

Analyse et calcul pour l'ingénieur-e (avancé)

Notes de cours

ENSIMAG 1A

Emmanuel Maître

Avertissement : Ce document a été rédigé à partir de notes de cours des années précédentes, et diverses sources (cf bibliographie). Les graphiques ont été construits à partir d'une première version générée en TikZ par IA. Des notes manuscrites ont été converties automatiquement en \LaTeX par IA. Des erreurs (que je mettrai sans hésiter sur le dos de l'IA) ont dû se glisser, vous pouvez me les signaler en écrivant à emmanuel.maitre@grenoble-inp.fr. Merci d'avance !

Un astérisque* signale un passage plus délicat et/ou qui peut être ignoré en première lecture.

Table des matières

1	Espaces métriques et normés, et point fixe	1
1.1	Motivations	1
1.2	Espaces métriques	2
1.2.1	Définition	2
1.2.2	Quelques notions de topologie	2
1.2.3	Exemples de métriques	3
1.2.4	Continuité d'une fonction	3
1.3	Compacité - premières notions	3
1.4	Cas particulier : Espaces Vectoriels Normés	5
1.4.1	Définitions et premières propriétés	5
1.4.2	Équivalence des normes	6
	Pour aller plus loin : Théorème de F. Riesz*	6
1.5	Espaces métriques complets, espaces de Banach	7
1.5.1	Définitions	7
1.5.2	Exemples	7
1.5.3	Applications linéaires continues	8
1.5.4	Séries dans les espaces de Banach	9
1.6	Théorème de point fixe contractant	10
1.6.1	Applications	10
1.6.2	Théorème de point fixe	10
1.7	Compacité et continuité	11
1.8	Meilleure approximation	11
1.9	Compacité dans les espaces de fonctions : Le Théorème d'Ascoli*	12
1.9.1	Équicontinuité	12
1.9.2	Le Théorème d'Ascoli-Arzelà	12
1.9.3	Démonstration	13
1.10	Quelques applications pratiques des espaces métriques	13
1.10.1	Application 1 : Codes correcteurs et Distance de Hamming	14
1.10.2	Application 2 : Analyse d'images et Transport Optimal	14
	Pour aller plus loin : quelques applications	15
	Pour aller plus loin : Flots de Gradient dans les Espaces Métriques*	16
2	Théorie de la mesure et Intégrale de Lebesgue	19
2.1	Espaces mesurables et Mesures	19
2.1.1	Tribus et Espaces mesurables	19
2.1.2	La Tribu borélienne	20
2.1.3	Mesures positives	20

TABLE DES MATIÈRES

2.1.4	Propriétés fondamentales	21
2.1.5	Ensembles négligeables	21
	Pour aller plus loin : Paradoxe de Banach-Tarski	21
	Pour aller plus loin : Le transport de mesure (Push-forward)*	22
2.2	Fonctions mesurables	22
2.2.1	Définition et caractérisations	22
2.2.2	Limites supérieure et inférieure*	23
2.2.3	Opérations sur les fonctions mesurables	24
2.2.4	Fonctions étagées et approximation	25
2.3	La mesure de Lebesgue	26
2.4	Construction de l'intégrale de Lebesgue	26
2.4.1	Étape 1 : Fonctions étagées positives	26
2.4.2	Étape 2 : Fonctions mesurables positives	27
2.4.3	Le Théorème de Convergence Monotone	28
2.4.4	Étape 3 : Fonctions mesurables de signe quelconque	31
2.5	L'espace des fonctions intégrables L^1	31
2.5.1	L'espace vectoriel \mathcal{L}^1	31
2.5.2	Le problème de la séparation et la relation « presque partout »	32
2.5.3	L'espace de Banach L^1	33
2.6	Les espaces L^p pour $1 < p \leq \infty$	34
2.6.1	Définitions et Normes pour $1 \leq p < \infty$	34
2.6.2	L'espace L^∞ (Fonctions essentiellement bornées)	34
2.6.3	Le cas particulier L^2 : Un espace de Hilbert	35
2.7	Les grands théorèmes	36
2.7.1	Le lemme de Fatou	36
2.7.2	Le Théorème de Convergence Dominée (TCD)	37
2.7.3	Variante : Le Théorème de Convergence Dominée généralisé*	38
2.7.4	Intégrales dépendant d'un paramètre	40
2.8	Mesures produits et Théorèmes de Fubini-Tonelli	42
2.8.1	La Mesure Produit	42
2.8.2	Le Théorème de Tonelli (Fonctions positives)	42
2.8.3	Le Théorème de Fubini (Fonctions intégrables)	43
2.8.4	Méthode pratique : Le Théorème de Fubini-Tonelli	44
2.8.5	Application : Le produit de convolution	44
2.9	Changement de variables dans \mathbb{R}^n	45
2.9.1	Difféomorphismes et résultat principal	46
2.9.2	Exemples	47
3	Calcul différentiel et Théorème des fonctions implicites	49
3.1	La notion de différentielle	49
3.1.1	Différentiabilité au sens de Fréchet	49
3.1.2	Lien avec les dérivées partielles	50
3.2	Opérations et calculs	52
3.2.1	Règles de calcul	52
3.2.2	Composition (Chain Rule)	52
3.2.3	Inégalité des Accroissements Finis (IAF)	53
3.3	Ordre supérieur : Théorème de Schwarz	53
	Pour aller plus loin : dérivées généralisées	55

3.4	Théorème des Fonctions Implicites	56
3.4.1	Interprétation et utilité	57
3.5	Extrema et Optimisation	57
	Pour aller plus loin : La Différentiation Automatique et l'IA	58
	Pour aller plus loin : Calcul des variations	60
	Pour aller plus loin : La méthode de l'État Adjoint	61
4	Espaces de Hilbert	65
4.1	Espaces préhilbertiens et de Hilbert	65
4.1.1	Produit scalaire	65
4.1.2	L'inégalité de Cauchy-Schwarz	66
4.1.3	Identités géométriques	66
4.1.4	Espaces de Hilbert	67
4.1.5	Exemples fondamentaux	67
4.2	Orthogonalité et Théorème de Projection	67
4.2.1	Orthogonalité et Pythagore	67
4.2.2	Projection sur un convexe fermé	68
4.2.3	Projection sur un sous-espace vectoriel fermé	68
4.2.4	Supplémentaire orthogonal	69
4.2.5	Application : La méthode des Moindres Carrés	70
4.3	Dualité et Théorème de représentation	71
4.3.1	Le Théorème de représentation de Riesz	71
4.3.2	Identification du dual	72
4.3.3	Application : L'opérateur adjoint	72
4.3.4	Convergence faible	73
4.4	Bases Hilbertiennes et Séries de Fourier	73
4.4.1	Familles orthonormales et Inégalité de Bessel	73
4.4.2	Définition d'une Base Hilbertienne	74
4.4.3	Exemples classiques	75
4.4.4	Exemple détaillé : Les Ondelettes de Haar	75
	Pour aller plus loin : Espaces à Noyau Reproduisant et Machine Learning	76
5	Fonctions périodiques et Séries de Fourier	79
5.1	Définitions et Coefficients de Fourier	79
5.1.1	Le cadre fonctionnel	79
5.1.2	Polynômes trigonométriques	80
5.1.3	Coefficients de Fourier (Forme exponentielle)	80
5.1.4	Coefficients de Fourier (Forme réelle)	80
5.1.5	Parité et Symétries	81
5.2	Convergence quadratique (Cadre L^2)	82
5.2.1	Approximation et Inégalité de Bessel	83
5.2.2	Théorème de Parseval-Plancherel	83
5.2.3	Convergence en moyenne quadratique	83
5.2.4	Exemple : Calcul de $\sum \frac{1}{n^2}$	84
5.3	Convergences Ponctuelle et Uniforme	84
5.3.1	Le Théorème de Dirichlet (Convergence ponctuelle)	84
5.3.2	Convergence Normale et Uniforme	86
5.4	Vitesse de convergence et Phénomène de Gibbs	86

TABLE DES MATIÈRES

5.4.1	Lien entre Régularité et Décroissance des coefficients	86
5.4.2	Le Phénomène de Gibbs	87
	Pour aller plus loin : Inégalité de Wirtinger	87
6	Méthodes numériques de base : Systèmes linéaires	91
6.1	Méthodes itératives pour des systèmes linéaires	91
6.1.1	Définitions et convergence	91
6.2	Méthodes par décomposition (Splitting)	92
6.2.1	Décomposition standard	92
6.2.2	La méthode de Jacobi	92
6.2.3	La méthode de Gauss-Seidel	93
6.2.4	Méthodes de relaxation (SOR)	93
6.3	Résultats de convergence	94
6.3.1	Matrices à diagonale dominante	94
6.3.2	Théorème d'Ostrowski (Matrices symétriques définies positives)	94
6.3.3	Comparaison des vitesses de convergence	95
6.4	Mise en œuvre pratique	95
6.4.1	Le résidu et le critère d'arrêt	95
6.4.2	Complexité algorithmique	96
6.5	Introduction aux méthodes de sous-espaces de Krylov	96
6.5.1	La méthode du gradient conjugué (CG)	96
7	Résolution numérique d'équations non linéaires	99
7.1	Exemples introductifs	99
7.1.1	Discretisation de l'équation de Poisson-Boltzmann	99
7.1.2	Modes de vibration d'un cristal	99
7.2	Méthode des approximations successives	100
7.2.1	Principe et définitions	100
7.3	La méthode de Newton	100
7.3.1	Construction de la fonction d'itération	100
7.3.2	Analyse de la convergence	101
7.4	La Méthode de Newton-Raphson	101
7.4.1	Construction de la méthode	101
7.4.2	Analyse de la convergence	102
7.4.3	Remarques sur la mise en œuvre	103
7.5	Variantes et méthodes Quasi-Newton	103
7.5.1	Approximation de la Jacobienne (Différences finies)	103
7.5.2	Méthode de Newton Modifiée (ou de la Corde)	104
7.5.3	Méthodes Quasi-Newton (Broyden)	104

Chapitre 1

Espaces métriques et normés, et point fixe

1.1 Motivations

- Un espace est un ensemble d'objets et une métrique sur cet espace est une fonction qui permet de mesurer l'écart entre deux objets. On appelle souvent **distance** cette métrique parce qu'elle généralise la distance qu'on connaît dans \mathbb{R}^3 : l'écart linéaire entre deux points.
- Pourtant, mesurer des distances n'est pas toujours aussi simple : cela dépend de la structure de l'espace. *Exemple* : Pour aller manger au RU (Restaurant Universitaire), la ligne droite n'est pas forcément possible, et la distance entre cette salle et le RU est plus complexe que la distance euclidienne.
- On aimerait aussi mesurer des distances entre des objets plus complexes que des points. Par exemple entre deux images. Comment peut-on définir une distance de sorte que l'image A soit plus proche de l'image B que de l'image C ?

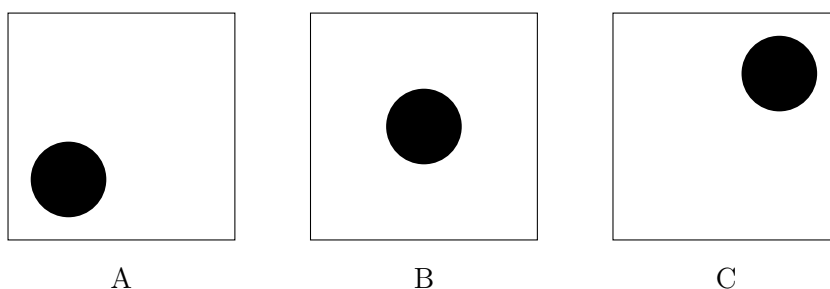


Figure 1.1 – Trois images comportant un disque disposé différemment.

- **Application** : La recherche d'image analogue à une image, c'est un sujet de recherche actuel de trouver des méthodes performantes pour comparer deux images, trouver le barycentre d'un ensemble d'images, etc... voir les travaux de Gabriel Peyré sur ces questions.
- L'objet de ce chapitre est d'extraire ce qu'il y a de fondamental dans une distance et un espace métrique et les propriétés qui en découlent, dans le but de les utiliser dans un cadre plus général.
- Comme propriétés minimales que devrait vérifier une distance, on peut penser à :
 - Être positive, et nulle si on mesure la distance entre un objet et lui-même.
 - Être symétrique : la distance d'ici au RU est la même que du RU à ici.

- La règle du détour (inégalité triangulaire) : la distance pour aller au RU via la laverie est plus grande (ou égale) que la distance d'ici au RU.

1.2 Espaces métriques

1.2.1 Définition

Soit E un ensemble quelconque. On dit que $d : E \times E \rightarrow \mathbb{R}^+$ est une **distance** sur E si elle vérifie :

(D1) $d(x, y) = 0 \iff x = y$ (Séparation)

(D2) $d(x, y) = d(y, x)$ (Symétrie)

(D3) $d(x, y) \leq d(x, z) + d(y, z)$ pour tous $x, y, z \in E$ (Inégalité triangulaire)

Muni de cette distance, E est appelé **espace métrique**, on note souvent (E, d) .

1.2.2 Quelques notions de topologie

Définition 1.1 (Convergence). Soit $(x_n)_{n \in \mathbb{N}}$ une suite d'éléments de E . On dit que (x_n) converge vers $x \in E$ si la suite réelle $(d(x_n, x))_{n \in \mathbb{N}}$ converge vers 0 dans \mathbb{R} . On note : $x_n \xrightarrow[n \rightarrow +\infty]{} x \iff d(x_n, x) \xrightarrow[n \rightarrow +\infty]{} 0$.

Définition 1.2 (Boules). Soit (E, d) un espace métrique, $x_0 \in E$ et $r > 0$. On appelle :

- **Boule ouverte** de centre x_0 et de rayon r l'ensemble $B(x_0, r) = \{x \in E \mid d(x, x_0) < r\}$.
- **Boule fermée** de centre x_0 et de rayon r l'ensemble $\bar{B}(x_0, r) = \{x \in E \mid d(x, x_0) \leq r\}$.

Définition 1.3 (Ouverts et Fermés). — Une partie $A \subset E$ est **ouverte** si $\forall x \in A, \exists r > 0, B(x, r) \subset A$.

- Une partie $B \subset E$ est **fermée** si son complémentaire $B^c = E \setminus B$ est ouvert.

Définition 1.4 (Voisinage). Soit A une partie de E , V est un voisinage de A s'il existe un ouvert U de E tel que $A \subset U \subset V$. Soit $a \in E$, V est un voisinage de a s'il existe un ouvert U contenant a tel que $U \subset V$ (cas $A = \{a\}$).

Proposition 1.1 (Caractérisation séquentielle des fermés). Une partie $B \subset E$ est fermée si et seulement si toute suite convergente d'éléments de B converge vers un élément de B . C'est-à-dire : $(x_n) \subset B$ et $x_n \rightarrow x \implies x \in B$.

Démonstration. **Sens \Rightarrow :** Soit B fermé. Si $x \notin B$, alors $x \in B^c$ qui est ouvert. Donc $\exists r > 0, B(x, r) \subset B^c$. Mais comme $x_n \rightarrow x$, pour n assez grand ($n \geq N$), $d(x_n, x) < r$. En particulier $x_N \in B(x, r) \subset B^c$, ce qui contredit $x_N \in B$. **Sens \Leftarrow :** Supposons que B contient ses limites de suites. Supposons B non fermé, donc B^c non ouvert. Il existe $x \in B^c$ tel que pour tout $r > 0$, $B(x, r) \not\subset B^c$, c'est-à-dire $B(x, r) \cap B \neq \emptyset$. Prenons $r = 1/n$. Il existe $x_n \in B(x, 1/n) \cap B$. On a construit une suite (x_n) d'éléments de B telle que $d(x_n, x) < 1/n \rightarrow 0$. Donc $x_n \rightarrow x$. Par hypothèse, $x \in B$, ce qui contredit $x \in B^c$. \square

Définition 1.5 (Adhérence). Soit $A \subset E$. On appelle adhérence (ou fermeture) de A dans E , notée \bar{A} , de manière équivalente :

- L'intersection des fermés contenant A .
- Le plus petit fermé contenant A .
- L'ensemble des limites de suites d'éléments de A .

Proposition 1.2. A est fermé $\iff A = \bar{A}$.

1.2.3 Exemples de métriques

1. \mathbb{R} muni de la distance $d(x, y) = |x - y|$. Les boules ouvertes sont les intervalles ouverts $]x - r, x + r[$.
2. Soit $\mathcal{F}_b(X, Y)$ l'ensemble des fonctions f d'un ensemble X dans un espace métrique (Y, δ) qui sont bornées (au sens où l'image $f(X)$ est incluse dans une boule $B(y_f, R_f)$). On peut munir cet ensemble de la distance de la convergence uniforme :

$$d(f, g) = \sup_{x \in X} \delta(f(x), g(x)).$$

3. **Distance triviale (ou discrète)** : Sur un ensemble E non vide, on définit : $d(x, y) = 0$ si $x = y$ et $d(x, y) = 1$ si $x \neq y$. Alors $B(x, 1/2) = \{x\} = \bar{B}(x, 1/2) = B(x, 1)$ et $\bar{B}(x, 1) = B(x, 2) = E$.

1.2.4 Continuité d'une fonction

Soient (E, d) et (E', d') deux espaces métriques. Une application $f : E \rightarrow E'$ est continue en $x_0 \in E$ si l'une des assertions suivantes (équivalentes) est vérifiée :

- (i) Pour tout voisinage W de $f(x_0)$ dans E' , il existe un voisinage V de x_0 dans E tel que $f(V) \subset W$.
- (ii) Pour tout voisinage W de $f(x_0)$ dans E' , $f^{-1}(W)$ est un voisinage de x_0 dans E .
- (iii) $\forall \varepsilon > 0, \exists \eta > 0, \forall x \in E, d(x, x_0) \leq \eta \implies d'(f(x), f(x_0)) \leq \varepsilon$.
- (iv) Pour toute suite $(x_n) \subset E$ telle que $x_n \rightarrow x_0$, alors $f(x_n) \rightarrow f(x_0)$.

Exemple 1.1. Si (E, d) est un espace métrique, alors pour $y \in E$, $x \rightarrow d(x, y)$ est continue de E dans \mathbb{R} .

1.3 Compacité - premières notions

Définition 1.6. Soit (E, d) un espace métrique. Une partie $K \subset E$ est **compacte** si de toute suite $(x_n)_{n \in \mathbb{N}}$ d'éléments de K , on peut extraire une sous-suite $(x_{\varphi(n)})$ qui converge vers une limite $x \in K$.

Dans un espace métrique, la définition séquentielle ci-dessus (Bolzano-Weierstrass) et la définition par recouvrement (Borel-Lebesgue) sont équivalentes.

Théorème 1.1 (Équivalence des compacités*). Soit (E, d) un espace métrique et $K \subset E$. Les deux propriétés suivantes sont équivalentes :

1. **Propriété de Bolzano-Weierstrass** : De toute suite $(x_n)_{n \in \mathbb{N}}$ de K , on peut extraire une sous-suite convergente dans K .
2. **Propriété de Borel-Lebesgue** : De tout recouvrement ouvert de K , on peut extraire un sous-recouvrement fini.

Démonstration. On se place dans l'espace métrique (K, d) , dont on rappelle que les ouverts sont les intersections des ouverts de E avec K .

Sens 2 \implies 1. Supposons que K vérifie la propriété de Borel-Lebesgue. Soit $(x_n)_{n \in \mathbb{N}}$ une suite de points de K , prenant un nombre infini de valeurs distinctes (sinon elle a *de facto* au moins une sous-suite convergente). Raisonnons par l'absurde en supposant que cette suite n'admet aucune sous-suite convergente dans K .

Cela implique que l'ensemble des valeurs de la suite, $A = \{x_n \mid n \in \mathbb{N}\}$, ne possède aucun point d'accumulation dans K . Autrement dit, pour tout $y \in K$, il existe un voisinage ouvert U_y de y qui ne contient qu'un nombre **fini** de termes de la suite (x_n) (sinon, y serait une valeur d'adhérence).

La famille $\mathcal{U} = (U_y)_{y \in K}$ constitue un recouvrement ouvert de K . D'après l'hypothèse de Borel-Lebesgue, on peut en extraire un sous-recouvrement fini :

$$K = \bigcup_{k=1}^p U_{y_k}.$$

Or, chaque ouvert U_{y_k} ne contient qu'un nombre fini d'indices n . Leur réunion finie ne peut donc contenir qu'un nombre fini de termes de la suite. Ceci est contradictoire avec le fait que la suite (x_n) a une infinité de termes distincts. L'hypothèse absurde est fausse : la suite admet nécessairement une sous-suite convergente.

Sens 1 \implies 2. Supposons que K est séquentiellement compact. La démonstration repose sur deux résultats intermédiaires.

Étape A : Le Lemme du nombre de Lebesgue. Soit $(O_i)_{i \in I}$ un recouvrement ouvert de K . Montrons qu'il existe $\delta > 0$ tel que toute boule de rayon δ soit incluse dans l'un des O_i . Par l'absurde, si ce n'est pas le cas : $\forall n \in \mathbb{N}^*$, il existe une boule $B(x_n, 1/n)$ qui n'est incluse dans aucun O_i . La suite (x_n) admet une sous-suite $(x_{\varphi(n)})$ convergeant vers $\ell \in K$. Puisque (O_i) recouvre K , $\exists i_0 \in I$ tel que $\ell \in O_{i_0}$. Comme O_{i_0} est ouvert, $\exists \varepsilon > 0, B(\ell, \varepsilon) \subset O_{i_0}$. Pour n assez grand, on a $d(x_{\varphi(n)}, \ell) < \varepsilon/2$ et $1/\varphi(n) < \varepsilon/2$. Par inégalité triangulaire, $B(x_{\varphi(n)}, 1/\varphi(n)) \subset B(\ell, \varepsilon) \subset O_{i_0}$, ce qui contredit l'hypothèse de départ. Le nombre (de Lebesgue) δ existe.

Étape B : Précompacité. Montrons que pour tout $\varepsilon > 0$, K peut être recouvert par un nombre fini de boules de rayon ε . Par l'absurde, supposons qu'il faille une infinité de boules. Soit x_0 quelconque, $B(x_0, \varepsilon)$ ne recouvre donc pas K . Prenons $x_1 \in K$ tel que $d(x_0, x_1) \geq \varepsilon$. $B(x_0, \varepsilon) \cup B(x_1, \varepsilon)$ ne recouvre pas K , donc il existe $x_2 \in K$ tel que $d(x_0, x_2) \geq \varepsilon$ et $d(x_1, x_2) \geq \varepsilon$. En procédant par récurrence, on obtient une suite (x_n) vérifiant $d(x_n, x_m) \geq \varepsilon$ pour $n \neq m$. Une telle suite ne peut pas avoir de sous-suite de Cauchy, donc pas de sous-suite convergente, ce qui contredit l'hypothèse de Bolzano-Weierstrass.

Étape C : Conclusion. Soit $(O_i)_{i \in I}$ un recouvrement ouvert quelconque de K .

1. D'après l'étape A, il existe un nombre de Lebesgue $\delta > 0$.
2. D'après l'étape B, on peut recouvrir E par un nombre fini de boules de rayon δ : $E \subset \bigcup_{k=1}^m B(z_k, \delta)$.
3. Pour chaque k , la boule $B(z_k, \delta)$ est incluse dans un certain ouvert O_{i_k} du recouvrement initial.

La famille finie $\{O_{i_1}, \dots, O_{i_m}\}$ recouvre donc E . □

Exemples :

- Dans \mathbb{R} : les segments $[a, b]$ sont compacts (Bolzano-Weierstrass).
- Dans \mathbb{R}^N : les pavés $\prod [a_i, b_i]$.
- Dans $\mathcal{C}^0([0, 1], \mathbb{R})$ muni de la norme infinie, la boule unité fermée n'est **pas** compacte. Contre-exemple : Soit la suite de fonction continues et affines par morceaux définie par $f_n(x) = 0$ en dehors de $]\frac{1}{n}, \frac{1}{n-1}[$ et $f_n(x) = 1$ au milieu de cet intervalle (voir figure 1.2). On a $f_n \in \bar{B}(0, 1)$ car $\|f\|_\infty = 1$, et $d(f_n, f_m) = \|f_n - f_m\|_\infty = 1$ pour $n \neq m$. Donc aucune sous-suite de (f_n) ne converge.

Définition 1.7 (Diamètre et Borné). Le diamètre d'un sous-ensemble $A \subset E$ est

$$\delta(A) = \sup_{(x,y) \in A^2} d(x, y) \in [0, +\infty].$$

On dit que A est bornée si $\delta(A) < +\infty$.

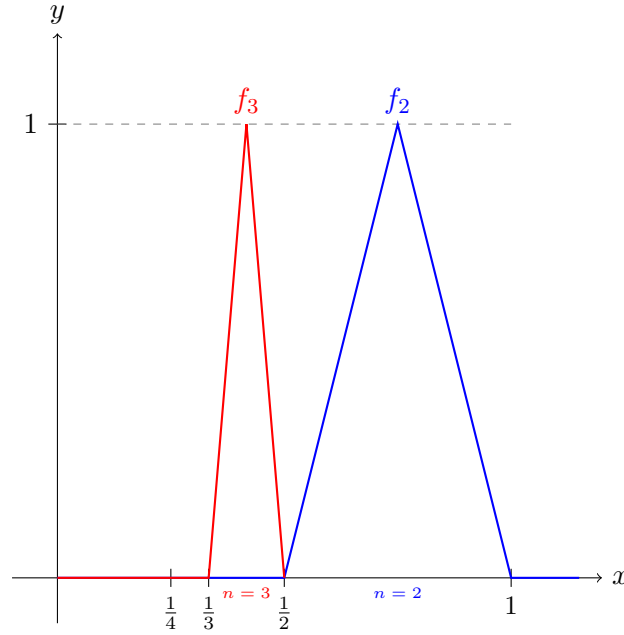


Figure 1.2 – Représentation des fonctions f_2 (bleu) et f_3 (rouge).

Proposition 1.3. Si A est compact, alors A est fermé et borné.

Démonstration. **Borné :** Par l'absurde, si $\delta(A) = +\infty$, on peut construire une suite (x_n, y_n) telle que $d(x_n, y_n) \rightarrow +\infty$. Quitte à extraire une sous-suite convergente on a, $x_{\varphi(n)} \rightarrow x$ et $y_{\varphi(n)} \rightarrow y$, donc $d(x_{\varphi(n)}, y_{\varphi(n)}) \rightarrow d(x, y) < +\infty$, absurde. **Fermé :** Si $(x_n) \subset A$ converge vers x , comme A est compact, une sous-suite converge dans A . Par unicité de la limite, $x \in A$. \square

Remarque 1.1. La réciproque (Fermé + Borné \implies Compact) est vraie dans \mathbb{R}^n mais **fausse** en général, notamment dans les EVN de dimension infinie (cf Théorème de Riesz).

1.4 Cas particulier : Espaces Vectoriels Normés

1.4.1 Définitions et premières propriétés

Soit E un espace vectoriel sur le corps \mathbb{K} (où $\mathbb{K} = \mathbb{R}$ ou \mathbb{C}).

Définition 1.8 (Norme). Une application $N : E \rightarrow \mathbb{R}_+$ est appelée une **norme** sur E si elle vérifie les trois propriétés suivantes pour tout $x, y \in E$ et tout $\lambda \in \mathbb{K}$:

1. **Séparation :** $N(x) = 0 \iff x = 0_E$.
2. **Homogénéité :** $N(\lambda x) = |\lambda|N(x)$.
3. **Inégalité triangulaire :** $N(x + y) \leq N(x) + N(y)$.

On note généralement $\|x\| = N(x)$. La paire $(E, \|\cdot\|)$ est un **Espace Vectoriel Normé (EVN)**.

Remarque 1.2. Si la propriété de séparation est seulement $\|x\| = 0 \Leftarrow x = 0$, on parle de semi-norme.

Les EVN sont un cas particuliers d'espaces métriques, avec comme distance $d(x, y) = \|x - y\|$. Tout ce qui a été énoncé ci-dessus peut donc se réécrire sous forme de norme.

Remarque 1.3. L'inégalité triangulaire implique l'inégalité triangulaire inversée : $|\|x\| - \|y\|| \leq \|x - y\|$. Cela montre en particulier que l'application norme $x \mapsto \|x\|$ est continue (car 1-lipschitzienne).

Exemple 1.2 (Normes usuelles sur \mathbb{R}^n). Pour $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, on définit classiquement :

- $\|x\|_1 = \sum_{i=1}^n |x_i|$ (norme de "Manhattan").
- $\|x\|_2 = (\sum_{i=1}^n |x_i|^2)^{1/2}$ (norme euclidienne).
- $\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$ (norme infinie ou du max).

Exemple 1.3 (Espaces de fonctions). Soit X un ensemble compact (par exemple un intervalle fermé borné $[a, b]$ de \mathbb{R}). On note $\mathcal{C}(X, \mathbb{K})$ l'espace vectoriel des fonctions continues de X dans \mathbb{K} . La norme la plus naturelle sur cet espace est la norme uniforme :

$$\|f\|_\infty = \sup_{t \in X} |f(t)|.$$

1.4.2 Équivalence des normes

La notion de convergence dépend a priori du choix de la norme. Cependant, dans certains cas, le choix de celle-ci n'influe pas sur la topologie (les ouverts, les limites).

Définition 1.9. Deux normes $\|\cdot\|_a$ et $\|\cdot\|_b$ sur un même espace vectoriel E sont dites **équivalentes** s'il existe deux constantes $c > 0$ et $C > 0$ telles que :

$$\forall x \in E, \quad c\|x\|_a \leq \|x\|_b \leq C\|x\|_a.$$

Si deux normes sont équivalentes, une suite converge pour l'une si et seulement si elle converge pour l'autre. Le résultat suivant est fondamental pour l'analyse numérique matricielle. Une démonstration se base sur la compacité de la boule unité fermée, voir 1.7.

Théorème 1.2 (Équivalence en dimension finie). Si E est un espace vectoriel de **dimension finie**, alors toutes les normes sur E sont équivalentes.

Remarque 1.4 (Attention en dimension infinie). Ce résultat est **faux** en dimension infinie. Par exemple, sur l'espace $\mathcal{C}([0, 1], \mathbb{R})$, la suite de fonctions $f_n(t) = t^n$ converge vers 0 pour la norme L^1 (l'aire sous la courbe tend vers 0), mais ne converge pas vers 0 pour la norme $\|\cdot\|_\infty$ (la valeur en 1 reste 1).

Remarque 1.5 (Le piège de la dimension infinie : Compacité). En dimension finie, les fermés bornés sont compacts (Théorème de Bolzano-Weierstrass). Ce n'est plus vrai dans un espace vectoriel normé de dimension infinie. **Théorème de Riesz :** La boule unité fermée $\bar{B}(0, 1) = \{x \in E \mid \|x\| \leq 1\}$ est compacte si et seulement si E est de dimension finie. C'est pourquoi, en analyse fonctionnelle, on ne peut pas simplement dire « la suite est bornée donc elle admet une sous-suite convergente ». Il faut souvent des outils plus puissants (convergence faible, etc.).



Pour aller plus loin : Théorème de F. Riesz*

Théorème 1.3. Soit E un E.V.N. Si la boule unité fermée est compacte, E est de dimension finie.

Démonstration.

Supposons $\bar{B}(0, 1)$ compacte. On a $\bar{B}(0, 1) \subset \bigcup_{x \in \bar{B}(0, 1)} B(x, 1)$. D'après la caractérisation de la compacité par la propriété de Borel-Lebesgue (théorème 1.1), on peut en extraire un recouvrement

fini :

$$\exists p \in \mathbb{N}, x_1, \dots, x_p \in \bar{B}(0, 1), \bar{B}(0, 1) \subset \bigcup_{j=1}^p B(x_j, 1)$$

Soit F l'espace vectoriel engendré par x_1, \dots, x_p . F est donc de dimension finie. Montrons que $E=F$. Sinon, soit $x_0 \in E \setminus F$. Comme F est fermé (car de dimension finie), il existe $y_0 \in F$ tel que $\|x_0 - y_0\| = \min_{y \in F} \|x_0 - y\|$ (projection sur un SEV de dimension finie). On a $d = \|x_0 - y_0\| > 0$, sinon $x_0 \in F$. Soit $z_0 = \frac{y_0 - x_0}{d}$, qui est de norme 1, donc appartient à $\bar{B}(0, 1)$. Pourtant, pour tout $z \in F$,

$$\|z_0 - z\| = \left\| \frac{y_0 - x_0}{d} - z \right\| = \left\| \frac{y_0 - dz - x_0}{d} \right\| \geq \frac{d}{d} = 1$$

car $y_0 - dz \in F$ et que tout élément de F est au moins à la distance d de x_0 . En particulier, $z_0 \notin B(x_j, 1)$ pour tout j , donc $z_0 \notin \bigcup_{j=1}^p B(x_j, 1)$. Contradiction. \square

1.5 Espaces métriques complets, espaces de Banach

En analyse, une technique très courante pour analyser un problème complexe est de le remplacer par une suite de problèmes plus simples qui convergent vers le problème compliqué. Imaginons qu'on souhaite résoudre $P(x) = 0$ qui est une équation très complexe, mais qu'on puisse approcher cette équation par une suite de problèmes plus simples (par exemple plus réguliers) $P_n(x) = 0$ dont on connaît des solutions x_n . On espère que la suite (x_n) converge vers x_* qui sera solution du problème difficile. Mais comment prouver cette convergence, puisqu'on ne connaît pas x_* , dont on cherche à prouver l'existence même ? Le critère de Cauchy est une manière d'étudier le comportement lorsque $n \rightarrow \infty$ d'une suite, sans connaître son éventuelle limite. C'est l'outil qu'il nous faut ! C'est une notion en général plus faible que la convergence, mais pour certains espaces, qui sont l'objet de cette section, c'est équivalent !

1.5.1 Définitions

Soit (E, d) un espace métrique.

Définition 1.10 (Suite de Cauchy). Une suite $(x_n)_{n \in \mathbb{N}}$ de E est dite de **Cauchy** si :

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall p, q \geq N, \quad d(x_p, x_q) \leq \varepsilon.$$

Proposition 1.4. Toute suite convergente est de Cauchy.

Démonstration. Si $x_n \rightarrow x$, alors $\forall \varepsilon > 0, \exists N, n \geq N \implies d(x_n, x) \leq \varepsilon/2$. Donc pour $p, q \geq N$, $d(x_p, x_q) \leq d(x_p, x) + d(x, x_q) \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$. \square

Définition 1.11 (Espace Complet). Un espace métrique est dit **complet** si toute suite de Cauchy est convergente. Un espace vectoriel normé complet est appelé **espace de Banach**.

1.5.2 Exemples

- \mathbb{R} est complet (pour la distance usuelle). En effet toute suite de Cauchy est bornée (exercice facile, fixer un des indices dans la définition), et toute suite bornée dans \mathbb{R} admet une sous-suite convergente (Bolzano-Weierstrass). Enfin toute suite de Cauchy ayant une sous-suite convergente est convergente (exercice facile aussi, dès que certains termes aussi loin qu'on veut restent à proximité d'un point fixe, comme tous les termes sont proches lorsque $n \rightarrow \infty$, ils restent proches de ce point).

- Tout espace vectoriel normé de **dimension finie** est un espace de Banach.
- Soit $E = C^0([a, b], \mathbb{R})$ muni de la norme uniforme $\|f\|_\infty = \sup_{t \in [a, b]} |f(t)|$. Alors $(E, \|\cdot\|_\infty)$ est un **espace de Banach**.

Démonstration. Soit (f_n) une suite de Cauchy dans E . $\forall \varepsilon > 0, \exists N, \forall p, q \geq N, \|f_p - f_q\|_\infty \leq \varepsilon$. Pour tout $x \in [a, b]$ fixé, $|f_p(x) - f_q(x)| \leq \|f_p - f_q\|_\infty \leq \varepsilon$. La suite réelle $(f_n(x))_n$ est de Cauchy dans \mathbb{R} complet, donc converge vers une limite notée $f(x)$. En passant à la limite quand $q \rightarrow +\infty$ dans l'inégalité, on obtient : $\forall p \geq N, \forall x, |f_p(x) - f(x)| \leq \varepsilon$. Ceci signifie que (f_n) converge uniformément vers f . Comme la limite uniforme d'une suite de fonctions continues est continue, $f \in E$. Donc E est complet. \square

- Soit $E = C^0([a, b], \mathbb{R})$ muni de la norme $L^1 : \|f\|_1 = \int_a^b |f(t)| dt$. Cet espace n'est **pas** complet. On peut construire une suite de fonctions continues (type affine par morceaux) qui converge en norme L^1 vers une fonction discontinue (en escalier), qui n'est pas dans E .

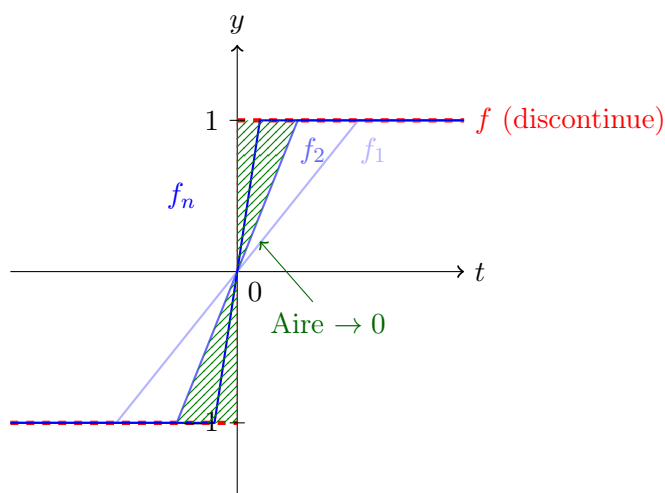


Figure 1.3 – Illustration de la non-complétude de $(C^0, \|\cdot\|_1)$. La suite de fonctions continues (f_n) (en bleu) converge en norme L^1 (l'aire hachurée tend vers 0) vers une fonction crêteau f (en rouge) qui est discontinue, donc hors de l'espace.

Proposition 1.5. — Si $F \subset E$ est fermé et E complet, alors (F, d) est complet.

- Si $F \subset E$ est complet, alors F est fermé dans E .

Démonstration. C'est un bon exercice! \square

Proposition 1.6. Soit (E, d) métrique.

1. Si $A \subset E$ est compact, alors (A, d) est complet.
2. Si $F \subset A$ avec A compact et F fermé, alors F est compact. (Un fermé d'un compact est compact).

1.5.3 Applications linéaires continues

En dimension infinie, une application linéaire n'est pas forcément continue.

Théorème 1.4 (Caractérisation de la continuité). Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux EVN et $L : E \rightarrow F$ une application linéaire. Les propriétés suivantes sont équivalentes :

1. L est continue sur E .
2. L est continue en 0.
3. L est Lipschitzienne, c'est-à-dire pour une application linéaire : $\exists K \geq 0, \forall x \in E, \|L(x)\|_F \leq K\|x\|_E$.

L'ensemble des applications linéaires continues de E dans F est noté $\mathcal{L}(E, F)$. On le munit de la **norme subordonnée** (ou **norme d'opérateur**) :

$$\|L\| = \sup_{x \neq 0} \frac{\|L(x)\|_F}{\|x\|_E} = \sup_{\|x\|_E \leq 1} \|L(x)\|_F.$$

Lorsque $F = \mathbb{R}$, on note $E' = \mathcal{L}(E, \mathbb{R})$, qu'on appelle le **dual topologique** de E .

Proposition 1.7 (Complétude de l'espace des applications linéaires continues). *Si F est un espace de Banach, alors $(\mathcal{L}(E, F), \|\cdot\|)$ est un espace de Banach (même si E ne l'est pas). En particulier, comme \mathbb{R} est complet, le dual topologique E' est complet pour tout EVN E . Ceci est à la base des notions de convergence faible qui permettent d'extraire des sous-suites faiblement convergentes de suites simplement bornées. Voir [15] pour plus de détail sur la notion de convergence faible.*

1.5.4 Séries dans les espaces de Banach

La complétude est très pratique pour étudier la convergence des séries.

Définition 1.12 (Convergence normale). Soit (x_n) une suite d'éléments d'un EVN E . On dit que la série $\sum x_n$ est **normalement convergente** si la série numérique des normes $\sum \|x_n\|$ est convergente.

Proposition 1.8. *Dans un espace de Banach, toute série normalement convergente est convergente.*

Cette proposition est souvent utilisée pour définir des fonctions d'opérateurs, comme l'exponentielle de matrice : $e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}$.

Démonstration. Soit $(x_n)_{n \in \mathbb{N}}$ une suite d'éléments de E telle que la série numérique $\sum \|x_n\|$ converge. Notons $S_n = \sum_{k=0}^n x_k$ la suite des sommes partielles de la série vectorielle. Puisque E est un espace de Banach (donc complet), pour montrer que la suite (S_n) converge, il suffit de montrer qu'elle est de Cauchy.

Soient p, q deux entiers tels que $q > p$. Calculons la norme de la différence des sommes partielles :

$$\|S_q - S_p\| = \left\| \sum_{k=0}^q x_k - \sum_{k=0}^p x_k \right\| = \left\| \sum_{k=p+1}^q x_k \right\|.$$

D'après l'inégalité triangulaire (étendue par récurrence à une somme finie), on a :

$$\left\| \sum_{k=p+1}^q x_k \right\| \leq \sum_{k=p+1}^q \|x_k\|.$$

Considérons maintenant la suite des sommes partielles de la série des normes : $\sigma_n = \sum_{k=0}^n \|x_k\|$. Par hypothèse, la série $\sum \|x_n\|$ converge dans \mathbb{R} , donc la suite (σ_n) converge. Donc (σ_n) est une suite de Cauchy. Ainsi, pour tout $\varepsilon > 0$, il existe un entier N tel que pour tout $q > p \geq N$:

$$|\sigma_q - \sigma_p| = \sum_{k=p+1}^q \|x_k\| < \varepsilon.$$

En combinant les inégalités, on obtient :

$$\forall q > p \geq N, \quad \|S_q - S_p\| \leq \sum_{k=p+1}^q \|x_k\| < \varepsilon.$$

La suite (S_n) est donc de Cauchy dans l'espace complet E , elle converge donc vers une limite $S \in E$. \square

Proposition 1.9. Soit $(E, \|\cdot\|)$ un espace de Banach, et $\text{id}_E \in \mathcal{L}(E)$ l'identité sur E . Alors pour tout $u \in \mathcal{L}(E)$ telle que $\|u\| < 1$, l'application $\text{id}_E - u$ est inversible.

Démonstration. En effet, soit $u \in \mathcal{L}(E)$ tel que $\|u\| < 1$. Soit la série $S = \sum_{k=0}^{+\infty} u^k$. Alors S est convergente car normalement convergente : si $S_n = \sum_{k=0}^n u^k$, on a $\|S_n\| \leq \sum_{k=0}^n \|u\|^k$ qui est une série convergente. Or $(\text{id}_E - u) \circ S_n = S_n \circ (\text{id}_E - u) = \text{id}_E - u^{n+1}$ donne à la limite lorsque $n \rightarrow +\infty$: $(\text{id}_E - u) \circ S = S \circ (\text{id}_E - u) = \text{id}_E$. \square

Remarque 1.6. Donc l'ensemble des endomorphismes bijectifs (ou automorphismes) de E est ouvert dans $\mathcal{L}(E)$. En effet d'après le résultat précédent, $u \in B(\text{id}_E, 1)$ est inversible puisque $\|\text{id}_E - u\| < 1$ implique $\text{id}_E - (\text{id}_E - u) = u$ inversible. Sur cet ouvert, l'application qui à un automorphisme associe son inverse est continue (cf exercice feuille de TD).

1.6 Théorème de point fixe contractant

Problème : Soit (E, d) un espace métrique, $T : E \rightarrow E$. Existe-t-il $x \in E$ tel que $T(x) = x$? Si oui, x est appelé un **point fixe** de T .

1.6.1 Applications

- Résolution de systèmes linéaires par méthode itérative.** Soit $A \in \mathcal{M}_n(\mathbb{R})$ et $b \in \mathbb{R}^n$. On veut résoudre $Ax = b$. Principe : Construire (x_n) avec $x_n \rightarrow x$. On décompose A sous la forme $A = M - N$ avec M facile à inverser (ex. diagonale). Alors $Ax = b \iff (M - N)x = b \iff Mx = Nx + b \iff x = M^{-1}Nx + M^{-1}b$. Soit $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ défini par $T(x) = M^{-1}Nx + M^{-1}b$. On cherche un point fixe de T . Algorithme : $x_{n+1} = T(x_n)$ en résolvant $Mx_{n+1} = Nx_n + b$. Question : quelles conditions sur M et N pour que cela converge ?
- Résolution d'équations intégrales / différentielles.** $y' = g(x, y(x))$ avec $y(a) = \alpha$. Ceci peut se réécrire sous forme intégrale : $y(x) = \alpha + \int_a^x g(t, y(t))dt$. C'est un problème de point fixe pour l'opérateur $T(y)(x) = \alpha + \int_a^x g(t, y(t))dt$.

1.6.2 Théorème de point fixe

Définition 1.13 (Application Lipschitzienne). Soit (E, d) un espace métrique, une application $f : A \subset E \rightarrow E$ est dite **k -lipschitzienne** (avec $k \geq 0$) si :

$$\forall x, y \in A, \quad d(f(x), f(y)) \leq kd(x, y).$$

Une application lipschitzienne est toujours uniformément continue. Si $k < 1$ l'application f est dite **contractante**.

Théorème 1.5 (Point fixe de Banach-Picard). Soit (E, d) un espace métrique **complet**, et $T : E \rightarrow E$ une application **contractante**. Alors :

1. Il existe un **unique** point fixe x^* de T dans E .
2. Pour tout $x_0 \in E$, la suite définie par $x_{n+1} = T(x_n)$ converge vers x^* .
3. On a l'estimation d'erreur : $d(x_n, x^*) \leq \frac{k^n}{1-k} d(x_1, x_0)$.

Démonstration. Soit $x_0 \in E$ et $x_{n+1} = T(x_n)$. Montrons que (x_n) est de Cauchy. $d(x_{n+1}, x_n) = d(T(x_n), T(x_{n-1})) \leq k d(x_n, x_{n-1}) \leq \dots \leq k^n d(x_1, x_0)$. Pour $p > 0$:

$$d(x_{n+p}, x_n) \leq \sum_{i=0}^{p-1} d(x_{n+i+1}, x_{n+i}) \leq \sum_{i=0}^{p-1} k^{n+i} d(x_1, x_0)$$

$$d(x_{n+p}, x_n) \leq k^n d(x_1, x_0) \sum_{i=0}^{p-1} k^i \leq k^n d(x_1, x_0) \frac{1}{1-k}.$$

Comme $k < 1$, $k^n \rightarrow 0$, donc (x_n) est de Cauchy. E complet $\implies x_n \rightarrow x^*$. T est contractante donc continue : $T(x_n) \rightarrow T(x^*)$. À la limite dans $x_{n+1} = T(x_n)$, on a $x^* = T(x^*)$. L'unicité vient de la contraction : si $T(x) = x$ et $T(y) = y$, alors $d(x, y) = d(T(x), T(y)) \leq k d(x, y)$. Comme $k < 1$, $d(x, y) = 0$ donc $x = y$. \square

1.7 Compacité et continuité

Proposition 1.10 (L'image continue d'un compact est compacte). Soient (E, d) et (E', d') deux espaces métriques et $f : E \rightarrow E'$ continue. Si $A \subset E$ est compact, alors l'image $f(A)$ est compacte dans E' .

Démonstration. Soit (y_n) une suite dans $f(A)$. Alors $y_n = f(x_n)$ avec $x_n \in A$. Comme A est compact, on peut extraire $x_{\varphi(n)} \rightarrow x \in A$. Par continuité de f , $y_{\varphi(n)} = f(x_{\varphi(n)}) \rightarrow f(x) \in f(A)$. \square

Théorème 1.6 (Extrema). Soit (E, d) un espace métrique, $X \subset E$ compact et $f : X \rightarrow \mathbb{R}$ continue. Alors f est bornée et atteint ses bornes dans X . Il existe x_{\min} et x_{\max} dans X tels que :

$$f(x_{\min}) = \inf_{x \in X} f(x) \quad \text{et} \quad f(x_{\max}) = \sup_{x \in X} f(x).$$

Démonstration. $f(X)$ est un compact de \mathbb{R} , donc un fermé borné. Il contient donc sa borne sup et sa borne inf. \square

Remarque 1.7. On peut se servir ces résultats, dans un EVN de dimension finie $(E, \|\cdot\|)$, pour prouver l'équivalence d'une norme N avec $\|\cdot\|$. Il suffit pour cela de montrer que N est continue par rapport $\|\cdot\|$, et qu'elle atteint donc ses bornes sur la boule unité fermée, qui est compacte. On conclut en utilisant l'homogénéité des normes.

1.8 Meilleure approximation

Soit (E, d) un espace métrique, $f \in E$ et $A \subset E$. On cherche à approcher au mieux f par un élément de A , c'est-à-dire trouver $g^* \in A$ tel que :

$$d(f, g^*) = d(f, A) = \inf_{g \in A} d(f, g).$$

Un tel g^* s'appelle une **meilleure approximation** de f sur A .

Proposition 1.11. Si A est **compact**, alors pour tout $f \in E$, il existe au moins une meilleure approximation de f dans A .

Démonstration. L'application $g \mapsto d(f, g)$ est continue. Sur le compact A , elle atteint son minimum. \square

Proposition 1.12. Si E est un espace vectoriel normé et V un sous-espace vectoriel de **dimension finie**, alors pour tout $f \in E$, il existe une meilleure approximation de f sur V .

Démonstration. Soit $d = d(f, V)$. On peut restreindre la recherche à l'intersection de V et de la boule fermée $\bar{B}(f, d + 1)$. Cette intersection est un fermé borné de V (qui est de dim finie), donc c'est un compact. On applique le résultat précédent. \square

1.9 Compacité dans les espaces de fonctions : Le Théorème d'Ascoli*

Dans les espaces vectoriels normés de dimension finie, les compacts sont exactement les fermés bornés (Théorème de Riesz). En dimension infinie, comme dans l'espace $\mathcal{C}(K, \mathbb{R})$, la boule unité n'est pas compacte. Il faut donc un critère plus fin pour caractériser les ensembles compacts de fonctions. C'est l'objet du théorème d'Ascoli, qui repose sur la notion d'équicontinuité.

1.9.1 Équicontinuité

Soit (K, d) un espace métrique compact (par exemple un segment $[a, b]$) et soit \mathcal{F} une famille de fonctions de K dans \mathbb{R} (ou \mathbb{C}).

Définition 1.14 (Équicontinuité). La famille de fonctions \mathcal{F} est dite **équicontinue** si, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que :

$$\forall x, y \in K, \quad d(x, y) < \delta \implies \forall f \in \mathcal{F}, \quad |f(x) - f(y)| < \varepsilon.$$

Remarque 1.8. La différence avec la simple continuité uniforme de chaque fonction réside dans la position du quantificateur $\forall f \in \mathcal{F}$. Le δ ne dépend pas de la fonction f choisie : il est **uniforme** pour toute la famille.

Exemple 1.4 (Familles Lipschitziennes). Si \mathcal{F} est une famille de fonctions k -lipschitziennes avec la **même** constante k pour toutes les fonctions (c'est-à-dire $\forall f \in \mathcal{F}, |f(x) - f(y)| \leq kd(x, y)$), alors \mathcal{F} est équicontinue. C'est souvent le cas pratique le plus rencontré (par exemple si les dérivées f' sont bornées uniformément par une constante M).

1.9.2 Le Théorème d'Ascoli-Arzelà

Ce théorème donne une condition nécessaire et suffisante de compacité relative pour la norme uniforme $\|\cdot\|_\infty$.

Théorème 1.7 (Théorème d'Ascoli). Soit K un espace métrique compact. Soit $\mathcal{F} \subset \mathcal{C}(K, \mathbb{R})$ un sous-ensemble de l'espace des fonctions continues muni de la norme uniforme. L'adhérence $\overline{\mathcal{F}}$ est compacte si et seulement si :

1. \mathcal{F} est **équicontinue**.

2. \mathcal{F} est **ponctuellement bornée** : pour tout $x \in K$, l'ensemble $\{f(x) \mid f \in \mathcal{F}\}$ est borné dans \mathbb{R} .

Corollaire 1.1 (Version séquentielle). De toute suite de fonctions (f_n) définies sur un compact K , qui est **bornée uniformément** et **équicontinue**, on peut extraire une sous-suite $(f_{\varphi(n)})$ qui converge uniformément sur K .

1.9.3 Démonstration

Nous démontrons l'implication directe (la suffisance), qui est le résultat le plus utile. La preuve repose sur le *procédé diagonal*.

Démonstration. Soit $(f_n)_{n \in \mathbb{N}}$ une suite équicontinue et ponctuellement bornée. Nous voulons en extraire une sous-suite uniformément convergente.

Étape 1 : Extraction d'une sous-suite convergeant sur un ensemble dense. L'espace métrique compact K est séparable. Il existe donc une partie dénombrable $D = \{x_k \mid k \in \mathbb{N}\}$ dense dans K .

- Pour x_0 , la suite numérique $(f_n(x_0))_n$ est bornée dans \mathbb{R} . D'après le théorème de Bolzano-Weierstrass, on peut extraire une sous-suite (notée ϕ_0) telle que $(f_{\phi_0(n)}(x_0))$ converge.
- Pour x_1 , la suite $(f_{\phi_0(n)}(x_1))$ est bornée. On extrait une sous-suite ϕ_1 de ϕ_0 telle que la suite converge en x_1 (et toujours en x_0).
- Par récurrence, on construit des extractrices emboîtées $\phi_0, \phi_1, \dots, \phi_k, \dots$ telles que la suite extraite par ϕ_k converge pour tous les points x_0, \dots, x_k .

On considère l'**extractrice diagonale** $\psi(n) = \phi_n(n)$. La suite $(g_n) = (f_{\psi(n)})$ converge simplement en tout point de D . Notons $g(y) = \lim g_n(y)$ pour $y \in D$.

Étape 2 : Passage de la convergence simple à uniforme via l'équicontinuité. Soit $\varepsilon > 0$. Par équicontinuité de la famille, il existe $\delta > 0$ tel que pour tous $x, y \in K$, $d(x, y) < \delta \implies |g_n(x) - g_n(y)| < \varepsilon/3$ pour tout n .

La famille de boules ouvertes $(B(x, \delta))_{x \in K}$ recouvre le compact K . On peut en extraire un sous-recouvrement fini : $K \subset \bigcup_{j=1}^p B(y_j, \delta)$. Quitte à bouger légèrement les centres, on peut supposer que les y_j sont dans l'ensemble dense D .

Pour tout $x \in K$, il existe un $y_j \in D$ tel que $d(x, y_j) < \delta$. Écrivons l'inégalité triangulaire pour deux indices n, m :

$$|g_n(x) - g_m(x)| \leq \underbrace{|g_n(x) - g_n(y_j)|}_{< \varepsilon/3 \text{ (équicont.)}} + \underbrace{|g_n(y_j) - g_m(y_j)|}_{\text{Conv. ponctuelle}} + \underbrace{|g_m(y_j) - g_m(x)|}_{< \varepsilon/3 \text{ (équicont.)}}.$$

Puisque l'ensemble des y_j est fini ($j \in \{1, \dots, p\}$), il existe un rang N tel que pour tout $n, m \geq N$, $|g_n(y_j) - g_m(y_j)| < \varepsilon/3$ pour tout $j = 1 \dots p$.

Ainsi, pour tout $n, m \geq N$ et pour tout $x \in K$, on a $|g_n(x) - g_m(x)| < \varepsilon$. La suite (g_n) est donc de Cauchy pour la norme uniforme. L'espace $\mathcal{C}(K, \mathbb{R})$ étant complet (c'est un Banach), la suite converge uniformément. \square

1.10 Quelques applications pratiques des espaces métriques

1.10.1 Application 1 : Codes correcteurs et Distance de Hamming

Dans la transmission numérique (Wifi, 5G, stockage SSD), les données sont représentées par des mots binaires. L'espace considéré est l'ensemble des mots de longueur n sur un alphabet fini (souvent $\mathbb{F}_2 = \{0, 1\}$). Notons cet espace $H_n = \{0, 1\}^n$.

Bien que H_n puisse être vu comme un espace vectoriel sur le corps fini \mathbb{F}_2 , la notion de norme est délicate. L'outil pertinent est la métrique.

Définition 1.15 (Distance de Hamming). Pour deux mots $x = (x_1, \dots, x_n)$ et $y = (y_1, \dots, y_n)$ de H_n , la distance de Hamming $d_H(x, y)$ est le nombre de positions où les bits diffèrent :

$$d_H(x, y) = \text{card}\{i \in \{1, \dots, n\} \mid x_i \neq y_i\}.$$

L'espace (H_n, d_H) est un espace métrique. Étant fini, il est trivialement complet (toute suite de Cauchy est stationnaire à partir d'un certain rang).

Intérêt pour l'ingénieur. Lors d'une transmission, un mot c (le code) est envoyé mais peut être altéré par du bruit. On reçoit un mot y . Le décodage par **maximum de vraisemblance** consiste à trouver le mot de code c_{est} le plus proche de y au sens de Hamming :

$$c_{est} = \operatorname{argmin}_{c \in \mathcal{C}} d_H(y, c).$$

Si les mots de code sont suffisamment espacés (c'est-à-dire si la distance minimale entre deux codes est grande), on peut garantir la correction d'erreurs. C'est le principe des sphères d'empilement dans un espace métrique.

1.10.2 Application 2 : Analyse d'images et Transport Optimal

Comment définir la distance entre deux images ? Si on représente deux images (en niveaux de gris) par des fonctions $f, g \in L^1(\Omega)$, la distance L^1 classique $\|f - g\|_1$ est souvent décevante : elle ne capture pas la géométrie. Si g est simplement l'image f décalée de quelques pixels, $\|f - g\|_1$ peut être très grande alors que les images sont perceptivement proches. À l'inverse la distance entre deux images peut être contre-intuitive. Si par exemple on considère les fonctions indicatrices $\mathbb{1}_A$, $\mathbb{1}_B$ et $\mathbb{1}_C$ des images A , B , C ci-dessous (c'est à dire les fonctions valant 1 sur les disques et 0 à l'extérieur de ceux-ci) et qu'on calcule leurs distances L^2 , on trouve qu'elles sont toutes les trois équidistantes : $\|\mathbb{1}_A - \mathbb{1}_B\|_1 = \|\mathbb{1}_A - \mathbb{1}_C\|_1 = \|\mathbb{1}_B - \mathbb{1}_C\|_1 = 2\text{Aire}(\text{disque})$.

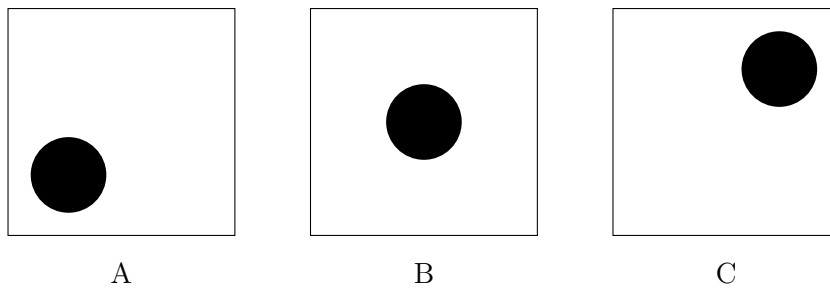


Figure 1.4 – Trois images comportant un disque disposé différemment.

L'approche moderne consiste à voir une image normalisée comme une mesure de probabilité (une distribution de masse de pixels) et à utiliser la théorie du **Transport Optimal**.

Définition 1.16 (Distance de Wasserstein W_1). Soient μ et ν deux distributions de probabilité sur \mathbb{R}^d . On appelle distance de Wasserstein-1 (ou *Earth Mover's Distance*) le coût minimal pour transporter la masse de μ vers la configuration ν :

$$W_1(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} \|x - y\| d\pi(x, y),$$

où l'infimum porte sur tous les plans de transport π (couplages) ayant μ et ν pour marginales.

Propriétés et Applications.

1. Dans l'exemple des trois images ci-dessous, l'image B sera à distance deux fois moindre de A et C qu'elles ne le sont entre elles.
2. L'espace des mesures de probabilité muni de W_1 est un espace métrique complet (Espace de Wasserstein).
3. **Interpolation d'images (Morphing)** : Contrairement à l'interpolation linéaire classique $(1 - t)f + tg$ qui crée un effet de fondu (superposition floue), l'interpolation géodésique dans l'espace de Wasserstein déplace physiquement les masses. Les pixels de l'image f "voyagent" pour former l'image g .
4. **Point fixe** : On utilise cette distance pour définir le barycentre d'un ensemble d'images (barycentre de Wasserstein). Le calcul de ce barycentre peut se formuler comme un problème de point fixe sur cet espace métrique, résolu par des algorithmes itératifs de type descente de gradient dans l'espace des mesures.

Ce cadre est aujourd'hui au cœur des algorithmes génératifs en Intelligence Artificielle (WGAN : Wasserstein Generative Adversarial Networks).

Pour aller plus loin : quelques applications

1. Compression d'images et Fractales (IFS)

Le théorème du point fixe n'agit pas seulement sur des nombres ou des vecteurs, mais aussi sur des ensembles (des images). Considérons l'espace \mathcal{K} des sous-ensembles compacts non vides de \mathbb{R}^2 (les "formes" bornées et fermées). On peut munir cet espace d'une distance, appelée **distance de Hausdorff**, qui en fait un espace métrique complet.

Un **Système de Fonctions Itérées (IFS)** est la donnée de N applications contractantes f_1, \dots, f_N de \mathbb{R}^2 dans \mathbb{R}^2 . On définit alors un opérateur global T sur l'espace des formes :

$$\forall K \in \mathcal{K}, \quad T(K) = f_1(K) \cup f_2(K) \cup \dots \cup f_N(K).$$

On peut démontrer que si les f_i sont contractantes, alors T est une contraction sur l'espace des formes. D'après le théorème du point fixe, il existe un unique compact K^* tel que $T(K^*) = K^*$. Cet ensemble K^* est souvent une **fractale**.

Exemple 1.5 (Le Tapis de Sierpinski). Prenez 3 homothéties de rapport $1/2$ centrées sur les sommets d'un triangle équilatéral. En partant de n'importe quelle forme initiale (un carré, un point...), en itérant l'opérateur T , la suite d'ensembles converge vers le triangle de Sierpinski. C'est le principe de la compression fractale d'images : au lieu de stocker les pixels, on stocke les quelques coefficients des fonctions f_i dont l'image est le point fixe.

2. Au cœur du moteur de recherche : PageRank

À la fin des années 90, Google a révolutionné la recherche web grâce à l'algorithme PageRank. L'idée est de mesurer l'importance d'une page web par la probabilité qu'un internaute aléatoire tombe dessus.

Soit n le nombre de pages web indexées (plusieurs milliards). On note $x \in \mathbb{R}^n$ le vecteur de probabilité ($x_i \geq 0, \sum x_i = 1$). Le modèle du "surfeur aléatoire" est le suivant :

- Avec une probabilité α (souvent 0.85), le surfeur clique sur un lien de la page actuelle.
- Avec une probabilité $1 - \alpha$, il s'ennuie et tape une nouvelle adresse au hasard (téléportation).

Cela se traduit par la recherche d'un vecteur x vérifiant :

$$x = \alpha Sx + (1 - \alpha) \frac{1}{n} \mathbf{1}$$

où S est la matrice de transition du graphe du web (adjacence normalisée) et $\mathbf{1}$ le vecteur ne contenant que des 1. Posons $f(x) = \alpha Sx + (1 - \alpha) \frac{1}{n} \mathbf{1}$. Si on munit l'espace de la norme $\|x\|_1 = \sum |x_i|$, on montre que :

$$\|f(x) - f(y)\|_1 \leq \alpha \|S(x - y)\|_1 \leq \alpha \|x - y\|_1.$$

Comme $\alpha \approx 0.85 < 1$, l'application est contractante ! Le classement PageRank est donc l'**unique point fixe** de cette transformation. Le calcul effectif de ce classement peut se faire simplement par itération $x_{k+1} = f(x_k)$, et le théorème du point fixe garantit que cela converge, quelle que soit l'initialisation.

Pour aller plus loin : Flots de Gradient dans les Espaces Métriques*

La notion de « descente de gradient » est intuitive dans \mathbb{R}^n ou dans un espace de Hilbert : pour minimiser une fonctionnelle (une énergie) \mathcal{E} , on suit la direction de plus grande pente. L'équation d'évolution est :

$$x'(t) = -\nabla \mathcal{E}(x(t)).$$

Mais comment définir une telle évolution si l'espace X n'est pas vectoriel, mais seulement un espace métrique (X, d) (comme une variété courbe, un graphe, ou l'espace des mesures de probabilité muni de la distance de Wasserstein) ? Sans structure linéaire, le vecteur gradient $\nabla \mathcal{E}$ n'existe pas.

La solution moderne, développée notamment par De Giorgi, puis Jordan, Kinderlehrer et Otto (JKO), consiste à définir l'évolution non pas par une équation différentielle, mais par un schéma de discrétisation temporelle variationnel.

Le schéma des Mouvements Minimisants (Minimizing Movements)

Rappelons le schéma d'Euler implicite pour l'équation $x' = -\nabla \mathcal{E}(x)$ dans \mathbb{R}^n avec un pas de temps $\tau > 0$:

$$\frac{x_{k+1} - x_k}{\tau} = -\nabla \mathcal{E}(x_{k+1}) \iff x_{k+1} + \tau \nabla \mathcal{E}(x_{k+1}) = x_k.$$

Ceci est exactement la condition d'optimalité du premier ordre pour le problème de minimisation suivant :

$$x_{k+1} = \operatorname{argmin}_{y \in \mathbb{R}^n} \left(\mathcal{E}(y) + \frac{1}{2\tau} \|y - x_k\|^2 \right).$$

Cette formulation ne fait intervenir que la fonction \mathcal{E} et la distance (au carré). Elle se généralise donc immédiatement à tout espace métrique (X, d) .

Définition 1.17 (Schéma JKO / Mouvements Minimisants). Soit (X, d) un espace métrique et $\mathcal{E} : X \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonctionnelle semi-continue inférieurement. Le schéma de **flot de gradient discret** construit une suite $(x_k^{(\tau)})$ définie récursivement par :

$$x_{k+1}^{(\tau)} \in \operatorname{argmin}_{y \in X} \left(\mathcal{E}(y) + \frac{1}{2\tau} d(y, x_k^{(\tau)})^2 \right).$$

L'intuition est un compromis : à chaque pas, on cherche un point y qui a une énergie $\mathcal{E}(y)$ plus faible, mais qui ne s'éloigne pas trop du point précédent x_k (pénalisation par la distance au carré, qui joue le rôle de l'énergie cinétique).

Application spectaculaire : L'équation de la chaleur

L'un des résultats fondateurs de cette théorie (Jordan-Kinderlehrer-Otto, 1998) concerne l'espace des mesures de probabilité $\mathcal{P}(\mathbb{R}^n)$ muni de la distance de transport optimal W_2 (distance de Wasserstein). Si l'on choisit comme énergie l'**Entropie** (l'opposé de l'entropie de Boltzmann-Shannon) :

$$\mathcal{E}(\rho) = \int_{\mathbb{R}^n} \rho(x) \ln(\rho(x)) dx,$$

alors le flot de gradient métrique associé à cette énergie dans l'espace de Wasserstein $(\mathcal{P}(\mathbb{R}^n), W_2)$ est exactement l'**équation de la chaleur** :

$$\frac{\partial \rho}{\partial t} = \Delta \rho.$$

Ce résultat a révolutionné l'analyse des EDP en permettant de voir des équations d'évolution classiques (chaleur, milieux poreux, Fokker-Planck) comme de simples descentes de gradient dans la "bonne" géométrie.

Remarque 1.9 (Références). Pour approfondir ce sujet de recherche très actif :

- **L. Ambrosio, N. Gigli, G. Savaré**, *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, Birkhäuser, 2005.
- **F. Santambrogio**, *Optimal Transport for Applied Mathematicians*, Birkhäuser, 2015.
- **C. Villani**, *Optimal Transport : Old and New*, Springer, 2009.

Anecdote : Le Café Écossais (source : [Wikipédia](#))

Les espaces de Banach ne sont pas nés dans un bureau austère, mais dans l'ambiance enfumée et bruyante d'un café de Lwów (Pologne, aujourd'hui Lviv en Ukraine) dans les années 1930 : le **Kawiarnia Szkocka** (Café Écossais).

Stefan Banach et ses collègues de l'École Mathématique de Lwów (Steinhaus, Ulam, Mazur...) avaient l'habitude de s'y retrouver pour débattre de mathématiques pendant des heures. Ils écrivaient leurs démonstrations directement sur les tables en marbre du café avec des crayons gras. Mais ces tables étaient régulièrement effacées par le personnel de ménage à la fermeture, perdant ainsi des théorèmes entiers !

Pour remédier à cela, l'épouse de Banach, Łucja, acheta un gros cahier d'écolier qu'elle confia au gérant du café : le célèbre **Livre Écossais**.

Ce livre recueillait les problèmes ouverts posés par les membres du groupe. Chaque problème était assorti d'une récompense offerte par l'auteur à celui qui trouverait la solution. Les prix variaient selon la difficulté : une tasse de café, une bière, un dîner au restaurant...

- Le problème n°153, posé par Stanisław Mazur en 1936 sur les bases dans les espaces de Banach, promettait une récompense insolite : **une oie vivante**.
- Il fallut attendre 1972 pour que le mathématicien suédois Per Enflo résolve le problème (par la négative). Stefan Banach étant décédé en 1945, c'est Mazur, alors très âgé, qui offrit personnellement l'oie vivante à Enflo lors d'une cérémonie à Varsovie !

Cette histoire nous rappelle que la recherche mathématique est avant tout une activité sociale, collaborative et vivante.

Chapitre 2

Théorie de la mesure et Intégrale de Lebesgue

L'objectif de ce chapitre est de construire une théorie de l'intégration plus robuste que celle de Riemann, capable de gérer des passages à la limite (interversion limite-intégrale) avec des conditions moins restrictives, et donnant lieu à des espaces de fonctions complets (les espaces de Banach L^p). Pour cela, nous devons d'abord formaliser la notion de « volume » ou de « masse » d'un ensemble : c'est l'objet de la théorie de la mesure.

2.1 Espaces mesurables et Mesures

La question fondamentale est : peut-on attribuer une longueur (ou un volume) à n'importe quel sous-ensemble de \mathbb{R} (ou \mathbb{R}^n) de manière cohérente ? La réponse est non (comme le montre le paradoxe de Banach-Tarski, voir encadré plus loin). On doit donc restreindre notre étude à une classe d'ensembles « raisonnables » : les ensembles mesurables, qui forment une structure de tribu.

2.1.1 Tribus et Espaces mesurables

Soit X un ensemble quelconque, on note $\mathcal{P}(X)$ l'ensemble des parties de X .

Définition 2.1 (Tribu / σ -algèbre). Une famille \mathcal{A} de sous-ensembles de X (i.e. $\mathcal{A} \subset \mathcal{P}(X)$) est appelée une **tribu** (ou σ -algèbre) si elle vérifie les propriétés suivantes :

1. **L'ensemble vide** : $\emptyset \in \mathcal{A}$.
2. **Stabilité par complémentaire** : Si $A \in \mathcal{A}$, alors son complémentaire $A^c = X \setminus A$ appartient à \mathcal{A} .
3. **Stabilité par union dénombrable** : Si $(A_n)_{n \in \mathbb{N}}$ est une suite d'éléments de \mathcal{A} , alors leur union $\bigcup_{n=0}^{\infty} A_n$ appartient à \mathcal{A} .

Le couple (X, \mathcal{A}) est appelé un **espace mesurable**. Les éléments de \mathcal{A} sont appelés les **ensembles mesurables**.

Remarque 2.1. Par passage au complémentaire, une tribu est aussi stable par intersection dénombrable ($\bigcap_{n=0}^{\infty} A_n \in \mathcal{A}$). En revanche, rien ne garantit la stabilité par union quelconque (non dénombrable).

Exemple 2.1.

1. **Tribu grossière** : $\mathcal{A} = \{\emptyset, X\}$. C'est la plus petite tribu possible.
2. **Tribu discrète** : $\mathcal{A} = \mathcal{P}(X)$. C'est la plus grande. Elle est utile pour les ensembles dénombrables (séries), mais trop riche pour définir la mesure de Lebesgue sur \mathbb{R} .

2.1.2 La Tribu borélienne

En analyse, on travaille souvent avec des ouverts (topologie). On souhaite donc pouvoir mesurer au minimum tous les ouverts (et donc les fermés).

Définition 2.2 (Tribu engendrée). Soit \mathcal{C} une collection de sous-ensembles de X . La **tribu engendrée** par \mathcal{C} , notée $\sigma(\mathcal{C})$, est la plus petite tribu contenant \mathcal{C} . C'est l'intersection de toutes les tribus contenant \mathcal{C} .

Définition 2.3 (Tribu de Borel). Soit X un espace topologique (par exemple \mathbb{R}^n muni de sa topologie usuelle). La **tribu borélienne** de X , notée $\mathcal{B}(X)$, est la tribu engendrée par les ouverts de X . Les éléments de $\mathcal{B}(X)$ sont appelés les **boréliens**.

Dans le cas de la droite réelle $X = \mathbb{R}$, la tribu borélienne $\mathcal{B}(\mathbb{R})$ est engendrée par les intervalles ouverts $]a, b[$. Elle contient donc :

- Les singletons $\{x\}$ (qui sont fermés).
- Les intervalles de la forme $[a, b]$, $[a, b[$, etc.
- Les unions dénombrables d'intervalles.

C'est la tribu « standard » de l'analyse.

2.1.3 Mesures positives

Une fois l'espace mesurable défini, on y place une mesure, qui généralise la notion de masse ou de volume.

Définition 2.4 (Mesure positive). Soit (X, \mathcal{A}) un espace mesurable. Une application $\mu : \mathcal{A} \rightarrow [0, +\infty]$ est appelée une **mesure positive** si :

1. $\mu(\emptyset) = 0$.
2. **σ -additivité** : Pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'ensembles mesurables **deux à deux disjoints**, on a :

$$\mu \left(\bigcup_{n=0}^{\infty} A_n \right) = \sum_{n=0}^{\infty} \mu(A_n).$$

Le triplet (X, \mathcal{A}, μ) est appelé un **espace mesuré**.

Remarque 2.2. Notez que la mesure peut prendre la valeur $+\infty$ (par exemple la longueur de la droite réelle tout entière). Les règles de calcul usuelles s'appliquent : $x + \infty = \infty$ pour $x \in \mathbb{R}$.

Exemple 2.2. 1. **Mesure de Dirac** : Soit $a \in X$. On définit $\delta_a(A) = 1$ si $a \in A$, et 0 sinon. C'est la modélisation d'une masse ponctuelle unité en a .

2. **Mesure de dénombrement** : Sur un ensemble X (par exemple \mathbb{N}), on définit $\mu(A) = \text{card}(A)$ (nombre d'éléments). C'est la mesure naturelle pour les séries.

3. **Mesure de Lebesgue (introductive)** : Sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, il existe une unique mesure λ telle que pour tout intervalle, $\lambda([a, b]) = b - a$. Sa construction rigoureuse sera l'objet de la section 3.

2.1.4 Propriétés fondamentales

La propriété de σ -additivité a des conséquences cruciales sur les limites d'ensembles. C'est la base de la souplesse de l'intégrale de Lebesgue.

Proposition 2.1 (Continuité de la mesure). Soit (X, \mathcal{A}, μ) un espace mesuré.

1. **Croissance** : Si $A \subset B$, alors $\mu(A) \leq \mu(B)$.
2. **Continuité croissante** : Si (A_n) est une suite croissante d'ensembles ($A_n \subset A_{n+1}$), alors :

$$\mu\left(\bigcup_{n=0}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} \mu(A_n).$$

3. **Continuité décroissante** : Si (A_n) est une suite décroissante d'ensembles ($A_{n+1} \subset A_n$) et si $\mu(A_0) < +\infty$, alors :

$$\mu\left(\bigcap_{n=0}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} \mu(A_n).$$

Idée de preuve. Pour la continuité croissante, on se ramène à une union disjointe en posant $B_0 = A_0$ et $B_n = A_n \setminus A_{n-1}$. Alors $\bigcup A_n = \bigcup B_n$ (union disjointe) et on applique la σ -additivité. \square

2.1.5 Ensembles négligeables

En intégration, ce qui se passera sur un ensemble « très petit » n'aura pas d'importance.

Définition 2.5 (Négligeable). Un ensemble $N \in \mathcal{A}$ est dit **négligeable** (pour la mesure μ) si $\mu(N) = 0$. On dit qu'une propriété est vraie μ **presque partout** (noté μ -p.p. ou a.e. *almost everywhere*) si l'ensemble des points où elle est fautive est négligeable.

Exemple 2.3. Pour la mesure de Lebesgue sur \mathbb{R} , tout singleton $\{x\}$ est négligeable. Par union dénombrable, \mathbb{Q} (l'ensemble des rationnels) est négligeable, bien qu'il soit dense dans \mathbb{R} . Cela signifie que « presque tous » les nombres réels sont irrationnels.

? Le Paradoxe de Banach-Tarski

Peut-on prendre une boule, la découper en morceaux, et réassembler ces morceaux pour obtenir... deux boules identiques à la première ? Intuitivement, c'est impossible car le volume (la masse) doit être conservé. Pourtant, en 1924, Stefan Banach et Alfred Tarski ont démontré un théorème étonnant :

Il est possible de découper la boule unité de \mathbb{R}^3 en un nombre fini de sous-ensembles (5 suffisent !), de les déplacer par des rotations et des translations, et de les réassembler pour former deux boules de rayon 1 disjointes.

Où est l'astuce ? Ce résultat repose sur l'**Axiome du Choix**. Les morceaux découpés sont des ensembles extrêmement complexes (des nuages de points fractals, de la « poussière » infinie) qui n'ont pas de volume défini. Ce sont des ensembles **non-mesurables**.

La leçon pour l'ingénieur : Ce paradoxe montre qu'il est impossible de définir une mesure de volume cohérente (invariante par déplacement) sur **tous** les sous-ensembles de l'espace. Si l'on veut conserver la physique (la masse d'un objet ne change pas quand on le bouge), on est obligé de restreindre la mesure et donc l'intégration à une classe d'ensembles « sages » : la **tribu des Boréliens**.

💡 **Pour aller plus loin : Le transport de mesure (Push-forward)***

En théorie de la mesure, il est fréquent de vouloir « transporter » une masse (ou une probabilité) d'un espace à un autre via une application. C'est la notion de **mesure image**.

Définition. Soient (X, \mathcal{A}, μ) un espace mesuré et (Y, \mathcal{B}) un espace mesurable. Soit $\phi : X \rightarrow Y$ une application mesurable. On appelle **mesure image** de μ par ϕ , notée $\phi_{\#}\mu$ (ou parfois $\mu \circ \phi^{-1}$), la mesure définie sur Y par :

$$\forall B \in \mathcal{B}, \quad (\phi_{\#}\mu)(B) := \mu(\phi^{-1}(B)).$$

Intuitivement, la masse de l'ensemble B dans l'espace d'arrivée est la masse de tous les points de X qui atterrissent dans B .

Le Théorème de Transfert. C'est la propriété fondamentale de la mesure image. Pour toute fonction mesurable positive (ou intégrable) $g : Y \rightarrow \mathbb{R}$, on a :

$$\int_Y g(y) d(\phi_{\#}\mu)(y) = \int_X g(\phi(x)) d\mu(x).$$

Cette formule dit simplement que « sommer les valeurs sur l'espace d'arrivée » revient à « sommer les valeurs sur l'espace de départ en suivant le transport ».

Liens avec d'autres domaines :

- **Probabilités** : Si $(\Omega, \mathcal{A}, \mathbb{P})$ est un espace de probabilité et $X : \Omega \rightarrow \mathbb{R}$ est une variable aléatoire réelle, la **loi de probabilité** de X (notée souvent P_X) n'est rien d'autre que la mesure image de \mathbb{P} par X :

$$P_X = X_{\#}\mathbb{P}.$$

Le théorème de transfert ci-dessus est alors exactement le théorème de transfert des probabilités : $\mathbb{E}[g(X)] = \int_{\mathbb{R}} g(x) dP_X(x)$.

- **Changement de variables** : Le théorème de changement de variables (Théorème 2.16) peut se réinterpréter ainsi : si ϕ est un difféomorphisme, la mesure image de la mesure de Lebesgue λ n'est pas directement λ , mais elle est modifiée par le facteur géométrique (le Jacobien). Plus précisément, si on pousse la mesure $d\mu = |\det J_{\phi}(x)| dx$ par ϕ , on obtient la mesure de Lebesgue dy sur l'arrivée.

2.2 Fonctions mesurables

Pour construire une intégrale $\int f d\mu$, nous avons besoin de mesurer des quantités liées aux valeurs prises par f (l'image) sur l'ensemble où on intègre (le domaine de départ). Contrairement à l'approche de Riemann qui découpe l'axe des x (le domaine de départ), l'approche de Lebesgue découpe l'axe des y (l'image). L'idée centrale est de regarder les lignes de niveau de la fonction. Pour que cela ait un sens, il faut que l'ensemble des points où la fonction dépasse une certaine valeur soit mesurable.

2.2.1 Définition et caractérisations

Soit (X, \mathcal{A}) un espace mesurable (l'espace de départ) et $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ l'espace d'arrivée muni de sa tribu borélienne.

Définition 2.6 (Fonction mesurable). Une application $f : X \rightarrow \mathbb{R}$ est dite **mesurable** (ou \mathcal{A} -mesurable) si l'image réciproque de tout ensemble borélien est mesurable :

$$\forall B \in \mathcal{B}(\mathbb{R}), \quad f^{-1}(B) = \{x \in X \mid f(x) \in B\} \in \mathcal{A}.$$

Plus généralement, une fonction $f : X \rightarrow \overline{\mathbb{R}}$ (pouvant donc prendre les valeurs $\pm\infty$) est **mesurable** si en plus de la condition ci-dessus, on a $f^{-1}(\pm\infty) \in \mathcal{A}$. On note $\mathcal{M}(X)$ ou simplement \mathcal{M} l'ensemble des fonctions mesurables sur X . On note $\mathcal{M}^+(X)$ l'ensemble des fonctions mesurables positives ou nulles sur X .

Vérifier cette condition pour tous les boréliens est impraticable. Heureusement, il suffit de tester des ensembles générateurs (comme les intervalles).

Proposition 2.2 (Critère de mesurabilité). Les assertions suivantes sont équivalentes :

1. $f : X \rightarrow \overline{\mathbb{R}}$ est mesurable.
2. Pour tout $a \in \mathbb{R}$, l'ensemble de sur-niveau $\{f > a\} = \{x \in X \mid f(x) > a\}$ appartient à \mathcal{A} .
3. Pour tout $a \in \mathbb{R}$, l'ensemble $\{f \leq a\}$ appartient à \mathcal{A} .

Remarque 2.3 (Interprétation physique). Si f représente une température sur une plaque X , dire que f est mesurable signifie simplement que la région où la température dépasse, par exemple, 20°C a une surface bien définie. C'est une condition très faible : toutes les fonctions rencontrées en physique ou en ingénierie sont mesurables.

Exemple 2.4. 1. Si $X = \mathbb{R}^n$ muni de sa tribu borélienne, toute fonction **continue** est mesurable. (Car l'image réciproque d'un ouvert est un ouvert, donc un borélien).
 2. La fonction indicatrice $\mathbf{1}_A$ d'un ensemble A est mesurable si et seulement si A est un ensemble mesurable ($A \in \mathcal{A}$).

2.2.2 Limites supérieure et inférieure*

Contrairement à la limite classique, qui n'existe pas toujours (pensez à $(-1)^n$), la limite supérieure et la limite inférieure existent pour **toute** suite numérique. C'est ce qui rend ces outils indispensables en intégration.

Définition 2.7 (Limsup et Liminf). Soit $(u_n)_{n \in \mathbb{N}}$ une suite d'éléments de $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$. On pose $v_n = \sup_{k \geq n} u_k$. La suite (v_n) est décroissante (car on prend le sup sur un ensemble d'indices de plus en plus petit), donc elle admet une limite dans $\overline{\mathbb{R}}$. On définit la **limite supérieure** de (u_n) par :

$$\limsup_{n \rightarrow \infty} u_n = \lim_{n \rightarrow \infty} \left(\sup_{k \geq n} u_k \right) = \inf_{n \in \mathbb{N}} \left(\sup_{k \geq n} u_k \right).$$

Symétriquement, on définit la **limite inférieure** par :

$$\liminf_{n \rightarrow \infty} u_n = \lim_{n \rightarrow \infty} \left(\inf_{k \geq n} u_k \right) = \sup_{n \in \mathbb{N}} \left(\inf_{k \geq n} u_k \right).$$

Exemple 2.5. Soit $u_n = (-1)^n$. La suite oscille et ne converge pas.

- Pour tout n , $\sup_{k \geq n} u_k = 1$. Donc $\limsup u_n = 1$.
- Pour tout n , $\inf_{k \geq n} u_k = -1$. Donc $\liminf u_n = -1$.

Proposition 2.3 (Propriétés fondamentales). Soit (u_n) une suite de $\overline{\mathbb{R}}$.

1. **Ordre :** On a toujours $\liminf_{n \rightarrow \infty} u_n \leq \limsup_{n \rightarrow \infty} u_n$.
2. **Convergence :** La suite (u_n) converge (dans $\overline{\mathbb{R}}$) si et seulement si :

$$\liminf_{n \rightarrow \infty} u_n = \limsup_{n \rightarrow \infty} u_n.$$

Dans ce cas, la limite est cette valeur commune.

3. **Caractérisation séquentielle :** $\limsup u_n$ est la plus grande valeur d'adhérence de la suite (la plus grande limite que l'on peut obtenir en extrayant une sous-suite convergente).

Remarque 2.4. Pour une suite de fonctions $f_n : X \rightarrow \mathbb{R}$, on définit $\limsup f_n$ ponctuellement : c'est la fonction définie par $x \mapsto \limsup_{n \rightarrow \infty} (f_n(x))$.

2.2.3 Opérations sur les fonctions mesurables

La classe des fonctions mesurables est remarquablement stable, beaucoup plus que celle des fonctions continues.

Proposition 2.4 (Stabilité algébrique). Si f et g sont des fonctions mesurables et $\lambda \in \mathbb{R}$, alors :

- $f + g$, $f \times g$, et λf sont mesurables.
- $\max(f, g)$ et $\min(f, g)$ sont mesurables.
- $|f|$ est mesurable.

La propriété la plus importante de la notion de mesurabilité est la stabilité par passage à la limite simple.

Théorème 2.1 (Stabilité par limite). Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions mesurables $X \rightarrow \mathbb{R}$. Alors les fonctions suivantes sont aussi mesurables :

$$\sup_n f_n, \quad \inf_n f_n, \quad \limsup_{n \rightarrow \infty} f_n, \quad \liminf_{n \rightarrow \infty} f_n.$$

En particulier, si la suite (f_n) converge simplement vers une fonction f , alors f est mesurable.

Démonstration. Rappelons qu'une fonction g est mesurable si et seulement si pour tout réel a , l'ensemble de sur-niveau strict $\{x \in X \mid g(x) > a\}$ appartient à la tribu \mathcal{A} .

1. **Cas du Supremum :** Soit $g(x) = \sup_{n \in \mathbb{N}} f_n(x)$. Pour un réel a fixé, un supremum est strictement supérieur à a si et seulement si au moins un des termes est strictement supérieur à a .

$$\sup_n f_n(x) > a \iff \exists n \in \mathbb{N}, f_n(x) > a.$$

En termes d'ensembles, cela se traduit par une union :

$$\{x \mid \sup_n f_n(x) > a\} = \bigcup_{n \in \mathbb{N}} \{x \mid f_n(x) > a\}.$$

Comme chaque fonction f_n est mesurable, les ensembles $\{f_n > a\}$ sont dans \mathcal{A} . Comme \mathcal{A} est une tribu, elle est stable par union dénombrable. Donc $\{g > a\} \in \mathcal{A}$, ce qui prouve que $\sup f_n$ est mesurable.

2. **Cas de l'Infimum :** identique au précédent.

3. **Cas des Limites sup et inf** : Par définition :

$$\limsup_{n \rightarrow \infty} f_n = \inf_{k \in \mathbb{N}} \left(\sup_{n \geq k} f_n \right) \quad \text{et} \quad \liminf_{n \rightarrow \infty} f_n = \sup_{k \in \mathbb{N}} \left(\inf_{n \geq k} f_n \right).$$

La fonction $g_k = \sup_{n \geq k} f_n$ est mesurable d'après le point 1. La \limsup est l'infimum (sur k) de ces fonctions g_k , elle est donc mesurable d'après le point 2. Le raisonnement est identique pour la \liminf .

4. **Cas de la limite simple** : Si la suite (f_n) converge simplement vers f , alors par définition de la convergence dans \mathbb{R} :

$$f(x) = \lim_{n \rightarrow \infty} f_n(x) = \limsup_{n \rightarrow \infty} f_n(x) = \liminf_{n \rightarrow \infty} f_n(x).$$

Puisque la \limsup est mesurable, f est mesurable.

□

2.2.4 Fonctions étagées et approximation

Pour définir l'intégrale, nous allons procéder par approximation. Les fonctions les plus simples à intégrer sont celles qui ne prennent qu'un nombre fini de valeurs.

Définition 2.8 (Fonction étagée). Une fonction $s : X \rightarrow \mathbb{R}$ est dite **étagée** (ou simple) si :

1. Elle est mesurable.
2. Elle ne prend qu'un nombre **fini** de valeurs distinctes.

Elle s'écrit alors sous la forme canonique $s(x) = \sum_{i=1}^k \alpha_i \mathbf{1}_{A_i}(x)$, où les A_i sont des ensembles mesurables disjoints et $\alpha_i \in \mathbb{R}$. On note \mathcal{E} l'espace vectoriel (le démontrer en exercice !) des fonctions étagées, et \mathcal{E}^+ l'ensemble de celles qui sont à valeurs positives.

Le théorème suivant est une des clés de voûte de la théorie. Il affirme qu'on peut reconstruire n'importe quelle fonction mesurable positive (même très irrégulière) comme une limite croissante de fonctions étagées.

Théorème 2.2 (Théorème d'approximation). Soit $f : X \rightarrow [0, +\infty]$ une fonction mesurable **positive**. Il existe une suite $(s_n)_{n \in \mathbb{N}}$ de fonctions étagées positives telle que :

1. La suite est croissante : $0 \leq s_n(x) \leq s_{n+1}(x) \leq f(x)$.
2. La suite converge simplement vers f : $\forall x \in X, \lim_{n \rightarrow \infty} s_n(x) = f(x)$.

NB : On notera cette convergence croissante par : $s_n \uparrow f$.

Idée de la construction. Pour construire s_n , on découpe l'axe des ordonnées (l'image) en intervalles de taille $1/2^n$. On définit :

$$s_n(x) = \sum_{k=0}^{n2^n-1} \frac{k}{2^n} \mathbf{1}_{\{\frac{k}{2^n} \leq f(x) < \frac{k+1}{2^n}\}}(x) + n \mathbf{1}_{\{f(x) \geq n\}}(x).$$

Cette fonction "quantifie" les valeurs de f par en dessous. Quand n augmente, la grille s'affine et s_n épouse le profil de f de plus en plus près (voir figure 2.1). □

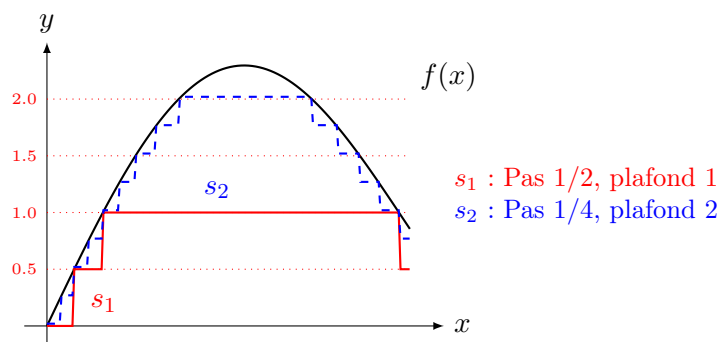


Figure 2.1 – Visualisation de la suite d'approximation (s_n) . La fonction s_1 (rouge) est une fonction en escalier grossière qui plafonne à 1. La fonction s_2 (bleu) a des marches deux fois plus fines et plafonne à 2, collant bien mieux à la courbe f .

2.3 La mesure de Lebesgue

Nous avons défini ce qu'est une mesure abstraite. Sur \mathbb{R}^n , nous voulons une mesure qui corresponde à notre intuition géométrique du volume : le volume d'un pavé doit être le produit des longueurs de ses côtés.

Théorème 2.3 (Mesure de Lebesgue). *Il existe une unique mesure positive λ sur la tribu borélienne $\mathcal{B}(\mathbb{R}^n)$ qui vérifie les propriétés suivantes :*

1. **Normalisation** : Pour tout pavé $P = [a_1, b_1] \times \cdots \times [a_n, b_n]$,

$$\lambda(P) = \prod_{i=1}^n (b_i - a_i).$$

2. **Invariance par translation** : Pour tout borélien A et tout vecteur $x \in \mathbb{R}^n$,

$$\lambda(A + x) = \lambda(A).$$

Cette mesure λ est appelée **Mesure de Lebesgue**.

Remarque 2.5. L'invariance par translation est une propriété cruciale en physique (l'expérience ne dépend pas du choix de l'origine). C'est ce qui distingue la mesure de Lebesgue des autres mesures (comme les probabilités gaussiennes).

2.4 Construction de l'intégrale de Lebesgue

L'objectif est de définir, pour un espace mesuré (X, \mathcal{A}, μ) , la quantité $\int_X f d\mu$. La construction se fait en trois étapes : d'abord pour les fonctions très simples (étagées positives), puis pour les fonctions positives (par approximation), et enfin pour les fonctions quelconques.

2.4.1 Étape 1 : Fonctions étagées positives

Soit $s : X \rightarrow [0, +\infty[$ une fonction étagée mesurable. Elle s'écrit comme une combinaison linéaire finie de fonctions indicatrices :

$$s(x) = \sum_{i=1}^k \alpha_i \mathbf{1}_{A_i}(x)$$

avec $\alpha_i \geq 0$ et $A_i \in \mathcal{A}$.

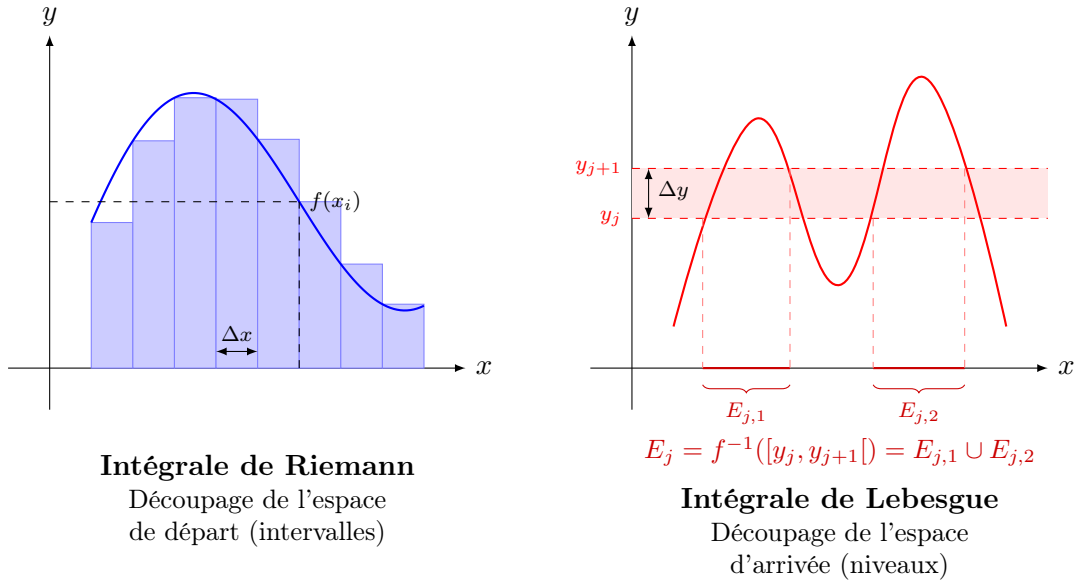


Figure 2.2 – Comparaison des approches d'intégration. À gauche (Riemann), on somme des rectangles de hauteur $f(x_i)$ et de base Δx . À droite (Lebesgue), on somme des tranches de hauteur Δy et de base la mesure de l'ensemble E_j (qui peut être une réunion d'intervalles disjoints).

Définition 2.9. On définit l'intégrale de la fonction étagée s par :

$$\int_X s \, d\mu = \sum_{i=1}^k \alpha_i \mu(A_i).$$

Remarque 2.6. Cette définition est cohérente : elle ne dépend pas de la façon dont on écrit s (la décomposition n'est pas unique). Si un A_i est de mesure infinie et que $\alpha_i > 0$, l'intégrale vaut $+\infty$. On adopte la convention $0 \times \infty = 0$ (si $\alpha_i = 0$ sur un ensemble de mesure infinie, cela ne contribue pas).

Proposition 2.5. L'intégrale ainsi définie est une forme linéaire sur l'ensemble des fonctions étagées.

2.4.2 Étape 2 : Fonctions mesurables positives

C'est ici que la magie de Lebesgue opère. Pour une fonction f quelconque, on l'approche par le bas avec des fonctions étagées (comme des rectangles sous la courbe).

Définition 2.10 (Intégrale d'une fonction positive). Soit $f : X \rightarrow [0, +\infty]$ une fonction mesurable positive. On définit son intégrale (qui peut valoir $+\infty$) par :

$$\int_X f \, d\mu = \sup \left\{ \int_X s \, d\mu \mid s \text{ étagée, } 0 \leq s \leq f \right\}.$$

Si cette intégrale est finie, on dit que la fonction est **intégrable**, ou parfois **sommable**.

Proposition 2.6 (Propriétés élémentaires de l'intégrale positive). Soient f et g deux fonctions mesurables positives et $\alpha \in \mathbb{R}^+$.

1. **Croissance (ou Monotonie) :** Si $f \leq g$ presque partout, alors

$$\int_X f \, d\mu \leq \int_X g \, d\mu.$$

2. **Homogénéité :**

$$\int_X \alpha f \, d\mu = \alpha \int_X f \, d\mu.$$

Démonstration. 1. **Croissance :** Supposons que $f(x) \leq g(x)$ pour tout $x \in X$ (si l'inégalité est vraie seulement presque partout, la démonstration est identique car l'intégrale ne voit pas les ensembles de mesure nulle). Par définition de l'intégrale ci-dessus :

$$\int_X f \, d\mu = \sup \left\{ \int_X \varphi \, d\mu \mid \varphi \in \mathcal{E}^+, 0 \leq \varphi \leq f \right\}.$$

Soit $\varphi \in \mathcal{E}^+$ telle que $0 \leq \varphi \leq f$. Comme $f \leq g$, on a a fortiori $0 \leq \varphi \leq g$. Ainsi, l'ensemble des fonctions étagées inférieures à f est inclus dans l'ensemble des fonctions étagées inférieures à g :

$$\{\varphi \in \mathcal{E}^+ \mid \varphi \leq f\} \subset \{\psi \in \mathcal{E}^+ \mid \psi \leq g\}.$$

En passant à la borne supérieure sur les intégrales de ces fonctions étagées, on obtient l'inégalité annoncée.

2. **Homogénéité :** Si $\alpha = 0$, l'égalité est triviale ($0 = 0$). Supposons $\alpha > 0$. On remarque que $\varphi \in \mathcal{E}^+$ et $\varphi \leq f$ équivaut à $\alpha\varphi \in \mathcal{E}^+$ et $\alpha\varphi \leq \alpha f$. En utilisant la linéarité de l'intégrale pour les fonctions étagées (déjà établie précédemment), on a $\int \alpha\varphi = \alpha \int \varphi$. D'où :

$$\begin{aligned} \int_X \alpha f \, d\mu &= \sup \left\{ \int_X \psi \, d\mu \mid \psi \in \mathcal{E}^+, \psi \leq \alpha f \right\} \\ &= \sup \left\{ \int_X \alpha\varphi \, d\mu \mid \varphi \in \mathcal{E}^+, \alpha\varphi \leq \alpha f \right\} \quad (\text{en posant } \psi = \alpha\varphi) \\ &= \sup \left\{ \alpha \int_X \varphi \, d\mu \mid \varphi \in \mathcal{E}^+, \varphi \leq f \right\} \\ &= \alpha \sup \left\{ \int_X \varphi \, d\mu \mid \varphi \in \mathcal{E}^+, \varphi \leq f \right\} \\ &= \alpha \int_X f \, d\mu. \end{aligned}$$

□

2.4.3 Le Théorème de Convergence Monotone

C'est le premier grand résultat de convergence, qui n'a pas d'équivalent aussi simple chez Riemann. Il affirme que pour les fonctions positives, on peut intervertir limite et intégrale sans condition restrictive.

Théorème 2.4 (Convergence Monotone / Beppo Levi). Soit $(f_n)_{n \in \mathbb{N}}$ une **suite croissante** de fonctions mesurables **positives** ($0 \leq f_n \leq f_{n+1}$). Alors la fonction limite $f = \lim_{n \rightarrow \infty} f_n$ (définie point par point) vérifie :

$$\int_X \left(\lim_{n \rightarrow \infty} f_n \right) \, d\mu = \lim_{n \rightarrow \infty} \int_X f_n \, d\mu.$$

Important : Ce théorème permet de justifier la linéarité de l'intégrale pour les fonctions positives ($\int (f + \lambda g) = \int f + \lambda \int g$) en utilisant des suites approximantes.

Démonstration. Notons $f = \lim_{n \rightarrow \infty} f_n$. Puisque la suite (f_n) est croissante, cette limite existe (éventuellement $+\infty$) en tout point.

Pour tout n , on a $f_n \leq f_{n+1} \leq f$. Par croissance de l'intégrale (proposition 2.6), on a :

$$\int_X f_n d\mu \leq \int_X f d\mu.$$

En passant à la limite (la suite des intégrales est croissante), on obtient :

$$\lim_{n \rightarrow \infty} \int_X f_n d\mu \leq \int_X f d\mu.$$

Il s'agit de montrer l'inégalité inverse. Par définition de l'intégrale de Lebesgue d'une fonction positive, $\int f = \sup\{\int s \mid 0 \leq s \leq f, s \text{ étagée}\}$. Il suffit donc de montrer que pour toute fonction étagée s telle que $0 \leq s \leq f$, on a $\int s \leq \lim \int f_n$.

Soit s une telle fonction étagée et soit $c \in]0, 1[$ un paramètre destiné à tendre vers 1. Posons $E_n = \{x \in X \mid f_n(x) \geq c \cdot s(x)\}$.

- La suite d'ensembles (E_n) est croissante car la suite de fonctions (f_n) l'est.
- La réunion des E_n est X tout entier. En effet, pour $x \in X$, si $s(x) = 0$, $x \in E_n$ pour tout n , trivialement. Si $s(x) > 0$, comme $f_n(x) \rightarrow f(x)$ et que $f(x) \geq s(x) > c \cdot s(x)$, il existe un rang m à partir duquel l'inégalité stricte est vérifiée, et donc $x \in E_m$.

Sur l'ensemble E_n , on a $f_n \geq c \cdot s$. Donc par croissance et homogénéité (proposition 2.6) :

$$\int_X f_n d\mu \geq \int_{E_n} f_n d\mu \geq c \int_{E_n} s d\mu.$$

L'application $A \mapsto \nu(A) = \int_A s d\mu$ est une mesure (car s est étagée positive). Par propriété de continuité croissante de la mesure, $\lim_{n \rightarrow \infty} \nu(E_n) = \nu(\bigcup E_n) = \nu(X) = \int_X s d\mu$.

En passant à la limite dans l'inégalité précédente :

$$\lim_{n \rightarrow \infty} \int_X f_n d\mu \geq c \int_X s d\mu.$$

Ceci est vrai pour tout $c < 1$. En faisant tendre c vers 1, on obtient :

$$\lim_{n \rightarrow \infty} \int_X f_n d\mu \geq \int_X s d\mu.$$

En prenant le supremum sur toutes les fonctions étagées $s \leq f$, on conclut :

$$\lim_{n \rightarrow \infty} \int_X f_n d\mu \geq \int_X f d\mu.$$

□

Proposition 2.7 (Additivité de l'intégrale sur \mathcal{M}^+). Soient f et g deux fonctions mesurables positives ($f, g \in \mathcal{M}^+$). Alors :

$$\int_X (f + g) d\mu = \int_X f d\mu + \int_X g d\mu.$$

Démonstration. La démonstration se fait en deux étapes : d'abord pour les fonctions étagées, puis pour les fonctions mesurables quelconques par passage à la limite via le Théorème de Convergence Monotone ci-dessus.

1. **Cas des fonctions étagées :** Supposons que f et g soient dans \mathcal{E}^+ . On peut toujours trouver une subdivision commune de l'espace pour écrire ces fonctions. Soient $f = \sum_{i=1}^n a_i \mathbf{1}_{A_i}$ et $g = \sum_{i=1}^n b_i \mathbf{1}_{A_i}$ où les (A_i) forment une partition de X (quitte à raffiner les partitions respectives de f et g). Alors $f + g = \sum_{i=1}^n (a_i + b_i) \mathbf{1}_{A_i}$ est encore une fonction étagée. Par définition de l'intégrale des fonctions étagées :

$$\int_X (f + g) d\mu = \sum_{i=1}^n (a_i + b_i) \mu(A_i) = \sum_{i=1}^n a_i \mu(A_i) + \sum_{i=1}^n b_i \mu(A_i) = \int_X f d\mu + \int_X g d\mu.$$

L'additivité est donc vraie sur \mathcal{E}^+ .

2. **Cas général (Passage à la limite) :** Soient maintenant $f, g \in \mathcal{M}^+$. Il existe deux suites croissantes de fonctions étagées positives $(\varphi_n)_{n \in \mathbb{N}}$ et $(\psi_n)_{n \in \mathbb{N}}$ telles que :

$$\varphi_n \uparrow f \quad \text{et} \quad \psi_n \uparrow g \quad \text{simplement.}$$

La suite somme $(\varphi_n + \psi_n)_{n \in \mathbb{N}}$ est alors une suite de fonctions étagées positives qui croît vers $f + g$.

D'après le **Théorème de Convergence Monotone** (appliqué à f , à g , et à $f + g$), on a :

$$\begin{aligned} \int_X (f + g) d\mu &= \lim_{n \rightarrow \infty} \int_X (\varphi_n + \psi_n) d\mu && \text{(par TCM sur la somme)} \\ &= \lim_{n \rightarrow \infty} \left(\int_X \varphi_n d\mu + \int_X \psi_n d\mu \right) && \text{(par additivité sur } \mathcal{E}^+) \\ &= \lim_{n \rightarrow \infty} \int_X \varphi_n d\mu + \lim_{n \rightarrow \infty} \int_X \psi_n d\mu && \text{(linéarité de la limite)} \\ &= \int_X f d\mu + \int_X g d\mu && \text{(par TCM sur } f \text{ et } g). \end{aligned}$$

□

Corollaire 2.1. Soit $f_n : X \rightarrow \overline{\mathbb{R}}_+$ pour tout n , on a :

$$\int_X \left(\sum_{n=0}^{+\infty} f_n \right) d\mu = \sum_{n=0}^{+\infty} \int_X f_n d\mu$$

Démonstration. On pose $g_j = \sum_{n=0}^j f_n$ et on applique Beppo-Levi sur la suite croissante (g_j) . □

Théorème 2.5. Soit $f : X \rightarrow \overline{\mathbb{R}}_+$. On a :

$$\int_X f(x) d\mu = 0 \Leftrightarrow f = 0 \text{ p.p. (soit } f(x) = 0 \text{ p.p.)}$$

Démonstration.

Posons $A = \{x \in X, f(x) \neq 0\}$.

$$\begin{aligned} \Leftrightarrow \forall x \in X f(x) &\leq \lim_{n \rightarrow \infty} n \mathbf{1}_A(x) \\ \int_X f(x) d\mu &\leq \int_X \lim_{n \rightarrow \infty} n \mathbf{1}_A(x) dx \stackrel{\text{th 2.4}}{=} \lim_{n \rightarrow \infty} n \underbrace{\int_X \mathbf{1}_A(x) d\mu}_{m(A)=0} \text{ car } n \mathbf{1}_A \text{ est croissante.} \end{aligned}$$

$$\begin{aligned} \Rightarrow \mathbb{1}_A(x) &\leq \lim_{n \rightarrow \infty} n f(x) \\ \mu(A) &= \int_X \mathbb{1}_A(x) d\mu \leq \int_X \lim_{n \rightarrow \infty} n f(x) d\mu \stackrel{th 2.4}{=} \lim_{n \rightarrow \infty} n \underbrace{\int_X f(x) d\mu}_{=0} = 0 \end{aligned}$$

□

2.4.4 Étape 3 : Fonctions mesurables de signe quelconque

Si $f : X \rightarrow \mathbb{R}$ (ou $\overline{\mathbb{R}}$) change de signe, on la décompose en sa partie positive et sa partie négative :

$$f^+(x) = \max(f(x), 0) \quad \text{et} \quad f^-(x) = \max(-f(x), 0).$$

On a alors $f = f^+ - f^-$ et $|f| = f^+ + f^-$. Ces deux fonctions sont mesurables et positives.

Définition 2.11 (Fonction intégrable). Une fonction mesurable f est dite **intégrable** (au sens de Lebesgue) si :

$$\int_X |f| d\mu < +\infty.$$

Cela équivaut à dire que $\int_X f^+ d\mu < +\infty$ et $\int_X f^- d\mu < +\infty$.

Définition 2.12 (L'intégrale de Lebesgue). Si f est intégrable, on définit son intégrale par :

$$\int_X f d\mu = \int_X f^+ d\mu - \int_X f^- d\mu.$$

Remarque 2.7 (Attention aux formes indéterminées). Si f n'est pas intégrable (par exemple si $\int f^+ = \infty$), on ne peut généralement pas définir $\int f$. La forme $\infty - \infty$ est interdite. C'est une différence avec les intégrales impropres de Riemann (comme $\int \frac{\sin x}{x} dx$) qui peuvent converger grâce à des compensations de signes sans que $\int |\frac{\sin x}{x}|$ soit finie. L'intégrale de Lebesgue est une intégrale absolue.

2.5 L'espace des fonctions intégrables L^1

Nous avons défini l'intégrale pour des fonctions mesurables vérifiant $\int |f| < +\infty$. Il est temps de structurer cet ensemble et de lui donner une structure d'espace vectoriel normé, cadre naturel de l'analyse fonctionnelle.

2.5.1 L'espace vectoriel \mathcal{L}^1

Définition 2.13 (L'espace \mathcal{L}^1). Soit (X, \mathcal{A}, μ) un espace mesuré. On note $\mathcal{L}^1(X, \mu)$ (ou simplement $\mathcal{L}^1(X)$ voire même \mathcal{L}^1) l'ensemble des fonctions mesurables $f : X \rightarrow \mathbb{R}$ telles que :

$$\int_X |f| d\mu < +\infty.$$

Proposition 2.8 (Structure vectorielle et Linéarité). $\mathcal{L}^1(X, \mu)$ est un **espace vectoriel** réel. De plus, l'application $f \mapsto \int_X f d\mu$ est une forme linéaire sur cet espace. Autrement dit, pour tous $f, g \in \mathcal{L}^1$ et tout $\lambda \in \mathbb{R}$:

1. $f + g \in \mathcal{L}^1$ et $\lambda f \in \mathcal{L}^1$.
2. $\int_X (f + g) d\mu = \int_X f d\mu + \int_X g d\mu$.

$$3. \int_X (\lambda f) d\mu = \lambda \int_X f d\mu.$$

$$4. \left| \int_X f d\mu \right| \leq \int_X |f| d\mu \quad \text{et} \quad \int_X |f+g| d\mu \leq \int_X |f| d\mu + \int_X |g| d\mu$$

Démonstration. La stabilité découle de l'inégalité triangulaire $|f+g| \leq |f|+|g|$. Pour la linéarité de l'intégrale, le cas positif a été traité grâce au théorème de convergence monotone (approximation par des fonctions étagées). Pour le cas général, si $h = f+g$, on écrit $h^+ - h^- = (f^+ - f^-) + (g^+ - g^-)$, ce qui donne $h^+ + f^- + g^- = h^- + f^+ + g^+$. Tous les termes étant positifs et intégrables, on applique la linéarité sur les fonctions positives :

$$\int h^+ + \int f^- + \int g^- = \int h^- + \int f^+ + \int g^+.$$

En regroupant les termes (ce qui est légitime car toutes les intégrales sont finies), on retrouve $\int h = \int f + \int g$. Les autres points se démontrent de manière analogue en décomposant en parties positives et négatives (exercice!). \square

Définition 2.14. Soit $A \subset X$ mesurable et $f : A \rightarrow \overline{\mathbb{R}}$. On note :

$$f_A : x \mapsto f_A(x) = \begin{cases} f(x) & \text{si } x \in A \\ 0 & \text{sinon} \end{cases}$$

qui est mesurable. On dit que f est intégrable sur A si f_A est intégrable sur X , et dans ce cas on pose

$$\int_A f d\mu = \int_X f_A d\mu$$

On note $\mathcal{L}^1(A, \mu)$ l'ensemble des applications intégrables sur A .

Remarque 2.8. On étend très simplement cette définition aux fonctions à valeurs dans $\overline{\mathbb{C}}$ en considérant les parties réelles et imaginaires. On dit que $f : X \rightarrow \overline{\mathbb{C}}$ est intégrable si $\operatorname{Im} f$ et $\operatorname{Re} f$ sont intégrables au sens de la définition ci-dessus et on pose alors

$$\int_X f d\mu = \int_X \operatorname{Re} f d\mu + i \int_X \operatorname{Im} f d\mu.$$

On a alors en module :

$$\left| \int_X f d\mu \right| \leq \int_X |f| d\mu \quad \text{et} \quad \int_X |f+g| d\mu \leq \int_X |f| d\mu + \int_X |g| d\mu$$

Cette définition sera utilisée dans le chapitre sur la transformée de Fourier. Néanmoins pour la suite par souci de simplicité, et comme nous venons de voir qu'on peut s'y ramener, nous considérons uniquement le cas des fonctions à valeurs réelles.

2.5.2 Le problème de la séparation et la relation « presque partout »

On souhaite munir \mathcal{L}^1 de la norme naturelle $\|f\|_1 = \int_X |f| d\mu$. Cependant, une propriété fondamentale des normes fait défaut : la séparation. En effet d'après le théorème 2.5, on a

$$\|f\|_1 = 0 \iff \int_X |f| d\mu = 0 \iff f = 0 \quad \text{presque partout.}$$

Cela n'implique pas que f est nulle partout (elle peut valoir 1 sur un ensemble négligeable). Ainsi, $\|\cdot\|_1$ n'est qu'une **semi-norme** sur \mathcal{L}^1 .

Pour résoudre ce problème, on identifie les fonctions qui ne diffèrent que sur un ensemble négligeable.

Définition 2.15 (Relation d'équivalence). On définit la relation \sim sur \mathcal{L}^1 par :

$$f \sim g \iff f = g \quad \mu\text{-presque partout} \iff \int_X |f - g| d\mu = 0.$$

C'est une relation d'équivalence compatible avec la structure d'espace vectoriel.

2.5.3 L'espace de Banach L^1

Définition 2.16 (L'espace quotient L^1). L'espace $L^1(X, \mu)$ est défini comme l'espace quotient :

$$L^1(X, \mu) = \mathcal{L}^1(X, \mu) / \sim.$$

Les éléments de L^1 sont des **classes d'équivalence** de fonctions.

Remarque 2.9 (Abus de langage). En pratique, on continue de parler de « la fonction f » dans L^1 et on écrit $f(x)$, mais on garde à l'esprit que f n'est définie qu'à un ensemble négligeable près. Parler de la valeur $f(x_0)$ en un point précis x_0 n'a généralement pas de sens dans L^1 (sauf si la mesure charge le point x_0 , ou si un représentant continu est dans la classe d'équivalence de f), mais parler de $\int f$ ou de $f > 0$ p.p. a un sens.

Théorème 2.6 (Riesz-Fischer). L'espace $(L^1(X, \mu), \|\cdot\|_1)$ muni de la norme :

$$\|f\|_1 = \int_X |f| d\mu$$

est un espace vectoriel normé **complet**. C'est donc un **espace de Banach**.

La complétude est un résultat profond qui signifie que toute suite de Cauchy pour la norme L^1 converge vers une fonction intégrable. C'est l'avantage décisif de l'intégrale de Lebesgue sur celle de Riemann.

Démonstration. Soit $(f_n)_{n \in \mathbb{N}}$ une suite de Cauchy dans $L^1(X, \mu)$. On cherche une fonction $f \in L^1$ telle que $\|f_n - f\|_1 \rightarrow 0$.

Étape 1 : Extraction d'une sous-suite à convergence rapide Puisque la suite est de Cauchy, on peut choisir des indices n_k de plus en plus grands tels que la distance entre deux termes consécutifs de la sous-suite soit très petite. Pour tout $k \in \mathbb{N}$, il existe un indice n_k tel que :

$$\|f_{n_{k+1}} - f_{n_k}\|_1 \leq \frac{1}{2^k}.$$

On considère la série des différences : $\sum_{k=0}^{+\infty} (f_{n_{k+1}} - f_{n_k})$. Si cette série converge, alors la suite télescopique (f_{n_k}) converge.

Étape 2 : Construction de la limite ponctuelle Considérons la série des valeurs absolues :

$$g(x) = \sum_{k=0}^{+\infty} |f_{n_{k+1}}(x) - f_{n_k}(x)|.$$

C'est une somme de fonctions positives. D'après le **Théorème de convergence monotone** (Beppo-Levi) appliqué aux sommes partielles, on peut intervertir intégrale et somme infinie :

$$\int_X g d\mu = \sum_{k=0}^{+\infty} \int_X |f_{n_{k+1}} - f_{n_k}| d\mu = \sum_{k=0}^{+\infty} \|f_{n_{k+1}} - f_{n_k}\|_1 \leq \sum_{k=0}^{+\infty} \frac{1}{2^k} = 2 < +\infty.$$

Puisque $\int g < +\infty$, la fonction g est finie presque partout (voir exercice en TD). Cela signifie que la série $\sum (f_{n_{k+1}}(x) - f_{n_k}(x))$ converge absolument pour presque tout x . La suite $(f_{n_k}(x))$ converge donc presque partout vers une limite que l'on note $f(x)$.

Étape 3 : Convergence en norme L^1 La limite ponctuelle f étant identifiée, il faut montrer que $f \in L^1$ et que la convergence a lieu en norme. On remarque que pour tout K , $|f_{n_K}(x)| \leq |f_{n_0}(x)| + \sum_{k=0}^{K-1} |f_{n_{k+1}} - f_{n_k}| \leq |f_{n_0}(x)| + g(x)$. En passant à la limite $K \rightarrow \infty$, on a $|f(x)| \leq |f_{n_0}(x)| + g(x)$ p.p. Comme f_{n_0} et g sont dans L^1 , f est bien dans L^1 .

Enfin, pour montrer que $\|f_{n_k} - f\|_1 \rightarrow 0$, on utilise le **Théorème de Convergence Dominée** :

- La suite $(f_{n_k} - f)$ tend vers 0 presque partout.
- Elle est dominée par une fonction intégrable : $|f_{n_k} - f| \leq |f_{n_k}| + |f| \leq 2(|f_{n_0}| + g) \in L^1$.

Donc $\|f_{n_k} - f\|_1 \rightarrow 0$. La sous-suite converge. Or, une suite de Cauchy qui admet une sous-suite convergente est elle-même convergente vers la même limite. Conclusion : $f_n \rightarrow f$ dans L^1 . \square

2.6 Les espaces L^p pour $1 < p \leq \infty$

Nous généralisons ici la construction de l'espace L^1 aux fonctions dont une puissance p -ième est intégrable. Ces espaces permettent de mesurer la taille des fonctions de manière plus fine (par exemple, pénaliser davantage les grandes valeurs).

2.6.1 Définitions et Normes pour $1 \leq p < \infty$

Soit p un nombre réel tel que $1 \leq p < \infty$.

Définition 2.17 (L'espace \mathcal{L}^p). On note $\mathcal{L}^p(X, \mu)$ l'ensemble des fonctions mesurables $f : X \rightarrow \mathbb{R}$ (ou \mathbb{C}) telles que $|f|^p$ soit intégrable :

$$\int_X |f(x)|^p d\mu(x) < +\infty.$$

C'est un espace vectoriel sur \mathbb{K} .

On définit sur cet espace une application qui vérifie les propriétés d'une norme (sauf la séparation) :

$$\|f\|_p = \left(\int_X |f(x)|^p d\mu(x) \right)^{1/p}.$$

L'inégalité triangulaire pour cette norme s'appelle l'**Inégalité de Minkowski** : $\|f + g\|_p \leq \|f\|_p + \|g\|_p$.

Comme pour L^1 , pour obtenir une vraie norme, il faut quotienter par la relation d'égalité presque partout.

Définition 2.18 (L'espace de Banach L^p). L'espace $L^p(X, \mu)$ est l'espace quotient $\mathcal{L}^p(X, \mu) / \sim$ (où $f \sim g \iff f = g$ p.p.). Muni de la norme $\|\cdot\|_p$, c'est un **espace de Banach** (il est complet).

2.6.2 L'espace L^∞ (Fonctions essentiellement bornées)

Le cas $p = \infty$ est particulier : il correspond aux fonctions bornées, mais en ignorant ce qui se passe sur les ensembles négligeables.

Définition 2.19 (Supremum essentiel). Soit f une fonction mesurable. On appelle **supremum essentiel** de $|f|$, noté $\|f\|_\infty$, la plus petite borne supérieure de $|f|$ en négligeant les ensembles de mesure nulle :

$$\|f\|_\infty = \inf\{C \geq 0 \mid |f(x)| \leq C \text{ presque partout}\}.$$

Si cet ensemble est vide, on pose $\|f\|_\infty = +\infty$.

Définition 2.20 (L'espace L^∞). L'espace $L^\infty(X, \mu)$ est l'ensemble des (classes de) fonctions mesurables telles que $\|f\|_\infty < +\infty$. Muni de la norme $\|\cdot\|_\infty$, c'est un **espace de Banach**.

Exemple 2.6. La fonction $f(x) = 1/x$ sur $]0, 1[$ n'est pas dans L^∞ . La fonction $f(x) = \sin(x)$ sur \mathbb{R} est dans L^∞ et $\|f\|_\infty = 1$.

2.6.3 Le cas particulier L^2 : Un espace de Hilbert

Parmi tous les espaces L^p , l'espace L^2 (fonctions de carré intégrable) joue un rôle central car sa norme provient d'un produit scalaire. C'est le seul espace L^p qui est un **espace de Hilbert**.

Définition 2.21 (Produit scalaire L^2). Pour deux fonctions $f, g \in L^2(X, \mu)$ (à valeurs réelles), on définit leur produit scalaire par :

$$\langle f, g \rangle = \int_X f(x)g(x) d\mu(x).$$

(Si les fonctions sont à valeurs complexes, $\langle f, g \rangle = \int f \bar{g} d\mu$).

Cette intégrale est bien définie grâce à l'inégalité de Cauchy-Schwarz (qui est le cas $p = 2$ de l'inégalité de Hölder) :

$$\left| \int_X fg d\mu \right| \leq \|f\|_2 \|g\|_2.$$

La norme associée est : $\|f\|_2 = \sqrt{\int |f|^2}$.

Remarque 2.10 (Importance pour l'analyse spectrale). L'espace L^2 est particulièrement important pour la transformation de Fourier, qui est une isométrie sur cet espace (cf cours 2A).

Approximation par des fonctions régulières

L'espace $L^1(\mathbb{R}^n)$ contient des fonctions très irrégulières (discontinues, non bornées, valant $+\infty$ sur des ensembles de mesure nulle...). On a donc l'impression qu'il va être compliqué de travailler dans cet espace, en comparaison avec le confort de manipuler des fonctions « lisses ». Bonne nouvelle : en fait on va voir dans le théorème suivant qu'on trouve des fonctions régulières aussi proche qu'on le souhaite d'une fonction quelconque de $L^1(\mathbb{R}^n)$.

On note $\mathcal{C}_c^0(\mathbb{R}^n)$ l'espace des fonctions continues sur \mathbb{R}^n à **support compact** (c'est-à-dire nulles en dehors d'une boule bornée).

Théorème 2.7 (Densité des fonctions continues). L'espace $\mathcal{C}_c^0(\mathbb{R}^n)$ est **dense** dans $L^1(\mathbb{R}^n)$.

Autrement dit, pour toute fonction intégrable $f \in L^1(\mathbb{R}^n)$ et pour tout $\varepsilon > 0$, il existe une fonction g continue à support compact telle que :

$$\|f - g\|_1 = \int_{\mathbb{R}^n} |f(x) - g(x)| dx < \varepsilon.$$

Idée de la preuve. La démonstration se fait en trois étapes d'approximation successives :

1. **Troncature** : On approche f par une fonction à support compact $f\mathbf{1}_{B(0,R)}$. Comme f est intégrable, le reste à l'infini est petit pour R assez grand.
2. **Fonctions étagées** : Sur ce compact, on approche f par une fonction étagée $s = \sum \alpha_i \mathbf{1}_{A_i}$ (Théorème d'approximation). Il suffit donc de savoir approcher une fonction indicatrice $\mathbf{1}_A$.
3. **Lissage** : Pour un ensemble mesurable borné A , on approche son indicatrice par des fonctions continues trapézoïdales (qui valent 1 sur A et descendent continûment vers 0 très vite à l'extérieur). Cela utilise la régularité de la mesure de Lebesgue (tout borélien peut être approché par des ouverts et des fermés).

□

Corollaire 2.2 (Densité des fonctions \mathcal{C}^∞). *L'espace des fonctions indéfiniment dérivables à support compact, noté $\mathcal{C}_c^\infty(\mathbb{R}^n)$ (ou $\mathcal{D}(\mathbb{R}^n)$), est également dense dans $L^1(\mathbb{R}^n)$.*

Remarque 2.11 (Conséquence pratique). Pour démontrer une propriété sur tout l'espace L^1 (comme par exemple le Lemme de Riemann-Lebesgue pour la transformée de Fourier, cf cours de deuxième année), il suffit souvent de :

1. Démontrer la propriété pour les fonctions $g \in \mathcal{C}_c^\infty$ (où les calculs sont faciles grâce à la régularité).
2. Étendre le résultat par densité à tout $f \in L^1$ si bien sûr cette propriété « passe à la limite » pour la norme $\|\cdot\|_1$.

2.7 Les grands théorèmes

Nous avons vu que pour des fonctions positives croissantes, la limite de l'intégrale est l'intégrale de la limite (Théorème de Convergence Monotone). Que se passe-t-il si la suite n'est pas monotone, ou si les fonctions changent de signe ?

2.7.1 Le lemme de Fatou

C'est un résultat d'inégalité, souvent utilisé comme première étape avant d'établir des égalités. Il s'applique aux fonctions positives sans hypothèse de monotonie.

Lemme 2.1 (Lemme de Fatou). *Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions mesurables **positives**. Alors :*

$$\int_X \left(\liminf_{n \rightarrow \infty} f_n \right) d\mu \leq \liminf_{n \rightarrow \infty} \left(\int_X f_n d\mu \right).$$

Démonstration. La démonstration repose sur l'utilisation du Théorème de Convergence Monotone appliqué à une suite auxiliaire bien choisie. Rappelons la définition de la limite inférieure ponctuelle :

$$f(x) = \liminf_{n \rightarrow \infty} f_n(x) = \sup_{k \in \mathbb{N}} \left(\inf_{n \geq k} f_n(x) \right).$$

Introduisons la suite de fonctions $(g_k)_{k \in \mathbb{N}}$ définie par :

$$g_k(x) = \inf_{n \geq k} f_n(x).$$

Analysons les propriétés de cette suite (g_k) : Comme les f_n sont mesurables, leur infimum dénombrable g_k est mesurable (voir Section 2). Comme $f_n \geq 0$, on a $g_k \geq 0$. La suite des ensembles sur

lesquels on prend l'infimum étant décroissante au sens de l'inclusion, la valeur de cet infimum est croissante.

La suite (g_k) est donc une suite croissante de fonctions mesurables positives qui converge simplement vers $f = \liminf f_n$. Nous pouvons appliquer le **Théorème de Convergence Monotone** à (g_k) :

$$\int_X \liminf_{n \rightarrow \infty} f_n d\mu = \int_X \lim_{k \rightarrow \infty} g_k d\mu = \lim_{k \rightarrow \infty} \int_X g_k d\mu.$$

Cherchons maintenant à majorer le terme de droite. Par définition, pour tout indice $n \geq k$, on a $g_k(x) \leq f_n(x)$. Par croissance de l'intégrale positive :

$$\forall n \geq k, \quad \int_X g_k d\mu \leq \int_X f_n d\mu.$$

Puisque cette inégalité est vraie pour tout $n \geq k$, l'intégrale de g_k est inférieure à la borne inférieure des intégrales :

$$\int_X g_k d\mu \leq \inf_{n \geq k} \left(\int_X f_n d\mu \right).$$

Passons à la limite quand $k \rightarrow \infty$ dans cette inégalité :

$$\lim_{k \rightarrow \infty} \int_X g_k d\mu \leq \lim_{k \rightarrow \infty} \left(\inf_{n \geq k} \int_X f_n d\mu \right).$$

Le membre de gauche est $\int \liminf f_n$ (d'après le point précédent) et le membre de droite est exactement la définition de $\liminf(\int f_n)$. On a donc bien démontré que :

$$\int_X \liminf_{n \rightarrow \infty} f_n d\mu \leq \liminf_{n \rightarrow \infty} \int_X f_n d\mu.$$

□

Exemple 2.7 (L'inégalité stricte est possible). Sur $X = \mathbb{R}$, prenons la suite de fonctions « bosse glissante » $f_n = \mathbf{1}_{[n, n+1]}$.

- Pour tout x fixé, $f_n(x)$ vaut 0 dès que $n > x$. Donc $\lim f_n(x) = 0$ et $\int \lim f_n = 0$.
- Pour tout n , $\int f_n = 1$. Donc $\lim \int f_n = 1$.

On a bien $0 \leq 1$. La masse peut « s'échapper à l'infini », ce que l'intégrale de la limite ne voit pas.

2.7.2 Le Théorème de Convergence Dominée (TCD)

C'est le théorème d'intégration le plus utilisé en analyse. Il donne une condition suffisante très maniable pour intervertir limite et intégrale : il suffit que la suite soit « contrôlée » par une fonction intégrable fixe.

Théorème 2.8 (Convergence Dominée de Lebesgue). Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions intégrables (à valeurs réelles ou complexes). Supposons que :

1. **Convergence simple** : La suite (f_n) converge simplement vers une fonction f presque partout sur X .
2. **Domination** : Il existe une fonction $g \in \mathcal{L}^1(X, \mu)$ (intégrable et positive) telle que :

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

Alors f est intégrable, et :

$$\lim_{n \rightarrow \infty} \int_X |f_n - f| d\mu = 0 \quad (\text{Convergence en moyenne}).$$

En particulier :

$$\lim_{n \rightarrow \infty} \int_X f_n d\mu = \int_X f d\mu.$$

Idée de la preuve. On applique le lemme de Fatou aux fonctions positives $h_n = 2g - |f_n - f|$. Comme $|f_n - f| \leq 2g$, ces fonctions sont positives. $\liminf \int (2g - |f_n - f|) \geq \int \liminf (2g - |f_n - f|) = \int 2g$. En simplifiant par $\int 2g$ (qui est fini!), on obtient $\limsup \int |f_n - f| \leq 0$, donc la convergence en limite. \square

2.7.3 Variante : Le Théorème de Convergence Dominée généralisé*

Un résultat plus général (cf Lieb & Loss [18]) permet de traiter des cas où la fonction dominante g dépend elle-même de n .

Théorème 2.9 (TCD Généralisé). Soit (f_n) une suite de fonctions intégrables convergeant p.p. vers f . Soit (g_n) une suite de fonctions intégrables convergeant p.p. vers g intégrable. Supposons que :

1. Pour tout n , $|f_n| \leq g_n$ p.p.
2. $\lim_{n \rightarrow \infty} \int_X g_n d\mu = \int_X g d\mu < +\infty$.

Alors $\int_X f_n d\mu \rightarrow \int_X f d\mu$.

Ce théorème est très utile quand g_n converge vers g dans L^1 mais n'est pas une domination uniforme fixe.

Démonstration. La preuve repose sur une double application du Lemme de Fatou. Rappelons les hypothèses : $f_n \rightarrow f$ p.p., $g_n \rightarrow g$ p.p., $|f_n| \leq g_n$ p.p., et $\int g_n \rightarrow \int g$.

1. **Première application (limite inférieure) :** Considérons la suite de fonctions $h_n = g_n + f_n$. Puisque $|f_n| \leq g_n$, on a $f_n \geq -g_n$, donc $h_n \geq 0$ presque partout. La suite (h_n) converge p.p. vers $g + f$. D'après le Lemme de Fatou :

$$\int_X (g + f) d\mu \leq \liminf_{n \rightarrow \infty} \int_X (g_n + f_n) d\mu.$$

Par linéarité de l'intégrale :

$$\int g + \int f \leq \liminf \left(\int g_n + \int f_n \right).$$

Puisque la suite numérique $\int g_n$ converge vers $\int g$, on a $\liminf (\int g_n + \int f_n) = \int g + \liminf \int f_n$. Comme $\int g$ est fini, on peut le simplifier de chaque côté :

$$\int_X f d\mu \leq \liminf_{n \rightarrow \infty} \int_X f_n d\mu.$$

2. **Seconde application (limite supérieure) :** Considérons maintenant la suite de fonctions $k_n = g_n - f_n$. On a aussi $k_n \geq 0$ p.p. (car $f_n \leq |f_n| \leq g_n$). La suite (k_n) converge p.p. vers $g - f$. D'après le Lemme de Fatou :

$$\int_X (g - f) d\mu \leq \liminf_{n \rightarrow \infty} \int_X (g_n - f_n) d\mu.$$

De même, par linéarité et convergence de $\int g_n$:

$$\int g - \int f \leq \int g + \liminf \left(- \int f_n \right).$$

Rappelons la propriété des limites : $\liminf(-u_n) = -\limsup(u_n)$. Donc :

$$- \int f \leq - \limsup \int f_n.$$

On obtient finalement :

$$\limsup_{n \rightarrow \infty} \int_X f_n d\mu \leq \int_X f d\mu.$$

Conclusion : En combinant les deux inégalités, on a encadré la limite :

$$\limsup \int f_n \leq \int f \leq \liminf \int f_n.$$

Cela implique que la suite $(\int f_n)$ converge et que sa limite est $\int f$. \square

Exemple 2.8 (Intérêt de la domination variable). Soit la suite de fonctions définies sur \mathbb{R} par $f_n(x) = \frac{1}{n} \mathbf{1}_{[n, n+1[}(x)$.

- La suite converge simplement vers 0.
- L'intégrale $\int f_n = 1/n$ converge bien vers 0.
- Cependant, le **TCD classique ne s'applique pas**. En effet, la borne supérieure des fonctions est $\sup_n |f_n(x)| \approx \frac{1}{x}$ pour x grand, fonction qui n'est pas intégrable (série harmonique).
- Le **TCD généralisé s'applique** en prenant simplement $g_n = f_n$, car on a bien la convergence des intégrales des majorants ($\int g_n \rightarrow 0$).

Ce théorème est particulièrement utile pour prouver le résultat suivant : si $f_n \rightarrow f$ p.p. et si $\|f_n\|_1 \rightarrow \|f\|_1$, alors $\|f_n - f\|_1 \rightarrow 0$ (convergence en moyenne).

Proposition 2.9 (Convergence ponctuelle et convergence des normes). Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions intégrables qui converge presque partout vers une fonction intégrable f . Si $\|f_n\|_1 \rightarrow \|f\|_1$ (c'est-à-dire $\int |f_n| \rightarrow \int |f|$), alors f_n converge vers f dans L^1 :

$$\lim_{n \rightarrow \infty} \|f_n - f\|_1 = \lim_{n \rightarrow \infty} \int_X |f_n - f| d\mu = 0.$$

Démonstration. L'idée est d'appliquer le Théorème de Convergence Dominée Généralisé à la suite de fonctions qui mesure l'écart entre f_n et f .

Posons $h_n = |f_n - f|$.

1. **Convergence simple :** Puisque $f_n \rightarrow f$ presque partout, on a clairement $h_n \rightarrow 0$ presque partout.
2. **Domination variable :** D'après l'inégalité triangulaire, on a :

$$h_n = |f_n - f| \leq |f_n| + |f|.$$

Posons $g_n = |f_n| + |f|$. C'est notre suite dominante.

3. **Convergence des intégrales des dominants** : On a évidemment $g_n \rightarrow 2|f|$ presque partout. De plus, par linéarité et par l'hypothèse de convergence des normes :

$$\int_X g_n d\mu = \int_X |f_n| d\mu + \int_X |f| d\mu \xrightarrow{n \rightarrow \infty} \int_X |f| d\mu + \int_X |f| d\mu = \int_X 2|f| d\mu.$$

La condition $\int g_n \rightarrow \int (\lim g_n)$ est donc vérifiée.

Nous sommes exactement dans les conditions d'application du Théorème de Convergence Dominée Généralisé (appliqué à la suite h_n dominée par g_n). On en déduit que :

$$\lim_{n \rightarrow \infty} \int_X h_n d\mu = \int_X (\lim_{n \rightarrow \infty} h_n) d\mu = \int_X 0 d\mu = 0.$$

C'est-à-dire $\|f_n - f\|_1 \rightarrow 0$. □

Remarque 2.12. Ce résultat est faux si on retire l'hypothèse $\int |f_n| \rightarrow \int |f|$. Reprenons l'exemple du « créneau glissant » $f_n = \mathbf{1}_{[n, n+1]}$. Elle converge vers 0 p.p., mais $\|f_n\|_1 = 1$ ne tend pas vers $\|0\|_1 = 0$. Et effectivement, $\|f_n - 0\|_1 = 1$ ne tend pas vers 0. L'hypothèse sur les normes empêche la masse de « s'échapper à l'infini » ou de se concentrer en un point de manière singulière.

2.7.4 Intégrales dépendant d'un paramètre

C'est l'application directe du TCD pour l'étude des fonctions définies par $F(t) = \int_X f(x, t) d\mu(x)$.

Théorème 2.10 (Continuité). Soit I un intervalle de \mathbb{R} et $f : X \times I \rightarrow \mathbb{R}$. Supposons que :

1. Pour tout $t \in I$, $x \mapsto f(x, t)$ est intégrable.
2. Pour presque tout $x \in X$, $t \mapsto f(x, t)$ est continue sur I .
3. Il existe une fonction $g \in \mathcal{L}^1(X)$ telle que :

$$\forall t \in I, \quad |f(x, t)| \leq g(x) \quad \text{p.p. } x \in X$$

Alors la fonction $F : t \mapsto \int_X f(x, t) d\mu(x)$ est continue sur I .

Démonstration. Soit t_* un point fixé dans l'intervalle I . Pour démontrer la continuité de F en t_* , nous allons utiliser la caractérisation séquentielle de la limite.

Soit $(t_n)_{n \in \mathbb{N}}$ une suite d'éléments de I qui converge vers t_* . On pose, pour tout $n \in \mathbb{N}$, la fonction $f_n : X \rightarrow \mathbb{R}$ définie par :

$$f_n(x) = f(x, t_n).$$

La suite de fonctions $(f_n)_{n \geq 0}$ satisfait les hypothèses du Théorème de Convergence Dominée :

1. **Convergence simple** : Pour presque tout $x \in X$, l'application $t \mapsto f(x, t)$ est continue sur I (hypothèse 2). Puisque $\lim_{n \rightarrow \infty} t_n = t_*$, on a par continuité :

$$\lim_{n \rightarrow \infty} f_n(x) = \lim_{n \rightarrow \infty} f(x, t_n) = f(x, t_*).$$

La suite (f_n) converge donc simplement presque partout vers la fonction $x \mapsto f(x, t_*)$.

2. **Domination** : D'après l'hypothèse 3, il existe une fonction $g \in \mathcal{L}^1(X)$ indépendante de t telle que pour tout $t \in I$, $|f(x, t)| \leq g(x)$ p.p. En particulier, pour tout $n \in \mathbb{N}$:

$$|f_n(x)| = |f(x, t_n)| \leq g(x) \quad \text{p.p.}$$

D'après le Théorème de Convergence Dominée, on peut passer à la limite sous l'intégrale :

$$\lim_{n \rightarrow \infty} \int_X f_n(x) d\mu(x) = \int_X \left(\lim_{n \rightarrow \infty} f_n(x) \right) d\mu(x).$$

Ce qui se réécrit, en utilisant la définition de F :

$$\lim_{n \rightarrow \infty} F(t_n) = \int_X f(x, t_*) d\mu(x) = F(t_*).$$

Ceci étant vrai pour toute suite (t_n) convergeant vers t_0 , on en déduit que F est continue en t_* . \square

Remarque 2.13. On peut remplacer la dernière hypothèse par une version localisée, en ayant toujours le même résultat de continuité sur I tout entier :

3. Pour tout compact $K \subset I$, il existe une fonction $g_K \in \mathcal{L}^1(X)$ telle que :

$$\forall t \in K, \quad |f(x, t)| \leq g_K(x) \quad \text{p.p. } x \in X$$

Théorème 2.11 (Dérivabilité sous le signe intégrale (Leibniz)). *Supposons que :*

1. Pour tout t , $x \mapsto f(x, t)$ est intégrable.
2. La dérivée partielle $\frac{\partial f}{\partial t}(x, t)$ existe pour tout t et presque tout x .
3. Il existe une fonction $g \in \mathcal{L}^1(X)$ telle que :

$$\forall t \in I, \quad \left| \frac{\partial f}{\partial t}(x, t) \right| \leq g(x) \quad \text{p.p. } x \in X$$

Alors F est dérivable sur I et :

$$F'(t) = \int_X \frac{\partial f}{\partial t}(x, t) d\mu(x).$$

Démonstration. Soit $t_* \in I$. Pour montrer que F est dérivable en t_* et calculer sa dérivée, nous étudions la limite du taux d'accroissement.

Soit $(t_n)_{n \in \mathbb{N}}$ une suite d'éléments de $I \setminus \{t_*\}$ convergeant vers t_* . Considérons la suite de fonctions $(\phi_n)_{n \in \mathbb{N}}$ définies sur X par le taux d'accroissement :

$$\phi_n(x) = \frac{f(x, t_n) - f(x, t_*)}{t_n - t_*}.$$

Nous allons appliquer le Théorème de Convergence Dominée à la suite (ϕ_n) .

1. **Convergence simple :** Pour presque tout $x \in X$, la fonction $t \mapsto f(x, t)$ est dérivable. Par définition de la dérivée partielle, on a pour presque tout x :

$$\lim_{n \rightarrow \infty} \phi_n(x) = \lim_{n \rightarrow \infty} \frac{f(x, t_n) - f(x, t_*)}{t_n - t_*} = \frac{\partial f}{\partial t}(x, t_*).$$

2. **Domination (via le Théorème des Accroissements Finis) :** Fixons x (presque partout) tel que $t \mapsto f(x, t)$ soit dérivable sur I . Pour tout n , appliquons le Théorème des Accroissements Finis (TAF) entre t_* et t_n . Il existe un réel $\theta_{n,x}$ compris strictement entre t_* et t_n tel que :

$$f(x, t_n) - f(x, t_*) = (t_n - t_*) \frac{\partial f}{\partial t}(x, \theta_{n,x}).$$

En divisant par $(t_n - t_*)$, on obtient $\phi_n(x) = \frac{\partial f}{\partial t}(x, \theta_{n,x})$.

D'après l'hypothèse 3 (domination uniforme de la dérivée partielle), on a alors :

$$|\phi_n(x)| = \left| \frac{\partial f}{\partial t}(x, \theta_{n,x}) \right| \leq g(x) \quad \text{p.p.}$$

La suite (ϕ_n) est donc dominée par la fonction intégrable g , indépendante de n .

D'après le Théorème de Convergence Dominée, la suite (ϕ_n) est intégrable et on peut intervertir limite et intégrale :

$$\lim_{n \rightarrow \infty} \int_X \phi_n(x) d\mu(x) = \int_X \left(\lim_{n \rightarrow \infty} \phi_n(x) \right) d\mu(x).$$

En réinjectant la définition de ϕ_n et sa limite, on obtient :

$$\lim_{n \rightarrow \infty} \frac{F(t_n) - F(t_0)}{t_n - t_0} = \int_X \frac{\partial f}{\partial t}(x, t_0) d\mu(x).$$

Ceci étant vrai pour toute suite (t_n) tendant vers t_0 , on conclut que F est dérivable en t_0 et que sa dérivée est donnée par l'intégrale de la dérivée partielle. \square

2.8 Mesures produits et Théorèmes de Fubini-Tonelli

Jusqu'à présent, nous avons intégré des fonctions sur un seul espace X . En pratique, on travaille souvent sur \mathbb{R}^2 ou \mathbb{R}^n , et l'on souhaite ramener le calcul d'une intégrale multiple à une succession d'intégrales simples (intégrales itérées). La théorie de Lebesgue apporte une réponse complète et plus souple que celle de Riemann pour intervertir l'ordre d'intégration.

2.8.1 La Mesure Produit

Soient (X, \mathcal{A}, μ) et (Y, \mathcal{B}, ν) deux espaces mesurés σ -finis (c'est-à-dire qu'on peut recouvrir l'espace par une union dénombrable d'ensembles de mesure finie, ce qui est toujours le cas pour \mathbb{R}^n avec la mesure de Lebesgue).

On munit l'espace produit $X \times Y$ de la **tribu produit** $\mathcal{A} \otimes \mathcal{B}$, engendrée par les « rectangles mesurables » $A \times B$ (où $A \in \mathcal{A}, B \in \mathcal{B}$).

Théorème 2.12 (Existence et Unicité de la mesure produit). *Il existe une unique mesure, notée $\mu \otimes \nu$, définie sur la tribu produit $\mathcal{A} \otimes \mathcal{B}$, telle que pour tout rectangle mesurable :*

$$(\mu \otimes \nu)(A \times B) = \mu(A) \times \nu(B).$$

Dans le cas où $X = Y = \mathbb{R}$ munis de la mesure de Lebesgue λ , la mesure produit $\lambda \otimes \lambda$ correspond bien à la mesure de Lebesgue bidimensionnelle sur \mathbb{R}^2 (celle qui donne l'aire).

2.8.2 Le Théorème de Tonelli (Fonctions positives)

Ce théorème concerne les fonctions positives. Il affirme que l'ordre d'intégration n'a aucune importance, même si l'intégrale vaut $+\infty$.

Théorème 2.13 (Tonelli). *Soit $f : X \times Y \rightarrow [0, +\infty]$ une fonction mesurable **positive**. Alors les fonctions partielles $x \mapsto \int_Y f(x, y) d\nu(y)$ et $y \mapsto \int_X f(x, y) d\mu(x)$ sont mesurables, et on a l'égalité :*

$$\begin{aligned} \int_{X \times Y} f(x, y) d(\mu \otimes \nu) &= \int_X \left(\int_Y f(x, y) d\nu(y) \right) d\mu(x) \\ &= \int_Y \left(\int_X f(x, y) d\mu(x) \right) d\nu(y). \end{aligned}$$

Remarque 2.14. Avec Tonelli, tout est permis : on peut intervertir dx et dy sans scrupule tant que la fonction est positive. Le résultat peut être fini ou infini.

💡 Application : Principe du “Mille-feuilles”

Il existe une jolie manière géométrique de voir l'intégrale de Lebesgue. En coupant le graphe de f horizontalement, on obtient la formule :

$$\int_X f(x) d\mu(x) = \int_0^{+\infty} \mu(\{x \in X \mid f(x) > t\}) dt.$$

L'intégrale de Lebesgue est simplement l'aire sous la courbe, calculée « tranche par tranche » horizontalement (intégrale de la fonction $t \mapsto \mu(\{f > t\})$).

Démonstration. La démonstration repose sur une astuce d'écriture fondamentale. Pour tout nombre réel $y \geq 0$, on peut écrire y comme l'intégrale de 1 entre 0 et y :

$$y = \int_0^y 1 dt = \int_0^{+\infty} \mathbf{1}_{\{t < y\}} dt.$$

Appliquons cela à $y = f(x)$ pour tout $x \in X$ fixé :

$$f(x) = \int_0^{+\infty} \mathbf{1}_{\{t < f(x)\}} dt.$$

Intégrons maintenant cette égalité par rapport à la mesure μ sur X :

$$\int_X f(x) d\mu(x) = \int_X \left(\int_0^{+\infty} \mathbf{1}_{\{f(x) > t\}} dt \right) d\mu(x).$$

On se retrouve avec une intégrale double (sur $X \times \mathbb{R}_+$). Puisque la fonction $(x, t) \mapsto \mathbf{1}_{\{f(x) > t\}}$ est positive et mesurable, on peut d'après le **théorème de Tonelli** intervertir l'ordre d'intégration :

$$\int_X f(x) d\mu(x) = \int_0^{+\infty} \left(\int_X \mathbf{1}_{\{f(x) > t\}} d\mu(x) \right) dt.$$

Or, l'intégrale intérieure n'est rien d'autre que la mesure de l'ensemble où la fonction est strictement supérieure à t :

$$\int_X \mathbf{1}_{\{f(x) > t\}} d\mu(x) = \mu(\{x \in X \mid f(x) > t\}).$$

D'où le résultat :

$$\int_X f(x) d\mu(x) = \int_0^{+\infty} \mu(\{f > t\}) dt.$$

□

2.8.3 Le Théorème de Fubini (Fonctions intégrables)

Pour les fonctions de signe quelconque, il faut être prudent. On ne peut intervertir les intégrales que si la fonction est **intégrable** par rapport à la mesure produit.

Théorème 2.14 (Fubini). Soit $f : X \times Y \rightarrow \mathbb{R}$ (ou \mathbb{C}) une fonction mesurable. Supposons que f soit intégrable pour la mesure produit, c'est-à-dire que :

$$\int_{X \times Y} |f(x, y)| d(\mu \otimes \nu) < +\infty.$$

Alors :

1. Pour presque tout $x \in X$, la fonction $y \mapsto f(x, y)$ est intégrable sur Y .
2. La fonction $x \mapsto \int_Y f(x, y) d\nu(y)$ (définie p.p.) est intégrable sur X .
3. On a l'égalité des intégrales itérées :

$$\int_{X \times Y} f d(\mu \otimes \nu) = \int_X \left(\int_Y f d\nu \right) d\mu = \int_Y \left(\int_X f d\mu \right) d\nu.$$

2.8.4 Méthode pratique : Le Théorème de Fubini-Tonelli

En pratique, pour calculer une intégrale sur un espace produit $\iint f$, on ne sait pas a priori si f est intégrable. On utilise donc les deux théorèmes conjointement selon la procédure suivante, souvent appelé « Théorème de Fubini-Tonelli » :

1. **Test d'intégrabilité (Tonelli) :** On calcule l'intégrale de la valeur absolue $|f|$ dans l'ordre le plus simple (par exemple d'abord en y , puis en x).

$$I_{abs} = \int_X \left(\int_Y |f(x, y)| d\nu(y) \right) d\mu(x).$$

2. **Conclusion :**

- Si $I_{abs} < +\infty$, alors $f \in L^1$. On a le droit d'appliquer Fubini à f (sans les valeurs absolues) et de calculer l'intégrale dans l'ordre que l'on veut (par exemple l'inverse de celui utilisé pour le test, si c'est plus simple).
- Si $I_{abs} = +\infty$, alors f n'est pas intégrable au sens de Lebesgue. L'intégrale double n'est pas définie (même si les intégrales itérées existent et se compensent, voir contre-exemple ci-dessous).

Exemple 2.9 (Contre-exemple : Pourquoi l'intégrabilité est cruciale). Considérons la fonction sur $]0, 1[^2$ définie par $f(x, y) = \frac{x^2 - y^2}{(x^2 + y^2)^2}$. On peut montrer par calcul direct (primitives de la forme $\frac{y}{x^2 + y^2}$) que :

$$\int_0^1 \left(\int_0^1 f(x, y) dy \right) dx = \frac{\pi}{4} \quad \text{mais} \quad \int_0^1 \left(\int_0^1 f(x, y) dx \right) dy = -\frac{\pi}{4}.$$

Fubini ne s'applique pas car $\iint |f| = +\infty$. La fonction f n'est donc pas intégrable sur le carré.

2.8.5 Application : Le produit de convolution

L'une des applications les plus importantes du théorème de Fubini-Tonelli en ingénierie (traitement du signal, automatique) et en probabilités (somme de variables aléatoires) est la convolution. C'est une opération qui « mélange » deux fonctions.

Définition 2.22 (Produit de convolution). Soient f et g deux fonctions de $L^1(\mathbb{R}^n)$. On appelle **produit de convolution** de f et g , noté $f * g$, la fonction définie par :

$$(f * g)(x) = \int_{\mathbb{R}^n} f(x - y)g(y) dy.$$

À première vue, il n'est pas évident que cette intégrale soit finie pour presque tout x . C'est le théorème suivant, conséquence directe de Fubini, qui l'assure.

Théorème 2.15 (Inégalité de Young pour L^1). Si $f \in L^1(\mathbb{R}^n)$ et $g \in L^1(\mathbb{R}^n)$, alors :

1. La fonction $y \mapsto f(x-y)g(y)$ est intégrable pour presque tout x .
2. La fonction $f * g$ appartient à $L^1(\mathbb{R}^n)$.
3. On a l'inégalité des normes :

$$\|f * g\|_1 \leq \|f\|_1 \|g\|_1.$$

Démonstration. On cherche à montrer que la fonction $x \mapsto (f * g)(x)$ est intégrable sur \mathbb{R}^n . Évaluons l'intégrale de sa valeur absolue en utilisant le théorème de **Tonelli** (car tout est positif) :

$$I = \int_{\mathbb{R}^n} |(f * g)(x)| dx = \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} f(x-y)g(y) dy \right| dx.$$

Par inégalité triangulaire :

$$I \leq \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |f(x-y)||g(y)| dy \right) dx.$$

On intervertit l'ordre d'intégration (Tonelli) :

$$I \leq \int_{\mathbb{R}^n} |g(y)| \left(\int_{\mathbb{R}^n} |f(x-y)| dx \right) dy.$$

Dans l'intégrale intérieure (en x), on effectue le changement de variable $z = x - y$ (à y fixé). Comme la mesure de Lebesgue est invariante par translation ($dx = dz$) :

$$\int_{\mathbb{R}^n} |f(x-y)| dx = \int_{\mathbb{R}^n} |f(z)| dz = \|f\|_1.$$

On réinjecte ce résultat :

$$I \leq \int_{\mathbb{R}^n} |g(y)| \|f\|_1 dy = \|f\|_1 \int_{\mathbb{R}^n} |g(y)| dy = \|f\|_1 \|g\|_1.$$

L'intégrale I est finie, donc $f * g$ est bien définie (p.p.) et est intégrable. \square

Remarque 2.15. Par un changement de variables très simple (cf section suivante), il est évident que la convolution est commutative : $f * g = g * f$.

Remarque 2.16 (Interprétation physique). — **Régularisation :** La convolution est une opération de « lissage ». Si g est une fonction très régulière (par exemple une gaussienne), alors $f * g$ sera régulière même si f est discontinue ou bruitée. On peut se servir de cette propriété pour démontrer les théorèmes de densité des espaces de fonctions continues dans les espaces L^p . Voir à ce sujet [15].

— **Signal :** Si f est un signal d'entrée et g la réponse impulsionnelle d'un système linéaire invariant, alors la sortie est $f * g$.

2.9 Changement de variables dans \mathbb{R}^n

Le calcul d'intégrales multiples est souvent simplifié par un changement de système de coordonnées (passer des cartésiennes aux polaires, par exemple) qui rend le domaine d'intégration plus simple (un carré ou un rectangle) ou la fonction plus symétrique.

2.9.1 Difféomorphismes et résultat principal

Soient U et V deux ouverts de \mathbb{R}^n . Soit $\phi : U \rightarrow V$ une application bijective. On dit que ϕ est un \mathcal{C}^1 -difféomorphisme si :

1. ϕ est de classe \mathcal{C}^1 sur U .
2. Son inverse ϕ^{-1} est de classe \mathcal{C}^1 sur V .

Cela implique que la matrice jacobienne de ϕ en tout point $x \in U$, notée $J_\phi(x)$, est inversible. On note $\det J_\phi(x)$ son déterminant (le jacobien).

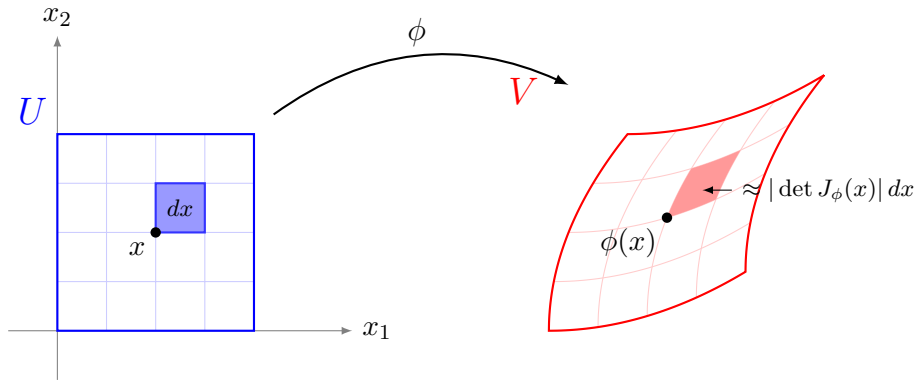


Figure 2.3 – Illustration du changement de variables. L'application ϕ transforme le domaine U (maillage cartésien) en un domaine V (maillage curviligne). Un petit élément de volume dx autour de x est transformé en un élément d'aire $|\det J_\phi(x)|dx$ autour de $\phi(x)$.

Théorème 2.16 (Formule de changement de variables). Soit $\phi : U \rightarrow V$ un \mathcal{C}^1 -difféomorphisme. Soit $f : V \rightarrow \mathbb{R}$ une fonction mesurable.

1. Si f est à valeurs positives, on a toujours :

$$\int_V f(y) dy = \int_U f(\phi(x)) |\det J_\phi(x)| dx.$$

2. Dans le cas général, f est intégrable sur V si et seulement si la fonction $(f \circ \phi) \cdot |\det J_\phi|$ est intégrable sur U . Dans ce cas, on a la même formule que ci-dessus.

Esquisse de la démonstration. La preuve se déroule en quatre étapes classiques, allant du cas géométrique simple au cas général fonctionnel.

Étape 1 : Le cas des fonctions indicatrices Supposons que $f = \mathbf{1}_K$ soit la fonction indicatrice d'un petit cube K inclus dans V (ou d'un borélien). Le membre de gauche est le volume de K : $\int_V \mathbf{1}_K(y) dy = \lambda(K)$. Le membre de droite est $\int_{\phi^{-1}(K)} |\det J_\phi(x)| dx$. Il s'agit donc de prouver que pour un domaine $A = \phi^{-1}(K) \subset U$, le volume de l'image $\phi(A)$ est donné par l'intégrale du jacobien sur A :

$$\lambda(\phi(A)) = \int_A |\det J_\phi(x)| dx.$$

- **Approximation affine** : Localement, autour d'un point x_0 , l'application ϕ se comporte comme son approximation affine : $\phi(x) \approx \phi(x_0) + J_\phi(x_0)(x - x_0)$.
- **Algèbre linéaire** : On sait (résultat d'algèbre) qu'une application linéaire L transforme un cube de volume v en un parallélépipède de volume $|\det L| \cdot v$.

- **Passage au global** : On découpe l'ouvert A en une réunion dénombrable de cubes disjoints arbitrairement petits. Sur chacun d'eux, on approxime ϕ par sa différentielle. En passant à la limite (somme de Riemann ou recouvrement de Vitali), on obtient la formule pour les volumes.

Étape 2 : Fonctions étagées (Linéarité) Si f est une fonction étagée positive, elle s'écrit comme une combinaison linéaire finie d'indicatrices : $f = \sum \alpha_i \mathbf{1}_{A_i}$. Par linéarité de l'intégrale (à gauche et à droite), la formule établie à l'étape 1 reste vraie pour f .

Étape 3 : Fonctions mesurables positives (Convergence Monotone) Soit f une fonction mesurable positive quelconque. D'après le **Théorème d'approximation** (Théorème 2.2), il existe une suite croissante de fonctions étagées positives (s_n) qui converge simplement vers f . On a l'égalité pour chaque s_n :

$$\int_V s_n(y) dy = \int_U s_n(\phi(x)) |\det J_\phi(x)| dx.$$

On applique le Théorème 2.4 de **Convergence Monotone** (ou Beppo-Levi) des deux côtés.

- À gauche : $\lim \int s_n = \int \lim s_n = \int f$.
- À droite : La suite $x \mapsto s_n(\phi(x)) |\det J_\phi(x)|$ est aussi croissante et converge vers $f(\phi(x)) |\det J_\phi(x)|$. L'intégrale converge donc vers l'intégrale de la limite.

La formule est donc vraie pour toute fonction mesurable positive (point 1 du théorème).

Étape 4 : Cas général (Fonctions intégrables) Soit f une fonction de signe quelconque. On décompose $f = f^+ - f^-$. La condition d'intégrabilité est $\int_V |f| < \infty$. D'après l'étape 3 appliquée à $|f|$ (qui est positive), cela équivaut à :

$$\int_U |f(\phi(x))| \cdot |\det J_\phi(x)| dx < \infty.$$

Ceci prouve l'équivalence des intégrabilités (point 2 du théorème). Si cette condition est vérifiée, on applique la formule à f^+ et f^- séparément, puis on soustrait les deux égalités (ce qui est légitime car les intégrales sont finies) pour obtenir le résultat final. \square

Remarque 2.17 (Attention à la valeur absolue!). Contrairement à l'intégrale simple de Riemann $\int_a^b \varphi(t) u'(t) dt$ où le signe change si l'on renverse les bornes ($a > b$), l'intégrale de Lebesgue se calcule sur un **ensemble** (sans orientation). Le terme $|\det J_\phi(x)|$ représente le facteur de dilatation du volume local : un petit cube de volume dx est envoyé sur un petit parallélépipède de volume $|\det J_\phi(x)| dx$. Ce volume doit être compté positivement.

Remarque 2.18 (Négligeabilité des bords). En pratique, on intègre souvent sur des domaines fermés (boules, cubes) et le difféomorphisme peut être singulier sur le bord (comme en polaires à l'origine). Comme les frontières de domaines réguliers dans \mathbb{R}^n sont de mesure de Lebesgue nulle, on peut restreindre l'intégration aux ouverts intérieurs sans changer la valeur de l'intégrale.

2.9.2 Exemples

2.9.2.1 Commutativité de la convolution

On a défini la convolution à la section précédente par

$$f * g(x) = \int_{\mathbb{R}^n} f(x - y) g(y) dy.$$

En opérant le changement de variable $z = x - y$, de Jacobien $(-1)^n$, on a

$$f * g(x) = \int_{\mathbb{R}^n} f(z) g(x - z) dz = g * f(x)$$

2.9.2.2 Coordonnées Polaires (Plan \mathbb{R}^2)

Pour intégrer sur un disque ou une portion de disque, on pose :

$$\begin{cases} x = r \cos \theta \\ y = r \sin \theta \end{cases}$$

L'application $\phi(r, \theta) = (r \cos \theta, r \sin \theta)$ envoie le rectangle $]0, R[\times]0, 2\pi[$ sur le disque épointé et fendu $D(0, R) \setminus ([0, R] \times \{0\})$. La matrice jacobienne est :

$$J_\phi(r, \theta) = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix}.$$

Son déterminant vaut $r(\cos^2 \theta + \sin^2 \theta) = r > 0$. La formule de changement de variable devient :

$$dx dy = r dr d\theta.$$

Exemple 2.10. Calcul de l'intégrale de Gauss $I = \int_{\mathbb{R}^2} e^{-(x^2+y^2)} dx dy$.

$$I = \int_0^{2\pi} \int_0^{+\infty} e^{-r^2} r dr d\theta = 2\pi \left[-\frac{1}{2} e^{-r^2} \right]_0^{+\infty} = \pi.$$

2.9.2.3 Coordonnées Sphériques (Espace \mathbb{R}^3)

Pour intégrer sur une boule, on utilise les coordonnées sphériques (rayon r , longitude φ , colatitude θ) :

$$\begin{cases} x = r \sin \theta \cos \varphi \\ y = r \sin \theta \sin \varphi \\ z = r \cos \theta \end{cases}$$

avec $(r, \theta, \varphi) \in]0, +\infty[\times]0, \pi[\times]0, 2\pi[$. Le calcul du déterminant jacobien (matrice 3×3) donne :

$$\det J_\phi(r, \theta, \varphi) = r^2 \sin \theta.$$

Comme $\theta \in]0, \pi[$, $\sin \theta > 0$, donc la valeur absolue est inutile. L'élément de volume est :

$$dx dy dz = r^2 \sin \theta dr d\theta d\varphi.$$

Exemple 2.11 (Volume de la boule). Calculons le volume de la boule unité B_3 . On intègre la fonction $f = 1$.

$$\text{Vol}(B_3) = \iiint_{B_3} 1 dx dy dz = \int_0^1 r^2 dr \times \int_0^\pi \sin \theta d\theta \times \int_0^{2\pi} d\varphi.$$

Les intégrales se séparent (domaine rectangle en sphérique) :

$$\text{Vol}(B_3) = \left[\frac{r^3}{3} \right]_0^1 \times [-\cos \theta]_0^\pi \times 2\pi = \frac{1}{3} \times 2 \times 2\pi = \frac{4}{3}\pi.$$

Chapitre 3

Calcul différentiel et Théorème des fonctions implicites

Le calcul différentiel généralise la notion de dérivée des fonctions d'une variable réelle aux fonctions définies sur des espaces vectoriels normés de dimension quelconque (finie ou infinie). L'idée centrale est l'**approximation linéaire** : on cherche à approcher localement une fonction non linéaire par une application linéaire continue. Ce chapitre pose les bases nécessaires à l'optimisation (conditions d'extremum) et à l'étude des contraintes (théorème des fonctions implicites), qui sont centrales pour de nombreux problèmes d'ingénierie.

Dans ce chapitre, E , F et G désignent des espaces de Banach sur \mathbb{R} . $\mathcal{L}(E, F)$ est l'espace des applications linéaires continues de E dans F , muni de la norme subordonnée.

3.1 La notion de différentielle

3.1.1 Différentiabilité au sens de Fréchet

Contrairement à la dérivée directionnelle (Gâteaux) qui regarde la variation le long d'une droite, la différentielle de Fréchet capture la variation de la fonction dans toutes les directions simultanément.

Définition 3.1 (Différentiabilité). Soit U un ouvert de E et $f : U \rightarrow F$. On dit que f est **différentiable** en un point $a \in U$ s'il existe une application linéaire continue $L \in \mathcal{L}(E, F)$ telle que :

$$f(a + h) = f(a) + L(h) + o(\|h\|) \quad \text{quand } h \rightarrow 0.$$

Autrement dit,

$$\lim_{h \rightarrow 0, h \neq 0} \frac{\|f(a + h) - f(a) - L(h)\|_F}{\|h\|_E} = 0.$$

Si une telle application L existe, elle est unique. On l'appelle la **différentielle** de f en a , notée $Df(a)$ ou $f'(a)$.

Proposition 3.1 (Unicité). *Si elle existe, la différentielle est unique.*

Démonstration. Supposons qu'il existe deux applications linéaires continues L_1 et L_2 vérifiant la définition. Par inégalité triangulaire :

$$\frac{\|L_1(h) - L_2(h)\|}{\|h\|} \leq \frac{\|f(a + h) - f(a) - L_1(h)\|}{\|h\|} + \frac{\|f(a + h) - f(a) - L_2(h)\|}{\|h\|}.$$

Quand $h \rightarrow 0$, le membre de droite tend vers 0. Ainsi, pour tout vecteur $v \in E$ non nul, en posant $h = tv$ avec $t \rightarrow 0$, on obtient $\|L_1(v) - L_2(v)\| = 0$, donc $L_1 = L_2$. \square

Remarque 3.1. L'objet $Df(a)$ est une *application linéaire*. La valeur de cette application pour un vecteur $h \in E$ est notée $Df(a)(h)$ ou $Df(a) \cdot h$. C'est un vecteur de F qui représente la variation de f au premier ordre dans la direction h .

Exemple 3.1. 1. **Fonction affine :** Si $f(x) = L(x) + c$ où $L \in \mathcal{L}(E, F)$ et $c \in F$, alors pour tout $a \in E$, $Df(a) = L$. La dérivée est constante.

2. **Forme quadratique :** Soit $E = \mathbb{R}^n$ et $f(x) = \langle Ax, x \rangle$ avec A une matrice symétrique. Alors $f(a+h) = \langle A(a+h), a+h \rangle = \langle Aa, a \rangle + 2\langle Aa, h \rangle + \langle Ah, h \rangle$. La partie linéaire en h est $2\langle Aa, h \rangle$, donc $Df(a)(h) = \langle 2Aa, h \rangle$.

Définition 3.2 (Fonctions de classe \mathcal{C}^1). Soit $f : U \rightarrow F$ une application définie sur un ouvert U de E . On dit que f est de **classe \mathcal{C}^1** sur U (ou continûment différentiable) si f est différentiable en tout point de U et si l'application dérivée :

$$\begin{aligned} Df : U &\rightarrow \mathcal{L}(E, F) \\ x &\mapsto Df(x) \end{aligned}$$

est continue sur U (pour la topologie de la norme des opérateurs). Rappelons que $\mathcal{L}(E, F)$ est un espace de Banach (puisque F l'est), pour la norme :

$$\|L\| = \sup_{x \neq 0} \frac{\|L(x)\|_F}{\|x\|_E} = \sup_{\|x\|_E \leq 1} \|L(x)\|_F.$$

3.1.2 Lien avec les dérivées partielles

Proposition 3.2. Soit $E = \mathbb{R}^n$ et $F = \mathbb{R}^m$. Si f est différentiable en a , alors ses dérivées partielles existent et la matrice de $Df(a)$ dans les bases canoniques est la **matrice jacobienne** $J_f(a) = \left(\frac{\partial f_i}{\partial x_j}(a) \right)_{i,j}$ c'est à dire :

$$J_f(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \dots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix}.$$

Démonstration. Soit (e_1, \dots, e_n) la base canonique de \mathbb{R}^n . Appliquons la définition de la différentiabilité avec $h = te_j$ ($t \in \mathbb{R}^*$) :

$$f(a + te_j) - f(a) = Df(a)(te_j) + o(t) = tDf(a)(e_j) + o(t).$$

En divisant par t et en faisant tendre t vers 0, on obtient :

$$\lim_{t \rightarrow 0} \frac{f(a + te_j) - f(a)}{t} = Df(a)(e_j).$$

Le terme de gauche est exactement la définition de la dérivée partielle $\frac{\partial f}{\partial x_j}(a)$. Ainsi, la j -ème colonne de la matrice de $Df(a)$ est constituée des dérivées partielles par rapport à x_j . \square

Remarque 3.2 (Attention). L'existence des dérivées partielles ne suffit pas à assurer la différentiabilité (sauf si les dérivées partielles sont continues, voir définition de la classe \mathcal{C}^1 ci-dessus).

Exemple 3.2 (Contre-exemple : Dérivées partielles sans différentiabilité). Considérons la fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par :

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

1. **Existence des dérivées partielles en $(0,0)$** : Regardons les restrictions de f aux axes.

$$f(x, 0) = \frac{x \cdot 0}{x^2 + 0} = 0 \quad \text{et} \quad f(0, y) = 0.$$

Ainsi, les fonctions partielles $x \mapsto f(x, 0)$ et $y \mapsto f(0, y)$ sont constantes (nulles). Leurs dérivées en 0 existent et valent 0 :

$$\frac{\partial f}{\partial x}(0, 0) = 0, \quad \frac{\partial f}{\partial y}(0, 0) = 0.$$

2. **Non-différentiabilité** : Si f était différentiable en $(0, 0)$, sa différentielle serait l'application linéaire nulle (car sa matrice dans la base canonique serait la matrice nulle des dérivées partielles). Cela impliquerait que f est continue en $(0, 0)$. Or, regardons la limite de f en $(0, 0)$ le long de la droite $y = x$. Pour $x \neq 0$:

$$f(x, x) = \frac{x^2}{x^2 + x^2} = \frac{1}{2}.$$

Quand $x \rightarrow 0$, $f(x, x) \rightarrow 1/2 \neq f(0, 0)$. La fonction n'est pas continue en $(0, 0)$, elle ne peut donc pas y être différentiable.

Ce contre-exemple montre que l'existence de variations selon les axes ne suffit pas à contrôler la variation dans toutes les directions. Nous avons en revanche le résultat suivant :

Théorème 3.1 (Condition suffisante de différentiabilité). Soit U un ouvert de \mathbb{R}^n et $f : U \rightarrow \mathbb{R}^m$. Si les dérivées partielles $\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}$ existent en tout point de U et sont **continues** sur U , alors f est différentiable sur U (et est donc de classe \mathcal{C}^1).

Démonstration. Pour simplifier les notations, plaçons-nous dans le cas où $m = 1$ (fonction à valeurs réelles). Le cas vectoriel se traite en appliquant le résultat composante par composante. Soit $a \in U$. Considérons un vecteur $h = (h_1, \dots, h_n)$ suffisamment petit pour que $a + h \in U$. Nous voulons estimer l'écart $f(a + h) - f(a)$. L'idée est de se déplacer de a à $a + h$ en changeant une coordonnée à la fois (chemin "en escalier"). Notons $v_0 = 0$ et $v_k = (h_1, \dots, h_k, 0, \dots, 0)$ pour $k = 1 \dots n$. Ainsi $a + v_0 = a$ et $a + v_n = a + h$. On écrit la somme télescopique :

$$f(a + h) - f(a) = \sum_{k=1}^n [f(a + v_k) - f(a + v_{k-1})].$$

Le terme général de cette somme ne fait varier que la k -ième coordonnée. En considérant la fonction d'une variable réelle $t \mapsto f(a + v_{k-1} + te_k)$, nous pouvons appliquer le Théorème des Accroissements Finis (TAF) entre 0 et h_k . Il existe un réel $\theta_k \in]0, 1[$ tel que :

$$f(a + v_k) - f(a + v_{k-1}) = \frac{\partial f}{\partial x_k}(c_k) \cdot h_k,$$

où $c_k = a + v_{k-1} + \theta_k h_k e_k$ est un point intermédiaire sur le segment reliant $a + v_{k-1}$ à $a + v_k$. On a donc :

$$f(a + h) - f(a) = \sum_{k=1}^n \frac{\partial f}{\partial x_k}(c_k) h_k.$$

Le candidat pour la différentielle est $L(h) = \sum_{k=1}^n \frac{\partial f}{\partial x_k}(a) h_k$. Estimons l'erreur :

$$f(a+h) - f(a) - L(h) = \sum_{k=1}^n \left(\frac{\partial f}{\partial x_k}(c_k) - \frac{\partial f}{\partial x_k}(a) \right) h_k.$$

Par l'inégalité de Cauchy-Schwarz (ou l'équivalence des normes) :

$$|f(a+h) - f(a) - L(h)| \leq \|h\|_\infty \sum_{k=1}^n \left| \frac{\partial f}{\partial x_k}(c_k) - \frac{\partial f}{\partial x_k}(a) \right|.$$

Lorsque $h \rightarrow 0$, tous les points intermédiaires c_k tendent vers a . Comme les dérivées partielles $\frac{\partial f}{\partial x_k}$ sont supposées **continues** en a , le terme $\left| \frac{\partial f}{\partial x_k}(c_k) - \frac{\partial f}{\partial x_k}(a) \right|$ tend vers 0. On a donc bien $f(a+h) = f(a) + L(h) + o(\|h\|)$, ce qui prouve que f est différentiable en a . \square

3.2 Opérations et calculs

3.2.1 Règles de calcul

Le calcul différentiel respecte les opérations algébriques usuelles.

Proposition 3.3. Soient $f, g : U \subset E \rightarrow F$ différentiables en a . $\forall \alpha, \beta \in \mathbb{R}$, $\alpha f + \beta g$ est différentiable en a et $D(\alpha f + \beta g)(a) = \alpha Df(a) + \beta Dg(a)$.

3.2.2 Composition (Chain Rule)

La formule de composition un outil puissant du calcul différentiel. En termes de matrices jacobiniennes, elle correspond simplement au produit matriciel.

Théorème 3.2. Soient $f : U \subset E \rightarrow F$ différentiable en a , et $g : V \subset F \rightarrow G$ différentiable en $b = f(a)$. Alors $g \circ f$ est différentiable en a et :

$$D(g \circ f)(a) = Dg(f(a)) \circ Df(a).$$

Démonstration. On écrit les développements limités : 1. $f(a+h) = f(a) + Df(a)(h) + \|h\|\varepsilon_f(h)$ avec $\varepsilon_f(h) \rightarrow 0$. 2. $g(b+k) = g(b) + Dg(b)(k) + \|k\|\varepsilon_g(k)$ avec $\varepsilon_g(k) \rightarrow 0$.

Posons $k = f(a+h) - f(a) = Df(a)(h) + \|h\|\varepsilon_f(h)$. Notons que $\|k\| = O(\|h\|)$ car $Df(a)$ est continue.

$$(g \circ f)(a+h) - (g \circ f)(a) = g(b+k) - g(b) = Dg(b)(k) + \|k\|\varepsilon_g(k).$$

En remplaçant k par son expression :

$$\begin{aligned} &= Dg(b)[Df(a)(h) + \|h\|\varepsilon_f(h)] + \|k\|\varepsilon_g(k) \\ &= [Dg(b) \circ Df(a)](h) + \|h\| \underbrace{\left(Dg(b)(\varepsilon_f(h)) + \frac{\|k\|}{\|h\|} \varepsilon_g(k) \right)}_{\rightarrow 0}. \end{aligned}$$

Le terme en accolade tend vers 0 car $Df(a)$ et $Dg(b)$ sont continues (donc bornées) et $\varepsilon_g(k) \rightarrow 0$ quand $h \rightarrow 0$ (car $k \rightarrow 0$). Cela prouve la différentiabilité et la formule. \square

3.2.3 Inégalité des Accroissements Finis (IAF)

En dimension supérieure à 1, il n'y a pas d'égalité des accroissements finis (pensez à l'hélice dans le plan complexe $t \mapsto e^{it}$), mais une inégalité fondamentale pour contrôler les variations de la fonction.

Théorème 3.3 (IAF). *Soit $f : U \rightarrow F$ différentiable sur un ouvert convexe U . Si la différentielle est bornée sur U , c'est-à-dire s'il existe $M \geq 0$ tel que $\forall x \in U, \|Df(x)\| \leq M$, alors :*

$$\forall a, b \in U, \|f(b) - f(a)\|_F \leq M\|b - a\|_E.$$

Démonstration. Soient $a, b \in U$. On définit $\phi : [0, 1] \rightarrow F$ par $\phi(t) = f((1-t)a + tb)$. Par composition, ϕ est dérivable sur $[0, 1]$ et $\phi'(t) = Df((1-t)a + tb) \cdot (b - a)$. On a donc $\|\phi'(t)\| \leq \|Df(\dots)\| \|b - a\| \leq M\|b - a\|$. Il suffit alors d'appliquer l'inégalité des accroissements finis pour les fonctions d'une variable réelle à valeurs vectorielles (démontrée par exemple par intégration ou par raisonnement direct sur la norme) : $\|\phi(1) - \phi(0)\| \leq \sup_{t \in [0, 1]} \|\phi'(t)\| |1 - 0|$. D'où $\|f(b) - f(a)\| \leq M\|b - a\|$. \square

Ce résultat permet notamment de montrer qu'une fonction dont la dérivée est nulle sur un connexe est constante.

3.3 Ordre supérieur : Théorème de Schwarz

Si l'application $x \mapsto Df(x)$ est elle-même différentiable, on peut définir la dérivée seconde $D^2f(a)$.

Définition 3.3 (Fonctions de classe \mathcal{C}^2). Soit $f : U \rightarrow F$ une application définie sur un ouvert U de E . On dit que f est de **classe \mathcal{C}^2** sur U si f est différentiable sur U et si l'application $Df : U \rightarrow \mathcal{L}(E, F)$ est elle-même de classe \mathcal{C}^1 . Cela équivaut à demander que f soit deux fois différentiable sur U et que l'application dérivée seconde :

$$\begin{aligned} D^2f : U &\rightarrow \mathcal{L}(E, \mathcal{L}(E, F)) \\ x &\mapsto D^2f(x) \end{aligned}$$

soit continue sur U .

Il est pratique d'identifier $\mathcal{L}(E, \mathcal{L}(E, F))$ à $\mathcal{L}(E \times E, F)$, c'est à dire qu'on peut considérer $D^2f(a)$ comme une application bilinéaire continue de $E \times E$ dans F . On note alors $D^2f(a)(h, k)$ la dérivée seconde appliquée aux vecteurs h et k .

Théorème 3.4 (Schwarz). *Si f est deux fois différentiable en a et que sa dérivée seconde est continue, alors la forme bilinéaire $D^2f(a)$ est **symétrique** :*

$$\forall h, k \in E, D^2f(a)(h, k) = D^2f(a)(k, h).$$

Démonstration. L'idée est d'utiliser les accroissements croisés. Pour t petit, considérons la quantité :

$$\Delta(t) = f(a + th + tk) - f(a + th) - f(a + tk) + f(a).$$

Fixons t et posons $\varphi(x) = f(x + tk) - f(x)$. Alors $\Delta(t) = \varphi(a + th) - \varphi(a)$. D'après l'IAF appliquée à φ (dont la dérivée est $\varphi'(x) \cdot u = Df(x + tk)u - Df(x)u$), on peut montrer que $\Delta(t)$ est proche de $t(\varphi'(a) \cdot h) = t(Df(a + tk)h - Df(a)h)$. En réutilisant la définition de la dérivée pour Df en a (appliquée à tk), on a $Df(a + tk)h - Df(a)h \approx D^2f(a)(k, h) \cdot t$. Finalement, $\Delta(t) \approx t^2 D^2f(a)(k, h)$. Comme $\Delta(t)$ est symétrique en h et k par construction, on en déduit en divisant par t^2 et $t \rightarrow 0$ que $D^2f(a)(k, h) = D^2f(a)(h, k)$. \square

Théorème 3.5 (Formule de Taylor-Young). Si f est de classe \mathcal{C}^2 au voisinage de a , alors :

$$f(a+h) = f(a) + Df(a)(h) + \frac{1}{2}D^2f(a)(h,h) + o(\|h\|^2).$$

Démonstration. Considérons la fonction d'une variable réelle $\phi : [0, 1] \rightarrow F$ définie par :

$$\phi(t) = f(a+th).$$

Puisque f est de classe \mathcal{C}^2 sur un ouvert contenant le segment $[a, a+h]$, ϕ est deux fois dérivable sur $[0, 1]$. En utilisant la règle de composition, on calcule ses dérivées successives :

- $\phi'(t) = Df(a+th)(h).$
- $\phi''(t) = D^2f(a+th)(h, h).$

Appliquons la formule de Taylor avec reste intégral à l'ordre 1 pour ϕ entre 0 et 1 :

$$\phi(1) = \phi(0) + \phi'(0) + \int_0^1 (1-t)\phi''(t) dt.$$

En revenant à f , cela s'écrit :

$$f(a+h) = f(a) + Df(a)(h) + \int_0^1 (1-t)D^2f(a+th)(h, h) dt.$$

Nous voulons comparer cette expression au développement polynomial attendu. Notons que le terme quadratique peut s'écrire comme une intégrale constante (car $\int_0^1 (1-t)dt = 1/2$) :

$$\frac{1}{2}D^2f(a)(h, h) = \int_0^1 (1-t)D^2f(a)(h, h) dt.$$

Formons la différence (le reste) :

$$R(h) = f(a+h) - f(a) - Df(a)(h) - \frac{1}{2}D^2f(a)(h, h).$$

En soustrayant les deux expressions intégrales, on obtient :

$$R(h) = \int_0^1 (1-t) [D^2f(a+th)(h, h) - D^2f(a)(h, h)] dt.$$

On majore la norme de ce reste :

$$\begin{aligned} \|R(h)\| &\leq \int_0^1 (1-t) \|D^2f(a+th) - D^2f(a)\| \|h\|^2 dt. \\ \|R(h)\| &\leq \|h\|^2 \left(\int_0^1 (1-t) dt \right) \sup_{t \in [0,1]} \|D^2f(a+th) - D^2f(a)\|. \end{aligned}$$

Comme f est \mathcal{C}^2 , l'application $x \mapsto D^2f(x)$ est continue en a . Ainsi, lorsque $h \rightarrow 0$, la quantité $\sup_{t \in [0,1]} \|D^2f(a+th) - D^2f(a)\|$ tend vers 0. On a donc bien $\|R(h)\| = o(\|h\|^2)$. \square

Cette formule est cruciale pour l'étude des extremums locaux (voir conditions du second ordre).

💡 Pour aller plus loin : dérivées généralisées

La formule de Taylor-Young nous dit que si une fonction est régulière (de classe \mathcal{C}^2), elle ressemble localement à un polynôme du second degré. Cependant, dans de nombreux problèmes modernes (équations aux dérivées partielles non linéaires, contrôle optimal, traitement d'image), les solutions que l'on cherche ne sont pas nécessairement différentiables partout.

Comment donner un sens à des équations faisant intervenir le gradient Du ou la hessienne D^2u si u n'est pas dérivable ? On verra l'année prochaine comment définir via les **distributions** une dérivée de fonctions non dérivables consistante avec la théorie classique. Une autre approche, la théorie des **solutions de viscosité** (due à M. Crandall et P.-L. Lions) répond à cette question pour certaines équations non linéaires en remplaçant la dérivée ponctuelle par des ensembles de "pentes" et de "courbures" admissibles, basés sur des inégalités de type Taylor.

On se place ici dans $E = \mathbb{R}^n$. Soit $u : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction continue (mais pas forcément dérivable).

Ordre 1 : Sous-différentiel et Sur-différentiel. Au lieu d'approcher u par un plan tangent (égalité), on cherche des plans qui restent en dessous ou au dessus du graphe de u au voisinage d'un point x .

Définition 3.4. — Le **sur-différentiel** de u en x , noté $D^+u(x)$, est l'ensemble des vecteurs $p \in \mathbb{R}^n$ tels que :

$$u(x+h) \leq u(x) + \langle p, h \rangle + o(\|h\|).$$

— Le **sous-différentiel** de u en x , noté $D^-u(x)$, est l'ensemble des vecteurs $p \in \mathbb{R}^n$ tels que :

$$u(x+h) \geq u(x) + \langle p, h \rangle + o(\|h\|).$$

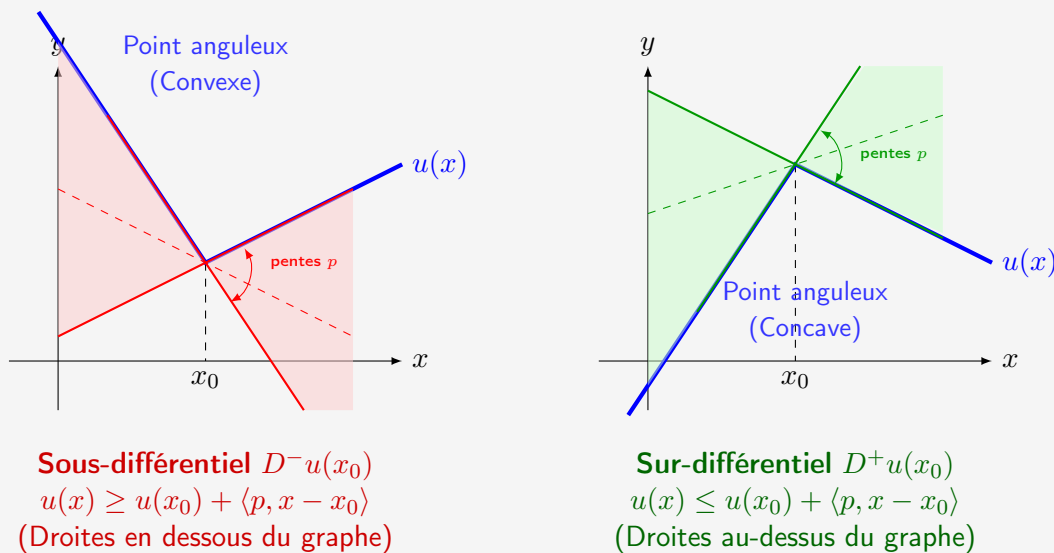


Figure 3.1 – Illustration des semi-différentiels en un point de non-différentiabilité x_0 . À gauche (D^-), la fonction fait un coin convexe ; toute droite passant par x_0 et restant dans la zone rouge (sous le graphe) définit un sous-gradient. À droite (D^+), la fonction fait un coin concave ; toute droite passant par x_0 et restant dans la zone verte (sur le graphe) définit un sur-gradient.

Si u est différentiable en x , alors $D^+u(x) = D^-u(x) = \{\nabla u(x)\}$. Si $u(x) = |x|$ en $x = 0$, on a $D^-u(0) = \bar{B}(0, 1)$ (la boule unité) et $D^+u(0) = \emptyset$. Une caractérisation plus graphique (et équivalente) du sur-différentiel consiste à écrire que $p \in D^+u(x)$ (resp. $p \in D^-u(x)$) ssi il existe une fonction

régulière $\phi \geq u$ (resp. $\phi \leq u$) telle que $\phi(x) = u(x)$ et $\nabla\phi(x) = p$ (faire un dessin si je ne l'ai pas fait moi-même !)

Ordre 2 : Les Jets. Pour traiter les équations du second ordre, on généralise cette idée en incluant un terme quadratique. On note \mathcal{S}_n l'ensemble des matrices symétriques réelles de taille $n \times n$.

Définition 3.5. — Le **sur-jet** d'ordre 2 de u en x , noté $J^{2,+}u(x)$, est l'ensemble des couples $(p, X) \in \mathbb{R}^n \times \mathcal{S}_n$ tels que :

$$u(x+h) \leq u(x) + \langle p, h \rangle + \frac{1}{2} \langle Xh, h \rangle + o(\|h\|^2).$$

— Le **sous-jet** d'ordre 2 de u en x , noté $J^{2,-}u(x)$, est l'ensemble des couples $(p, X) \in \mathbb{R}^n \times \mathcal{S}_n$ tels que :

$$u(x+h) \geq u(x) + \langle p, h \rangle + \frac{1}{2} \langle Xh, h \rangle + o(\|h\|^2).$$

Intuitivement, un élément $(p, X) \in J^{2,+}u(x)$ signifie qu'on peut toucher le graphe de u par le dessus au point x avec une parabole de gradient p et de hessienne X . On a la même caractérisation plus géométrique avec ces fonctions "test" : $(p, X) \in J^{2,+}u(x)$ (resp. $p \in J^{2,-}u(x)$) ssi il existe une fonction régulière $\phi \geq u$ (resp. $\phi \leq u$) telle que $\phi(x) = u(x)$, $\nabla\phi(x) = p$ et $\nabla^2\phi(x) = X$.

La cohérence avec l'analyse classique est assurée par le résultat suivant :

Proposition 3.4. Si u est de classe \mathcal{C}^2 au voisinage de x , alors les jets sont réduits à un singleton donné par la formule de Taylor-Young :

$$J^{2,+}u(x) = J^{2,-}u(x) = \{(\nabla u(x), \nabla^2 u(x))\}.$$

Ces objets permettent de définir ce qu'est une solution u d'une équation $F(x, u, \nabla u, \nabla^2 u) = 0$ au sens de viscosité : on teste l'équation sur les éléments des jets (qui sont des vecteurs et des matrices, ou des gradients et hessiennes des fonctions régulières ϕ) aux endroits où ils existent.

Références : le polycopié d'Alberto Bressan [14], et [13] (mais pour les longues soirées d'hiver).

3.4 Théorème des Fonctions Implicites

Théorème 3.6. Soient E, F, G des espaces de Banach, $f : E \times F \rightarrow G$ de classe \mathcal{C}^1 , et (x_0, y_0) tel que $f(x_0, y_0) = 0$. Si $\partial_y f(x_0, y_0)$ est un isomorphisme de F sur G , alors il existe des voisinages U de x_0 , V de y_0 et une unique application $\varphi : U \rightarrow V$ telle que $f(x, \varphi(x)) = 0$. De plus φ est \mathcal{C}^1 sur U .

Démonstration. Sans perte de généralité, supposons $(x_0, y_0) = (0, 0)$ et notons $A = \partial_y f(0, 0)$. A est inversible. L'équation $f(x, y) = 0$ est équivalente à $Ay - (Ay - f(x, y)) = 0$, soit :

$$y = A^{-1}(Ay - f(x, y)).$$

Pour x fixé (petit), définissons l'application $\Phi_x : F \rightarrow F$ par $\Phi_x(y) = y - A^{-1}f(x, y)$. Chercher y tel que $f(x, y) = 0$ revient à chercher un point fixe de Φ_x . Calculons sa différentielle par rapport à y :

$$\partial_y \Phi_x(y) = \text{Id}_F - A^{-1} \partial_y f(x, y) = A^{-1}(A - \partial_y f(x, y)).$$

Comme f est \mathcal{C}^1 et $\partial_y f(0, 0) = A$, par continuité, il existe des boules fermées $\bar{B}_E(0, r)$ et $\bar{B}_F(0, \rho)$ telles que pour tout (x, y) dans le produit cartésien de ces boules :

$$\|\text{Id} - A^{-1} \partial_y f(x, y)\| \leq \frac{1}{2}.$$

Par l'IAF, cela implique que pour tout x fixé, $y \mapsto \Phi_x(y)$ est $\frac{1}{2}$ -contractante sur la boule $V = \bar{B}_F(0, \rho)$. Pour appliquer le théorème du point fixe (Chapitre 2), il faut que Φ_x envoie V dans V . Or $\Phi_x(0) = -A^{-1}f(x, 0)$. Comme $f(0, 0) = 0$, par continuité, pour x assez petit (disons dans U), $\|\Phi_x(0)\| \leq \frac{\rho}{2}$. Alors pour $y \in V$: $\|\Phi_x(y)\| \leq \|\Phi_x(y) - \Phi_x(0)\| + \|\Phi_x(0)\| \leq \frac{1}{2}\|y\| + \frac{\rho}{2} \leq \rho$. Φ_x est donc une contraction de V (Banach) dans lui-même. Elle admet un unique point fixe $y = \varphi(x)$. La régularité \mathcal{C}^1 de φ s'obtient ensuite en écrivant les différentielles. \square

3.4.1 Interprétation et utilité

- **Géométrie** : Ce théorème signifie que l'ensemble des solutions de $f(x, y) = 0$ ressemble localement au graphe d'une fonction $y = \varphi(x)$. C'est ce qui permet de définir la notion de sous-variété (courbes, surfaces régulières).
- **Ingénierie** : En mécanique ou en théorie du contrôle, on a souvent des équations d'état $f(u, x) = 0$ où u est une commande et x l'état. Le théorème garantit que pour de petites variations de la commande u , l'état x s'adapte de manière régulière.

3.5 Extrema et Optimisation

Terminons par l'application directe du calcul différentiel à la recherche d'optimums, sujet central en ingénierie. Soit $J : U \subset E \rightarrow \mathbb{R}$ une fonctionnelle à minimiser, différentiable.

Proposition 3.5 (Condition nécessaire). *Si J admet un extremum local en a , alors $DJ(a) = 0$.*

Démonstration. Supposons $DJ(a) \neq 0$. Il existe $h \in E$ tel que $DJ(a)(h) < 0$ (sinon, par linéarité, on prend $-h$). Par définition de la différentielle : $J(a + th) = J(a) + tDJ(a)(h) + o(t) = J(a) + t[DJ(a)(h) + \varepsilon(t)]$. Pour $t > 0$ assez petit, le crochet est strictement négatif, donc $J(a + th) < J(a)$. De même, pour $t < 0$ assez petit, $J(a + th) > J(a)$ (si on change le signe). Donc a n'est ni un minimum ni un maximum local. \square

Pour distinguer minimum et maximum, on regarde l'ordre 2.

Proposition 3.6 (Condition suffisante du second ordre). *Soit a un point critique de J , c'est à dire tel que $DJ(a) = 0$.*

- Si $D^2J(a)$ est **coercive** (ou définie positive en dimension finie), c'est-à-dire s'il existe $\alpha > 0$ tel que $D^2J(a)(h, h) \geq \alpha\|h\|^2$, alors a est un minimum local strict.
- Si $D^2J(a)$ est définie négative, a est un maximum local strict.

Démonstration. Utilisons la formule de Taylor-Young à l'ordre 2 :

$$J(a + h) = J(a) + \underbrace{DJ(a)(h)}_0 + \frac{1}{2}D^2J(a)(h, h) + o(\|h\|^2).$$

Comme la forme quadratique est coercive :

$$J(a + h) - J(a) \geq \frac{\alpha}{2}\|h\|^2 - |o(\|h\|^2)| = \|h\|^2 \left(\frac{\alpha}{2} - \varepsilon(h) \right).$$

Pour $\|h\|$ assez petit, le terme entre parenthèses reste positif (disons $\geq \alpha/4$), donc $J(a+h) > J(a)$ pour $h \neq 0$. \square

💡 Pour aller plus loin : La différentiation automatique et l'IA

L'une des applications les plus spectaculaires du calcul différentiel au XXI^e siècle réside dans l'apprentissage profond (Deep Learning). Entraîner un réseau de neurones consiste à minimiser une fonction de coût (ou *Loss function*) $J(\theta)$ dépendant d'un nombre gigantesque de paramètres θ (les poids du réseau, parfois plusieurs milliards).

Pour utiliser des algorithmes de descente de gradient, il faut calculer le gradient $\nabla J(\theta)$. Comment faire cela efficacement sur ordinateur ?

Pourquoi les méthodes classiques échouent.

- **Différences finies** : On pourrait approximer la dérivée partielle par $\frac{J(\theta + \varepsilon e_i) - J(\theta)}{\varepsilon}$. Cependant, si $\theta \in \mathbb{R}^N$, cela nécessite $N + 1$ évaluations de J . Avec $N \approx 10^9$, c'est impossible. De plus, cette méthode est numériquement instable.
- **Calcul symbolique** : C'est l'approche des logiciels comme Maple ou Mathematica. Elle manipule les formules mathématiques. Le problème est l'explosion des expressions : la dérivée d'une fonction composée complexe peut devenir exponentiellement longue à écrire et à stocker.

La solution : La Différentiation Automatique (AD). L'approche retenue par les librairies modernes (PyTorch, TensorFlow, JAX) est la différentiation automatique. Elle ne manipule ni des approximations (comme les différences finies), ni des formules (comme le calcul symbolique), mais le **programme** qui calcule la fonction.

L'idée clé repose sur le théorème de dérivation des fonctions composées (Chain Rule). Tout programme informatique calculant une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ peut être vu comme une composition d'opérations élémentaires (additions, multiplications, sinus, exp...) :

$$f = f_k \circ f_{k-1} \circ \dots \circ f_1.$$

On représente cela par un *graphe de calcul*. La différentiation automatique applique la règle de la chaîne sur ce graphe. Il existe deux modes :

1. **Le mode direct (Forward mode)** : On calcule la dérivée en même temps que l'on évalue la fonction, en propageant les variations de l'entrée vers la sortie.

$$\text{Coût} \propto \text{Dimension de l'entrée.}$$

Ce mode est efficace si l'on a peu d'entrées et beaucoup de sorties.

2. **Le mode inverse (Reverse mode)** : C'est celui utilisé en IA.

(a) *Passe avant* : On évalue la fonction $x \mapsto f_1(x) \mapsto \dots \mapsto f_k(\dots)$ et on stocke les valeurs intermédiaires.

(b) *Passe arrière* : On part de la fin et on propage le gradient vers l'arrière en utilisant la transposée de la jacobienne (règle de la chaîne à l'envers).

$$\text{Coût} \propto \text{Dimension de la sortie.}$$

La Rétropropagation (Backpropagation). Dans l'apprentissage supervisé, la fonction de coût J est scalaire (sortie de dimension 1), mais l'entrée θ est de dimension immense. C'est le cas idéal pour le **mode inverse**. Le coût de calcul du gradient complet $\nabla J(\theta)$ est alors indépendant de la dimension N des paramètres ! Il est typiquement de l'ordre de 2 à 4 fois le coût d'une simple évaluation de la fonction J .

L'algorithme de **Backpropagation**, popularisé dans les années 80 pour les réseaux de neurones, n'est rien d'autre que l'implémentation algorithmique efficace du théorème de dérivation des fonctions composées en mode inverse. C'est grâce à cette efficacité mathématique que l'IA a pu passer à l'échelle.

Détail technique : L'algorithme de Rétropropagation

Pour comprendre pourquoi le mode inverse est si efficace, formalisons le calcul d'un réseau de neurones (ou de tout programme différentiable). Considérons une suite de transformations :

$$x_0 \xrightarrow{f_1} x_1 \xrightarrow{f_2} x_2 \cdots \xrightarrow{f_{k-1}} x_{k-1} \xrightarrow{f_k} x_k = J.$$

Ici, x_0 représente les paramètres (poids) du modèle, les f_i sont les couches successives (linéaires, ReLU, sigmoïde...), et $x_k = J \in \mathbb{R}$ est la fonction de coût scalaire à minimiser.

Notre objectif est de calculer le gradient de J par rapport à l'entrée x_0 , c'est-à-dire le vecteur $\nabla_{x_0} J$.

1. La Règle de la Chaîne (Point de vue matriciel). D'après la règle de composition, la différentielle de J par rapport à x_0 est le produit des différentielles intermédiaires :

$$DJ(x_0) = Df_k(x_{k-1}) \circ Df_{k-1}(x_{k-2}) \circ \cdots \circ Df_1(x_0).$$

En termes de matrices jacobiniennes, le gradient (qui est la transposée de la différentielle dans un espace euclidien) s'écrit en renversant l'ordre :

$$\nabla_{x_0} J = (J_{f_1}(x_0))^T \cdot (J_{f_2}(x_1))^T \cdots (J_{f_k}(x_{k-1}))^T \cdot 1.$$

2. L'Algorithme. L'algorithme de rétropropagation calcule ce produit de la droite vers la gauche (de la fin vers le début). On introduit la variable adjointe $\delta_i = \nabla_{x_i} J$, qui représente la sensibilité du coût final par rapport à la variable intermédiaire x_i .

1. **Passé Avant (Forward Pass) :** On calcule successivement les états x_i en partant de x_0 .

$$\text{Pour } i = 1 \dots k, \quad x_i = f_i(x_{i-1}).$$

Important : On doit mémoriser les valeurs des x_{i-1} car la dérivée Df_i en dépend souvent (exemple : la dérivée de $x \mapsto x^2$ dépend de x).

2. **Passé Arrière (Backward Pass) :** On initialise $\delta_k = 1$ (car $\nabla_{x_k} x_k = 1$). Ensuite, on remonte la chaîne. La relation de récurrence liant l'erreur à l'étape i à celle de l'étape $i - 1$ est donnée par :

$$\delta_{i-1} = (Df_i(x_{i-1}))^T \cdot \delta_i.$$

C'est un produit **Matrice-Vecteur**.

3. Complexité et Produit Jacobienne-Vecteur (VJP). Pourquoi est-ce efficace ?

- Si l'on faisait le produit de gauche à droite (mode direct), on multiplierait des matrices Jacobiennes matrice par matrice. Si $x_i \in \mathbb{R}^{1000}$, on manipulerait des matrices 1000×1000 .
- En mode arrière, on multiplie une matrice par un vecteur δ_i . Le résultat est un vecteur. On ne stocke jamais les énormes matrices Jacobiennes explicitement ! On implémente directement l'opération " $v \mapsto (Df)^T \cdot v$ " (appelée *Vector-Jacobian Product* ou VJP).

C'est cette astuce d'algèbre linéaire (l'associativité du produit matriciel permettant de prioriser les produits matrice-vecteur) qui rend l'entraînement des réseaux profonds réalisable.



Pour aller plus loin : Calcul des variations et équations d'Euler-Lagrange

Le calcul différentiel dans les espaces de dimension infinie (espaces de fonctions) est l'outil central du **Calcul des Variations**. Le problème type est de trouver une fonction u qui minimise une certaine énergie $J(u)$ définie par une intégrale.

Le cadre général

Soit E un espace de fonctions dérivables sur un intervalle $[a, b]$ (par exemple $E = \mathcal{C}^1([a, b], \mathbb{R})$) avec des conditions aux bords fixées $u(a) = \alpha, u(b) = \beta$. On considère la fonctionnelle :

$$J(u) = \int_a^b L(t, u(t), u'(t)) dt$$

où $L : \mathbb{R} \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ est une fonction régulière appelée le **Lagrangien**.

Si u est un minimum local de J , alors sa différentielle doit être nulle : $DJ(u)(h) = 0$ pour toute perturbation admissible h (c'est-à-dire telle que $h(a) = h(b) = 0$). Calculons cette dérivée directionnelle (différentielle de Gâteaux) :

$$DJ(u)(h) = \lim_{\varepsilon \rightarrow 0} \frac{J(u + \varepsilon h) - J(u)}{\varepsilon} = \int_a^b \left[\frac{\partial L}{\partial u} h(t) + \frac{\partial L}{\partial u'} h'(t) \right] dt.$$

Une intégration par parties sur le second terme (en utilisant $h(a) = h(b) = 0$) donne :

$$\int_a^b \frac{\partial L}{\partial u'} h'(t) dt = \left[\frac{\partial L}{\partial u'} h(t) \right]_a^b - \int_a^b \frac{d}{dt} \left(\frac{\partial L}{\partial u'} \right) h(t) dt = - \int_a^b \frac{d}{dt} \left(\frac{\partial L}{\partial u'} \right) h(t) dt.$$

La condition d'optimalité $DJ(u)(h) = 0$ pour tout h implique alors (par le lemme fondamental du calcul des variations) que u doit vérifier l'**Équation d'Euler-Lagrange** :

$$\frac{\partial L}{\partial u} - \frac{d}{dt} \left(\frac{\partial L}{\partial u'} \right) = 0.$$

Application 1 : Mécanique et Principe de Moindre Action

En mécanique classique, l'évolution d'un système au cours du temps n'est pas arbitraire. Le **Principe de Moindre Action** (Maupertuis, Hamilton) stipule que la trajectoire réelle $q(t)$ suivie par un

système entre deux instants minimise une quantité appelée Action :

$$S(q) = \int_{t_1}^{t_2} (T - V) dt$$

où $L = T - V$ est la différence entre l'énergie cinétique $T = \frac{1}{2}m(\dot{q})^2$ et l'énergie potentielle $V(q)$. Appliquons l'équation d'Euler-Lagrange avec $u = q$ et $u' = \dot{q}$:

$$- \frac{\partial L}{\partial q} = - \frac{\partial V}{\partial q} = F(q) \text{ (la force dérivant du potentiel).}$$

$$- \frac{\partial L}{\partial \dot{q}} = \frac{\partial}{\partial \dot{q}} \left(\frac{1}{2} m \dot{q}^2 \right) = m \dot{q}.$$

L'équation devient $F(q) - \frac{d}{dt}(m\dot{q}) = 0$, soit $F = ma$. On retrouve ainsi la loi fondamentale de la dynamique de Newton comme conséquence d'un principe d'optimisation fonctionnelle.

Application 2 : Analyse d'images et Régularisation

En traitement d'image, on cherche souvent à reconstruire une image propre $u(x, y)$ à partir d'une observation bruitée $u_0(x, y)$. Le problème est mal posé. Pour le résoudre, on minimise une énergie composée d'un terme de fidélité et d'un terme de régularité (lissage). Soit $\Omega \subset \mathbb{R}^2$ le domaine de l'image. On considère l'énergie (modèle de Tikhonov) :

$$J(u) = \frac{1}{2} \int_{\Omega} (u - u_0)^2 dx + \frac{\lambda}{2} \int_{\Omega} \|\nabla u\|^2 dx.$$

Le terme $\|\nabla u\|^2$ pénalise les variations trop brusques (le bruit). Pour trouver l'équation d'Euler-Lagrange associée, on calcule la variation de J par rapport à u dans la direction h (une fonction test nulle au bord) :

$$DJ(u)(h) = \int_{\Omega} (u - u_0)h dx + \lambda \int_{\Omega} \nabla u \cdot \nabla h dx.$$

La formule de Green (l'équivalent 2D de l'intégration par parties) permet de transformer le second terme : $\int_{\Omega} \nabla u \cdot \nabla h = - \int_{\Omega} (\Delta u)h$. Ainsi :

$$DJ(u)(h) = \int_{\Omega} (u - u_0 - \lambda \Delta u)h dx.$$

L'image optimale u est donc solution de l'équation aux dérivées partielles (EDP) elliptique :

$$u - \lambda \Delta u = u_0.$$

C'est une version stationnaire de l'équation de la chaleur : pour débruiter l'image u_0 , on lui applique une diffusion contrôlée par λ . Des modèles plus avancés (comme la Variation Totale de Rudin-Osher-Fatemi) remplacent $\|\nabla u\|^2$ par $\|\nabla u\|$ pour préserver les contours nets (les "sauts" dans l'image).

Pour aller plus loin : La méthode de l'état adjoint

En ingénierie, on cherche souvent à optimiser un système régi par des équations physiques (EDP, système dynamique). Le problème se pose ainsi : on veut minimiser une fonction coût J dépendant de paramètres de contrôle u et de l'état du système y . L'état y n'est pas libre, il est déterminé par u via une *équation d'état* $E(u, y) = 0$.

Le problème abstrait

Soient U, Y, Z trois espaces de Banach (respectivement espace des contrôles, des états, et des contraintes).

- L'équation d'état est $E(u, y) = 0$, où $E : U \times Y \rightarrow Z$ est de classe \mathcal{C}^1 .
- La fonction coût est $J(u, y) \in \mathbb{R}$, où $J : U \times Y \rightarrow \mathbb{R}$ est différentiable.

On suppose que pour tout u , l'équation $E(u, y) = 0$ admet une unique solution $y(u)$ (hypothèse du Théorème des Fonctions Implicites, en supposant $\partial_y E$ inversible). On cherche à calculer le gradient de la fonctionnelle réduite $\hat{J}(u) = J(u, y(u))$ pour appliquer un algorithme de descente.

La méthode directe (coûteuse)

Par la règle de composition (Chain Rule), la différentielle de \hat{J} dans une direction $h \in U$ est :

$$D\hat{J}(u) \cdot h = \partial_u J(u, y) \cdot h + \partial_y J(u, y) \cdot (Dy(u) \cdot h).$$

Le terme difficile est $Dy(u) \cdot h$, qui représente la sensibilité de l'état par rapport au contrôle. Pour l'obtenir, on différencie l'équation d'état $E(u, y(u)) = 0$ par rapport à u dans la direction h :

$$\partial_u E(u, y) \cdot h + \partial_y E(u, y) \cdot (Dy(u) \cdot h) = 0.$$

On en déduit $Dy(u) \cdot h = -[\partial_y E(u, y)]^{-1} \circ [\partial_u E(u, y) \cdot h]$. En injectant cela dans l'expression du gradient :

$$D\hat{J}(u) \cdot h = \partial_u J(u, y) \cdot h - \underbrace{\partial_y J(u, y) \circ [\partial_y E(u, y)]^{-1}}_{\text{Terme complexe}} \circ [\partial_u E(u, y) \cdot h].$$

Si la dimension de l'espace de contrôle U est grande (ex : 10^6 paramètres), il faudrait résoudre l'équation linéarisée pour **chaque** direction de base h afin de construire le gradient complet. C'est inenvisageable.

L'astuce de l'État Adjoint

L'idée est de changer l'ordre des calculs en introduisant un vecteur intermédiaire $p \in Z'$ (le dual de Z), appelé **état adjoint**. On définit p comme la solution de l'équation suivante (équation adjointe) :

$$[\partial_y E(u, y)]^* p = -[\partial_y J(u, y)]^T$$

où A^* désigne l'opérateur adjoint de A . Une fois p calculé (ce qui ne nécessite de résoudre qu'un seul système linéaire, quelle que soit la dimension de U), on peut réécrire le gradient très simplement. En effet, en notant $\langle \cdot, \cdot \rangle$ le crochet de dualité :

$$\begin{aligned} \partial_y J \cdot (Dy \cdot h) &= \langle (\partial_y J)^T, Dy \cdot h \rangle \\ &= \langle -(\partial_y E)^* p, Dy \cdot h \rangle \quad (\text{par définition de } p) \\ &= \langle p, -\partial_y E \cdot (Dy \cdot h) \rangle \quad (\text{définition de l'adjoint}) \\ &= \langle p, \partial_u E \cdot h \rangle \quad (\text{par l'équation d'état différenciée}). \end{aligned}$$

Finalement, le gradient s'exprime sans jamais calculer $Dy(u)$:

$$\nabla \hat{J}(u) = \partial_u J(u, y)^T + [\partial_u E(u, y)]^* p.$$

Exemple : Contrôle d'un système linéaire

Soit à inverser le système $Ay = u$ (où A est une matrice inversible) pour minimiser $J(u, y) = \frac{1}{2}\|y - y_{cible}\|^2 + \frac{\alpha}{2}\|u\|^2$.

1. **État direct** : $E(u, y) = Ay - u = 0$. Donc $\partial_y E = A$ et $\partial_u E = -Id$.
2. **Gradient coût** : $\partial_y J = (y - y_{cible})^T$ et $\partial_u J = \alpha u^T$.
3. **État adjoint** : L'équation $A^*p = -(\partial_y J)^T$ devient $A^T p = -(y - y_{cible})$. On résout ce système linéaire pour trouver p .
4. **Gradient final** : $\nabla \hat{J}(u) = \alpha u + (-Id)^* p = \alpha u - p$.

Le coût de calcul du gradient est donc équivalent à deux résolutions de système linéaire (une pour y , une pour p), indépendamment du nombre de paramètres u .

Chapitre 4

Espaces de Hilbert

Les espaces de Banach, vus aux chapitres précédents, offrent un cadre général pour l'analyse (convergence, complétude). Cependant, il leur manque une structure essentielle pour faire de la géométrie : la notion d'**angle** et donc d'orthogonalité. Dans \mathbb{R}^2 ou \mathbb{R}^3 , l'angle et l'orthogonalité sont définis grâce au **produit scalaire**.

L'objectif de ce chapitre est d'étendre cette géométrie euclidienne aux espaces de dimension infinie. Nous allons définir les espaces de Hilbert, qui sont à la fois des espaces complets (pour l'analyse) et munis d'un produit scalaire (pour la géométrie). C'est le cadre naturel de la mécanique quantique, du traitement du signal (Séries de Fourier) et des méthodes d'éléments finis.

Dans tout le chapitre, \mathbb{K} désigne le corps des réels \mathbb{R} ou des complexes \mathbb{C} .

4.1 Espaces préhilbertiens et de Hilbert

4.1.1 Produit scalaire

Définition 4.1 (Produit scalaire). Soit H un espace vectoriel sur \mathbb{K} . On appelle **produit scalaire** une application $\varphi : H \times H \rightarrow \mathbb{K}$, notée généralement $\langle x, y \rangle$, vérifiant les propriétés suivantes :

1. **Linéarité à gauche** : Pour tout $y \in H$, l'application $x \mapsto \langle x, y \rangle$ est linéaire.

$$\forall x_1, x_2 \in H, \forall \lambda \in \mathbb{K}, \quad \langle \lambda x_1 + x_2, y \rangle = \lambda \langle x_1, y \rangle + \langle x_2, y \rangle.$$

2. **Symétrie hermitienne** : Pour tout $x, y \in H$,

$$\langle y, x \rangle = \overline{\langle x, y \rangle}.$$

(Si $\mathbb{K} = \mathbb{R}$, c'est la symétrie usuelle).

3. **Positivité** : Pour tout $x \in H$, $\langle x, x \rangle$ est un réel positif ou nul.

4. **Caractère défini** : $\langle x, x \rangle = 0 \iff x = 0_H$.

Un espace vectoriel muni d'un produit scalaire est appelé **espace préhilbertien**.

Remarque 4.1 (Convention). La symétrie hermitienne implique que le produit scalaire est **anti-linéaire** par rapport à la seconde variable (si $\mathbb{K} = \mathbb{C}$) :

$$\langle x, \lambda y \rangle = \overline{\lambda} \langle x, y \rangle.$$

4.1.2 L'inégalité de Cauchy-Schwarz

C'est sans doute l'inégalité la plus importante de l'analyse fonctionnelle.

Théorème 4.1 (Inégalité de Cauchy-Schwarz). *Soit $(H, \langle \cdot, \cdot \rangle)$ un espace préhilbertien. Pour tout $x, y \in H$, on a :*

$$|\langle x, y \rangle| \leq \sqrt{\langle x, x \rangle} \sqrt{\langle y, y \rangle}.$$

De plus, il y a égalité si et seulement si la famille (x, y) est liée (les vecteurs sont colinéaires).

Démonstration. Traitons le cas réel ($\mathbb{K} = \mathbb{R}$). Pour tout $\lambda \in \mathbb{R}$, par positivité du produit scalaire :

$$P(\lambda) = \langle x + \lambda y, x + \lambda y \rangle \geq 0.$$

En développant par bilinéarité et symétrie :

$$P(\lambda) = \|y\|^2 \lambda^2 + 2\langle x, y \rangle \lambda + \|x\|^2 \geq 0.$$

C'est un polynôme du second degré en λ qui est toujours positif. Son discriminant réduit doit donc être négatif ou nul :

$$\Delta' = (\langle x, y \rangle)^2 - \|x\|^2 \|y\|^2 \leq 0,$$

ce qui donne l'inégalité cherchée. Le cas d'égalité correspond au cas où le discriminant est nul, c'est-à-dire quand le polynôme admet une racine λ_0 , donc $x + \lambda_0 y = 0$. \square

Cette inégalité permet de définir une norme naturelle.

Proposition 4.1 (Norme induite). *L'application $x \mapsto \|x\| = \sqrt{\langle x, x \rangle}$ est une **norme** sur H .*

Démonstration. La seule propriété non triviale est l'inégalité triangulaire $\|x + y\| \leq \|x\| + \|y\|$. Elle découle directement de Cauchy-Schwarz :

$$\|x + y\|^2 = \|x\|^2 + 2\operatorname{Re}(\langle x, y \rangle) + \|y\|^2 \leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.$$

\square

4.1.3 Identités géométriques

Dans un espace préhilbertien, la norme provient d'un produit scalaire. Elle vérifie donc des identités spécifiques que ne vérifient pas les normes quelconques (comme la norme L^1 ou L^∞).

Proposition 4.2 (Identité du parallélogramme). *Pour tout $x, y \in H$:*

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

Cette identité caractérise les normes hilbertiennes. Si une norme ne la vérifie pas (par exemple $\|(1, 0) + (0, 1)\|_1^2 + \|(1, 0) - (0, 1)\|_1^2 = 2^2 + 2^2 = 8 \neq 2(1^2 + 1^2)$), elle ne peut pas provenir d'un produit scalaire.

Proposition 4.3 (Identités de polarisation). *Elles permettent de retrouver le produit scalaire à partir de la norme.*

- Cas réel : $\langle x, y \rangle = \frac{1}{4} (\|x + y\|^2 - \|x - y\|^2)$.
- Cas complexe : $\langle x, y \rangle = \frac{1}{4} \sum_{k=0}^3 i^k \|x + i^k y\|^2$.

4.1.4 Espaces de Hilbert

Définition 4.2 (Espace de Hilbert). Un **espace de Hilbert** est un espace préhilbertien $(H, \langle \cdot, \cdot \rangle)$ qui est **complet** pour la norme induite.

Autrement dit, c'est un espace de Banach dont la norme découle d'un produit scalaire.

4.1.5 Exemples fondamentaux

Exemple 4.1 (Dimension finie). L'espace \mathbb{K}^n muni du produit scalaire canonique :

$$\langle x, y \rangle = \sum_{k=1}^n x_k \overline{y_k}$$

est un espace de Hilbert (euclidien si \mathbb{R} , hermitien si \mathbb{C}).

Exemple 4.2 (Suites de carré sommable). L'espace $\ell^2(\mathbb{N})$ des suites $u = (u_n)$ telles que $\sum |u_n|^2 < +\infty$ est un espace de Hilbert pour le produit scalaire :

$$\langle u, v \rangle = \sum_{n=0}^{\infty} u_n \overline{v_n}.$$

C'est le prototype des espaces de Hilbert séparables de dimension infinie.

Exemple 4.3 (Fonctions de carré intégrable). Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré. L'espace de Lebesgue $L^2(\Omega)$, quotienté par la relation d'égalité presque partout, est un espace de Hilbert pour le produit scalaire :

$$\langle f, g \rangle = \int_{\Omega} f(t) \overline{g(t)} d\mu(t).$$

La complétude de cet espace est assurée par le théorème de Riesz-Fischer (Chapitre 2).

4.2 Orthogonalité et Théorème de Projection

L'intérêt majeur des espaces de Hilbert est de permettre de faire de la géométrie euclidienne (angles droits, projections) en dimension infinie.

4.2.1 Orthogonalité et Pythagore

Définition 4.3 (Orthogonalité). Deux vecteurs x et y de H sont dits **orthogonaux**, noté $x \perp y$, si leur produit scalaire est nul :

$$\langle x, y \rangle = 0.$$

Deux sous-ensembles A et B sont dits orthogonaux si $\forall a \in A, \forall b \in B, a \perp b$.

Proposition 4.4 (Théorème de Pythagore). Si $x \perp y$, alors :

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

Plus généralement, si (x_1, \dots, x_n) est une famille orthogonale (i.e. $\langle x_i, x_j \rangle = 0$ pour $i \neq j$), alors :

$$\left\| \sum_{i=1}^n x_i \right\|^2 = \sum_{i=1}^n \|x_i\|^2.$$

4.2.2 Projection sur un convexe fermé

C'est le résultat fondamental de ce chapitre. Il affirme qu'on peut toujours trouver l'élément d'un ensemble convexe fermé le plus proche d'un vecteur donné de l'espace. Cela a des applications importantes en théorie de l'approximation.

Soit K une partie de H . On rappelle que la distance d'un point $x \in H$ à l'ensemble K est définie par $d(x, K) = \inf_{y \in K} \|x - y\|$.

Théorème 4.2 (Projection sur un convexe fermé). *Soit H un espace de Hilbert et K une partie de H qui est :*

1. **Convexe** (si $u, v \in K$, alors le segment $[u, v] = \{tu + (1 - t)v, t \in [0, 1]\}$ est inclus dans K) ;
2. **Fermée** ;
3. **Non vide**.

Alors, pour tout $x \in H$, il existe un **unique** point $y^* \in K$ qui réalise la distance minimale, c'est-à-dire :

$$\|x - y^*\| = d(x, K) = \min_{y \in K} \|x - y\|.$$

Ce point y^* est appelé la **projection** de x sur K et noté $P_K(x)$.

Démonstration (Idée). **Existence :** On considère une suite minimisante (y_n) de K telle que $\|x - y_n\| \rightarrow d(x, K)$. En utilisant l'identité du parallélogramme, on montre que cette suite est de Cauchy. L'espace H étant complet et K étant fermé, la suite converge vers un élément $y^* \in K$ qui réalise le minimum. **Unicité :** On utilise la stricte convexité de la boule (aussi via l'identité du parallélogramme) : si deux points réalisaient le minimum, leur milieu serait strictement plus proche, ce qui est absurde. \square

Comment caractériser ce point y^* sans avoir à calculer toutes les distances ? Géométriquement, l'angle entre le vecteur erreur $x - P_K(x)$ et n'importe quel vecteur partant vers l'intérieur du convexe doit être obtus (voir figure 4.1).

Théorème 4.3 (Caractérisation variationnelle). *Soit K un convexe fermé non vide et $x \in H$. Un point $y \in K$ est la projection $P_K(x)$ si et seulement si :*

$$\forall z \in K, \quad \operatorname{Re}\langle x - y, z - y \rangle \leq 0.$$

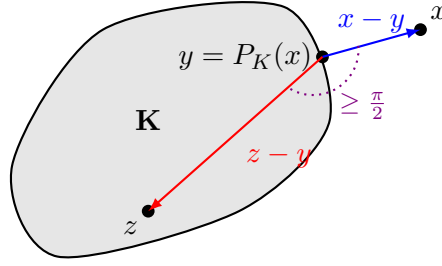
4.2.3 Projection sur un sous-espace vectoriel fermé

Un cas particulier très fréquent est celui où l'ensemble K est un sous-espace vectoriel F . Un sous-espace est toujours convexe. S'il est fermé (toujours vrai s'il est de dimension finie, même si l'espace ambiant ne l'est pas), le théorème précédent s'applique, mais l'inégalité devient une égalité (orthogonalité), voir figure 4.2.

Proposition 4.5 (Caractérisation par l'orthogonalité). *Soit F un sous-espace vectoriel fermé de H et $x \in H$. La projection $y = P_F(x)$ est l'unique élément de F tel que son écart à x soit **orthogonal** à F :*

$$x - y \perp F \quad \Longleftrightarrow \quad \forall z \in F, \langle x - y, z \rangle = 0.$$

Remarque 4.2. L'application $P_F : H \rightarrow H$ est alors une application **linéaire**, continue, de norme 1 (si $F \neq \{0\}$). On l'appelle projecteur orthogonal sur F .



Caractérisation : $\forall z \in K, \langle x - y, z - y \rangle \leq 0$
(L'angle entre $x - y$ et $z - y$ est obtus)

Figure 4.1 – Illustration de la projection d'un point x sur un convexe fermé K . Le point $y = P_K(x)$ est le point de K le plus proche de x . Il est caractérisé par le fait que le vecteur $x - y$ forme un angle obtus avec tout vecteur $z - y$ partant de y vers l'intérieur du convexe.

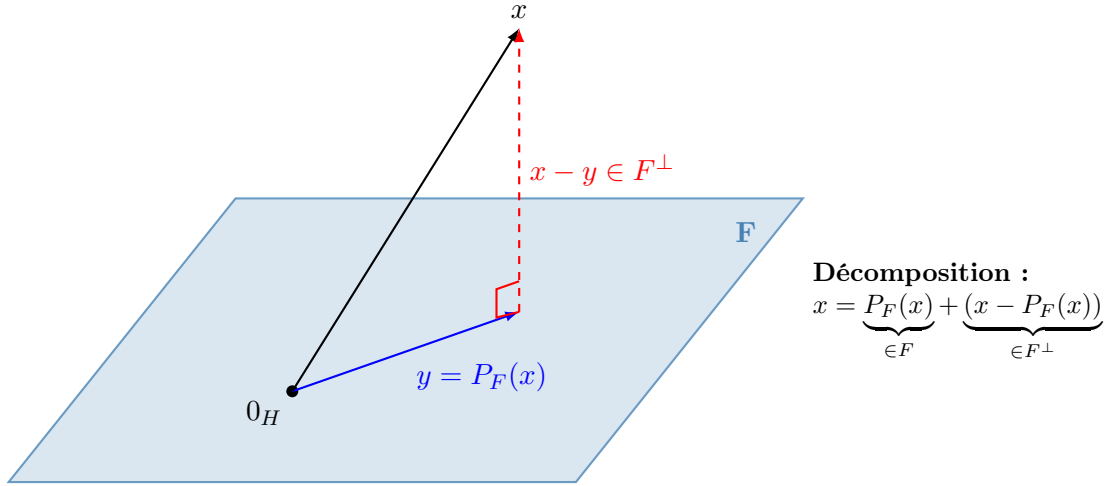


Figure 4.2 – Projection orthogonale sur un sous-espace fermé F . Le vecteur erreur $x - P_F(x)$ est orthogonal à tout vecteur du plan F .

4.2.4 Supplémentaire orthogonal

Définition 4.4. Soit F un sous-ensemble de H . L'**orthogonal** de F , noté F^\perp , est l'ensemble des vecteurs orthogonaux à tous les éléments de F :

$$F^\perp = \{x \in H \mid \forall y \in F, \langle x, y \rangle = 0\}.$$

F^\perp est toujours un sous-espace vectoriel fermé (même si F ne l'est pas).

Théorème 4.4 (Décomposition orthogonale). Soit F un sous-espace vectoriel **fermé** de H . Alors :

$$H = F \oplus F^\perp.$$

Cela signifie que tout vecteur $x \in H$ se décompose de manière unique sous la forme $x = u + v$ avec $u \in F$ et $v \in F^\perp$. On a alors $u = P_F(x)$ et $v = P_{F^\perp}(x)$.

Exemple 4.4 (Séries de Fourier). Dans $L^2([0, 2\pi])$, si on note F_n le sous-espace engendré par les polynômes trigonométriques de degré n , la projection d'une fonction f sur F_n est simplement sa somme de Fourier tronquée $S_n(f)$. Le théorème de projection dit que $S_n(f)$ est la « meilleure approximation » de f par un polynôme trigonométrique au sens de l'énergie (moindres carrés).

4.2.5 Application : La méthode des Moindres Carrés

Un problème fréquent en ingénierie consiste à résoudre un système linéaire « surdéterminé » : on a beaucoup plus d'équations (de mesures) que d'inconnues (de paramètres). Soit $A \in \mathcal{M}_{m,n}(\mathbb{R})$ une matrice rectangulaire avec $m \gg n$, et $b \in \mathbb{R}^m$ un vecteur de mesures. Le système $Ax = b$ n'a généralement **aucune solution** exacte (le vecteur b n'est pas dans l'image de A).

L'approche des **Moindres Carrés** consiste à chercher le vecteur \hat{x} qui minimise l'erreur quadratique :

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|^2.$$

Interprétation Géométrique

Plaçons-nous dans l'espace de Hilbert $H = \mathbb{R}^m$ muni du produit scalaire canonique. L'ensemble des vecteurs de la forme Ax (quand x parcourt \mathbb{R}^n) est exactement l'image de la matrice A , notée $\text{Im}(A)$. C'est un sous-espace vectoriel de \mathbb{R}^m .

Le problème revient à trouver le point $p \in \text{Im}(A)$ qui est le plus proche de b . D'après le **Théorème de la projection** (vu en Section 4.2), ce point existe et est unique : c'est la projection orthogonale de b sur $\text{Im}(A)$.

$$A\hat{x} = P_{\text{Im}(A)}(b).$$

Les Équations Normales

Comment calculer \hat{x} en pratique ? La caractérisation de la projection nous dit que le vecteur erreur $b - A\hat{x}$ doit être **orthogonal** au sous-espace $\text{Im}(A)$.

$$\forall y \in \text{Im}(A), \quad \langle b - A\hat{x}, y \rangle = 0.$$

Tout vecteur y de l'image s'écrit $y = Az$ pour un certain $z \in \mathbb{R}^n$.

$$\forall z \in \mathbb{R}^n, \quad \langle b - A\hat{x}, Az \rangle = 0.$$

Utilisons la définition de l'adjoint (ici la transposée) :

$$\forall z \in \mathbb{R}^n, \quad \langle A^T(b - A\hat{x}), z \rangle_{\mathbb{R}^n} = 0.$$

Si un vecteur est orthogonal à tout l'espace, il est nul. On obtient donc l'**équation normale** :

$$A^T(b - A\hat{x}) = 0 \iff A^T A\hat{x} = A^T b.$$

Si la matrice A est de rang plein (ses colonnes sont indépendantes), alors la matrice carrée $A^T A$ est inversible (symétrique définie positive), et la solution est unique :

$$\hat{x} = (A^T A)^{-1} A^T b.$$

La matrice $A^+ = (A^T A)^{-1} A^T$ est appelée **Pseudo-inverse** de Moore-Penrose.

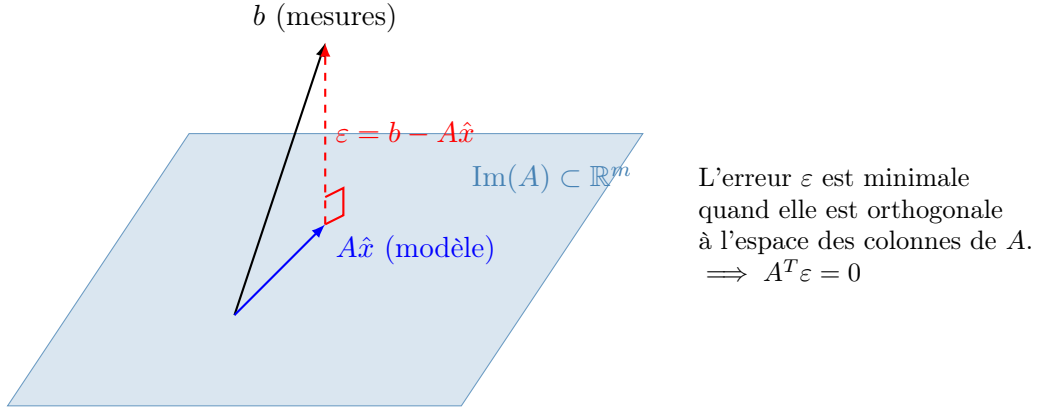


Figure 4.3 – Interprétation géométrique des moindres carrés : on projette le vecteur des mesures b sur l'espace engendré par les colonnes de la matrice A .

Exemple : La Régression Linéaire

Supposons que l'on dispose de m points de mesure (t_i, y_i) et que l'on cherche la droite $y = \alpha t + \beta$ qui passe « au mieux » par ces points. Le système s'écrit :

$$\begin{cases} \alpha t_1 + \beta = y_1 \\ \vdots \\ \alpha t_m + \beta = y_m \end{cases} \iff \underbrace{\begin{pmatrix} t_1 & 1 \\ \vdots & \vdots \\ t_m & 1 \end{pmatrix}}_A \underbrace{\begin{pmatrix} \alpha \\ \beta \end{pmatrix}}_x = \underbrace{\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}}_b.$$

En résolvant le système 2×2 donné par $A^T A \hat{x} = A^T b$, on trouve les coefficients optimaux $\hat{\alpha}$ et $\hat{\beta}$ de la régression linéaire.

4.3 Dualité et Théorème de représentation

Dans le chapitre 1, nous avons défini le dual topologique E' d'un espace de Banach comme l'espace des formes linéaires continues sur E . Dans le cas d'un espace de Hilbert, ce dual s'identifie parfaitement à l'espace lui-même grâce au produit scalaire.

4.3.1 Le Théorème de représentation de Riesz

Si on fixe un vecteur $y \in H$, l'application $L_y : x \mapsto \langle x, y \rangle$ est clairement linéaire. Elle est aussi continue car d'après Cauchy-Schwarz :

$$|L_y(x)| = |\langle x, y \rangle| \leq \|y\| \|x\|.$$

Donc $L_y \in H'$ et sa norme est $\|L_y\|_{H'} \leq \|y\|$. (En prenant $x = y$, on voit que la norme est exactement $\|y\|$).

La question inverse est plus profonde : est-ce que **toute** forme linéaire continue s'exprime comme un produit scalaire ?

Théorème 4.5 (Représentation de Riesz). Soit H un espace de Hilbert et $\phi : H \rightarrow \mathbb{K}$ une forme linéaire continue (i.e. $\phi \in H'$). Alors, il existe un **unique** vecteur $y \in H$ tel que :

$$\forall x \in H, \quad \phi(x) = \langle x, y \rangle.$$

De plus, l'application $\Phi : y \mapsto \langle \cdot, y \rangle$ est une isométrie de H sur H' . On a :

$$\|\phi\|_{H'} = \|y\|_H.$$

Démonstration. Unicité : Si $\langle x, y_1 \rangle = \langle x, y_2 \rangle$ pour tout x , alors $\langle x, y_1 - y_2 \rangle = 0$. En choisissant $x = y_1 - y_2$, on obtient $\|y_1 - y_2\|^2 = 0$, donc $y_1 = y_2$.

Existence : Si ϕ est la forme nulle, alors $y = 0$ convient. Supposons $\phi \neq 0$. Soit $N = \ker \phi$. Comme ϕ est continue, N est un sous-espace vectoriel **fermé**. Puisque $\phi \neq 0$, $N \neq H$, donc son orthogonal N^\perp n'est pas réduit à $\{0\}$. Il existe donc un vecteur $z \in N^\perp$ avec $z \neq 0$. On peut le normaliser pour avoir $\|z\| = 1$. Construisons le vecteur y . Pour tout $x \in H$, on peut décomposer x en utilisant la valeur de $\phi(x)$. Considérons le vecteur $u = \phi(x)z - \phi(z)x$. On vérifie que $u \in N$:

$$\phi(u) = \phi(x)\phi(z) - \phi(z)\phi(x) = 0.$$

Puisque $z \in N^\perp$ et $u \in N$, on a $\langle u, z \rangle = 0$.

$$\langle \phi(x)z - \phi(z)x, z \rangle = 0 \implies \phi(x)\langle z, z \rangle - \phi(z)\langle x, z \rangle = 0.$$

Comme $\langle z, z \rangle = \|z\|^2 = 1$, on obtient :

$$\phi(x) = \phi(z)\langle x, z \rangle = \langle x, \overline{\phi(z)}z \rangle.$$

Le vecteur cherché est donc $y = \overline{\phi(z)}z$. □

4.3.2 Identification du dual

Ce théorème permet d'**identifier** H et son dual H' . C'est une propriété très forte, appelée **réflexivité**.

- Dans L^2 , le dual est L^2 . (Alors que pour L^p avec $p \neq 2$, le dual est L^q avec $\frac{1}{p} + \frac{1}{q} = 1$.)
- Cela simplifie énormément les notations : on ne manipule plus de formes linéaires abstraites, mais simplement des vecteurs.

4.3.3 Application : L'opérateur adjoint

Soit $T \in \mathcal{L}(H)$ un opérateur linéaire continu de H dans H . Fixons $y \in H$. L'application $x \mapsto \langle Tx, y \rangle$ est une forme linéaire continue sur H . D'après le théorème de Riesz, il existe un unique vecteur $z \in H$ tel que cette forme s'écrive $\langle x, z \rangle$. Ce vecteur z dépend de y : on le note T^*y .

Définition 4.5 (Adjoint). Soit $T \in \mathcal{L}(H)$. On appelle **opérateur adjoint** de T , noté T^* , l'unique opérateur vérifiant :

$$\forall x, y \in H, \quad \langle Tx, y \rangle = \langle x, T^*y \rangle.$$

Proposition 4.6. 1. T^* est linéaire et continu, et $\|T^*\| = \|T\|$.

2. $(T^*)^* = T$.

3. $(ST)^* = T^*S^*$.

4. Si $H = \mathbb{C}^n$ et T est représenté par une matrice A , alors T^* est représenté par la matrice adjointe $A^* = \overline{A}^T$ (transposée conjuguée).

4.3.4 Convergence faible

La réflexivité des espaces de Hilbert permet de définir une notion de convergence moins exigeante que la convergence en norme.

Définition 4.6 (Convergence faible). Une suite (x_n) de H converge **faiblement** vers x (noté $x_n \rightharpoonup x$) si :

$$\forall y \in H, \quad \lim_{n \rightarrow \infty} \langle x_n, y \rangle = \langle x, y \rangle.$$

Exemple 4.5. Dans un espace de Hilbert de dimension infinie, une base orthonormée (e_n) converge faiblement vers 0. En effet, pour tout y , la suite des coefficients de Fourier $\langle e_n, y \rangle$ tend vers 0 (lemme de Riemann-Lebesgue, conséquence de Bessel, cf section ci-dessous). Pourtant, $\|e_n - 0\| = 1$, donc elle ne converge pas fortement. La convergence faible correspond à la convergence « coordonnée par coordonnée », tandis que la convergence forte correspond à la convergence de l'énergie totale.

Proposition 4.7 (Convergence faible vs forte). Une suite (x_n) de H converge **fortement** vers x si et seulement si elle converge **faiblement** vers x et $\|x_n\| \rightarrow \|x\|$.

Démonstration. En effet supposons que $x_n \rightarrow x$. Alors par Cauchy-Schwarz il est clair que $x_n \rightharpoonup x$. De plus $|\|x_n\| - \|x\|| \leq \|x_n - x\| \rightarrow 0$. Inversement supposons $x_n \rightharpoonup x$ et $\|x_n\| \rightarrow \|x\|$. Alors $\|x_n - x\|^2 = \|x_n\|^2 + \|x\|^2 - 2\operatorname{Re} \langle x_n, x \rangle \rightarrow \|x\|^2 + \|x\|^2 - 2\operatorname{Re} \langle x, x \rangle = 0$. \square

4.4 Bases Hilbertiennes et Séries de Fourier

Dans un espace vectoriel de dimension finie \mathbb{K}^n , le choix d'une base orthonormée (e_1, \dots, e_n) permet de ramener tous les calculs sur les vecteurs à des calculs sur leurs coordonnées :

$$x = \sum_{k=1}^n \langle x, e_k \rangle e_k \quad \text{et} \quad \|x\|^2 = \sum_{k=1}^n |\langle x, e_k \rangle|^2.$$

L'objectif est d'étendre cette décomposition aux espaces de dimension infinie (comme les espaces de fonctions). On parlera alors de **série de Fourier généralisée**.

Dans toute cette section, on suppose que l'espace de Hilbert H est **séparable** (il admet une famille dénombrable dense), ce qui permet de manipuler des bases dénombrables (suites), suffisantes pour les applications usuelles.

4.4.1 Familles orthonormales et Inégalité de Bessel

Définition 4.7 (Système orthonormal). Une famille (suite) de vecteurs $(e_n)_{n \in \mathbb{N}}$ de H est dite **orthonormale** si :

$$\forall n, m \in \mathbb{N}, \quad \langle e_n, e_m \rangle = \delta_{nm} = \begin{cases} 1 & \text{si } n = m \\ 0 & \text{si } n \neq m \end{cases}$$

Soit $(e_n)_{n \in \mathbb{N}}$ un système orthonormal. Pour tout $x \in H$, on appelle **coefficients de Fourier** de x les scalaires $c_n(x) = \langle x, e_n \rangle$. Considérons la projection orthogonale de x sur le sous-espace F_N engendré par les $N + 1$ premiers vecteurs. D'après la structure orthonormée, elle vaut :

$$S_N(x) = \sum_{n=0}^N c_n(x) e_n.$$

Par le théorème de Pythagore, comme $(x - S_N(x)) \perp S_N(x)$, on a :

$$\|x\|^2 = \|S_N(x)\|^2 + \|x - S_N(x)\|^2 = \sum_{n=0}^N |c_n(x)|^2 + \|x - S_N(x)\|^2.$$

On en déduit immédiatement l'inégalité fondamentale suivante :

Proposition 4.8 (Inégalité de Bessel). *Pour tout $x \in H$, la série des carrés des coefficients de Fourier converge, et sa somme est majorée par la norme du vecteur :*

$$\sum_{n=0}^{\infty} |\langle x, e_n \rangle|^2 \leq \|x\|^2.$$

Corollaire 4.1 (Lemme de Riemann-Lebesgue). *Les coefficients de Fourier tendent vers 0 :*

$$\lim_{n \rightarrow \infty} \langle x, e_n \rangle = 0.$$

4.4.2 Définition d'une Base Hilbertienne

Attention, la notion de base en dimension infinie est subtile. Une base algébrique (toute famille finie est liée) n'a pas grand sens ici. On utilise une notion topologique (limite de séries).

Définition 4.8 (Base Hilbertienne). Soit $(e_n)_{n \in \mathbb{N}}$ une famille orthonormale de H . On dit que c'est une **base hilbertienne** (ou une base orthonormée complète) si elle vérifie l'une des propriétés équivalentes suivantes :

1. **Densité** : L'espace vectoriel engendré par la famille (les combinaisons linéaires *finies*) est dense dans H .
2. **Parseval** : Pour tout $x \in H$, on a égalité dans Bessel :

$$\|x\|^2 = \sum_{n=0}^{\infty} |\langle x, e_n \rangle|^2.$$

3. **Totalité** : Le vecteur nul est le seul vecteur orthogonal à tous les e_n :

$$(\forall n \in \mathbb{N}, \langle x, e_n \rangle = 0) \implies x = 0.$$

Remarque 4.3 (Attention au vocabulaire). En dimension infinie, il ne faut pas confondre :

- **Base algébrique (de Hamel)** : Famille génératrice par combinaisons linéaires *finies*. Elle est généralement indénombrable et inutilisable en analyse.
- **Base hilbertienne** : Famille orthonormée génératrice par combinaisons linéaires *infinies* (séries convergentes). C'est celle que nous utilisons ici.

Une base hilbertienne n'est **pas** une base au sens algébrique (car il faut en général combiner une infinité de l'éléments de la famille pour reconstruire un vecteur).

Théorème 4.6 (Décomposition en série de Fourier). *Si (e_n) est une base hilbertienne de H , alors tout vecteur $x \in H$ est somme de sa série de Fourier (convergence en norme) :*

$$x = \sum_{n=0}^{\infty} \langle x, e_n \rangle e_n \quad \text{au sens où} \quad \lim_{N \rightarrow \infty} \left\| x - \sum_{n=0}^N \langle x, e_n \rangle e_n \right\| = 0.$$

Remarque 4.4 (Vocabulaire). Le terme "Base Hilbertienne" est le plus correct. On évite de dire simplement "Base" pour ne pas confondre avec les bases de Hamel (algébriques).

4.4.3 Exemples classiques

Exemple 4.6 (Séries de Fourier classiques). Dans l'espace $L^2([0, 2\pi])$ des fonctions 2π -périodiques de carré intégrable, la famille

$$e_n(t) = \frac{1}{\sqrt{2\pi}} e^{int}, \quad n \in \mathbb{Z}$$

forme une base hilbertienne. L'égalité de Parseval s'écrit :

$$\frac{1}{2\pi} \int_0^{2\pi} |f(t)|^2 dt = \sum_{n=-\infty}^{+\infty} |c_n(f)|^2.$$

Exemple 4.7 (Polynômes orthogonaux). Dans $L^2([-1, 1])$, la famille des polynômes $1, X, X^2, \dots$ est libre mais pas orthogonale. En appliquant le procédé d'orthogonalisation de Gram-Schmidt, on obtient une base hilbertienne constituée des **Polynômes de Legendre** (normalisés).

4.4.4 Exemple détaillé : Les Ondelettes de Haar

Contrairement aux bases de Fourier (sinus/cosinus) qui sont délocalisées (elles oscillent sur tout \mathbb{R}), les ondelettes sont des fonctions à support compact capables d'analyser le signal localement.

Définition 4.9 (Ondelette Mère de Haar). On appelle **ondelette mère** de Haar la fonction $\psi \in L^2(\mathbb{R})$ définie par :

$$\psi(t) = \begin{cases} 1 & \text{si } 0 \leq t < 1/2 \\ -1 & \text{si } 1/2 \leq t < 1 \\ 0 & \text{sinon.} \end{cases}$$

C'est une oscillation élémentaire : son intégrale est nulle ($\int \psi = 0$) et sa norme L^2 vaut 1.

À partir de cette fonction, on construit une base en la translatant (paramètre k) et en la dilatant (paramètre j).

Proposition 4.9 (Base de Haar). Pour tout $j, k \in \mathbb{Z}$, on définit la fonction :

$$\psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k).$$

- L'indice j est l'**échelle** (ou la fréquence). Si j est grand, la fonction est très comprimée (support de taille 2^{-j}) et haute.
- L'indice k est la **position** (ou le temps). La fonction est centrée autour de $k2^{-j}$.
- Le facteur $2^{j/2}$ assure la normalisation : $\|\psi_{j,k}\|_2 = 1$.

La famille $\{\psi_{j,k}\}_{j,k \in \mathbb{Z}}$ forme une **base hilbertienne** de $L^2(\mathbb{R})$.

Cela signifie que toute fonction f d'énergie finie se décompose comme la somme de ses détails à différentes échelles :

$$f(t) = \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} \langle f, \psi_{j,k} \rangle \psi_{j,k}(t).$$

Remarque 4.5 (Lien avec le traitement du signal). Cette base est orthogonale car :

- Si les échelles j sont différentes, l'ondelette la plus fine oscille sur une zone où l'ondelette la plus large est constante. Comme $\int \psi = 0$, le produit scalaire est nul.

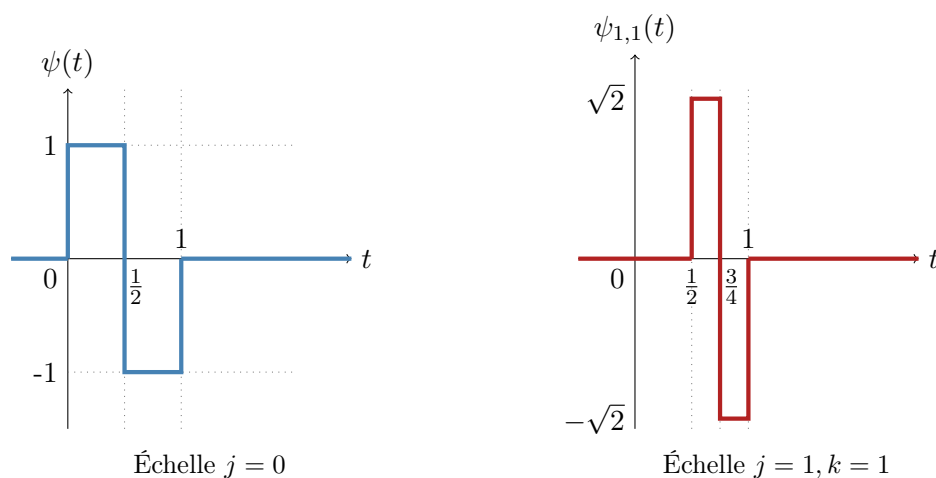


Figure 4.4 – Visualisation des ondelettes de Haar. À gauche, l'ondelette mère. À droite, une version compressée ($j = 1$) et translatée ($k = 1$). Notez que le support est réduit de moitié et l'amplitude augmentée pour conserver la norme L^2 .

- Si les échelles sont identiques mais les positions k différentes, les supports sont disjoints, donc le produit est nul.

C'est le principe de la compression d'image (JPEG 2000) : on ne garde que les coefficients $\langle f, \psi_{j,k} \rangle$ importants, qui correspondent aux zones de variation brusque (contours). Pour en savoir plus, un livre de référence est le livre de Stéphane Mallat [25]. Les cours de traitement d'images à l'école abordent aussi ces aspects.

💡 Pour aller plus loin : Espaces à Noyau Reproductif et Machine Learning

Dans ce chapitre, nous avons vu que les espaces de Hilbert généralisent la géométrie euclidienne. Une classe particulière d'espaces de Hilbert joue un rôle central en apprentissage statistique (Machine Learning) : les RKHS (Reproducing Kernel Hilbert Spaces).

Le problème de la classification non linéaire

Supposons que l'on veuille séparer deux classes de points (des ronds rouges et des croix bleues) dans \mathbb{R}^2 .

- Si les points sont séparables par une droite, un produit scalaire classique suffit pour définir l'hyperplan séparateur (c'est le principe du Perceptron ou du SVM linéaire).
- Si les points ne sont pas linéairement séparables (par exemple, les ronds sont au centre et les croix autour), il faut trouver une séparation courbe.

L'astuce (l'*astuce du noyau* ou *Kernel Trick*) consiste à envoyer les données dans un espace de Hilbert H de dimension beaucoup plus grande (voire infinie) via une application $\Phi : \mathbb{R}^n \rightarrow H$, où elles deviendront linéairement séparables. Cette idée d'augmenter la dimension est assez couramment utilisée en analyse pour rendre un problème mieux posé, ou plus facilement attaquant : par exemple la méthode Level Set pour la représentation implicite de courbes et surfaces, ou encore le doublement de variables présent sous diverses formes (passage au dual).

Définition d'un RKHS

Soit X un ensemble (l'espace des données). Un **Espace de Hilbert à Noyau Reproductif (RKHS)** est un espace de Hilbert H de fonctions $f : X \rightarrow \mathbb{R}$ tel que, pour tout $x \in X$, l'application d'évaluation :

$$\delta_x : f \mapsto f(x)$$

est une forme linéaire **continue**.

D'après le **théorème de représentation de Riesz**, puisque δ_x est continue, il existe un unique vecteur $K_x \in H$ tel que :

$$\forall f \in H, \quad f(x) = \langle f, K_x \rangle_H.$$

La fonction $K(x, y) = \langle K_x, K_y \rangle_H = K_x(y)$ est appelée le **Noyau Reproductif** de l'espace.

L'Astuce du Noyau (Kernel Trick)

Pourquoi est-ce utile ? Dans les algorithmes comme les SVM (Support Vector Machines), on cherche à calculer des produits scalaires entre les données transformées $\langle \Phi(x), \Phi(y) \rangle_H$. Si l'on choisit H comme étant le RKHS associé à un noyau K , et $\Phi(x) = K_x$, alors :

$$\langle \Phi(x), \Phi(y) \rangle_H = \langle K_x, K_y \rangle_H = K(x, y).$$

Un théorème stipule qu'il existe toujours un tel espace de Hilbert dès lors que K est symétrique défini positif. **Conclusion** : On peut calculer le produit scalaire dans l'espace de grande dimension H sans jamais connaître explicitement Φ ni H ! Il suffit d'évaluer la fonction noyau $K(x, y)$.

Exemple 4.8 (Noyau Gaussien RBF). Le noyau le plus célèbre est le noyau gaussien :

$$K(x, y) = \exp \left(-\frac{\|x - y\|^2}{2\sigma^2} \right).$$

Il correspond à un produit scalaire dans un espace de Hilbert de dimension infinie. Utiliser ce noyau dans un SVM permet de tracer des frontières de décision très complexes et lisses, simplement en calculant des exponentielles.

Source : adapté de [Wikipedia](#)

Chapitre 5

Fonctions périodiques et Séries de Fourier

L'idée fondamentale de l'analyse de Fourier est de décomposer un signal complexe (une fonction périodique) en une superposition de signaux élémentaires purs (des fonctions sinusoïdales). C'est l'analogie mathématique du prisme de Newton qui décompose la lumière blanche en un spectre de couleurs, ou de l'oreille humaine qui décompose une onde sonore en fréquences distinctes (hauteurs de notes).

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

5.1 Définitions et Coefficients de Fourier

5.1.1 Le cadre fonctionnel

Nous nous intéressons aux fonctions définies sur \mathbb{R} , à valeurs dans \mathbb{K} , et T -périodiques (où $T > 0$ est la période). On définit la **pulsation** :

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

Pour appliquer la théorie des espaces de Hilbert vue au chapitre précédent, nous devons définir un espace de fonctions muni d'un produit scalaire.

Définition 5.1 (Espace L^2_{per}). On note L^2_T (ou L^2_{per}) l'espace des fonctions T -périodiques $f : \mathbb{R} \rightarrow \mathbb{C}$ qui sont de carré intégrable sur une période (c'est-à-dire $\int_0^T |f(t)|^2 dt < +\infty$). On munit cet espace du **produit scalaire** hermitien normalisé :

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

La norme associée est la norme quadratique :

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

Remarque 5.1 (Normalisation). Le facteur $\frac{1}{T}$ est là pour que la norme d'une fonction constante $f(t) = 1$ soit égale à 1 (et non \sqrt{T}). Il permet de simplifier les formules de conservation d'énergie.

Nous utiliserons aussi l'espace \mathcal{CM}_{per} des fonctions T -périodiques, **continues par morceaux** sur \mathbb{R} (c'est-à-dire qu'elles n'ont qu'un nombre fini de discontinuités de première espèce sur une période).

5.1.2 Polynômes trigonométriques

Les « briques élémentaires » de notre théorie sont les oscillations harmoniques. Pour tout entier $n \in \mathbb{Z}$, on définit la fonction e_n par :

$$e_n(t) = e^{in\omega t} = \cos(n\omega t) + i \sin(n\omega t).$$

Proposition 5.1 (Orthonormalité). La famille $(e_n)_{n \in \mathbb{Z}}$ est une famille **orthonormale** dans L_T^2 .

Démonstration. Calculons le produit scalaire $\langle e_n, e_m \rangle$:

$$\langle e_n, e_m \rangle = \frac{1}{T} \int_0^T e^{in\omega t} \overline{e^{im\omega t}} dt = \frac{1}{T} \int_0^T e^{i(n-m)\omega t} dt.$$

Si $n = m$, on intègre la constante 1, on trouve $\frac{1}{T} \times T = 1$. Si $n \neq m$, on intègre une exponentielle 2π -périodique sur un nombre entier de périodes, l'intégrale est nulle. \square

Définition 5.2 (Polynôme trigonométrique). On appelle polynôme trigonométrique toute combinaison linéaire *finie* des fonctions e_n . L'ensemble de ces polynômes, noté \mathcal{P}_T , est un sous-espace vectoriel de L_T^2 .

5.1.3 Coefficients de Fourier (Forme exponentielle)

Puisque (e_n) est une famille orthonormale, on peut projeter n'importe quelle fonction f sur les vecteurs de cette famille.

Définition 5.3 (Coefficients de Fourier complexes). Pour $f \in L_{per}^1$ (et a fortiori pour $f \in L_{per}^2$), on définit les coefficients de Fourier complexes $(c_n(f))_{n \in \mathbb{Z}}$ par :

$$c_n(f) = \langle f, e_n \rangle = \frac{1}{T} \int_0^T f(t) e^{-in\omega t} dt.$$

On associe alors à la fonction f sa **Série de Fourier** :

$$S(f)(t) = \sum_{n=-\infty}^{+\infty} c_n(f) e^{in\omega t}.$$

Attention, pour l'instant cette notation est formelle, nous ne savons pas si cette série converge, ni si elle converge vers f . C'est l'objet des sections suivantes.

Remarque 5.2 (Interprétation physique : Le Spectre). L'ensemble des nombres complexes $\{c_n(f)\}_{n \in \mathbb{Z}}$ constitue le **spectre fréquentiel** du signal.

- $|c_n(f)|$ est l'amplitude de l'harmonique de rang n (fréquence n/T).
- $\arg(c_n(f))$ est la phase à l'origine.

5.1.4 Coefficients de Fourier (Forme réelle)

Si f est à valeurs réelles, il est souvent plus naturel de décomposer f en sinus et cosinus. En regroupant les termes n et $-n$, et en utilisant les formules d'Euler, on obtient :

Définition 5.4 (Coefficients de Fourier réels). Pour $f \in L^1_{per}$, on définit pour $n \in \mathbb{N}$:

$$a_n(f) = \frac{2}{T} \int_0^T f(t) \cos(n\omega t) dt \quad (n \geq 1), \quad a_0(f) = \frac{1}{T} \int_0^T f(t) dt.$$

$$b_n(f) = \frac{2}{T} \int_0^T f(t) \sin(n\omega t) dt \quad (n \geq 1), \quad b_0(f) = 0.$$

La série de Fourier s'écrit alors sous forme réelle :

$$S(f)(t) \sim a_0(f) + \sum_{n=1}^{+\infty} [a_n(f) \cos(n\omega t) + b_n(f) \sin(n\omega t)].$$

Proposition 5.2 (Relations Coefficients Complexes \leftrightarrow Réels).

- $c_0 = a_0$ (valeur moyenne du signal).
- Pour $n \geq 1$: $c_n = \frac{a_n - ib_n}{2}$ et $c_{-n} = \frac{a_n + ib_n}{2}$.
- Inversement : $a_n = c_n + c_{-n}$ et $b_n = i(c_n - c_{-n})$.

5.1.5 Parité et Symétries

Les calculs d'intégrales peuvent être grandement simplifiés si la fonction présente des symétries. Pour cela, il peut être intéressant d'intégrer sur l'intervalle symétrique $[-T/2, T/2]$ plutôt que $[0, T]$.

Proposition 5.3. Soit f une fonction T -périodique à valeurs réelles.

1. **Fonction Paire** ($f(-t) = f(t)$) :
 - Tous les coefficients sinus sont nuls : $\forall n, b_n(f) = 0$.
 - $a_n(f) = \frac{4}{T} \int_0^{T/2} f(t) \cos(n\omega t) dt$.
2. **Fonction Impaire** ($f(-t) = -f(t)$) :
 - Tous les coefficients cosinus (y compris la moyenne) sont nuls : $\forall n, a_n(f) = 0$.
 - $b_n(f) = \frac{4}{T} \int_0^{T/2} f(t) \sin(n\omega t) dt$.
3. **Réalité du signal** : Si f est à valeurs réelles, les coefficients complexes vérifient la symétrie hermitienne :

$$c_{-n}(f) = \overline{c_n(f)}.$$

Exemple 5.1 (Calcul des coefficients du signal Carré). Considérons la fonction « créneau » f , 2π -périodique, définie sur $] -\pi, \pi]$ par :

$$f(t) = \begin{cases} -1 & \text{si } t \in] -\pi, 0[\\ 1 & \text{si } t \in]0, \pi[\\ 0 & \text{si } t = 0 \text{ ou } t = \pi \end{cases}$$

Ici la période est $T = 2\pi$, donc la pulsation est $\omega = 1$.

1. **Analyse de la parité** : On remarque que pour tout $t \notin \{k\pi\}$, $f(-t) = -f(t)$. La fonction est **impaire**. Par conséquent, tous les coefficients cosinus sont nuls, ainsi que la valeur moyenne :

$$a_n(f) = 0 \quad \text{pour tout } n \geq 0.$$

Il ne reste qu'à calculer les coefficients $b_n(f)$.

2. **Calcul des coefficients b_n** : D'après la formule pour les fonctions impaires :

$$b_n(f) = \frac{4}{T} \int_0^{T/2} f(t) \sin(nt) dt = \frac{2}{\pi} \int_0^{\pi} 1 \cdot \sin(nt) dt.$$

Calculons l'intégrale :

$$\int_0^{\pi} \sin(nt) dt = \left[-\frac{\cos(nt)}{n} \right]_0^{\pi} = -\frac{1}{n} (\cos(n\pi) - \cos(0)) = \frac{1 - (-1)^n}{n}.$$

On distingue alors deux cas selon la parité de n :

- Si n est pair ($n = 2k$) : $1 - (-1)^{2k} = 0$, donc $b_{2k} = 0$.
- Si n est impair ($n = 2k + 1$) : $1 - (-1)^{2k+1} = 2$.

$$b_{2k+1} = \frac{2}{\pi} \times \frac{2}{2k+1} = \frac{4}{\pi(2k+1)}.$$

3. **Écriture de la Série de Fourier** : La série de Fourier associée à f est donc :

$$S(f)(t) = \sum_{k=0}^{+\infty} \frac{4}{\pi(2k+1)} \sin((2k+1)t) = \frac{4}{\pi} \left(\sin t + \frac{\sin 3t}{3} + \frac{\sin 5t}{5} + \dots \right).$$

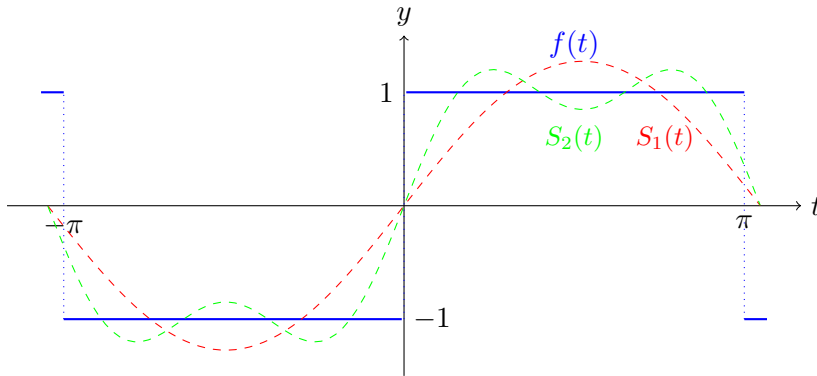


Figure 5.1 – Le signal carré (bleu) et les deux premiers termes de sa série de Fourier (rouge et vert).

Remarque 5.3 (Vitesse de décroissance). On observe que les coefficients b_n décroissent en $1/n$. Cette décroissance lente est caractéristique des fonctions discontinues.

5.2 Convergence quadratique (Cadre L^2)

Dans cette section, nous nous plaçons dans l'espace de Hilbert $H = L^2_{per}$ muni de la norme $\|f\|_2 = \sqrt{\frac{1}{T} \int_0^T |f(t)|^2 dt}$. Nous avons vu que la famille des exponentielles $(e_n)_{n \in \mathbb{Z}}$ est une famille orthonormale. La question centrale est de savoir si elle permet de reconstruire n'importe quelle fonction (c'est-à-dire si c'est une base hilbertienne).

5.2.1 Approximation et Inégalité de Bessel

Si l'on cherche à approcher f par un polynôme trigonométrique P_N de degré N (somme de $-N$ à N), la meilleure approximation au sens de l'énergie (moindres carrés) est donnée par la projection orthogonale, c'est-à-dire la somme partielle de la série de Fourier :

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

D'après l'inégalité de Bessel (vue au Chapitre 4), l'énergie de l'approximation est toujours inférieure à l'énergie du signal :

Proposition 5.4 (Inégalité de Bessel). *Pour toute fonction $f \in L^2_{per}$, la série des carrés des modules des coefficients de Fourier converge, et :*

$$\sum_{n=-\infty}^{+\infty} |c_n(f)|^2 \leq \|f\|_2^2 = \frac{1}{T} \int_0^T |f(t)|^2 dt.$$

Corollaire 5.1 (Lemme de Riemann-Lebesgue). *Les coefficients de Fourier tendent vers 0 à l'infini :*

$$\lim_{|n| \rightarrow \infty} c_n(f) = 0 \quad \text{et} \quad \lim_{n \rightarrow \infty} a_n(f) = \lim_{n \rightarrow \infty} b_n(f) = 0.$$

Cela signifie que les très hautes fréquences contiennent peu d'énergie.

5.2.2 Théorème de Parseval-Plancherel

Le résultat fondamental est que la famille (e_n) est non seulement orthonormale, mais **totale** (dense). C'est donc une base hilbertienne de L^2_{per} . Cela traduit la conservation exacte de l'énergie.

Théorème 5.1 (Identité de Parseval). *Pour toute fonction $f \in L^2_{per}$, on a l'égalité :*

$$\frac{1}{T} \int_0^T |f(t)|^2 dt = \sum_{n=-\infty}^{+\infty} |c_n(f)|^2.$$

Remarque 5.4 (Formule avec coefficients réels). Si f est à valeurs réelles, l'identité s'écrit avec les coefficients a_n et b_n :

$$\frac{1}{T} \int_0^T (f(t))^2 dt = a_0(f)^2 + \frac{1}{2} \sum_{n=1}^{+\infty} (a_n(f)^2 + b_n(f)^2).$$

On voit ici l'importance du facteur $1/2$ pour les harmoniques $n \geq 1$, qui vient du fait que la valeur efficace d'une sinusoïde est $1/\sqrt{2}$. Le terme a_0^2 (composante continue) n'a pas ce facteur.

5.2.3 Convergence en moyenne quadratique

Puisque (e_n) est une base hilbertienne, la série de Fourier converge vers la fonction pour la norme L^2 .

Théorème 5.2 (Convergence L^2). *Pour toute fonction $f \in L^2_{per}$, la suite des sommes partielles $S_N(f)$ converge vers f en moyenne quadratique :*

$$\lim_{N \rightarrow \infty} \|f - S_N(f)\|_2 = \lim_{N \rightarrow \infty} \sqrt{\frac{1}{T} \int_0^T |f(t) - S_N(f)(t)|^2 dt} = 0.$$

Remarque 5.5. La convergence en moyenne quadratique n'implique pas la convergence ponctuelle. Il est possible que $S_N(f)(t)$ diverge en certains points t , même si l'intégrale de l'erreur tend vers 0. C'est l'objet de la section suivante.

5.2.4 Exemple : Calcul de $\sum \frac{1}{n^2}$

L'identité de Parseval est un outil puissant pour calculer des sommes de séries numériques. Reprenons l'exemple du signal carré f (2π -périodique, impair, vaut 1 sur $]0, \pi[$).

- **Énergie temporelle** : $|f(t)|^2 = 1$ partout (sauf en 0), donc $\frac{1}{2\pi} \int_{-\pi}^{\pi} 1 dt = 1$.
- **Énergie fréquentielle** : Nous avons calculé $a_n = 0$ et $b_{2k} = 0$, $b_{2k+1} = \frac{4}{\pi(2k+1)}$. La formule de Parseval donne :

$$1 = 0^2 + \frac{1}{2} \sum_{k=0}^{\infty} \left(\frac{4}{\pi(2k+1)} \right)^2 = \frac{8}{\pi^2} \sum_{k=0}^{\infty} \frac{1}{(2k+1)^2}.$$

On en déduit la somme des inverses des carrés impairs : $\sum_{k=0}^{\infty} \frac{1}{(2k+1)^2} = \frac{\pi^2}{8}$. (On peut ensuite en déduire $\sum \frac{1}{n^2} = \frac{\pi^2}{6}$ en séparant pairs et impairs).

5.3 Convergences Ponctuelle et Uniforme

La convergence en moyenne quadratique (L^2) est très satisfaisante pour l'énergie, mais elle ne dit rien sur ce qui se passe en un point précis t . Il est possible que la série de Fourier diverge en un point, ou converge vers une valeur différente de $f(t)$. Les théorèmes de cette section donnent des conditions suffisantes pour assurer que $S_N(f)(t) \rightarrow f(t)$.

5.3.1 Le Théorème de Dirichlet (Convergence ponctuelle)

Ce théorème est le plus utile en pratique car il s'applique aux fonctions « standard » de l'ingénieur, qui sont souvent régulières par morceaux (comme le signal carré ou la dent de scie).

Définition 5.5 (Fonction \mathcal{C}^1 par morceaux). Une fonction T -périodique f est dite \mathcal{C}^1 **par morceaux** si, sur une période $[0, T]$, elle est continue et dérivable sauf peut-être en un nombre fini de points t_k , et si en ces points elle admet des limites finies à gauche et à droite (ainsi que sa dérivée).

On note $f(t^+) = \lim_{\varepsilon \rightarrow 0^+} f(t + \varepsilon)$ et $f(t^-) = \lim_{\varepsilon \rightarrow 0^+} f(t - \varepsilon)$. On appelle **régularisée** de f la fonction \tilde{f} définie par la moyenne du saut :

$$\tilde{f}(t) = \frac{f(t^+) + f(t^-)}{2}.$$

Si f est continue en t , alors $\tilde{f}(t) = f(t)$.

Théorème 5.3 (Théorème de Dirichlet). Soit f une fonction T -périodique et \mathcal{C}^1 par morceaux. Alors, pour tout réel t , la série de Fourier de f converge vers sa régularisée :

$$\lim_{N \rightarrow \infty} S_N(f)(t) = \sum_{n=-\infty}^{+\infty} c_n(f) e^{in\omega t} = \frac{f(t^+) + f(t^-)}{2}.$$

Démonstration (Cas $T = 2\pi$). La preuve repose sur l'utilisation du **Noyau de Dirichlet**.

Étape 1 : Expression intégrale de la somme partielle La somme partielle d'ordre N s'écrit :

$$S_N(f)(t) = \sum_{n=-N}^N c_n(f) e^{int} = \sum_{n=-N}^N \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} f(u) e^{-inu} du \right) e^{int}.$$

Par linéarité de l'intégrale, on rentre la somme :

$$S_N(f)(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(u) \left(\sum_{n=-N}^N e^{in(t-u)} \right) du.$$

Posons $x = u - t$. Par périodicité, on peut intégrer sur $[-\pi, \pi]$:

$$S_N(f)(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t+x) D_N(x) dx$$

où $D_N(x) = \sum_{n=-N}^N e^{-inx}$ est le **Noyau de Dirichlet**. C'est une somme géométrique. Pour $x \notin 0[2\pi]$, on montre facilement que :

$$D_N(x) = \frac{\sin((N+1/2)x)}{\sin(x/2)}.$$

Étape 2 : Isoler la singularité On sait que $\frac{1}{2\pi} \int_{-\pi}^{\pi} D_N(x) dx = 1$. Par symétrie du noyau (fonction paire), on a $\frac{1}{2\pi} \int_0^{\pi} D_N(x) dx = 1/2$. Nous voulons montrer que $S_N(f)(t)$ tend vers $\ell = \frac{f(t^+) + f(t^-)}{2}$. Écrivons la différence (l'erreur) :

$$S_N(f)(t) - \ell = \frac{1}{2\pi} \int_{-\pi}^0 (f(t+x) - f(t^-)) D_N(x) dx + \frac{1}{2\pi} \int_0^{\pi} (f(t+x) - f(t^+)) D_N(x) dx.$$

Traisons la seconde intégrale (la première se traite de façon symétrique). Notons-la I_N .

$$I_N = \frac{1}{2\pi} \int_0^{\pi} \underbrace{\left(\frac{f(t+x) - f(t^+)}{\sin(x/2)} \right)}_{g(x)} \sin((N+1/2)x) dx.$$

Étape 3 : Application du Lemme de Riemann-Lebesgue La fonction $g(x)$ définie ci-dessus pose question en $x = 0$. Cherchons sa limite quand $x \rightarrow 0^+$:

$$g(x) = \frac{f(t+x) - f(t^+)}{x} \times \frac{x}{\sin(x/2)}.$$

Quand $x \rightarrow 0$, $\frac{x}{\sin(x/2)} \rightarrow 2$. Le terme $\frac{f(t+x) - f(t^+)}{x}$ tend vers la dérivée à droite $f'_d(t)$, qui existe car f est supposée \mathcal{C}^1 par morceaux. Ainsi, la fonction g est continue par morceaux sur $[0, \pi]$. L'intégrale I_N est donc un coefficient de Fourier (en sinus) de la fonction g . D'après le **Lemme de Riemann-Lebesgue** (conséquence de Bessel), ce coefficient tend vers 0 quand la fréquence $(N+1/2)$ tend vers l'infini.

Conclusion : Les deux intégrales d'erreur tendent vers 0. On a donc bien $\lim S_N(f)(t) = \ell$. \square

Exemple 5.2 (Retour sur le signal carré). Pour la fonction créneau (impair, vaut 1 sur $]0, \pi[$), le théorème assure que :

- Pour $t \in]0, \pi[$, f est continue, donc la série converge vers $f(t) = 1$.
- En $t = 0$, la fonction a un saut entre -1 et 1 . La série doit converger vers $\frac{1+(-1)}{2} = 0$. C'est bien le cas car tous les termes $\sin(k \cdot 0)$ sont nuls.

5.3.2 Convergence Normale et Uniforme

La convergence ponctuelle peut être lente ou irrégulière. La convergence uniforme est beaucoup plus forte : elle assure que l'erreur maximale $\sup_t |f(t) - S_N(f)(t)|$ tend vers 0.

Rappelons que la série de Fourier converge **normalement** sur \mathbb{R} si la série numérique des normes $\sum \|c_n e^{in\omega t}\|_\infty = \sum |c_n|$ est convergente. La convergence normale implique la convergence uniforme.

Théorème 5.4 (Convergence Normale). Soit f une fonction T -périodique. Si :

1. f est **continue** sur \mathbb{R} ,
2. f est \mathcal{C}^1 par morceaux,

alors la série de Fourier de f converge **normalement** (et donc uniformément) vers f sur \mathbb{R} .

Idée de la preuve. Si f est continue et \mathcal{C}^1 par morceaux, sa dérivée f' est définie presque partout et est dans L^2 . On peut relier les coefficients : $c_n(f') = in\omega c_n(f)$. Donc $|c_n(f)| = \frac{1}{n\omega} |c_n(f')|$. Comme $f' \in L^2$, la série $\sum |c_n(f')|^2$ converge (Bessel). On utilise l'inégalité de Cauchy-Schwarz pour les séries :

$$\sum |c_n(f)| = \frac{1}{\omega} \sum \frac{1}{n} |c_n(f')| \leq \frac{1}{\omega} \sqrt{\sum \frac{1}{n^2}} \sqrt{\sum |c_n(f')|^2} < +\infty.$$

La série des coefficients est absolument convergente, donc la convergence est normale. \square

Remarque 5.6. Si f vérifie ces hypothèses (continue et \mathcal{C}^1 pm), il n'y a pas de phénomène d'oscillation résiduelle aux abords des points difficiles : l'approximation colle parfaitement à la courbe quand N est grand.

5.4 Vitesse de convergence et Phénomène de Gibbs

Pour les applications numériques (compression, transmission), il est fondamental de savoir à quelle vitesse la série de Fourier converge. Doit-on garder 10 termes ou 10 000 pour avoir une bonne approximation ? La réponse dépend directement de la régularité de la fonction.

5.4.1 Lien entre Régularité et Décroissance des coefficients

Plus une fonction est régulière (lisse), plus ses coefficients de Fourier décroissent vite vers 0. Inversement, une décroissance lente signale des singularités (sauts, pointes).

On rappelle la relation obtenue par intégration par parties : si f est dérivable, $c_n(f') = in\omega c_n(f)$. Ainsi, $|c_n(f)| = \frac{1}{|n|\omega} |c_n(f')|$.

Théorème 5.5 (Régularité et décroissance). Soit f une fonction T -périodique.

1. **Fonction discontinue (saut) :** Si f présente des discontinuités de première espèce (comme le créneau), alors les coefficients décroissent lentement, en $\frac{1}{n}$.

$$c_n(f) = O\left(\frac{1}{n}\right).$$

2. **Fonction continue, dérivée discontinue (coin) :** Si f est continue mais que f' a des sauts (comme le signal triangulaire), alors les coefficients décroissent en $\frac{1}{n^2}$.

$$c_n(f) = O\left(\frac{1}{n^2}\right).$$

3. **Fonction de classe \mathcal{C}^k** : Si f est k fois continûment dérivable, alors :

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

4. **Fonction \mathcal{C}^∞** : Si f est indéfiniment dérivable, les coefficients décroissent plus vite que n'importe quelle puissance de n (décroissance rapide).

Exemple 5.3. — **Signal Carré (discontinu)** : $b_n \sim \frac{1}{n}$. Convergence lente. Il faut beaucoup d'harmoniques pour reconstruire le front raide.

— **Signal Triangle (continu)** : On intègre le signal carré. Les coefficients sont divisés par n , donc $b_n \sim \frac{1}{n^2}$. Convergence rapide. Quelques harmoniques suffisent pour une très bonne approximation visuelle.

5.4.2 Le Phénomène de Gibbs

Que se passe-t-il exactement lorsqu'on approxime une fonction discontinue par une somme finie de sinus (qui sont des fonctions très lisses) ? Le théorème de Dirichlet nous dit qu'il y a convergence simple (point par point). Mais la convergence n'est pas uniforme au voisinage du saut.

On observe une oscillation parasite qui ne disparaît pas, même quand le nombre de termes N tend vers l'infini. C'est le **Phénomène de Gibbs**.

Proposition 5.5 (Description du phénomène). *Considérons le signal créneau f sautant de -1 à 1 en $t = 0$. Notons $S_N(f)$ sa somme partielle d'ordre N .*

- *Quand $N \rightarrow \infty$, $S_N(f)(t)$ tend vers $f(t)$ pour tout t fixé.*
- *Cependant, l'amplitude maximale de l'oscillation près du saut ne tend pas vers 1, mais vers une valeur plus grande :*

$$\lim_{N \rightarrow \infty} \left(\max_t S_N(f)(t) \right) = \frac{2}{\pi} \int_0^\pi \frac{\sin x}{x} dx \approx 1.17898.$$

*Cela signifie qu'il y a un **dépassement** (overshoot) systématique d'environ 9% de la hauteur du saut (1.18 au lieu de 1, soit 18% de l'amplitude totale du saut).*

Remarque 5.7 (Conséquence pratique : Artefacts de compression). Ce phénomène a des conséquences concrètes en traitement d'image (JPEG). Les contours nets d'une image correspondent à des sauts de luminosité (discontinuités). La compression, qui supprime les hautes fréquences, est équivalente à une troncature de la série de Fourier. Cela crée des « échos » visuels ou des ondulations parasites (ringing artifacts) autour des bords des objets.

Pour aller plus loin : Inégalité de Wirtinger et Isopérimétrie

À quoi servent les séries de Fourier en dehors du traitement du signal ? Elles permettent de résoudre des problèmes géométriques complexes avec une élégance rare. Voici l'exemple le plus célèbre.

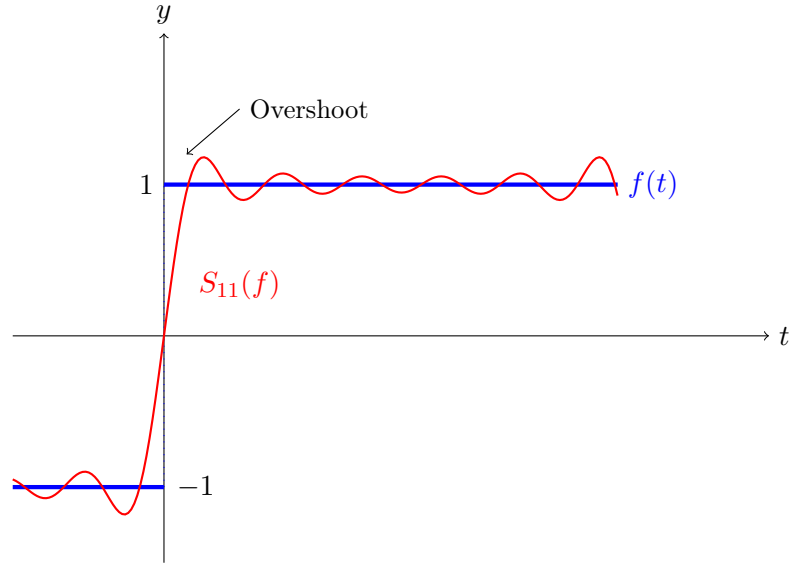


Figure 5.2 – Illustration du phénomène de Gibbs pour $N = 11$. On observe le dépassement caractéristique au voisinage de la discontinuité en 0.

L'inégalité de Wirtinger

Ce résultat relie la norme L^2 d'une fonction à celle de sa dérivée. Intuitivement, si une fonction oscille beaucoup (dérivée grande), elle s'éloigne de sa moyenne.

Théorème 5.6 (Wirtinger). Soit f une fonction 2π -périodique, de classe \mathcal{C}^1 , et de **moyenne nulle** (i.e. $\int_0^{2\pi} f(t) dt = 0$). Alors :

$$\int_0^{2\pi} |f(t)|^2 dt \leq \int_0^{2\pi} |f'(t)|^2 dt.$$

De plus, l'égalité a lieu si et seulement si f est une sinusoïde pure : $f(t) = A \cos(t) + B \sin(t)$.

Démonstration. La démonstration est une application directe de l'identité de Parseval. Soient $(c_n(f))$ les coefficients de Fourier de f .

1. Puisque la moyenne est nulle, $c_0(f) = 0$.
2. La relation dérivation/Fourier donne $c_n(f') = in c_n(f)$.
3. D'après Parseval :

$$\frac{1}{2\pi} \int |f|^2 = \sum_{n \in \mathbb{Z}^*} |c_n|^2 \quad \text{et} \quad \frac{1}{2\pi} \int |f'|^2 = \sum_{n \in \mathbb{Z}^*} |n c_n|^2 = \sum_{n \in \mathbb{Z}^*} n^2 |c_n|^2.$$

Pour tout entier non nul $n \in \mathbb{Z}^*$, on a $n^2 \geq 1$. D'où le résultat.

L'égalité n'est possible que si les coefficients sont nuls pour tout $|n| > 1$, c'est-à-dire si f est un polynôme trigonométrique de degré 1. \square

Application : Le problème de Didon (Isopérimétrie)

La légende raconte que la reine Didon fonda Carthage en délimitant un terrain avec une peau de bœuf découpée en fines lanières. Pour avoir le plus grand terrain possible, elle forma un cercle. Mathématiquement : *Parmi toutes les courbes fermées simples de longueur ℓ fixée, laquelle délimite la plus grande aire \mathcal{A} ?*

Théorème 5.7 (Inégalité Isopérimétrique). *Pour toute courbe fermée simple de classe \mathcal{C}^1 de longueur ℓ et d'aire \mathcal{A} , on a :*

$$\mathcal{A} \leq \frac{\ell^2}{4\pi}.$$

L'égalité n'est atteinte que pour le cercle.

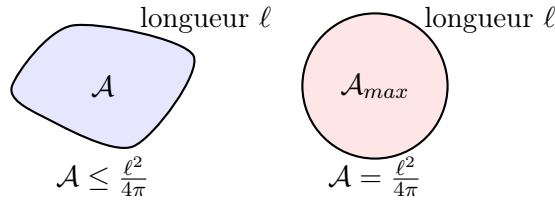


Figure 5.3 – L'inégalité isopérimétrique : le cercle maximise l'aire pour un périmètre donné.

Esquisse de la preuve via Fourier. On cherche à relier l'aire \mathcal{A} et la longueur ℓ .

Soit γ une courbe fermée simple de classe \mathcal{C}^1 . On la paramètre par une variable $t \in [0, 2\pi]$ proportionnelle à l'abscisse curviligne. Si s est l'abscisse curviligne ($s \in [0, \ell]$), on pose $t = \frac{2\pi s}{\ell}$. On écrit $\gamma(t) = (x(t), y(t))$. La norme du vecteur dérivé, tangent à la courbe, est constante :

$$x'(t)^2 + y'(t)^2 = \left(\frac{ds}{dt}\right)^2 = \left(\frac{\ell}{2\pi}\right)^2.$$

En intégrant cette constante sur $[0, 2\pi]$, on a donc :

$$\int_0^{2\pi} x'(t)^2 + y'(t)^2 dt = \frac{\ell^2}{2\pi}. \quad (*)$$

Quitte à traduire la courbe, ce qui ne change ni sa longueur ni son aire, on peut que la valeur moyenne de x est nulle, ce qui permettra d'appliquer l'inégalité de Wirtinger à x . D'après la formule de Green-Riemann, l'aire algébrique est donnée par $\mathcal{A} = \int_{\gamma} x dy = \int_0^{2\pi} x(t)y'(t) dt$. Utilisons l'inégalité élémentaire $ab \leq \frac{1}{2}(a^2 + b^2)$:

$$\mathcal{A} = \int_0^{2\pi} x(t)y'(t) dt \leq \frac{1}{2} \int_0^{2\pi} (x(t)^2 + y'(t)^2) dt.$$

$$\mathcal{A} \leq \frac{1}{2} \int_0^{2\pi} x(t)^2 dt + \frac{1}{2} \int_0^{2\pi} y'(t)^2 dt.$$

Comme x est \mathcal{C}^1 , 2π -périodique et de moyenne nulle, l'**inégalité de Wirtinger** assure que :

$$\int_0^{2\pi} x(t)^2 dt \leq \int_0^{2\pi} x'(t)^2 dt.$$

On a donc finalement :

$$\mathcal{A} \leq \frac{1}{2} \int_0^{2\pi} (x'(t)^2 + y'(t)^2) dt.$$

En utilisant l'égalité (*) établie à l'étape 1 on obtient l'inégalité isopérimétrique :

$$\mathcal{A} \leq \frac{1}{2} \cdot \frac{L^2}{2\pi} = \frac{\ell^2}{4\pi}.$$

Pour avoir égalité, il faut avoir égalité à chaque étape :

— Dans Wirtinger : $x(t)$ doit être de la forme $A \cos t + B \sin t$.

— Dans $ab \leq \frac{1}{2}(a^2 + b^2)$: il faut $x(t) = y'(t)$ d'où $y(t) = A \sin t - B \cos t + y_c$, $y_c \in \mathbb{R}$.

On a donc $x(t)^2 + (y(t) - y_c)^2 = A^2 + B^2$. La courbe est un cercle. □

Source : Vous trouverez plus de détails et une démonstration détaillée qui a inspiré celle-ci, au début du chapitre 4 du livre de E.M. Stein et R. Shakarchi [26].

Chapitre 6

Méthodes numériques de base : Systèmes linéaires

6.1 Méthodes itératives pour des systèmes linéaires

On sait aujourd'hui résoudre numériquement des systèmes linéaires de l'ordre du million d'inconnues (et d'équations). Pour des systèmes **creux** (c'est-à-dire lorsque la matrice du système possède beaucoup de coefficients nuls), on arrive à une centaine de millions d'inconnues.

Les systèmes pleins font appel à des **méthodes directes** qui donnent la solution exacte (aux erreurs d'arrondi près) en un nombre fini d'opérations, et seront décrites dans un chapitre ultérieur.

Pour les très grands systèmes creux (comme ceux issus de la discrétisation par différences finies de problèmes aux limites pour des équations aux dérivées partielles, par exemple en dimension 3 : N^3 points de discrétisation $\implies N = 100 \rightarrow 10^6$ inconnues), on utilise des **méthodes itératives**.

Le principe est de construire une suite de vecteurs qui converge vers la solution. L'intérêt est que ces méthodes ne manipulent pas la matrice (ce qui économiserait de la mémoire), mais seulement une fonction qui définit une suite par récurrence.

6.1.1 Définitions et convergence

Définition 6.1. Soit $A \in \mathcal{M}_n(\mathbb{R})$ inversible et $b \in \mathbb{R}^n$. On appelle **méthode itérative** de résolution du système linéaire $Ax = b$ ($x \in \mathbb{R}^n$) une méthode qui construit une suite récurrente $(x_k)_{k \geq 0}$ telle que :

$$(x_k \rightarrow x) \implies Ax = b.$$

Une méthode itérative est **convergente** pour toute condition initiale x_0 si pour tout $x_0 \in \mathbb{R}^n$, la suite (x_k) converge vers $x = A^{-1}b$.

La forme la plus générale d'une telle méthode est $x_{k+1} = \Phi(x_k)$ où $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ est une fonction continue vérifiant $\Phi(x) = x \iff Ax = b$.

Nous nous intéresserons aux **méthodes linéaires**, de la forme :

$$x_{k+1} = Bx_k + c$$

où $B \in \mathcal{M}_n(\mathbb{R})$ est la **matrice d'itération** et $c \in \mathbb{R}^n$ un vecteur. Pour que la méthode soit consistante (i.e. $x = Bx + c \iff Ax = b$), il faut un lien entre (B, c) et (A, b) .

Proposition 6.1 (Erreur et Convergence). Soit x la solution de $Ax = b$. On définit l'erreur à l'étape k par $e_k = x_k - x$. On a la relation :

$$e_{k+1} = x_{k+1} - x = (Bx_k + c) - (Bx + c) = B(x_k - x) = Be_k.$$

Par récurrence, $e_k = B^k e_0$. La méthode converge pour tout x_0 si et seulement si $\lim_{k \rightarrow \infty} B^k = 0$.

Théorème 6.1 (Condition spectrale). La méthode itérative $x_{k+1} = Bx_k + c$ converge pour tout choix initial x_0 si et seulement si le rayon spectral de B est strictement inférieur à 1 :

$$\rho(B) < 1.$$

Remarque 6.1. Rappelons que $\rho(B) = \max\{|\lambda|; \lambda \in \text{Sp}(B)\}$. On a toujours $\rho(B) \leq |||B|||$ pour toute norme matricielle subordonnée. Une condition suffisante de convergence est donc $|||B||| < 1$.

6.2 Méthodes par décomposition (Splitting)

La plupart des méthodes classiques sont basées sur une décomposition de la matrice A sous la forme :

$$A = M - N$$

où M est une matrice **inversible** ("facile" à inverser, par exemple diagonale ou triangulaire). Le système $Ax = b$ s'écrit $(M - N)x = b \iff Mx = Nx + b$. On définit l'algorithme :

$$Mx_{k+1} = Nx_k + b \iff x_{k+1} = M^{-1}Nx_k + M^{-1}b.$$

C'est une méthode linéaire avec $B = M^{-1}N$ et $c = M^{-1}b$. La méthode converge si et seulement si $\rho(M^{-1}N) < 1$.

6.2.1 Décomposition standard

On décompose la matrice A en trois parties : sa diagonale (D), sa partie triangulaire inférieure stricte ($-E$) et sa partie triangulaire supérieure stricte ($-F$). Attention aux signes, on écrit :

$$A = D - E - F$$

avec :

$$D = \begin{pmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix}, \quad E = - \begin{pmatrix} 0 & & 0 \\ a_{21} & 0 & \\ \vdots & \ddots & \ddots \\ a_{n1} & \dots & a_{n,n-1} & 0 \end{pmatrix}, \quad F = - \begin{pmatrix} 0 & a_{12} & \dots & a_{1n} \\ & 0 & \ddots & \vdots \\ & & \ddots & a_{n-1,n} \\ 0 & & & 0 \end{pmatrix}.$$

On suppose dans toute la suite que $a_{ii} \neq 0$ pour tout i (sinon on permute des lignes/colonnes). Donc D est inversible.

6.2.2 La méthode de Jacobi

On choisit $M = D$ et $N = E + F$. L'itération est $Dx_{k+1} = (E + F)x_k + b$. Composante par composante, cela s'écrit :

$$a_{ii}x_i^{(k+1)} = b_i - \sum_{j \neq i} a_{ij}x_j^{(k)} \implies x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j \neq i} a_{ij}x_j^{(k)} \right).$$

La matrice d'itération est $J = D^{-1}(E + F)$.

6.2.3 La méthode de Gauss-Seidel

L'idée est d'utiliser les composantes de x_{k+1} dès qu'elles sont calculées. Pour calculer $x_i^{(k+1)}$, on utilise $x_1^{(k+1)}, \dots, x_{i-1}^{(k+1)}$ qui viennent d'être mis à jour, et $x_{i+1}^{(k)}, \dots, x_n^{(k)}$.

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j < i} a_{ij} x_j^{(k+1)} - \sum_{j > i} a_{ij} x_j^{(k)} \right).$$

Matriciellement, cela revient à choisir $M = D - E$ (triangulaire inférieure) et $N = F$.

$$(D - E)x_{k+1} = Fx_k + b.$$

La matrice d'itération est $\mathcal{L}_1 = (D - E)^{-1}F$.

6.2.4 Méthodes de relaxation (SOR)

On introduit un paramètre $\omega \in \mathbb{R}$ (paramètre de relaxation). L'idée est de définir x_{k+1} comme une moyenne pondérée entre ce que donnerait Gauss-Seidel (notons le \hat{x}) et l'ancien itéré x_k .

$$x_i^{(k+1)} = (1 - \omega)x_i^{(k)} + \omega\hat{x}_i^{(k+1)}.$$

En remplaçant \hat{x}_i par la formule de Gauss-Seidel :

$$x_i^{(k+1)} = (1 - \omega)x_i^{(k)} + \frac{\omega}{a_{ii}} \left(b_i - \sum_{j < i} a_{ij} x_j^{(k+1)} - \sum_{j > i} a_{ij} x_j^{(k)} \right).$$

On peut réécrire cela sous forme matricielle. En multipliant par $\frac{a_{ii}}{\omega}$ et en regroupant les termes en $k + 1$ à gauche :

$$\left(\frac{D}{\omega} - E \right) x_{k+1} = \left(\frac{1 - \omega}{\omega} D + F \right) x_k + b.$$

Cela correspond au splitting $A = M - N$ avec :

$$M = \frac{D}{\omega} - E, \quad N = \frac{1 - \omega}{\omega} D + F.$$

La matrice d'itération est $\mathcal{L}_\omega = (D - \omega E)^{-1}((1 - \omega)D + \omega F)$.

- Si $\omega = 1$, on retrouve Gauss-Seidel.
- Si $\omega < 1$, on parle de sous-relaxation.
- Si $\omega > 1$, on parle de sur-relaxation (Successive Over Relaxation - SOR).

Le saviez-vous ? D'où vient le mot "Relaxation" ?

Le terme **relaxation** peut sembler étrange pour désigner un algorithme de calcul intensif. Son origine est en fait mécanique et précède l'invention de l'ordinateur.

Dans les années 1930 et 1940, l'ingénieur britannique **Sir Richard Southwell** développa ces méthodes pour calculer les contraintes dans les structures métalliques (comme les charpentes d'avions ou de dirigeables).

- Il modélisait la structure comme un réseau de nœuds.
- Le **résidu** représentait les **forces résiduelles** (les tensions) s'exerçant sur les nœuds si l'on bloquait la structure dans une position fautive.
- L'étape de calcul consistait à « libérer » (ou relâcher) un nœud pour qu'il se déplace sous l'effet

de ces forces jusqu'à l'équilibre local, réduisant ainsi la tension globale.

On « relaxait » donc littéralement les contraintes mécaniques du système, nœud après nœud. À l'époque, ces calculs étaient effectués à la main par des équipes entières de calculateurs humains (souvent des calculatrices) remplissant des grilles géantes sur papier !

Source : [Wikipedia](#)

6.3 Résultats de convergence

6.3.1 Matrices à diagonale dominante

Définition 6.2. Une matrice A est dite **strictement diagonale dominante (s.d.d.)** si pour tout $i = 1 \dots n$:

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}|.$$

Proposition 6.2. Si A est strictement diagonale dominante, alors A est inversible.

Démonstration. Si A n'est pas inversible, il existe $x \neq 0$ tel que $Ax = 0$. Soit k l'indice tel que $|x_k| = \|x\|_\infty = \max_i |x_i|$. La ligne k de $Ax = 0$ donne $a_{kk}x_k = -\sum_{j \neq k} a_{kj}x_j$. En passant aux valeurs absolues :

$$|a_{kk}||x_k| \leq \sum_{j \neq k} |a_{kj}||x_j| \leq \left(\sum_{j \neq k} |a_{kj}| \right) |x_k|.$$

En divisant par $|x_k|$ (qui est non nul), on obtient $|a_{kk}| \leq \sum_{j \neq k} |a_{kj}|$, ce qui contredit l'hypothèse de dominance stricte. \square

Théorème 6.2 (Convergence pour les matrices s.d.d.). Si A est strictement diagonale dominante, alors :

1. La méthode de **Jacobi** converge ($\rho(J) < 1$).
2. La méthode de **Gauss-Seidel** converge ($\rho(\mathcal{L}_1) < 1$).

Éléments de preuve pour Jacobi. La matrice de Jacobi est $J = D^{-1}(E + F)$. Ses coefficients sont $J_{ij} = -\frac{a_{ij}}{a_{ii}}$ si $i \neq j$ et 0 sinon. On calcule la norme infinie de la matrice (max de la somme des modules sur les lignes) :

$$\|J\|_\infty = \max_i \sum_j |J_{ij}| = \max_i \sum_{j \neq i} \left| \frac{a_{ij}}{a_{ii}} \right| = \max_i \frac{1}{|a_{ii}|} \sum_{j \neq i} |a_{ij}|.$$

L'hypothèse s.d.d. implique exactement que cette quantité est strictement inférieure à 1. Donc $\rho(J) \leq \|J\|_\infty < 1$. \square

6.3.2 Théorème d'Ostrowski (Matrices symétriques définies positives)

Le résultat suivant est fondamental pour la méthode de relaxation. Il donne une condition nécessaire sur ω , et une condition nécessaire et suffisante si A est symétrique définie positive.

Théorème 6.3 (Householder-John). Pour que la méthode de relaxation (SOR) converge (c'est-à-dire $\rho(\mathcal{L}_\omega) < 1$), il est **nécessaire** que :

$$\omega \in]0, 2[.$$

Démonstration. Le terme constant du polynôme caractéristique de \mathcal{L}_ω est $\det(\mathcal{L}_\omega) = \det(M^{-1}N) = \frac{\det N}{\det M}$. Les matrices M et N sont triangulaires (ou presque), leur déterminant est le produit des termes diagonaux. $\det M = \prod \frac{a_{ii}}{\omega}$ et $\det N = \prod \frac{a_{ii}(1-\omega)}{\omega}$. Donc $\det(\mathcal{L}_\omega) = (1-\omega)^n$. Or le déterminant est le produit des valeurs propres : $|\det \mathcal{L}_\omega| = \prod |\lambda_i| \leq (\rho(\mathcal{L}_\omega))^n$. Si la méthode converge, $\rho < 1$, donc $|1 - \omega|^n < 1$, ce qui implique $|1 - \omega| < 1$, soit $0 < \omega < 2$. \square

Théorème 6.4 (Ostrowski). *Si A est **symétrique définie positive**, alors la méthode de relaxation converge **si et seulement si** :*

$$\omega \in]0, 2[.$$

Corollaire 6.1. *Si A est **symétrique définie positive**, alors la méthode de **Gauss-Seidel** converge (cas $\omega = 1$).*

6.3.3 Comparaison des vitesses de convergence

Il existe des cas où A est **symétrique définie positive** et où la méthode de Jacobi ne converge pas (alors que Gauss-Seidel converge toujours d'après le corollaire précédent).

Proposition 6.3 (Cas Tridiagonal). *Si A est **tridiagonale** (et D inversible), on peut montrer la relation suivante entre les rayons spectraux :*

$$\rho(\mathcal{L}_1) = (\rho(J))^2.$$

Conséquences :

- La méthode de Gauss-Seidel converge si et seulement si celle de Jacobi converge.
- Si elles convergent, Gauss-Seidel converge **deux fois plus vite** que Jacobi (asymptotiquement).
- Pour ce type de matrices, si les valeurs propres de J sont réelles, le paramètre optimal de relaxation ω_{opt} (celui qui minimise $\rho(\mathcal{L}_\omega)$) est donné par :

$$\omega_{opt} = \frac{2}{1 + \sqrt{1 - \rho(J)^2}}.$$

On a alors $1 < \omega_{opt} < 2$ (sur-relaxation).

6.4 Mise en œuvre pratique

6.4.1 Le résidu et le critère d'arrêt

Dans la pratique, on ne peut pas calculer l'erreur $e_k = x_k - x$ puisque la solution exacte x est inconnue. Pour contrôler la convergence, on utilise le **vecteur résidu** :

$$r_k = b - Ax_k.$$

Notez que $r_k = Ax - Ax_k = A(x - x_k) = -Ae_k$. Si la matrice A est bien conditionnée, une norme faible du résidu implique une norme faible de l'erreur. Un algorithme itératif s'arrête généralement lorsque le résidu relatif passe sous un seuil de tolérance ε (par exemple 10^{-6}) :

$$\frac{\|r_k\|}{\|b\|} \leq \varepsilon.$$

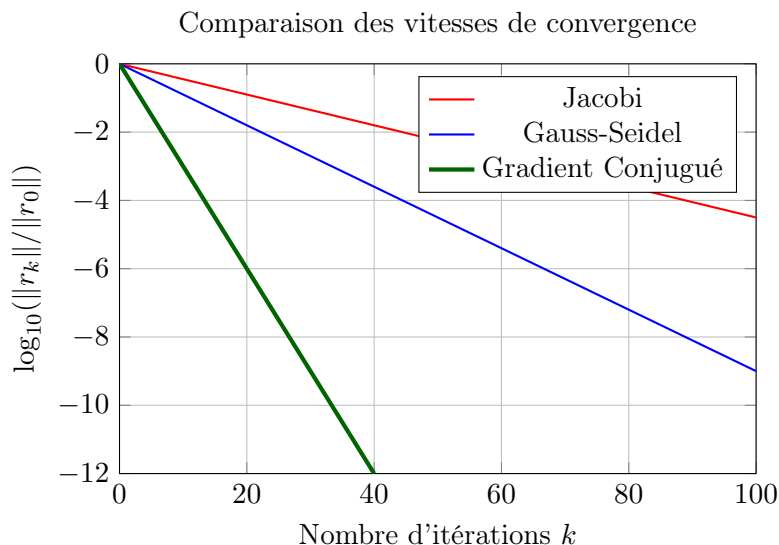


Figure 6.1 – Allure typique de la convergence du résidu pour les méthodes itératives. On remarque que Gauss-Seidel a une pente double de celle de Jacobi, tandis que le Gradient Conjugué (méthode de Krylov) est bien plus efficace.

6.4.2 Complexité algorithmique

Pour les méthodes de splitting (Jacobi, GS, SOR), le coût principal par itération est le produit matrice-vecteur. Si A est une matrice creuse contenant N_z éléments non nuls, le coût est de l'ordre de $O(N_z)$ opérations par itération. C'est très efficace comparé aux $O(N^3)$ d'une méthode directe (Gauss) pour une matrice pleine.

6.5 Introduction aux méthodes de sous-espaces de Krylov

Les méthodes précédentes (stationnaires) construisent x_{k+1} uniquement à partir de x_k . Les méthodes modernes cherchent x_{k+1} dans un espace plus grand engendré par les itérés précédents et le résidu, appelé **espace de Krylov** : $\mathcal{K}_k(A, r_0) = \text{Vect}(r_0, Ar_0, A^2r_0, \dots, A^{k-1}r_0)$.

6.5.1 La méthode du gradient conjugué (CG)

Si la matrice A est **symétrique définie positive**, résoudre $Ax = b$ équivaut à minimiser la fonctionnelle quadratique (voir Chapitre 3) :

$$J(y) = \frac{1}{2} \langle Ay, y \rangle - \langle b, y \rangle.$$

Le gradient de J en y est $\nabla J(y) = Ay - b = -r(y)$.

- **Descente de gradient (pas optimal)** : On pourrait définir $x_{k+1} = x_k + \alpha_k r_k$, où $r_k = r(x_k)$. Cela ressemble aux méthodes précédentes et converge lentement.
- **Gradient Conjugué** : L'idée est de choisir des directions de descente d_k qui sont A -orthogonales (conjuguées) entre elles, c'est-à-dire $\langle Ad_i, d_j \rangle = 0$ pour $i \neq j$.

Algorithme (résumé) : On part de x_0 , $r_0 = b - Ax_0$, $d_0 = r_0$. Tant que $\|r_k\| > \varepsilon$:

1. Pas optimal : $\alpha_k = \frac{\|r_k\|^2}{\langle Ad_k, d_k \rangle}$.

2. Mise à jour : $x_{k+1} = x_k + \alpha_k d_k$ et $r_{k+1} = r_k - \alpha_k A d_k$.
3. Nouvelle direction (orthogonalisation) : $\beta_k = \frac{\|r_{k+1}\|^2}{\|r_k\|^2}$, puis $d_{k+1} = r_{k+1} + \beta_k d_k$.

Théorème 6.5 (Convergence du GC). *En arithmétique exacte, le Gradient Conjugué converge vers la solution exacte en au plus n itérations (où n est la dimension du système). En pratique (avec erreurs d'arrondi), c'est une méthode itérative très rapide dont la vitesse dépend du conditionnement $\kappa(A) = \lambda_{\max}/\lambda_{\min}$.*

Des références classiques pour les méthodes itératives sont les livres [22] et [23], et plus récemment celui de Youcef Saad [24], de l'université du Minnesota, mathématicien appliqué algérien spécialiste mondial des ces questions, qui a d'ailleurs fait sa thèse à Grenoble.

Chapitre 7

Résolution numérique d'équations non linéaires

De nombreux problèmes issus de la physique ou de l'ingénierie conduisent à la résolution d'équations de la forme :

$$F(x) = 0$$

où $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ est une fonction non linéaire, supposée de classe \mathcal{C}^1 .

7.1 Exemples introductifs

7.1.1 Discrétisation de l'équation de Poisson-Boltzmann

Ce modèle décrit le potentiel électrique ϕ dans une solution ionique (par exemple autour d'une molécule d'ADN ou d'une protéine). En dimension 1, l'équation s'écrit (sous forme adimensionnée) :

$$\Delta\phi = \kappa^2 \sinh(\phi)$$

avec des conditions aux limites, par exemple $\phi(0) = 0$ et $\phi(1) = 1$. En discrétisant l'intervalle $[0, 1]$ avec un pas $h = \frac{1}{N+1}$ et les points $x_i = ih$, on approxime le laplacien $\Delta\phi(x_i) = \phi''(x_i)$ par différences finies :

$$\frac{\phi_{i+1} - 2\phi_i + \phi_{i-1}}{h^2} = \kappa^2 \sinh(\phi_i), \quad \text{pour } i = 1, \dots, N.$$

En notant $x = (\phi_1, \dots, \phi_N)^T$ le vecteur des inconnues, on obtient un système non linéaire de taille N . La i -ème équation est :

$$-\phi_{i-1} + 2\phi_i - \phi_{i+1} + h^2\kappa^2 \sinh(\phi_i) = 0.$$

7.1.2 Modes de vibration d'un cristal

Considérons un système de particules en interaction décrit par un Hamiltonien :

$$H(p, q) = \sum_n \frac{p_n^2}{2m} + \sum_{n \neq k} V(\|q_n - q_k\|).$$

Les positions d'équilibre correspondent aux minima de l'énergie potentielle, c'est-à-dire aux points où le gradient du potentiel s'annule :

$$\nabla_q V_{total}(q) = 0.$$

Si le potentiel V n'est pas quadratique (anharmonique), cela conduit à un système d'équations non linéaires couplées.

7.2 Méthode des approximations successives

7.2.1 Principe et définitions

Soit $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ une fonction continue. On appelle **point fixe** de Φ un point $a \in \mathbb{R}^n$ tel que :

$$\Phi(a) = a.$$

La méthode des approximations successives consiste à construire une suite $(x_k)_{k \in \mathbb{N}}$ par la récurrence :

$$\begin{cases} x_0 \in \mathbb{R}^n & (\text{donné}) \\ x_{k+1} = \Phi(x_k) \end{cases}$$

Si la suite (x_k) converge vers une limite ℓ , alors par continuité de Φ , on a $\ell = \Phi(\ell)$. La limite est donc nécessairement un point fixe de Φ .

D'après le théorème du point fixe contractant vu au chapitre 1, si Φ est contractante, elle admet un **unique** point fixe $a \in \mathbb{R}^n$ et pour tout choix initial $x_0 \in \mathbb{R}^n$, la suite des approximations successives $x_{k+1} = \Phi(x_k)$ converge vers a .

Remarque 7.1 (Vitesse de convergence). On a l'estimation d'erreur suivante :

$$\|x_k - a\| \leq \frac{K^k}{1 - K} \|x_1 - x_0\|.$$

La convergence est dite **linéaire** (ou géométrique), la vitesse étant pilotée par la constante de contraction K . Plus K est petit, plus la convergence est rapide.

Corollaire 7.1 (Convergence locale). Si Φ est de classe \mathcal{C}^1 et si a est un point fixe tel que $\|D\Phi(a)\| < 1$ (pour une norme subordonnée), alors il existe un voisinage de a sur lequel Φ est contractante. La méthode converge donc pour tout x_0 choisi assez près de a .

7.3 La méthode de Newton

7.3.1 Construction de la fonction d'itération

On revient au problème de trouver une racine de $F(x) = 0$. L'idée est de transformer cette équation en un problème de point fixe $x = \Phi(x)$ de sorte que la méthode des approximations successives converge **très vite** (plus vite que linéairement).

Si l'on écrit le développement de Taylor de F autour d'une racine a :

$$0 = F(a) = F(x) + DF(x)(a - x) + o(\|a - x\|).$$

Cela suggère que $a \approx x - [DF(x)]^{-1}F(x)$. On introduit donc la fonction d'itération de Newton :

$$\Phi(x) = x - [DF(x)]^{-1}F(x).$$

Chercher $F(x) = 0$ équivaut bien à chercher un point fixe de Φ (si $DF(x)$ est inversible). L'algorithme de Newton est donc une méthode d'approximations successives appliquée à cette fonction Φ :

$$x_{k+1} = x_k - [DF(x_k)]^{-1}F(x_k).$$

7.3.2 Analyse de la convergence

Pourquoi cette méthode est-elle plus efficace ? Regardons la dérivée de Φ au point fixe a . Calculons la dérivée de $\Phi(x)$ (en dimension 1 pour simplifier, le résultat est identique en dimension n) :

$$\Phi'(x) = 1 - \frac{d}{dx} \left(\frac{F(x)}{F'(x)} \right) = 1 - \frac{F'(x)^2 - F(x)F''(x)}{F'(x)^2} = \frac{F(x)F''(x)}{F'(x)^2}.$$

Au point fixe a (la solution), on a $F(a) = 0$. Donc :

$$\Phi'(a) = 0.$$

Comme la dérivée est nulle (donc très petite, bien inférieure à 1), la contraction est extrêmement forte au voisinage de la solution.

Théorème 7.1 (Convergence quadratique). *Si F est de classe \mathcal{C}^2 , si $F(a) = 0$ et si $DF(a)$ est inversible, alors la méthode de Newton converge **quadratiquement** localement. C'est-à-dire qu'il existe $C > 0$ tel que :*

$$\|x_{k+1} - a\| \leq C\|x_k - a\|^2.$$

Remarque 7.2. Alors que la méthode du point fixe classique donne une erreur en K^k , Newton donne une erreur en K^{2^k} . Le nombre de décimales exactes double à chaque itération.

7.4 La Méthode de Newton-Raphson

7.4.1 Construction de la méthode

Soit $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ une fonction de classe \mathcal{C}^1 . On cherche $x^* \in \mathbb{R}^n$ tel que $F(x^*) = 0$. Supposons que l'on dispose d'une approximation $x^{(k)}$ de la solution. Nous cherchons une correction $\delta \in \mathbb{R}^n$ telle que $x^{(k)} + \delta$ soit une meilleure approximation, idéalement une solution exacte :

$$F(x^{(k)} + \delta) = 0.$$

Effectuons un développement de Taylor de F au voisinage de $x^{(k)}$:

$$F(x^{(k)} + \delta) = F(x^{(k)}) + DF(x^{(k)}) \cdot \delta + o(\|\delta\|).$$

La matrice $DF(x^{(k)})$ est la **matrice Jacobienne** de F au point $x^{(k)}$, notée J_k . C'est la matrice carrée $n \times n$ dont les coefficients sont :

$$(J_k)_{ij} = \frac{\partial F_i}{\partial x_j}(x^{(k)}).$$

En négligeant le terme d'ordre supérieur $o(\|\delta\|)$, l'équation $F(x^{(k)} + \delta) = 0$ devient une équation linéaire en δ (l'équation tangente) :

$$F(x^{(k)}) + J_k \delta = 0.$$

Si la matrice Jacobienne J_k est inversible, on peut calculer la correction δ_k :

$$\delta_k = -J_k^{-1} F(x^{(k)}).$$

Le nouvel itéré est alors défini par $x^{(k+1)} = x^{(k)} + \delta_k$.

Algorithme 1 : Algorithme de Newton-Raphson

Données : F , sa Jacobienne J_F , point initial $x^{(0)}$, tolérance ε ;

$k \leftarrow 0$;

tant que $\|F(x^{(k)})\| > \varepsilon$ **et** $k < k_{max}$ **faire**

1. Calculer la matrice $A = J_F(x^{(k)})$;
2. Calculer le second membre $b = -F(x^{(k)})$;
3. Résoudre le système linéaire $A\delta = b$;
4. Mettre à jour $x^{(k+1)} = x^{(k)} + \delta$;
- $k \leftarrow k + 1$;

fin

7.4.2 Analyse de la convergence

La caractéristique principale de la méthode de Newton est sa vitesse de convergence.

Définition 7.1 (Ordre de convergence). Soit $(x^{(k)})$ une suite convergeant vers x^* . On dit que la convergence est d'**ordre** p s'il existe une constante $C > 0$ telle que pour k assez grand :

$$\|x^{(k+1)} - x^*\| \leq C \|x^{(k)} - x^*\|^p.$$

- Si $p = 1$ (et $C < 1$), la convergence est **linéaire** (ex : méthode de Picard, gradient à pas fixe).
- Si $p = 2$, la convergence est **quadratique**. Le nombre de décimales exactes double à chaque itération.

Théorème 7.2 (Convergence locale de Newton). Soit $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ une fonction de classe \mathcal{C}^2 . Soit x^* un zéro de F (i.e. $F(x^*) = 0$). Supposons que la matrice Jacobienne $J_F(x^*)$ est **inversible**.

Alors, il existe un voisinage \mathcal{V} de x^* tel que pour tout point initial $x^{(0)} \in \mathcal{V}$, la suite générée par la méthode de Newton est bien définie et converge vers x^* . De plus, la convergence est **quadratique**.

Démonstration. La démonstration repose sur la formule de Taylor avec reste intégral (ou l'inégalité des accroissements finis appliquée à la dérivée). Posons $e_k = x^{(k)} - x^*$ l'erreur à l'étape k . Par définition de l'algorithme :

$$x^{(k+1)} = x^{(k)} - J_F(x^{(k)})^{-1} F(x^{(k)}).$$

En soustrayant x^* des deux côtés :

$$e_{k+1} = e_k - J_F(x^{(k)})^{-1} (F(x^{(k)}) - \underbrace{F(x^*)}_0).$$

Factorisons $J_F(x^{(k)})^{-1}$:

$$e_{k+1} = J_F(x^{(k)})^{-1} \left[J_F(x^{(k)}) e_k - (F(x^{(k)}) - F(x^*)) \right].$$

Utilisons le développement de Taylor de F en $x^{(k)}$ pour évaluer le terme entre crochets. Comme F est \mathcal{C}^2 , on a :

$$F(x^*) = F(x^{(k)}) + J_F(x^{(k)})(x^* - x^{(k)}) + O(\|x^* - x^{(k)}\|^2).$$

Comme $x^* - x^{(k)} = -e_k$, cela s'écrit :

$$0 = F(x^{(k)}) - J_F(x^{(k)})e_k + O(\|e_k\|^2).$$

Donc le terme entre crochets dans l'expression de e_{k+1} est exactement un terme d'ordre 2 :

$$J_F(x^{(k)})e_k - (F(x^{(k)}) - F(x^*)) = O(\|e_k\|^2).$$

Ainsi :

$$\|e_{k+1}\| \leq \|J_F(x^{(k)})^{-1}\| \cdot C' \|e_k\|^2.$$

Puisque $x^{(k)} \rightarrow x^*$ et que $x \mapsto J_F(x)$ est continue et inversible en x^* , la norme $\|J_F(x^{(k)})^{-1}\|$ est bornée par une constante M au voisinage de x^* . On obtient finalement :

$$\|e_{k+1}\| \leq C \|e_k\|^2.$$

Cela prouve la convergence quadratique, à condition que l'erreur initiale e_0 soit suffisamment petite pour que $C\|e_0\| < 1$. □

7.4.3 Remarques sur la mise en œuvre

1. **Choix du point initial** : La contrainte locale est forte. Si $x^{(0)}$ est trop loin, la Jacobienne peut être singulière ou la méthode peut osciller.
2. **Critère d'arrêt** : On utilise généralement le résidu relatif :

$$\frac{\|F(x^{(k)})\|}{\|F(x^{(0)})\|} < \varepsilon \quad \text{ou} \quad \|x^{(k+1)} - x^{(k)}\| < \varepsilon.$$

3. **Coût de calcul** : La résolution du système $J_k \delta = -F_k$ coûte $O(n^3)$ opérations (factorisation LU). C'est le goulot d'étranglement pour les grands systèmes.

Avantages et Inconvénients :

- ✓ **Vitesse** : La convergence quadratique est extrêmement rapide. Une fois dans le bassin d'attraction, le nombre de chiffres significatifs double à chaque itération.
- ✗ **Localité** : La méthode ne converge que si l'initialisation $x^{(0)}$ est « assez proche » de la solution. Sinon, elle peut diverger.
- ✗ **Coût** : À chaque itération, il faut :
 1. Évaluer la matrice Jacobienne (coût n^2 évaluations de fonctions).
 2. Résoudre un système linéaire plein (coût $O(n^3)$ par Gauss).

7.5 Variantes et méthodes Quasi-Newton

Pour pallier le coût de calcul élevé, plusieurs variantes existent.

7.5.1 Approximation de la Jacobienne (Différences finies)

Si la forme analytique des dérivées partielles $\frac{\partial F_i}{\partial x_j}$ est inconnue ou trop complexe, on peut les approcher par différences finies :

$$\frac{\partial F_i}{\partial x_j}(x) \approx \frac{F_i(x + \delta e_j) - F_i(x)}{\delta}$$

avec δ petit. Cela évite le calcul formel mais ne réduit pas le coût algorithmique.

7.5.2 Méthode de Newton Modifiée (ou de la Corde)

Pour éviter de résoudre un système linéaire différent à chaque étape, on peut décider de geler la Jacobienne à sa valeur initiale $A = J_F(x^{(0)})$. L'algorithme devient :

$$x^{(k+1)} = x^{(k)} - A^{-1}F(x^{(k)}).$$

Avantage : On ne calcule la factorisation LU de A qu'une seule fois au début. À chaque itération, on ne fait qu'une descente-remontée (coût $O(n^2)$). **Inconvénient** : La convergence n'est plus que **linéaire** (comme une méthode de point fixe classique). On perd la vitesse quadratique.

7.5.3 Méthodes Quasi-Newton (Broyden)

Ces méthodes (vues en TD, comme la méthode de **Broyden**) cherchent un compromis. L'idée est de mettre à jour une approximation de la Jacobienne (ou de son inverse) à chaque itération en utilisant les variations de F , via des formules de rang 1, sans jamais recalculer de dérivées ni résoudre de système complet. Elles offrent une convergence **super-linéaire** (plus vite que linéaire, moins que quadratique) pour un coût modéré.

Références

- [1] J-P. Ansel et Y. Ducel, *Exercices corrigés en théorie de la mesure et de l'intégration*, Ellipses
- [2] N. Boccara, *Intégration*, Mathématiques pour l'ingénieur, Ellipses.
- [3] H. Cartan, *Calcul différentiel*, Hermann
- [4] G. Choquet, *Cours de topologie*, Dunod
- [5] C. Gasquet et P. Witomski, *Analyse de Fourier et applications : Filtrage, Calcul numérique, Ondelettes*, Dunod, 2000.
- [6] G. Gasquet, P. Witomski, *Analyse de Fourier et Applications*, Masson.
- [7] M. Mamode, *Mathématiques pour la physique*, Universités physique, Ellipses.
- [8] H. Reinhard, *Eléments de Mathématiques du signal. Tome 1 : signaux déterministes*, Dunod.
- [9] F. Roddier, *Distributions et transformée de Fourier à l'usage des physiciens et des ingénieurs*, Mc Graw Hill
- [10] E. Maitre et V. Perrier, *Analyse pour l'Ingénieur-e, Notes de cours ENSIMAG 1A*.
- [11] G. Carlier, *Calcul Différentiel et Optimisation*, cours Paris Dauphine : <https://www.ceremade.dauphine.fr/~carlier/calculdiff.pdf>
- [12] F. Rouvière, *Petit guide de calcul différentiel*, Cassini.
- [13] M. Crandall, H. Ishii, P.-L. Lions, *User's guide to viscosity solutions of second order partial differential equations*, <https://arxiv.org/abs/math/9207212> (pour les courageuses et courageux) !
- [14] A. Bressan, *Viscosity Solutions of Hamilton-Jacobi Equations and Optimal Control Problems*, Cours Penn State University : <https://sites.psu.edu/bressan/3-lecture-notes/>. Plus pédagogique que Crandall-Ishii-Lions.
- [15] H. Brezis, *Analyse fonctionnelle : Théorie et applications*, Dunod. (Référence classique pour les espaces de Banach).
- [16] G. Carlier, *Calcul Différentiel et Optimisation*, Notes de cours. (Utilisé pour le contexte optimisation/-calcul diff).
- [17] T. Gallouët et R. Herbin, *Mesure, Intégration, Probabilités*. (Pour les exemples sur les espaces L^p).
- [18] E. H. Lieb et M. Loss, *Analysis*, Graduate Studies in Mathematics, Vol. 14, American Mathematical Society, 2nd edition, 2001.
- [19] W. Rudin, *Analyse réelle et complexe*, Dunod, 3ème édition, 1998.

- [20] M. Briane et G. Pagès, *Théorie de l'intégration : Convolution et transformée de Fourier*, Vuibert, 6ème édition, 2015.
- [21] H. Lebesgue, *Leçons sur l'intégration et la recherche des fonctions primitives*, Gauthier-Villars, 1904.
- [22] P.G. Ciarlet, *Introduction à l'analyse numérique matricielle et à l'optimisation*, Masson.
- [23] P. Lascaux et R. Théodor, *Analyse numérique matricielle appliquée à l'art de l'ingénieur*, Masson.
- [24] Y. Saad, *Iterative Methods for Sparse Linear Systems*, 2nd edition, SIAM, 2003.
- [25] S. Mallat, *Une exploration des signaux en ondelettes*, Éditions de l'École Polytechnique, 2000.
- [26] E.M. Stein et R. Shakarchi, *Fourier Analysis : An Introduction*, Princeton University Press, 2003. (Voir Chapitre 4).