} \int_0^T S_N(t) e^{-i\omega n t} dt = \lim_{N \rightarrow +\infty} c_n = c_n. \quad \square \end{aligned}$$

**Proposition 3.21.** Si la suite des dérivées  $\left( S'_N := \sum_{n=-N}^N i\omega n c_n e_n \right)$  converge normalement vers  $g$ , alors

la suite  $\left( S_N := \sum_{n=-N}^N c_n e_n \right)$  converge normalement vers  $f$  avec  $f$  de classe  $\mathcal{C}^1$  et  $f' = g$ .

*Démonstration.* Pour  $n \neq 0$ , on a

$$\|c_n e_n\|_\infty = |c_n| \leq |n c_n| = \|i\omega n c_n e_n\|_\infty.$$

La convergence normale de la suite des dérivées  $S'_N$  implique donc la convergence normale de la suite  $S_N$ . On conclut grâce à la Proposition 1.31.  $\square$

### 3.1.8 Objectif du chapitre

La question fondamentale est de savoir si la série de Fourier d'une fonction  $f$  converge, autrement dit, est-ce qu'on peut écrire en un certain sens que

$$f(x) = \sum_{n \in \mathbb{Z}} c_n(f) e^{i\omega n x} \quad ?$$

On va voir que la réponse dépend de la *régularité* de  $f$ . Nous verrons en particulier que si  $f \in \mathcal{C}_T$  est une fonction de classe  $\mathcal{C}^1$ , alors  $f$  est la somme de sa série de Fourier.

Attention, il se peut que  $f$  soit continue et que sa série de Fourier ne converge pas normalement. En fait, il se peut même que sa série de Fourier ne converge pas.

## 3.2 Assimiler

### 3.2.1 Convergence uniforme et convergence ponctuelle

#### La norme uniforme

On peut introduire sur  $\mathcal{C}_T$  la *norme uniforme*  $\|\cdot\|_\infty$ , où

$$\forall f \in \mathcal{C}_T, \quad \|f\|_\infty = \sup_{x \in \mathbb{R}} |f(x)|.$$

En effet, si une fonction  $f$  est continue  $T$ -périodique, alors l'ensemble des valeurs qu'elle prend coïncide avec l'ensemble des valeurs qu'elle prend sur  $[0, T]$  (on utilise ici la  $T$ -périodicité) :

$$f(\mathbb{R}) = \{f(x); x \in \mathbb{R}\} = f([0, T]).$$

Sur le segment  $[0, T]$ , la fonction  $f$  est continue, donc est bornée et atteint ses bornes. Autrement dit, il existe  $a, b \in [0, T]$  tels que pour tout  $x \in [0, T]$ ,

$$f(a) \leq f(x) \leq f(b).$$

Cette double inégalité reste vraie pour tout  $x \in \mathbb{R}$ . On peut donc définir

$$\|f\|_\infty = \max_{x \in [0, T]} |f(x)| = \max_{x \in \mathbb{R}} |f(x)|.$$

On aura donc ici  $\|f\|_\infty = \max(|f(a)|, |f(b)|)$ .

Pour en savoir plus sur la notion de norme, on se reportera à la section *Approfondir*.

### 3.2.2 Produit de convolution

Commençons par manipuler les sommes partielles de la série de Fourier pour les écrire sous une forme plus agréable à manipuler :

$$\begin{aligned} S_N(f)(x) &= \sum_{n=-N}^N c_n(f) e_n(x) \\ &= \sum_{n=-N}^N \left( \frac{1}{T} \int_0^T f(t) e^{-i\omega n t} dt \right) e^{i\omega n x} = \frac{1}{T} \int_0^T f(t) \cdot \left( \sum_{n=-N}^N e^{i\omega n(x-t)} \right) dt. \end{aligned}$$

On a donc :

$$S_N(f)(x) = \frac{1}{T} \int_0^T f(t) \cdot D_N(x-t) dt \quad \text{avec} \quad D_N(t) = \sum_{n=-N}^N e_n.$$

Une notion s'impose alors naturellement, la notion de produit de convolution.

**Définition 3.22.** Si  $f \in \mathcal{C}_T$  et  $g \in \mathcal{C}_T$ , le produit de convolution de  $f$  par  $g$  est la fonction :

$$f * g : x \mapsto \frac{1}{T} \int_0^T f(t) \cdot g(x-t) dt.$$

Nous allons dans un premier temps étudier quelques propriétés de ce produit de convolution.

**Proposition 3.23.** Soient  $f, g$  des fonctions de  $\mathcal{C}_T$ . Alors :

1.  $f * g \in \mathcal{C}_T$ .
2.  $f * g = g * f$ .
3. L'application  $h \mapsto f * h$  est une application linéaire de  $\mathcal{C}_T$  dans  $\mathcal{C}_T$ .
4. Pour tout  $n \in \mathbb{Z}$ ,  $f * e_n = c_n(f) \cdot e_n$ .

*Démonstration.* 1. La périodicité de  $f * g$  découle de la périodicité de  $g$  :  $g(x + T - t) = g(x - t)$ .  
On admet la continuité de  $f * g$  qui découle d'un théorème de continuité d'une fonction définie par une intégrale : la fonction  $(x, t) \mapsto f(t)g(x - t)$  est continue sur  $[0, T] \times [0, T]$  et donc,  $x \mapsto \int_0^T f(t)g(x - t)dt$  est continue sur  $[0, T]$  (donc sur  $\mathbb{R}$  par périodicité).  
2. L'égalité  $f * g = g * f$  s'obtient en opérant le changement de variable  $u = x - t$  et en utilisant la périodicité de  $f$  et  $g$  :

$$\begin{aligned} f * g(x) &= \frac{1}{T} \int_0^T f(t)g(x - t)dt \\ &= -\frac{1}{T} \int_x^{x-T} f(x - u)g(u)du \\ &= \frac{1}{T} \int_0^T g(u)f(x - u)du = g * f(x). \end{aligned}$$

3. La linéarité de l'application  $h \mapsto f * h$  est une conséquence immédiate de la linéarité de l'intégrale.
4. Si  $n \in \mathbb{Z}$  :

$$f * e_n(x) = \frac{1}{T} \int_0^T f(t)e^{i\omega n(x-t)}dt = \left( \frac{1}{T} \int_0^T f(t)e^{-i\omega nt}dt \right) e^{i\omega nx} = c_n(f) \cdot e_n(x). \quad \square$$

**Corollaire 3.24.** Si  $P := \sum_{n=-N}^N a_n e_n$  est un polynôme trigonométrique, alors  $f * P$  est également un polynôme trigonométrique.

### 3.2.3 Unité approchée

Un résultat fondamental est celui du produit de convolution avec une unité approchée de  $\mathcal{C}_T$ .

**Définition 3.25.** Une unité approchée de  $\mathcal{C}_T$  est une suite  $(h_n)$  de fonctions positives de  $\mathcal{C}_T$  telle que :

- pour tout  $0 < \delta < T/2$  et tout  $\varepsilon > 0$ , on a  $0 \leq h_n \leq \varepsilon$  sur l'intervalle  $[\delta, T - \delta]$  pour  $n$  assez grand et
- $\frac{1}{T} \int_0^T h_n = 1$  pour tout  $n$ .

Notons que si  $(h_n)$  est une unité approchée de  $\mathcal{C}_T$ , toute la masse de l'intégrale de  $h_n$  se concentre autour des multiples de  $T$  : pour tout  $\delta \in ]0, T/2[$ , on a :

$$\lim_{n \rightarrow +\infty} \frac{1}{T} \int_\delta^{T-\delta} h_n = 0 \quad \text{et} \quad \lim_{n \rightarrow +\infty} \frac{1}{T} \int_{-\delta}^\delta h_n = 1.$$

Exercice 3.26. Pour  $n \geq 0$ , posons :

$$h_n(t) := \frac{1}{A_n} \cos^{2n}(\omega t/2) \quad \text{avec} \quad A_n := \frac{1}{T} \int_0^T \cos^{2n}(\omega t/2) dt.$$

Montrer que  $h_n$  est un polynôme trigonométrique et que  $(h_n)$  est une unité approchée de  $\mathcal{C}_T$ .

**Proposition 3.27.** Si  $f \in \mathcal{C}_T$  et si  $(h_n)$  est une unité approchée de  $\mathcal{C}_T$ , alors la suite  $(f * h_n)$  converge uniformément vers  $f$  sur  $\mathbb{R}$  : pour tout  $\varepsilon > 0$ , on a

$$|f * h_n - f| \leq \varepsilon \quad \text{pour } n \text{ assez grand.}$$

*Démonstration.* Si  $f$  est identiquement nulle,  $f * h_n$  est identiquement nulle pour tout  $n$ , et il n'y a rien à montrer. Supposons donc que  $f$  n'est pas identiquement nulle, ce qui implique  $2c_0(|f|) > 0$ .

Fixons  $\varepsilon > 0$ . La fonction  $f$  est continue et périodique sur  $\mathbb{R}$ . Elle est donc uniformément continue. Par conséquent, il existe  $\delta \in ]0, T/2[$  tel que :

$$|f(x) - f(y)| \leq \varepsilon \quad \text{dès que} \quad |x - y| \leq 2\delta.$$

Choisissons  $n_0$  suffisamment grand pour que :

$$h_n(t) \leq \frac{\varepsilon}{2c_0(|f|)} \quad \text{dès que} \quad n \geq n_0 \quad \text{et} \quad t \in [\delta, T - \delta].$$

Alors, pour  $n \geq n_0$  :

$$\begin{aligned} |f * h_n(x) - f(x)| &= \left| \frac{1}{T} \int_0^T f(x-t)h_n(t)dt - \frac{1}{T} \int_0^T f(x)h_n(t)dt \right| \\ &\leq \frac{1}{T} \int_0^T |f(x-t) - f(t)| \cdot h_n(t)dt \\ &= \frac{1}{T} \int_{-\delta}^{\delta} \underbrace{|f(x-t) - f(t)|}_{\leq \varepsilon} \cdot h_n(t)dt \\ &\quad + \frac{1}{T} \int_{\delta}^{T-\delta} |f(x-t) - f(t)| \cdot \underbrace{h_n(t)}_{\leq \varepsilon/2c_0(|f|)} dt \\ &\leq \underbrace{\varepsilon \frac{1}{T} \int_{-\delta}^{\delta} h_n(t)dt}_{\leq 1} + \frac{\varepsilon}{2c_0(|f|)} \underbrace{\frac{1}{T} \int_{\delta}^{T-\delta} |f(x-t)| + |f(t)| dt}_{\leq 2c_0(|f|)} \leq 2\varepsilon. \quad \square \end{aligned}$$

### 3.2.4 Formule de Parseval

Avant de régler le cas de la convergence de la série de Fourier, nous allons d'ores et déjà utiliser le résultat que nous venons d'établir pour montrer un théorème de Weierstraß.

**Théorème 3.28** (Weierstraß). Toute fonction continue  $f \in \mathcal{C}_T$  est limite uniforme sur  $\mathbb{R}$  de polynômes trigonométriques.

*Démonstration.* Posons

$$h_N(t) = \frac{1}{A_N} \cos^{2N}(\omega t/2) \quad \text{avec} \quad A_N = \frac{1}{T} \int_0^T \cos^{2N}(\omega t/2) dt.$$

Les  $h_N$  sont des polynômes trigonométriques, et donc  $h_N * f$  est un polynôme trigonométrique. La suite  $(h_N)$  est une unité approchée de  $\mathcal{C}_T$  et donc on a la convergence uniforme  $h_N * f \rightarrow f$ .  $\square$

On peut munir  $\mathcal{C}_T$  du produit scalaire complexe :

$$\langle f, g \rangle := \frac{1}{T} \int_0^T \bar{f}(t)g(t)dt.$$

Pour en savoir plus sur la notion de produit scalaire complexe, on se reportera à la section *Approfondir*.

Alors, pour  $n \in \mathbb{Z}$ , les exponentielles  $e_n$  forment une famille orthonormale de  $\mathcal{C}_T$ . En effet, d'après le Lemme 3.3, pour tout  $m, n \in \mathbb{Z}$  :

$$\langle e_n, e_m \rangle = \frac{1}{T} \int_0^T e^{i\omega(m-n)t} dt = \frac{1}{T} \int_0^T e_{m-n}(t) dt = \begin{cases} 1 & \text{si } m = n \\ 0 & \text{si } m \neq n. \end{cases}$$

Si  $f \in \mathcal{C}_T$  et si  $J$  est une partie finie de  $\mathbb{Z}$ , le projeté orthogonal de  $f$  sur l'espace vectoriel engendré par  $\{e_j\}_{j \in J}$  est donc donné par :

$$\sum_{j \in J} c_j(f) \cdot e_j \quad \text{avec} \quad c_j(f) := \langle e_j, f \rangle.$$

La série de Fourier  $S_N(f)$  est donc le projeté orthogonal de  $f$  sur l'espace vectoriel

$$\text{Vect}(e_{-N}, \dots, e_{-1}, e_0, e_1, \dots, e_N).$$

La densité des polynômes trigonométriques permet d'établir l'égalité suivante, qui est une généralisation en dimension infinie du théorème de Pythagore (dont nous ne donnerons pas la démonstration).

**Théorème 3.29** (Formule de Parseval). *Pour toute fonction  $f \in \mathcal{C}_T$ , on a :*

$$\frac{1}{T} \int_0^T |f|^2 = \sum_{n \in \mathbb{Z}} |c_n(f)|^2.$$

*Exemple 3.30.* Dans le cas de la fonction  $f \in \mathcal{C}_{2\pi}$  qui coïncide avec  $x \mapsto |x|$  sur  $[-\pi, \pi]$ , on a  $c_0(f) = \frac{\pi}{2}$  et pour  $p \geq 1$ ,  $c_{\pm 2p}(f) = 0$  et  $c_{\pm(2p-1)}(f) = -\frac{2}{\pi(2p-1)^2}$ . On a donc :

$$\frac{\pi^2}{4} + 2 \sum_{p=0}^{+\infty} \frac{4}{\pi^2(2p+1)^4} = \frac{1}{2\pi} \int_{-\pi}^{\pi} |t|^2 dt = \frac{\pi^2}{3}.$$

On en déduit que :

$$\frac{8}{\pi^2} \sum_{p=0}^{+\infty} \frac{1}{(2p+1)^4} = \frac{\pi^2}{3} - \frac{\pi^2}{4} = \frac{\pi^2}{12} \quad \text{et} \quad \sum_{p=0}^{+\infty} \frac{1}{(2p+1)^4} = \frac{\pi^4}{96}.$$

En corollaire, on obtient la Proposition 3.16 que nous devons démontrer.

**Corollaire 3.31.** *Si  $f \in \mathcal{C}_T$ , alors  $c_n(f) \rightarrow 0$  quand  $n \rightarrow \pm\infty$ .*

*Démonstration.* La série  $\sum_{n \in \mathbb{Z}} |c_n(f)|^2$  converge. □

**Corollaire 3.32.** *Si  $f \in \mathcal{C}_T$  et  $g \in \mathcal{C}_T$  ont les mêmes coefficients de Fourier, alors  $f = g$ .*

*Démonstration.* Les coefficients de  $f - g$  sont nuls et donc

$$\frac{1}{T} \int_0^T (f(t) - g(t))^2 dt = 0.$$

On déduit que  $f - g = 0$ . □

### 3.2.5 Théorème de Dirichlet

Nous avons mentionné plus haut que la somme partielle  $S_N(f)$  peut s'écrire :

$$S_N(f) = f * D_N \quad \text{avec} \quad D_N = \sum_{n=-N}^N e_n.$$

Les fonctions  $D_N$  s'appellent les noyaux de Dirichlet. Il est possible de donner une valeur explicite des noyaux de Dirichlet car il s'agit en fait d'une série géométrique :

$$\begin{aligned} D_N(t) &= \sum_{n=-N}^N e^{i\omega n t} = e^{-i\omega N t} \frac{1 - e^{i\omega(2N+1)t}}{1 - e^{i\omega t}} \\ &= \frac{e^{i\omega(N+1/2)t} - e^{-i\omega(N+1/2)t}}{e^{i\omega t/2} - e^{-i\omega t/2}} = \frac{\sin(\omega(N+1/2)t)}{\sin(\omega t/2)}. \end{aligned}$$

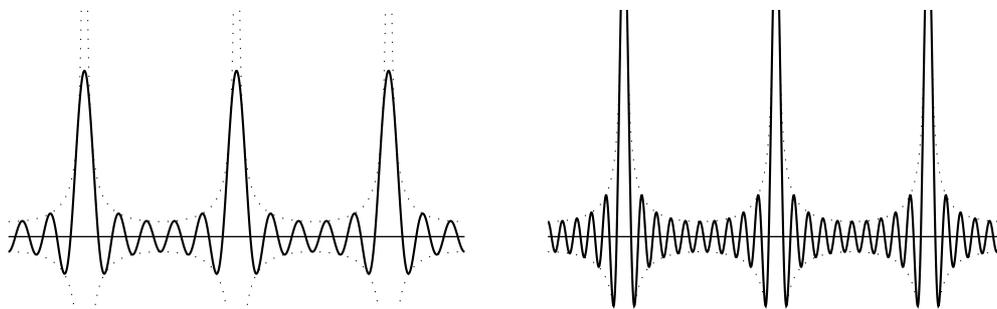


FIGURE 3.2 – Les noyaux de Dirichlet pour  $N = 5$  et  $N = 10$ .

Une fonction  $f \in \mathcal{C}_T$  n'est pas nécessairement la somme de sa série de Fourier car la suite  $(D_N)$  n'est pas une unité approchée. Les oscillations du noyau de Dirichlet  $D_N$  ne deviennent pas petites quand  $N \rightarrow +\infty$ . Bien au contraire, la fonction  $D_N$  oscille entre les graphes des fonctions  $1/\sin(\omega t/2)$  et  $-1/\sin(\omega t/2)$ .

**Théorème 3.33** (Dirichlet). *Si  $f \in \mathcal{C}_T$  est de classe  $\mathcal{C}^1$ , alors  $f$  est la somme de sa série de Fourier :*

$$\forall x \in \mathbb{R}, \quad \lim_{N \rightarrow +\infty} S_N(f)(x) = f(x).$$

*Démonstration.* Observons d'abord que d'après le Lemme 3.3,

$$\frac{1}{T} \int_0^T D_N(t) dt = \sum_{n=-N}^N \frac{1}{T} \int_0^T e_n(t) dt = 1.$$

Par conséquent :

$$S_N(f)(x) - f(x) = f * D_N(x) - f(x) = \frac{1}{T} \int_0^T f(x-t) D_N(t) dt - \frac{1}{T} \int_0^T f(x) D_N(t) dt.$$

On a donc :

$$S_N(f)(x) - f(x) = \frac{1}{T} \int_0^T (f(x-t) - f(x)) \cdot D_N(t) dt.$$

Nous souhaitons montrer que cette quantité tend vers 0 quand  $N$  tend vers  $+\infty$ .

Remplaçons  $D_N(t)$  par la valeur explicite calculée plus haut. On a :

$$\frac{1}{T} \int_0^T (f(x-t) - f(x)) \cdot D_N(t) dt = \frac{1}{T} \int_0^T \frac{f(x-t) - f(x)}{\sin(\omega t/2)} \cdot \sin\left(\omega\left(N + \frac{1}{2}\right)t\right) dt.$$

La fonction  $t \mapsto \frac{f(x-t) - f(x)}{\sin(\omega t/2)}$  est périodique de période  $T$ , continue en dehors des multiples de  $T$  (là où le dénominateur s'annule). Comme  $f$  est  $\mathcal{C}^1$ , elle se prolonge par continuité en  $t = 0$  (et donc aux multiples de  $T$  par périodicité). En effet :

$$\frac{f(x-t) - f(x)}{\sin(\omega t/2)} = \frac{f(x-t) - f(x)}{t} \cdot \frac{t}{\sin(\omega t/2)} \xrightarrow{t \rightarrow 0} \frac{2f'(x)}{\omega}.$$

Notons  $g : \mathbb{R} \rightarrow \mathbb{C}$  le prolongement par continuité.

On a alors :

$$S_N(f)(x) - f(x) = \frac{1}{T} \int_0^T g(t) \cdot \frac{e^{i\omega(N+1/2)t} - e^{-i\omega(N+1/2)t}}{2i} dt = c_{-N}(g_1) - c_N(g_2)$$

avec  $g_1 : \mathbb{R} \rightarrow \mathbb{C}$  et  $g_2 : \mathbb{R} \rightarrow \mathbb{C}$  les fonctions de  $\mathcal{C}_T$  définies par :

$$g_1(t) = \frac{1}{2i} g(t) e^{i\omega t/2} \quad \text{et} \quad g_2(t) = \frac{1}{2i} g(t) e^{-i\omega t/2}.$$

Comme  $g_1 \in \mathcal{C}_T$  et  $g_2 \in \mathcal{C}_T$ , les coefficients de Fourier  $c_{-N}(g_1)$  et  $c_N(g_2)$  tendent vers 0 quand  $N$  tend vers l'infini. Cela montre bien que :

$$\lim_{N \rightarrow +\infty} S_N(f)(x) - f(x) = \lim_{N \rightarrow +\infty} c_{-N}(g_1) - c_N(g_2) = 0. \quad \square$$

*Exemple 3.34.* Dans le cas de la fonction  $f \in \mathcal{C}_{2\pi}$  qui coïncide avec  $x \mapsto |x|$  sur  $[-\pi, \pi]$ , on a  $c_0(f) = \frac{\pi}{2}$  et pour  $p \geq 1$ ,  $c_{\pm 2p}(f) = 0$  et  $c_{\pm(2p-1)}(f) = -\frac{2}{\pi(2p-1)^2}$ . On a donc :

$$\pi = f(\pi) = \frac{\pi}{2} + 2 \sum_{p=0}^{+\infty} \frac{2}{\pi(2p+1)^2}.$$

On en déduit que :

$$\sum_{p=0}^{+\infty} \frac{1}{(2p+1)^2} = \frac{\pi}{8}.$$

### 3.2.6 Théorème de Féjer

À la fin du 19-ème siècle, l'italien Cesàro a l'idée de rendre convergentes des suites divergentes  $(u_n)$  en considérant les moyennes arithmétiques :

$$v_n = \frac{u_1 + \dots + u_n}{n}.$$

Par exemple, si l'on applique ce procédé à la suite  $(u_n)$  définie par  $u_n := (-1)^n$ , la suite  $(v_n)$  converge vers 0. Si la suite  $(u_n)$  converge, il est rassurant de constater que la suite  $v_n$  converge vers la même limite.

Si l'on applique le procédé de Cesàro à la suite des sommes partielles  $(S_N(f))$  de la série de Fourier de  $f$ , on obtient une nouvelle suite :

$$T_N(f) := \frac{1}{N} (S_0(f) + S_1(f) + \dots + S_{N-1}(f)) = \sum_{n=-N}^N \left(1 - \frac{|n|}{N}\right) c_n(f) \cdot e_n.$$

Comme on l'a fait pour les sommes partielles de la série de Fourier, on peut écrire  $T_N(f)$  à l'aide d'un produit de convolution :

$$T_N(f) = \frac{1}{N} (f * D_0 + f * D_1 + \dots + f * D_{N-1}) = f * F_N$$

où les fonctions

$$F_N = \frac{1}{N} (D_0 + D_1 + \dots + D_{N-1})$$

portent le nom de noyaux de Féjer.

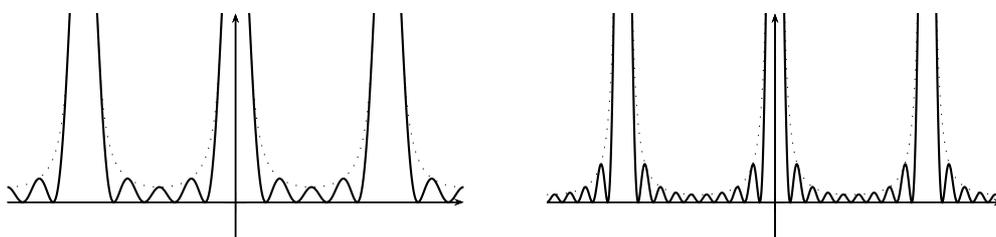


FIGURE 3.3 – Les noyaux de Féjer pour  $N = 5$  et  $N = 10$ .

**Théorème 3.35** (Féjer). Si  $f \in C_T$ , les sommes de Féjer  $T_N(f) := \frac{1}{N} (S_0(f) + \dots + S_{N-1}(f))$  convergent vers  $f$ .

*Démonstration.* On a

$$\begin{aligned} F_N(t) &= \frac{\sin(\omega t/2) + \sin(3\omega t/2) + \dots + \sin((2N+1)\omega t/2)}{N \sin(\omega t/2)} \\ &= \operatorname{Im} \left( \frac{e^{i\omega t/2} (1 + e^{i\omega t} + \dots + e^{i\omega(N-1)t})}{N \sin(\omega t/2)} \right) \\ &= \frac{1}{N \sin(\omega t/2)} \operatorname{Im} \left( e^{i\omega t/2} \frac{1 - e^{i\omega N t}}{1 - e^{i\omega t}} \right). \end{aligned}$$

On a donc

$$F_N(t) = \frac{1 - \cos(\omega N t)}{2N \sin^2(\omega t/2)} = \frac{1}{N} \left( \frac{\sin(\omega N t/2)}{\sin(\omega t/2)} \right)^2.$$

En particulier,  $F_N$  est positive. De plus,

$$\frac{1}{T} \int_0^T F_N(t) dt = \frac{1}{N} \sum_{k=0}^{N-1} \frac{1}{T} \int_0^T D_k(t) dt = 1.$$

Enfin, pour tout  $\delta \in ]0, T/2[$ ,  $F_N$  converge uniformément vers 0 sur  $[\delta, T - \delta]$  puisque :

$$0 \leq F_N(t) \leq \frac{1}{N \sin^2(\omega t/2)}.$$

Par conséquent, la suite  $(F_N)$  des noyaux de Féjer est une unité périodique approchée.  $\square$

### 3.2.7 Fonctions continues par morceaux

**Définition 3.36.** Soit  $[a, b]$  un segment de  $\mathbb{R}$  et  $f : [a, b] \rightarrow \mathbb{C}$  une fonction.

1. On dit que  $f$  est continue par morceaux sur  $[a, b]$  s'il existe une subdivision  $a_0 = a < a_1 < \dots < a_m = b$  telle que pour tout  $i = 0, \dots, m-1$ ,  $f|_{]a_i, a_{i+1}[}$  se prolonge en une fonction continue à  $[a_i, a_{i+1}]$ , autrement dit :  $f|_{]a_i, a_{i+1}[}$  est continue et les limites  $\lim_{x \rightarrow a_i^+} f(x)$ ,  $\lim_{x \rightarrow a_{i+1}^-} f(x)$  existent et sont finies.
2. On dit que  $f$  est  $\mathcal{C}^1$  par morceaux sur  $[a, b]$  s'il existe une subdivision  $a_0 = a < a_1 < \dots < a_m = b$  telle que pour tout  $i = 0, \dots, m-1$ ,  $f|_{]a_i, a_{i+1}[}$  se prolonge en une fonction  $\mathcal{C}^1$  à  $[a_i, a_{i+1}]$ , autrement dit :  $f|_{]a_i, a_{i+1}[}$  est  $\mathcal{C}^1$  et les limites

$$\lim_{x \rightarrow a_i^+} f(x), \quad \lim_{x \rightarrow a_{i+1}^-} f(x), \quad \lim_{x \rightarrow a_i^+} f'(x), \quad \lim_{x \rightarrow a_{i+1}^-} f'(x)$$

existent et sont finies.

**Définition 3.37.** Soit  $I$  un intervalle quelconque de  $\mathbb{R}$  et  $f : I \rightarrow \mathbb{C}$  une fonction. On dit que  $f$  est continue (resp.  $\mathcal{C}^1$ ) par morceaux sur  $I$  si la restriction de  $f$  à tout segment  $[a, b] \subset I$  est continue (resp.  $\mathcal{C}^1$ ) par morceaux sur  $[a, b]$ .

*Méthode 3.38.* Pour montrer qu'une fonction  $f : \mathbb{R} \rightarrow \mathbb{C}$  périodique de période  $T$  est continue par morceaux,

1. on fait la liste des discontinuités de  $f$  dans  $[0, T]$ ,
2. on montre qu'en chacune de ces discontinuités,  $f$  a une limite finie à gauche et à droite,
3. on montre que  $f$  a une limite finie à droite en 0 et à gauche en  $T$ .

Si  $f$  est paire ou impaire, il suffit d'étudier les limites à gauche et à droite aux discontinuités contenues dans  $[0, T/2]$ , ainsi qu'en 0 et en  $T/2$ . Pour montrer que  $f$  est  $\mathcal{C}^1$  par morceaux, il faudra faire le même travail avec  $f'$ .

On peut définir l'intégrale d'une fonction continue par morceaux en utilisant l'intégrale des fonctions continues de la manière suivante. Soit  $f$  une fonction continue par morceaux sur  $[a, b]$ . Notons  $a_0 = a < a_1 < \dots < a_m = b$  ses points de discontinuité : pour tout  $i = 0, \dots, m-1$ ,  $f|_{]a_i, a_{i+1}[}$  se prolonge en une fonction continue à  $[a_i, a_{i+1}]$  mais la limite à gauche ou la limite à droite en  $a_i$  (ou les deux) ne coïncident pas avec  $f(a_i)$ .

On définit alors

$$\int_a^b f(t) dt = \sum_{i=0}^{m-1} \int_{a_i}^{a_{i+1}} f(t) dt.$$

Noter que chaque terme de la somme de droite est bien défini. En effet, comme  $f|_{]a_i, a_{i+1}[}$  se prolonge en une fonction continue à  $[a_i, a_{i+1}]$ , on sait donner un sens à l'intégrale  $\int_{a_i}^{a_{i+1}} f(t) dt$  (intégrale d'une fonction continue sur un segment).

Autrement dit, pour calculer l'intégrale d'une fonction continue par morceaux sur un segment, on découpe ce segment en sous-intervalles sur lesquels la fonction est continue.

Les coefficients de Fourier  $a_n(f)$ ,  $b_n(f)$  pour  $n \in \mathbb{N}$  et  $c_n(f)$  pour  $n \in \mathbb{Z}$  d'une fonction continue par morceaux  $f$  se définissent exactement de la même manière que pour les fonctions continues.

*Exercice 3.39.* On considère la fonction  $f : \mathbb{R} \rightarrow \mathbb{R}$  qui est  $2\pi$  périodique, impaire, et qui vaut 1 sur l'intervalle  $]0, \pi[$ .

1. Que vaut  $f$  en 0? en  $\pi$ ?
2. Dessiner le graphe de  $f$  sur trois périodes.
3. Justifier que  $f$  est une fonction  $\mathcal{C}^1$  par morceaux.

4. Calculer les coefficients de Fourier de  $f$ .

Considérons à présent une fonction  $f$  qui est  $\mathcal{C}^1$  par morceaux sur un segment  $[a, b]$  : il existe une subdivision  $a_0 = a < a_1 < \dots < a_m = b$  telle que pour tout  $i = 0, \dots, m-1$ ,  $f|_{]a_i, a_{i+1}[}$  se prolonge en une fonction  $\mathcal{C}^1$  à  $[a_i, a_{i+1}]$ . A priori, la dérivée  $f'$  est seulement définie sur  $[a, b] \setminus \{a_0, \dots, a_m\}$ , mais pas nécessairement aux points  $a_i$ . On peut seulement définir les limites à gauche ou à droite de  $f'$  en chaque  $a_i$ . En notant  $g_i$  le prolongement par continuité de  $f'|_{]a_i, b_i[}$  à  $[a_i, b_i]$ , on peut aussi définir

$$\int_{a_i}^{a_{i+1}} f'(t) dt = \int_{a_i}^{a_{i+1}} g_i(t) dt.$$

On peut enfin considérer l'intégrale de  $f'$  sur  $[a, b]$  :

$$\int_a^b f'(t) dt = \sum_{i=0}^{m-1} \int_{a_i}^{a_{i+1}} f'(t) dt.$$

*Exercice 3.40.* Montrer que la formule  $c_n(f') = inc_n(f)$  reste encore valable lorsque  $f$  est continue  $2\pi$  périodique et  $\mathcal{C}^1$  par morceaux.

### Le théorème de Dirichlet

Lorsque  $f$  est seulement continue par morceaux mais pas continue,  $S_n(f)$  ne peut pas converger uniformément vers  $f$ , car une limite uniforme de fonctions continues est continue. Par contre, on dispose du théorème suivant (dont on admet la preuve) :

**Théorème 3.41** (Théorème de Dirichlet). *Soit  $f : \mathbb{R} \rightarrow \mathbb{C}$  une fonction  $T$ -périodique  $\mathcal{C}^1$  par morceaux. Alors pour tout  $x \in \mathbb{R}$ ,*

$$\lim_{n \rightarrow +\infty} S_n(f)(x) = \frac{1}{2}(f(x^-) + f(x^+)),$$

où  $f(x^-)$  (resp.  $f(x^+)$ ) est la limite à gauche (resp. à droite) de  $f$  en  $x$ .

En particulier, si  $f$  est continue en  $x \in \mathbb{R}$ , alors  $\lim_{n \rightarrow +\infty} S_n(f)(x) = f(x)$ .

### 3.2.8 Formule de Parseval

Sur l'ensemble des fonctions  $T$ -périodiques et continues par morceaux, une autre quantité joue un rôle important pour *mesurer* les fonctions :

$$\|f\|_2 := \sqrt{\frac{1}{T} \int_0^T |f(t)|^2 dt}.$$

L'application  $f \mapsto \|f\|_2$  s'appelle la norme 2, ou encore norme quadratique (là encore, le lecteur intéressé par les normes pourra se reporter à la section suivante).

La formule de Parseval reste valide pour les fonctions continues par morceaux.

**Théorème 3.42** (Formule de Parseval). *Soit  $f : \mathbb{R} \rightarrow \mathbb{C}$  une fonction  $T$ -périodique et continue par morceaux. Alors*

$$\|f\|_2^2 = \sum_{n \in \mathbb{Z}} |c_n(f)|^2.$$

Lorsque  $f$  est simplement continue ou continue par morceaux, on ne peut pas affirmer que sa série de Fourier converge uniformément, ni même simplement. Par contre, on peut dire qu'elle converge pour la norme 2, au sens suivant :

**Corollaire 3.43.** Soit  $f : \mathbb{R} \rightarrow \mathbb{C}$  une fonction  $T$ -périodique et continue par morceaux. Alors

$$\lim_{N \rightarrow +\infty} \|f - S_N(f)\|_2 = 0.$$

*Démonstration.* Étant donné  $N \geq 1$ , on pose

$$g_N = f - S_N(f).$$

Alors

$$c_n(g) = \begin{cases} 0 & \text{si } |n| \leq N \\ c_n(f) & \text{si } |n| > N. \end{cases}$$

D'après la formule de Parseval,

$$\|f - S_N(f)\|_2^2 = \|g_N\|_2^2 = \sum_{|n| > N} |c_n(f)|^2 \xrightarrow{N \rightarrow +\infty} 0. \quad \square$$

Le sens de ce corollaire est que l'énergie de la partie négligée d'un signal (le reste de la série) peut être rendue arbitrairement petite, lorsqu'on néglige de moins en moins de termes dans le signal transmis.

*Méthode 3.44.* Pour développer une fonction  $T$ -périodique  $f : \mathbb{R} \rightarrow \mathbb{R}$  en série de Fourier, on pourra procéder comme suit :

1. Tracer le graphe de  $f$  sur plusieurs périodes (au moins 3).
2. Déterminer la classe de  $f$  : continue par morceaux ou continue,  $\mathcal{C}^1$  par morceaux ou  $\mathcal{C}^1$ .
3. Calculer les coefficients de Fourier de  $f$  :  $a_n(f)$  et  $b_n(f)$ , ou  $c_n(f)$ . On utilisera ici la parité éventuelle de  $f$ .
4. Appliquer
  - (a) le théorème de Parseval : convergence en  $\|\cdot\|_2$  pour les fonctions continues par morceaux,
  - (b) le théorème de Dirichlet : convergence en tout point de continuité de  $f$  pour les fonctions  $\mathcal{C}^1$  par morceaux,

*Exercice 3.45* (Contrôle continu 2018). Soit  $f : \mathbb{R} \rightarrow \mathbb{R}$  la fonction  $2\pi$ -périodique et paire telle que

$$\forall x \in [0, \pi], \quad f(x) = \left|x - \frac{\pi}{2}\right|.$$

1. Dessiner le graphe de  $f$  sur au moins trois périodes.
2. Montrer que  $f$  est continue et  $\mathcal{C}^1$  par morceaux.
3. Calculer les coefficients de Fourier de  $f$ .
4. Est-ce que la série de Fourier  $(S_N(f))_{N \in \mathbb{N}}$  de  $f$  converge uniformément ? On justifiera la réponse.
5. Est-ce que

$$\lim_{N \rightarrow +\infty} \int_0^{2\pi} |S_N(f)(t) - f(t)|^2 dt = 0 \quad ?$$

On justifiera la réponse.

## 3.3 Approfondir

### 3.3.1 Notion de norme

On étudiera en détail le concept de norme dans le chapitre suivant. On se contente pour l'instant de donner la définition :

**Définition 3.46.** On appelle norme sur un espace vectoriel  $E$  toute application  $N : E \rightarrow \mathbb{R}^+$  vérifiant

1. (séparation)  $\forall x \in E, N(x) = 0$  si et seulement si  $x = 0$ ,
2. (positive homogénéité)  $\forall x \in E, \forall \lambda \in \mathbb{R}, N(\lambda x) = |\lambda|N(x)$ ,
3. (inégalité triangulaire)  $\forall x, y \in E, N(x + y) \leq N(x) + N(y)$ .

**Lemme 3.47.** L'application

$$f \in \mathcal{C}_T \mapsto \|f\|_\infty = \max_{x \in [0, T]} |f(x)|.$$

est une norme.

*Démonstration.* D'abord, cette application est bien à valeurs dans  $\mathbb{R}^+$  et vaut 0 lorsque  $f$  est la fonction constante nulle. De plus, si  $\|f\|_\infty = 0$ , cela signifie que  $\max_{x \in [0, T]} |f(x)| = 0$ , donc  $f = 0$  sur  $[0, T]$  et par  $T$ -périodicité, sur  $\mathbb{R}$ . La propriété de séparation est ainsi établie.

Pour montrer la positive homogénéité, soient  $f \in \mathcal{C}_T$  et  $\lambda \in \mathbb{R}$ . Alors pour tout  $x \in [0, T]$ ,

$$|\lambda f(x)| = |\lambda| |f(x)| \leq |\lambda| \|f\|_\infty$$

d'où l'on déduit  $\|\lambda f\|_\infty \leq |\lambda| \|f\|_\infty$ . De même, si  $\lambda \neq 0$ , pour tout  $x \in [0, T]$ ,

$$|f(x)| = \frac{1}{|\lambda|} |\lambda| |f(x)| = \frac{1}{|\lambda|} |\lambda f(x)| \leq \frac{1}{|\lambda|} \|\lambda f\|_\infty$$

et donc  $\|f\|_\infty \leq \frac{1}{|\lambda|} \|\lambda f\|_\infty$ , ce qui implique  $|\lambda| \|f\|_\infty \leq \|\lambda f\|_\infty$ , et cette inégalité reste vraie pour  $\lambda = 0$ . Par double inégalité, on a donc bien  $\|\lambda f\|_\infty = |\lambda| \|f\|_\infty$  et la positive homogénéité est démontrée.

Enfin, pour l'inégalité triangulaire, soient  $f, g \in \mathcal{C}_T$ . On observe que pour tout  $x \in [0, T]$ ,

$$|(f + g)(x)| = |f(x) + g(x)| \leq |f(x)| + |g(x)| \leq \|f\|_\infty + \|g\|_\infty,$$

d'où l'on déduit que  $\|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$ , ce qui achève la preuve de l'inégalité triangulaire.  $\square$

On a introduit dans ce chapitre une autre norme, la norme 2 :

$$\forall f \in \mathcal{C}_T, \quad \|f\|_2 := \sqrt{\frac{1}{T} \int_0^T |f(t)|^2 dt}.$$

Pour vérifier qu'il s'agit bien d'une norme au sens de la Définition 3.46, il est commode de commencer par introduire la notion de produit scalaire complexe. La définition diffère légèrement du produit scalaire réel.

**Définition 3.48.** Soit  $E$  un espace vectoriel sur  $\mathbb{C}$ . Un produit scalaire complexe (ou hermitien) est une application  $\langle \cdot, \cdot \rangle : E \times E \rightarrow \mathbb{C}$  telle que

1. (antilinearité par rapport à la première variable) pour tout  $w \in E$ , la fonction  $v \in E \mapsto \langle v, w \rangle \in \mathbb{C}$  est antilinéaire<sup>1</sup>,

---

1. Cela signifie que pour tout  $v, v' \in E$ , pour tout  $\lambda \in \mathbb{C}$ ,  $\langle v + \lambda v', w \rangle = \langle v, w \rangle + \bar{\lambda} \langle v', w \rangle$ .

2. (linéarité par rapport à la deuxième variable) pour tout  $v \in E$ , la fonction  $w \in E \mapsto \langle v, w \rangle \in \mathbb{R}$  est linéaire,
3. (antisymétrie) pour tout  $v, w \in E$ ,  $\langle v, w \rangle = \overline{\langle w, v \rangle}$ ,
4. (défini positif) pour tout  $v \in E$ ,  $\langle v, v \rangle \geq 0$  avec égalité si et seulement si  $v = 0$ .

On peut montrer que les produits scalaires complexes vérifient les principales propriétés des produits scalaires réels :

1. l'application  $\| \cdot \| : E \rightarrow \mathbb{R}$  définie par

$$\|v\| := \sqrt{\langle v, v \rangle} \quad , v \in E$$

est une norme sur  $E$  au sens de la définition 3.46,

2. pour tout  $v, w \in E$ ,

$$\langle v + w, v + w \rangle = \|v\|^2 + \|w\|^2 + 2 \operatorname{Re} \langle v, w \rangle,$$

3. l'inégalité de Cauchy-Schwarz : pour tout  $v, w \in E$

$$|\langle v, w \rangle| \leq \|v\| \|w\|,$$

4. pour deux vecteurs orthogonaux  $v$  et  $w$ , c'est-à-dire  $\langle v, w \rangle = 0$ , on a la relation de Pythagore :

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2.$$

Si  $f, g \in \mathcal{C}_T$ , on définit

$$\langle f, g \rangle = \frac{1}{T} \int_0^T \overline{f(t)} g(t) dt.$$

**Proposition 3.49.** L'application  $\langle \cdot, \cdot \rangle$  est un produit scalaire complexe sur  $\mathcal{C}_T$ .

*Démonstration.* Par linéarité de l'intégrale, on a bien antilinéarité par rapport à la première variable  $f$ , linéarité par rapport à la seconde variable  $g$ . De plus,

$$\overline{\langle f, g \rangle} = \overline{\int_a^b \overline{f(t)} g(t) dt} = \int_a^b \overline{\overline{f(t)} g(t)} dt = \int_a^b g(t) \overline{f(t)} dt = \langle g, f \rangle,$$

d'où l'on déduit l'antisymétrie. Enfin,

$$\langle f, f \rangle = \int_a^b \overline{f(t)} f(t) dt = \int_a^b |f(t)|^2 dt \geq 0$$

avec égalité si, et seulement si,  $f = 0$  (on utilise ici que l'intégrale d'une fonction continue positive est nulle si, et seulement si, la fonction est partout nulle).  $\square$

La norme associée à ce produit scalaire est

$$\|f\|_2 := \sqrt{\frac{1}{T} \int_0^T |f(t)|^2 dt}.$$

Dans la section précédente de ce chapitre, on a vu que la série de Fourier d'une fonction  $f$  convergeait pour la norme quadratique dès que  $f$  est continue (ou continue par morceaux). En revanche, la convergence uniforme exige des hypothèses supplémentaires (par exemple,  $f$  continue et  $\mathcal{C}^1$  par morceaux).

On pourra remarquer que pour tout  $f \in \mathcal{C}_T$ ,

$$\|f\|_2^2 = \frac{1}{2\pi} \int_0^{2\pi} |f(t)|^2 dt \leq \frac{1}{2\pi} \int_0^{2\pi} \|f\|_\infty^2 dt \leq \|f\|_\infty^2$$

et donc  $\| \cdot \|_2 \leq \| \cdot \|_\infty$ . Cela implique que la convergence uniforme implique la convergence quadratique. En revanche, la réciproque est fautive.

# Chapitre 4

## Normes

### 4.1 Découvrir

#### 4.1.1 Normes et suites convergentes

Commençons par le cas  $n = 2$ . Etant donnés deux points  $A$  et  $B$  dans le plan, on a l'habitude de dire que le chemin le plus court entre  $A$  et  $B$  est la ligne droite. Qu'entend-on exactement par *le plus court*? Comment mesure-t-on cette distance? Identifions le plan muni d'un repère à  $\mathbb{R}^2$ . Dans ce repère, les deux points  $A$  et  $B$  ont respectivement pour coordonnées  $(x_A, y_A)$  et  $(x_B, y_B)$ . La distance entre ces deux points est généralement définie comme la norme du vecteur  $\vec{AB} = (x_B - x_A, y_B - y_A)$ , à savoir

$$\text{dist}(A, B) = \|\vec{AB}\| = \sqrt{\langle \vec{AB}, \vec{AB} \rangle} = \sqrt{(x_B - x_A)^2 + (y_B - y_A)^2}.$$

Mais il existe en fait bien d'autres manières de mesurer la norme d'un vecteur (et donc aussi la distance entre deux points). La norme décrite précédemment est dite *euclidienne*, parce qu'elle est définie à partir d'un produit scalaire. On la notera  $\|\cdot\|_2$  (prononcer *norme 2*) dans la suite.

Imaginons maintenant que pour rejoindre les points  $A$  et  $B$ , on ne puisse suivre que des lignes horizontales ou verticales. Alors partant de  $A$ , on suivra par exemple la ligne horizontale passant par  $A$  jusqu'au point intermédiaire  $C$  de coordonnées  $(x_B, y_A)$  puis on remontera la verticale jusqu'au point  $B$ . On aura parcouru la distance  $|x_A - x_B| + |y_A - y_B|$ . Tout autre chemin entre  $A$  et  $B$  et suivant uniquement des horizontales ou verticales aurait au moins la même longueur. On obtient de cette manière une autre norme du vecteur  $\vec{AB}$  notée  $\|\cdot\|_1$  (prononcer *norme 1*), et définie par

$$\|\vec{AB}\|_1 = |x_A - x_B| + |y_A - y_B|.$$

Il y a en fait une infinité de manières de mesurer la norme d'un vecteur, selon les critères que l'on veut privilégier. Pourtant, toutes ces normes sont comparables. Ainsi, pour tous points  $A$  et  $B$ ,

$$\|\vec{AB}\|_2 \leq \|\vec{AB}\|_1 \leq \sqrt{2} \|\vec{AB}\|_2.$$

On dit que ces normes sont équivalentes.

On peut introduire d'autres normes sur  $\mathbb{R}^2$ . En fait, sur chaque espace vectoriel, on peut définir une infinité de normes possibles, qui servent toutes à *mesurer* (en un certain sens) les vecteurs de cet espace vectoriel. Dans le chapitre précédent, on a déjà vu deux exemples de normes sur l'espace des fonctions continues  $T$ -périodiques. Elles nous ont permis de mesurer l'écart entre un signal  $f$  et son approximation  $S_N(f)$  (la somme partielle de sa série de Fourier).

Donnons la définition générale d'une norme sur un espace vectoriel<sup>1</sup>  $E$ .

---

1. On a exceptionnellement rajouté une quatrième section à ce chapitre, qui récapitule tout ce que vous devez savoir sur les espaces vectoriels.

**Définition 4.1.** On appelle norme sur un espace vectoriel  $E$  toute application  $N : E \rightarrow \mathbb{R}^+$  vérifiant

1. (séparation)  $\forall x \in E, N(x) = 0$  si et seulement si  $x = 0$ ,
2. (positive homogénéité)  $\forall x \in E, \forall \lambda \in \mathbb{R}, N(\lambda x) = |\lambda|N(x)$ ,
3. (inégalité triangulaire)  $\forall x, y \in E, N(x + y) \leq N(x) + N(y)$ .

Comme exemple d'espace vectoriel  $E$ , on considèrera souvent  $\mathbb{R}^n$  (avec  $n \geq 1$ ), ou bien un espace de matrices, ou encore un espace de fonctions. Souvent, les normes sont notées par une double barre  $\|\cdot\|$ . Alors,  $\|x\|$  désigne la norme d'un vecteur  $x \in E$  et  $\|\cdot\|$  doit être compris comme la fonction qui à un vecteur  $x$  associe sa norme  $\|x\|$ .

Vérifions que sur  $\mathbb{R}^2$ , l'application  $\|\cdot\|_1 : (x_1, x_2) \in \mathbb{R}^2 \mapsto |x_1| + |x_2|$  est bien une norme. On observe d'abord que c'est une fonction qui ne prend que des valeurs positives et qui s'annule pour  $(x_1, x_2) = (0, 0)$ . Réciproquement si  $(x_1, x_2) \in \mathbb{R}^2$  est tel que  $\|(x_1, x_2)\| = 0$ , alors  $|x_1| + |x_2| = 0$  et donc  $|x_1| = -|x_2|$ . Donc  $|x_1|$  est à la fois positif (comme valeur absolue d'un nombre) et négatif (comme opposé de la valeur absolue de  $x_2$ ). Donc  $|x_1| = 0$ , ce qui implique  $x_1 = 0$ . De même,  $x_2 = 0$ . On a donc vérifié la propriété de séparation.

Vérifions la positive homogénéité. Soit  $(x_1, x_2) \in \mathbb{R}^2$  et  $\lambda \in \mathbb{R}$ . Alors

$$\begin{aligned} \|\lambda(x_1, x_2)\|_1 &= \|(\lambda x_1, \lambda x_2)\|_1 = |\lambda x_1| + |\lambda x_2| \\ &= |\lambda||x_1| + |\lambda||x_2| = |\lambda|(|x_1| + |x_2|) = |\lambda|\|(x_1, x_2)\|_1. \end{aligned}$$

Cela montre la positive homogénéité.

Enfin, montrons l'inégalité triangulaire. Soient  $(x_1, x_2), (y_1, y_2) \in \mathbb{R}^2$ . Alors

$$\begin{aligned} \|(x_1, x_2) + (y_1, y_2)\|_1 &= \|(x_1 + y_1, x_2 + y_2)\|_1 = |x_1 + y_1| + |x_2 + y_2| \\ &\leq (|x_1| + |y_1|) + (|x_2| + |y_2|) = (|x_1| + |x_2|) + (|y_1| + |y_2|) \\ &= \|(x_1, x_2)\|_1 + \|(y_1, y_2)\|_1, \end{aligned}$$

ce qui montre l'inégalité triangulaire. On peut conclure que  $\|\cdot\|_1$  est bien une norme sur  $\mathbb{R}^2$ .

*Exercice 4.2.* On considère le cas de l'espace vectoriel  $\mathbb{R}$ . Montrer que l'application  $x \in \mathbb{R} \mapsto |x| \in \mathbb{R}^+$  est une norme sur  $\mathbb{R}$ .

*Exercice 4.3.* Est-ce que l'application  $x \in \mathbb{R} \mapsto x \in \mathbb{R}$  est une norme sur  $\mathbb{R}$  ?

*Exercice 4.4.* Sur  $\mathbb{R}^n$ , on considère l'application

$$(x_1, \dots, x_n) \mapsto |x_1| + \dots + |x_n|.$$

Montrer que c'est une norme.

*Exercice 4.5.* Sur  $\mathbb{R}^n$ , on considère l'application

$$(x_1, \dots, x_n) \mapsto \max(|x_1|, \dots, |x_n|).$$

Montrer que c'est une norme (on pourra commencer avec le cas  $n = 2$ ).

*Exercice 4.6.* Sur  $\mathbb{R}^n$ , on considère l'application

$$(x_1, \dots, x_n) \mapsto \sqrt{x_1^2 + \dots + x_n^2}.$$

Montrer que c'est une norme.

*Indication :* pour l'inégalité triangulaire, on pourra utiliser l'inégalité de Cauchy-Schwarz :

$$\sum_{i=1}^n x_i y_i \leq \sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}.$$

Quelques propriétés sont vérifiées par toutes les normes :

**Proposition 4.7.** Soit  $N : E \rightarrow \mathbb{R}^+$  une norme.

1. Pour tout  $x \in \mathbb{R}^n$ ,  $N(-x) = N(x)$ .
2. Pour tout  $x_1, \dots, x_k \in E$ ,  $N(x_1 + \dots + x_k) \leq N(x_1) + \dots + N(x_k)$ .
3. Pour tout  $x, y \in E$ ,  $|N(x) - N(y)| \leq N(x - y)$ .

Preuve : La première propriété est une conséquence de la positive homogénéité :

$$N(-x) = N((-1) \cdot x) = |(-1)|N(x) = N(x).$$

Montrons la deuxième propriété par récurrence sur  $k \geq 1$ . L'inégalité est vraie pour  $k = 1$ . Supposons-la vraie pour un entier  $k \geq 1$  et montrons-la pour  $k+1$ . Soient  $x_1, \dots, x_{k+1} \in E$ . Posons  $y = x_1 + \dots + x_k$ . Alors par l'inégalité triangulaire, on a

$$N(y + x_{k+1}) \leq N(y) + N(x_{k+1}).$$

Par hypothèse de récurrence,

$$N(y) = N(x_1 + \dots + x_k) \leq N(x_1) + \dots + N(x_k).$$

Ainsi,

$$N(x_1 + \dots + x_{k+1}) = N(y + x_{k+1}) \leq N(y) + N(x_{k+1}) \leq N(x_1) + \dots + N(x_{k+1}).$$

La propriété est donc démontrée pour  $k + 1$ . On peut conclure qu'elle est vraie pour tout  $k \geq 1$ .

Montrons la dernière propriété. Soient  $x, y \in E$ . Comme  $x = y + (x - y)$ , l'inégalité triangulaire donne

$$N(x) \leq N(y) + N(x - y),$$

d'où l'on déduit

$$N(x) - N(y) \leq N(x - y).$$

En écrivant maintenant  $y = x + (y - x)$ , on obtient

$$N(y) - N(x) \leq N(y - x) = N(x - y).$$

Ainsi, le nombre  $N(x) - N(y)$  et son opposé, qui est  $N(y) - N(x)$ , sont tous les deux inférieurs à  $N(x - y)$ . On en déduit que

$$|N(x) - N(y)| \leq N(x - y).$$

□

On se souvient sans doute de la notion de convergence pour une suite de nombres réels  $(x_k)_{k \in \mathbb{N}}$  : on dit qu'elle converge vers un nombre  $x \in \mathbb{R}$ , si pour tout  $\varepsilon > 0$ , il existe  $K \in \mathbb{N}$  tel que pour tout  $k \geq K$ ,

$$|x_k - x| \leq \varepsilon.$$

On peut reformuler la conclusion ci-dessus en écrivant  $-\varepsilon \leq x_k - x \leq \varepsilon$ , ou encore  $x_k \in [x - \varepsilon, x + \varepsilon]$ . La valeur absolue est donc utilisée ici pour mesurer la *distance* entre un élément  $x_k$  de la suite et la limite  $x$ . En fait, on a vu dans l'exercice 4.2 que la valeur absolue est une norme sur  $\mathbb{R}$ .

Maintenant qu'on dispose de la notion de norme dans l'espace vectoriel  $E$ , on peut de la même manière mesurer des distances entre deux éléments de  $E$ , et aussi étendre la notion de convergence d'une suite à ce contexte.

**Définition 4.8.** Soit  $N : E \rightarrow \mathbb{R}^+$  une norme. On dit alors qu'une suite  $(x_k \in E)_{k \in \mathbb{N}}$  converge vers  $x$  si pour tout  $\varepsilon > 0$ , il existe  $K \in \mathbb{N}$  tel que pour tout  $k \geq K$ ,

$$N(x_k - x) \leq \varepsilon.$$

De manière équivalente, cela revient à demander que la suite de réels  $(N(x_k - x))_{k \in \mathbb{N}}$  converge vers 0 :

$$\lim_{k \rightarrow +\infty} N(x_k - x) = 0.$$

*Remarque 4.9.* Comme pour les suites réelles, la limite d'une suite, lorsqu'elle existe, est unique. En effet, si  $x$  et  $x'$  sont deux limites d'une suite  $(x_k)_{k \in \mathbb{N}}$ , on a

$$N(x - x') \leq N(x - x_k) + N(x' - x_k).$$

En prenant la limite quand  $k \rightarrow +\infty$  de l'inégalité précédente, on obtient

$$N(x - x') = 0,$$

et donc  $x = x'$ . Cela prouve l'unicité de la limite.

## 4.1.2 Normes équivalentes

La convergence d'une suite est directement liée à la norme qu'on considère sur  $E$ . D'où la question : si on se donne une autre norme  $N'$  sur  $E$ , est-ce qu'une suite qui convergeait au sens de la norme  $N$  va encore converger pour la norme  $N'$  ? La notion d'équivalence des normes permet justement de donner une condition suffisante pour que ce soit le cas.

**Définition 4.10.** On dit que deux normes  $N$  et  $N'$  sont équivalentes quand il existe  $C > 0$  et  $C' > 0$  tels que pour tout  $x \in E$ ,

$$N(x) \leq CN'(x) \quad , \quad N'(x) \leq C'N(x).$$

Insistons : les constantes  $C$  et  $C'$  sont indépendantes de  $x$  !

On verra dans la section *Assimiler* que si deux normes  $N$  et  $N'$  sont équivalentes, toute suite qui converge pour l'une, converge aussi pour l'autre (et les limites sont les mêmes).

Sur un espace vectoriel de dimension finie, on dispose du théorème fondamental suivant.

**Théorème 4.11.** Sur un espace vectoriel de dimension finie  $E$ , toutes les normes sont équivalentes.

La preuve de ce théorème est difficile et on l'admettra.

## 4.2 Assimiler

### 4.2.1 Applications de l'équivalence des normes

On a vu dans la partie précédente la notion d'*équivalence des normes* et le théorème très important que deux normes quelconques sur un espace vectoriel de dimension finie sont toujours équivalentes.

Cela permet de montrer que la convergence des suites ne dépend pas de la norme choisie :

**Proposition 4.12.** Soient  $N$  et  $N'$  deux normes équivalentes sur  $E$ . Alors toute suite qui converge pour l'une converge vers la même limite pour l'autre.

*Démonstration.* Soit  $(x_k)_{k \in \mathbb{N}} \subset E$  une suite convergente pour  $N$ . Appelons  $x$  sa limite :

$$\lim_{k \rightarrow +\infty} N(x_k - x) = 0.$$

On sait qu'il existe  $C > 0$  tel que  $N'(y) \leq CN(y)$  pour tout  $y \in E$ . Alors pour tout  $k \in \mathbb{N}$ ,

$$0 \leq N'(x_k - x) \leq CN(x_k - x).$$

On en déduit que  $\lim_{k \rightarrow +\infty} N'(x_k - x) = 0$ . Donc  $(x_k)_{k \in \mathbb{N}}$  converge vers  $x$  pour  $N'$ .

Symétriquement, on montre qu'une suite convergente pour  $N'$  converge pour  $N$ . □

Selon le problème, il sera commode d'utiliser plutôt une norme qu'une autre. Par exemple,

**Exercice 4.13.** Soit  $(u_n)_{n \in \mathbb{N}}$  une suite de  $\mathbb{R}^2$ . Pour tout  $n \in \mathbb{N}$ , on note  $v_n$  la première coordonnée de  $u_n$  et  $w_n$  la seconde coordonnée de  $u_n$  :  $u_n = (v_n, w_n)$ . Montrer alors que  $(u_n)_{n \in \mathbb{N}}$  converge si et seulement si les suites  $(v_n)_{n \in \mathbb{N}}$  et  $(w_n)_{n \in \mathbb{N}}$  convergent.

Une fois n'est pas coutume, on présente la solution de l'exercice...

**Solution :** Comme la convergence d'une suite ne dépend pas de la norme équivalente choisie et que toutes les normes sont équivalentes dans  $\mathbb{R}^2$ , on peut démontrer chaque implication avec la norme qui nous convient le mieux.

Supposons d'abord que  $(u_n)_{n \in \mathbb{N}}$  converge et montrons que  $(v_n)_{n \in \mathbb{N}}$  et  $(w_n)_{n \in \mathbb{N}}$  convergent. Notons  $\ell = (\ell_1, \ell_2) \in \mathbb{R}^2$  la limite de la suite  $(u_n)_{n \in \mathbb{N}}$ . On choisit sur  $\mathbb{R}^2$  la norme  $\|\cdot\|_\infty$ . Comme pour tout  $n \in \mathbb{N}$ ,  $0 \leq |v_n - \ell_1| \leq \|u_n - \ell\|_\infty$ , on en déduit  $\lim_{n \rightarrow +\infty} |v_n - \ell_1| = 0$ . Ainsi,  $(v_n)_{n \in \mathbb{N}}$  converge vers  $\ell_1$ . De même,  $(w_n)_{n \in \mathbb{N}}$  converge vers  $\ell_2$ . Cela prouve l'implication directe.

Réciproquement, supposons  $(v_n)_{n \in \mathbb{N}}$  et  $(w_n)_{n \in \mathbb{N}}$  convergentes, vers des limites qu'on note  $a$  et  $b$  respectivement. Montrons alors que  $(u_n)_{n \in \mathbb{N}}$  converge vers  $(a, b)$ . On choisit sur  $\mathbb{R}^2$  la norme  $\|\cdot\|_1$ . Comme pour tout  $n \in \mathbb{N}$ ,  $\|u_n - (a, b)\|_1 = |v_n - a| + |w_n - b|$  et que les deux termes du membre de droite tendent vers 0, on en déduit

$$\lim_{n \rightarrow +\infty} \|u_n - (a, b)\|_1 = 0.$$

Cela prouve l'implication réciproque. □

## 4.2.2 Boules et sphères

**Définition 4.14.** Soit  $N$  une norme sur  $E$ . On définit alors

1. la boule ouverte de centre  $a$  et de rayon  $r > 0$

$$B(a, r) := \{x \in E : N(x - a) < r\},$$

2. la boule fermée de centre  $a$  et de rayon  $r > 0$

$$\overline{B(a, r)} := \{x \in E : N(x - a) \leq r\},$$

3. la sphère de centre  $a$  et de rayon  $r > 0$

$$S(a, r) := \{x \in E : N(x - a) = r\}.$$

**Exercice 4.15.** Sur  $\mathbb{R}$  muni de la valeur absolue (qui est bien une norme !), dessiner  $B(1, 2)$ ,  $\overline{B(1, 2)}$  et  $S(1, 2)$ .

**Exercice 4.16.** Dessiner les boules ouvertes centrées en 0 et de rayon 1 pour les normes  $\|\cdot\|_1$ ,  $\|\cdot\|_2$  et  $\|\cdot\|_\infty$  dans  $\mathbb{R}^2$ .

**Solution :** Par définition

$$B_{\|\cdot\|_2}(0, 1) = \{x = (x_1, x_2) \in \mathbb{R}^2 : \sqrt{x_1^2 + x_2^2} < 1\} = \{x = (x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 < 1\}.$$

On reconnaît le disque (ouvert) centré en 0 de rayon 1.

Ensuite, pour dessiner

$$B_{\|\cdot\|_1}(0, 1) = \{x = (x_1, x_2) \in \mathbb{R}^2 : |x_1| + |x_2| < 1\},$$

on commence par considérer les points  $(x_1, x_2)$  dans cette boule tels que  $x_1 > 0, x_2 > 0$ . On trace la droite  $x_1 + x_2 = 1$  et on garde donc tous les points au-dessous de cette droite dans le quadrangle supérieur droit. On répète la même opération pour les autres quadrangles, ou bien on procède par symétrie. On obtient le carré de sommets  $(\pm 1, 0)$  et  $(0, \pm 1)$ .

Pour la boule

$$B_{\|\cdot\|_\infty}(0, 1) = \{x = (x_1, x_2) \in \mathbb{R}^2 : \max(|x_1|, |x_2|) < 1\},$$

on suit la même stratégie, et on obtient le carré de sommets  $(\pm 1, \pm 1)$ .

Une boule, ça peut donc être un carré!

On munit un espace vectoriel  $E$  d'une norme  $N$ , ce qui permet de définir les boules (ouvertes ou fermées) associées à  $N$ .

**Définition 4.17.** Soit  $A$  une partie non vide de  $E$  et  $a \in E$ .

1. On dit que  $a$  est intérieur à  $A$  s'il existe  $r > 0$  tel que  $B(a, r) \subset A$ .
2. On dit que  $a$  est adhérent à  $A$  si pour tout  $r > 0$ ,  $B(a, r) \cap A \neq \emptyset$ .

Un point intérieur à  $A$  est nécessairement dans  $A$ . De manière explicite,  $a$  est intérieur à  $A$  s'il existe  $r > 0$  tel que

$$\{x \in E : N(x - a) < r\} \subset A.$$

Tous les points de  $A$  sont adhérents à  $A$  mais un point adhérent à  $A$  n'est pas forcément dans  $A$ . Par exemple, dans  $\mathbb{R}$ , les bornes d'un intervalle ouvert sont adhérentes à cet intervalle et pourtant n'appartiennent pas à cet intervalle.

*Exemple 4.18.* On considère le cas  $n = 1$  avec  $N$  la valeur absolue. Parmi les ensembles suivants, indiquer ceux pour lesquels 0 est un point intérieur, puis un point adhérent :

$$\mathbb{R}, \quad ]-1, 1[, \quad ]0, 1[, \quad ]1, 2[.$$

### 4.2.3 Equivalence des normes et inclusions des boules

**Proposition 4.19.** L'équivalence des normes se traduit géométriquement en termes d'inclusions de boules.

Soient  $N$  et  $N'$  deux normes équivalentes sur un espace vectoriel  $E$  : il existe  $C, C' > 0$  telles que

$$N \leq CN' \quad , \quad N' \leq C'N.$$

Alors en notant

$$B_N(a, r) = \{x \in E : N(x - a) < r\} \quad , \quad B_{N'}(a, r) = \{x \in E : N'(x - a) < r\},$$

on a

$$B_N(a, r) \subseteq B_{N'}(a, C'r) \quad , \quad B_{N'}(a, r) \subseteq B_N(a, Cr).$$

En effet, montrons par exemple la première inclusion. Soit  $x \in B_N(a, r)$ . Alors  $N(x - a) < r$ . On en déduit que

$$N'(x - a) \leq C'N(x - a) < C'r.$$

Ainsi,  $x \in B_{N'}(a, C'r)$ . Comme le choix de  $x$  était arbitraire, on conclut que  $B_N(a, r) \subseteq B_{N'}(a, C'r)$ .

**Proposition 4.20.** Soient  $N$  et  $N'$  deux normes équivalentes sur un espace vectoriel  $E$  et  $A$  une partie de  $E$ . Alors un vecteur  $a \in E$  est intérieur à  $A$  pour la norme  $N$  si et seulement si il est intérieur à  $A$  pour la norme  $N'$ .

*Démonstration.* Soit  $a$  un point de  $E$  intérieur à  $A$  pour la norme  $N$ . Par définition, cela signifie qu'il existe  $r > 0$  tel que

$$\{x \in E : N(x - a) < r\} \subset A.$$

Comme  $N$  et  $N'$  sont équivalentes, il existe  $C > 0$  tel que pour tout  $y \in E$ ,  $N(y) \leq CN'(y)$ . En particulier, pour tout  $x \in E$ ,  $N(x - a) \leq CN'(x - a)$ . On en déduit que

$$\{x \in E : N'(x - a) < \frac{r}{C}\} \subseteq \{x \in E : N(x - a) < r\} \subset A.$$

Ainsi,  $a$  est intérieur à  $A$  pour la norme  $N'$ . La réciproque se démontre de manière analogue.  $\square$

En particulier, en dimension finie, la notion de point intérieur ne dépend pas de la norme choisie. On notera alors  $\text{int } A$  l'ensemble des points intérieurs à  $A$ , qu'on appelle l'intérieur de  $A$ .

On montrerait avec des arguments similaires que la notion de point adhérent ne dépend pas de la norme équivalente choisie. On notera  $\text{adh } A$  l'ensemble des points adhérents à  $A$ , qu'on appelle l'adhérence de  $A$ .

La notion de point intérieur sera très importante dans les deux chapitres suivants lorsqu'on introduira la continuité et la différentiabilité des fonctions de plusieurs variables. C'est déjà le cas pour les fonctions d'une seule variable, lorsqu'on définit la dérivabilité. Ainsi, étant donné une fonction  $f$  et un réel  $a$ , pour définir la dérivée de  $f$  en  $a$ , on forme le taux d'accroissement

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

pour  $x \neq a$ . Puis on prend la limite de cette quantité lorsque  $x$  tend vers  $a$ . On a besoin de pouvoir approcher  $a$  par la gauche ou par la droite, et vérifier que les limites obtenues sont égales. Pour que ce taux d'accroissement soit bien défini, il est donc nécessaire que  $f$  soit elle-même définie au moins sur un petit intervalle à droite de  $a$ , et sur un petit intervalle à gauche de  $a$ . C'est la raison pour laquelle on définit la dérivée de  $f$  en  $a$  lorsque  $a$  est à l'intérieur du domaine de définition de  $f$ . Et pour cette raison, on définit la dérivée de  $f$  sur un intervalle lorsque cet intervalle est ouvert.

## 4.3 Approfondir

### 4.3.1 Quelques normes en dimension finie

Dans cette section, on présente des exemples de normes sur un espace vectoriel de dimension finie qui n'est pas l'espace  $\mathbb{R}^n$ .

#### Espace de matrices

Considérons l'ensemble des matrices  $\mathcal{M}_2(\mathbb{R})$  à deux lignes et deux colonnes. Dans l'espace  $\mathcal{M}_2(\mathbb{R})$ , on introduit les quantités

$$\forall A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad N_\infty(A) = \max(|a|, |b|, |c|, |d|), \quad N_2(A) = \left( a^2 + b^2 + c^2 + d^2 \right)^{\frac{1}{2}}.$$

*Exercice 4.21.* Montrer que  $N_\infty$  et  $N_2$  sont des normes sur  $\mathcal{M}_2(\mathbb{R})$ .

Un exemple important de norme sur l'ensemble des matrices carré est donné par la notion de norme d'opérateur.

**Définition 4.22.** *Etant donné une norme  $\|\cdot\|$  sur  $\mathbb{R}^n$ , on appelle norme d'opérateur associée la norme sur  $\mathcal{M}_n(\mathbb{R})$  définie par*

$$\|A\| = \sup_{\substack{x \in \mathbb{R}^n \\ x \neq 0}} \frac{\|Ax\|}{\|x\|}.$$

*Remarque 4.23.* On a aussi

$$\|A\| = \sup_{\substack{x \in \mathbb{R}^n \\ \|x\|=1}} \|Ax\|. \quad (4.1)$$

En effet, comme  $\{x \in \mathbb{R}^n \mid \|x\|=1\} \subset \{x \in \mathbb{R}^n \mid x \neq 0\}$  et que  $\frac{\|Ax\|}{\|x\|} = \|Ax\|$  lorsque  $\|x\|=1$ , on a

$$\sup_{\substack{x \in \mathbb{R}^n \\ \|x\|=1}} \|Ax\| \leq \|A\|.$$

Ensuite, pour tout  $y \in \mathbb{R}^n$ ,  $y \neq 0$ , posons  $x = \frac{y}{\|y\|}$ . Alors  $\|x\|=1$  et  $\frac{\|Ay\|}{\|y\|} = \|A(y/\|y\|)\| = \|Ax\|$ . On en déduit

$$\frac{\|Ay\|}{\|y\|} \leq \sup_{\substack{x \in \mathbb{R}^n \\ \|x\|=1}} \|Ax\|.$$

Finalement, on a bien l'égalité (4.1).

**Proposition 4.24.** *La norme d'opérateur est bien une norme!*

*Démonstration.* Observons d'abord que l'application  $A \in \mathcal{M}_n(\mathbb{R}) \mapsto \|A\|$  vaut 0 en 0 et si  $A \in \mathcal{M}_n(\mathbb{R})$  vérifie  $\|A\|=0$ , alors pour tout  $x \in \mathbb{R}^n$ ,  $Ax=0$ . En prenant pour  $x$  les vecteurs de la base canonique  $e_i = (0, \dots, 0, 1, 0, \dots, 0)$  (où le 1 est en  $i^{\text{ème}}$  position), on voit que chaque colonne de  $A$  est nulle. Ainsi  $A=0$ , ce qui achève la preuve de la séparation.

Par ailleurs, pour tout  $\lambda \in \mathbb{R}$ , pour tout  $x \in \mathbb{R}^n$  de norme 1,

$$\|(\lambda A)x\| = |\lambda| \|Ax\| \leq |\lambda| \sup_{\substack{y \in \mathbb{R}^n \\ \|y\|=1}} \|Ay\|,$$

d'où l'on déduit que

$$\sup_{\substack{x \in \mathbb{R}^n \\ \|x\|=1}} \|(\lambda A)x\| \leq |\lambda| \sup_{\substack{y \in \mathbb{R}^n \\ \|y\|=1}} \|Ay\|.$$

Pour montrer que cette inégalité est en fait une égalité, on observe que pour tout  $\lambda \neq 0$ , pour tout  $x \in \mathbb{R}^n$  de norme 1,

$$\|Ax\| = \frac{1}{|\lambda|} \left( |\lambda| \|Ax\| \right) \leq \frac{1}{|\lambda|} \sup_{\substack{y \in \mathbb{R}^n \\ \|y\|=1}} |\lambda| \|Ay\|.$$

Donc

$$\sup_{\substack{x \in \mathbb{R}^n \\ \|x\|=1}} \|Ax\| \leq \frac{1}{|\lambda|} \sup_{\substack{y \in \mathbb{R}^n \\ \|y\|=1}} |\lambda| \|Ay\|.$$

En multipliant par  $|\lambda|$ , il vient

$$|\lambda| \sup_{\substack{x \in \mathbb{R}^n \\ \|x\|=1}} \|Ax\| \leq \sup_{\substack{y \in \mathbb{R}^n \\ \|y\|=1}} |\lambda| \|Ay\|.$$

Cette égalité reste vraie pour  $\lambda = 0$ . On a donc l'égalité

$$\sup_{\substack{x \in \mathbb{R}^n \\ \|x\|=1}} \|(\lambda A)x\| = |\lambda| \sup_{\substack{x \in \mathbb{R}^n \\ \|x\|=1}} \|Ax\|,$$

et donc  $\| \lambda A \| = |\lambda| \|A\|$ , ce qui prouve la positive homogénéité.

Enfin, pour tout  $A, B \in \mathcal{M}_n(\mathbb{R})$ , pour tout  $x \in \mathbb{R}^n$  tel que  $\|x\| = 1$ ,

$$\|(A+B)x\| = \|Ax+Bx\| \leq \|Ax\| + \|Bx\| \leq \|A\| + \|B\|.$$

On en déduit  $\|A+B\| \leq \|A\| + \|B\|$ , c'est-à-dire l'inégalité triangulaire.

Conclusion : la norme opérateur est bien une norme. □

**Proposition 4.25.** Soient  $A, B \in \mathcal{M}_n(\mathbb{R})$  et  $x \in \mathbb{R}^n$ . Alors

1.  $\|Ax\| \leq \|A\| \|x\|$ ,
2.  $\|AB\| \leq \|A\| \|B\|$ .

*Démonstration.* Pour tout  $x \in \mathbb{R}^n \setminus \{0\}$ ,  $\|A\| \geq \frac{\|Ax\|}{\|x\|}$ . Donc

$$\|Ax\| \leq \|A\| \|x\|.$$

Ensuite, pour tout  $x \in \mathbb{R}^n$ ,

$$\|ABx\| = \|A(Bx)\| \leq \|A\| \|Bx\| \leq \|A\| \|B\| \|x\|.$$

Donc  $\|AB\| \leq \|A\| \|B\|$ . □

*Exercice 4.26.* On considère la matrice

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

Calculer la norme opérateur de  $A$  lorsqu'on prend sur  $\mathbb{R}^2$  la norme  $\|\cdot\|_1$ , puis la norme  $\|\cdot\|_\infty$ .

*Exercice 4.27.* Sur  $\mathcal{M}_n(\mathbb{R})$ , on définit l'application  $N$  par : pour tout  $A = (a_{ij})_{\substack{1 \leq i \leq n, \\ 1 \leq j \leq n}}$ ,

$$N(A) = \max_{\substack{1 \leq i \leq n, \\ 1 \leq j \leq n}} |a_{ij}|.$$

1. Montrer que  $N$  est une norme sur  $\mathcal{M}_n(\mathbb{R})$ .
2. Montrer qu'il n'existe pas de norme  $\|\cdot\|$  sur  $\mathbb{R}^n$  telle que  $N$  soit la norme opérateur associée à  $\|\cdot\|$ .

*Exemple 4.28.* Soit  $A \in \mathcal{M}_2(\mathbb{R})$  une matrice qui a deux valeurs propres distinctes de module  $< 1$ . Alors la suite  $(A^n)_{n \in \mathbb{N}}$  tend vers 0.

Comme  $A$  a deux valeurs propres distinctes (a priori complexes), elle est diagonalisable sur  $\mathbb{C}$  : il existe  $P \in GL_2(\mathbb{C})$  et une matrice diagonale  $D \in \mathcal{M}_2(\mathbb{R})$  tels que  $A = PDP^{-1}$ . Les éléments diagonaux de  $D$  sont les valeurs propres  $\lambda_1$  et  $\lambda_2$  de  $A$ . Alors pour tout  $n \in \mathbb{N}$ ,  $A^n = PD^nP^{-1}$ .

On introduit sur  $\mathcal{M}_2(\mathbb{R})$  une norme bien adaptée à notre problème. On considère d'abord la norme  $N_1$  définie sur  $\mathcal{M}_2(\mathbb{C})$  par :

$$\forall C = (c_{ij})_{1 \leq i, j \leq 2} \in \mathcal{M}_2(\mathbb{C}), \quad N_1(C) = \sum_{1 \leq i, j \leq 2} |c_{ij}|.$$

On définit maintenant la norme  $N(C) = N_1(P^{-1}CP)$  (on admet ici qu'il s'agit bien d'une norme).

Il reste à observer que

$$N(A^n) = N(PD^nP^{-1}) = N_1(P^{-1}(PD^nP^{-1})P) = N_1(D^n) = |\lambda_1|^n + |\lambda_2|^n.$$

Comme  $|\lambda_1| < 1$  et  $|\lambda_2| < 1$ , on en déduit  $\lim_{n \rightarrow +\infty} N(A^n) = 0$ , q.e.d.

## Espace de polynômes

Considérons à présent l'ensemble des polynômes de degré inférieur ou égal à 2, c'est-à-dire l'ensemble des fonctions  $f : \mathbb{R} \rightarrow \mathbb{R}$  de la forme  $x \mapsto ax^2 + bx + c$ .

Une base de cet ensemble est donnée par les trois éléments :

$$f_0 : x \mapsto x^2, \quad f_1 : x \mapsto x, \quad f_2 : x \mapsto 1.$$

*Exercice 4.29.* Montrer que l'application qui à un polynôme de degré  $\leq 2$  :  $f : x \mapsto ax^2 + bx + c$ , associe la quantité  $N_1(f) = |a| + |b| + |c|$  est une norme sur l'espace des polynômes de degré  $\leq 2$ .

## Espaces de dimension finie

*Exercice 4.30.* Soit  $E$  un espace vectoriel de dimension finie  $n \geq 1$  et  $\{\varepsilon_1, \dots, \varepsilon_n\}$  une base de  $E$ . On note  $\{e_1, \dots, e_n\}$  la base canonique de  $\mathbb{R}^n$ .

1. Justifier qu'il existe une unique application linéaire  $\Phi : E \rightarrow \mathbb{R}^n$  telle que  $\Phi(\varepsilon_i) = e_i$  pour tout  $1 \leq i \leq n$ .
2. Soit  $N$  une norme sur  $\mathbb{R}^n$ . Montrer que l'application  $x \in E \mapsto N(\Phi(x))$  est une norme sur  $E$ .
3. Soit  $N'$  une norme sur  $E$ . Montrer qu'il existe une norme  $N$  sur  $\mathbb{R}^n$  telle que  $N(\Phi(x)) = N'(x)$  pour tout  $x \in E$ .

### 4.3.2 Des exemples de normes en dimension infinie

Les espaces vectoriels de fonctions sont (généralement) des espaces vectoriels de dimension infinie. Par exemple, étant donné un intervalle non vide  $I$  de  $\mathbb{R}$ , on peut munir l'ensemble  $\mathcal{F}(I; \mathbb{R})$  des fonctions définies sur  $I$  et à valeurs dans  $\mathbb{R}$ , de deux opérations naturelles : pour tout  $f, g \in \mathcal{F}(I; \mathbb{R})$  et  $\lambda \in \mathbb{R}$ , on définit les fonctions  $f + g$  et  $\lambda \cdot f$  par

$$\forall x \in I, (f + g)(x) = f(x) + g(x) \quad , \quad (\lambda \cdot f)(x) = \lambda f(x).$$

Muni de ces deux opérations, l'ensemble  $\mathcal{F}(I; \mathbb{R})$  devient un espace vectoriel sur  $\mathbb{R}$ . Plus généralement, en notant  $\mathbb{K}$  l'ensemble des nombres réels ou complexes (autrement dit,  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{K} = \mathbb{C}$ ), l'ensemble  $\mathcal{F}(A; E)$  des fonctions d'un ensemble quelconque  $A$  à valeurs dans un espace vectoriel  $E$  sur  $\mathbb{K}$ , muni des deux opérations naturelles d'addition et de multiplication par un nombre *point par point*, est un espace vectoriel sur  $\mathbb{K}$ .

Un autre exemple est l'ensemble des fonctions continues  $\mathcal{C}^0([a, b]; \mathbb{R})$  du segment  $[a, b]$  à valeurs réelles. En fait, il s'agit d'un sous-espace vectoriel de  $\mathcal{F}([a, b]; \mathbb{R})$ . Pour le voir, observons que l'ensemble  $\mathcal{C}^0([a, b]; \mathbb{R})$  est non vide, par exemple parce qu'il contient les fonctions constantes. Pour tout  $f, g \in \mathcal{C}^0([a, b]; \mathbb{R})$  et pour tout  $\lambda \in \mathbb{R}$ ,  $\lambda f$  est continu (comme produit d'une fonction continue par un nombre) puis  $\lambda f + g$  est continue (comme somme de deux fonctions continues). Ainsi,  $\mathcal{C}^0([a, b]; \mathbb{R})$  est stable pour l'addition et la multiplication par les nombres. En conclusion,  $\mathcal{C}^0([a, b]; \mathbb{R})$  est un sous-espace vectoriel de  $\mathcal{F}([a, b]; \mathbb{R})$  et donc un espace vectoriel.

**Proposition 4.31.** Sur l'espace vectoriel  $\mathcal{C}^0([a, b], \mathbb{C})$ , on considère la fonction

$$f \mapsto \int_a^b |f(t)| dt.$$

Il s'agit d'une norme. On la note  $\|f\|_1$ .

*Démonstration.* D'abord,  $f \mapsto \int_a^b |f|$  est une application bien définie sur  $\mathcal{C}^0([a, b], \mathbb{C})$  et à valeurs positives. Ensuite,  $\int_a^b |0| = 0$  et si  $f \in \mathcal{C}^0([a, b], \mathbb{C})$  est telle que  $\int_a^b |f(t)| dt = 0$ , alors la fonction continue positive  $t \mapsto |f(t)|$  est d'intégrale nulle, ce qui implique que  $|f(t)| = 0$  pour tout  $t \in [a, b]$ . On a donc  $f \equiv 0$ . Cela montre la *séparation*.

Pour montrer la *positive homogénéité*, soient  $f \in \mathcal{C}^0([a, b], \mathbb{C})$  et  $\lambda \in \mathbb{R}$ . Pour tout  $t \in [a, b]$ ,  $|(\lambda f)(t)| = |\lambda| |f(t)|$ . En intégrant cette identité sur  $[a, b]$  et en utilisant la linéarité de l'intégrale, on obtient bien

$$\int_a^b |(\lambda f)(t)| dt = \int_a^b |\lambda| |f(t)| dt = |\lambda| \int_a^b |f(t)| dt.$$

Enfin, pour l'*inégalité triangulaire*, soient  $f, g \in \mathcal{C}^0([a, b], \mathbb{C})$ . Alors pour tout  $t \in [a, b]$ ,

$$|(f + g)(t)| = |f(t) + g(t)| \leq |f(t)| + |g(t)|.$$

En intégrant cette inégalité et en utilisant la monotonie puis la linéarité de l'intégrale, on obtient bien

$$\int_a^b |(f + g)(t)| dt \leq \int_a^b (|f(t)| + |g(t)|) dt = \int_a^b |f(t)| dt + \int_a^b |g(t)| dt.$$

On peut conclure que  $f \mapsto \int_a^b |f|$  est une norme. □

Il y a beaucoup d'autres normes sur l'espace  $\mathcal{C}^0([a, b]; \mathbb{C})$ . On en a déjà vu deux au chapitre précédent. La première était la norme infinie :

$$\forall f \in \mathcal{C}^0([a, b]; \mathbb{C}), \quad \|f\|_\infty = \sup_{x \in [a, b]} |f(x)| = \sup\{|f(x)|, x \in [a, b]\}.$$

La seconde était la norme  $\|\cdot\|_2$ , qui intervient naturellement dans le théorème de Parseval. On a également évoqué son lien avec le produit scalaire complexe. De manière générale, les produits scalaires donnent naturellement naissance à des normes. On consacre à ces notions le paragraphe suivant, dans le cadre réel plutôt que complexe.

## Norme et produit scalaire

On rappelle d'abord la définition d'un produit scalaire.

**Définition 4.32.** Une application  $\langle \cdot, \cdot \rangle : E \times E \rightarrow \mathbb{R}$  est un produit scalaire si elle vérifie les propriétés suivantes :

1. (linéarité par rapport à la première variable) pour tout  $w \in E$ , la fonction  $v \in E \mapsto \langle v, w \rangle \in \mathbb{R}$  est linéaire,
2. (linéarité par rapport à la deuxième variable) pour tout  $v \in E$ , la fonction  $w \in E \mapsto \langle v, w \rangle \in \mathbb{R}$  est linéaire,
3. (symétrie) pour tout  $v, w \in E$ ,  $\langle v, w \rangle = \langle w, v \rangle$ ,
4. (défini positif) pour tout  $v \in E$ ,  $\langle v, v \rangle \geq 0$  avec égalité si et seulement si  $v = 0$ .

*Exemple 4.33.* Sur  $\mathbb{R}^n$ , on définit  $\langle x, y \rangle = \sum_{j=1}^n x_j y_j$  en notant  $x = (x_1, \dots, x_n)$ ,  $y = (y_1, \dots, y_n)$ . On vérifie qu'il s'agit d'un produit scalaire.

*Exemple 4.34.* Dans  $\mathcal{C}^0([a, b]; \mathbb{R})$ ,  $(f, g) \mapsto \int_a^b f(t)g(t) dt$  est un produit scalaire.

En effet, par linéarité de l'intégrale, on a bien linéarité par rapport à la première variable  $f$  et linéarité par rapport à la seconde variable  $g$ . Pour montrer la symétrie, on observe que

$$\int_a^b f(t)g(t) dt = \int_a^b g(t)f(t) dt.$$

Enfin,

$$\int_a^b f(t)f(t) dt = \int_a^b f(t)^2 dt \geq 0$$

avec égalité si et seulement si  $f = 0$  (on utilise ici que l'intégrale d'une fonction continue positive est nulle si et seulement si la fonction est partout nulle). Cela montre que l'application est bien définie positive.

**Proposition 4.35** (Inégalité de Cauchy-Schwarz). *Soit  $E$  un espace vectoriel muni du produit scalaire  $\langle \cdot, \cdot \rangle$ . Alors, pour tout  $v, w \in E$ ,*

$$|\langle v, w \rangle| \leq \sqrt{\langle v, v \rangle} \sqrt{\langle w, w \rangle}.$$

*Démonstration.* Soit  $t \in \mathbb{R}$ . Alors

$$0 \leq \langle v + tw, v + tw \rangle = t^2 \langle w, w \rangle + 2t \langle v, w \rangle + \langle v, v \rangle.$$

Comme le trinôme du second degré du membre de droite ne change pas de signe, c'est que son discriminant est  $\leq 0$  :

$$\langle v, w \rangle^2 \leq \langle v, v \rangle \langle w, w \rangle.$$

On en déduit l'inégalité désirée en prenant les racines carrées de chaque membre. □

L'inégalité de Cauchy-Schwarz nous permet de définir une norme sur  $E$ .

**Proposition 4.36.** *Soit  $E$  un espace vectoriel muni du produit scalaire  $\langle \cdot, \cdot \rangle$ . L'application*

$$v \in E \mapsto \sqrt{\langle v, v \rangle}$$

*est une norme sur  $E$ .*

*Démonstration.* Il s'agit bien d'une application de  $E$  dans  $\mathbb{R}^+$ . Ensuite,  $\sqrt{\langle 0, 0 \rangle} = 0$ . Réciproquement, soit  $v \in E$  tel que  $\sqrt{\langle v, v \rangle} = 0$ . Comme le produit scalaire est défini positif, cela montre que  $v = 0$ . On en déduit la séparation.

Pour la positive homogénéité, soient  $v \in E$  et  $\lambda \in \mathbb{R}$ . Alors,

$$\sqrt{\langle \lambda v, \lambda v \rangle} = \sqrt{\lambda^2 \langle v, v \rangle} = \sqrt{\lambda^2} \sqrt{\langle v, v \rangle} = |\lambda| \sqrt{\langle v, v \rangle},$$

comme attendu.

Enfin, pour l'inégalité triangulaire, soient  $v, w \in E$ . Alors par bilinéarité et symétrie du produit scalaire,

$$\langle v + w, v + w \rangle = \langle v, v \rangle + \langle w, w \rangle + 2\langle v, w \rangle.$$

Par l'inégalité de Cauchy-Schwarz,

$$\langle v + w, v + w \rangle \leq \langle v, v \rangle + \langle w, w \rangle + 2\langle v, v \rangle \langle w, w \rangle = (\sqrt{\langle v, v \rangle} + \sqrt{\langle w, w \rangle})^2.$$

En prenant les racines carrées, il vient

$$\sqrt{\langle v + w, v + w \rangle} \leq \sqrt{\langle v, v \rangle} + \sqrt{\langle w, w \rangle},$$

ce qui permet de conclure. □

Sur l'espace  $\mathcal{C}^0([a, b]; \mathbb{R})$ , on a déjà introduit le produit scalaire

$$(f, g) \mapsto \int_a^b f(t)g(t) dt.$$

On en déduit une nouvelle norme sur  $\mathcal{C}^0([a, b]; \mathbb{R})$  :

$$f \in \mathcal{C}^0([a, b]; \mathbb{R}) \mapsto \sqrt{\int_a^b f(t)^2 dt}.$$

On la note  $\|f\|_2$ .

Grâce à l'inégalité de Cauchy-Schwarz, on peut contrôler la norme  $\|\cdot\|_1$  par la norme  $\|\cdot\|_2$ .

*Exemple 4.37.* Pour tout  $f \in \mathcal{C}^0([a, b]; \mathbb{R})$ ,  $\|f\|_1 \leq \sqrt{b-a}\|f\|_2$ , autrement dit

$$\int_a^b |f(t)| dt \leq \sqrt{b-a} \sqrt{\int_a^b |f(t)|^2 dt}.$$

Pour le voir, on applique l'inégalité de Cauchy-Schwarz aux fonctions  $t \mapsto |f(t)|$  et  $t \mapsto 1$ , ce qui donne

$$\int_a^b |f(t)| dt \leq \sqrt{\int_a^b 1^2 dt} \sqrt{\int_a^b |f(t)|^2 dt} = \sqrt{b-a} \sqrt{\int_a^b |f(t)|^2 dt}.$$

On peut également contrôler la norme  $\|\cdot\|_1$  et la norme  $\|\cdot\|_2$  par la norme  $\|\cdot\|_\infty$ .

*Exemple 4.38.* Pour tout  $f \in \mathcal{C}^0([a, b]; \mathbb{R})$ ,

$$\|f\|_1 \leq (b-a)\|f\|_\infty, \quad \|f\|_2 \leq \sqrt{b-a}\|f\|_\infty.$$

En effet, par définition de  $\|f\|_\infty$ , on sait que  $\|f\|_\infty$  est un majorant de l'ensemble  $\{|f(x)| : x \in [a, b]\}$ . Ainsi, pour tout  $x \in [a, b]$ , on a

$$|f(x)| \leq \|f\|_\infty. \tag{4.2}$$

En intégrant l'inégalité précédente, il vient

$$\int_a^b |f(x)| dx \leq \int_a^b \|f\|_\infty dx = (b-a)\|f\|_\infty,$$

c'est-à-dire  $\|f\|_1 \leq (b-a)\|f\|_\infty$ . On déduit aussi de (4.2) que pour tout  $x \in [a, b]$ ,

$$|f(x)|^2 \leq \|f\|_\infty^2.$$

Par intégration sur  $[a, b]$ , il vient

$$\int_a^b |f(x)|^2 dx \leq (b-a)\|f\|_\infty^2$$

d'où, en prenant les racines carrées,  $\|f\|_2 \leq \sqrt{b-a}\|f\|_\infty$ .

Au regard de ces deux exemples, on peut se demander si les normes  $\|\cdot\|_1$ ,  $\|\cdot\|_2$  et  $\|\cdot\|_\infty$  sont équivalentes. Comme on va le voir dans la section suivante, il n'en est rien.

## Non équivalence des normes

*Méthode 4.39.* Pour montrer que deux normes  $N$  et  $N'$  ne sont pas équivalentes, il suffit de trouver une suite  $(x_n)_{n \in \mathbb{N}}$  dans  $E$  qui converge vers 0 pour  $N$  mais pas pour  $N'$ .

Pour justifier cette méthode, rappelons que deux normes  $N$  et  $N'$  sont équivalentes sur  $E$  s'il existe deux constantes  $C > 0$  et  $C' > 0$  telles que pour tout  $x \in E$ ,  $N(x) \leq CN'(x)$  et  $N'(x) \leq C'N(x)$ .

Supposons avoir trouvé une suite  $(x_n)_{n \in \mathbb{N}}$  dans  $E$  qui converge vers 0 pour  $N$  mais pas pour  $N'$ . Alors, il ne peut exister  $C' > 0$  tel que  $N'(x) \leq C'N(x)$  pour tout  $x \in E$  car sinon, pour tout  $n \in \mathbb{N}$ ,

$$N'(x_n) \leq C'N(x_n)$$

et puisque  $\lim_{n \rightarrow +\infty} N(x_n) = 0$ , on en déduirait que  $\lim_{n \rightarrow +\infty} N'(x_n) = 0$ , une contradiction. Ainsi,  $N$  et  $N'$  ne peuvent pas être équivalentes.

*Exemple 4.40.* Sur l'espace vectoriel  $\mathbb{R}[X]$  des polynômes à coefficients réels, on définit les deux applications  $N$  et  $N'$  suivantes : étant donné un polynôme  $P(x) = a_0 + a_1x + \dots + a_nx^n$ ,  $x \in \mathbb{R}$ ,

$$N(P) := \max_{0 \leq i \leq n} |a_i| \quad , \quad N'(P) := \sum_{0 \leq i \leq n} |a_i|.$$

Alors  $N$  et  $N'$  sont des normes sur  $\mathbb{R}[X]$  qui ne sont pas équivalentes.

Admettons ici que  $N$  et  $N'$  sont des normes sur  $\mathbb{R}[X]$  et montrons qu'elles ne sont pas équivalentes. Pour tout  $n \in \mathbb{N}$ , on définit le polynôme

$$Q_n(x) = \frac{1}{n+1}(1 + x + \dots + x^n).$$

Alors  $N(Q_n) = \frac{1}{n+1}$  tandis que  $N'(Q_n) = 1$ . On en déduit que la suite  $(Q_n)_{n \in \mathbb{N}}$  converge vers 0 pour la norme  $N$  mais pas pour la norme  $N'$ .

*Exemple 4.41.* On considère sur  $\mathcal{C}^0([0, 1]; \mathbb{R})$  les deux normes suivantes :

$$\|f\|_1 = \int_0^1 |f(t)| dt \quad , \quad \|f\|_2 = \sqrt{\int_0^1 |f(t)|^2 dt}.$$

Ces deux normes ne sont pas équivalentes.

Pour montrer que ces deux normes ne sont pas équivalentes, il suffit de trouver une suite  $(f_n)_{n \in \mathbb{N}^*} \subset \mathcal{C}^0([0, 1]; \mathbb{R})$  telle que  $\|f_n\|_1 \rightarrow 0$  alors que  $\|f_n\|_2 \not\rightarrow 0$ . On prend

$$f_n(x) = \begin{cases} -n^{3/2}(x - 1/n) & \text{si } 0 \leq x \leq \frac{1}{n}, \\ 0 & \text{sinon.} \end{cases}$$

Alors  $\|f_n\|_1 = \frac{1}{2\sqrt{n}}$  (il s'agit de l'aire sous la courbe) et donc  $(f_n)_{n \in \mathbb{N}^*}$  tend vers 0 pour la norme  $\|\cdot\|_1$ . Pour estimer  $\|f_n\|_2$ , on observe que

$$\|f_n\|_2^2 = \frac{1}{3}$$

En particulier,  $(\|f_n\|_2)_{n \in \mathbb{N}^*}$  ne tend pas vers 0.

Conclusion : ces deux normes ne sont pas équivalentes sur  $\mathcal{C}^0([0, 1]; \mathbb{R})$ .

*Exercice 4.42.* Montrer que les normes  $\|\cdot\|_1$  et  $\|\cdot\|_\infty$  ne sont pas équivalentes.

### 4.3.3 Ouverts, fermés, bornés

Soit  $E$  un espace vectoriel muni d'une norme  $N$ . Cela permet en particulier de définir des boules (ouvertes ou fermées sur  $E$ ).

**Définition 4.43.** Soit  $A \subset E$ .

1. L'intérieur de  $A$  est l'ensemble des points intérieurs à  $A$ . On dit que  $A$  est ouvert s'il est égal à son intérieur.
2. L'adhérence de  $A$  est l'ensemble des points adhérents à  $A$ . On dit que  $A$  est fermé s'il est égal à son adhérence.

L'adhérence du complémentaire est le complémentaire de l'intérieur :

**Proposition 4.44.** Soit  $A$  une partie de  $E$ . Alors

$$\text{adh}(\mathbb{R}^n \setminus A) = \mathbb{R}^n \setminus (\text{int } A).$$

*Démonstration.* Un point  $x$  est adhérent à  $\mathbb{R}^n \setminus A$  si et seulement si

$$\forall r > 0, \quad B(x, r) \cap (\mathbb{R}^n \setminus A) \neq \emptyset.$$

De manière équivalente,

$$\forall r > 0, \quad B(x, r) \not\subset A.$$

En d'autres termes,

$$x \notin \text{int } A.$$

Cela revient à dire que  $x \in \mathbb{R}^n \setminus (\text{int } A)$ . □

**Proposition 4.45.** Une partie  $F$  de  $E$  est fermée si et seulement si son complémentaire est ouvert.

*Démonstration.* La partie  $F$  est fermée si et seulement si  $\text{adh } F = F$ . Son complémentaire est ouvert si et seulement si  $\text{int}(\mathbb{R}^n \setminus F) = \mathbb{R}^n \setminus F$ . Par ailleurs, en appliquant la proposition précédente à  $A = \mathbb{R}^n \setminus F$ , on a

$$\text{adh } F = \mathbb{R}^n \setminus (\text{int}(\mathbb{R}^n \setminus F)).$$

Donc

$$F = \text{adh } F \iff F = \mathbb{R}^n \setminus (\text{int}(\mathbb{R}^n \setminus F)) \iff \mathbb{R}^n \setminus F = \text{int}(\mathbb{R}^n \setminus F),$$

ce qui achève la preuve. □

**Proposition 4.46.** Une boule ouverte est un ouvert.

*Démonstration.* Soit  $B(a, r)$  une boule ouverte de  $\mathbb{R}^n$ . Montrons que c'est un ouvert. Soit  $x \in B(a, r)$ . Il s'agit de voir que  $x$  est dans l'intérieur de  $B(a, r)$ . Posons  $s := r - \|x - a\|$  (observer que  $s > 0$ ) et vérifions que  $B(x, s) \subset B(a, r)$ . Soit  $y \in B(x, s)$ . Alors  $\|y - x\| < s$  et donc par l'inégalité triangulaire

$$\|y - a\| \leq \|y - x\| + \|x - a\| < s + \|x - a\| = r.$$

Ainsi,  $y \in B(a, r)$ . Comme  $y$  est arbitraire, cela montre que  $B(x, s) \subset B(a, r)$ . En conclusion,  $B(a, r)$  est un ouvert. □

**Proposition 4.47.** 1. La réunion de deux ouverts (ou même d'une infinité d'ouverts) est un ouvert.

2. L'intersection de deux ouverts (ou même d'un nombre fini d'ouverts) est un ouvert.

*Exercice 4.48.* \* Soit  $\Omega \subset E$ . Alors  $\Omega$  est un ouvert si et seulement si  $\Omega$  s'écrit comme une réunion infinie de boules ouvertes.

**Solution :** Si  $\Omega$  est un ouvert, alors pour tout  $x \in \Omega$ , il existe  $r_x > 0$  tel que  $B(x, r_x) \subset \Omega$ . On observe alors que

$$\Omega = \cup_{x \in \Omega} B(x, r_x).$$

En effet, l'inclusion  $\subset$  provient du fait que chaque  $x \in \Omega$  appartient à  $B(x, r_x)$  et donc appartient à la réunion  $\cup_{x \in \Omega} B(x, r_x)$ . L'inclusion  $\supset$  est conséquence du fait que  $B(x, r_x) \subset \Omega$  pour chaque  $x \in \Omega$ . L'égalité est démontrée et  $\Omega$  est donc bien une réunion de boules ouvertes.

Dans l'autre sens, on utilise qu'une réunion infinie d'ouverts est un ouvert.

En dimension finie, la notion d'ouvert ne dépend pas de la norme qu'on met sur l'espace :

**Proposition 4.49.** On considère deux normes  $N_1$  et  $N_2$  sur un espace vectoriel de dimension finie  $E$ . Soit  $\Omega \subset E$ . Si  $\Omega$  est un ouvert pour la norme  $N_1$ , alors  $\Omega$  est un ouvert pour la norme  $N_2$ .

*Démonstration.* Soit  $\Omega$  un ouvert pour la norme  $N_1$ . Soit  $x \in \Omega$ . Comme  $\Omega$  est un ouvert pour  $N_1$ , il existe  $r_x > 0$  tel que

$$B_{N_1}(x, r_x) = \{y \in E : N_1(y - x) < r_x\} \subset \Omega.$$

Comme  $N_1$  et  $N_2$  sont équivalentes, il existe  $C > 0$  tel que  $N_1 \leq CN_2$ . Par la remarque 4.19, on en déduit que

$$B_{N_2}(x, \frac{r_x}{C}) \subset B_{N_1}(x, r_x) \subset \Omega.$$

Donc  $\Omega$  est un ouvert pour la norme  $N_2$ . □

*Exemple 4.50.* Un demi-plan dans  $\mathbb{R}^2$  (et plus généralement un demi-espace dans  $\mathbb{R}^n$ ) est un ouvert.

On considère un demi-plan de  $\mathbb{R}^2$  : une droite de  $\mathbb{R}^2$  sépare un plan en deux parties qui ne contiennent pas cette droite. C'est ce qu'on appelle (ici) les demi-plans associés à cette droite.

Montrons qu'un tel demi-plan est ouvert (pour n'importe quelle norme, puisqu'on a vu que la notion d'ouvert ne dépendait pas de la norme équivalente qu'on choisissait et que toutes les normes sont équivalentes dans  $\mathbb{R}^2$ ). On va prendre dans la suite la norme  $\|\cdot\|_2$ . On se donne donc une droite  $\Delta$  et on note  $\Pi$  l'un des demi-plans définis par cette droite. Soit  $x \in \Pi$ . On note  $y$  la projection orthogonale de  $x$  sur  $\Delta$ . Comme  $x \notin \Delta$ ,  $y \neq x$  et donc  $\|x - y\|_2 > 0$ . Montrons que

$$B_{\|\cdot\|_2}(x, \|x - y\|_2) \subset \Pi.$$

Soit  $z \notin \Pi$ . La demi-droite de sommet  $x$  passant par  $z$  coupe  $\Delta$  en  $z'$ . Alors

$$\|x - z\|_2^2 \geq \|x - z'\|_2^2 = \|x - y\|_2^2 + \|y - z'\|_2^2.$$

La dernière égalité résulte du théorème de Pythagore. On en déduit

$$\|x - z\|_2 \geq \|x - y\|_2,$$

et donc  $z \notin B_{\|\cdot\|_2}(x, \|x - y\|_2)$ . Par contraposée, on a montré que pour tout  $z \in B_{\|\cdot\|_2}(x, \|x - y\|_2)$ , on a  $z \in \Pi$ . Conclusion :  $B_{\|\cdot\|_2}(x, \|x - y\|_2) \subset \Pi$ . On en déduit que  $\Pi$  est un ouvert. □

*Exemple 4.51.* Comme exemples de fermés, on peut citer :

1. une boule fermée dans  $\mathbb{R}^n$  munie d'une norme  $N$ ,
2. une droite dans  $\mathbb{R}^2$ .

1. Soit  $\overline{B(a, r)}$  une boule fermée dans un e.v.n.  $(E, N)$ . Montrons que c'est un fermé, c'est-à-dire que son complémentaire est un ouvert. Soit  $x \notin \overline{B(a, r)}$ . Donc  $N(x - a) > r$ . Soit  $r_x = N(x - a) - r$  et montrons que  $B(x, r_x) \subset E \setminus \overline{B(a, r)}$ . Soit  $y \in B(x, r_x)$ . Alors

$$N(y - a) \geq N(x - a) - N(y - x) > N(x - a) - r_x = r.$$

Donc  $y \notin \overline{B(a, r)}$ . Donc  $B(x, r_x) \subset E \setminus \overline{B(a, r)}$ . Donc  $E \setminus \overline{B(a, r)}$  est un ouvert, ce qui montre que  $\overline{B(a, r)}$  est un fermé.

2. Soit  $\Delta$  une droite de  $\mathbb{R}^2$ . Le complémentaire de cette droite est constitué de deux demi-plans, dont on a montré qu'ils étaient ouverts. Comme la réunion de deux ouverts est un ouvert, on en déduit que le complémentaire d'une droite est un ouvert. Donc une droite est un fermé.

**Définition 4.52.** Soit  $A$  une partie d'un espace vectoriel normé  $E$ . La frontière de  $A$ , notée  $\partial A$ , est l'ensemble des points  $x$  de  $E$  tels que pour tout  $r > 0$ ,

$$B(x, r) \cap A \neq \emptyset \quad , \quad B(x, r) \setminus A \neq \emptyset.$$

Autrement dit,  $\partial A = \text{adh } A \setminus \text{int } A$ .

Comme d'habitude, cette notion ne dépend pas de la norme équivalente choisie.

**Exercice 4.53.** \* Soit  $x \in \mathbb{R}^n$  et  $r > 0$ . Alors montrer que

1. la sphère  $S(x, r)$  est la frontière de la boule ouverte  $B(x, r)$  et aussi la frontière de la boule fermée  $\overline{B(x, r)}$ .
2. la boule fermée  $\overline{B(x, r)}$  est l'adhérence de la boule ouverte  $B(x, r)$ ,
3. la boule ouverte  $B(x, r)$  est l'intérieur de la boule fermée  $\overline{B(x, r)}$ .

**Définition 4.54.** 1. Une partie  $A \subset E$  est dite bornée si elle est contenue dans une boule (ouverte ou fermée) centrée en 0 : il existe  $r > 0$  tel que  $A \subset B(0, r)$ .

2. Une suite d'éléments de  $E$  est dite bornée si tous ses termes sont contenus dans une partie bornée de  $E$ .

**Exercice 4.55.** Soient  $N$  et  $N'$  deux normes sur un espace vectoriel de dimension finie  $E$ . Soit  $A$  une partie de  $E$ . Montrer que si  $A$  est bornée pour  $N$ , alors  $A$  est bornée pour  $N'$ .

**Exercice 4.56.** \* Soit  $N$  une norme sur  $E$ . Soit  $A$  une partie de  $E$ . On suppose que  $A$  n'est pas bornée. Montrer qu'il existe une suite  $(x_n)_{n \in \mathbb{N}}$  contenue dans  $A$  et telle que  $\lim_{n \rightarrow +\infty} N(x_n) = +\infty$ . La réciproque est-elle vraie ?

**Exercice 4.57.** Une suite de  $E$  convergente est bornée.

**Solution :** Soit  $(x_m)_{m \in \mathbb{N}}$  une suite convergente. Notons  $\ell \in E$  sa limite. Alors il existe  $m_0 \in \mathbb{N}$  tel que pour tout  $m \geq m_0$ ,  $\|x_m - \ell\| \leq 1$ . On a alors  $\|x_m\| \leq \|x_m - \ell\| + \|\ell\| \leq 1 + \|\ell\|$ . Notons  $M := 1 + \|\ell\| + \max_{m < m_0} \|x_m\|$ . En distinguant les cas  $m < m_0$  et  $m \geq m_0$ , on vérifie que pour tout  $m \in \mathbb{N}$ ,

$$\|x_m\| \leq M.$$

Ainsi, la suite  $(x_m)_{m \in \mathbb{N}}$  est contenue dans la boule  $\overline{B(0, M)}$ . Elle est donc bornée. □

## Critères séquentiels

On dispose d'une caractérisation séquentielle (i.e. faisant intervenir les suites) des fermés :

**Exercice 4.58.** Soit  $E$  un espace vectoriel muni d'une norme. Une partie  $F \subset E$  est fermée si et seulement si toute suite d'éléments de  $F$  qui converge dans  $E$  a sa limite dans  $F$ .

**Solution :** Supposons que  $F$  est fermé. Soit  $(x_n)_{n \in \mathbb{N}}$  une suite d'éléments de  $F$  qui converge vers un vecteur  $\ell \in \mathbb{R}^n$ . Montrons que  $\ell \in F$ . Supposons par l'absurde que  $\ell$  appartienne au complémentaire de  $F$  qui est un ouvert. Par définition des ouverts, il existe  $r > 0$  tel que  $B(\ell, r) \subset E \setminus F$ . Comme  $(x_n)_{n \in \mathbb{N}}$  tend vers  $\ell$ , il existe  $n_0$  tel que  $x_{n_0} \in B(\ell, r)$  et donc  $x_{n_0} \notin F$  : contradiction. Ainsi,  $\ell \in F$ .

Réciproquement, supposons que toute suite d'éléments de  $F$  qui converge dans  $E$  a sa limite dans  $F$  et montrons que  $F$  est fermé, autrement dit que  $E \setminus F$  est ouvert. Soit  $\ell \in E \setminus F$  et montrons qu'il existe  $r > 0$  tel que  $B(\ell, r) \subset E \setminus F$  en procédant par l'absurde. Sinon, pour tout  $n \in \mathbb{N}$ , il existe  $x_n \in B(\ell, \frac{1}{n+1}) \cap F$ . Alors la suite  $(x_n)_{n \in \mathbb{N}}$  tend vers  $\ell$ . Par hypothèse,  $\ell \in F$  : contradiction. On en déduit l'existence de  $r > 0$  tel que  $B(\ell, r) \subset E \setminus F$ , donc que  $E \setminus F$  est ouvert, donc que  $F$  est fermé. □

## 4.4 Rappels sur les espaces vectoriels

### 4.4.1 Espaces vectoriels

Dans tout ce qui suit,  $\mathbb{K}$  désigne l'ensemble  $\mathbb{R}$  ou  $\mathbb{C}$ .

**Définition 4.59.** Soit  $E$  un ensemble sur lequel on a défini deux opérations :

- une addition, notée  $+$ , qui à deux éléments  $x$  et  $y$  de  $E$ , associe un troisième élément de  $E$  :  $x + y$ ,
- une multiplication, notée  $\cdot$ , qui à un élément  $x$  de  $E$  et un nombre  $\lambda \in \mathbb{K}$ , associe un autre élément de  $E$  :  $\lambda \cdot x$ .

On dit alors que  $E$  est un espace vectoriel sur  $\mathbb{K}$  lorsque l'addition et la multiplication vérifient les propriétés suivantes :

1. Pour tout  $x, y, z \in E$ ,

$$x + (y + z) = (x + y) + z \quad , \quad x + y = y + x.$$

2. Il existe un élément de  $E$ , noté  $0_E$ , tel que pour tout  $x \in E$ ,

$$x + 0_E = 0_E + x = x.$$

3. Pour tout  $x \in E$ , il existe un élément de  $E$ , noté  $-x$ , tel que

$$x + (-x) = (-x) + x = 0_E.$$

4. Pour tout  $x, y \in E$ , pour tout  $\lambda \in \mathbb{K}$ ,

$$\lambda \cdot (x + y) = (\lambda \cdot x) + (\lambda \cdot y) \quad , \quad (\lambda + \mu) \cdot x = (\lambda \cdot x) + (\mu \cdot x),$$

$$\lambda \cdot (\mu \cdot x) = (\lambda\mu) \cdot x \quad , \quad 1 \cdot x = x.$$

L'exemple le plus simple d'espace vectoriel sur  $\mathbb{R}$  est  $\mathbb{R}$  lui-même, muni de l'addition et de la multiplication usuelles. L'ensemble  $\mathbb{R}^2$  des couples de réels, muni de l'addition et de la multiplication usuelles définies coordonnées par coordonnées :

$$\forall (x_1, x_2), (y_1, y_2) \in \mathbb{R}^2, (x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2) \quad , \quad \lambda \cdot (x_1, y_1) = (\lambda x_1, \lambda y_1)$$

est aussi un espace vectoriel. Plus généralement, l'ensemble des  $n$ -uplets  $\mathbb{R}^n$  est un espace vectoriel lorsqu'on le munit de l'addition et de la multiplication coordonnées par coordonnées. En fait, le produit de deux espaces vectoriels quelconques est encore un espace vectoriel, pour l'addition et la multiplication définies coordonnées par coordonnées.

On note  $\mathcal{M}_{n,p}(\mathbb{K})$  l'ensemble des matrices à  $n$  lignes et  $p$  colonnes à coefficients dans  $\mathbb{K}$ . C'est donc l'ensemble des *tableaux* à  $n$  lignes et  $p$  colonnes dont chaque case est remplie par un nombre (i.e. un élément de  $\mathbb{K}$ ). On considère deux matrices  $A = (a_{ij})_{\substack{1 \leq i \leq n, \\ 1 \leq j \leq p}}$ , et  $B = (b_{ij})_{\substack{1 \leq i \leq n, \\ 1 \leq j \leq p}}$ . On peut encore écrire

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{np} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & \cdots & b_{1p} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{np} \end{pmatrix}.$$

Alors l'addition de  $A$  et  $B$  est définie coefficients par coefficients :  $A + B$  est la matrice  $C = (c_{ij})_{\substack{1 \leq i \leq n, \\ 1 \leq j \leq p}}$ , où pour tout  $(i, j) \in \llbracket 1, n \rrbracket \times \llbracket 1, p \rrbracket$ ,  $c_{ij} = a_{ij} + b_{ij}$ . Autrement dit,

$$A + B = \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1p} + b_{1p} \\ \vdots & \ddots & \vdots \\ a_{n1} + b_{n1} & \cdots & a_{np} + b_{np} \end{pmatrix}.$$

La multiplication de la matrice  $A$  par un nombre  $\lambda \in \mathbb{K}$  est la matrice  $D = (d_{ij})_{\substack{1 \leq i \leq n, \\ 1 \leq j \leq p}}$ , où  $d_{ij} = \lambda a_{ij}$ . Autrement dit,

$$\lambda A = \begin{pmatrix} \lambda a_{11} & \cdots & \lambda a_{1p} \\ \vdots & \ddots & \vdots \\ \lambda a_{n1} & \cdots & \lambda a_{np} \end{pmatrix}.$$

Muni de ces deux opérations, l'ensemble des matrices  $\mathcal{M}_{n,p}(\mathbb{K})$  est un espace vectoriel.

#### 4.4.2 Sous-espaces vectoriels

Soit  $E$  un espace vectoriel sur  $\mathbb{K}$  muni des opérations  $+$  et  $\cdot$ . Soit  $F$  une partie non vide de  $E$ . Supposons que  $F$  soit stable pour l'addition et la multiplication, autrement dit, pour tout  $x, y \in F$ , pour tout  $\lambda \in \mathbb{K}$ ,

$$x + y \in F, \quad \lambda \cdot x \in F.$$

Alors on peut définir la *restriction* de l'addition et de la multiplication à  $F$ , c'est-à-dire considérer  $+$  comme une application qui à deux éléments de  $F$  associe un élément de  $F$ , et considérer  $\cdot$  comme une application qui à un élément de  $F$  et un nombre de  $\mathbb{K}$  associe un élément de  $F$ . L'ensemble  $F$  est alors muni de deux opérations.

**Proposition 4.60.** *Une partie non vide  $F$  d'un espace vectoriel  $E$  qui est stable pour l'addition et la multiplication par les nombres est un espace vectoriel.*

**Définition 4.61.** *On dit alors que  $F$  est un sous-espace vectoriel de  $E$ .*

On en déduit la méthode la plus utilisée pour montrer qu'un ensemble est un espace vectoriel :

*Méthode 4.62.* Etant donné un ensemble  $F$ , pour montrer que  $F$  est un espace vectoriel sur  $\mathbb{K}$  associé à deux opérations  $+$  et  $\cdot$ ,

1. on identifie un espace vectoriel  $E$  pour ces deux lois qui contient  $F$ ,
2. on montre que  $F$  est non vide,
3. on montre que  $F$  est stable pour l'addition et la multiplication, par exemple en montrant que pour tout  $x, y \in F$ , pour tout  $\lambda \in \mathbb{K}$ ,  $\lambda x + y \in F$ .

### 4.4.3 Dimension

Soit  $E$  un espace vectoriel.

**Définition 4.63.** Soient  $(x_1, \dots, x_n)$  une famille d'éléments de  $E$ .

1. On dit que cette famille est *génératrice* si tout élément  $x \in E$  peut s'écrire comme une combinaison linéaire d'éléments de  $E$ , autrement dit, il existe  $\lambda_1, \dots, \lambda_n \in \mathbb{K}$  tels que

$$x = \lambda_1 x_1 + \dots + \lambda_n x_n.$$

2. On dit que cette famille est *libre* quand pour tout  $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ , si

$$\lambda_1 x_1 + \dots + \lambda_n x_n = 0,$$

alors  $\lambda_1 = \dots = \lambda_n = 0$ .

3. On dit que cette famille est une *base* de  $E$  si elle est à la fois libre et génératrice.

**Définition 4.64.** On dit que  $E$  est de *dimension finie* lorsqu'il existe  $n \in \mathbb{N}^*$  et une famille  $(x_1, \dots, x_n)$  d'éléments de  $E$  qui est une famille génératrice de  $E$ .

Il faut retenir les résultats suivants :

**Théorème 4.65.** 1. Tout espace de dimension finie admet une base.

2. De toute famille génératrice de  $E$ , on peut extraire une base de  $E$ .
3. Toute famille libre de  $E$  peut être complétée en une base.

**Théorème 4.66.** Soit  $E$  un espace vectoriel de dimension finie. Toutes les bases de  $E$  ont même cardinal  $n$ . L'entier  $n$  s'appelle la *dimension* de  $E$  (par convention, si  $E = \{0\}$ , on dit que  $E$  est de dimension 0).

**Théorème 4.67.** Soit  $E$  un espace vectoriel de dimension  $n \in \mathbb{N}^*$ . Alors

1. Toute famille libre a au plus  $n$  éléments et une famille libre qui a exactement  $n$  éléments est une base.
2. Toute famille génératrice a au moins  $n$  éléments et une famille génératrice qui a exactement  $n$  éléments est une base.

**Exemple 4.68.** L'ensemble  $\mathbb{R}[X]$  des polynômes à coefficients dans  $\mathbb{R}$  est un espace vectoriel sur  $\mathbb{R}$  (on vérifie que c'est un sous-espace vectoriel de  $\mathcal{F}(\mathbb{R}; \mathbb{R})$ ). Pour tout  $k \in \mathbb{N}$ , notons  $P_k$  le polynôme  $P_k : x \mapsto x^k$ . Alors

1. pour tout  $n \in \mathbb{N}$ , la famille  $(P_0, P_1, \dots, P_n)$  est libre,
2. l'espace  $\mathbb{R}[X]$  n'est pas de dimension finie,
3. pour tout  $n \in \mathbb{N}$ , l'espace  $\mathbb{R}_n[X]$  des polynômes de degré  $\leq n$  est un espace vectoriel de dimension finie égale à  $n + 1$ .

Montrons que la famille  $(P_0, P_1, \dots, P_n)$  est libre. Soient  $\lambda_0, \dots, \lambda_n \in \mathbb{R}$  et supposons que  $\lambda_0 P_0 + \lambda_1 P_1 + \dots + \lambda_n P_n = 0$ . Alors pour tout  $x \in \mathbb{R}$ ,

$$\lambda_0 P_0(x) + \lambda_1 P_1(x) + \dots + \lambda_n P_n(x) = 0.$$

Ainsi,

$$\lambda_0 + \lambda_1 x + \dots + \lambda_n x^n = 0.$$

Posons  $f(x) := \lambda_0 + \lambda_1 x + \dots + \lambda_n x^n, x \in \mathbb{R}$ . Alors pour tout  $k \in \llbracket 0, n \rrbracket$ ,  $f^{(k)}(0) = k! \lambda_k$ . Comme par ailleurs  $f$  est la fonction nulle, toutes ses dérivées sont nulles, en particulier en 0. Ainsi pour tout  $k \in \llbracket 0, n \rrbracket$ ,  $k! \lambda_k = 0$  et donc  $\lambda_k = 0$ . Conclusion : la famille  $(P_0, P_1, \dots, P_n)$  est libre.

Cela permet d'en déduire que  $\mathbb{R}[X]$  n'est pas de dimension finie. Pour le voir, procédons par l'absurde. Si  $\mathbb{R}[X]$  était de dimension finie, il existerait un entier  $n \in \mathbb{N}$  tel que  $\mathbb{R}[X]$  serait de dimension  $n$ . Alors toute famille libre aurait au plus  $n$  éléments. Or, d'après le premier point, la famille  $(P_0, \dots, P_n)$  qui possède  $n + 1$  éléments, est libre : contradiction. Conclusion :  $\mathbb{R}[X]$  n'est pas de dimension finie.

Pour tout  $n \in \mathbb{N}$ , l'espace  $\mathbb{R}_n[X]$  des polynômes de degré  $\leq n$  est un espace vectoriel. Pour le voir, montrons que c'est un sous-espace vectoriel de l'espace vectoriel  $\mathbb{R}[X]$ . D'abord,  $\mathbb{R}_n[X]$  est non vide, car il contient le polynôme nul. Ensuite, si  $P, Q \in \mathbb{R}_n[X]$  et  $\lambda \in \mathbb{R}$ , alors  $\lambda P$  a le même degré que  $P$  si  $\lambda \neq 0$  et est le polynôme nul si  $\lambda = 0$ . Dans les deux cas,  $\lambda P \in \mathbb{R}_n[X]$ . De plus,  $\lambda P + Q$  est de degré  $\leq \max(\deg \lambda P, \deg Q)$ . Donc  $\lambda P + Q \in \mathbb{R}_n[X]$ . Ainsi,  $\mathbb{R}_n[X]$  est stable par addition et multiplication par les nombres. C'est donc un sous-espace vectoriel de  $\mathbb{R}[X]$  et donc un espace vectoriel.

Montrons que  $\mathbb{R}_n[X]$  est de dimension  $n+1$ . Pour cela, il suffit de vérifier que la famille  $(P_0, \dots, P_n)$ , qui est bien contenue dans  $\mathbb{R}_n[X]$ , est une base de  $\mathbb{R}_n[X]$ . On sait déjà qu'elle est libre. Soit  $P \in \mathbb{R}_n[X]$ . Il existe  $a_0, \dots, a_n \in \mathbb{R}$  tels que

$$P(x) = a_0 + a_1 x + \dots + a_n x^n \quad , \quad \forall x \in \mathbb{R}.$$

On en déduit que  $P = a_0 P_0 + \dots + a_n P_n$ . Autrement dit  $P$  s'écrit comme combinaison linéaire de  $P_0, \dots, P_n$ . Cela montre que  $(P_0, \dots, P_n)$  est une famille génératrice de  $\mathbb{R}_n[X]$ . Finalement,  $(P_0, \dots, P_n)$  est une base de  $\mathbb{R}_n[X]$ , ce qui implique que  $\mathbb{R}_n[X]$  est de dimension  $n + 1$ .



## Chapitre 5

# Continuité des fonctions de plusieurs variables

### 5.1 Découvrir

#### 5.1.1 Limite et continuité pour les fonctions d'une variable

Dans tout ce paragraphe, on considère une fonction  $f$  définie sur une partie non vide  $I$  de  $\mathbb{R}$  et à valeurs dans  $\mathbb{R}$ . On rappelle que l'adhérence de  $I$ , notée  $\text{adh } I$ , est l'ensemble des  $x \in \mathbb{R}$  tels que pour tout  $r > 0$ ,  $]x - r, x + r[ \cap I \neq \emptyset$ . Intuitivement, l'adhérence de  $I$  est l'ensemble des points de  $\mathbb{R}$  qu'on peut approcher aussi près qu'on veut par des points de  $I$ .

**Définition 5.1.** Soit  $a \in \text{adh } I$ .

1. On dit que  $f$  tend vers un nombre  $\ell \in \mathbb{R}$  en  $a$  si pour tout  $\varepsilon > 0$ , il existe  $\eta > 0$  tel que pour tout  $x \in ]a - \eta, a + \eta[ \cap I$ , avec  $x \neq a$ , on a

$$|f(x) - \ell| \leq \varepsilon.$$

2. On dit que  $f$  tend vers  $+\infty$  en  $a$  si pour tout  $M \in \mathbb{R}$ , il existe  $\eta > 0$  tel que pour tout  $x \in ]a - \eta, a + \eta[ \cap I$ , avec  $x \neq a$ , on a

$$f(x) \geq M.$$

On rappelle que si une limite existe, alors elle est unique. De manière analogue au point 2, on peut définir la limite  $-\infty$ .

*Exercice 5.2.* Rappeler les opérations possibles sur les limites et les formes indéterminées.

**Définition 5.3.** On dit que  $f$  est continue en un point  $a \in I$  si  $f$  tend vers  $f(a)$  en  $a$ . On dit que  $f$  est continue sur  $I$  si  $f$  est continue en tout point  $a$  de  $I$ .

*Exercice 5.4.* \* La fonction  $f : I \rightarrow \mathbb{R}$  est continue en  $a \in I$  si et seulement si pour toute suite  $(x_n)_{n \in \mathbb{N}} \subset I$  qui tend vers  $a$ , la suite  $(f(x_n))_{n \in \mathbb{N}}$  tend vers  $f(a)$ .

On rappelle que la somme, le produit ou la composée de deux fonctions continues est continue. Il en est de même de l'inverse d'une fonction continue qui ne s'annule pas.

*Exercice 5.5.* \* Soit  $f : \mathbb{R} \rightarrow \mathbb{R}$  une fonction continue en 0 et qui vérifie la propriété suivante :

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

Montrer alors que  $f$  est constante. (Indication : on cherchera à prouver que  $f(x) = f(0)$  pour tout  $x \in \mathbb{R}$ ).

Dans les deux exercices suivants, on pourra chercher à appliquer le théorème des valeurs intermédiaires.

*Exercice 5.6.* Soit  $f : I \rightarrow \mathbb{R}$  une fonction continue. Montrer que  $f$  est injective si et seulement si  $f$  est strictement monotone.

*Exercice 5.7.* Soit  $f : I \rightarrow \mathbb{Z}$  une fonction continue sur un intervalle et à valeurs entières. Montrer que  $f$  est constante.

On rappelle que si  $f : [c, d] \rightarrow \mathbb{R}$  est une fonction continue, alors elle est bornée et elle atteint ses bornes, ce qui signifie qu'il existe  $x_-$  et  $x_+$  dans  $[c, d]$  tels que

$$f(x_-) = \inf\{f(x) : x \in [c, d]\} \quad , \quad f(x_+) = \sup\{f(x) : x \in [c, d]\}.$$

*Exercice 5.8.* Soit  $f : \mathbb{R} \rightarrow \mathbb{R}$  une fonction continue telle que les deux limites  $\lim_{x \rightarrow \pm\infty} f(x)$  existent dans  $\mathbb{R}$ . Prouver que  $f$  est bornée. Les bornes sont-elles atteintes ?

### 5.1.2 Limite et continuité pour les fonctions de plusieurs variables à valeurs réelles

Dans toute cette section, on se donne une norme  $\|\cdot\|$  sur  $\mathbb{R}^n$ ,  $n \geq 1$ . On s'intéresse à des fonctions définies sur  $\mathbb{R}^n$  (ou plus généralement sur une partie  $A$  de  $\mathbb{R}^n$ ) à valeurs dans  $\mathbb{R}$ .

Soit donc  $A \subset \mathbb{R}^n$  et  $f : A \rightarrow \mathbb{R}$ . Ainsi, à tout vecteur  $x \in A$ , on associe un (unique) nombre noté  $f(x)$ .

Lorsque  $n = 2$ , on peut représenter le graphe d'une telle fonction en associant à tout point  $(x, y) \in \mathbb{R}^2$ , le point de coordonnées  $(x, y, f(x, y))$  de  $\mathbb{R}^3$ .

#### Définition d'une limite

On rappelle qu'un point  $a$  est adhérent à  $A$ , et on note  $a \in \text{adh } A$ , lorsque pour tout  $r > 0$ ,  $B(a, r) \cap A \neq \emptyset$ .

**Définition 5.9.** Soit  $a \in \text{adh } A$ .

1. On dit que  $f$  tend vers  $\ell \in \mathbb{R}$  en  $a$  si pour tout  $\varepsilon > 0$ , il existe  $\eta > 0$  tel que pour tout  $y \in B(a, \eta) \cap A$ , avec  $y \neq a$ ,

$$|f(y) - \ell| \leq \varepsilon.$$

2. On dit que  $f$  tend vers  $+\infty$  en  $a$  si pour tout  $M \in \mathbb{R}$ , il existe  $\eta > 0$  tel que pour tout  $y \in B(a, \eta) \cap A$ , avec  $y \neq a$ ,

$$f(y) \geq M.$$

Dans le premier cas, on note  $\lim_{x \rightarrow a} f(x) = \ell$ . Dans le second cas, on note  $\lim_{x \rightarrow a} f(x) = +\infty$ .

Cette définition prolonge la définition de la limite pour les fonctions définies sur un intervalle de  $\mathbb{R}$  et à valeurs dans  $\mathbb{R}$  (simplement, les intervalles sont remplacés par des boules).

Remarquer que comme  $a$  est adhérent à  $A$ , il existe toujours des vecteurs  $y$  dans l'intersection  $A \cap B(a, \eta)$ , quelle que soit la valeur de  $\eta$ . Ainsi,  $a$  peut être approché d'aussi près qu'on veut par des éléments de  $A$ .

Par ailleurs,  $\ell$  n'est pas forcément égal à  $f(a)$ . En fait, il est possible que  $a$  ne soit pas nécessairement dans  $A$  alors que  $f$  n'est définie que sur  $A$ .

Toutes les propriétés vues sur les limites pour les fonctions de  $\mathbb{R}$  dans  $\mathbb{R}$  restent vraies ici, notamment l'unicité de la limite, et les limites d'une somme, d'un produit par un scalaire, d'une composée de fonctions.

*Exercice 5.10.* Soit  $N$  une norme sur  $\mathbb{R}^2$ . Quelle est la limite de la fonction  $N$  en 0 ?

### Définition de la continuité

La définition suivante généralise la définition de continuité pour les fonctions définies sur un intervalle de  $\mathbb{R}$  et à valeurs dans  $\mathbb{R}$ .

**Définition 5.11.** On dit qu'une fonction  $f : A \rightarrow \mathbb{R}$  est continue en un point  $a \in A$  si  $\lim_{x \rightarrow a} f(x) = f(a)$ .

Avec des quantificateurs, cela s'écrit : pour tout  $\varepsilon > 0$ , il existe  $\eta > 0$  tel que pour tout  $x \in B(a, \eta) \cap A$ ,

$$|f(x) - f(a)| \leq \varepsilon.$$

Intuitivement, cela signifie que l'on peut rendre  $f(x)$  aussi proche que l'on veut de  $f(a)$  pourvu que l'on prenne  $x$  suffisamment proche de  $a$ . Noter qu'ici (contrairement à la définition des limites), le point  $a$  est dans  $A$ .

L'inégalité  $|f(x) - f(a)| \leq \varepsilon$  se réécrit  $f(a) - \varepsilon \leq f(x) \leq f(a) + \varepsilon$ , autrement dit

$$f(x) \in [f(a) - \varepsilon, f(a) + \varepsilon].$$

Ainsi, dire que pour tout  $x \in B(a, \eta) \cap A$ ,  $|f(x) - f(a)| \leq \varepsilon$  revient à dire que

$$f\left(B(a, \eta) \cap A\right) \subset [f(a) - \varepsilon, f(a) + \varepsilon].$$

**Proposition 5.12.** La continuité d'une fonction en un point est invariante par changement de norme.

Cela signifie que si on prend une autre norme  $N_1$  sur  $\mathbb{R}^n$ , équivalente à  $N$ , alors toute fonction continue lorsqu'on prend la norme  $N$  sera continue lorsqu'on prend la norme  $N_1$ .

Preuve : Soit une fonction  $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$  continue en  $a \in A$  pour la norme  $N$  sur  $\mathbb{R}^n$ . Soit  $\varepsilon > 0$ . Il existe  $\eta > 0$  tel que pour tout  $y \in B_N(a, \eta) \cap A$ ,

$$|f(y) - f(a)| \leq \varepsilon.$$

Comme  $N$  et  $N_1$  sont équivalentes, il existe  $C > 0$  tel que pour tout  $y \in \mathbb{R}^n$ ,  $N(y) \leq CN_1(y)$ . Posons  $\eta' = \eta/C$ . Alors pour tout  $y \in B_{N_1}(a, \eta') \cap A$ , on a  $N_1(y - a) < \eta/C$ , donc  $N(y - a) < \eta$ , ce qui implique  $|f(y) - f(a)| \leq \varepsilon$ . Cela montre la continuité de  $f$  pour la norme  $N_1$ . □

**Définition 5.13** (Continuité sur un ensemble). On dit que  $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$  est continue sur  $A$  quand  $f$  est continue en tout point  $a \in A$ .

*Exemple 5.14.* La fonction  $f : (x_1, \dots, x_n) \in \mathbb{R}^n \mapsto x_1$  est continue sur  $\mathbb{R}^n$ .

Preuve : Soit  $a = (a_1, \dots, a_n) \in \mathbb{R}^n$ . Montrons que  $f$  est continue en  $a$ . Prenons sur  $\mathbb{R}^n$  la norme  $\|\cdot\|_\infty$ .

Soit  $\varepsilon > 0$ . Posons  $\eta = \varepsilon$ . Alors pour tout  $x = (x_1, \dots, x_n) \in B(a, \eta)$ , on a

$$|f(x) - f(a)| = |x_1 - a_1| \leq \|x - a\|_\infty \leq \eta = \varepsilon.$$

On en déduit que  $f$  est bien continue en  $a$ .

## 5.2 Assimiler

### 5.2.1 Continuité et critère séquentiel

La proposition qui suit donne un critère séquentiel de continuité.

**Proposition 5.15.** Soit  $f : A \rightarrow \mathbb{R}$  et  $a \in A$ . Alors  $f$  est continue en  $a$  si et seulement si pour toute suite  $(x_n)_{n \in \mathbb{N}}$  dans  $A$  qui tend vers  $a$ , la suite  $(f(x_n))_{n \in \mathbb{N}}$  tend vers  $f(a)$ .

Preuve : Supposons que  $f$  soit continue en  $a$ . Soit  $(x_n)_{n \in \mathbb{N}}$  une suite dans  $A$  qui tend vers  $a$ . Alors pour tout  $\varepsilon > 0$ , il existe  $\eta > 0$  tel que pour tout  $x \in A \cap B(a, \eta)$ ,

$$|f(x) - f(a)| \leq \varepsilon.$$

De plus, il existe  $n_0 \in \mathbb{N}$  tel que pour tout  $n \geq n_0$ ,  $x_n \in A \cap B(a, \eta)$  et donc

$$|f(x_n) - f(a)| \leq \varepsilon.$$

Cela montre que la suite  $(f(x_n))_{n \in \mathbb{N}}$  tend vers  $f(a)$ .

Réciproquement, supposons que pour toute suite  $(x_n)_{n \in \mathbb{N}}$  dans  $A$  qui tend vers  $a$ , la suite  $(f(x_n))_{n \in \mathbb{N}}$  tend vers  $f(a)$ . Supposons par l'absurde que  $f$  ne tende pas vers  $f(a)$  en  $a$ . Alors il existe  $\varepsilon > 0$  tel que pour tout  $\eta > 0$ , il existe  $x \in A \cap B(a, \eta)$  tel que  $|f(x) - f(a)| > \varepsilon$ . On applique ceci pour chaque  $\eta = \frac{1}{n+1}$ , lorsque  $n$  parcourt  $\mathbb{N}$ . On en déduit une suite  $(x_n)_{n \in \mathbb{N}}$  telle que pour tout  $n \in \mathbb{N}$ ,

$$x_n \in A \cap B\left(a, \frac{1}{n+1}\right) \quad \text{et} \quad |f(x_n) - f(a)| > \varepsilon.$$

Donc  $(x_n)_{n \in \mathbb{N}}$  tend vers  $a$  et pourtant  $(f(x_n))_{n \in \mathbb{N}}$  ne tend pas vers  $f(a)$  : contradiction. Donc  $f$  est continue en  $a$ . □

On utilise souvent la proposition précédente pour montrer qu'une fonction n'est pas continue en un point. Voici un exemple :

*Exemple 5.16.* Soit

$$f : (x, y) \in \mathbb{R}^2 \mapsto \begin{cases} \frac{x^2 + y^2}{3x^2 + 4y^2} & \text{si } (x, y) \neq (0, 0), \\ (0, 0) & \text{si } (x, y) = (0, 0). \end{cases}$$

Alors  $f$  n'est pas continue en  $(0, 0)$ .

En effet, on observe que pour tout  $n \in \mathbb{N}^*$ ,

$$f\left(\frac{1}{n}, 0\right) = \frac{1}{3}.$$

De plus, la suite  $\left(\left(\frac{1}{n}, 0\right)\right)_{n \in \mathbb{N}^*}$  converge vers  $(0, 0)$ . Comme la suite  $\left(f\left(\frac{1}{n}, 0\right)\right)_{n \in \mathbb{N}^*}$  ne converge pas vers  $f(0, 0)$ , on déduit de la proposition précédente que  $f$  n'est pas continue en  $(0, 0)$ .

### 5.2.2 Opérations sur les fonctions continues

Comme pour les fonctions d'une seule variable, la continuité est préservée par les opérations usuelles :

**Proposition 5.17.** Soient  $f, g : A \rightarrow \mathbb{R}$  deux fonctions continues en  $a \in A$  et  $\lambda \in \mathbb{R}$ . Alors les fonctions  $f + g$ ,  $fg$  et  $\lambda f$  sont continues en  $a$ .

*Démonstration.* Soit  $\varepsilon > 0$ . Comme  $f$  est continue en  $a$ , il existe  $\eta > 0$  tel que pour tout  $x \in A \cap B(a, \eta)$ ,  $|f(x) - f(a)| \leq \varepsilon$ . De même, comme  $g$  est continue en  $a$ , il existe  $\eta' > 0$  tel que pour tout  $x \in A \cap B(a, \eta')$ ,  $|g(x) - g(a)| \leq \varepsilon$ . On en déduit en posant  $\eta'' = \min(\eta, \eta')$  que pour tout  $x \in A \cap B(a, \eta'')$ ,

$$\begin{aligned} |(f+g)(x) - (f+g)(a)| &= |(f(x) - f(a)) + (g(x) - g(a))| \\ &\leq |f(x) - f(a)| + |g(x) - g(a)| \\ &\leq \varepsilon + \varepsilon = 2\varepsilon. \end{aligned}$$

Comme  $\varepsilon$  est arbitraire, cela montre la continuité de  $f+g$  en  $a$ . Par ailleurs,

$$\begin{aligned} |(fg)(x) - (fg)(a)| &= |f(x)g(x) - f(a)g(a)| = |f(x)(g(x) - g(a)) + g(a)(f(x) - f(a))| \\ &\leq |f(x)(g(x) - g(a))| + |g(a)(f(x) - f(a))| \\ &= |f(x)||g(x) - g(a)| + |g(a)||f(x) - f(a)| \\ &\leq |f(x)|\varepsilon + |g(a)|\varepsilon \\ &\leq (|f(x) - f(a)| + |f(a)|)\varepsilon + |g(a)|\varepsilon \\ &\leq \varepsilon^2 + (|f(a) + g(a)|)\varepsilon. \end{aligned}$$

Comme  $\varepsilon$  est arbitraire, cela montre la continuité de  $fg$  en  $a$ . On laisse la continuité de  $\lambda f$  en exercice.  $\square$

*Exemple 5.18.* Une forme linéaire sur  $\mathbb{R}^n$  est continue.

*Démonstration.* En effet, si  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  est une forme linéaire, alors il existe  $a_1, \dots, a_n \in \mathbb{R}$  tels que

$$\forall x = (x_1, \dots, x_n) \in \mathbb{R}^n, \quad f(x) = a_1x_1 + \dots + a_nx_n.$$

Par l'Exemple 5.14, on sait que pour chaque  $i \in \{1, \dots, n\}$ , la fonction  $x \mapsto x_i$  est continue. Donc par produit avec une constante,  $x \mapsto a_ix_i$  est continue. Par somme de fonctions continues, on en déduit enfin que  $f$  est continue.  $\square$

*Exemple 5.19.*

$$g(x, y) = \begin{cases} \frac{x^2e^x + y^2}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0), \\ 1 & \text{si } (x, y) = (0, 0). \end{cases}$$

Sur  $\mathbb{R}^2 \setminus \{(0, 0)\}$ , la fonction  $g$  est continue comme quotient de fonctions continues, le dénominateur ne s'annulant pas. Montrons la continuité de  $g$  en  $(0, 0)$ . Pour tout  $(x, y) \neq (0, 0)$ ,

$$|g(x, y) - 1| = \frac{|e^x - 1|x^2|}{x^2 + y^2} \leq |e^x - 1|.$$

Soit  $\varepsilon > 0$ . Par continuité de  $x \mapsto e^x$  en 0, il existe  $\eta > 0$  tel que pour tout  $|x| < \eta$ , on a  $|e^x - 1| < \varepsilon$ . Donc pour tout  $\|(x, y)\|_\infty < \eta$ ,  $|g(x, y) - 1| < \varepsilon$ . Ainsi,  $g$  est bien continue en  $(0, 0)$ .

*Exemple 5.20.* Une fonction de la forme  $x = (x_1, \dots, x_n) \in \mathbb{R}^n \mapsto \lambda x_1^{\alpha_1} \dots x_n^{\alpha_n}$ , où  $\lambda \in \mathbb{R}$  et  $\alpha_1, \dots, \alpha_n \in \mathbb{N}$  est un monôme. Un polynôme est une somme de monômes. Tous les polynômes sont continus.

En effet, chaque fonction  $x \mapsto x_i$  est continue. Donc par produit de fonctions continues,  $x \mapsto x_i^{\alpha_i}$  est continue. Donc chaque monôme est continu comme produit de fonctions continues. Donc chaque polynôme est continu comme somme de fonctions continues.

Exercice 5.21. Justifier que l'application suivante

$$f : (x, y) \in \mathbb{R}^2 \mapsto x^2 + 4y^5 \in \mathbb{R}$$

est continue sur  $\mathbb{R}^2$ .

Exercice 5.22. Soit  $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction continue. Montrer que

1. tout point de l'ensemble  $E_f = \{x \in A : f(x) > 0\}$  est dans l'intérieur de  $E_f$ ,
2. tout point dans l'adhérence de l'ensemble  $F_f = \{x \in A : f(x) \geq 0\}$  est dans  $F_f$ .

### 5.2.3 Continuité des fonctions à valeurs dans $\mathbb{R}^p$

Dans ce paragraphe, on considère des fonctions définies sur une partie  $A \subset \mathbb{R}^n$  et à valeurs dans un espace  $\mathbb{R}^p$  :

$$f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^p.$$

Pour tout  $x \in A$ ,  $f(x)$  est donc un vecteur dans  $\mathbb{R}^p$ , dont on note  $(f_1(x), \dots, f_p(x))$  les coordonnées. On associe ainsi pour tout  $1 \leq i \leq p$ , à chaque  $x \in A$  la  $i$ ème composante de  $f(x)$ , notée  $f_i(x)$ .

Par exemple, à tout point de l'espace, on associe la température et la pression en ce point. On obtient ainsi une fonction définie sur  $\mathbb{R}^3$  (ou une partie de  $\mathbb{R}^3$ , par exemple la salle de cours, où tous les points seraient repérés par leurs coordonnées dans un grand système de coordonnées), et à valeurs dans  $\mathbb{R}^2$ . La première composante de la fonction donne la température en ce point, la seconde donne la pression en ce même point<sup>1</sup>.

**Définition 5.23.** On dit que  $f$  a une limite  $\ell = (\ell_1, \dots, \ell_p)$  en un point  $a \in \text{adh } A$  si chaque composante  $f_i$  de  $f$  a pour limite  $\ell_i$  en  $a$ .

Autrement dit,

$$\lim_{x \rightarrow A} f(x) = \ell \iff \forall i \in \{1, \dots, p\}, \quad \lim_{x \rightarrow A} f_i(x) = \ell_i.$$

**Définition 5.24.** Soit  $a \in A$ . On dit que  $f$  est continue en  $a$  si  $\lim_{x \rightarrow a} f(x) = f(a)$ .

Autrement dit,  $f$  est continue en  $a$  si pour tout  $1 \leq i \leq p$ ,  $f_i$  est continue en  $a$ .

**Proposition 5.25.** Soit  $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$  et  $g : B \subset \mathbb{R}^m \rightarrow \mathbb{R}^\ell$  deux fonctions continues. On suppose que  $f(A)$  est contenu dans  $B$  de sorte que la composée  $g \circ f$  est bien définie. Alors  $g \circ f$  est continue sur  $A$ .

On pourra trouver la preuve de cette proposition dans la section *Approfondir*.

*Remarque 5.26.* En particulier, si  $f$  est continue et à valeurs dans  $\mathbb{R}^*$ , alors la proposition précédente appliquée à la fonction continue  $g : x \in \mathbb{R}^* \mapsto \frac{1}{x}$  implique que  $\frac{1}{f}$  est encore continue.

*Exemple 5.27.* Toutes les applications linéaires de  $\mathbb{R}^n$  dans  $\mathbb{R}^p$  sont continues.

Preuve : Il suffit de montrer que chaque composante de  $\Phi$  est continue. Or, les composantes de  $\Phi$  sont des formes linéaires. Le résultat est donc une conséquence de l'Exemple 5.18. □

**Proposition 5.28.** Soit  $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$  une fonction continue en  $a = (a_1, \dots, a_n)$ . Alors toutes les restrictions de  $f$  par rapport à 1, 2, ..., ou  $n - 1$  variables sont continues.

1. Sur la surface de la Terre, on peut montrer *mathématiquement* qu'il existe deux points antipodaux où à la fois la pression et la température sont égales.

*Démonstration.* On montre que la restriction par rapport à la première variable  $x_1 \in \mathbb{R} \mapsto f(x_1, a_2, \dots, a_n)$  est continue (les autres assertions se démontrent de manière analogue). Il s'agit de la composée de  $x_1 \mapsto (x_1, a_2, \dots, a_n)$  qui est continue en  $a_1$  et de  $f$  qui est continue en  $a$ . On conclut grâce à la proposition sur la continuité des composées.  $\square$

*Remarque 5.29.* Attention : la réciproque est fautive. Par exemple pour  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , ce n'est pas parce que  $x \mapsto f(x, a_2)$  et  $y \mapsto f(a_1, y)$  sont continues que  $f$  est continue en  $(a_1, a_2)$ .

*Exemple 5.30.*

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

Les fonctions  $f(0, \cdot)$  et  $f(\cdot, 0)$  sont constamment égales à 0 et en particulier continues. Pourtant, pour tout  $y \neq 0$ ,  $f(y, y) = \frac{1}{2}$ . En particulier,  $|f(0, 0) - f(y, y)| \geq \frac{1}{2}$  et donc  $f$  n'est pas continu en  $(0, 0)$ .

## 5.3 Approfondir

### 5.3.1 Retour sur les limites et la continuité des fonctions à valeurs dans $\mathbb{R}^p$

On se donne une norme  $N$  sur  $\mathbb{R}^n$  et une norme  $N'$  sur  $\mathbb{R}^p$ , ce qui permet de considérer des boules  $B_N(a, r)$  dans  $\mathbb{R}^n$  et des boules  $B_{N'}(b, s)$  dans  $\mathbb{R}^p$ . Le plus souvent, on se contentera de noter  $B(a, r)$  au lieu de  $B_N(a, r)$ , et  $B(b, s)$  au lieu de  $B_{N'}(b, s)$ .

Dans la proposition suivante, on présente une autre manière de formuler l'existence d'une limite pour les fonctions à valeurs dans  $\mathbb{R}^p$ , à comparer avec la Définition 5.23.

**Proposition 5.31.** Soient  $f : A \rightarrow \mathbb{R}^p$  et  $a \in \text{adh } A$ . Alors  $f$  admet une limite  $\ell \in \mathbb{R}^p$  en  $a$  si et seulement si pour tout  $\varepsilon > 0$ , il existe  $\eta > 0$  tel que pour tout  $x \in A \cap B_N(a, \eta)$  avec  $a \neq x$ ,

$$f(x) \in B_{N'}(\ell, \varepsilon).$$

*Démonstration.* Supposons que  $f$  ait une limite  $\ell = (\ell_1, \dots, \ell_p) \in \mathbb{R}^p$  en  $a$ . Par définition, cela signifie que chaque composante  $f_i$  a une limite  $\ell_i$  en  $a$ .

Soit alors  $\varepsilon > 0$ . Sur  $\mathbb{R}^p$ , la norme  $N'$  est équivalente à la norme  $\|\cdot\|_1$ , donc il existe  $C > 0$  tel que  $N' \leq C\|\cdot\|_1$ .

Pour chaque  $i \in \{1, \dots, p\}$ , il existe  $\eta_i > 0$  tel que pour tout  $x \in B_N(a, \eta_i) \cap A$  avec  $x \neq a$ , on a  $|f_i(x) - \ell_i| < \varepsilon/(Cp)$ . Posons  $\eta = \min(\eta_1, \dots, \eta_p)$ . Alors pour tout  $x \in B_N(a, \eta) \cap A$  avec  $x \neq a$ , on a

$$\|f(x) - \ell\|_1 \leq \sum_{i=1}^p |f_i(x) - \ell_i| < \varepsilon/C.$$

Donc

$$N'(f(x) - \ell) \leq C\|f(x) - \ell\|_1 < \varepsilon.$$

Cela montre que  $f(x) \in B_{N'}(\ell, \varepsilon)$ . D'où l'implication directe.

Pour montrer l'implication réciproque, établissons par exemple que  $\lim_{x \rightarrow a} f_1(x) = \ell_1$ . Soit  $\varepsilon > 0$ . Par équivalence des normes sur  $\mathbb{R}^p$ , il existe  $C' > 0$  tel que  $\|\cdot\|_\infty \leq C'N'$ .

Alors par hypothèse, il existe  $\eta > 0$  tel que pour tout  $x \in A \cap B_N(a, \eta)$ ,  $a \neq x$ ,  $f(x) \in B_{N'}(\ell, \varepsilon/C')$ , autrement dit,  $N'(f(x) - \ell) < \varepsilon/C'$ . Donc

$$\|f(x) - \ell\|_\infty \leq C'N'(f(x) - \ell) < \varepsilon.$$

En particulier,

$$|f_1(x) - \ell_1| \leq \|f(x) - \ell\|_\infty < \varepsilon.$$

Cela montre que  $\lim_{x \rightarrow a} f_1(x) = \ell_1$ , comme attendu.  $\square$

A partir de la proposition précédente, on peut donc aussi reformuler la continuité d'une application  $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$ . Elle sera continue en  $a$  si et seulement si pour tout  $\varepsilon > 0$ , il existe  $\eta > 0$  tel que

$$\forall x \in B_N(a, \eta) \cap A, \quad f(x) \in B_{N'}(f(a), \varepsilon).$$

On en déduit une preuve élémentaire de la Proposition 5.25 :

Preuve : Pour simplifier les notations, on ne rappelle pas les normes en indices lorsqu'on écrit une boule. Soit  $a \in A$  et  $\varepsilon > 0$ . Comme  $g$  est continue en  $f(a)$ , il existe  $\eta > 0$  tel que

$$\forall y \in B(f(a), \eta) \cap B, \quad g(y) \in B(g(f(a)), \varepsilon).$$

Comme  $f$  est continue en  $a$ , il existe  $\eta' > 0$  tel que

$$\forall x \in B(a, \eta') \cap A, \quad f(x) \in B(f(a), \eta).$$

Comme  $f(A) \subset B$ , on en déduit que pour tout  $x \in B(a, \eta')$ ,  $g(f(x)) \in B(g(f(a)), \varepsilon)$ , ce qui montre la continuité de  $g \circ f$  en  $a$ . Comme  $a$  est arbitraire, cela implique la continuité de  $g \circ f$  sur  $A$ . □

### 5.3.2 Fonctions continues sur un fermé borné

En utilisant que de toute suite contenue dans une partie bornée et fermée de  $\mathbb{R}^n$ , on peut extraire une sous-suite convergente, on déduit le théorème suivant :

**Théorème 5.32.** *Toute fonction continue sur une partie fermée et bornée est bornée et atteint ses bornes.*

Cela signifie que si  $A \subset \mathbb{R}^n$  est une partie fermée et bornée et si  $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$  est continue sur  $A$ , alors il existe  $M > 0$  tel que pour tout  $x \in A$ , on a  $|f(x)| \leq M$ . De plus, il existe  $x_- \in A$  et  $x_+ \in A$  tels que

$$\forall x \in A, \quad f(x_-) \leq f(x) \leq f(x_+).$$

Autrement dit,  $x_-$  est un point de minimum de  $f$  sur  $A$ , c'est-à-dire,

$$f(x_-) = \min\{f(x) : x \in A\}.$$

On dit aussi que  $f$  atteint son minimum en  $x_-$ . La valeur minimale de  $f$  sur  $A$  est  $f(x_-)$ .

De même,  $x_+$  est un point de maximum de  $f$  sur  $A$ .

## Chapitre 6

# Différentiabilité des fonctions de plusieurs variables

### 6.1 Découvrir

Soient  $n, p \geq 1$ . L'objet de ce chapitre est de définir (une généralisation de) la dérivabilité pour les fonctions définies sur une partie de  $\mathbb{R}^n$  et prenant leurs valeurs dans  $\mathbb{R}^p$ . Les applications linéaires constituent l'exemple le plus important de telles fonctions mais toutes les fonctions de  $\mathbb{R}^n$  dans  $\mathbb{R}^p$  ne sont pas linéaires.

Pour définir rigoureusement cette notion de dérivabilité, nous ferons intervenir des normes sur  $\mathbb{R}^n$  et  $\mathbb{R}^p$ . Comme il s'agit d'espaces de dimension finie, toutes les normes sont équivalentes. On rappelle que deux normes  $N_1$  et  $N_2$  sont équivalentes s'il existe deux constantes  $C_1 > 0$  et  $C_2 > 0$  telles que pour tout vecteur  $x$ ,  $N_1(x) \leq C_1 N_2(x)$  et  $N_2(x) \leq C_2 N_1(x)$ .

Lorsque deux normes sont équivalentes, si l'une vérifie une inégalité, alors l'autre vérifie la même inégalité à une constante multiplicative près. Aussi, beaucoup des notions qui vont suivre, faisant intervenir des inégalités, ne dépendront pas de la norme choisie. C'est la raison pour laquelle, en général, nous ne préciserons pas la norme que nous utilisons, et nous la noterons  $\|\cdot\|$ . Les boules  $B(0, r)$ ,  $B(0, \eta)$ , ... sont définies à partir de cette norme.

On commence par une section de révision sur la dérivation des fonctions d'une variable.

#### 6.1.1 Fonctions dérivables de $\mathbb{R}$ dans $\mathbb{R}$

Soit  $I$  un intervalle ouvert, non vide de  $\mathbb{R}$  et  $f : I \rightarrow \mathbb{R}$ .

**Définition 6.1.** Soit  $a \in I$ . On dit que  $f$  est dérivable en  $a$  si le taux d'accroissement

$$x \mapsto \frac{f(x) - f(a)}{x - a}$$

a une limite finie en  $a$ .

Noter que le taux d'accroissement est bien défini sur  $I \setminus \{a\}$ . La limite s'appelle le nombre dérivé de  $f$  en  $a$ . On dit que  $f$  est dérivable sur  $I$  si  $f$  est dérivable en tout point de  $I$ . On dispose alors de la fonction dérivée

$$f' : x \in I \rightarrow f'(x) \in \mathbb{R}.$$

Rappelons qu'on a l'habitude de noter  $o(1)$ ,  $x \rightarrow a$  n'importe quelle fonction, définie sur un ensemble contenant  $]a - \varepsilon, a[ \cup ]a, a + \varepsilon[$ , avec  $\varepsilon > 0$ , et qui tend vers 0 en  $a$ . Ainsi, si  $f$  est dérivable en  $a$ , alors la fonction

$$x \mapsto \frac{f(x) - f(a)}{x - a} - f'(a)$$

tend vers 0 en  $a$  et on peut donc noter

$$\frac{f(x) - f(a)}{x - a} - f'(a) = o(1), \quad x \rightarrow a.$$

En multipliant à gauche et à droite par  $x - a$ , on obtient

$$f(x) - f(a) - f'(a)(x - a) = (x - a)o(1), \quad x \rightarrow a.$$

On note  $(x - a)o(1)$  sous la forme  $o(x - a)$ . C'est donc une notation qui désigne une fonction de la forme  $(x - a)h(x)$ , où  $h$  est une fonction qui tend vers 0 en  $a$ . Ainsi, on a

$$f(x) = f(a) + f'(a)(x - a) + o(x - a), \quad x \rightarrow a.$$

*Exercice 6.2.* Soient  $I$  un intervalle ouvert de  $\mathbb{R}$ ,  $f : I \rightarrow \mathbb{R}$  et  $a \in I$ . On suppose qu'il existe  $\ell \in \mathbb{R}$  tel que pour tout  $x \in I$ ,

$$f(x) = f(a) + \ell(x - a) + o(x - a).$$

Montrer que  $f$  est dérivable en  $a$  et  $f'(a) = \ell$ .

*Exercice 6.3.* Rappeler les opérations usuelles sur les fonctions dérivables.

Parmi les théorèmes importants, citons l'inégalité des accroissements finis :

**Théorème 6.4.** Soit  $f : [a, b] \rightarrow \mathbb{R}$  une fonction continue sur  $[a, b]$ , dérivable sur  $]a, b[$ . Alors

$$\frac{|f(b) - f(a)|}{|b - a|} \leq \sup_{x \in ]a, b[} |f'(x)|.$$

## 6.1.2 Les dérivées partielles

Dans cette section, on considère une fonction  $g : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$  sur une partie  $D$  de  $\mathbb{R}^n$ . On fixe aussi un point  $a = (a_1, \dots, a_n)$  dans l'intérieur de  $D$ . On rappelle que cela signifie qu'il existe  $\varepsilon > 0$  tel que  $B(a, \varepsilon) \subset D$ . Dans toute cette section, par commodité, on travaillera avec la norme  $\|\cdot\|_\infty$  pour définir les boules :

$$\begin{aligned} B(a, \varepsilon) &= ]a_1 - \varepsilon, a_1 + \varepsilon[ \times \dots \times ]a_n - \varepsilon, a_n + \varepsilon[ \\ &= \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_i \in ]a_i - \varepsilon, a_i + \varepsilon[, \forall i = 1, \dots, n\}. \end{aligned}$$

Pour tout  $1 \leq j \leq n$ , on va considérer la fonction d'une variable

$$x_j \mapsto g(a_1, \dots, a_{j-1}, x_j, a_{j+1}, \dots, a_n).$$

Cette fonction est bien sûr définie sur

$$D_j = \{t \in \mathbb{R} : (a_1, \dots, a_{j-1}, t, a_{j+1}, \dots, a_n) \in D\}.$$

Observez que si  $|t - a_j| < \varepsilon$ , alors

$$\|(a_1, \dots, a_{j-1}, t, a_{j+1}, \dots, a_n) - (a_1, \dots, a_n)\|_\infty = |t - a_j| < \varepsilon$$

et donc  $(a_1, \dots, a_{j-1}, t, a_{j+1}, \dots, a_n) \in D$ . Ainsi, la fonction

$$t \mapsto g(a_1, \dots, a_{j-1}, t, a_{j+1}, \dots, a_n)$$

est définie sur  $]a_j - \varepsilon, a_j + \varepsilon[$ .

Une généralisation possible de la dérivation pour les fonctions de plusieurs variables consiste à *geler* toutes les variables sauf une, et à dériver la fonction d'une variable ainsi obtenue. On est alors conduit à définir la notion de *dérivée partielle*.

**Définition 6.5.** On dit que  $g$  admet une dérivée partielle par rapport à la  $j$  ème variable en  $a$  quand la fonction

$$t \mapsto g(a_1, \dots, a_{j-1}, t, a_{j+1}, \dots, a_n)$$

est dérivable en  $a_j$ . On note

$$\begin{aligned} \partial_j g(a) &= \lim_{\substack{t \rightarrow a_j \\ t \neq a_j}} \frac{g(a_1, \dots, a_{j-1}, t, a_{j+1}, \dots, a_n) - g(a_1, \dots, a_n)}{t - a_j} \\ &= \lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{g(a_1, \dots, a_{j-1}, a_j + h, a_{j+1}, \dots, a_n) - g(a_1, \dots, a_n)}{h}. \end{aligned}$$

Même si pour calculer  $\partial_j g(a)$ , on gèle les variables  $x_1 = a_1, \dots, x_{j-1} = a_{j-1}, x_{j+1} = a_{j+1}, \dots, x_n = a_n$ , et on dérive par rapport à la  $j$  ème, le résultat  $\partial_j g(a)$  dépend de toutes les coordonnées de  $a$  et pas seulement de  $a_j$ .

*Exercice 6.6.* Calculer les dérivées partielles de la fonction

$$(x, y) \mapsto x^2 e^y.$$

### 6.1.3 Gradient et différentielle d'une fonction

**Définition 6.7.** Soit  $g$  une fonction sur un ensemble  $D$  et  $a$  un point intérieur à  $D$ . On suppose que toutes les dérivées partielles de  $g$  existent en  $a$ .

1. On appelle gradient de  $g$  en  $a$  le vecteur de  $\mathbb{R}^n$

$$\nabla g(a) = \begin{pmatrix} \partial_1 g(a) \\ \vdots \\ \partial_n g(a) \end{pmatrix}.$$

2. On appelle différentielle de  $g$  en  $a$  la forme linéaire de  $\mathbb{R}^n$  dans  $\mathbb{R}$  définie par

$$h = (h_1, \dots, h_n) \in \mathbb{R}^n \mapsto \langle \nabla g(a), h \rangle = h_1 \partial_1 g(a) + \dots + h_n \partial_n g(a).$$

La différentielle de  $g$  en  $a$  est notée  $Dg(a)$ . On a donc  $Dg(a) : h \mapsto \langle \nabla g(a), h \rangle$ . On utilisera la notation  $Dg(a)[h]$  plutôt que  $Dg(a)(h)$ .

Si  $(e_1, \dots, e_n)$  désigne la base canonique de  $\mathbb{R}^n$ ,

$$Dg(a)[e_i] = \langle \nabla g(a), e_i \rangle = \partial_i g(a).$$

La matrice de l'application linéaire  $Dg(a)$  dans les bases canoniques est

$$\begin{pmatrix} \partial_1 g(a) & \partial_2 g(a) & \cdots & \partial_n g(a) \end{pmatrix}.$$

Il s'agit de la transposée du vecteur gradient.

### Interprétation géométrique du gradient

Pour une fonction dérivable  $f : \mathbb{R} \rightarrow \mathbb{R}$ , le nombre dérivé de  $f$  en  $a$  est la pente de la tangente au graphe de  $f$  au point  $(a, f(a))$ . Un vecteur directeur de cette tangente est  $(1, f'(a))$ . Un vecteur normal est donc  $(f'(a), -1)$ .

Considérons une fonction  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ . Le graphe de  $f$  est donc l'ensemble des points de  $\mathbb{R}^3$  de la forme  $(x_1, x_2, f(x_1, x_2))$  lorsque  $(x_1, x_2) \in \mathbb{R}^2$ . En un tel point  $(a_1, a_2, f(a_1, a_2))$ , on peut définir le

plan tangent au graphe. Pour cela, on regarde toutes les courbes dessinées sur le graphe et qui passent par le point  $(a_1, a_2, f(a_1, a_2))$ . Par exemple, l'image de la fonction  $t \mapsto (t, a_2, f(t, a_2))$  est une telle courbe (cette courbe passe par  $(a_1, a_2, f(a_1, a_2))$  à l'instant  $t = a_1$ ). Un vecteur tangent à cette courbe est donné par la dérivée de cette fonction :  $(1, 0, \partial_1 f(a_1, a_2))$ . Une autre courbe dessinée sur le graphe de  $f$  est l'image de la fonction  $t \mapsto (a_1, t, f(a_1, t))$ . Un vecteur tangent à cette courbe en  $(a_1, a_2, f(a_1, a_2))$  est le vecteur  $(0, 1, \partial_2 f(a_1, a_2))$ . On obtient le plan tangent au graphe de  $f$  en prenant l'union des droites tangent à toutes les courbes dessinées sur le graphe passant par le point  $(a_1, a_2, f(a_1, a_2))$ .

Voici une manière plus précise de définir le plan tangent au graphe en  $(a_1, a_2, f(a_1, a_2))$  : c'est le plan de  $\mathbb{R}^3$  qui passe par le point  $(a_1, a_2, f(a_1, a_2))$  et qui admet pour vecteurs directeurs  $V_1 = (1, 0, \partial_1 f(a))$  et  $V_2 = (0, 1, \partial_2 f(a))$ , où  $a = (a_1, a_2) \in \mathbb{R}^2$ . Autrement dit, ce plan tangent est

$$\{y = (y_1, y_2, y_3) \in \mathbb{R}^3 : \exists \alpha, \beta \in \mathbb{R} \text{ tels que } y - (a, f(a)) = \alpha V_1 + \beta V_2\}.$$

Un vecteur normal à ce plan est orthogonal à tous les vecteurs directeurs du plan, en particulier à  $V_1$  et  $V_2$ . Un exemple de vecteur normal est donc

$$(\partial_1 f(a), \partial_2 f(a), -1) = (\nabla f(a), -1).$$

Le gradient de  $f$  en  $a$  permet donc de construire la normale au plan tangent du graphe de  $f$  en  $(a, f(a))$ .

## 6.2 Assimiler

### 6.2.1 Fonctions de classe $C^1$

On suppose désormais que l'ensemble  $D$  est un ouvert de  $\mathbb{R}^n$ , autrement dit que tous les points de  $D$  sont dans l'intérieur de  $D$ .

**Définition 6.8.** On dit que  $g$  est de classe  $C^1$  sur  $D$  quand

1.  $\forall a \in D, \forall 1 \leq j \leq n, g$  admet une dérivée partielle  $\partial_j g(a)$ ,
2.  $\forall 1 \leq j \leq n, a \mapsto \partial_j g(a)$  est continue sur  $D$ .

*Exercice 6.9.* Justifier que la fonction

$$g : (x, y) \in \mathbb{R}^2 \mapsto y^2 \sin(xy) + x$$

est de classe  $C^1$  sur  $\mathbb{R}^2$ .

*Remarque 6.10.* On peut montrer (voir la partie III de ce chapitre) qu'une fonction de classe  $C^1$  est automatiquement continue.

Comme pour les fonctions dérivables, on peut ajouter, multiplier les fonctions différentiables et  $C^1$  (on verra dans la section suivante qu'on peut également les composer) :

**Proposition 6.11.** L'existence de dérivées partielles ou le fait d'être de classe  $C^1$  sont compatibles avec la somme, le produit, la multiplication par un nombre, l'inverse lorsqu'il est défini. De plus on a le formulaire :

$$\partial_j(g_1 + g_2)(a) = \partial_j g_1(a) + \partial_j g_2(a) \quad , \quad \partial_j(g_1 g_2)(a) = \partial_j g_1(a) g_2(a) + \partial_j g_2(a) g_1(a),$$

$$\partial_j(\lambda g_1)(a) = \lambda \partial_j g_1(a) \quad , \quad \partial_j \left( \frac{1}{g_1} \right) = -\partial_j g_1(a) \frac{1}{g_1(a)^2}.$$

La proposition précédente est la méthode la plus fréquente pour justifier qu'une fonction est  $C^1$ .

*Exemple 6.12.* Une forme linéaire est  $C^1$  sur  $\mathbb{R}^n$ .

Preuve : Une forme linéaire est une fonction de la forme  $\ell : x \mapsto a_1x_1 + \dots + a_nx_n$ . Alors pour tout  $1 \leq i \leq n$ ,

$$\partial_i \ell(x) = a_i$$

et les fonctions (constantes)  $x \mapsto a_i$  sont continues ! Donc  $\ell$  est de classe  $C^1$ .

*Exercice 6.13.* Justifier que la fonction suivante est  $C^1$ .

$$g : (x, y) \in \mathbb{R}^2 \mapsto \frac{\sin(x^3 - y^5)}{1 + x^2 + y^2}.$$

*Exercice 6.14.* Soit  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  définie par

$$f(x, y) = \begin{cases} \frac{(\sin x)^4}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0), \\ 0 & \text{sinon} \end{cases}.$$

Montrer que  $f$  est de classe  $C^1$  sur  $\mathbb{R}^2$ .

### Composition de fonctions de classe $C^1$

On rappelle que si  $f, g : \mathbb{R} \rightarrow \mathbb{R}$  sont deux fonctions dérivables, alors  $g \circ f$  est dérivable et pour tout  $a \in \mathbb{R}$ ,

$$(g \circ f)'(a) = g'(f(a)) \cdot f'(a).$$

On dispose de formules analogues pour les fonctions de plusieurs variables.

Considérons d'abord une fonction  $f : I \rightarrow \mathbb{R}^n$ , où  $I$  est un intervalle de  $\mathbb{R}$ . On note alors  $f(t) = (f_1(t), \dots, f_n(t))$  pour tout  $t \in I$ , les fonctions  $f_i : I \rightarrow \mathbb{R}$  étant les composantes de  $f$ . On dit alors que  $f$  est dérivable (respectivement  $C^1$ ) si chaque composante  $f_i$  l'est. Dans ce cas, la dérivée de  $f$  est

$$\forall t \in I, \quad f'(t) = (f'_1(t), \dots, f'_p(t)).$$

On se donne à présent une fonction  $g : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ , où  $D$  est un ouvert de  $\mathbb{R}^n$ . On suppose que  $f(I) \subset D$ , de sorte qu'on peut considérer la composée :  $G = g \circ f$ .

**Proposition 6.15.** *On suppose que  $f$  et  $g$  sont  $C^1$ . Alors la fonction  $G = g \circ f$  est  $C^1$  sur  $I$  et on a*

$$G'(t) = \partial_1 g(f(t)) f'_1(t) + \dots + \partial_n g(f(t)) f'_n(t) = Dg(f(t))[f'(t)].$$

On pourra trouver une preuve de cette proposition dans la partie *Approfondir* de ce chapitre.

*Exercice 6.16.* Soit  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$  une fonction  $C^1$ . Justifier que la fonction

$$t \mapsto g(t^2, t^3)$$

est  $C^1$  et calculer sa dérivée en fonction des dérivées partielles de  $g$ .

## 6.2.2 Dérivées partielles et différentielles des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}^p$

### Définitions

Soient  $D \subset \mathbb{R}^n$  ouvert et  $f : D \rightarrow \mathbb{R}^p$ . On note  $(f_1, \dots, f_p)$  les fonctions coordonnées de  $f$  :

$$x = (x_1, \dots, x_n) \mapsto \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_p(x_1, \dots, x_n) \end{pmatrix}.$$

Dans l'écriture précédente, on a utilisé une notation en vecteurs colonne, qu'on aurait pu aussi écrire :

$$(f_1(x_1, \dots, x_n), \dots, f_p(x_1, \dots, x_n)).$$

**Définition 6.17.** On dit que  $f$  admet une dérivée partielle par rapport à la  $j$ ème variable en  $a = (a_1, \dots, a_n) \in D$  si pour tout  $1 \leq i \leq p$ ,  $f_i$  admet une dérivée partielle par rapport à la  $j$ ème variable en  $a$ . On note

$$\partial_j f(a) = \begin{pmatrix} \partial_j f_1(a) \\ \vdots \\ \partial_j f_p(a) \end{pmatrix}.$$

Si  $f$  admet des dérivées partielles en  $a$  par rapport à toutes les variables, on peut alors définir la fonction

$$h = (h_1, \dots, h_n) \in \mathbb{R}^n \mapsto \begin{pmatrix} \sum_{j=1}^n h_j \partial_j f_1(a) \\ \vdots \\ \sum_{j=1}^n h_j \partial_j f_p(a) \end{pmatrix} = \begin{pmatrix} Df_1(a)[h] \\ \vdots \\ Df_p(a)[h] \end{pmatrix}.$$

Dans le vecteur colonne de droite, chaque composante  $Df_i(a)[h]$  est un réel. Comme il y en a  $p$ , elles forment un vecteur de  $\mathbb{R}^p$ . Comme chaque  $Df_i(a)$  est une application linéaire de  $\mathbb{R}^n$  dans  $\mathbb{R}$ , on voit que la fonction ci-dessus est une application linéaire de  $\mathbb{R}^n$  dans  $\mathbb{R}^p$ . On la note  $Df(a)$  et on a donc pour tout  $h \in \mathbb{R}^n$  :

$$Df(a)[h] = \begin{pmatrix} Df_1(a)[h] \\ \vdots \\ Df_p(a)[h] \end{pmatrix}.$$

On peut aussi écrire

$$Df(a)[h] = \begin{pmatrix} \sum_{j=1}^n h_j \partial_j f_1(a) \\ \vdots \\ \sum_{j=1}^n h_j \partial_j f_p(a) \end{pmatrix} = \begin{pmatrix} \partial_1 f_1(a) & \dots & \partial_n f_1(a) \\ \vdots & \ddots & \vdots \\ \partial_1 f_p(a) & \dots & \partial_n f_p(a) \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}.$$

La matrice

$$\begin{pmatrix} \partial_1 f_1(a) & \dots & \partial_n f_1(a) \\ \vdots & \ddots & \vdots \\ \partial_1 f_p(a) & \dots & \partial_n f_p(a) \end{pmatrix}$$

s'appelle la matrice jacobienne de  $f$  en  $a$ . On la notera  $M_{Df(a)}$  dans la suite. Observer que  $\partial_j f_k(a)$  est le coefficient placé sur la  $j$ ème colonne et la  $k$ ème ligne de la matrice jacobienne.

Puisque  $Df(a)$  est une application linéaire, pour tout  $h, k \in \mathbb{R}^n$  et pour tout  $\lambda \in \mathbb{R}$ ,

$$Df(a)[\lambda h + k] = \lambda Df(a)[h] + Df(a)[k].$$

*Exercice 6.18.* Calculer les dérivées partielles et la différentielle de la fonction suivante :

$$f : (x, y) \mapsto (e^{x^2+y^3}, xy)$$

*Exercice 6.19.* Calculer la matrice jacobienne pour la fonction de l'exercice 6.18.

**Définition 6.20.** On dit que  $f$  est de classe  $C^1$  sur  $D$  si pour tout  $1 \leq i \leq p$ ,  $f_i : D \rightarrow \mathbb{R}$  est de classe  $C^1$ .

En travaillant composante par composante, on obtient

**Proposition 6.21.** *L'existence d'une dérivée partielle ou le caractère  $C^1$  sont compatibles avec la somme, la multiplication par un nombre et on a le formulaire :  $\forall a \in D$ ,*

$$\partial_j(f+h)(a) = \partial_j f(a) + \partial_j h(a) \quad , \quad \partial_j(\lambda f)(a) = \lambda \partial_j f(a).$$

### Composition des fonctions $C^1$ pour les fonctions vectorielles

Soient  $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$  et  $g : E \subset \mathbb{R}^p \rightarrow \mathbb{R}^q$  deux fonctions  $C^1$  sur les ouverts  $D$  et  $E$  respectivement :

$$f(x_1, \dots, x_n) = \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_p(x_1, \dots, x_n) \end{pmatrix} \quad , \quad g(y_1, \dots, y_p) = \begin{pmatrix} g_1(y_1, \dots, y_p) \\ \vdots \\ g_q(y_1, \dots, y_p) \end{pmatrix}.$$

On suppose que  $f(D) \subset E$ , de sorte que la composée  $g \circ f$  est bien définie.

**Proposition 6.22.** *La fonction  $g \circ f : D \rightarrow \mathbb{R}^q$  est  $C^1$  et de plus, pour tout  $1 \leq j \leq n$ , pour tout  $a \in D$ ,*

$$\partial_j(g \circ f)(a) = \partial_1 g(f(a)) \partial_j f_1(a) + \dots + \partial_p g(f(a)) \partial_j f_p(a).$$

On trouvera une preuve de cette proposition dans la section *Approfondir* de ce chapitre.

L'égalité précédente est une égalité entre vecteurs de  $\mathbb{R}^q$ , qu'on peut encore expliciter : pour tout  $1 \leq k \leq q$ ,

$$\partial_j(g_k \circ f)(a) = \partial_1 g_k(f(a)) \partial_j f_1(a) + \dots + \partial_p g_k(f(a)) \partial_j f_p(a).$$

Le nombre de gauche est le coefficient de la matrice jacobienne  $M_{D(g \circ f)(a)}$  placé sur la  $k$ ème ligne et la  $j$ ème colonne.

On reconnaît dans le membre de droite le coefficient qu'on obtient sur la  $k$ ème ligne et la  $j$ ème colonne lorsqu'on fait le produit de la matrice  $M_{Dg(f(a))}$  avec la matrice  $M_{Df(a)}$ . Ainsi,

$$M_{D(g \circ f)(a)} = M_{Dg(f(a))} M_{Df(a)}.$$

Le produit  $M_{Dg(f(a))} M_{Df(a)}$  est la matrice de la composée  $Dg(f(a)) \circ Df(a)$ . Donc

$$\boxed{D(g \circ f)(a) = Dg(f(a)) \circ Df(a).} \quad (6.1)$$

*Exemple 6.23.* On considère les fonctions  $C^1$  :

$$h : (r, \theta) \mapsto (r \cos \theta, r \sin \theta) \quad , \quad f : (x, y) \mapsto (f_1(x, y), f_2(x, y), f_3(x, y)).$$

On pose  $G = f \circ h$ .

Alors

$$\partial_r G = \begin{pmatrix} \partial_1 f_1(r \cos \theta, r \sin \theta) \cos \theta \\ \partial_1 f_2(r \cos \theta, r \sin \theta) \cos \theta \\ \partial_1 f_3(r \cos \theta, r \sin \theta) \cos \theta \end{pmatrix} + \begin{pmatrix} \partial_2 f_1(r \cos \theta, r \sin \theta) \sin \theta \\ \partial_2 f_2(r \cos \theta, r \sin \theta) \sin \theta \\ \partial_2 f_3(r \cos \theta, r \sin \theta) \sin \theta \end{pmatrix},$$

ce qui correspond à

$$\partial_r G = \cos \theta \partial_x f(r \cos \theta, r \sin \theta) + \sin \theta \partial_y f(r \cos \theta, r \sin \theta).$$

De même,

$$\partial_\theta G = -r \sin \theta \partial_x f(r \cos \theta, r \sin \theta) + r \cos \theta \partial_y f(r \cos \theta, r \sin \theta).$$

On vérifie que

$$M_{DG(r, \theta)} = M_{Df(h(r, \theta))} M_{Dh(r, \theta)}.$$

## 6.3 Approfondir

### 6.3.1 Dérivées directionnelles

Considérons une fonction  $g$  définie sur un ensemble  $D \subset \mathbb{R}^n$  et prenons un point  $a = (a_1, \dots, a_n)$  dans l'intérieur de  $D$ . Notons  $(e_1, \dots, e_n)$  la base canonique de  $\mathbb{R}^n$ . Alors pour tout  $t \in \mathbb{R}$ ,  $a + te_1 = (a_1 + t, a_2, \dots, a_n)$  et le nombre  $\partial_1 g(a)$  est la limite (si elle existe) :

$$\lim_{t \rightarrow 0} \frac{g(a + te_1) - g(a)}{t}.$$

Autrement dit, on fait varier  $g$  dans la direction de la première coordonnée. De même,  $\partial_i g$  est obtenu en calculant les variations de  $g$  le long de la  $i^{\text{ème}}$  coordonnée. Mais on pourrait considérer la dérivée dans n'importe quelle autre direction, et obtenir ainsi une nouvelle généralisation de la dérivation des fonctions définies sur  $\mathbb{R}$ . On se donne un vecteur  $e$  non nul et on cherche à calculer, si elle existe, la limite

$$\lim_{t \rightarrow 0} \frac{g(a + te) - g(a)}{t}.$$

Lorsqu'elle existe, on note  $\partial_e g(a)$  la *dérivée directionnelle* de  $g$  dans la direction  $e$ . Lorsque  $e$  est un vecteur  $e_i$  de la base canonique, alors  $\partial_e g(a)$  coïncide avec  $\partial_i g(a)$ . Autrement dit, la notation  $\partial_i g(a)$  est un raccourci pour  $\partial_{e_i} g(a)$ .

Considérons l'exemple  $g : (x, y) \in \mathbb{R}^2 \mapsto x^2 e^y$ . Cherchons à calculer la dérivée de  $g$  en  $(1, 0)$  selon la direction  $e = (1, 1)$  : d'abord, pour tout  $t \in \mathbb{R}^*$ ,

$$\frac{g((1, 0) + te) - g(1, 0)}{t} = \frac{g(1+t, 0+t) - g(1, 0)}{t} = \frac{(1+t)^2 e^t - 1}{t}.$$

On reconnaît dans le membre de droite le taux d'accroissement de la fonction  $h : t \mapsto (1+t)^2 e^t$  en 0. On en déduit qu'il tend vers  $h'(0) = 3$ . Ainsi,  $\partial_e g(1, 0) = 3$ . On observe que

$$\partial_e g(1, 0) = \partial_1 g(1, 0) + \partial_2 g(1, 0).$$

Plus généralement, soit  $e' = \alpha e_1 + \beta e_2$ , où  $\alpha$  et  $\beta$  sont deux réels quelconques qui ne sont pas tous les deux nuls. On cherche encore à calculer  $\partial_{e'} g(1, 0)$ . Pour tout  $t \in \mathbb{R}^*$ ,

$$\frac{g((1, 0) + te') - g(1, 0)}{t} = \frac{g(1+t\alpha, 0+t\beta) - g(1, 0)}{t} = \frac{(1+t\alpha)^2 e^{t\beta} - 1}{t}.$$

Ici encore, pour calculer la limite lorsque  $t \rightarrow 0$ , on peut interpréter le membre de droite comme le taux d'accroissement de la fonction  $k : t \mapsto (1+t\alpha)^2 e^{t\beta}$  en 0. On en déduit qu'il tend vers  $k'(0) = 2\alpha + \beta$ . Ainsi,  $\partial_{e'} g(1, 0) = 2\alpha + \beta$ . On observe une fois de plus que

$$\partial_{e'} g(1, 0) = \alpha \partial_1 g(1, 0) + \beta \partial_2 g(1, 0).$$

*Exercice 6.24.* Reproduire les calculs précédents avec la fonction

$$h : (x, y) \in \mathbb{R}^2 \mapsto y^2 \sin(xy) + x.$$

Au regard de cet exemple, on pourrait être amené à conclure que pour calculer les dérivées directionnelles dans n'importe quelle direction, il suffit de connaître les dérivées partielles dans les deux directions de la base canonique. Ensuite, la décomposition d'un vecteur  $e = \sum_{i=1}^n \alpha_i e_i$  permet d'obtenir directement la dérivée directionnelle dans la direction  $e$  grâce à la formule :

$$\forall a \in \mathbb{R}^n, \quad \partial_e g(a) = \sum_{i=1}^n \alpha_i \partial_i g(a).$$

Cette formule n'est pourtant pas vraie en général. Considérons la fonction  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  telle que pour tout  $(x, y) \in \mathbb{R}^2$ ,

$$f(x, y) = \begin{cases} x & \text{si } y = 0, \\ 0 & \text{si } y \neq 0. \end{cases} \quad (6.2)$$

Alors, en  $(0, 0)$ , la fonction  $f$  admet des dérivées directionnelles dans toutes les directions. En effet, si  $e$  est un vecteur colinéaire à  $e_1$ , disons  $e = \alpha e_1$ , avec  $\alpha \neq 0$ , alors la fonction  $t \mapsto f(te) = f(\alpha t, 0) = t\alpha$  a pour dérivée  $\alpha$  en 0. Autrement dit,  $\partial_e f(0, 0) = \alpha$ . Si maintenant  $e$  est un vecteur non colinéaire à  $e_1$ , il s'écrit sous la forme  $e = \alpha e_1 + \beta e_2$ , avec  $\beta \neq 0$ . Alors pour tout  $t \in \mathbb{R}^*$ ,

$$\frac{f(te) - f(0, 0)}{t} = \frac{f(t\alpha, t\beta) - f(0, 0)}{t} = \frac{0}{t} = 0.$$

On en déduit que  $f$  admet une dérivée directionnelle selon  $e$ . Pourtant,  $\partial_1 f(0, 0) = 1$ ,  $\partial_2 f(0, 0) = 0$  et on n'a pas (pour  $\alpha \neq 0$ ) l'identité :

$$\partial_e f(0, 0) (= 0) = \alpha \partial_1 f(0, 0) + \beta \partial_2 f(0, 0) (= \alpha).$$

Ainsi, même si une fonction admet des dérivées directionnelles dans toutes les directions, il n'est pas toujours possible de les calculer linéairement à partir des dérivées partielles dans les directions de la base canonique.

### 6.3.2 Différentiabilité

Cependant, il existe une grande classe de fonctions pour lesquelles on peut exprimer les dérivées directionnelles linéairement à partir des dérivées partielles. Il s'agit des *fonctions différentiables*, qu'on peut considérer comme la *bonne* généralisation des fonctions dérivables dans le contexte des fonctions de plusieurs variables.

**Définition 6.25.** On dit que  $g : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$  est différentiable en un point  $a$  intérieur à  $D$  s'il existe une application linéaire  $L \in \mathcal{L}(\mathbb{R}^n; \mathbb{R})$  telle que

$$\lim_{x \rightarrow a} \frac{g(x) - g(a) - L(x - a)}{\|x - a\|} = 0.$$

La fonction  $x \mapsto \frac{g(x) - g(a) - L(x - a)}{\|x - a\|}$  est définie sur  $D \setminus \{a\}$ . Elle ressemble à un taux d'accroissement, à ceci près qu'au dénominateur, on prend la norme de  $x - a$ , au lieu de  $x - a$  (on ne sait pas diviser par un vecteur).

**Lemme 6.26.** L'application linéaire  $L$ , lorsqu'elle existe, est nécessairement unique.

*Démonstration.* En effet, s'il existait une autre application linéaire  $\tilde{L}$  vérifiant la même propriété, on aurait pour tout  $x \in D \setminus \{a\}$ ,

$$\begin{aligned} \left| L \left( \frac{x - a}{\|x - a\|} \right) - \tilde{L} \left( \frac{x - a}{\|x - a\|} \right) \right| &\leq \left| \frac{-g(x) + g(a)}{\|x - a\|} + \frac{L(x - a)}{\|x - a\|} \right| + \left| \frac{g(x) - g(a)}{\|x - a\|} - \frac{\tilde{L}(x - a)}{\|x - a\|} \right| \\ &\leq \left| \frac{g(x) - g(a) - L(x - a)}{\|x - a\|} \right| + \left| \frac{g(x) - g(a) - \tilde{L}(x - a)}{\|x - a\|} \right|. \end{aligned}$$

Par hypothèse, on sait que le membre de droite tend vers 0 lorsque  $x$  tend vers  $a$ .

Soit maintenant  $v \in \mathbb{R}^n$ ,  $\|v\| = 1$ . Comme  $a$  est un point intérieur à  $D$ , il existe  $\eta > 0$  tel que pour tout  $t \in ]-\eta, \eta[$ ,  $a + tv \in D$  (parce que  $\|(a + tv) - a\| = |t|\|v\| < \eta$ ). En prenant  $x = a + tv$  dans le

calcul précédent, le membre de gauche devient  $|L(v) - \tilde{L}(v)|$ , tandis que le membre de droite tend vers 0 lorsque  $t$  tend vers 0. On en déduit que  $|L(v) - \tilde{L}(v)| = 0$ , i.e.  $L(v) = \tilde{L}(v)$ . Ceci est vrai pour tout vecteur de norme 1, et donc aussi par linéarité, pour tout vecteur de  $\mathbb{R}^n$ . En conclusion,  $L = \tilde{L}$ , ce qui démontre l'unicité de l'application linéaire  $L$ .  $\square$

**Lemme 6.27.** *On suppose que  $g : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$  est différentiable en  $a \in D$  et on note  $L$  l'application correspondante comme dans la Définition 6.25. Alors*

1. toutes les dérivées directionnelles de  $g$  en  $a$  existent et pour tout  $e \in \mathbb{R}^n \setminus \{0\}$ ,

$$\partial_e g(a) = L[e].$$

2. Pour tout  $v = (v_1, \dots, v_n)$ ,

$$L[v] = v_1 \partial_1 g(a) + \dots + v_n \partial_n g(a).$$

*Démonstration.* Pour le premier point, on observe que pour tout  $t \in \mathbb{R}^*$ ,

$$\begin{aligned} \frac{g(a + te) - g(a)}{t} &= \frac{\|te\|}{t} \frac{g(a + te) - g(a) - Dg(a)[te]}{\|te\|} + \frac{Dg(a)[te]}{t} \\ &= \frac{|t|}{t} \|e\| \frac{g(a + te) - g(a) - Dg(a)[te]}{\|a + te - a\|} + Dg(a)[e]. \end{aligned}$$

Le premier terme tend vers 0 par définition de la différentielle. On en déduit que  $g$  admet une dérivée directionnelle selon  $e$ , et qui vaut  $Dg(a)[e]$ . Cela démontre le premier point. De plus, si  $e$  est un vecteur de la base canonique  $e_i$ , on obtient

$$Dg(a)[e_i] = \partial_i g(a).$$

On en déduit par linéarité de  $Dg(a)$  que pour tout  $v = (v_1, \dots, v_n)$ ,

$$\begin{aligned} \partial_v g(a) &= Dg(a)[v] = Dg(a)[v_1 e_1 + \dots + v_n e_n] \\ &= v_1 Dg(a)[e_1] + \dots + v_n Dg(a)[e_n] \\ &= v_1 \partial_1 g(a) + \dots + v_n \partial_n g(a). \end{aligned}$$

Cela achève la preuve.  $\square$

On reconnaît donc dans l'application linéaire  $L$  la différentielle de  $g$  en  $a$  qu'on avait déjà introduit dans la section *Assimiler*. Il faut prendre garde ici qu'on peut *définir la différentielle* dès que toutes les dérivées partielles de  $g$  existe. Cela n'implique pas pour autant que  $g$  soit *différentiable* (la fonction  $f$  définie par 6.2 en est un contreexemple). En revanche, le lemme précédent affirme que si  $g$  est différentiable, alors  $g$  admet des dérivées partielles dans toutes les directions.

Le lemme précédent affirme aussi que lorsqu'une fonction  $g$  est différentiable, on peut effectivement calculer toutes ses dérivées directionnelles en fonction des dérivées partielles.

*Exercice 6.28.* En admettant que la fonction

$$(x, y) \mapsto y^2 \sin(xy) + x$$

est différentiable sur  $\mathbb{R}^2$ , exprimer sa différentielle en tout point.

*Exercice 6.29.* Montrer qu'une fonction différentiable en un point est continue en ce point.

**Solution :** soit  $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$  qu'on suppose différentiable en un point  $a \in D$ . Alors pour tout  $x \neq a$ ,

$$\begin{aligned} |f(x) - f(a)| &\leq |f(x) - f(a) - Df(a)[x - a]| + \|Df(a)[x - a]\| \\ &\leq \frac{|f(x) - f(a) - Df(a)[x - a]|}{\|x - a\|} \|x - a\| + \|Df(a)\| \|x - a\|. \end{aligned}$$

Les deux termes du membre de gauche tendent vers 0 quand  $x \rightarrow a$ .

**Proposition 6.30.** Soit  $f : I \rightarrow \mathbb{R}^n$ . Alors  $f$  est dérivable si et seulement si  $f$  est différentiable.

*Démonstration.* Notons  $f = (f_1, \dots, f_n)$  les composantes de  $f$  : pour tout  $t \in I$ , on a donc  $f(t) = (f_1(t), \dots, f_n(t))$ . Dire que  $f$  est dérivable signifie que chaque composante  $f_i$  est dérivable. Si c'est le cas, alors

$$\begin{aligned} & \lim_{h \rightarrow 0} \frac{f(t+h) - f(t) - h(f'_1(t), \dots, f'_n(t))}{h} \\ &= \left( \lim_{h \rightarrow 0} \frac{f_1(t+h) - f_1(t) - hf'_1(t)}{h}, \dots, \lim_{h \rightarrow 0} \frac{f_n(t+h) - f_n(t) - hf'_n(t)}{h} \right) = (0, \dots, 0). \end{aligned}$$

On en déduit que  $f$  est différentiable en  $t$  et  $Df(t)$  est l'application linéaire

$$h \in \mathbb{R}^n \mapsto h(f'_1(t), \dots, f'_n(t)) \in \mathbb{R}^n.$$

Réciproquement, si  $f$  est différentiable en  $t$  alors il existe une application linéaire  $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$  telle que

$$\lim_{h \rightarrow 0} \frac{f(t+h) - f(t) - L[h]}{h} = 0.$$

On en déduit que le taux d'accroissement  $\frac{f(t+h)-f(t)}{h}$  a une limite, qui vaut  $L[1]$ . Ainsi,  $f$  est dérivable en  $t$ , de dérivée  $L(1)$ .  $\square$

La preuve montre donc

$$Df(t) : h \mapsto Df(t)[h] = h(f'_1(t), \dots, f'_n(t)).$$

*Exercice 6.31.* Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  une forme linéaire. Montrer que  $f$  est différentiable sur  $\mathbb{R}^n$  et que pour tout  $a \in \mathbb{R}^n$ ,  $Df(a) = f$ .

*Exercice 6.32.* Soit  $A$  une matrice symétrique. On note  $\langle \cdot, \cdot \rangle$  le produit scalaire canonique sur  $\mathbb{R}^n$ . On définit la fonction

$$f : x \mapsto \langle x, Ax \rangle.$$

Montrer que  $f$  est différentiable sur  $\mathbb{R}^n$ .

**Solution :** On forme pour tout  $x, h \in \mathbb{R}^n$  la différence

$$f(x+h) - f(x) = \langle x+h, A(x+h) \rangle - \langle x, Ax \rangle = \langle h, Ax \rangle + \langle x, Ah \rangle + \langle h, Ah \rangle.$$

Comme  $A$  est symétrique, on a  $\langle h, Ax \rangle + \langle x, Ah \rangle = 2\langle Ax, h \rangle$ . Pour tout  $x \in \mathbb{R}^n$ , la fonction  $L : h \mapsto 2\langle Ax, h \rangle$  est linéaire. C'est donc un bon candidat pour être la différentielle de  $f$  en  $x$ .

On a par ailleurs  $|\langle h, Ah \rangle| \leq \|h\|_2 \|Ah\|_2$  par l'inégalité de Cauchy-Schwarz, puis

$$\|Ah\|_2 \leq \sum_{i=1}^n |h_i| \|Ae_i\|_2 \leq \left( \max_{1 \leq i \leq n} \|Ae_i\|_2 \right) \|h\|_1.$$

Notons  $M := \left( \max_{1 \leq i \leq n} \|Ae_i\|_2 \right)$ . On a donc montré :

$$\left| \frac{f(x+h) - f(x) - L[h]}{\|h\|_2} \right| = M \|h\|_1.$$

On en déduit que le taux d'accroissement du membre de gauche tend vers 0 quand  $h$  tend vers 0. Ainsi,  $f$  est différentiable de différentielle  $L$ .

Exercice 6.33. Soit

$$f : (x, y) \in \mathbb{R}^2 \mapsto \begin{cases} (x^2 + y^2) \sin \frac{1}{\sqrt{x^2 + y^2}}, & \text{si } (x, y) \neq (0, 0), \\ 0, & \text{sinon.} \end{cases}$$

Montrer que les dérivées partielles de  $f$  en  $(0, 0)$  existent et les calculer. Est-ce que  $f$  est différentiable en  $(0, 0)$  ?

**Théorème 6.34.** Si  $g$  est de classe  $C^1$  sur  $D$ , alors  $g$  est différentiable sur  $D$ .

On pourra admettre la preuve de ce théorème.

*Démonstration.* On prouve le résultat pour  $n = 2$  (le cas général est analogue mais moins lisible). On utilise la norme  $\|\cdot\|_\infty$ . On fixe  $a = (a_1, a_2) \in \mathbb{R}^2$ . Soit  $\varepsilon > 0$ . Comme  $\partial_1 g$  et  $\partial_2 g$  sont continues en  $a$ , il existe  $\eta > 0$  tel que pour tout  $\|(x_1, x_2) - (a_1, a_2)\|_\infty < \eta$ ,

$$|\partial_1 g(x_1, x_2) - \partial_1 g(a_1, a_2)| \leq \varepsilon \quad , \quad |\partial_2 g(x_1, x_2) - \partial_2 g(a_1, a_2)| \leq \varepsilon. \quad (6.3)$$

Pour tout  $(x_1, x_2) \in B((a_1, a_2), \eta)$ , on écrit

$$g(x_1, x_2) - g(a_1, a_2) = (g(x_1, x_2) - g(x_1, a_2)) + (g(x_1, a_2) - g(a_1, a_2)).$$

Par le théorème fondamental du calcul différentiel appliqué à la fonction  $x_2 \mapsto g(x_1, x_2)$ ,

$$g(x_1, x_2) - g(x_1, a_2) = \int_0^1 \frac{d}{dt} (g(x_1, a_2 + t(x_2 - a_2))) dt = \int_0^1 \partial_2 g(x_1, a_2 + t(x_2 - a_2))(x_2 - a_2) dt.$$

On en déduit

$$|g(x_1, x_2) - g(x_1, a_2) - \partial_2 g(a_1, a_2)(x_2 - a_2)| \leq |x_2 - a_2| \int_0^1 |\partial_2 g(x_1, a_2 + t(x_2 - a_2)) - \partial_2 g(a_1, a_2)| dt.$$

Pour tout  $t \in [0, 1]$ , on a  $\|(x_1, a_2 + t(x_2 - a_2)) - (a_1, a_2)\|_\infty < \eta$ . Donc par (6.3),

$$|\partial_2 g(x_1, a_2 + t(x_2 - a_2)) - \partial_2 g(a_1, a_2)| \leq \varepsilon.$$

Ainsi,

$$|g(x_1, x_2) - g(x_1, a_2) - \partial_2 g(a_1, a_2)(x_2 - a_2)| \leq \varepsilon |x_2 - a_2|.$$

On montre de même

$$|g(x_1, a_2) - g(a_1, a_2) - \partial_1 g(a_1, a_2)(x_1 - a_1)| \leq \varepsilon |x_1 - a_1|.$$

Il suit que

$$|g(x_1, x_2) - g(a_1, a_2) - (x_1 - a_1)\partial_1 g(a_1, a_2) - (x_2 - a_2)\partial_2 g(a_1, a_2)| \leq \varepsilon(|x_1 - a_1| + |x_2 - a_2|) = 2\varepsilon \|(x_1 - a_1, x_2 - a_2)\|_\infty.$$

Pour résumer, on a montré que si  $0 < \|(x_1 - a_1, x_2 - a_2)\|_\infty < \eta$ , alors

$$\left| \frac{g(x_1, x_2) - g(a_1, a_2) - (x_1 - a_1)\partial_1 g(a_1, a_2) - (x_2 - a_2)\partial_2 g(a_1, a_2)}{\|(x_1 - a_1, x_2 - a_2)\|_\infty} \right| \leq 2\varepsilon.$$

Ainsi,

$$\lim_{(x_1, x_2) \rightarrow (a_1, a_2)} \frac{g(x_1, x_2) - g(a_1, a_2) - (x_1 - a_1)\partial_1 g(a_1, a_2) - (x_2 - a_2)\partial_2 g(a_1, a_2)}{\|(x_1 - a_1, x_2 - a_2)\|_\infty} = 0.$$

Cela montre que  $f$  est différentiable en  $(a_1, a_2)$ , de différentielle

$$(h_1, h_2) \mapsto h_1 \partial_1 g(a_1, a_2) + h_2 \partial_2 g(a_1, a_2).$$

Cela conclut la preuve du théorème. □

*Remarque 6.35.* On peut voir que lorsque  $g$  est de classe  $C^1$  sur  $D$ , alors l'application

$$a \in D \mapsto Dg(a) \in \mathcal{L}(\mathbb{R}^n; \mathbb{R})$$

est continue sur  $D$  (pour n'importe quelle norme sur  $\mathcal{L}(\mathbb{R}^n; \mathbb{R})$  puisqu'elles sont toutes équivalentes).

## Gradients et variation d'une fonction

On dit parfois que le gradient d'une fonction différentiable  $g : D \rightarrow \mathbb{R}$  en un point  $a$  intérieur à  $D$  indique *la direction de plus grande pente*. Cette phrase peut être comprise dans le sens suivant. Etant donné un vecteur  $e \in \mathbb{R}^n$  qu'on suppose unitaire pour la norme  $\|\cdot\|_2$ , c'est-à-dire  $\|e\|_2 = 1$ , la variation infinitésimale de  $g$  dans la direction  $e$  est donnée par

$$\lim_{t \rightarrow 0} \frac{g(a + te) - g(a)}{t} = \partial_e g(a) = Dg(a)[e] = \langle \nabla g(a), e \rangle.$$

Pour maximiser cette variation infinitésimale, on cherche la direction  $e$  pour laquelle  $\langle \nabla g(a), e \rangle$  est maximale, autrement dit :

$$\sup_{\substack{e \in \mathbb{R}^n, \\ \|e\|=1}} \langle \nabla g(a), e \rangle.$$

Par le théorème de Cauchy-Schwarz, pour tout  $e \in \mathbb{R}^n$ , on a  $\langle \nabla g(a), e \rangle \leq \|\nabla g(a)\| \|e\|$ , avec égalité ssi  $e$  est positivement colinéaire à  $\nabla g(a)$ . Ainsi, le vecteur  $e$  qui maximise les variations infinitésimales est  $\nabla g(a) / \|\nabla g(a)\|$  (pourvu que  $\nabla g(a) \neq 0$ ).

### 6.3.3 Différentiabilité des fonctions vectorielles

Comme pour les dérivées partielles, on peut étendre aux fonctions à valeurs vectorielles la notion de différentiabilité introduite précédemment pour les fonctions à valeurs dans  $\mathbb{R}$ .

**Définition 6.36.** On dit que  $f$  est différentiable en  $a = (a_1, \dots, a_n) \in D$  si pour tout  $1 \leq i \leq p$ ,  $f_i$  est différentiable en  $a$ . On note

$$h \in \mathbb{R}^n \mapsto Df(a)[h] = \begin{pmatrix} Df_1(a)[h] \\ \vdots \\ Df_p(a)[h] \end{pmatrix}.$$

La différentiabilité d'une fonction  $f$  peut être établie par l'utilisation de taux d'accroissements :

**Proposition 6.37.** Une fonction  $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$  est différentiable en un point  $a \in D$  si et seulement si il existe une application linéaire  $L : \mathbb{R}^n \rightarrow \mathbb{R}^p$  telle

$$\lim_{x \rightarrow a} \frac{\|f(x) - f(a) - L[x - a]\|}{\|x - a\|} = 0.$$

*Démonstration.* Notons  $f_1, \dots, f_p$  les composantes de  $f$  et supposons qu'elles sont différentiables en  $a$ . Sur  $\mathbb{R}^p$ , on utilise la norme 1. Notons  $L[h] = (Df_1(a)[h], \dots, Df_p(a)[h])$ . Alors

$$f(x) - f(a) - L[x - a] = (f_1(x) - f_1(a) - Df_1(a)[x - a], \dots, f_p(x) - f_p(a) - Df_p(a)[x - a]).$$

Donc pour  $x \neq a$ ,

$$\frac{\|f(x) - f(a) - L[x - a]\|_1}{\|x - a\|} = \frac{\sum_{j=1}^p |f_j(x) - f_j(a) - Df_j(a)[x - a]|}{\|x - a\|}$$

Pour tout  $1 \leq j \leq p$ , on a

$$\lim_{x \rightarrow a} \frac{|f_j(x) - f_j(a) - Df_j(a)[x - a]|}{\|x - a\|} = 0.$$

On en déduit que

$$\lim_{x \rightarrow a} \frac{\|f(x) - f(a) - L[x - a]\|_1}{\|x - a\|} = 0.$$

Réciproquement, s'il existe une application linéaire  $L : \mathbb{R}^n \rightarrow \mathbb{R}^p$  telle

$$\lim_{x \rightarrow a} \frac{\|f(x) - f(a) - L[x - a]\|}{\|x - a\|} = 0,$$

montrons que chaque  $f_j$ , par exemple  $f_1$  est différentiable en  $a$ . Notons  $L_j$  la  $j^{\text{ème}}$  composante de  $L$ , qui est une application linéaire de  $\mathbb{R}^n$  dans  $\mathbb{R}$ . On a

$$|f_1(x) - f_1(a) - L_1[x - a]| \leq \sum_{j=1}^p |f_j(x) - f_j(a) - L_j[x - a]| = \|f(x) - f(a) - L[x - a]\|_1.$$

On en déduit que

$$\lim_{x \rightarrow a} \frac{|f_1(x) - f_1(a) - L_1[x - a]|}{\|x - a\|} = 0.$$

Donc  $f_1$  est différentiable en  $a$  et  $Df_1(a) = L_1$ , ce qui conclut la preuve.  $\square$

La proposition précédente montre alors que la fonction

$$F : x \mapsto \begin{cases} \frac{f(x) - f(a) - Df(a)[x - a]}{\|x - a\|} & \text{si } x \neq a, \\ 0 & \text{si } x = a \end{cases}$$

est continue si et seulement si  $f$  est différentiable en  $a$ .

On dit que  $f : D \rightarrow \mathbb{R}^p$  est différentiable sur un ouvert  $D$  si elle est différentiable en tout point de  $D$ .

A partir du Théorème 6.34, on obtient le résultat fondamental (en travaillant composante par composante) :

**Théorème 6.38.** Soit  $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$  une fonction de classe  $C^1$ . Alors  $f$  est différentiable sur  $D$ .

### 6.3.4 Composition de fonctions différentiables

Soient  $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$  et  $g : E \subset \mathbb{R}^p \rightarrow \mathbb{R}^q$  deux fonctions différentiables sur les ouverts  $D$  et  $E$  respectivement.

**Théorème 6.39.** La fonction  $g \circ f$  est différentiable sur  $D$  et

$$D(g \circ f)(a) = Dg(f(a)) \circ Df(a) \quad , \quad a \in D.$$

On se souvient que  $Df(a)$  est une application linéaire de  $\mathbb{R}^n$  vers  $\mathbb{R}^p$  et que  $Dg(f(a))$  est une application linéaire de  $\mathbb{R}^p$  vers  $\mathbb{R}^q$ . La composée est donc une application linéaire de  $\mathbb{R}^n$  vers  $\mathbb{R}^q$ , ce qui est aussi le cas de  $D(g \circ f)(a)$ .

On pourra admettre le théorème...

*Démonstration.* Introduisons les deux fonctions

$$F(x) = \begin{cases} \frac{f(x) - f(a) - Df(a)[x - a]}{\|x - a\|} & \text{si } x \neq a, \\ 0 & \text{si } x = a, \end{cases}$$

$$G(y) = \begin{cases} \frac{g(y) - g(f(a)) - Dg(f(a))[y - f(a)]}{\|y - f(a)\|} & \text{si } y \neq f(a), \\ 0 & \text{si } y = f(a). \end{cases}$$

Comme  $f$  est différentiable en  $a$ ,  $F$  est continue en  $a$ . De même, comme  $g$  est différentiable en  $f(a)$ ,  $G$  est continue en  $f(a)$ . Notre but est de montrer que

$$\lim_{x \rightarrow a} \frac{g \circ f(x) - g \circ f(a) - Dg(f(a)) \circ Df(a)[x - a]}{\|x - a\|} = 0.$$

On observe que pour tout  $x \neq a$ ,

$$\begin{aligned} \frac{g \circ f(x) - g \circ f(a) - Dg(f(a)) \circ Df(a)[x - a]}{\|x - a\|} &= \frac{g(f(x)) - g(f(a)) - Dg(f(a))[Df(a)[x - a]]}{\|x - a\|} \\ &= \frac{g(f(x)) - g(f(a)) - Dg(f(a))[f(x) - f(a)]}{\|x - a\|} \\ &\quad + \frac{Dg(f(a))[f(x) - f(a) - Df(a)[x - a]]}{\|x - a\|}. \end{aligned}$$

Si  $f(x) \neq f(a)$ , on a

$$\begin{aligned} \frac{g(f(x)) - g(f(a)) - Dg(f(a))[f(x) - f(a)]}{\|x - a\|} &= \frac{g(f(x)) - g(f(a)) - Dg(f(a))[f(x) - f(a)]}{\|f(x) - f(a)\|} \frac{\|f(x) - f(a)\|}{\|x - a\|} \\ &= G(f(x)) \frac{\|f(x) - f(a)\|}{\|x - a\|}. \end{aligned}$$

Cette égalité reste vraie si  $f(x) = f(a)$ . Par ailleurs,

$$\frac{Dg(f(a))[f(x) - f(a) - Df(a)[x - a]]}{\|x - a\|} = Dg(f(a)) \left[ \frac{f(x) - f(a) - Df(a)[x - a]}{\|x - a\|} \right] = Dg(f(a))[F(x)].$$

On a donc montré que

$$\frac{g \circ f(x) - g \circ f(a) - Dg(f(a)) \circ Df(a)[x - a]}{\|x - a\|} = G(f(x)) \frac{\|f(x) - f(a)\|}{\|x - a\|} + Dg(f(a))[F(x)].$$

Ainsi,

$$\frac{\|g \circ f(x) - g \circ f(a) - Dg(f(a)) \circ Df(a)[x - a]\|}{\|x - a\|} \leq \|G(f(x))\| \frac{\|f(x) - f(a)\|}{\|x - a\|} + \|Dg(f(a))[F(x)]\|.$$

Comme  $Dg(a)$  est linéaire, elle est continue. Or,  $F$  est continue en  $a$ . Par composition,  $\lim_{x \rightarrow a} Dg(f(a))[F(x)] = 0$ . En ce qui concerne le premier terme, on sait déjà que  $G$  est continue en  $f(a)$ . Comme  $f$  est continue en  $a$ , on en déduit que  $\lim_{x \rightarrow a} G(f(x)) = 0$ . Il reste à estimer le terme

$$\frac{\|f(x) - f(a)\|}{\|x - a\|} \leq \frac{\|f(x) - f(a) - Df(a)(x - a)\|}{\|x - a\|} + \frac{\|Df(a)(x - a)\|}{\|x - a\|} = \|F(x)\| + \left\| Df(a) \left( \frac{x - a}{\|x - a\|} \right) \right\|.$$

On a  $\lim_{x \rightarrow a} F(x) = 0$  et

$$\left\| Df(a) \left( \frac{x - a}{\|x - a\|} \right) \right\| \leq \|Df(a)\|.$$

Dans le membre de droite,  $\|Df(a)\|$  désigne la norme opérateur de l'application linéaire  $Df(a)$ , voir au besoin la fin du chapitre 2. On en déduit que

$$\lim_{x \rightarrow a} \|G(f(x))\| \frac{\|f(x) - f(a)\|}{\|x - a\|} = 0.$$

Finalement, on a bien montré que

$$\lim_{x \rightarrow a} \frac{g \circ f(x) - g \circ f(a) - Dg(f(a)) \circ Df(a)[x - a]}{\|x - a\|} = 0. \quad \square$$

La traduction matricielle du théorème s'écrit en termes de matrices jacobiniennes :

$$M_{D(g \circ f)(a)} = M_{Dg(f(a))} M_{Df(a)}.$$

A partir du Théorème 6.39, on peut démontrer les Proposition 6.15 et 6.22.

### Preuve de la Proposition 6.15

Si  $f : I \rightarrow \mathbb{R}^n$  est  $C^1$ , alors elle est dérivable et donc différentiable par la Proposition 6.30. De plus, on a pour tous  $t, h \in \mathbb{R}$  :

$$Df(t)[h] = hf'(t) = (hf'_1(t), \dots, hf'_n(t)).$$

Alors par le Théorème 6.39, on en déduit que  $g \circ f$  est différentiable, et sa différentielle vérifie

$$\begin{aligned} D(g \circ f)(t)[h] &= (Dg(f(t)) \circ Df(t))[h] = Dg(f(t))[Df(t)[h]] \\ &= Dg(f(t))[hf'(t)] \\ &= Dg(f(t))[(hf'_1(t), \dots, hf'_n(t))]. \end{aligned}$$

Le vecteur  $(hf'_1(t), \dots, hf'_n(t))$  peut se réécrire  $\sum_{i=1}^n hf'_i(t)e_i$ , en notant  $e_i$  les vecteurs de la base canonique de  $\mathbb{R}^n$ . Donc par linéarité de  $Dg(f(t))$

$$D(g \circ f)(t)[h] = \sum_{i=1}^n hf'_i(t)Dg(f(t))[e_i] = h \sum_{i=1}^n f'_i(t)\partial_i g(f(t)).$$

Par ailleurs, en appliquant à nouveau la Proposition 6.30,  $g \circ f$  est dérivable et

$$h(g \circ f)'(t) = D(g \circ f)(t)[h].$$

Donc

$$h(g \circ f)'(t) = h \sum_{i=1}^n f'_i(t)\partial_i g(f(t)).$$

Comme l'égalité précédente est vraie pour tout  $h \in \mathbb{R}$ , elle l'est en particulier pour  $h = 1$ . On a donc montré que

$$(g \circ f)'(t) = \sum_{i=1}^n f'_i(t)\partial_i g(f(t)).$$

Comme chaque  $f_i$  et  $g$  sont  $C^1$ , le membre de droite de l'égalité est continue. Donc  $(g \circ f)'$  est continue. Cela montre que  $g \circ f$  est  $C^1$ , avec l'expression attendue de sa dérivée. □

### Preuve de la Proposition 6.15

Par le Théorème 6.39, on sait que  $g \circ f$  est différentiable, avec pour tout  $a \in D$  et pour tout  $h \in \mathbb{R}^n$  :

$$D(g \circ f)(a)[h] = Dg(f(a))[Df(a)[h]].$$

Appliquons cette inégalité à  $h = e_j$  (le  $j^{\text{ème}}$  vecteur de la base canonique de  $\mathbb{R}^n$ ) :

$$D(g \circ f)(a)[e_j] = Dg(f(a))[Df(a)[e_j]].$$

De manière équivalente,

$$\partial_j(g \circ f)(a) = Dg(f(a))[\partial_j f(a)].$$

Notons  $\varepsilon_i$  le  $i^{\text{ème}}$  vecteur de la base canonique de  $\mathbb{R}^n$ . En écrivant  $\partial_j f(a) = (\partial_j f_1(a), \dots, \partial_j f_p(a)) = \sum_{i=1}^p \partial_j f_i(a)\varepsilon_i$ , il vient

$$\partial_j(g \circ f)(a) = \sum_{i=1}^p \partial_j f_i(a)Dg(f(a))[\varepsilon_i] = \sum_{i=1}^p \partial_j f_i(a)\partial_i g(f(a)).$$

C'est la formule attendue. Comme  $f$  et  $g$  sont supposés  $C^1$ , le membre de droite de l'égalité précédente est continue. Il en est donc de même du membre de gauche, ce qui prouve que  $g \circ f$  est aussi  $C^1$ .  $\square$

### 6.3.5 Théorème des accroissements finis

Considérons une fonction  $f : D \rightarrow \mathbb{R}$  une fonction  $C^1$ . Alors pour tout  $a, b \in D$  tels que

$$[a, b] = \{(1-t)a + tb; t \in [0, 1]\} \subset D,$$

on peut considérer la fonction  $g : t \in [0, 1] \mapsto f((1-t)a + tb)$ . Par composition, il s'agit d'une fonction  $C^1$  sur  $[0, 1]$ . Par le théorème des accroissements finis appliqué à  $g$ , on a

$$|g(1) - g(0)| \leq \sup_{c \in ]0, 1[} |g'(c)|.$$

Ainsi,

$$|f(b) - f(a)| \leq \sup_{c \in ]0, 1[} |g'(c)|.$$

Par la règle de la chaîne, on calcule

$$g'(t) = Df((1-t)a + tb)(b-a) = \langle \nabla f((1-t)a + tb), b-a \rangle.$$

Par le théorème de Cauchy-Schwarz,

$$|\langle \nabla f((1-t)a + tb), b-a \rangle| \leq \|\nabla f((1-t)a + tb)\|_2 \|b-a\|_2.$$

On en déduit

#### **Théorème 6.40.**

$$|f(b) - f(a)| \leq \sup_{x \in [a, b]} \|\nabla f(x)\|_2 \|b-a\|_2.$$

*Exercice 6.41.* Soit  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  une fonction  $C^1$  telle que pour tout  $(x, y) \in \mathbb{R}^2$ ,

$$|\partial_1 f(x, y)| + |\partial_2 f(x, y)| \leq 1.$$

Montrer que pour tout  $(x, y) \in \mathbb{R}^2$ ,

$$|f(x, y)| \leq |f(0, 0)| + \max(|x|, |y|).$$



## Chapitre 7

# Quelques compléments sur les fonctions de plusieurs variables

Dans ce chapitre, on poursuit l'étude de la différentiabilité des fonctions de plusieurs variables. On commence par le *théorème d'inversion globale* qui permet de justifier facilement des changements de variables. Puis on introduit les *dérivées partielles d'ordre supérieur*. Elles nous permettront d'identifier les *minima* d'une fonction.

### 7.1 Découvrir

#### 7.1.1 Changement de variables

Dans cette section,  $D$  et  $\Omega$  désignent deux ouverts de  $\mathbb{R}^n$  (pour rappel, cela signifie que tous les points de  $D$  sont intérieurs à  $D$  et de même pour  $\Omega$ ).

**Définition 7.1.** On dit que  $f : D \rightarrow \mathbb{R}^n$  est un  $C^1$  difféomorphisme de  $D$  sur  $\Omega$  si

1.  $f$  est  $C^1$ ,
2.  $f$  est une bijection de  $D$  sur  $\Omega$ ,
3.  $f^{-1}$  est de classe  $C^1$  dans  $\Omega$ .

Il est facile de voir que si  $f$  est un  $C^1$  difféomorphisme, alors sa différentielle en tout point (qui est une application linéaire) est inversible. C'est l'objet de la proposition suivante.

**Proposition 7.2.** Soit  $f : D \rightarrow \mathbb{R}^n$  un  $C^1$  difféomorphisme sur  $\Omega$ . Notons  $h = f^{-1}$  son inverse. Alors pour tout  $a \in D$ ,  $Df(a)$  est une application linéaire inversible et

$$[Df(a)]^{-1} = Dh(f(a)).$$

Preuve : Pour tout  $a \in D$ , on a  $h \circ f(a) = a$ . Les dérivées partielles de l'application linéaire  $i : a \in \mathbb{R}^n \mapsto a \in \mathbb{R}^n$  sont  $\partial_j i(a) = e_j$  (le  $j^{\text{ème}}$  vecteur de la base canonique). Comme  $h \circ f(a) = a$ , on obtient donc  $\partial_j (h \circ f)(a) = e_j$ . En répétant ce calcul pour tous les  $j$ , on en déduit que  $D(h \circ f)(a)$  est l'application identité de  $\mathbb{R}^n$  dans lui-même. Par la règle de calcul de la dérivée partielle d'une composée (voir (6.1)), il vient donc  $Dh(f(a)) \circ Df(a) = Id$ . Ainsi,  $Df(a)$  est une application linéaire inversible, d'inverse  $Dh(f(a))$ . □

La proposition a une réciproque. Pour l'énoncer, nous aurons besoin de la notion de jacobien :

**Définition 7.3.** On note  $J_f(a) = \det M_{Df(a)}$  le jacobien de  $f$  en  $a$ .

Le théorème suivant, appelé théorème d'inversion globale, donne une condition nécessaire et suffisante pour que  $f$  soit un difféomorphisme.

**Théorème 7.4.** Soit  $f : D \rightarrow \mathbb{R}^n$  une application  $C^1$ . On suppose que  $f$  est injective et que

$$\forall a \in D, \quad J_f(a) = \det Df(a) \neq 0.$$

Alors l'ensemble  $\Omega = f(D)$  est un ouvert et  $f$  est un  $C^1$  difféomorphisme de  $D$  sur  $\Omega$ .

## 7.1.2 Dérivées partielles d'ordre supérieur

Dans cette section,  $D$  est un ouvert de  $\mathbb{R}^n$  et  $f : D \rightarrow \mathbb{R}^p$ .

**Définition 7.5.** On dit que  $f$  est de classe  $C^2$  si  $f$  est  $C^1$  et toutes les dérivées partielles de  $f$ , i.e. les fonctions  $\partial_j f : D \rightarrow \mathbb{R}^p$  sont aussi  $C^1$ .

On note

$$\partial_i(\partial_j f) = \partial_{ij}^2 f.$$

**Théorème 7.6** (Théorème de Schwarz). Si  $f$  est de classe  $C^2$  sur  $D$ , alors

$$\partial_{ij}^2 f = \partial_{ji}^2 f.$$

*Exercice 7.7.* Soit  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  définie par

$$f(x, y) = \begin{cases} \frac{xy(x^2 - y^2)}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0), \\ 0 & \text{sinon.} \end{cases}$$

Montrer que  $\partial_{12} f(0, 0)$  et  $\partial_{21} f(0, 0)$  existent. Déterminer ces deux nombres. Qu'en déduire ?

**Définition 7.8.** On définit par récurrence les fonctions de classe  $C^n$  :  $f$  est de classe  $C^n$  si  $f$  est  $C^1$  et toutes les dérivées partielles de  $f$  sont  $C^{n-1}$ . De manière équivalente,  $f$  est  $C^n$  si  $f$  est  $C^{n-1}$  et les dérivées partielles de  $f$  d'ordre  $n - 1$  sont  $C^1$ .

## 7.2 Assimiler

### 7.2.1 Quelques exemples de changements de variables

*Exemple 7.9* (Coordonnées polaires dans le plan). La fonction

$$\mathcal{C} : (r, \theta) \in D = ]0, +\infty[ \times ]-\pi, \pi[ \mapsto (r \cos \theta, r \sin \theta) \in \Omega = \mathbb{R}^2 \setminus \{(x, 0), x \leq 0\}$$

est un  $C^1$  difféomorphisme de  $D$  sur  $\Omega$ .

Pour justifier cette affirmation, commençons par observer que  $\mathcal{C}$  est bien à valeurs dans  $\Omega$ . En effet, pour tout  $(r, \theta) \in D$ , si  $r \sin \theta = 0$ , c'est que  $\theta = 0$ , et donc  $r \cos \theta = r > 0$ . Donc  $\mathcal{C}(r, \theta) \in \Omega$ .

Ensuite,  $\mathcal{C}$  est  $C^1$  (produit de fonction  $C^1$ ). Montrons que  $\mathcal{C} : D \rightarrow \Omega$  est surjective. Pour tout  $(x, y) \in \Omega$ , posons  $r = \|(x, y)\|_2 = \sqrt{x^2 + y^2}$ . Alors le vecteur  $(\frac{x}{r}, \frac{y}{r})$  est de norme 2 égale à 1. On se souvient que l'ensemble des vecteurs  $(a, b)$  de  $\mathbb{R}^2$  tels que  $\|(a, b)\|_2 = 1$  est le cercle centré en  $(0, 0)$  et de rayon 1. Il existe donc  $\theta \in ]-\pi, \pi[$  défini par

$$\cos \theta = \frac{x}{r}, \quad \sin \theta = \frac{y}{r}.$$

Comme  $(x/r, y/r) \neq (-1, 0)$ , on peut bien exclure  $\theta = \pm\pi$ . Alors  $\mathcal{C}(r, \theta) = (x, y)$  donc  $\mathcal{C}$  est surjective de  $D$  sur  $\Omega$ .

Par ailleurs,  $\mathcal{C}$  est injective : en effet, si  $(r \cos \theta, r \sin \theta) = (r' \cos \theta', r' \sin \theta')$ , alors en prenant la norme 2 de ces vecteurs, on trouve  $r = r'$ , puis par injectivité de  $\theta \mapsto (\cos \theta, \sin \theta)$  sur  $] - \pi, \pi[$ , on en déduit  $\theta = \theta'$ . Ainsi,  $\mathcal{C}$  est une bijection de  $D$  sur  $\Omega$ .

Enfin, on calcule  $J_{\mathcal{C}}(x, y) = r \neq 0$ . Le théorème s'applique :  $\mathcal{C}$  est un  $C^1$  difféomorphisme de  $D$  sur  $\Omega$ . On en déduit que  $\mathcal{C}^{-1}$  est également  $C^1$ .

*Exemple 7.10.* En revanche,

$$\tilde{\mathcal{C}} : (r, \theta) \in [0, +\infty[ \times ] - \pi, \pi[ \mapsto (r \cos \theta, r \sin \theta) \in \mathbb{R}^2 \setminus \{(x, 0), x < 0\}$$

n'est pas injectif de  $[0, +\infty[ \times ] - \pi, \pi[$  sur son image, et n'est donc pas un difféomorphisme.

*Exemple 7.11.* La fonction

$$\psi : (x, y) \in D = \mathbb{R} \times \mathbb{R}^{*+} \mapsto \left( \frac{x}{y}, x^2 + y^2 \right) \in \Omega = \mathbb{R} \times \mathbb{R}^{*+}.$$

est un  $C^1$  difféomorphisme de  $D$  sur  $\Omega$ .

Comme  $x^2 + y^2 > 0$  pour tout  $(x, y) \in D$ , on vérifie que  $\psi(D) \subset \Omega$ . On observe aussi que les fonctions  $x \mapsto \frac{x}{y}$  et  $(x, y) \mapsto x^2 + y^2$  sont  $C^1$  sur  $D$  comme somme, produit et inverse de fonctions  $C^1$ , le dénominateur ne s'annulant pas. On peut alors calculer la jacobienne de  $\psi$  en un point  $(x, y) \in D$  :

$$J_{\psi}(x, y) = \begin{pmatrix} \frac{1}{y} & \frac{-x}{y^2} \\ 2x & 2y \end{pmatrix}$$

d'où l'on déduit le jacobien  $J_{\psi}(x, y) = 2(1 + \frac{x^2}{y^2}) > 0$ . Soit  $(u, v) \in \Omega$ . On cherche à résoudre le système

$$\begin{cases} \frac{x}{y} = u \\ x^2 + y^2 = v \end{cases} \iff \begin{cases} x = yu \\ y^2(u^2 + 1) = v \end{cases}$$

Comme  $v > 0$ , on obtient l'unique solution  $(x, y) = (u\sqrt{\frac{v}{1+u^2}}, \sqrt{\frac{v}{1+u^2}})$  qui appartient à  $D$ . L'existence d'une solution garantit la surjectivité de  $\psi$  tandis que l'unicité garantit l'injectivité de  $\psi$ . On peut conclure par le théorème d'inversion globale que  $\psi$  est un  $C^1$  difféomorphisme de  $D$  sur  $\Omega$ .

Dans les deux exercices qui suivent, on utilise le théorème d'inversion globale de la manière suivante. On se donne un difféomorphisme  $\psi$  d'un ouvert  $D \subset \mathbb{R}^n$  sur un ouvert  $\Omega \subset \mathbb{R}^n$ . On note  $C^1(D)$  l'ensemble des applications de classe  $C^1$  sur  $D$  à valeurs dans  $\mathbb{R}$ , et de même pour  $C^1(\Omega)$ . Alors l'application

$$f \in C^1(\Omega) \mapsto f \circ \psi \in C^1(D)$$

est bijective, d'inverse

$$g \in C^1(D) \mapsto g \circ \psi^{-1} \in C^1(\Omega).$$

*Exercice 7.12.* Résoudre l'équation aux dérivées partielles suivante :

$$\partial_1 f(x, y) - \partial_2 f(x, y) = 0$$

On pourra faire le changement de coordonnées :  $u = x + y, v = x - y$ .

*Exercice 7.13.* Soient  $U = \{(x, y) \in \mathbb{R}^2 : x < y\}$  et

$$\varphi : (x, y) \in \mathbb{R}^2 \mapsto (x^2 - 2xy - y^2, y)$$

1. Montrer que  $\varphi$  est un  $C^1$  difféomorphisme de  $U$  sur  $V = \{(u, v) \in \mathbb{R}^2 : u + 2v^2 > 0\}$  (pour la surjectivité, on pourra dresser le tableau de variations de la fonction  $x \mapsto x^2 - 2vx - v^2 - u$ ).
2. Résoudre l'équation aux dérivées partielles suivantes d'inconnue  $f : U \rightarrow \mathbb{R}$  :

$$(x + y)\partial_1 f(x, y) + (x - y)\partial_2 f(x, y) = 0.$$

## 7.2.2 Formule de Taylor-Young et extrema

On commence par présenter une version de la formule de Taylor-Young pour les fonctions de plusieurs variables.

**Théorème 7.14** (Formule de Taylor-Young). Soient  $f : D \rightarrow \mathbb{R}$  de classe  $C^2$  et  $a$  un point dans l'intérieur de  $D$ . Alors il existe  $\delta : D \rightarrow \mathbb{R}$  tel que  $\lim_{x \rightarrow a} \delta(x) = 0$  et tel que pour tout  $x \in D$ ,

$$f(x) = f(a) + Df(a)[x - a] + \frac{1}{2} \sum_{1 \leq i, j \leq n} \partial_{ij}^2 f(a)(x_i - a_i)(x_j - a_j) + \|x - a\|^2 \delta(x).$$

On pourra trouver une preuve de la formule de Taylor-Young dans la section *Approfondir* de ce chapitre.

**Exercice 7.15.** 1. Dessiner l'allure approchée des courbes de niveau de la fonction  $f : (x, y) \mapsto x^2y + 3yx^4 - 2xy^3$  au voisinage du point  $(1, 1)$ .

2. Même question pour la fonction

$$f : (x, y) \mapsto x^4 + y^4 - 4x^3 - 4y^3 + 2x^2y^2 - 4xy^2 - 4yx^2 + 8x^2 + 8y^2 + 8xy - 8x - 8y + 4$$

au voisinage du point  $(1, 1)$ .

Dans toute la suite de cette section,  $D$  est un ouvert de  $\mathbb{R}^2$  et  $g : D \rightarrow \mathbb{R}$  une fonction de classe  $C^1$ .

**Définition 7.16.** 1. On dit que  $g$  admet un maximum local en  $(x_0, y_0)$  s'il existe  $r > 0$  tel que

$$\forall (x, y) \in B((x_0, y_0), r), \quad g(x, y) \leq g(x_0, y_0),$$

2. minimum local :  $\geq$

3. maximum local strict :  $<$

4. minimum local strict :  $>$

On remarquera que cette définition ne dépend pas de la norme choisie pour définir les boules  $B((x_0, y_0), r)$ .

**Définition 7.17.** 1.  $g$  admet un maximum global en  $(x_0, y_0)$  si pour tout  $(x, y) \in D$ ,

$$g(x, y) \leq g(x_0, y_0),$$

2.  $g$  admet un minimum global en  $(x_0, y_0)$  si pour tout  $(x, y) \in D$ ,

$$g(x, y) \geq g(x_0, y_0).$$

Ces définitions généralisent les notions de maximum et minimum pour les fonctions d'une seule variable. Quelques rappels dans ce cadre : soit  $f : \mathbb{R} \rightarrow \mathbb{R}$  une fonction au moins  $C^2$ . Si  $x_0$  est un minimum local de  $f$ , alors  $f'(x_0) = 0$ . On peut le voir par exemple en considérant les taux d'accroissements : pour tout  $h > 0$  assez petit,

$$\frac{f(x_0 + h) - f(x_0)}{h} \geq 0.$$

On en déduit en passant à la limite pour  $h \rightarrow 0^+$  que  $f'(x_0) \geq 0$ . Maintenant, pour tout  $h < 0$ ,

$$\frac{f(x_0 + h) - f(x_0)}{h} \leq 0.$$

En passant à la limite pour  $h \rightarrow 0^-$ , on trouve cette fois  $f'(x_0) \leq 0$ . Finalement,  $f'(x_0) = 0$ .

On peut également obtenir une information sur la dérivée seconde. Dans cet objectif, on écrit la formule de Taylor-Young : il existe  $r > 0$  et  $\delta : ]x_0 - r, x_0 + r[ \rightarrow \mathbb{R}$  une fonction qui tend vers 0 en  $x_0$  tels que

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + \frac{1}{2}(x - x_0)^2 f''(x_0) + (x - x_0)^2 \delta(x) \quad , \quad x \in ]x_0 - r, x_0 + r[.$$

Comme  $f'(x_0) = 0$ , il vient

$$f(x) - f(x_0) = \frac{1}{2}(x - x_0)^2 f''(x_0) + (x - x_0)^2 \delta(x).$$

Comme  $x_0$  est un minimum local, on en déduit que pour tout  $x$  proche de  $x_0$ ,

$$\frac{1}{2}(x - x_0)^2 f''(x_0) + (x - x_0)^2 \delta(x) \geq 0$$

et donc en simplifiant par  $(x - x_0)^2$ ,

$$f''(x_0) + 2\delta(x) \geq 0.$$

En faisant tendre  $x$  vers  $x_0$ , on obtient finalement que  $f''(x_0) \geq 0$ . Pour résumer, on a montré que si  $x_0$  est un minimum local de  $f$ , alors nécessairement  $f'(x_0) = 0$  et  $f''(x_0) \geq 0$ . Ces conditions  $f'(x_0) = 0$  et  $f''(x_0) \geq 0$  sont *presque* suffisantes pour que  $x_0$  soit un minimum local de  $f$ .

Plus précisément, si  $x_0 \in \mathbb{R}$  est tel que  $f'(x_0) = 0$  et  $f''(x_0) > 0$ , alors  $x_0$  est un minimum local (on pourra aisément deviner un énoncé analogue pour les maxima locaux). La preuve repose sur la formule de Taylor-Young (pour les fonctions d'une variable) : il existe  $r > 0$  et  $\delta : ]-r, r[ \rightarrow \mathbb{R}$  une fonction qui tend vers 0 en  $x_0$  tels que

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + \frac{1}{2}(x - x_0)^2 f''(x_0) + (x - x_0)^2 \delta(x) \quad , \quad x \in ]x_0 - r, x_0 + r[.$$

Comme  $f'(x_0) = 0$ , on en déduit

$$f(x) = f(x_0) + \frac{1}{2}(x - x_0)^2 f''(x_0) + (x - x_0)^2 \delta(x) \quad , \quad x \in ]x_0 - r, x_0 + r[.$$

Quitte à réduire  $r > 0$ , on peut supposer que  $|\delta(x)| \leq \frac{1}{4}f''(x_0)$  lorsque  $x \in ]x_0 - r, x_0 + r[$  (ici, on utilise l'hypothèse que  $\lim_{x \rightarrow x_0} \delta(x) = 0$ ). Alors pour tout  $x \in ]x_0 - r, x_0 + r[$ ,

$$f(x) \geq f(x_0) + \frac{1}{2}(x - x_0)^2 f''(x_0) - \frac{1}{4}f''(x_0)(x - x_0)^2 = f(x_0) + \frac{1}{4}(x - x_0)^2 f''(x_0) \geq f(x_0)$$

et donc  $x_0$  est un minimum local.

Ces conditions nécessaires ou suffisantes d'extrémalité peuvent se généraliser aux fonctions de plusieurs variables. On se limite dans la suite à deux variables.

**Proposition 7.18.** *Si  $g$  présente un minimum ou un maximum local en  $(x_0, y_0)$  et si  $g$  est  $C^1$ , alors*

$$\partial_1 g(x_0, y_0) = 0 = \partial_2 g(x_0, y_0).$$

**Définition 7.19.** *Si  $Dg(x_0, y_0) = 0$ , on dit que  $(x_0, y_0)$  est un point critique.*

Preuve : Les fonctions  $x \mapsto g(x, y_0)$  et  $y \mapsto g(x_0, y)$  admettent toutes les deux un maximum ou un minimum local en  $x_0$  et  $y_0$  respectivement. Donc

$$\partial_1 g(x_0, y_0) = 0 = \partial_2 g(x_0, y_0).$$

□

*Remarque 7.20.* Les extremums sont à chercher parmi les points critiques, mais tous les points critiques ne sont pas forcément des extremums.

On introduit les notations :

$$r = \partial_{11}^2 g(x_0, y_0), \quad s = \partial_{12}^2 g(x_0, y_0), \quad t = \partial_{22}^2 g(x_0, y_0).$$

Pour pouvoir déterminer si un point critique est un extremum ou pas, on a besoin des dérivées partielles d'ordre 2, et plus précisément de la formule de Taylor-Young d'ordre 2. Au voisinage d'un point critique  $(x_0, y_0)$  (où les dérivés partielles s'annulent), la formule de Taylor-Young s'écrit :

$$g(x_0 + h_1, y_0 + h_2) = g(x_0, y_0) + \frac{1}{2} (r^2 h_1^2 + t^2 h_2^2 + 2s^2 h_1 h_2) + \|(h_1, h_2)\|_2^2 \delta(h_1, h_2),$$

avec  $\lim_{(h_1, h_2) \rightarrow 0} \delta(h_1, h_2) = 0$ . En utilisant la notation matricielle,

$$g(x_0 + h_1, y_0 + h_2) - g(x_0, y_0) = \frac{1}{2} \begin{pmatrix} h_1 & h_2 \end{pmatrix} \begin{pmatrix} r & s \\ s & t \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} + \|(h_1, h_2)\|_2^2 \delta(h_1, h_2). \quad (7.1)$$

Supposons par exemple que  $(x_0, y_0)$  soit un minimum local. Alors pour tout  $(h_1, h_2)$ ,  $g(x_0 + h_1, y_0 + h_2) - g(x_0, y_0) \geq 0$ , ce qui d'après la formule précédente implique :

$$\begin{pmatrix} h_1 & h_2 \end{pmatrix} \begin{pmatrix} r & s \\ s & t \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} + 2\|(h_1, h_2)\|_2^2 \delta(h_1, h_2) \geq 0.$$

On aimerait se débarrasser du reste. Pour cela, on peut procéder ainsi. On se donne un vecteur  $z = (z_1, z_2)$  et on applique ce qui précède au vecteur  $h = \lambda z = (\lambda z_1, \lambda z_2)$ , où  $\lambda$  est un nombre destiné à tendre vers 0 :

$$\lambda^2 \begin{pmatrix} z_1 & z_2 \end{pmatrix} \begin{pmatrix} r & s \\ s & t \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} + 2\lambda^2 \|(z_1, z_2)\|_2^2 \delta(\lambda z) \geq 0.$$

On divise par  $\lambda^2$  et on fait tendre  $\lambda$  vers 0. Comme  $\lim_{\lambda \rightarrow 0} \delta(\lambda z) = 0$ , il vient

$$\begin{pmatrix} z_1 & z_2 \end{pmatrix} \begin{pmatrix} r & s \\ s & t \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \geq 0.$$

Comme c'est vrai pour tout  $(z_1, z_2)$ , on en déduit que la *matrice hessienne de g en  $(x_0, y_0)$*  définie par

$$\nabla^2 g(x_0, y_0) = \begin{pmatrix} r & s \\ s & t \end{pmatrix}$$

est positive. Comme elle est symétrique, ses valeurs propres sont positives. On a donc montré que si  $(x_0, y_0)$  est un minimum, alors la hessienne en  $(x_0, y_0)$  est une matrice positive. La réciproque est *presque vraie*.

**Théorème 7.21.** Soient  $g : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}$  une fonction de classe  $C^2$ ,  $(x_0, y_0) \in D$  un point critique de  $g$ . On note  $r = \partial_{11}^2 g(x_0, y_0)$ ,  $s = \partial_{12}^2 g(x_0, y_0)$ ,  $t = \partial_{22}^2 g(x_0, y_0)$ .

1.  $(x_0, y_0)$  est un minimum local strict si  $rt - s^2 > 0$  et  $r > 0$ ,
2.  $(x_0, y_0)$  est un maximum local strict si  $rt - s^2 > 0$  et  $r < 0$ .

Dans le cas où  $rt - s^2 < 0$ , on dit que  $(x_0, y_0)$  est un point selle. En notant  $\lambda_1, \lambda_2$  les valeurs propres de  $\nabla^2 g(x_0, y_0)$ , alors le signe du déterminant  $rt - s^2 = \lambda_1 \lambda_2$  montre que  $\lambda_1$  et  $\lambda_2$  sont de signe opposé. Supposons par exemple  $\lambda_1 < 0 < \lambda_2$ . Notons  $\varepsilon_1$  un vecteur propre associé à  $\lambda_1$  et  $\varepsilon_2$  un vecteur propre associé à  $\lambda_2$ . Alors on peut montrer que 0 est un minimum local de  $t \mapsto f((x_0, y_0) + t\varepsilon_1)$  et un maximum local de  $t \mapsto f((x_0, y_0) + t\varepsilon_2)$ .

*Exemple 7.22.* Le point  $(0, 0)$  est un point critique de la fonction  $(x, y) \mapsto x^2 - y^2$ . Mais ce n'est ni un maximum local ni un minimum local.

*Méthode 7.23.* Pour étudier les extrema locaux d'une fonction  $f : U \rightarrow \mathbb{R}$  de classe  $C^2$  sur un ouvert  $U \subset \mathbb{R}^n$ ,

1. on calcule les points critiques de  $f$ , c'est-à-dire les  $x \in U$  tels que  $\nabla f(x) = 0$ ,
2. on calcule la hessienne de  $f$  en chaque point critique,
3. on détermine s'il s'agit d'extrema (minima ou maxima locaux) en calculant les valeurs propres de ces matrices hessiennes.

*Exercice 7.24.* Chercher les extrema locaux des fonctions suivantes sur leur domaine de définition

1.  $(x, y) \mapsto \cos x \sin y$ ,
2.  $(x, y) \mapsto xy$ ,
3.  $(x, y) \mapsto x^2 - 2y^2 - 2x + 1$ .
4.  $(x, y) \mapsto x^3 + \frac{4}{3}y^3 + 3y^2 - 3x + 1$ .

## 7.3 Approfondir

### 7.3.1 Preuve de la formule de Taylor-Young : Théorème 7.14

On travaille avec la norme  $\|\cdot\|_1$  sur  $\mathbb{R}^n$  (changer de norme revient à changer de fonction  $\delta$ , mais pas le fait que cette fonction tende vers 0 en  $a$ ). On introduit la fonction

$$\varphi(x) := f(x) - \sum_{i=1}^n \partial_i f(a)(x_i - a_i) - \frac{1}{2} \sum_{1 \leq i, j \leq n} \partial_{ij}^2 f(a)(x_i - a_i)(x_j - a_j).$$

Posons pour tout  $x \in D$ ,

$$\delta(x) = \begin{cases} \frac{1}{\|x-a\|_1^2} \varphi(x) & \text{si } x \neq a, \\ 0 & \text{si } x = a. \end{cases}$$

Il s'agit de montrer que  $\lim_{x \rightarrow a} \delta(x) = 0$ . Pour tout  $k \in \llbracket 1, n \rrbracket$ ,

$$\begin{aligned} \partial_k \varphi(x) &= \partial_k f(x) - \partial_k f(a) - \frac{1}{2} \sum_{i=1}^n \partial_{ik}^2 f(a)(x_i - a_i) - \frac{1}{2} \sum_{j=1}^n \partial_{kj}^2 f(a)(x_j - a_j) \\ &= \partial_k f(x) - \partial_k f(a) - \sum_{i=1}^n \partial_i(\partial_k f)(a)(x_i - a_i). \end{aligned}$$

Soit  $\varepsilon > 0$ . Comme  $\partial_k f$  est  $C^1$  pour tout  $k \in \llbracket 1, n \rrbracket$ , il existe  $\eta > 0$  tel que pour tout  $x \in B(a, \eta)$ , pour tout  $k \in \llbracket 1, n \rrbracket$ ,  $|\partial_k \varphi(x)| \leq \varepsilon \|x - a\|_1$ . Cela implique que pour tout  $y \in B(a, \eta)$ , pour tout  $e = (e_1, \dots, e_n)$ ,

$$|D\varphi(y)[e]| = \left| \sum_{k=1}^n \partial_k \varphi(y) e_k \right| \leq \sum_{k=1}^n |\partial_k \varphi(y)| |e_k| \leq \varepsilon \|y - a\|_1 \sum_{k=1}^n |e_k| = \varepsilon \|y - a\|_1 \|e\|_1$$

et donc

$$\|D\varphi(y)\|_1 = \sup_{\|e\|_1=1} |D\varphi(y)[e]| \leq \varepsilon \|y - a\|_1.$$

En particulier, pour tout  $y$  dans le segment  $[a, x] = \{ta + (1-t)x : t \in [0, 1]\}$ , on a

$$\|D\varphi(y)\|_1 \leq \varepsilon \|y - a\|_1 \leq \varepsilon \|x - a\|_1.$$

Alors, par le théorème des accroissements finis (Théorème 6.40), on en déduit que pour tout  $x \in B(a, \eta)$ ,

$$|\varphi(x) - \varphi(a)| \leq \varepsilon \|x - a\|_1^2,$$

ce qui se réécrit

$$\left| f(x) - f(a) - Df(a)[x - a] - \frac{1}{2} \sum_{1 \leq i, j \leq n} \partial_{ij}^2 f(a)(x_i - a_i)(x_j - a_j) \right| \leq \varepsilon \|x - a\|_1^2.$$

Cela montre bien que la fonction

$$\delta(x) = \frac{1}{\|x - a\|_1^2} \left( f(x) - f(a) - Df(a)[x - a] - \frac{1}{2} \sum_{1 \leq i, j \leq n} \partial_{ij}^2 f(a)(x_i - a_i)(x_j - a_j) \right), \quad x \neq a,$$

tend vers 0 en  $a$ .

□

### 7.3.2 Preuve du Théorème de Monge : Théorème 7.21

Avant de donner la preuve de l'énoncé proprement dit, on aura besoin du résultat suivant en algèbre linéaire.

**Lemme 7.25.** *Si  $A \in \mathcal{M}_2(\mathbb{R})$  est une matrice symétrique définie positive, alors il existe  $\alpha > 0$  tel que pour tout  $h \in \mathbb{R}^2$ ,*

$$h^T A h \geq \alpha \|h\|_2^2.$$

Preuve du lemme : Noter que  $Ah$  est un vecteur (colonne) de  $\mathbb{R}^n$  et  $h^T Ah$  est un vecteur colonne de  $\mathbb{R}$ , autrement dit un nombre réel !

Comme  $A$  est symétrique, il existe une matrice orthogonale  $P$  et une matrice diagonale  $D$  telle que  $A = P^T D P$ . Les éléments diagonaux de  $D$  sont les valeurs propres  $\lambda_1$  et  $\lambda_2$  de  $A$  :  $D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ .

Alors

$$h^T A h = h^T P^T D P h = k^T D k$$

en posant  $k = P h$ . En notant  $k = \begin{pmatrix} k_1 \\ k_2 \end{pmatrix}$ , il vient

$$h^T A h = \lambda_1 k_1^2 + \lambda_2 k_2^2 \geq \min(\lambda_1, \lambda_2)(k_1^2 + k_2^2) = \min(\lambda_1, \lambda_2) \|k\|_2^2.$$

Comme  $P$  est orthogonale,  $\|k\|_2 = \|h\|_2$ . Ainsi,

$$h^T A h \geq \min(\lambda_1, \lambda_2) \|h\|_2^2.$$

On pose  $\alpha = \min(\lambda_1, \lambda_2)$ . Comme  $A$  est définie positive,  $\lambda_1 > 0$  et  $\lambda_2 > 0$ , donc  $\alpha > 0$ . Cela permet de conclure la preuve du lemme.

□

On peut alors se lancer dans la preuve du théorème :

Preuve : On prouve le théorème dans le cas  $r > 0$ . Le cas  $r < 0$  se déduit du précédent en considérant  $-f$  au lieu de  $f$ .

La matrice

$$A = \begin{pmatrix} r & s \\ s & t \end{pmatrix}$$

est symétrique, donc diagonalisable dans une base orthonormée. Notons  $\lambda_1$  et  $\lambda_2$  ses valeurs propres. Le déterminant de  $A$  qui vaut  $rt - s^2 > 0$  est le produit des valeurs propres  $\lambda_1\lambda_2$ . On en déduit que  $\lambda_1$  et  $\lambda_2$  sont de même signe. Comme  $r > 0$  et  $rt - s^2 > 0$ , alors  $t > s^2/r > 0$ . La trace de  $A$ , qui est la somme des valeurs propres  $\lambda_1 + \lambda_2$ , vaut  $r + t > 0$ . On en déduit que  $\lambda_1$  et  $\lambda_2$  sont toutes les deux  $> 0$ . Autrement dit,  $A$  est une matrice symétrique définie positive.

D'après le lemme, il existe  $\alpha > 0$  tel que pour tout  $h \in \mathbb{R}^2$ ,

$$h^T Ah \geq \alpha \|h\|_2^2.$$

Par la formule (7.1), on obtient :

$$\begin{aligned} g(x_0 + h_1, y_0 + h_2) &= g(x_0, y_0) + \frac{1}{2} h^T Ah + \|(h_1, h_2)\|_2^2 \delta(h_1, h_2) \\ &\geq g(x_0, y_0) + \frac{\alpha}{2} \|(h_1, h_2)\|_2^2 + \|(h_1, h_2)\|_2^2 \delta(h_1, h_2) \\ &\geq g(x_0, y_0) + \left(\frac{\alpha}{2} + \delta(h_1, h_2)\right) \|(h_1, h_2)\|_2^2. \end{aligned}$$

Comme  $\lim_{(h_1, h_2) \rightarrow (0,0)} \delta(h_1, h_2) = 0$ , il existe  $r > 0$  tel que pour tout  $(h_1, h_2) \in B(0, r)$ ,  $|\delta(h_1, h_2)| \leq \frac{\alpha}{2}$ . Alors pour de tels vecteurs  $(h_1, h_2)$ ,

$$g(x_0 + h_1, y_0 + h_2) \geq g(x_0, y_0),$$

ce qui montre bien que  $(x_0, y_0)$  est un minimum local de  $g$ . □

**Exercice 7.26.** \*\* Soit  $f : \mathbb{R} \rightarrow \mathbb{R}$  une fonction de classe  $C^1$ . On suppose qu'il existe  $\kappa \in ]0, 1[$  tel que  $\forall t \in \mathbb{R}, |f'(t)| \leq \kappa$ . On définit alors la fonction :

$$F : (x, y) \in \mathbb{R}^2 \mapsto (x + f(y), y + f(x)).$$

1. Montrer que  $|f(t) - f(s)| \leq \kappa|t - s|$ , et en déduire que  $F$  est injective sur  $\mathbb{R}^2$ .
2. Montrer que  $F(\mathbb{R}^2)$  est un ouvert.
3. L'objet de cette question est en fait de montrer que  $F(\mathbb{R}^2) = \mathbb{R}^2$ .
  - (a) Montrer que  $|f(t)| \leq |f(0)| + \kappa|t|$ . En déduire que  $\|F(x, y)\|_1 \geq (1 - \kappa)\|(x, y)\|_1 - 2|f(0)|$  puis que  $\lim_{\|(x, y)\| \rightarrow +\infty} \|F(x, y)\|_2 = +\infty$ .
  - (b) Soit  $(x_0, y_0) \in \mathbb{R}^2$ . On définit la fonction

$$G : (x, y) \mapsto \|F(x, y) - (x_0, y_0)\|_2.$$

Montrer que  $\lim_{\|(x, y)\| \rightarrow +\infty} G(x, y) = +\infty$ . En déduire que  $G$  admet un minimum global sur  $\mathbb{R}^2$  (on pourra utiliser une suite  $(x_n, y_n)_{n \in \mathbb{N}}$  telle que

$$\lim_{n \rightarrow +\infty} G(x_n, y_n) = \inf_{(x, y) \in \mathbb{R}^2} G(x, y)$$

et montrer qu'elle est bornée, avant d'en extraire une sous-suite convergente).

- (c) En déduire que la fonction  $G^2$  admet un minimum global sur  $\mathbb{R}^2$ . On note  $(x_*, y_*)$  un point où  $G^2$  admet son minimum.
- (d) Conclure en calculant  $\nabla G^2(x_*, y_*)$ .

**Exercice 7.27.** \*\* Soient  $n \in \mathbb{N}^*$ ,  $\mathbb{R}^n$  muni du produit scalaire usuel  $\langle \cdot, \cdot \rangle$  et de la norme euclidienne  $\|\cdot\|_2$ ,  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  de classe  $C^1$ . On suppose qu'il existe  $\kappa > 0$  tel que pour tout  $x, y \in \mathbb{R}^n$ ,

$$\langle F(x) - F(y), x - y \rangle \geq \kappa \|x - y\|^2. \quad (7.2)$$

1. Justifier que  $F$  est injective.
2. Montrer que pour tout  $e \in \mathbb{R}^n$ , pour tout  $y \in \mathbb{R}^n$ ,  $\langle DF(y)[e], e \rangle \geq \kappa \|e\|^2$  (on pourra prendre  $x = y + te$  dans (7.2) puis faire tendre  $t$  vers 0).
3. Etablir que  $F$  est un  $\mathcal{C}^1$  difféomorphisme de  $\mathbb{R}^n$  sur  $F(\mathbb{R}^n)$ .
4. Montrer que  $\lim_{\|x\|_2 \rightarrow +\infty} \|F(x)\|_2 = +\infty$ .
5. Soit  $a \in \mathbb{R}^n$ . Montrer que  $\lim_{\|x\|_2 \rightarrow +\infty} \|F(x) - a\|_2 = +\infty$ .
6. En déduire que la fonction  $g : x \mapsto \|F(x) - a\|_2^2$  admet un minimum global sur  $\mathbb{R}^n$ . En calculant son gradient en ce point de minimum, montrer qu'il existe  $x_a \in \mathbb{R}^n$  tel que  $F(x_a) = a$ .
7. Conclure que  $F(\mathbb{R}^n) = \mathbb{R}^n$ .

# Chapitre 8

## Probabilités discrètes

### 8.1 Événements aléatoires

#### 8.1.1 Issues et événements

L'ensemble des issues possibles d'une expérience aléatoire est appelé l'univers, et souvent noté  $\Omega$ . Les issues, c'est-à-dire les éléments de  $\Omega$ , seront souvent désignées par la lettre  $\omega$ .

*Exemple 8.1.* On peut modéliser

1. un lancer de dé en prenant  $\Omega = \{1, 2, \dots, 6\}$ ,
2. deux lancers successifs d'un même dé par  $\Omega = \{1, 2, \dots, 6\}^2$ . Une issue est donc un couple  $(i, j)$  où  $i$  désigne le résultat du premier lancer, et  $j$  le résultat du deuxième lancer,
3. un lancer de deux dés, l'un rouge, l'autre vert, avec  $\Omega = \{1, 2, \dots, 6\}^2$ . Une issue est un couple de nombres, le premier indiquant le résultat du dé rouge, le second indiquant le résultat du dé vert,
4. un lancer de deux dés identiques par  $\Omega := \{\{i, j\} : 1 \leq i, j \leq 6\}$ . Les issues ne sont plus des couples, mais des ensembles qui sont soit des singletons :  $\{1\}, \dots, \{6\}$  (c'est le cas si les deux dés indiquent le même résultat), soit des ensembles à deux éléments :  $\{i, j\}$  avec  $i \neq j$ .
5. On lance un dé une infinité de fois. Une issue sera alors la suite des résultats obtenus, c'est-à-dire une suite numérique compris entre 1 et 6. L'univers est alors  $\{1, \dots, 6\}^{\mathbb{N}}$ .
6. On lance une pièce de monnaie et on s'arrête dès qu'on obtient un pile. Chaque issue correspondra à un nombre de lancers supérieur ou égal à 1. Dans ce cas l'univers s'identifie à  $\mathbb{N}^* \cup \{\infty\}$  (on rajoute  $\{\infty\}$  pour l'issue où on n'obtient jamais pile).

On note  $\mathcal{P}(\Omega)$  l'ensemble des parties de  $\Omega$ . Les *éléments* de  $\mathcal{P}(\Omega)$  sont donc des *parties* de  $\Omega$ . On considère un sous-ensemble  $\mathcal{E}$  de  $\mathcal{P}(\Omega)$  qu'on appelle l'ensemble des *événements* de l'expérience aléatoire. Un événement est donc une partie de  $\Omega$ , autrement dit un ensemble d'issues de l'expérience aléatoire.

*Exemple 8.2.* 1. Lors du lancer d'un dé, on pourra considérer l'événement *le nombre obtenu est 1*, ou bien *le nombre obtenu est pair*.

2. Lors du lancer répété d'une pièce de monnaie qu'on arrête dès qu'on obtient pile, on pourra considérer l'événement *on obtient pile avant le cinquième lancer*.

On choisira toujours l'ensemble des événements  $\mathcal{E}$  de sorte que si  $A$  et  $B$  sont deux événements, alors l'ensemble

- $A \cup B = \{\omega \in \Omega : \omega \in A \text{ ou } \omega \in B\}$  est un événement, appelé *événement A ou B*,
- $A \cap B = \{\omega \in \Omega : \omega \in A \text{ et } \omega \in B\}$  est un événement, appelé *événement A et B*,
- $\Omega \setminus A = \{\omega \in \Omega : \omega \notin A\}$ , noté encore  $A^c$ , est un événement, appelé *événement contraire de A*.

On supposera aussi que l'ensemble vide, noté  $\emptyset$ , qui est la partie de  $\Omega$  qui ne contient aucun élément, est un événement. On l'appelle l'événement impossible.

On peut considérer des suites d'événements : pour chaque entier  $n \in \mathbb{N}$ , on se donne un événement  $A_n$ . On note  $(A_n)_{n \in \mathbb{N}}$  cette suite d'événements.

**Définition 8.3.** La réunion d'une suite d'événements  $(A_n)_{n \in \mathbb{N}}$  est définie par

$$\cup_{n \in \mathbb{N}} A_n = \{\omega \in \Omega : \text{il existe } n \in \mathbb{N} \text{ tel que } \omega \in A_n\}.$$

On dit qu'une suite d'événements est complète lorsque

1. les événements sont deux à deux disjoints :

$$\forall n, m \in \mathbb{N}, \quad A_n \cap A_m = \emptyset,$$

2. de réunion  $\Omega$  :

$$\cup_{n \in \mathbb{N}} A_n = \Omega.$$

On supposera toujours que la réunion d'une suite d'événements est encore un événement.

Enfin, on dit que deux événements  $A$  et  $B$  sont incompatibles si  $A \cap B = \emptyset$ .

**Exemple 8.4.** 1. Dans le cas du dé, l'événement *le nombre obtenu est divisible par 7* est impossible. Les événements *le nombre obtenu est pair* et *le nombre obtenu est impair* sont incompatibles.

2. Dans le cas du lancer de pièce, pour chaque  $n \in \mathbb{N}^*$ , on peut définir l'événement  $A_n$  comme *on obtient pile avant le lancer n*. Alors les événements  $A_n$  ne sont pas deux à deux disjoints. Si on considère l'événement  $B_n$  défini par *on obtient pile au lancer n*, alors les  $B_n$  sont deux à deux disjoints.

3. Toujours dans le cas du lancer de pièce, on note  $C_n$  l'événement *on obtient pile après le lancer  $2^{n-1} - 1$  et avant le lancer  $2^n$* , pour tout  $n \in \mathbb{N}^*$ , et  $C_0$  l'événement *on n'obtient jamais pile*, alors la suite d'événements  $(C_n)_{n \in \mathbb{N}}$  est complète.

## 8.1.2 Probabilités

Une mesure de probabilité vise à donner une idée quantitative de l'importance de chaque événement. Elle traduit par un nombre la chance que l'événement a de se réaliser.

**Définition 8.5.** Une mesure de probabilité  $\mathbb{P}$  sur  $\Omega$  est une application de l'ensemble des événements  $\mathcal{E}$  à valeurs dans  $[0, 1]$  telle que

1.  $\mathbb{P}(\Omega) = 1, \mathbb{P}(\emptyset) = 0$ ,
2. si  $A$  et  $B$  sont deux événements tels que  $A \subset B$ , alors  $\mathbb{P}(A) \leq \mathbb{P}(B)$ ,
3. si  $A$  et  $B$  sont deux événements tels que  $A \cap B = \emptyset$ , alors  $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ ,
4. si  $(A_n)_{n \in \mathbb{N}}$  est une suite d'événements deux à deux disjoints, alors

$$\mathbb{P}(\cup_{n \in \mathbb{N}} A_n) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n).$$

Le dernier axiome signifie que la série à termes positifs  $\sum_n \mathbb{P}(A_n)$  converge vers le nombre  $\mathbb{P}(\cup_{n \in \mathbb{N}} A_n)$ .

**Remarque 8.6.** En fait, on peut montrer que les axiomes 1 et 4 dans la définition 8.5 impliquent les axiomes 2 et 3.

**Proposition 8.7.** Si  $\mathbb{P}$  est une mesure de probabilité, et  $A, B \subset \Omega$ ,

1.  $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$ ,
2.  $\mathbb{P}(B) = \mathbb{P}(A) - \mathbb{P}(A \setminus B)$  si  $B \subset A$ ,
3.  $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ .

Preuve : En écrivant  $\Omega = A \cup A^c$ , on obtient

$$1 = \mathbb{P}(\Omega) = \mathbb{P}(A \cup A^c) = \mathbb{P}(A) + \mathbb{P}(A^c)$$

d'où la première relation. En écrivant  $A = B \cup (A \setminus B)$ , on obtient la deuxième égalité. Enfin, en observant que  $A \cup B = A \cup (B \setminus (A \cap B))$ , la troisième égalité découle de la deuxième.  $\square$

*Exemple 8.8.* 1. Si  $\Omega$  est un univers fini (comme pour le lancer de dé), on définit la mesure de probabilité uniforme sur  $\Omega$  par

$$\forall A \subset \Omega, \quad \mathbb{P}(A) = \frac{|A|}{|\Omega|}.$$

Ici,  $|A|$  désigne le cardinal de  $A$

2. Si  $\Omega = \{0, 1\}$ , on définit la mesure de Bernoulli de paramètre  $p \in [0, 1]$  par

$$\mathbb{P}(\{1\}) = p, \quad \mathbb{P}(\{0\}) = 1 - p.$$

3. Si  $\Omega$  est  $\{0, \dots, N\}$ , la mesure binomiale est définie par :

$$\forall k \in \{0, \dots, N\}, \quad \mathbb{P}(\{k\}) = C_N^k p^k (1-p)^{N-k}.$$

**Définition 8.9.** Soient  $A$  et  $B$  deux événements. On suppose que  $\mathbb{P}(B) > 0$ . La probabilité conditionnelle de  $A$  sachant  $B$  est

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

**Lemme 8.10.** Soit  $\{A_n\}_{n \in \mathbb{N}}$  une suite d'événements disjoints deux à deux. Alors pour tout événement  $B \subset \Omega$ , la série  $\sum_n \mathbb{P}(A_n \cap B)$  converge et

$$\sum_{n=0}^{+\infty} \mathbb{P}(A_n \cap B) = \mathbb{P}((\cup_{n \in \mathbb{N}} A_n) \cap B).$$

On pourra admettre la preuve.

Preuve : Observer que

$$(\cup_{n \in \mathbb{N}} A_n) \cap B = \cup_{n \in \mathbb{N}} (A_n \cap B)$$

(appartenir à l'ensemble de gauche signifie qu'on appartient à  $B$  et à l'un des  $A_n$ , tandis qu'appartenir à l'ensemble de droite signifie qu'il existe  $n$  pour lequel on appartient à  $B$  et à  $A_n$  : il s'agit donc du même ensemble!). De plus, les  $A_n$  étant deux à deux disjoints, il en est de même des  $A_n \cap B$  :

$$(A_n \cap B) \cap (A_m \cap B) = A_n \cap A_m \cap B = (A_n \cap A_m) \cap B = \emptyset \cap B = \emptyset.$$

On en déduit que la série  $\sum_n \mathbb{P}(A_n \cap B)$  converge et sa limite est  $\mathbb{P}(\cup_{n \in \mathbb{N}} (A_n \cap B))$ . Finalement,

$$\mathbb{P}((\cup_{n \in \mathbb{N}} A_n) \cap B) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n \cap B).$$

$\square$

Ce lemme a deux conséquences. La première est le fait que les probabilités conditionnelles sont des mesures de probabilité :

**Proposition 8.11.** On suppose que  $\mathbb{P}(B) > 0$ . L'application  $P(\cdot|B) : \mathcal{E} \rightarrow \mathbb{R}$  est une mesure de probabilités sur  $\Omega$ .

On pourra admettre la preuve.

Preuve : Par définition de la probabilité conditionnelle,

$$\mathbb{P}(\emptyset|B) = \frac{\mathbb{P}(\emptyset \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\emptyset)}{\mathbb{P}(B)} = 0, \mathbb{P}(\Omega|B) = \frac{\mathbb{P}(\Omega \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1.$$

De plus, si  $(A_n)_{n \in \mathbb{N}}$  est une suite d'événements deux à deux disjoints,

$$\mathbb{P}(\cup_{n \in \mathbb{N}} A_n | B) = \frac{\mathbb{P}((\cup_{n \in \mathbb{N}} A_n) \cap B)}{\mathbb{P}(B)}.$$

D'après le lemme,

$$\mathbb{P}((\cup_{n \in \mathbb{N}} A_n) \cap B) = \sum_{n=0}^{\infty} \mathbb{P}(A_n \cap B)$$

d'où

$$\mathbb{P}(\cup_{n \in \mathbb{N}} A_n | B) = \frac{\sum_{n=0}^{\infty} \mathbb{P}(A_n \cap B)}{\mathbb{P}(B)} = \sum_{n=0}^{\infty} \frac{\mathbb{P}(A_n \cap B)}{\mathbb{P}(B)} = \sum_{n=0}^{\infty} P(A_n | B).$$

Conclusion :  $P(\cdot|B)$  est bien une probabilité. □

La deuxième conséquence du lemme est la formule des probabilités totales.

**Proposition 8.12** (Formule des probabilités totales). Si  $\{A_n\}_n$  est une suite complète d'événements, alors pour tout événement  $B \subset \Omega$  :

$$\mathbb{P}(B) = \sum_n \mathbb{P}(B|A_n)\mathbb{P}(A_n)$$

(avec la convention  $\mathbb{P}(B|A_n) \mathbb{P}(A_n) = 0$  si  $\mathbb{P}(A_n) = 0$ ).

Preuve : D'après le lemme,  $\sum_n \mathbb{P}(B \cap A_n)$  converge vers

$$\mathbb{P}(B \cap (\cup_n A_n)) = \mathbb{P}(B \cap \Omega) = \mathbb{P}(B).$$

On conclut en utilisant que  $\mathbb{P}(B|A_n)\mathbb{P}(A_n) = \mathbb{P}(B \cap A_n)$ . □

On en déduit aussitôt :

*Remarque 8.13* (Formule de Bayes). Sous l'hypothèse de la proposition précédente et si  $\mathbb{P}(B) > 0$ ,

$$\mathbb{P}(A_1|B) = \frac{\mathbb{P}(A_1 \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B|A_1)\mathbb{P}(A_1)}{\sum_n \mathbb{P}(B|A_n)\mathbb{P}(A_n)}.$$

Pour calculer des probabilités conditionnelles, on utilise souvent un arbre dont les branches portent les probabilités conditionnelles tandis que les noeuds portent les probabilités d'événements :

*Exemple 8.14.* Le dépistage d'une maladie conduit à un taux de détection de 90 pour 100. Le taux de faux positif est de 5 pour 100. Le taux de détection de 90 pour 100 signifie que si 100 individus malades passent ce test, en moyenne 90 seront détectés comme malade et 10 ne le seront pas. Parallèlement, le taux de faux positif de 5 pour 100 signifie que si 100 individus sains passent ce test, en moyenne 5 seront identifiés à tort comme malades. Sachant que la probabilité qu'un individu soit malade est environ de 1/650, calculer la probabilité qu'un individu soit sain sachant qu'il a été déclaré malade.

Modélisons : l'ensemble des issues est constitué par l'ensemble des individus. Notons  $A$  l'événement : l'individu est sain, et  $B$  l'événement : le test est positif. On traduit l'énoncé :  $\mathbb{P}(B|A^c) = 0.9$ ,  $\mathbb{P}(B|A) = 0.05$ ,  $\mathbb{P}(A^c) = \frac{1}{650}$ . Par la formule de Bayes, on en déduit

$$\begin{aligned}\mathbb{P}(A|B) &= \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B|A)\mathbb{P}(A) + \mathbb{P}(B|A^c)\mathbb{P}(A^c)} \\ &= \frac{0.05 \times \left(1 - \frac{1}{650}\right)}{0.05 \times \left(1 - \frac{1}{650}\right) + 0.9 \times \frac{1}{650}} = \frac{0.05 \times 649}{0.05 \times 649 + 0.9} = 0,97.\end{aligned}$$

## 8.2 Variables aléatoires discrètes

Dans toute cette section, on se donne un univers  $\Omega$  et un ensemble d'événements  $\mathcal{E}$  sur  $\Omega$ . On se donne de plus un ensemble  $\mathcal{X} \subset \mathbb{R}$  de la forme

$$\mathcal{X} = \{x_i, i \in I\}$$

où  $I$  est soit  $\mathbb{N}$  soit un ensemble fini. Dans cette écriture, on suppose implicitement que les  $x_i$  sont distincts (autrement dit, l'application  $i \in I \mapsto x_i \in \mathcal{X}$  est injective).

**Définition 8.15.** Une variable aléatoire sur  $\mathcal{X}$  est une fonction  $X : \Omega \rightarrow \mathcal{X}$  telle que pour tout  $x \in \mathcal{X}$ , l'ensemble

$$X^{-1}(\{x\}) = \{\omega \in \Omega : X(\omega) = x\}$$

est un événement.

Pour  $x \in \mathcal{X}$ , on notera souvent  $[X = x]$  l'ensemble  $X^{-1}(\{x\})$ . Si  $A \subset \mathcal{X}$ ,  $[X \in A]$  désignera l'événement  $X^{-1}(A) = \{\omega \in \Omega : X(\omega) \in A\}$ .

**Définition 8.16.** La loi d'une variable aléatoire est la fonction  $x \in \mathcal{X} \mapsto p(x) := \mathbb{P}([X = x])$ .

*Exemple 8.17.* 1. Si  $\mathcal{X} = \{0, 1\}$  et  $p \in [0, 1]$ , on dit que  $X$  suit une loi de Bernoulli lorsque  $p(1) = p$  et  $p(0) = 1 - p$ .

2. Si  $\mathcal{X} = \{1, \dots, N\}$ , on dit que  $X$  suit une loi uniforme si  $p(i) = \frac{1}{N}$  pour tout  $i \in \{1, \dots, N\}$ .

*Remarque 8.18.* Si  $A \subset \Omega$ , on note  $1_A$  la fonction indicatrice de  $A$  définie par

$$1_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A, \\ 0 & \text{si } \omega \in A^c. \end{cases}$$

L'indicatrice de l'événement  $A$  suit une loi de Bernoulli de paramètre  $\mathbb{P}(A)$ .

**Proposition 8.19.** Pour toute partie  $\mathcal{X}' \subset \mathcal{X}$ ,

$$\mathbb{P}([X \in \mathcal{X}']) = \sum_{x \in \mathcal{X}'} p(x).$$

En particulier,

$$\sum_{x \in \mathcal{X}} p(x) = 1.$$

On pourra admettre la preuve.

*Preuve :* L'ensemble  $\mathcal{X}'$  est de la forme  $\{x_j : j \in J\}$  où  $J$  est une partie de  $I$ . Pour tout  $j \in J$ , notons  $A_j$  l'événement  $[X = x_j]$ . Alors les  $A_j$  sont deux à deux disjoints et de réunion  $[X \in \mathcal{X}']$ . On en déduit

$$\sum_{x \in \mathcal{X}'} p(x) = \sum_{j \in J} \mathbb{P}(A_j) = \mathbb{P}(\cup_{j \in J} A_j) = \mathbb{P}([X \in \mathcal{X}']),$$

ce qui montre la première assertion. Dans le cas où  $\mathcal{X}' = \mathcal{X}$ , on obtient

$$\sum_{x \in \mathcal{X}} p(x) = \mathbb{P}([X \in \mathcal{X}]) = \mathbb{P}(\Omega) = 1,$$

ce qui prouve la seconde assertion. □

**Définition 8.20.** La fonction de répartition associée à une variable aléatoire  $X : \Omega \rightarrow \mathcal{X}$  est la fonction  $F_X : \mathbb{R} \rightarrow [0, 1]$  définie par

$$F_X(x) = \mathbb{P}([X \leq x]).$$

Une issue  $\omega$  appartient à l'ensemble  $[X \leq x]$  si et seulement si  $X(\omega) \leq x$ . De manière équivalente, il existe  $k \in I$  tel que  $X(\omega) = x_k$  et  $x_k \leq x$ . Ainsi

$$[X \leq x] = \cup_{k: x_k \leq x} [X = x_k].$$

Dans le membre de droite de l'égalité ci-dessus, on prend la réunion sur tous les indices  $k$  tels que  $x_k \leq x$ . On en déduit :

$$F_X(x) = \mathbb{P}(\cup_{k: x_k \leq x} [X = x_k]) = \sum_{k: x_k \leq x} p(x_k). \quad (8.1)$$

**Proposition 8.21.** La fonction de répartition  $F_X$  est une fonction croissante. De plus,

$$\lim_{x \rightarrow -\infty} F_X(x) = 0 \quad , \quad \lim_{x \rightarrow +\infty} F_X(x) = 1.$$

On pourra admettre la preuve.

Preuve : La monotonie de  $F_X$  provient de la monotonie de la mesure  $\mathbb{P}$  par rapport à l'inclusion (cela correspond à l'axiome 2 de la définition d'une probabilité). Alternativement, on peut utiliser (8.1). Montrons maintenant la dernière propriété (la limite en  $-\infty$  se démontre de manière analogue). Comme  $F_X$  est croissante et bornée par 1, sa limite en  $+\infty$  existe et est  $\leq 1$ . Montrons que cette limite vaut 1. Soit  $\varepsilon > 0$ . Alors il existe  $k_0 \in \mathbb{N}$  tel que

$$\sum_{k \leq k_0} p(x_k) \geq 1 - \varepsilon.$$

Ici, on a utilisé que la série dans le membre de gauche converge vers 1, par la proposition 8.19. Alors d'après la formule (8.1), en posant  $y = \max\{x_k : k \leq k_0\}$ ,

$$F(y) = \sum_{k: x_k \leq y} p(x_k) \geq \sum_{k \leq k_0} p(x_k) \geq 1 - \varepsilon.$$

Comme  $F_X$  est croissante, pour tout  $x \geq y$ ,  $F_X(x) \geq F_X(y) \geq 1 - \varepsilon$ . Cela prouve que  $\lim_{x \rightarrow +\infty} F_X(x) = 1$  comme attendu. □

## 8.3 Moments des variables aléatoires discrètes réelles

La notion d'espérance est la traduction mathématique de l'idée de moyenne des résultats obtenus en répétant une même expérience dans des conditions identiques et indépendantes.

### 8.3.1 Espérance

Soit  $X$  une variable aléatoire sur  $\Omega$  à valeurs dans  $\mathcal{X} = \{x_i, i \in I\} \subset \mathbb{R}$ , où  $I$  est soit  $\mathbb{N}$ , soit un ensemble fini. Chaque  $x_i$  est un nombre réel.

**Définition 8.22.** Si  $\sum_{i \in I} |x_i| \mathbb{P}(X = x_i) < \infty$ , on dit que  $X$  est intégrable et on définit l'espérance de  $X$  par

$$\mathbb{E}(X) = \sum_{i \in I} x_i \mathbb{P}(X = x_i).$$

Noter que si  $I$  est fini, alors  $X$  est automatiquement intégrable.

*Remarque 8.23.* Soit  $A$  un événement. L'indicatrice de  $A$  prend exactement deux valeurs 0 et 1. On en déduit

$$\mathbb{E}(1_A) = 1 \cdot \mathbb{P}([1_A = 1]) = \mathbb{P}(A).$$

Pour une variable aléatoire  $X$ , on peut donc exprimer son espérance par la formule suivante :

$$\mathbb{E}(X) = \sum_{i \in I} x_i \mathbb{P}([X = x_i]) = \sum_{i \in I} x_i \mathbb{E}(1_{[X=x_i]}).$$

### 8.3.2 Composition et linéarité

**Proposition 8.24.** [Théorème de transfert] Soient  $X : \Omega \rightarrow \mathcal{X}$  est une variable aléatoire et  $f : \mathcal{X} \rightarrow \mathbb{R}$  une fonction. Si  $\sum_{i \in I} |f(x_i)| \mathbb{P}([X = x_i])$  converge, alors  $f(X)$  est intégrable et :

$$\mathbb{E}(f(X)) = \sum_{i \in I} f(x_i) \mathbb{P}([X = x_i]).$$

On pourra admettre la preuve.

Preuve : On prouve ce résultat dans deux cas particuliers. On commence par faire l'hypothèse supplémentaire que  $f$  est injective. Dans ce cas,  $f(X)$  prend les valeurs distinctes  $f(x_i), i \in I$ . On en déduit par définition de l'espérance

$$\mathbb{E}(f(X)) = \sum_{i \in I} f(x_i) \mathbb{P}([f(X) = f(x_i)]).$$

Comme  $f$  est injective,  $[f(X) = f(x_i)] = [X = x_i]$  et donc  $\mathbb{P}([f(X) = f(x_i)]) = \mathbb{P}([X = x_i])$ . On en déduit

$$\mathbb{E}(f(X)) = \sum_{i \in I} f(x_i) \mathbb{P}([X = x_i]),$$

ce qui conclut la preuve dans ce cas là.

On ne suppose plus que  $f$  est injective, mais que  $\mathcal{X}$  est fini. On a toujours  $f(\mathcal{X}) = \{f(x_i) : i \in I\}$  mais comme  $f$  n'est plus supposée injective, il se peut que les  $f(x_i)$  ne soit pas distincts. Introduisons alors  $K \subset I$  tel que  $\{f(x_k) : k \in K\}$  soit l'ensemble des valeurs distinctes prises par  $f$ . Alors par définition de l'espérance,

$$\mathbb{E}(f(X)) = \sum_{k \in K} f(x_k) \mathbb{P}([f(X) = f(x_k)]).$$

Pour tout  $k \in K$ , notons  $I_k := \{i \in I : f(x_i) = f(x_k)\}$ . Alors

$$\mathbb{P}([f(X) = f(x_k)]) = \mathbb{P}(\cup_{i \in I_k} [X = x_i]) = \sum_{i \in I_k} \mathbb{P}([X = x_i]).$$

En reportant dans l'inégalité précédente, il vient

$$\mathbb{E}(f(X)) = \sum_{k \in K} f(x_k) \sum_{i \in I_k} \mathbb{P}([X = x_i]) = \sum_{k \in K} \sum_{i \in I_k} f(x_k) \mathbb{P}([X = x_i]) = \sum_{k \in K} \sum_{i \in I_k} f(x_i) \mathbb{P}([X = x_i]).$$

Pour obtenir la dernière égalité, on a utilisé que  $f(x_i) = f(x_k)$  lorsque  $i \in I_k$ . Comme  $I = \cup_{k \in K} I_k$  et qu'il s'agit d'ensembles finis, on obtient finalement

$$\mathbb{E}(f(X)) = \sum_{i \in I} f(x_i) \mathbb{P}([X = x_i])$$

ce qui achève la preuve. □

En particulier, si  $X$  est une variable aléatoire intégrable, alors il en est de même de  $\lambda X$  pour tout  $\lambda \in \mathbb{R}$  et de plus

$$\mathbb{E}(\lambda X) = \lambda \mathbb{E}(X).$$

Un autre cas particulier implique l'inégalité triangulaire :

**Proposition 8.25** (Inégalité triangulaire). *Soit  $X$  une variable aléatoire intégrable. Alors  $|X|$  est intégrable et*

$$|\mathbb{E}(X)| \leq \mathbb{E}(|X|).$$

Preuve : En utilisant l'inégalité triangulaire pour les séries numériques et en appliquant la proposition 8.24 à  $f = |\cdot|$ , on obtient :

$$|\mathbb{E}(X)| = \left| \sum_{i \in I} x_i \mathbb{P}([X = x_i]) \right| \leq \sum_{i \in I} |x_i| \mathbb{P}([X = x_i]) = \mathbb{E}(|X|),$$

ce qui achève la preuve. □

**Proposition 8.26.** *Si  $X$  et  $Y$  sont deux variables aléatoires intégrables à valeurs dans  $\mathcal{X}$  et  $\mathcal{Y}$  respectivement, alors on peut définir l'espérance de  $X + Y$  et*

$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y).$$

On admettra cette proposition.

Il suit de la définition que si  $X$  est une variable aléatoire positive (i.e. ne prenant que des valeurs positives) et intégrable, alors  $\mathbb{E}(X) \geq 0$ . En utilisant la linéarité de l'espérance, on obtient :

**Proposition 8.27** (Monotonie). *Soient  $X, Y$  deux variables aléatoires telles que pour tout  $\omega \in \Omega$ ,  $X(\omega) \leq Y(\omega)$ . Si  $X$  et  $Y$  sont intégrables, alors*

$$\mathbb{E}(X) \leq \mathbb{E}(Y).$$

### 8.3.3 Variance

**Définition 8.28** (Moments d'ordre  $m$ ). *Soit  $X$  une variable aléatoire. Si  $\sum_i |x_i|^m \mathbb{P}([X = x_i])$  converge, on définit le moment d'ordre  $m$  de  $X$  par  $\mathbb{E}(X^m)$ .*

Une variable prenant un nombre fini de valeurs a tous ses moments finis.

Comme conséquence de la proposition 8.24, on obtient :

$$\mathbb{E}(X^m) = \sum_{k \in I} x_k^m p(x_k).$$

La proposition suivante peut être vue comme un cas particulier de l'inégalité de Cauchy-Schwarz :

**Proposition 8.29.** *Soit  $X$  une variable aléatoire de second moment fini. Alors  $X$  est intégrable et*

$$\mathbb{E}(|X|) \leq [\mathbb{E}(|X|^2)]^{1/2}.$$

Une variable de second moment fini est donc de premier moment fini. L'écart-type permet de mesurer la distance à la moyenne :

**Définition 8.30.** Soit  $X$  une variable aléatoire de second moment fini. Alors on appelle variance de  $X$  la quantité

$$\mathbb{V}(X) = \mathbb{E}[(X - \mathbb{E}(X))^2] = \mathbb{E}(X^2) - (\mathbb{E}(X))^2.$$

La racine carrée de la variance s'appelle l'écart-type.

Pour justifier l'égalité dans l'énoncé précédent, observons que

$$\mathbb{E}[(X - \mathbb{E}(X))^2] = \mathbb{E}(X^2 + (\mathbb{E}(X))^2 - 2\mathbb{E}(X)X).$$

Lorsqu'on écrit dans la ligne ci-dessus  $(\mathbb{E}(X))^2$ , il s'agit ici de la fonction constante égale au nombre  $(\mathbb{E}(X))^2$ . Alors par linéarité,

$$\begin{aligned} \mathbb{E}[(X - \mathbb{E}(X))^2] &= \mathbb{E}(X^2) + \mathbb{E}((\mathbb{E}(X))^2) - 2\mathbb{E}(\mathbb{E}(X)X) \\ &= \mathbb{E}(X^2) + (\mathbb{E}(X))^2\mathbb{E}(1) - 2\mathbb{E}(X)\mathbb{E}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2. \end{aligned}$$

Dans la ligne ci-dessus, on a utilisé le fait que l'espérance de la fonction constante égale à 1 est, par définition de l'espérance,  $\mathbb{E}(1) = 1 \times \mathbb{P}(\Omega) = 1$ .

*Remarque 8.31.* La variance n'est pas linéaire : si  $\lambda \in \mathbb{R}$ ,

$$\mathbb{V}(\lambda X + \mu) = \lambda^2 \mathbb{V}(X).$$

Preuve :

$$\begin{aligned} \mathbb{V}(\lambda X + \mu) &= \mathbb{E}((\lambda X + \mu)^2) - (\mathbb{E}(\lambda X + \mu))^2 = \mathbb{E}(\lambda^2 X^2 + \mu^2 + 2\lambda\mu X) - (\lambda\mathbb{E}(X) + \mu)^2 \\ &= \lambda^2 \mathbb{E}(X^2) + \mu^2 + 2\lambda\mu \mathbb{E}(X) - \lambda^2 (\mathbb{E}(X))^2 - \mu^2 - 2\lambda\mu \mathbb{E}(X) \\ &= \lambda^2 \mathbb{V}(X). \end{aligned}$$

□

*Remarque 8.32.* Si  $\mathbb{V}(X) = 0$ , la variable  $X$  est constante en dehors d'un événement de probabilité nulle.

Preuve : Par hypothèse et en utilisant la proposition 8.24,  $0 = \mathbb{V}(X) = \sum_{i \in I} (x_i - \mathbb{E}(X))^2 p(x_i)$ .

Comme il s'agit d'une série de nombres positifs, cela implique que pour tout  $i \in I$ ,  $x_i = \mathbb{E}(X)$  ou  $p(x_i) = 0$ . On en déduit

$$\mathbb{P}(\cup_{i \in I: x_i \neq \mathbb{E}(X)} [X = x_i]) = \sum_{i \in I: x_i \neq \mathbb{E}(X)} p(x_i) = 0.$$

De plus,  $X = \mathbb{E}(X)$  sur  $\Omega \setminus \cup_{i \in I: x_i \neq \mathbb{E}(X)} [X = x_i]$ , ce qui conclut la preuve.

□

*Exemple 8.33.* Si  $X$  est une variable de Bernoulli de paramètre  $p$ , alors

$$\mathbb{E}(X) = 1 \times p = p, \quad \mathbb{V}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = 1^2 \times p - p^2 = p(1 - p).$$

*Exemple 8.34.* Si  $X$  suit une loi binomiale de paramètre  $p$  et  $n > 1$ , alors

$$\mathbb{E}(X) = np.$$

En effet,

$$\mathbb{E}(X) = \sum_{k=0}^n k C_n^k p^k (1-p)^{n-k} = \sum_{k=1}^n k C_n^k p^k (1-p)^{n-k}.$$

On remarque ensuite

$$k C_n^k = k \frac{n!}{k!(n-k)!} = n \frac{(n-1)!}{(k-1)!((n-1)-(k-1))!} = n C_{n-1}^{k-1}.$$

Il vient alors

$$\begin{aligned} \mathbb{E}(X) &= n \sum_{k=1}^n C_{n-1}^{k-1} p^{k-1} (1-p)^{n-k} = np \sum_{k=1}^n C_{n-1}^{k-1} p^{k-1} (1-p)^{(n-1)-(k-1)} \\ &= np \sum_{k=0}^{n-1} C_{n-1}^k p^k (1-p)^{(n-1)-k} = np(p + (1-p))^{n-1} = np. \end{aligned}$$

*Exercice 8.35.* Calculer la variance d'une variable aléatoire qui suit une loi binomiale de paramètres  $n$  et  $p$ .

Les deux inégalités suivantes sont très souvent utilisées dans la théorie des probabilités.

**Proposition 8.36.** 1. **Inégalité de Markov** Si  $X$  est une variable aléatoire intégrable à valeurs positives et  $a > 0$ , alors

$$\mathbb{P}([X \geq a]) \leq \frac{\mathbb{E}(X)}{a}.$$

2. **Inégalité de Bienaymé-Tchebychev** Si  $X$  est une variable aléatoire de second moment fini et  $a > 0$ , alors

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq a) \leq \frac{\mathbb{V}(X)}{a^2}.$$

Preuve : Pour tout  $\omega \in \Omega$ ,

$$1 \geq 1_{[X \geq a]}(\omega),$$

ce qui implique

$$X(\omega) \geq X(\omega) 1_{[X \geq a]}(\omega) \geq a 1_{[X \geq a]}(\omega). \quad (8.2)$$

La dernière inégalité résulte du fait que si  $X(\omega) < a$ , alors  $1_{[X \geq a]}(\omega) = 0$ . Comme (8.2) est vraie pour tout  $\omega$ , on a donc  $X \geq a 1_{[X \geq a]}$ . Alors par monotonie de l'espérance,

$$\mathbb{E}(X) \geq \mathbb{E}(a 1_{[X \geq a]}) = a \mathbb{E}(1_{[X \geq a]}) = a \mathbb{P}([X \geq a]),$$

ce qui prouve l'inégalité de Markov. En appliquant l'inégalité de Markov à  $(X - \mathbb{E}(X))^2$  et  $a^2$ , il vient

$$\mathbb{P}((X - \mathbb{E}(X))^2 \geq a^2) \leq \frac{\mathbb{E}((X - \mathbb{E}(X))^2)}{a^2}.$$

Pour conclure, il suffit d'observer que  $[(X - \mathbb{E}(X))^2 \geq a^2] = [|X - \mathbb{E}(X)| \geq a]$  et donc  $\mathbb{P}((X - \mathbb{E}(X))^2 \geq a^2) = \mathbb{P}(|X - \mathbb{E}(X)| \geq a)$ .

**Définition 8.37.** Si  $X$  et  $Y$  sont deux variables aléatoires de second moment fini, la covariance de  $X$  et  $Y$  est

$$\text{Cov}(X, Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).$$

Lorsque  $\mathbb{V}(X) > 0$  et  $\mathbb{V}(Y) > 0$ , le coefficient de corrélation est donné par

$$\text{Cor}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\mathbb{V}(X)\mathbb{V}(Y)}}.$$

*Exercice 8.38.* Montrer la propriété d'invariance d'échelle :

$$\forall a > 0, c > 0, b \in \mathbb{R}, d \in \mathbb{R}, \quad \text{Cor}(X, Y) = \text{Cor}(aX + b, cY + d).$$

## 8.4 Indépendance

### 8.4.1 Indépendance d'événements

**Définition 8.39** (Indépendance de deux événements). Deux événements  $A, B \subset \Omega$  sont dit indépendants si  $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$ .

*Remarque 8.40.* Si  $\mathbb{P}(B) > 0$ , alors  $A$  et  $B$  sont indépendants ssi  $\mathbb{P}(A|B) = \mathbb{P}(A)$ .

*Exemple 8.41.* Lorsqu'on lance deux dés simultanément, les deux résultats obtenus sont indépendants. Mathématiquement, l'univers est  $\Omega = \{1, \dots, 6\}^2$  muni de la probabilité uniforme et pour tout  $A, B \subset \{1, \dots, 6\}$ , les événements  $A \times \{1, \dots, 6\}$  et  $\{1, \dots, 6\} \times B$  sont indépendants :

$$\begin{aligned} \mathbb{P}((A \times \{1, \dots, 6\}) \cap (\{1, \dots, 6\} \times B)) &= \mathbb{P}(A \times B) = \frac{1}{36}(\text{card } A \times B) \\ &= \frac{1}{36}(\text{card } A)(\text{card } B) = \mathbb{P}(A \times \{1, \dots, 6\})\mathbb{P}(\{1, \dots, 6\} \times B). \end{aligned}$$

### 8.4.2 Indépendance de variables aléatoires

**Définition 8.42.** On dit que  $n$  variables aléatoires  $X_1, \dots, X_n$  à valeurs dans  $\mathcal{X}_1, \dots, \mathcal{X}_n$  sont indépendantes si pour tout  $A_1 \subset \mathcal{X}_1, \dots, A_n \subset \mathcal{X}_n$ ,

$$\mathbb{P}([X_1 \in A_1] \cap \dots \cap [X_n \in A_n]) = \mathbb{P}([X_1 \in A_1]) \dots \mathbb{P}([X_n \in A_n]).$$

On en déduit facilement :

*Remarque 8.43.* Si  $X$  et  $Y$  sont deux variables aléatoires indépendantes, alors pour toutes fonctions  $f$  et  $g$  de  $\mathbb{R}$  dans  $\mathbb{R}$ , les variables  $f(X)$  et  $g(Y)$  sont indépendantes.

**Proposition 8.44.** Les variables aléatoires discrètes  $X_1, \dots, X_n$  sont indépendantes si et seulement si pour tout  $x_1 \in \mathcal{X}_1, \dots, x_n \in \mathcal{X}_n$ ,

$$\mathbb{P}([X_1 = x_1, X_2 = x_2, \dots, X_n = x_n]) = \mathbb{P}([X_1 = x_1])\mathbb{P}([X_2 = x_2]) \dots \mathbb{P}([X_n = x_n]).$$

On pourra admettre la preuve.

Preuve (dans le cas où chaque  $\mathcal{X}_j$  est fini) : Si les  $X_1, \dots, X_n$  sont indépendantes, alors pour tout  $x_1 \in \mathcal{X}_1, \dots, x_n \in \mathcal{X}_n$ , notons  $A_j$  l'événement  $[X_j = x_j]$ ,  $1 \leq j \leq n$ . Alors

$$\begin{aligned} \mathbb{P}([X_1 = x_1, X_2 = x_2, \dots, X_n = x_n]) &= \mathbb{P}([X_1 \in A_1] \cap \dots \cap [X_n \in A_n]) \\ &= \mathbb{P}([X_1 \in A_1]) \dots \mathbb{P}([X_n \in A_n]) \\ &= \mathbb{P}([X_1 = x_1]) \dots \mathbb{P}([X_n = x_n]) \end{aligned}$$

ce qui montre l'implication directe.

Pour l'implication réciproque, soient  $A_1 \subset \mathcal{X}_1, \dots, A_n \subset \mathcal{X}_n$ . Pour chaque  $1 \leq j \leq n$ , on note  $x_{ij}$  les éléments de  $A_j$ , où  $i$  parcourt l'ensemble fini d'indices  $I_j$ . On a donc

$$A_j = \{x_{ij} : i \in I_j\}.$$

Alors

$$\begin{aligned} [X_1 \in A_1] \cap \dots \cap [X_n \in A_n] &= (\cup_{i_1 \in I_1} [X_1 = x_{i_1 1}]) \cap \dots \cap (\cup_{i_n \in I_n} [X_n = x_{i_n n}]) \\ &= \cup_{(i_1, \dots, i_n) \in I_1 \times \dots \times I_n} [X_1 = x_{i_1 1}] \cap \dots \cap [X_n = x_{i_n n}] \\ &= \cup_{(i_1, \dots, i_n) \in I_1 \times \dots \times I_n} [X_1 = x_{i_1 1}, \dots, X_n = x_{i_n n}]. \end{aligned}$$

On en déduit

$$\mathbb{P}([X_1 \in A_1] \cap \dots \cap [X_n \in A_n]) = \sum_{(i_1, \dots, i_n) \in I_1 \times \dots \times I_n} \mathbb{P}([X_1 = x_{i_1 1}, \dots, X_n = x_{i_n n}]).$$

En utilisant l'hypothèse de l'énoncé, on obtient

$$\begin{aligned}\mathbb{P}([X_1 \in A_1] \cap \dots \cap [X_n \in A_n]) &= \sum_{(i_1, \dots, i_n) \in I_1 \times \dots \times I_n} \mathbb{P}([X_1 = x_{i_1 1}]) \dots \mathbb{P}([X_n = x_{i_n n}]) \\ &= \sum_{i_1 \in I_1} \mathbb{P}([X_1 = x_{i_1 1}]) \dots \sum_{i_n \in I_n} \mathbb{P}([X_n = x_{i_n n}]) \\ &= P([X_1 \in A_1]) \dots P([X_n \in A_n]),\end{aligned}$$

ce qui montre que les variables aléatoires  $X_1, \dots, X_n$  sont indépendantes et achève la preuve.  $\square$

*Exemple 8.45.* Deux variables de Bernoulli  $B_1$  et  $B_2$  sont indépendantes si et seulement si les événements  $[B_1 = 1]$  et  $[B_2 = 1]$  sont indépendants.

En effet, d'après la proposition précédente, il suffit de montrer que pour tout  $\varepsilon_1, \varepsilon_2 \in \{0, 1\}$ , on a

$$\mathbb{P}([B_1 = \varepsilon_1] \cap [B_2 = \varepsilon_2]) = \mathbb{P}([B_1 = \varepsilon_1])\mathbb{P}([B_2 = \varepsilon_2]).$$

Considérons le cas  $\varepsilon_1 = \varepsilon_2 = 1$ . Alors par hypothèse,

$$\mathbb{P}([B_1 = \varepsilon_1] \cap [B_2 = \varepsilon_2]) = \mathbb{P}([B_1 = 1] \cap [B_2 = 1]) = \mathbb{P}([B_1 = 1])\mathbb{P}([B_2 = 1]).$$

Considérons à présent le cas  $\varepsilon_1 = 0, \varepsilon_2 = 1$ . Alors on écrit

$$\begin{aligned}[B_1 = \varepsilon_1] \cap [B_2 = \varepsilon_2] &= [B_1 = 0] \cap [B_2 = 1] = (\Omega \setminus [B_1 = 1]) \cap [B_2 = 1] \\ &= (\Omega \cap [B_2 = 1]) \setminus ([B_1 = 1] \cap [B_2 = 1]) = [B_2 = 1] \setminus ([B_1 = 1] \cap [B_2 = 1]).\end{aligned}$$

Alors

$$\begin{aligned}\mathbb{P}([B_1 = 0] \cap [B_2 = 1]) &= \mathbb{P}([B_2 = 1]) - \mathbb{P}([B_1 = 1] \cap [B_2 = 1]) \\ &= \mathbb{P}([B_2 = 1]) - \mathbb{P}([B_1 = 1])\mathbb{P}([B_2 = 1]) = (1 - \mathbb{P}([B_1 = 1]))\mathbb{P}([B_2 = 1]) \\ &= (\mathbb{P}(\Omega) - \mathbb{P}([B_1 = 1]))\mathbb{P}([B_2 = 1]) = \mathbb{P}(\Omega \setminus [B_1 = 1])\mathbb{P}([B_2 = 1]) \\ &= \mathbb{P}([B_1 = 0])\mathbb{P}([B_2 = 1]).\end{aligned}$$

Les autres cas se montrent de manière tout à fait similaire.

*Exercice 8.46.* Le lancer de deux dés est modélisé par l'univers  $\Omega = \{1, \dots, 6\}^2$  muni de la probabilité uniforme, chaque lancer correspondant à une variable aléatoire :  $X_i : (\omega_1, \omega_2) \in \Omega \mapsto \omega_i, i = 1, 2$ . Alors  $X_1$  et  $X_2$  sont indépendantes.

On admet la proposition suivante :

**Proposition 8.47.** Si  $X_1, \dots, X_k, Y_1, \dots, Y_\ell$  sont des variables indépendantes et  $f : \mathbb{R}^k \rightarrow \mathbb{R}$  est une fonction continue, alors  $f(X_1, \dots, X_k), Y_1, \dots, Y_\ell$  sont indépendantes.

### 8.4.3 Espérance et indépendance

Soient  $A, B \subset \Omega$  deux événements indépendants. En termes de fonctions indicatrices, l'indépendance de  $A$  et  $B$  s'écrit

$$\mathbb{E}(1_{A \cap B}) = \mathbb{E}(1_A)\mathbb{E}(1_B).$$

Comme  $1_{A \cap B} = 1_A 1_B$ , on peut aussi écrire

$$\mathbb{E}(1_A 1_B) = \mathbb{E}(1_A)\mathbb{E}(1_B).$$

Autrement dit, l'espérance du produit est égale au produit des espérances. Ce phénomène est tout à fait général sous l'hypothèse d'indépendance convenable :

**Proposition 8.48.** Soient  $X$  et  $Y$  deux variables aléatoires à valeurs dans  $\mathcal{X}$  et  $\mathcal{Y}$  respectivement. Soient  $f : \mathcal{X} \rightarrow \mathbb{R}$  et  $g : \mathcal{Y} \rightarrow \mathbb{R}$  deux fonctions telles que  $f(X)$  et  $g(Y)$  sont intégrables. Si  $X$  et  $Y$  sont indépendantes, alors le produit  $f(X)g(Y)$  est intégrable et

$$\mathbb{E}(f(X)g(Y)) = \mathbb{E}(f(X))\mathbb{E}(g(Y))$$

On pourra admettre la preuve.

Preuve (dans le cas où  $\mathcal{X}$  et  $\mathcal{Y}$  sont finis) : Notons  $\mathcal{X} = \{x_i : i \in I\}$  et  $\mathcal{Y} = \{y_j : j \in J\}$ . Alors la fonction  $\omega \in \Omega \mapsto f(X(\omega))g(Y(\omega))$  prend la valeur  $f(x_i)g(y_j)$  sur l'ensemble  $[X = x_i, Y = y_j]$ . On en déduit en suivant les arguments de la preuve de la proposition 8.24 que

$$\mathbb{E}(f(X)g(Y)) = \sum_{(i,j) \in I \times J} f(x_i)g(y_j)\mathbb{P}([X = x_i, Y = y_j]).$$

Comme  $X$  et  $Y$  sont indépendantes, il suit que

$$\begin{aligned} \mathbb{E}(f(X)g(Y)) &= \sum_{(i,j) \in I \times J} f(x_i)g(y_j)\mathbb{P}([X = x_i])\mathbb{P}([Y = y_j]) \\ &= \sum_{i \in I} f(x_i)\mathbb{P}([X = x_i]) \sum_{j \in J} g(y_j)\mathbb{P}([Y = y_j]) = \mathbb{E}(f(X))\mathbb{E}(g(Y)). \end{aligned}$$

□

Comme cas particulier de la proposition précédente, on a

**Corollaire 8.49.** Si les variables  $X_1, X_2$  sont indépendantes et intégrables, alors

$$\mathbb{E}(X_1X_2) = \mathbb{E}(X_1)\mathbb{E}(X_2).$$

Si on suppose de plus que  $X_1$  et  $X_2$  sont de second moment fini, alors

$$\mathbb{V}(X_1 + X_2) = \mathbb{V}(X_1) + \mathbb{V}(X_2).$$

En effet, en utilisant la proposition 8.49, on a

$$\mathbb{E}((X_1 + X_2)^2) = \mathbb{E}((X_1)^2) + \mathbb{E}((X_2)^2) + 2\mathbb{E}(X_1X_2) = \mathbb{E}((X_1)^2) + \mathbb{E}((X_2)^2) + 2\mathbb{E}(X_1)\mathbb{E}(X_2).$$

Par ailleurs,

$$(\mathbb{E}(X_1 + X_2))^2 = (\mathbb{E}(X_1) + \mathbb{E}(X_2))^2 = (\mathbb{E}(X_1))^2 + (\mathbb{E}(X_2))^2 + 2\mathbb{E}(X_1)\mathbb{E}(X_2).$$

On en déduit

$$\mathbb{V}(X_1 + X_2) = \mathbb{E}((X_1 + X_2)^2) - (\mathbb{E}(X_1 + X_2))^2 = \mathbb{V}(X_1) + \mathbb{V}(X_2),$$

comme attendu. Comme conséquence de la définition de la covariance, on a également :

**Proposition 8.50.** La covariance de deux variables indépendantes de second moment fini est nulle.

*Remarque 8.51.* Cette condition n'est pas suffisante.

## 8.5 Processus de Bernoulli

On considère dans cette section une suite  $X_i, i \in \mathbb{N}^*$  de variables indépendantes suivant une loi de Bernoulli de paramètre  $p$ . L'indépendance signifie ici que, pour tout  $n \geq 1$ , les variables  $X_i, 1 \leq i \leq n$ , sont indépendantes.

Pour toute issue  $\omega \in \Omega$ , on note

$$T(\omega) = \min\{i \in \mathbb{N}^* : X_i(\omega) = 1\}.$$

Autrement dit,  $T$  est la variable aléatoire qui donne l'indice du premier 1. Si l'ensemble du membre de droite est vide, alors on convient de poser  $T(\omega) = \infty$ . En fait, ce dernier cas advient si pour tout  $i \in \mathbb{N}^*$ ,  $X_i(\omega) = 0$ . Cet événement est de probabilité nulle : pour tout  $N \in \mathbb{N}^*$ ,

$$\mathbb{P}(\cap_{i \in \mathbb{N}^*} [X_i = 0]) \leq \mathbb{P}(\cap_{1 \leq i \leq N} [X_i = 0]) \leq \prod_{i=1}^N \mathbb{P}([X_i = 0]),$$

où la dernière inégalité résulte de l'indépendance des  $X_i$ . Ainsi,

$$\mathbb{P}(\cap_{i \in \mathbb{N}^*} [X_i = 0]) \leq \prod_{i=1}^N (1-p) = (1-p)^N.$$

Comme ceci doit être vrai pour tout  $N \in \mathbb{N}^*$  et que  $\lim_{N \rightarrow +\infty} (1-p)^N = 0$ , on en déduit

$$\mathbb{P}(\cap_{i \in \mathbb{N}^*} [X_i = 0]) = 0.$$

Conclusion : hors d'un ensemble de probabilité nulle,  $T$  prend ses valeurs dans  $\mathbb{N}^*$ .

Fixons  $n \in \mathbb{N}^*$ . Alors une issue  $\omega \in \Omega$  vérifie  $T(\omega) = n$  (le premier 1 apparaît en numéro  $n$ ) si et seulement si  $X_i(\omega) = 0$  pour tout  $1 \leq i \leq n-1$  (les  $n-1$  premiers numéros ont donné 0) et  $X_n(\omega) = 1$ . Ainsi,

$$[T = n] = [X_n = 1] \cap \cap_{1 \leq i \leq n-1} [X_i = 0].$$

Par indépendance, il suit que

$$\mathbb{P}([T = n]) = \mathbb{P}([X_n = 1]) \prod_{i=1}^{n-1} \mathbb{P}([X_i = 0]) = p(1-p)^{n-1}.$$

De même,

$$[T > n] = \cap_{1 \leq i \leq n} [X_i = 0]$$

et donc

$$\mathbb{P}([T > n]) = (1-p)^n.$$

**Définition 8.52.** On dit qu'une variable aléatoire  $T : \Omega \rightarrow \mathbb{N}^*$  telle que  $\mathbb{P}([T = n]) = p(1-p)^{n-1}$  suit une loi géométrique de paramètre  $p$ .

On note  $\mathcal{G}(p)$  la loi géométrique de paramètre  $p$ .

**Proposition 8.53.** Si  $T : \Omega \rightarrow \mathbb{N}^*$  suit une loi géométrique de paramètre  $p$ , alors pour tout  $k, j \in \mathbb{N}^*$ ,

$$\mathbb{P}([T > j+k] | [T > j]) = \mathbb{P}([T > k]).$$

Preuve : On revient à la définition de la probabilité conditionnelle :

$$\begin{aligned} \mathbb{P}([T > j+k] | [T > j]) &= \frac{\mathbb{P}([T > j+k] \cap [T > j])}{\mathbb{P}([T > j])} = \frac{\mathbb{P}([T > j+k])}{\mathbb{P}([T > j])} \\ &= \frac{(1-p)^{j+k}}{(1-p)^j} = (1-p)^k \\ &= \mathbb{P}([T > k]), \end{aligned}$$

comme attendu. □

L'espérance d'une variable aléatoire suivant une loi géométrique est

$$\mathbb{E}(T) = \sum_{n \geq 1} n \mathbb{P}([T = n]) = \sum_{n \geq 1} np(1-p)^{n-1} = p \sum_{n \geq 1} n(1-p)^{n-1}.$$

La série entière  $\sum_{n \geq 0} x^n$  converge sur  $] -1, 1[$  vers  $\frac{1}{1-x}$ . Elle est donc dérivable terme à terme sur cet intervalle et on a

$$\frac{1}{(1-x)^2} = \sum_{n \geq 0} nx^{n-1} = \sum_{n \geq 1} nx^{n-1}.$$

En appliquant cette égalité à  $x = 1 - p$ , on obtient

$$\mathbb{E}(T) = p \frac{1}{p^2} = \frac{1}{p}.$$

*Exercice 8.54.* Montrer que la variance d'une variable aléatoire  $T$  suivant une loi géométrique de paramètre  $p$  est

$$\mathbb{V}(T) = \frac{1-p}{p^2}.$$

On introduit à présent pour tout  $n \geq 1$ , la variable aléatoire :

$$S_n = X_1 + \dots + X_n.$$

Pour tout  $\omega \in \Omega$ ,  $0 \leq S_n(\omega) \leq n$ . Pour tout  $k \in \{0, \dots, n\}$ ,  $S_n(\omega) = k$  si exactement  $k$  des variables aléatoires  $X_1, \dots, X_n$  prennent la valeur 1 en  $\omega$  :

$$[S_n = k] = \bigcup_{\substack{I \subset \{1, 2, \dots, n\}, \\ |I|=k}} (\cap_{i \in I} [X_i = 1]) \cap (\cap_{i \notin I} [X_i = 0]).$$

Dans le membre de droite, on prend l'union sur tous les sous-ensembles  $I$  de  $\{1, 2, \dots, n\}$  qui sont de cardinal  $k$ . Ainsi, comme il s'agit d'une union disjointe et en utilisant l'indépendance des  $X_j$ ,

$$\mathbb{P}([S_n = k]) = \sum_{\substack{I \subset \{1, 2, \dots, n\}, \\ |I|=k}} \prod_{i \in I} \mathbb{P}([X_i = 1]) \prod_{i \notin I} \mathbb{P}([X_i = 0]) = \sum_{\substack{I \subset \{1, 2, \dots, n\}, \\ |I|=k}} p^k (1-p)^{n-k}.$$

Comme il y a  $C_n^k$  façons de choisir  $k$  éléments parmi  $n$ , c'est-à-dire  $C_n^k$  sous-ensembles  $I$  possibles, on en déduit

$$\mathbb{P}([S_n = k]) = C_n^k p^k (1-p)^{n-k}.$$

Ainsi, la loi de  $S_n$  est la loi binomiale de paramètres  $n$  et  $p$ .

**Proposition 8.55.** La somme de deux variables binomiales indépendantes de paramètres  $(k, p)$  et  $(\ell, p)$  respectivement est une variable binomiale de paramètres  $(k + \ell, p)$ .

*Preuve :* Notons  $S : \Omega \rightarrow \{0, \dots, k\}$  une variable binomiale de paramètres  $(k, p)$  et  $T : \Omega \rightarrow \{0, \dots, \ell\}$  une variable binomiale de paramètres  $(\ell, p)$ . Supposons que  $S$  et  $T$  soient indépendantes. Alors  $S + T : \Omega \rightarrow \{0, \dots, k + \ell\}$  vérifie

$$\forall i \in \{0, \dots, k + \ell\}, \quad [S + T = i] = \bigcup_{j=0}^i [S = j] \cap [T = i - j].$$

Comme les événements  $[S = j] \cap [T = i - j]$  sont disjoints pour deux valeurs distinctes de  $j$ ,

$$\mathbb{P}([S + T = i]) = \sum_{j=0}^i \mathbb{P}([S = j] \cap [T = i - j]).$$

Comme  $S$  et  $T$  sont indépendantes,

$$\mathbb{P}([S + T = i]) = \sum_{j=0}^i \mathbb{P}([S = j])\mathbb{P}([T = i - j]).$$

On utilise maintenant que  $S$  et  $T$  sont binomiales :

$$\mathbb{P}([S + T = i]) = \sum_{j=0}^i C_k^j p^j (1-p)^{k-j} C_\ell^{i-j} p^{i-j} (1-p)^{\ell+j-i} = p^i (1-p)^{\ell+k-i} \sum_{j=0}^i C_k^j C_\ell^{i-j}.$$

Pour calculer  $\sum_{j=0}^i C_k^j C_\ell^{i-j}$ , on calcule de deux manières différentes la quantité  $(1+x)^k (1+x)^\ell$ . D'une part,

$$(1+x)^k (1+x)^\ell = (1+x)^{k+\ell} = \sum_{i=0}^{k+\ell} C_{k+\ell}^i x^i.$$

D'autre part,

$$\begin{aligned} (1+x)^k (1+x)^\ell &= \left( \sum_{j=0}^k C_k^j x^j \right) \left( \sum_{j'=0}^{\ell} C_\ell^{j'} x^{j'} \right) \\ &= \sum_{j=0}^k \sum_{j'=0}^{\ell} x^{j+j'} C_k^j C_\ell^{j'} \\ &= \sum_{i=0}^{k+\ell} x^i \sum_{j+j'=i} C_k^j C_\ell^{j'} = \sum_{i=0}^{k+\ell} x^i \sum_{j=0}^i C_k^j C_\ell^{i-j}. \end{aligned}$$

Par identification des coefficients de ce polynôme, on obtient donc

$$\sum_{j=0}^i C_k^j C_\ell^{i-j} = C_{k+\ell}^i.$$

Finalement,

$$\mathbb{P}([S + T = i]) = C_{k+\ell}^i p^i (1-p)^{\ell+k-i},$$

ce qui prouve le résultat attendu. □

*Remarque 8.56.* Si les variables binomiales indépendantes  $S$  et  $T$  de paramètres  $(k, p)$  et  $(\ell, p)$  sont obtenues comme somme de variables de Bernoulli indépendantes de paramètre  $p$ , alors  $S + T$  est une somme de  $k + \ell$  variables de Bernoulli indépendantes de paramètre  $p$ , et suit donc une loi binomiale de paramètres  $(k + \ell, p)$ , ce qui corrobore le résultat précédent.

**Définition 8.57.** On dit qu'une variable aléatoire  $X : \Omega \rightarrow \mathbb{N}$  suit une loi de Poisson de paramètre  $\lambda > 0$ , notée  $\mathcal{P}(\lambda)$ , si

$$\forall k \in \mathbb{N}, \quad \mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}.$$

*Exercice 8.58.* Une variable  $X$  suivant une loi de Poisson de paramètre  $\lambda$  est intégrable et de second moment fini. De plus,

$$\mathbb{E}(X) = \lambda, \quad \mathbb{V}(X) = \lambda.$$

Dans la proposition suivante, on montre que sous certaines conditions, une suite de variables aléatoires suivant une loi binomiale converge en un certain sens vers une variable aléatoire suivant une loi de Poisson.

**Proposition 8.59.** Soit  $(p_n)_{n \geq 1}$  une suite de réels strictement positifs telle que  $p_n \sim \frac{\lambda}{n}$  lorsque  $n \rightarrow +\infty$ . Soit  $(X_n)_{n \geq 1}$  une suite de lois binomiales de paramètres  $n$  et  $p_n$ . Alors

$$\forall k \in \mathbb{N}, \quad C_n^k p_n^k (1 - p_n)^{n-k} \rightarrow e^{-\lambda} \frac{\lambda^k}{k!}$$

Preuve : On fixe  $k \in \mathbb{N}$ . On écrit

$$C_n^k p_n^k (1 - p_n)^{n-k} = \frac{1}{k!} \frac{n!}{(n-k)! n^k} (np_n)^k \left(1 - \frac{np_n}{n}\right)^{n-k}.$$

Par hypothèse,  $\lim_{n \rightarrow +\infty} (np_n)^k = \lambda^k$ . De plus,

$$\begin{aligned} \left(1 - \frac{np_n}{n}\right)^{n-k} &= e^{(n-k) \ln\left(1 - \frac{np_n}{n}\right)} \\ &= e^{(n+O(1)) \ln\left(1 - \frac{\lambda}{n} + o\left(\frac{1}{n}\right)\right)} \\ &= e^{(n+O(1))\left(-\frac{\lambda}{n} + o\left(\frac{1}{n}\right)\right)} \\ &= e^{(-\lambda + o(1))}. \end{aligned}$$

Ainsi,  $\lim_{n \rightarrow +\infty} \left(1 - \frac{np_n}{n}\right)^{n-k} = e^{-\lambda}$ . Enfin,

$$\frac{n!}{(n-k)! n^k} = \frac{n(n-1) \cdots (n-k+1)}{n^k} \sim \frac{n \cdot n \cdots n}{n^k}, \quad n \rightarrow +\infty.$$

On en déduit  $\lim_{n \rightarrow +\infty} \frac{n!}{(n-k)! n^k} = 1$ . Finalement,

$$\lim_{n \rightarrow +\infty} C_n^k p_n^k (1 - p_n)^{n-k} = e^{-\lambda} \frac{\lambda^k}{k!},$$

ce qu'il fallait démontrer. □

La dernière proposition peut recevoir l'interprétation suivante. La somme  $S_n$  d'un grand nombre de variables de Bernoulli indépendantes de petit paramètre suit approximativement la loi de Poisson de paramètre  $\mathbb{E}(S_n)$ .

La proposition précédente fait partie de la grande famille des *théorèmes centraux limites*, qui sont l'un des objets principaux de la section suivante.

## 8.6 Statistiques

Soit  $X$  une variable aléatoire (discrète ou à densité). Soient  $X_i, i \in \mathbb{N}^*$  des variables aléatoires indépendantes de même loi que  $X$ . On considère les variables aléatoires

$$\bar{X}_n = \frac{1}{n}(X_1 + \cdots + X_n) = \frac{S_n}{n}.$$

$$\mathbb{E}(\bar{X}_n) = \mathbb{E}(X), \quad \mathbb{V}(\bar{X}_n) = \frac{1}{n}\mathbb{V}(X).$$

### 8.6.1 Loi faible des grands nombres

**Proposition 8.60.** Soient  $X_1, \dots, X_n$  des variables aléatoires indépendantes de même loi de second moment fini. Alors pour tout  $\varepsilon > 0$ ,

$$\mathbb{P} [ |\bar{X}_n - \mathbb{E}(X)| > \varepsilon ] \leq \frac{\mathbb{V}(X)}{\varepsilon^2 n}.$$

Preuve : on calcule  $\mathbb{V}(\bar{X}_n - \mathbb{E}(X))$  puis on applique l'inégalité de Bienaymé-Tchebychev. □

*Remarque 8.61* (Intervalle de confiance). La probabilité que l'intervalle

$$\left[ \bar{X}_n - \sqrt{\frac{\mathbb{V}(X)}{na}}, \bar{X}_n + \sqrt{\frac{\mathbb{V}(X)}{na}} \right]$$

contienne  $\mathbb{E}(X)$  est supérieure ou égale à  $1 - a$ .

*Remarque 8.62.* Application pratique : la variance est souvent inconnue mais on peut la majorer :

1. si  $|X| \leq M$ , alors  $\mathbb{V}(X) \leq M^2$ ,
2. si  $X$  est de Bernoulli,  $\mathbb{V}(X) \leq 1/4$ .

Application numérique : pour  $n = 1000$ , au seuil de confiance 90% (=  $1 - a$ ), l'incertitude est de 5% (=  $\sqrt{\frac{\mathbb{V}(X_1)}{na}}$ ) pour une variable de Bernoulli.

### 8.6.2 Théorème de Moivre-Laplace

Nous admettrons le dernier résultat de ce chapitre :

**Théorème 8.63** (Théorème de Moivre-Laplace). Soit  $(X_n)_{n \geq 1}$  une suite de variables aléatoires de Bernoulli indépendantes<sup>1</sup> de même paramètre  $0 < p < 1$ . Alors pour tout  $a < b$ ,

$$\lim_{n \rightarrow +\infty} \mathbb{P} \left( a \leq \frac{\sqrt{n}}{\sqrt{p(1-p)}} (\bar{X}_n - p) \leq b \right) = \int_a^b \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx.$$

Comme la variable aléatoire

$$S_n : \omega \mapsto \sum_{k=1}^n X_k(\omega)$$

suit une loi binomiale de paramètre  $(n, p)$ , on peut reformuler cet énoncé en termes de variables aléatoires suivant une loi binomiale. En effet, soit  $(Z_n)_{n \geq 1}$  une suite de variables aléatoires telles que pour tout  $n \geq 1$ ,  $Z_n$  suit une loi binomiale de paramètres  $(n, p)$  : pour tout  $k \in \mathbb{N}$ , si  $k > n$ , alors  $\mathbb{P}(Z_n = k) = 0$  tandis que si  $0 \leq k \leq n$ ,  $\mathbb{P}(Z_n = k) = C_n^k p^k (1-p)^{n-k}$ . L'espérance de  $Z_n$  est  $np$  et son écart-type est  $\sqrt{np(1-p)}$ . Alors pour tout  $\alpha < \beta$ , pour tout  $n \geq 1$ ,

$$\mathbb{P}(\alpha \leq Z_n \leq \beta) = \mathbb{P}(\alpha \leq S_n \leq \beta).$$

En prenant  $\alpha = a\sqrt{np(1-p)} + np$  et  $\beta = b\sqrt{np(1-p)} + np$ , on a donc

$$\mathbb{P} \left( a \leq \frac{Z_n - \mathbb{E}(Z_n)}{\sqrt{\mathbb{V}(Z_n)}} \leq b \right) = \mathbb{P} \left( a \leq \frac{S_n - np}{\sqrt{p(1-p)}\sqrt{n}} \leq b \right).$$

---

1. Ce qui signifie que pour tout  $n \geq 1$ ,  $X_1, \dots, X_n$  sont indépendantes.

Le théorème de Moivre-Laplace s'écrit donc

$$\lim_{n \rightarrow +\infty} \mathbb{P} \left( a \leq \frac{Z_n - \mathbb{E}(Z_n)}{\sqrt{\mathbb{V}(Z_n)}} \leq b \right) = \int_a^b \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{x^2}{2} \right) dx.$$

Le premier intérêt de ce théorème est de donner une estimation de  $\mathbb{P}(\alpha \leq Z_n \leq \beta)$  lorsque  $Z_n$  suit une loi binomiale, qui serait impossible à calculer en pratique si on utilisait la formule

$$\mathbb{P}(\alpha \leq \frac{1}{\sigma_n} (Z_n - \mathbb{E}(Z_n)) \leq \beta) = \sum_{k=\alpha}^{\beta} C_n^k p^k (1-p)^{n-k}.$$

Pour mettre en évidence un autre intérêt de ce théorème, on peut imaginer un lancer de pièce tombant sur pile avec une probabilité  $p$  et tombant sur face avec une probabilité  $(1-p)$ . On cherche à donner une bonne estimation de  $p$ . Soit  $0 < \varepsilon < 1$ . Comme la fonction

$$\eta \in [0, +\infty[ \mapsto \int_{-\eta}^{\eta} \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{x^2}{2} \right) dx$$

est continue, strictement croissante (car l'intégrande est  $> 0$  en tout point), s'annule en 0 et tend vers 1 en  $+\infty$ , on en déduit à l'aide du théorème des valeurs intermédiaires qu'il existe un unique  $\eta > 0$  tel que

$$\int_{-\eta}^{\eta} \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{x^2}{2} \right) dx = 1 - \varepsilon.$$

Alors avec une probabilité asymptotique de  $1 - \varepsilon$ , la moyenne empirique  $\bar{X}_n$  s'écarte du paramètre  $p$  de moins de  $\eta \frac{\sqrt{p(1-p)}}{\sqrt{n}} \leq \frac{\eta}{2\sqrt{n}}$ , ce qui sera une approximation d'autant meilleure que  $n$  est grand. Naturellement, plus on est exigeant sur la probabilité (autrement dit, plus on réduit  $\varepsilon$ ), plus le paramètre  $\eta$  grandit, et donc plus on doit faire de lancers (i.e. plus on doit augmenter  $n$ ) pour préserver une marge d'erreur identique. On a ainsi un moyen d'obtenir une bonne approximation de  $p$  avec une grande probabilité.



## Chapitre 9

# Intégrales à paramètres et intégrales multiples

Ici,  $\mathbb{K}$  désigne  $\mathbb{R}$  ou  $\mathbb{C}$ .

### 9.1 Régularité des intégrales à paramètres

Soient  $I$  et  $J$  deux intervalles de  $\mathbb{R}$ . Ces intervalles ne sont pas nécessairement des segments, c'est-à-dire des intervalles fermés bornés, de la forme  $[a, b]$ , où  $a < b$  sont deux réels. Lorsqu'on considère des intégrales sur  $J$  dans le cas où  $J$  n'est pas un segment, on fait référence à la théorie des intégrales sur un intervalle quelconque étudiée au semestre précédent.

#### 9.1.1 Rappels sur l'intégrale sur un intervalle quelconque

On suppose ici connue la définition de l'intégrale d'une fonction continue (ou continue par morceaux) sur un segment.

On se donne un intervalle quelconque  $J$  dont on note  $-\infty \leq \alpha < \beta \leq +\infty$  les extrémités. On a donc  $J = [\alpha, \beta]$  (dans ce cas  $J$  est un segment), ou  $J = ]\alpha, \beta]$ , ou  $J = [\alpha, \beta[$  ou  $J = ]\alpha, \beta[$ .

1. On dit qu'une fonction continue et positive  $\varphi : J \rightarrow \mathbb{R}^+$  est intégrable sur  $J$  s'il existe  $M > 0$  tel que pour tout segment  $[a, b] \subset J$ ,

$$\int_{[a,b]} \varphi(t) dt \leq M.$$

Si  $J$  est un segment, l'intégrabilité est automatique.

2. On dit qu'une fonction continue  $\varphi : J \rightarrow \mathbb{R}$  est intégrable sur  $J$  si  $|f|$  (qui est une fonction continue positive) est intégrable sur  $J$ .
3. Dans ce cas, pour tout  $x_0 \in J$ , les limites suivantes existent :

$$\lim_{y \rightarrow \beta} \int_{x_0}^y \varphi(t) dt \quad , \quad \lim_{x \rightarrow \alpha} \int_x^{x_0} \varphi(t) dt$$

et on note

$$\int_{\alpha}^{\beta} \varphi(t) dt = \lim_{x \rightarrow \alpha} \int_x^{x_0} \varphi(t) dt + \lim_{y \rightarrow \beta} \int_{x_0}^y \varphi(t) dt.$$

4. Il se peut que les limites précédentes existent sans que  $\varphi$  ne soit intégrable. Dans ce cas, on note encore

$$\int_{\alpha}^{\beta} f(t) dt = \lim_{x \rightarrow \alpha} \int_x^{x_0} \varphi(t) dt + \lim_{y \rightarrow \beta} \int_{x_0}^y \varphi(t) dt$$

et on dit que l'intégrale est semi-convergente.

En pratique, étant donné une fonction  $\varphi$  définie sur un intervalle  $J \subset \mathbb{R}$ , on procède selon la démarche suivante :

1. On s'assure que la fonction est continue par morceaux sur l'intervalle  $J$ .
2. Si l'extrémité gauche de  $J$  n'est pas  $-\infty$  et n'appartient pas à  $J$  (autrement dit,  $J$  est de la forme  $] \alpha, \beta[$  ou  $] \alpha, \beta]$  avec  $\alpha \in \mathbb{R}$ ), on étudie la possibilité de prolonger  $\varphi$  par continuité en  $\alpha$ . Si  $\varphi$  est continue en  $\alpha$  ou peut être prolongée par continuité en  $\alpha$ , alors  $\varphi$  est automatiquement intégrable sur tout segment de la forme  $] \alpha, x_0]$ , pour tout  $x_0 \in J$ .
3. On fait le même travail en  $\beta$ .
4. Si  $\alpha = -\infty$  ou bien s'il n'a pas été possible de prolonger  $\varphi$  par continuité en  $\alpha$ , on fixe un  $x_0 \in ] \alpha, \beta[$  et on se donne un  $a \in ] \alpha, x_0[$ . Comme  $\varphi$  est continue par morceaux sur le segment  $[a, x_0]$ , on peut considérer les intégrales  $\int_a^{x_0} |\varphi|$  et  $\int_a^{x_0} \varphi$ .
  - (a) S'il existe un majorant de  $\int_a^{x_0} |\varphi|$  qui soit indépendant de  $a$ , alors  $\varphi$  sera intégrable sur  $] \alpha, x_0]$ . C'est en fait le cas si et seulement si la fonction  $a \mapsto \int_a^{x_0} |\varphi|$  a une limite quand  $a$  tend vers  $\alpha$ . Dans ce cas, on peut affirmer que  $\varphi$  est intégrable sur  $] \alpha, x_0]$  et

$$\int_{\alpha}^{x_0} \varphi = \lim_{a \rightarrow \alpha} \int_a^{x_0} \varphi.$$

- (b) Si  $\varphi$  n'est pas intégrable sur  $] \alpha, x_0]$ , on regarde si la fonction  $a \mapsto \int_a^{x_0} \varphi$  a une limite quand  $a$  tend vers  $\alpha$ . Si c'est le cas, l'intégrale  $\int_{\alpha}^{x_0} \varphi$  est semi-convergente.

5. On fait le même travail en  $\beta$ .
6. Si  $\varphi$  est intégrable sur  $] \alpha, x_0]$  et sur  $[x_0, \beta[$ , alors  $\varphi$  est intégrable sur  $] \alpha, \beta[$  et

$$\int_{\alpha}^{\beta} \varphi = \int_{\alpha}^{x_0} \varphi + \int_{x_0}^{\beta} \varphi.$$

Si  $\int_{\alpha}^{x_0} \varphi$  et  $\int_{x_0}^{\beta} \varphi$  sont semi-convergentes, alors  $\int_{\alpha}^{\beta} \varphi$  est semi-convergente et

$$\int_{\alpha}^{\beta} \varphi = \int_{\alpha}^{x_0} \varphi + \int_{x_0}^{\beta} \varphi.$$

*Exemple 9.1.* La fonction  $t \mapsto \frac{1}{t^2}$  est absolument intégrable sur l'intervalle  $[1, +\infty[$  mais ne l'est pas sur l'intervalle  $]0, +\infty[$ .

En effet, la fonction est continue et positive sur  $]0, +\infty[$ . De plus, pour tout  $b > 1$ ,

$$\int_1^b \frac{dt}{t^2} = 1 - \frac{1}{b}.$$

Cette quantité est majorée par 1 (qui ne dépend pas de  $b$ ). Donc la fonction  $t \mapsto \frac{1}{t^2}$  est intégrable sur  $[1, +\infty[$ , d'intégrale :

$$\int_1^{+\infty} \frac{dt}{t^2} = \lim_{b \rightarrow +\infty} \left(1 - \frac{1}{b}\right) = 1.$$

Maintenant, pour tout  $\varepsilon \in ]0, 1[$ ,

$$\int_{\varepsilon}^1 \frac{dt}{t^2} = \frac{1}{\varepsilon} - 1.$$

Cette quantité tend vers  $+\infty$  quand  $\varepsilon$  tend vers 0. En particulier, elle ne peut être majorée indépendamment de  $\varepsilon$ . On en déduit que  $t \mapsto \frac{1}{t^2}$  n'est pas intégrable sur  $]0, +\infty[$ .

*Exemple 9.2.* La fonction  $t \mapsto \frac{\cos t}{t^2}$  est intégrable sur  $[1, +\infty[$  mais pas sur  $]0, +\infty[$ .

La fonction  $\varphi : t \mapsto \frac{\cos t}{t^2}$  est continue sur  $]0, +\infty[$ . Soit  $b > 1$ . Comme pour tout  $t \in [1, +\infty[$ ,  $|\varphi(t)| \leq 1/t^2$ ,

$$\int_1^b |\varphi(t)| dt \leq \int_1^b \frac{dt}{t^2}.$$

La fonction  $t \mapsto \frac{1}{t^2}$  est intégrable sur  $[1, +\infty[$  d'après l'exemple précédent. Donc  $\int_1^b \frac{dt}{t^2}$  peut être majoré indépendamment de  $b$ , par exemple (l'intégrande  $t \mapsto 1/t^2$  étant à valeurs positives) par  $\int_1^{+\infty} \frac{dt}{t^2}$ .

On en déduit que  $|\varphi|$  est intégrable sur  $[1, +\infty[$  et donc  $\varphi$  aussi. Montrons à présent que  $\varphi$  n'est pas intégrable sur  $]0, 1]$ . Comme la fonction  $\cos$  est décroissante sur  $[0, 1]$ , pour tout  $t \in [0, 1]$ ,

$$\frac{|\cos t|}{t^2} = \frac{\cos t}{t^2} \geq \frac{\cos 1}{t^2}.$$

On en déduit que pour tout  $\varepsilon \in ]0, 1[$ ,

$$\int_\varepsilon^1 \frac{|\cos t|}{t^2} dt \geq \int_\varepsilon^1 \frac{\cos 1}{t^2} dt = \left(\frac{1}{\varepsilon} - 1\right) \cos 1.$$

Donc  $\lim_{\varepsilon \rightarrow 0} \int_\varepsilon^1 |\varphi| = +\infty$ , ce qui montre que  $\varphi$  n'est pas intégrable sur  $]0, 1]$ . En fait, l'intégrale de  $\varphi$  n'est même pas semi-convergente sur  $]0, 1]$ . En effet, par intégration par parties,

$$\begin{aligned} \int_\varepsilon^1 \frac{\cos t}{t^2} dt &= \left[ \frac{-\cos t}{t} \right]_\varepsilon^1 - \int_\varepsilon^1 \frac{\sin t}{t} dt \\ &= \frac{\cos \varepsilon}{\varepsilon} - \cos 1 - \int_\varepsilon^1 \frac{\sin t}{t} dt. \end{aligned}$$

D'une part,

$$\lim_{\varepsilon \rightarrow 0} \frac{\cos \varepsilon}{\varepsilon} = +\infty.$$

D'autre part, la fonction  $t \mapsto \frac{\sin t}{t}$  est une fonction continue sur  $]0, 1]$  qui se prolonge par continuité en 0. Donc c'est une fonction intégrable sur  $[0, 1]$  et

$$\lim_{\varepsilon \rightarrow 0} \int_\varepsilon^1 \frac{\sin t}{t} dt = \int_0^1 \frac{\sin t}{t} dt.$$

On en déduit que

$$\lim_{\varepsilon \rightarrow 0} \int_\varepsilon^1 \frac{\cos t}{t^2} dt = +\infty.$$

Conclusion : l'intégrale de  $\varphi$  n'est pas semi-convergente sur  $]0, 1]$ .

*Exemple 9.3.* La fonction  $\psi : t \mapsto \frac{\sin t}{t}$  est d'intégrale convergente sur l'intervalle  $]0, +\infty[$  mais elle n'est pas absolument convergente sur ce même intervalle.

La fonction  $\psi$  est continue sur  $]0, +\infty[$  et se prolonge par continuité en 0. Il suffit donc d'étudier son intégrabilité en  $+\infty$ . Soit  $k \in \mathbb{N}^*$ . On écrit

$$\int_0^{2k\pi} \frac{|\sin t|}{t} dt = \sum_{\ell=0}^{k-1} \int_{2\ell\pi}^{2(\ell+1)\pi} \frac{|\sin t|}{t} dt.$$

Pour tout  $\ell \in \{0, \dots, k-1\}$ , par  $2\pi$  périodicité de  $\sin$ ,

$$\int_{2\ell\pi}^{2(\ell+1)\pi} \frac{|\sin t|}{t} dt = \int_0^{2\pi} \frac{|\sin t|}{t + 2\ell\pi} dt.$$

L'intégrande étant positif,

$$\int_{2\ell\pi}^{2(\ell+1)\pi} \frac{|\sin t|}{t} dt \geq \int_{\pi/4}^{3\pi/4} \frac{|\sin t|}{t + 2\ell\pi} dt = \int_{\pi/4}^{3\pi/4} \frac{\sin t}{t + 2\ell\pi} dt.$$

Sur l'intervalle  $[\pi/4, 3\pi/4]$ , la fonction  $\sin$  est minorée par  $\frac{\sqrt{2}}{2}$ . Donc

$$\begin{aligned} \int_{\pi/4}^{3\pi/4} \frac{\sin t}{t + 2\ell\pi} dt &\geq \frac{\sqrt{2}}{2} \int_{\pi/4}^{3\pi/4} \frac{dt}{t + 2\ell\pi} \\ &= \frac{\sqrt{2}}{2} \ln \left( \frac{\frac{3\pi}{4} + 2\ell\pi}{\frac{\pi}{4} + 2\ell\pi} \right). \end{aligned}$$

Introduisons

$$u_\ell = \ln \left( \frac{\frac{3\pi}{4} + 2\ell\pi}{\frac{\pi}{4} + 2\ell\pi} \right), \quad \ell \geq 0.$$

Ainsi

$$\int_{2\ell\pi}^{2(\ell+1)\pi} \frac{|\sin t|}{t} dt \geq \frac{\sqrt{2}}{2} u_\ell.$$

On a donc montré que

$$\int_0^{2k\pi} \frac{|\sin t|}{t} dt \geq \frac{\sqrt{2}}{2} \sum_{\ell=0}^{k-1} u_\ell. \quad (9.1)$$

Montrons que la série  $\sum_{\ell} u_\ell$  diverge. D'abord,  $u_\ell \geq 0$ . Ensuite,

$$\begin{aligned} u_\ell &= \ln \left( \frac{1 + \frac{3}{8\ell}}{1 + \frac{1}{8\ell}} \right) \\ &= \ln \left( 1 + \frac{3}{8\ell} \right) - \ln \left( 1 + \frac{1}{8\ell} \right) \\ &= \left( \frac{3}{8\ell} + O\left(\frac{1}{\ell^2}\right) \right) - \left( \frac{1}{8\ell} + O\left(\frac{1}{\ell^2}\right) \right) \\ &= \frac{1}{4\ell} + O\left(\frac{1}{\ell^2}\right). \end{aligned}$$

On en déduit que  $u_\ell \sim \frac{1}{4\ell}$  pour  $\ell \rightarrow +\infty$ . Par comparaison de séries à termes positifs, la série  $\sum_{\ell} u_\ell$  diverge et on a

$$\lim_{k \rightarrow +\infty} \sum_{\ell=0}^{k-1} u_\ell = +\infty.$$

Revenant à (9.1), on obtient

$$\lim_{k \rightarrow +\infty} \int_0^{2k\pi} \frac{|\sin t|}{t} dt = +\infty.$$

Il est donc impossible de majorer  $\int_0^b \frac{|\sin t|}{t} dt$  indépendamment de  $b$ . Conclusion : La fonction  $\psi$  n'est pas intégrable sur  $[0, +\infty[$ .

On montre à présent que  $\psi$  a une intégrale semi-convergente sur  $[1, +\infty[$ . Par intégration par parties,

$$\begin{aligned} \int_1^b \frac{\sin t}{t} dt &= \left[ \frac{-\cos t}{t} \right]_1^b - \int_1^b \frac{\cos t}{t^2} dt \\ &= \cos 1 - \frac{\cos b}{b} - \int_1^b \frac{\cos t}{t^2} dt. \end{aligned}$$

Par l'exemple précédent, la fonction  $t \mapsto \frac{\cos t}{t^2}$  est intégrable sur  $[1, +\infty[$  donc  $b \mapsto \int_1^b \frac{\cos t}{t^2} dt$  a une limite en  $+\infty$  et

$$\lim_{b \rightarrow +\infty} \int_1^b \frac{\cos t}{t^2} dt = \int_1^{+\infty} \frac{\cos t}{t^2} dt.$$

Comme  $|\cos b| \leq 1$ ,  $\lim_{b \rightarrow +\infty} \frac{\cos b}{b} = 0$ . On en déduit que  $\psi$  a une intégrale semi-convergente sur  $[1, +\infty[$  et

$$\lim_{b \rightarrow +\infty} \int_1^b \frac{\sin t}{t} dt = \cos 1 - \int_1^{+\infty} \frac{\cos t}{t^2} dt,$$

ce qu'il fallait démontrer.

### 9.1.2 Continuité sous le signe intégral

Dans la suite, on considère une fonction de deux variables  $f : (x, t) \in I \times J \mapsto \mathbb{K}$  et on s'intéresse à la fonction

$$F : x \mapsto \int_J f(x, t) dt.$$

**Théorème 9.4** (Théorème de continuité sous le signe intégral). *On suppose que*

1. (continuité) la fonction  $f$  est continue sur  $I \times J$ ,
2. (domination) il existe une fonction  $\varphi : J \rightarrow \mathbb{R}$  continue et absolument intégrable sur  $J$  telle que

$$\forall (x, t) \in I \times J, |f(x, t)| \leq \varphi(t).$$

Alors la fonction  $F$  est bien définie et continue sur  $I$ .

Dans le théorème précédent, si  $J$  est un segment  $[a, b]$ , alors l'hypothèse de domination est superflue (en fait, elle est automatiquement vérifiée sur tout ensemble de la forme  $I' \times J$ , pour chaque segment  $I' \subset I$ , ce qui est suffisant pour obtenir la conclusion).

*Exercice 9.5.* Soit  $f : [a, b] \times [c, d] \rightarrow \mathbb{K}$  une fonction continue. Montrer que la fonction

$$F : (x, y) \in [a, b] \times [c, d] \mapsto \int_a^y f(x, t) dt$$

est continue sur  $[a, b] \times [c, d]$ .

Notons  $\| \cdot \|$  une norme quelconque sur  $\mathbb{R}^2$  (elles sont toutes équivalentes). Soit  $(x_0, y_0) \in [a, b] \times [c, d]$  et montrons que  $F$  est continue en  $(x_0, y_0)$ . Soit  $\varepsilon > 0$ . Pour tout  $(x, y) \in [a, b] \times [c, d]$ ,

$$F(x, y) - F(x_0, y_0) = F(x, y) - F(x, y_0) + F(x, y_0) - F(x_0, y_0).$$

Par continuité de  $f$  en  $(x_0, y_0)$ , il existe  $\eta_1 > 0$  tel que pour tout  $(x, y) \in [a, b] \times [c, d]$ , si  $\|(x, y) - (x_0, y_0)\| \leq \eta_1$ , alors

$$|f(x, y) - f(x_0, y_0)| \leq \varepsilon$$

et donc

$$|f(x, y)| \leq \varepsilon + |f(x_0, y_0)|.$$

Pour un tel couple  $(x, y)$ , on a donc

$$\begin{aligned} |F(x, y) - F(x_0, y_0)| &= \left| \int_a^y f(x, t) dt - \int_a^{y_0} f(x, t) dt \right| \\ &\leq \left| \int_{y_0}^y |f(x, t)| dt \right| \\ &\leq (\varepsilon + |f(x_0, y_0)|) |y - y_0|. \end{aligned}$$

Si on exige de plus  $|y - y_0| \leq \frac{\varepsilon}{2(\varepsilon + |f(x_0, y_0)|)}$ , alors

$$|F(x, y) - F(x_0, y_0)| \leq \frac{\varepsilon}{2}.$$

Comme la fonction  $(x, t) \mapsto f(x, t)$  est continue sur  $[a, b] \times [c, d]$  et qu'on intègre sur le segment  $[a, y_0]$ , on peut appliquer le théorème de continuité sous le signe intégral à la fonction  $x \mapsto F(x, y_0) = \int_a^{y_0} f(x, t) dt$ . Il existe donc  $\eta_2 > 0$  tel que pour tout  $x \in [a, b]$ , si  $|x - x_0| \leq \eta_2$ , alors

$$|F(x, y_0) - F(x_0, y_0)| \leq \frac{\varepsilon}{2}.$$

Récapitulons : notant  $\eta = \min\left(\eta_1, \eta_2, \frac{\varepsilon}{2(\varepsilon + |f(x_0, y_0)|)}\right)$ , si  $\|(x, y) - (x_0, y_0)\| \leq \eta$ , alors  $|y - y_0| \leq \frac{\varepsilon}{2(\varepsilon + |f(x_0, y_0)|)}$  et  $|x - x_0| \leq \eta_2$ . Donc

$$|F(x, y) - F(x_0, y_0)| \leq |F(x, y) - F(x, y_0)| + |F(x, y_0) - F(x_0, y_0)| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

ce qui achève la preuve de la continuité de  $F$  en  $(x_0, y_0)$ .

### 9.1.3 Dérivabilité sous le signe intégral

**Théorème 9.6.** *On suppose que*

1. (continuité de la fonction et de sa dérivée partielle) la fonction  $f$  est continue sur  $I \times J$  et admet une dérivée partielle  $\frac{\partial f}{\partial x}$  par rapport à la première variable qui est aussi continue sur  $I \times J$ ,
2. (domination) il existe une fonction  $\varphi : J \rightarrow \mathbb{R}$  continue et absolument intégrable sur  $J$  telle que

$$\forall (x, t) \in I \times J, |f(x, t)| + \left| \frac{\partial f}{\partial x}(x, t) \right| \leq \varphi(t).$$

Alors la fonction  $F$  est bien définie et  $C^1$  sur  $I$ , de dérivée

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

Dans le théorème précédent, si  $J$  est un segment  $[a, b]$ , alors l'hypothèse de domination est superflue (en fait, elle est automatiquement vérifiée sur tout ensemble de la forme  $I' \times J$ , pour chaque segment  $I' \subset I$ , ce qui est suffisant pour obtenir la conclusion).

**Exercice 9.7.** 1. Montrer que pour tout  $t \in \mathbb{R}$ , l'intégrale de la fonction  $x \mapsto e^{-x^2} \cos(tx)$  converge sur  $\mathbb{R}$ .

2. Montrer que la fonction

$$f : t \mapsto \int_{\mathbb{R}} e^{-x^2} \cos(tx) dx$$

est continue sur  $\mathbb{R}$ .

3. Montrer que  $f$  est  $C^1$  sur  $\mathbb{R}$  et que

$$f'(t) = - \int_{\mathbb{R}} x e^{-x^2} \sin(tx) dx.$$

4. Montrer que  $f$  est solution de l'équation  $y'(t) + \frac{t}{2}y(t) = 0$ .

5. On admet que  $\int_{\mathbb{R}} e^{-x^2} dx = \sqrt{\pi}$ . En déduire une expression explicite de  $f$ .

**Exercice 9.8.** Soit  $I$  un intervalle ouvert,  $a, b \in \mathbb{R}$  et  $f : I \times [a, b] \rightarrow \mathbb{R}$  une fonction continue admettant une dérivée partielle  $\frac{\partial f}{\partial x}$  continue sur  $I \times [a, b]$ . Soient  $g_1, g_2 : I \rightarrow \mathbb{R}$  deux fonctions  $C^1$  sur  $I$  telles que pour tout  $x \in I$ ,  $g_1(x), g_2(x) \in ]a, b[$ . Alors

$$F : x \in I \mapsto \int_{g_1(x)}^{g_2(x)} f(x, t) dt$$

est  $C^1$  et

$$F'(x) = g_2'(x)f(x, g_2(x)) - g_1'(x)f(x, g_1(x)) + \int_{g_1(x)}^{g_2(x)} \frac{\partial f}{\partial x}(x, t) dt.$$

Introduisons la fonction

$$G(x, y, z) = \int_y^z f(x, t) dt$$

définie pour tout  $y, z \in [a, b]$  et  $x \in I$ . Le nombre  $G(x, y, z)$  est bien défini comme intégrale de la fonction continue  $t \mapsto f(x, t)$  sur le segment d'extrémités  $y$  et  $z$ . De plus,

$$F(x) = G(x, g_1(x), g_2(x)).$$

Pour montrer que  $F$  est  $C^1$  sur  $I$ , il suffit donc de montrer que  $G$  est une fonction  $C^1$  sur  $I \times ]a, b[ \times ]a, b[$ . Fixons  $t_0 \in [a, b]$ . Alors par la relation de Chasles, pour tout  $x \in I$ ,  $y, z \in [a, b]$ ,

$$G(x, y, z) = G(x, y, t_0) + G(x, t_0, z).$$

Par l'exercice 9.5, la fonction

$$(x, z) \in I \times [a, b] \mapsto \int_{t_0}^z f(x, t) dt = G(x, t_0, z)$$

est continue en ses deux variables. Il en est de même de  $(x, y) \mapsto G(x, y, t_0)$ . On en déduit que la fonction  $G$  est continue.

Fixons  $z \in [a, b]$ . La fonction  $(x, t) \mapsto f(x, t)$  est continue et admet une dérivée partielle continue sur  $I \times [a, b]$ . Par intégration sur le segment d'extrémités  $t_0$  et  $z$ , le théorème de dérivation sous le signe intégral implique que

$$x \mapsto \int_{t_0}^z f(x, t) dt$$

est  $C^1$  de dérivée

$$x \mapsto \int_{t_0}^z \frac{\partial f}{\partial x}(x, t) dt.$$

De même, pour  $y$  fixé, on montre que la fonction

$$x \mapsto \int_y^{t_0} f(x, t) dt$$

est  $C^1$  de dérivée

$$x \mapsto \int_y^{t_0} \frac{\partial f}{\partial x}(x, t) dt.$$

Cela prouve que  $G$  admet une dérivée partielle par rapport à  $x$  qui est continue sur  $I \times [a, b] \times [a, b]$  et

$$\frac{\partial G}{\partial x}(x, y, z) = \int_y^z \frac{\partial f}{\partial x}(x, t) dt.$$

Par ailleurs, à  $x$  fixé, la fonction

$$z \mapsto \int_{t_0}^z f(x, t) dt$$

est une primitive de la fonction continue  $t \mapsto f(x, t)$ . Elle est donc  $C^1$  sur  $]a, b[$  de dérivée

$$z \mapsto f(x, z).$$

De même, la fonction

$$y \mapsto \int_y^{t_0} f(x, t) dt$$

est  $C^1$  sur  $]a, b[$  de dérivée

$$y \mapsto -f(x, y).$$

On en déduit que les dérivées partielles de  $G$  par rapport à  $y$  et  $z$  existent et sont continues. Récapitulons :

$$\frac{\partial G}{\partial x}(x, y, z) = \int_y^z \frac{\partial f}{\partial x}(x, t) dt, \quad \frac{\partial G}{\partial y}(x, y, z) = -f(x, y), \quad \frac{\partial G}{\partial z}(x, y, z) = f(x, z).$$

Par la règle de différentiation des fonctions composées, on en déduit que  $F$  est  $C^1$  et de dérivée

$$\begin{aligned} F'(x) &= \frac{\partial G}{\partial x}(x, g_1(x), g_2(x)) + \frac{\partial G}{\partial y}(x, g_1(x), g_2(x))g_1'(x) + \frac{\partial G}{\partial z}(x, g_1(x), g_2(x))g_2'(x) \\ &= \int_{g_1(x)}^{g_2(x)} \frac{\partial f}{\partial x}(x, t) dt - f(x, g_1(x))g_1'(x) + f(x, g_2(x))g_2'(x), \end{aligned}$$

ce qu'il fallait démontrer.

## 9.1.4 Différentielle totale exacte

On sait que toute fonction continue sur un intervalle de  $\mathbb{R}$  admet une primitive sur ce même intervalle. Est-ce que ce résultat admet une généralisation pour les fonctions de 2 variables ?

Plus précisément, on considère le problème suivant : soient  $f, g : \mathbb{R}^2 \rightarrow \mathbb{R}$  deux fonctions de classe  $C^1$ . A quelles conditions sur  $f$  et  $g$  existe-t-il une fonction  $H$  de classe  $C^1$  sur  $\mathbb{R}^2$  telle que

$$\frac{\partial H}{\partial x} = f, \quad \frac{\partial H}{\partial y} = g ?$$

Autrement dit, quand peut-on trouver une primitive de la fonction

$$(x, y) \in \mathbb{R}^2 \mapsto (f(x, y), g(x, y)) ?$$

Pour répondre à cette question, on procède par analyse et synthèse. Autrement dit, on suppose que  $H$  existe, et on déploie toutes les conséquences possibles sur  $f$  et  $g$  en espérant mettre la main sur des conditions suffisantes pour garantir l'existence de  $H$ .

**Etape d'analyse** Si  $H$  existe, alors ses dérivées partielles existent et sont  $C^1$  et donc  $H$  est  $C^2$ . Par le théorème de Schwarz, ses dérivées partielles secondes commutent, et donc

$$\frac{\partial f}{\partial y} = \frac{\partial^2 H}{\partial y \partial x} = \frac{\partial^2 H}{\partial x \partial y} = \frac{\partial g}{\partial x}$$

et on obtient donc une première condition nécessaire, qui porte sur  $f$  et  $g$  :

$$\frac{\partial f}{\partial y} = \frac{\partial g}{\partial x}.$$

Fixons  $(x, y) \in \mathbb{R}^2$ . Introduisons la fonction

$$h : t \mapsto H(tx, ty).$$

La fonction  $h$  est  $C^1$  par composition de fonctions  $C^1$  et de plus

$$h'(t) = x \frac{\partial H}{\partial x}(tx, ty) + y \frac{\partial H}{\partial y}(tx, ty) = xf(tx, ty) + yg(tx, ty).$$

En écrivant

$$h(t) - h(0) = \int_0^1 h'(s) ds,$$

on obtient donc

$$H(x, y) - H(0, 0) = \int_0^1 (xf(sx, sy) + yg(sx, sy)) ds,$$

ce qui constitue une expression explicite de la fonction  $H$  cherchée (à une constante additive près).

**Synthèse** Supposons que

$$\frac{\partial f}{\partial y} = \frac{\partial g}{\partial x} \tag{9.2}$$

et posons pour tout  $(x, y) \in \mathbb{R}^2$ ,

$$H(x, y) = \int_0^1 (xf(sx, sy) + yg(sx, sy)) ds.$$

Pour  $(x, y)$  fixé, la fonction  $s \mapsto xf(sx, sy) + yg(sx, sy)$  est continue sur  $\mathbb{R}$  donc le nombre  $H(x, y)$  (obtenu comme l'intégrale d'une fonction continue sur un segment) est bien défini.

Pour  $x$  fixé, la fonction  $\rho : (y, s) \mapsto xf(sx, sy) + yg(sx, sy)$  est continue sur  $\mathbb{R}^2$ , admet une dérivée partielle par rapport à  $y$  :

$$\frac{\partial \rho}{\partial y}(y, s) = sx \frac{\partial f}{\partial y}(sx, sy) + sy \frac{\partial g}{\partial y}(sx, sy) + g(sx, sy)$$

qui est une fonction continue sur  $\mathbb{R}^2$ . Par intégration sur le segment  $[0, 1]$ , le théorème de dérivation sous le signe intégral implique que  $H$  admet une dérivée partielle par rapport à  $y$  qui est continue par rapport à  $(x, y)$  et

$$\begin{aligned} \frac{\partial H}{\partial y}(x, y) &= \int_0^1 \frac{\partial \rho}{\partial y}(y, s) ds \\ &= \int_0^1 (sx \frac{\partial f}{\partial y}(sx, sy) + sy \frac{\partial g}{\partial y}(sx, sy) + g(sx, sy)) ds. \end{aligned}$$

En utilisant l'hypothèse (9.2), il vient

$$\frac{\partial H}{\partial y}(x, y) = \int_0^1 (sx \frac{\partial g}{\partial x}(sx, sy) + sy \frac{\partial g}{\partial y}(sx, sy) + g(sx, sy)) ds.$$

Or

$$x \frac{\partial g}{\partial x}(sx, sy) + y \frac{\partial g}{\partial y}(sx, sy) = \frac{\partial}{\partial s}(g(sx, sy)).$$

On en déduit

$$\frac{\partial H}{\partial y}(x, y) = \int_0^1 s \frac{\partial}{\partial s}(g(sx, sy)) + g(sx, sy) ds.$$

Par intégration par parties

$$\frac{\partial H}{\partial y}(x, y) = \left[ sg(sx, sy) \right]_0^1 - \int_0^1 g(sx, sy) ds + \int_0^1 g(sx, sy) ds = g(x, y).$$

On montre de même que la dérivée partielle de  $H$  par rapport à  $x$  existe et vaut

$$\frac{\partial H}{\partial x}(x, y) = f(x, y).$$

Ainsi la fonction  $H$  a bien pour dérivées partielles  $f$  et  $g$ . Comme il s'agit de fonctions continues,  $H$  est  $C^1$  sur  $\mathbb{R}^2$ . On peut écrire

$$DH(x, y) = f(x, y)dx + g(x, y)dy, \quad \forall (x, y) \in \mathbb{R}^2,$$

en notant  $dx : \mathbb{R}^2 \rightarrow \mathbb{R}$  la projection sur la première coordonnée, et  $dy : \mathbb{R}^2 \rightarrow \mathbb{R}$  la projection sur la deuxième coordonnée.

### 9.1.5 Théorème de Fubini

Soient  $[a, b]$  et  $[c, d]$  deux segments de  $\mathbb{R}$ . On considère une fonction  $f : [a, b] \times [c, d] \rightarrow \mathbb{K}$ .

**Théorème 9.9.** *On suppose que  $f$  est continue sur  $[a, b] \times [c, d]$ . Alors*

$$\int_a^b \left( \int_c^d f(x, t) dt \right) dx = \int_c^d \left( \int_a^b f(x, t) dx \right) dt.$$

Preuve : Par le théorème de continuité sous le signe intégral, la fonction  $x \mapsto \int_c^d f(x, t) dt$  est continue sur  $[a, b]$ . On peut donc considérer son intégrale sur  $[a, b]$  et le membre de gauche de l'identité à démontrer est donc bien définie. Il en est de même du membre de droite.

Introduisons la fonction  $F(x, y) = \int_c^y f(x, t) dt$ . Alors  $F$  est continue sur  $[a, b] \times [c, d]$ .

De plus, par le théorème fondamental du calcul différentiel,  $F$  admet une dérivée partielle par rapport à  $x$  :  $\frac{\partial F}{\partial x}(x, y) = f(x, y)$ , qui est continue sur  $[a, b] \times [c, d]$ .

Par le théorème de dérivation sous le signe intégral (cas où le domaine d'intégration est un segment, on n'a donc pas besoin de vérifier l'hypothèse de domination), on en déduit que la fonction

$$G : y \in [c, d] \mapsto \int_a^b F(x, y) dx$$

est bien définie et  $C^1$  sur  $[c, d]$ , de dérivée

$$G'(y) = \int_a^b \frac{\partial F}{\partial y}(x, y) dx = \int_a^b f(x, y) dx.$$

On peut à présent conclure. D'une part, par définition de  $G$ ,

$$G(d) - G(c) = G(d) = \int_a^b \int_c^d f(x, y) dy dx.$$

D'autre part,

$$G(d) - G(c) = \int_c^d G'(y) dy = \int_c^d \int_a^b f(x, y) dx dy.$$

Ainsi on a bien

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

□