

Méthodes mathématiques pour la physique

Yaël Moussouni*

*Observatoire astronomique de Strasbourg, UMR 7550,
CNRS/Université de Strasbourg, 11 rue de l'Université, 67000 Strasbourg, France*

Révisé et augmenté par Loris Delafosse

*Laboratoire de Chimie Quantique de Strasbourg, UMR 7177 - Institute of Chemistry,
CNRS/University of Strasbourg, 4 rue Blaise Pascal, 67000 Strasbourg, France and
Quantum dynamics of nano-objects, Département d'Optique ultrarapide et Nanophotonique,
UMR 7504 - IPCMS, CNRS/University of Strasbourg,
23 rue du Loess, 67034 Strasbourg Cedex 2 BP 43, France*

Notes inspirées par le cours de méthodes mathématiques pour la physique (L3S5) de Hervé Molique (MCF) et Eric Baussan (MCF) en 2022.

Licence: En soumettant ces notes de cours aux *Strasbourg Students Physical Letters*, les auteurs certifient qu'ils ont obtenu l'accord de l'enseignant, et acceptent eux-mêmes qu'elles soit distribuées sous la licence CC BY-SA 4.0 (Distribution libre - Attribution - Partage sous les mêmes conditions) sur le site des 2SPL (www.2spl.fr).

Avertissement: Ces notes n'ont pas été corrigées par l'enseignant et peuvent contenir des erreurs ou dévier du contenu du cours. L'enseignant ne peut être tenu pour responsable d'éventuelles erreurs ou inexactitudes contenues dans ces notes.

CONTENTS

Bibliographie	2
I. Espaces de HILBERT	2
A. Anneaux et corps	2
B. Espace vectoriel	3
C. Injection, surjection, bijection et applications linéaires	3
D. Produit scalaire hermitien	3
E. Espace préhilbertien, hermitien, euclidien	4
F. Inégalités de CAUCHY-SCHWARZ et triangulaire	4
G. Métrique et distance	5
H. Norme et espace vectoriel normé	5
I. Orthogonalité	6
1. Définition	6
2. Cas de la dimension finie	6
3. Cadre général en dimension infinie	6
II. Opérateurs	8
A. Définition générale	8
B. Opérations sur les opérateurs	8
C. Opérateurs particuliers	9
D. Projecteur	10
E. Valeur propre et vecteur propre d'un opérateur	10
1. Cadre général	10
2. Cas de la dimension infinie	12
III. Distributions	12
A. Espace des fonctions tests et espace dual	12
B. Distributions particulières	13
C. Opérations sur les distributions	13

* yael.moussouni@astro.unistra.fr

D. Suites et séries de distribution, convergence	14
E. Produit de convolution de distributions	15
F. Valeur principale de CAUCHY d'une intégrale	15
IV. Séries de FOURIER et transformées de FOURIER	16
A. Séries de FOURIER	16
B. Développements en séries de FOURIER	17
C. Transformées de FOURIER de fonctions	17
D. Transformées de FOURIER de distributions	18
V. Probabilité	19
A. Éléments mathématiques	19
B. Probabilités	20
1. Définition générale	20
2. Probabilités conditionnelles	20
3. Indépendance stochastique	21
4. Loi binomiale	21
5. Formule de BAYES	22
C. Variables aléatoires	23
1. Caractéristiques numériques	23
2. Les fonctions de répartition, génératrices et caractéristiques	25
3. Les lois classiques	25
D. Systèmes de variables aléatoires	28
E. Fonctions de variables aléatoires	30
F. Limites de probabilités	32

BIBLIOGRAPHIE

- W. Appel, *Mathématiques pour la physique et les physiciens*, 5e ed. (H & K, 2017).
- C. Cohen-Tannoudji, B. Diu, F. Laloë, *Mécanique quantique*, Nouvelle ed., Vol. I (EDP Sciences, 2018).
- L. Delafosse, A construction of the quantum state space, Stras. Stud. Phys. Lett. **1**, 0202 (2025).

I. ESPACES DE HILBERT

A. Anneaux et corps

Définition I.1 (anneaux). Soit E un ensemble d'éléments $\{x_1, x_2, \dots\}$. On suppose qu'à tout couple d'éléments de E on peut associer un troisième élément de E de deux manières différentes, l'addition $+$ et la multiplication $\cdot$ qui satisfont les propriétés suivantes :

- $\forall (x_1, x_2, x_3) \in E^3$, $x_1 + (x_2 + x_3) = (x_1 + x_2) + x_3$ (associativité de l'addition) ;
- $\forall (x_1, x_2) \in E^2$, $x_1 + x_2 = x_2 + x_1$ (commutativité de l'addition) ;
- $\exists 0 \in E$, $\forall x \in E$, $x + 0 = x$ (il existe un élément neutre 0 pour l'addition) ;
- $\forall x \in E$, $\exists (-x) \in E$, $x + (-x) = 0$ (il existe un opposé $-x$ de x) ;
- $\forall (x_1, x_2, x_3) \in E^3$, $x_1 \cdot (x_2 \cdot x_3) = (x_1 \cdot x_2) \cdot x_3$ (associativité de la multiplication) ;
- $\forall (x_1, x_2, x_3) \in E^3$, $x_1 \cdot (x_2 + x_3) = x_1 \cdot x_2 + x_1 \cdot x_3$ et $(x_1 + x_2) \cdot x_3 = x_1 \cdot x_3 + x_2 \cdot x_3$ (distributivités gauche et droite de la multiplication par rapport à l'addition).

Alors E est un **anneau**.

Remarque I.2. Si l'anneau vérifie de plus la propriété $x_1 \cdot x_2 = x_2 \cdot x_1$, alors il est dit **commutatif**.

On fixe $(x_1, x_2) \in E^2$, alors l'équation $\forall x \in E$, $x_1 \cdot x = x_2$ n'a pas forcément de solution dans le cas général.

Définition I.3 (corps). Soit $\mathbf{K}$ un anneau commutatif vérifiant :

$$\forall x_1 \in \mathbf{K}, \forall x_2 \in \mathbf{K} \setminus \{0\}, \exists x \in \mathbf{K}, x_2 \cdot x = x_1 \quad (1)$$

alors $\mathbf{K}$ est un **corps** et x est appelé le quotient de x_1 par x_2 .

Remarque I.4. L'élément unité du corps $e \in \mathbf{K}$ vérifie $\forall x \in \mathbf{K}$, $x \cdot e = x$. L'élément inverse $x^{-1} \in \mathbf{K}$ de $x \in \mathbf{K}$ est défini par $x^{-1} \cdot x = e$.

B. Espace vectoriel

Définition I.5. Soit $\mathbf{K}$ un corps commutatif composé d'éléments $\{m_1, m_2, \dots\}$ appelés **scalaires** et $\mathcal{E}$ un ensemble d'éléments $\{X_1, X_2, \dots\}$. On peut définir l'addition de deux éléments de $\mathcal{E}$, $X_1 \in \mathcal{E}$, $X_2 \in \mathcal{E}$, $X_1 + X_2$ par :

- $\forall (X_1, X_2) \in \mathcal{E}^2$, $X_1 + X_2 \in \mathcal{E}$;
- $\forall (X_1, X_2, X_3) \in \mathcal{E}^3$, $X_1 + (X_2 + X_3) = (X_1 + X_2) + X_3$ (associativité) ;
- $\forall (X_1, X_2) \in \mathcal{E}^2$, $X_1 + X_2 = X_2 + X_1$ (commutativité) ;
- $\exists O_{\mathcal{E}} \in \mathcal{E}$, $\forall X \in \mathcal{E}$, $X + O_{\mathcal{E}} = X$ (existence d'un élément nul de l'ensemble) ;
- $\forall X \in \mathcal{E}$, $\exists (-X) \in \mathcal{E}$, $X + (-X) = 0$ (existence d'un opposé).

On définit la multiplication d'un élément de $\mathcal{E}$ par un scalaire de $\mathbf{K}$, $X \in \mathcal{E}$, $m \in \mathbf{K}$, $m \cdot X$ par :

- $\forall m \in \mathbf{K}$, $\forall X \in \mathcal{E}$, $m \cdot X \in \mathcal{E}$;
- $\forall m \in \mathbf{K}$, $\forall (X_1, X_2) \in \mathcal{E}^2$, $m \cdot (X_1 + X_2) = m \cdot X_1 + m \cdot X_2$ (distributivité par rapport à l'addition vectorielle) ;
- $\forall (m_1, m_2) \in \mathbf{K}^2$, $\forall X \in \mathcal{E}$, $(m_1 + m_2) \cdot X = m_1 \cdot X + m_2 \cdot X$ (distributivité par rapport à l'addition scalaire) ;
- $\forall (m_1, m_2) \in \mathbf{K}$, $\forall X \in \mathcal{E}$, $(m_1 \cdot m_2) \cdot X = m_1 \cdot (m_2 \cdot X)$ (compatibilité des multiplications) ;
- $\forall X \in \mathcal{E}$, $1 \cdot X = X$ ($1 \in \mathbf{K}$) (existence d'un élément neutre).

L'ensemble $\mathcal{E}$ muni des lois $+$ et $\cdot$ est un **espace vectoriel**, et $X \in \mathcal{E}$ est un **vecteur** de cet espace.

C. Injection, surjection, bijection et applications linéaires

Définition I.6. Soient A et B deux ensembles. Par définition :

- une *injection* de A dans B est une fonction telle que tout élément de B possède un ou zero antécédent dans A ;
- une *surjection* de A dans B est une fonction telle que tout élément de B possède au moins un antécédent dans A ;
- une *bijection* de A dans B fait correspondre à tout élément de B un unique antécédent dans A .

Soient $\mathcal{E}$ et $\mathcal{F}$ deux espaces vectoriels sur le corps $\mathbf{K}$.

Définition I.7 (linéarité). Une application $f : \mathcal{E} \rightarrow \mathcal{F}$ est dite *linéaire* si :

$$\forall \lambda \in \mathbf{K}, \forall (\phi, \psi) \in \mathcal{E}^2, f(\phi + \lambda\psi) = f(\phi) + \lambda f(\psi) \quad (2)$$

Définition I.8. Une application linéaire de $\mathcal{E}$ dans $\mathbf{K}$ est une **forme linéaire** sur $\mathcal{E}$.

Définition I.9. L'application linéaire $f : \mathcal{E} \rightarrow \mathcal{F}$ est :

- un *endomorphisme* si $\mathcal{E} = \mathcal{F}$;
- un *isomorphisme* si f est bijective ;
- un *automorphisme* si $\mathcal{E} = \mathcal{F}$ et f est bijective.

D. Produit scalaire hermitien

Définition I.10 (p.s.h.). Une forme hermitienne sur E est une application $\langle \cdot | \cdot \rangle$ de $E^2 \rightarrow \mathbf{C}$ telle que :

- $\forall (\phi, \psi) \in E^2$, $\langle \phi | \psi \rangle = \langle \psi | \phi \rangle^*$;
- $\forall (\phi, \psi, \chi) \in E^3$, $\langle \phi | \psi + \chi \rangle = \langle \phi | \psi \rangle + \langle \phi | \chi \rangle$, $\langle \phi + \chi | \psi \rangle = \langle \phi | \psi \rangle + \langle \chi | \psi \rangle$;
- $\forall \lambda \in \mathbf{C}$, $\forall (\phi, \psi) \in E^2$, $\langle \phi | \lambda\psi \rangle = \lambda \langle \phi | \psi \rangle$ (linéarité en la 2^e variable) ;
- $\forall \lambda \in \mathbf{C}$, $\forall (\phi, \psi) \in E^2$, $\langle \lambda\phi | \psi \rangle = \lambda^* \langle \phi | \psi \rangle$ (anti-linéarité en la 1^{re} variable).

Une forme hermitienne est dite *positive* si $\forall \phi \in E$, $\langle \phi | \phi \rangle \geq 0$. Elle est dite *définie positive* si de plus $\forall \phi \in E$, $\langle \phi | \phi \rangle = 0 \Leftrightarrow \phi = O_E$ l'élément nul de l'ensemble E .

On appelle **produit scalaire hermitien** (p.s.h.) une forme linéaire hermitienne définie positive.

Remarque I.11. On dit que le p.s.h. est *sesquilinéaire à gauche*, car anti-linéaire par rapport à la première variable et linéaire par rapport à la seconde.

1. On note $\forall z \in \mathbf{C}$, z^* est le conjugué complexe de z .

E. Espace préhilbertien, hermitien, euclidien

Définition I.12. *Un espace préhilbertien complexe (resp. réel) est un espace vectoriel complexe (resp. réel) muni d'un produit scalaire hermitien (resp. euclidien). Un espace préhilbertien de dimension finie est un espace hermitien (resp. euclidien).*

F. Inégalités de CAUCHY-SCHWARZ et triangulaire

Théorème I.13. *Soit $\mathcal{E}$ un espace préhilbertien, l'inégalité de CAUCHY-SCHWARZ est :*

$$\forall (\phi, \psi) \in \mathcal{E}^2, |\langle \phi | \psi \rangle|^2 \leq \langle \phi | \phi \rangle \cdot \langle \psi | \psi \rangle \quad (3)$$

Avec la norme euclidienne $\|\cdot\|$ définie $\forall \varphi \in \mathcal{E}, \|\varphi\| = \sqrt{\langle \varphi | \varphi \rangle}$:

$$\forall \varphi \in \mathcal{E}, |\langle \phi | \psi \rangle| \leq \|\phi\| \cdot \|\psi\| \quad (4)$$

Démonstration. Soient $\lambda \in \mathbf{K}$, $(\phi, \psi) \in \mathcal{E}^2$, on calcule le produit scalaire :

$$\begin{aligned} \langle \phi + \lambda\psi | \phi + \lambda\psi \rangle &= \langle \phi | \phi \rangle + \lambda^* \langle \psi | \phi \rangle + \lambda \langle \phi | \psi \rangle + \lambda\lambda^* \langle \psi | \psi \rangle \\ &= |\lambda|^2 \langle \psi | \psi \rangle + (\lambda \langle \phi | \psi \rangle)^* + \lambda \langle \phi | \psi \rangle + \langle \phi | \phi \rangle \\ &= |\lambda|^2 \langle \psi | \psi \rangle + 2 \operatorname{Re}(\lambda \langle \phi | \psi \rangle) + \langle \phi | \phi \rangle \end{aligned} \quad (5)$$

Si $\langle \psi | \psi \rangle = 0$, alors $\psi = 0_{\mathcal{E}}$, et l'inégalité est vérifiée.

On considère $\langle \psi | \psi \rangle \neq 0$. On pose dans ce cas :

$$\lambda \equiv -\frac{\langle \phi | \psi \rangle^*}{\langle \psi | \psi \rangle} \quad (6)$$

On peut alors écrire le produit scalaire précédent :

$$\begin{aligned} \langle \phi + \lambda\psi | \phi + \lambda\psi \rangle &= \left| -\frac{\langle \phi | \psi \rangle^*}{\langle \psi | \psi \rangle} \right|^2 \langle \psi | \psi \rangle + 2 \operatorname{Re} \left(-\frac{\langle \phi | \psi \rangle^*}{\langle \psi | \psi \rangle} \langle \phi | \psi \rangle \right) + \langle \phi | \phi \rangle \\ &= \frac{|\langle \phi | \psi \rangle|^2}{\langle \psi | \psi \rangle} - 2 \frac{|\langle \phi | \psi \rangle|^2}{\langle \psi | \psi \rangle} + \langle \phi | \phi \rangle \frac{\langle \psi | \psi \rangle}{\langle \psi | \psi \rangle} \\ &= \frac{\langle \phi | \phi \rangle \cdot \langle \psi | \psi \rangle - |\langle \phi | \psi \rangle|^2}{\langle \psi | \psi \rangle} \end{aligned} \quad (7)$$

Or, $\langle \phi + \lambda\psi | \phi + \lambda\psi \rangle \geq 0$ donc $\frac{\langle \phi | \phi \rangle \cdot \langle \psi | \psi \rangle - |\langle \phi | \psi \rangle|^2}{\langle \psi | \psi \rangle} \geq 0$. Comme $\langle \psi | \psi \rangle \geq 0$, le numérateur est positif ou nul, donc :

$$|\langle \phi | \psi \rangle|^2 \leq \langle \phi | \phi \rangle \cdot \langle \psi | \psi \rangle \quad (8)$$

On retrouve bien l'inégalité de CAUCHY-SCHWARZ. $\square$

Théorème I.14. *Pour tout couple (ϕ, ψ) de vecteurs d'un espace préhilbertien, on vérifie l'inégalité triangulaire :*

$$\|\phi + \psi\| \leq \|\phi\| + \|\psi\| \quad (9)$$

Démonstration. On peut établir les équivalences suivantes (notées $(*)$) :

$$\begin{aligned} (*) &\Leftrightarrow \langle \phi + \psi | \phi + \psi \rangle^{\frac{1}{2}} \leq \langle \phi | \phi \rangle^{\frac{1}{2}} + \langle \psi | \psi \rangle^{\frac{1}{2}} \\ &\Leftrightarrow \langle \phi + \psi | \phi + \psi \rangle \leq (\langle \phi | \phi \rangle^{\frac{1}{2}} + \langle \psi | \psi \rangle^{\frac{1}{2}})^2 \\ &\Leftrightarrow \langle \phi + \psi | \phi + \psi \rangle \leq \langle \phi | \phi \rangle + \langle \psi | \psi \rangle + 2(\langle \phi | \phi \rangle \cdot \langle \psi | \psi \rangle)^{\frac{1}{2}} \end{aligned} \quad (10)$$

On peut écrire le membre de gauche comme :

$$\langle \phi + \psi | \phi + \psi \rangle = \langle \phi | \phi \rangle + \langle \psi | \psi \rangle + \langle \phi | \psi \rangle + \langle \psi | \phi \rangle \quad (11)$$

que l'on peut mettre sous la forme :

$$\langle \phi + \psi | \phi + \psi \rangle = \langle \phi | \phi \rangle + \langle \psi | \psi \rangle + 2 \operatorname{Re}(\langle \phi | \psi \rangle) \quad (12)$$

On a alors l'équivalence :

$$(*) \Leftrightarrow \operatorname{Re}(\langle \phi | \psi \rangle) \leq (\langle \phi | \phi \rangle \cdot \langle \psi | \psi \rangle)^{\frac{1}{2}} \quad (13)$$

Si $\operatorname{Re}(\langle \phi | \psi \rangle) < 0$, le résultat est pédestre.

Si $\operatorname{Re}(\langle \phi | \psi \rangle) \geq 0$, on peut alors écrire l'équivalence :

$$(*) \Leftrightarrow [\operatorname{Re}(\langle \phi | \psi \rangle)]^2 \leq \langle \phi | \phi \rangle \cdot \langle \psi | \psi \rangle \quad (14)$$

Or la valeur absolue de la partie réelle d'un nombre complexe est toujours inférieure au module de ce nombre, donc :

$$[\operatorname{Re}(\langle \phi | \psi \rangle)]^2 \leq |\langle \phi | \psi \rangle|^2 \quad (15)$$

Avec l'inégalité de CAUCHY-SCHWARZ, on a alors :

$$|\langle \phi | \psi \rangle|^2 \leq \langle \phi | \phi \rangle \cdot \langle \psi | \psi \rangle \quad (16)$$

Donc $(*)$ est vérifiée dans tous les cas, on retrouve le théorème par équivalences. $\square$

G. Métrique et distance

Définition I.15. Soit un ensemble E quelconque. Soit d une fonction numérique positive sur E^2 telle que :

- $\forall (x, y) \in E^2, d(x, y) = d(y, x)$ (symétrie) ;
- $\forall (x, y) \in E^2, d(x, y) = 0 \Leftrightarrow x = y$;
- $\forall (x, y, z) \in E^3, d(x, z) \leq d(x, y) + d(y, z)$ (inégalité triangulaire).

On appelle d une **distance** sur E . (E, d) est un **espace métrique**, dont les éléments sont appelés **points**.

Note I.16. Si l'on abandonne la deuxième condition, alors d est un écart plutôt qu'une distance, et E est dit **pseudo-métrique**.

H. Norme et espace vectoriel normé

Définition I.17. Soit un espace vectoriel $\mathcal{E}$ (complexe ou réel) et $\mathbf{K}$ un corps. Soit N une fonction numérique positive sur E telle que :

- $\forall x \in \mathcal{E}, N(x) = 0 \Leftrightarrow x = 0$ (0 est l'élément neutre de $\mathcal{E}$) ;
- $\forall x \in \mathcal{E}, \forall a \in \mathbf{K}, N(ax) = |a|N(x)$;
- $\forall (x, y) \in \mathcal{E}^2, N(x + y) \leq N(x) + N(y)$.

N définit une **norme** sur $\mathcal{E}$. (E, N) est un **espace vectoriel normé**. La norme peut être notée $\|\cdot\|_{\mathcal{E}}$ ou simplement $\|\cdot\|$ quand il n'y a pas d'ambiguïté.

Note I.18. Si l'on abandonne la première condition, N est alors une **semi-norme** sur E .

On peut définir dans un espace vectoriel normé (E, N) une distance d avec :

$$\forall (x, y) \in \mathcal{E}^2, d(x, y) \equiv N(x - y) \quad (17)$$

On dit que deux normes sont *équivalentes* si et seulement si les distances correspondantes le sont.

Pour un espace préhilbertien $\mathcal{E}$, on peut définir la **norme euclidienne** à partir du produit scalaire par :

$$\forall \phi \in \mathcal{E}, \|\phi\| = \sqrt{\langle \phi | \phi \rangle} \quad (18)$$

Tout espace préhilbertien est donc un espace vectoriel normé.

I. Orthogonalité

1. Définition

Soient ϕ et ψ deux éléments d'un espace préhilbertien donné. ($n \in \mathbf{N}$)

- ϕ et ψ sont dits orthogonaux si $\langle \phi | \psi \rangle = 0$;
- une famille d'éléments de cet espace est dite orthogonale si ses éléments sont orthogonaux deux à deux ;
- une famille orthogonale $(\phi_i)_{i \in \llbracket 1, n \rrbracket}$ est dite orthonormée si $\forall i \in \llbracket 1, n \rrbracket, \langle \phi_i | \phi_i \rangle = 1$.

2. Cas de la dimension finie

Soit $\mathbf{R}_n[X]$ l'espace euclidien de dimension $n + 1$ des polynômes de degré inférieur ou égal à $n \in \mathbf{N} : \forall P \in \mathbf{R}_n[X], \deg(P) \leq n$. On considère deux polynômes $(P, Q) \in \mathcal{E}_n^2$:

$$P = \sum_{i=0}^n a_i X^i \quad \text{et} \quad Q = \sum_{i=0}^n b_i X^i \quad (19)$$

On peut définir plusieurs produits scalaires entre ces polynômes, en particulier :

$$\begin{aligned} \langle P | Q \rangle_1 &\equiv \sum_{i=1}^n a_i b_i \\ \langle P | Q \rangle_2 &\equiv \int_{-1}^1 dx P(x) Q(x) \\ \langle P | Q \rangle_3 &\equiv \int_{-\infty}^{+\infty} dx e^{-x^2} P(x) Q(x) \end{aligned} \quad (20)$$

Pour chacun de ces produits scalaires, on peut construire une base orthogonale, en particulier :

$$\begin{aligned} \langle P | Q \rangle_1 &\longrightarrow (1, x, x^2, \dots, x^n) \\ \langle P | Q \rangle_2 &\longrightarrow \text{polynômes de LEGENDRE} \\ \langle P | Q \rangle_3 &\longrightarrow \text{polynômes d'HERMITE} \end{aligned} \quad (21)$$

De manière plus générale, un système de polynômes est dit orthogonal sur l'intervalle $[a, b]$ par rapport au poids W lorsque :

$$\forall (m, n) \in \llbracket 1, n \rrbracket^2, m \neq n \Rightarrow \int_a^b dx W(x) P_m(x) P_n(x) = 0 \quad (22)$$

Les données relatives à un certain nombre de ces polynômes régulièrement utilisés peuvent être trouvées dans la littérature, en particulier le *Handbook of Mathematical Functions* de M. ABRAMOWITZ et I. A. STEGUN . On peut entre autre y trouver les valeurs des poids w_i et des nœuds x_i des quadratures de GAUSS de la forme :

$$\int_a^b dx f(x) \approx \sum_{i=1}^n w_i f(x_i) \quad (23)$$

3. Cadre général en dimension infinie

On considère ici des familles de polynômes déjà vues, mais où le degré des polynômes n'est pas limité. On construit alors des séries convergentes :

$$P = \sum_i a_i x^i \quad \text{etc.} \quad (24)$$

Définition I.19 (ensemble complet). *Un espace topologique est dit **complet** si et seulement si toute suite de CAUCHY formée d'éléments de cet espace converge vers un élément de ce même espace.*

Définition I.20 (espace de BANACH). *Un espace vectoriel normé complet est appelé **espace de BANACH**.*

Définition I.21 (espace de HILBERT). *Un espace préhilbertien complet pour la norme associée au produit scalaire est appelé **espace de HILBERT**. Il s'agit donc aussi d'un espace de BANACH muni d'un produit scalaire.*

On considère l'intervalle $]a, b[$, $(a, b) \in \overline{\mathbf{R}}^2$ et $W :]a, b[\rightarrow]0, +\infty[$ continue. L'espace des fonctions de carré intégrable pour la mesure $W(x) dx$ noté $L^2(]a, b[, W(x))$ des fonctions de x à valeurs dans $\mathbf{R}$ telles que :

$$\int |f(x)|^2 W(x) dx \quad (25)$$

est un espace de HILBERT pour le produit scalaire :

$$\langle f|g \rangle = \int_a^b W(x) f(x) g(x) dx \quad (26)$$

Définition I.22. *Soit un espace préhilbertien $\mathcal{E}$ donné et $(e_n)_{n>0}$ une suite orthonormée de vecteurs. $\forall x \in \mathcal{E}$, on définit les coefficients de FOURIER de x par rapport à la suite orthonormée $(e_n)_{n>0}$:*

$$c_n(x) \equiv \langle e_n | x \rangle \quad (27)$$

On a l'inégalité de BESSEL :

$$\sum_{n=1}^{+\infty} |c_n(x)|^2 \leq \|x\|^2 \quad (28)$$

Théorème I.23. *Il y a équivalence entre :*

- la famille $(e_n)_{n>0}$ est une **base hilbertienne** de l'espace préhilbertien $\mathcal{E}$ (définition) ;
- la somme $\sum_{n=1}^m c_n(x) e_n$ converge vers x au sens de la convergence dans $\mathcal{E}$ définie précédemment ;
- $\forall x \in \mathcal{E}$, $x = \sum_{n=1}^{+\infty} c_n(x) e_n$;
- $\forall x \in \mathcal{E}$, $\sum_{n=1}^{+\infty} |c_n(x)|^2 = \|x\|^2$ (égalité de PARSEVAL²) ;
- $\forall (x, y) \in \mathcal{E}^2$, $\langle x | y \rangle = \sum_{n=1}^{+\infty} \langle x | e_n \rangle \langle e_n | y \rangle$ (relation de fermeture).

Définition I.24. *On considère désormais un espace métrique $\mathcal{E}$. On appelle **boule ouverte** de centre x_0 et de rayon $R > 0$ l'ensemble des points x tels que $d(x, x_0) < R$. Un ensemble A est dit dense dans $\mathcal{E}$ si toute boule ouverte de $\mathcal{E}$ contient un élément de A , i.e. on approche d'aussi près que l'on veut n'importe quel point de $\mathbf{E}$ pour un point de A . $\mathcal{E}$ est dit séparable s'il contient un ensemble dénombrable dense.*

Remarque I.25. *Un espace préhilbertien de dimension finie est automatiquement de HILBERT et séparable. Un espace de HILBERT de dimension infinie est séparable si, et seulement si, il possède une base hilbertienne.*

Définition I.26. *L'espace dual de $\mathcal{E}$ noté $\mathcal{E}^*$ est l'espace des formes linéaires sur $\mathcal{E}$. C'est un espace vectoriel pour les opérations habituelles. Cela signifie :*

$$\forall (\mu, \nu) \in (\mathcal{E}^*)^2, \forall \lambda \in \mathbf{K}, \forall \phi \in \mathcal{E}, [\lambda\mu + \nu](\phi) = \lambda\mu(\phi) + \nu(\phi) \quad (29)$$

Définition I.27. *L'espace dual $\mathcal{E}$ est muni de la norme naturelle $\|\cdot\|_\infty$ telle que :*

$$\forall \mu \in \mathcal{E}^*, \|\mu\|_\infty = \sup_{\|\phi\| \leq 1} |\mu(\phi)| \quad (30)$$

Cette borne supérieure est parfois infinie, et la forme linéaire est dite non normalisable. Les formes linéaires non normalisables sont exactement les formes linéaires non continues sur $\mathcal{H}$.

Si on considère désormais un espace de HILBERT $\mathcal{H}$, on peut faire correspondre à tout vecteur ψ de cet espace une forme linéaire $\psi_\mu \in \mathcal{H}^*$ donnant, pour tout vecteur ϕ de l'espace, le produit scalaire hermitien entre ψ et ϕ , i.e. :

$$\forall \psi \in \mathcal{H}, \exists \psi_\mu \in \mathcal{H}^*, \forall \phi \in \mathcal{H}, \psi_\mu(\phi) = \langle \psi | \phi \rangle \quad (31)$$

En utilisant les notations de mécanique quantique, on a $\psi_\mu = \langle \psi | \in \mathcal{H}^*$. Ainsi, tout vecteur de $\mathcal{H}$ peut être identifié avec un élément du dual $\mathcal{H}^*$. La réciproque est fautive : en dimension infinie, les formes linéaires non normalisables (ou de façon équivalente, non continues) ne sont associées à aucun vecteur de $\mathcal{H}$.

2. À ne pas confondre avec PERCEVAL.

Théorème I.28 (FRÉCHET-RIESZ). Soit μ une forme linéaire continue sur $\mathcal{H}$. Il existe alors un unique vecteur $\psi \in \mathcal{H}$ tel que :

$$\forall \phi \in \mathcal{H}, \mu(\phi) = \langle \psi | \phi \rangle \quad (32)$$

De plus :

$$\|\mu\|_{\infty} = \|\psi\| \quad (33)$$

On dit que tout espace de Hilbert est isométriquement isomorphe à son dual topologique (l'ensemble des formes linéaires continues). Le dual topologique est donc également un espace de HILBERT.

Ce théorème justifie ainsi la notation de DIRAC, dans laquelle un état quantique est représentée indifféremment par un ket (un vecteur) ou par son bra associé (une forme linéaire continue). Les formes linéaires non normalisables sont les **kets généralisés** (deltas de DIRAC, ondes planes...) qui n'appartiennent pas à l'espace de HILBERT. Ref.³ fournit de plus amples détails sur ce point.

II. OPÉRATEURS

A. Définition générale

Définition II.1. Un opérateur linéaire $\hat{O}$ sur un espace vectoriel $\mathcal{H}$ est une transformation $\hat{O} : \mathcal{H} \rightarrow \mathcal{H}$, $\phi \mapsto \phi' = \hat{O}|\phi\rangle$ linéaire, c'est à dire que :

$$\forall (|\phi_1\rangle, |\phi_2\rangle) \in \mathcal{H}, \forall (\lambda_1, \lambda_2) \in \mathbf{K}, \hat{O}(\lambda_1 |\phi_1\rangle + \lambda_2 |\phi_2\rangle) = \lambda_1 \hat{O}|\phi_1\rangle + \lambda_2 \hat{O}|\phi_2\rangle \quad (34)$$

On définit de la même manière un opérateur anti-linéaire $\hat{K}$, vérifiant cette fois :

$$\forall (|\phi_1\rangle, |\phi_2\rangle) \in \mathcal{H}, \forall (\lambda_1, \lambda_2) \in \mathbf{K}, \hat{K}(\lambda_1 |\phi_1\rangle + \lambda_2 |\phi_2\rangle) = \lambda_1^* \hat{K}|\phi_1\rangle + \lambda_2^* \hat{K}|\phi_2\rangle \quad (35)$$

Définition II.2. $\hat{O}$ un opérateur sur $\mathcal{H}$ est dit **borné** si :

$$\exists M \in \mathbf{K}, \forall |\phi\rangle \in \mathcal{H}, \|\hat{O}|\phi\rangle\| \leq M\|\phi\| \quad (36)$$

La **norme** de $\hat{O}$ est la plus petite valeur de M vérifiant la condition précédente.

B. Opérations sur les opérateurs

Définition II.3. Soient $\hat{O}_1$, $\hat{O}_2$ et $\hat{O}$ des opérateurs d'un espace de HILBERT $\mathcal{H}$. On définit la somme de deux opérateurs :

$$\forall |\phi\rangle \in \mathcal{H}, [\hat{O}_1 + \hat{O}_2]|\phi\rangle = \hat{O}_1|\phi\rangle + \hat{O}_2|\phi\rangle \quad (37)$$

De même, on définit le produit d'un opérateur par un scalaire par :

$$\forall |\phi\rangle \in \mathcal{H}, [\lambda \hat{O}]|\phi\rangle = \lambda \hat{O}|\phi\rangle \quad (38)$$

On définit la composée de deux opérateurs :

$$\forall |\phi\rangle \in \mathcal{H}, (\hat{O}_1 \hat{O}_2)|\phi\rangle = (\hat{O}_1 \circ \hat{O}_2)|\phi\rangle = \hat{O}_1(\hat{O}_2|\phi\rangle) \quad (39)$$

Le commutateur de deux opérateurs est défini par :

$$[\hat{O}_1, \hat{O}_2] = \hat{O}_1 \hat{O}_2 - \hat{O}_2 \hat{O}_1 \quad (40)$$

3. L. Delafosse, A construction of the quantum state space, *Stras. Stud. Phys. Lett.* **1**, 0202 (2025).

Deux opérateurs commutent si et seulement si $[\hat{O}_1, \hat{O}_2] = 0$. L'anti-commutateur de deux opérateurs est défini par :

$$\{\hat{O}_1, \hat{O}_2\} = \hat{O}_1\hat{O}_2 + \hat{O}_2\hat{O}_1 \quad (41)$$

La puissance d'un opérateur est définie par :

$$\hat{O}^n = \underbrace{\hat{O} \circ \dots \circ \hat{O}}_{n \text{ fois}} \quad (42)$$

L'exponentielle d'un opérateur est définie par :

$$\exp(\hat{O}) = e^{\hat{O}} = \sum_{n=0}^{+\infty} \frac{1}{n!} \hat{O}^n \quad (43)$$

Définition II.4. L'opérateur $\hat{O}$ d'un espace de HILBERT $\mathcal{H}$ est dit inversible si et seulement si il existe son inverse $\hat{O}^{-1}$ tel que :

$$\hat{O}\hat{O}^{-1} = \hat{O}^{-1}\hat{O} = \hat{1} \quad (44)$$

Définition II.5. L'opérateur adjoint $\hat{O}^\dagger$ de $\hat{O}$ est l'opérateur vérifiant :

$$\forall(\phi, \psi) \in \mathcal{H}^2, \quad \langle \phi | \hat{O}^\dagger | \psi \rangle = \langle \psi | \hat{O} | \phi \rangle^* \quad (45)$$

En notation matricielle, on a $\hat{O}^\dagger = O^{T*}$.

Proposition II.6. On a :

- $(\hat{O}^\dagger)^\dagger = \hat{O}$;
- $\forall \lambda \in \mathbf{K}, (\lambda \hat{O})^\dagger = \lambda^* \hat{O}^\dagger$;
- $(\hat{O}_1 + \hat{O}_2)^\dagger = \hat{O}_1^\dagger + \hat{O}_2^\dagger$;
- $(\hat{O}_1 \hat{O}_2)^\dagger = \hat{O}_2^\dagger \hat{O}_1^\dagger$;
- $(|\phi\rangle\langle\psi|)^\dagger = |\psi\rangle\langle\phi|$.

En pratique, $\forall |\psi\rangle \in \mathcal{H}$, avec $|\psi'\rangle = \hat{O} |\psi\rangle$, on a pour le bra associé $\langle\psi'| = \langle\psi| \hat{O}^\dagger$. Ainsi :

$$\forall(\phi, \psi) \in \mathcal{E}^2, \quad \langle \hat{O}^\dagger \phi | \psi \rangle = \langle \phi | \hat{O} \psi \rangle = \langle \phi | \hat{O} | \psi \rangle \quad (46)$$

C. Opérateurs particuliers

Définition II.7. Un opérateur $\hat{A}$ est dit auto-adjoint si :

$$\hat{A}^\dagger = \hat{A} \quad (47)$$

Un opérateur $\hat{B}$ est dit isométrique si :

$$\hat{B}^\dagger \hat{B} = \hat{1} \quad (48)$$

Il est dit co-isométrique si :

$$\hat{B} \hat{B}^\dagger = \hat{1} \quad (49)$$

Un opérateur $\hat{U}$ est dit unitaire s'il est isométrique et co-isométrique, i.e. :

$$\hat{U}^\dagger \hat{U} = \hat{U} \hat{U}^\dagger = \hat{1} \quad \text{ou} \quad \hat{U}^\dagger = \hat{U}^{-1} \quad (50)$$

Un opérateur $\hat{B}$ isométrique conserve le produit scalaire (c'est en particulier le cas pour un opérateur unitaire), i.e. :

$$\langle B\phi | B\psi \rangle = \langle \phi | B^\dagger B | \psi \rangle = \langle \phi | \psi \rangle \quad (51)$$

Définition II.8. Un opérateur $\hat{C}$ est dit anti-unitaire s'il est anti-linéaire et $\hat{C}^\dagger = \hat{C}^{-1}$.

D. Projecteur

Définition II.9. On considère un ensemble de vecteurs orthonormés $\{|\phi_i\rangle\}_{i \in \llbracket 1, N \rrbracket}$, l'opérateur de projection sur le sous-espace lié au N vecteurs $\{|\phi_i\rangle\}_{i \in \llbracket 1, N \rrbracket}$ est :

$$\hat{P} = \sum_{i=1}^N |\phi_i\rangle\langle\phi_i| \quad (52)$$

On a donc $\hat{P}^2 = \hat{P}$

Démonstration.

$$\hat{P}^2 = \sum_{i=1}^n \sum_{j=1}^N |\phi_i\rangle\langle\phi_i|\phi_j\rangle\langle\phi_j| = \sum_{i=1}^N \sum_{j=1}^N |\phi_i\rangle\langle\phi_j| \delta_{ij} = \sum_{i=1}^N |\phi_i\rangle\langle\phi_i| = \hat{P} \quad (53)$$

□

E. Valeur propre et vecteur propre d'un opérateur

1. Cadre général

Définition II.10. Soit $\hat{O}$ un opérateur sur $\mathcal{H}$. Une équation aux valeurs propres est de la forme :

$$\exists \lambda \in \mathbf{K}, \exists |\phi\rangle \in \mathcal{H}, |\phi\rangle \neq 0, \hat{O}|\phi\rangle = \lambda|\phi\rangle \quad (54)$$

avec λ une valeur propre de $\hat{O}$ et $|\phi\rangle$ un vecteur propre de $\hat{O}$ pour la valeur propre λ . L'ensemble des vecteurs propres liés à une valeur propre λ est le sous espace propre de la valeur propre $E_\lambda = \text{Ker}(\lambda\hat{1} - \hat{O})$. Si la dimension de cet espace est 1, alors la valeur propre est simple. Si la dimension de cet espace est $n > 1$, alors la valeur propre est n fois dégénérée. La somme directe des sous-espaces propres correspond à l'espace $\mathcal{H}$. L'ensemble des valeurs propres de l'opérateur est appelé le spectre de cet opérateur $\text{Sp}(\hat{O})$, avec $\dim[\text{Sp}(\hat{O})] \leq \dim(\mathcal{H})$.

On peut citer comme exemple d'équation aux valeurs propres intervenant en mécanique quantique l'équation de SCHRÖDINGER indépendante du temps :

$$\mathcal{H}|\phi_n\rangle = E_n|\phi_n\rangle \quad (55)$$

Théorème II.11. Les valeurs propres d'un opérateur auto-adjoint sont réelles. Ses vecteurs propres correspondant à deux valeurs propres distinctes sont orthogonaux.

Démonstration. On considère $|\phi\rangle \in \mathcal{H}$ vecteur propre de l'opérateur auto-adjoint $\hat{O}$ pour la valeur propre λ . Alors :

$$\langle\phi|\hat{O}|\phi\rangle = \lambda\langle\phi|\phi\rangle = \lambda\|\phi\|^2 \quad (56)$$

et :

$$\langle\phi|\hat{O}|\phi\rangle = \langle\hat{O}^\dagger\phi|\phi\rangle = \langle\phi|\hat{O}^\dagger|\phi\rangle^* = \langle\phi|\hat{O}|\phi\rangle^* = \lambda^*\langle\phi|\phi\rangle^* = \lambda^*\|\phi\|^2 \quad (57)$$

d'où :

$$\lambda = \lambda^* \Rightarrow \lambda \in \mathbf{R} \quad (58)$$

□

Démonstration. On considère Considérons $|\phi\rangle$ et $|\psi\rangle$ deux vecteurs propres de $\hat{O}$ aux valeurs propres λ et μ différentes donc $\hat{O}|\phi\rangle = \lambda|\phi\rangle$ et $\hat{O}|\psi\rangle = \mu|\psi\rangle$. On a :

$$\langle\phi|\hat{O}|\psi\rangle = \mu\langle\phi|\psi\rangle \quad (59)$$

Par ailleurs :

$$\langle \phi | \hat{O} | \psi \rangle = \langle \phi | \hat{O}^\dagger | \psi \rangle = \langle \psi | \hat{O} | \phi \rangle^* = \lambda^* \langle \psi | \phi \rangle^* = \lambda \langle \phi | \psi \rangle \quad (60)$$

On en déduit :

$$(\mu - \lambda) \langle \phi | \psi \rangle = 0 \quad (61)$$

Or, $\mu \neq \lambda$ donc :

$$\langle \phi | \psi \rangle = 0 \quad (62)$$

□

Théorème II.12. *Pour un opérateur unitaire, les valeurs propres sont des nombres complexes de module unité. Les vecteurs propres correspondant à deux valeurs propres différentes sont orthogonaux.*

Démonstration. Soit $|\psi\rangle$ un vecteur propre d'un opérateur unitaire $\hat{U}$ pour la valeur propre u . On a :

$$\langle \psi | \hat{U}^\dagger \hat{U} | \psi \rangle = uu^* \langle \psi | \psi \rangle = uu^* \|u\|^2 \quad (63)$$

D'autre part, comme $\hat{U}^\dagger \hat{U} = \hat{1}$, on a :

$$\langle \psi | \hat{U}^\dagger \hat{U} | \psi \rangle = \langle \psi | \psi \rangle = \|\psi\|^2 \quad (64)$$

d'où $uu^* = \|u\|^2 = 1$ donc les valeurs propres sont biens des nombres complexes de module unité. □

Démonstration. Considérons $|\phi\rangle$ et $|\psi\rangle$ deux vecteurs propres de $\hat{U}$ correspondant aux valeurs μ et ν différentes donc $\hat{U}|\phi\rangle = \mu|\phi\rangle$ et $\hat{U}|\psi\rangle = \nu|\psi\rangle$. On a :

$$\langle \phi | \psi \rangle = \langle \phi | \hat{U}^\dagger \hat{U} | \psi \rangle = \mu^* \nu \langle \phi | \psi \rangle \quad (65)$$

Or $\mu \neq \nu$, par conséquent, $\langle \phi | \psi \rangle = 0$. □

Théorème II.13 (décomposition spectrale). *En notant a_m les n valeurs propres d'un opérateur $\hat{O}$ de multiplicité respectives α_m , la relation de décomposition spectrale de $\hat{O}$ s'écrit :*

$$\hat{O} = \sum_{m=1}^n \sum_{r=1}^{\alpha_m} |m, r\rangle a_m \langle m, r| \quad (66)$$

Démonstration. Soit $\{|m, r\rangle\}$ l'ensemble des vecteurs propres de $\hat{O}$ où $m \in \llbracket 1, n \rrbracket$ et $r \in \llbracket 1, \alpha_m \rrbracket$. Cet ensemble de vecteurs sous-tend complètement l'espace $\mathcal{H}$ dont il est une base. Par conséquent, on pourra écrire la relation de fermeture :

$$\sum_{m=1}^n \sum_{r=1}^{\alpha_m} |m, r\rangle \langle m, r| = \hat{1}_H \quad (67)$$

En introduisant le projecteur P_m sur le sous-espace engendré par le vecteur relatif à la valeur propre a_m :

$$\hat{P}_m = \sum_{r=1}^{\alpha_m} |m, r\rangle \langle m, r| \quad (68)$$

On peut alors écrire la relation de fermeture précédente comme :

$$\sum_{m=1}^n \hat{P}_m = 1_H \quad (69)$$

Ainsi, pour tout vecteur $|\phi\rangle$ de l'espace $\mathcal{H}$, on a :

$$\hat{O}|\phi\rangle = \hat{O} \sum_{m=1}^n \hat{P}_m |\phi\rangle = \sum_{m=1}^n \hat{O} \hat{P}_m |\phi\rangle = \sum_{m=1}^n \hat{O} \sum_{r=1}^{\alpha_m} |m, r\rangle \langle m, r| \phi\rangle \quad (70)$$

Soit :

$$\hat{O}|\phi\rangle = \sum_{m=1}^n \sum_{r=1}^{\alpha_m} \hat{O}|m, r\rangle \langle m, r|\phi\rangle = \sum_{m=1}^n \sum_{r=1}^{\alpha_m} a_m |m, r\rangle \langle m, r|\phi\rangle \quad (71)$$

Ainsi :

$$\hat{O}|\phi\rangle = \sum_{m=1}^n \sum_{r=1}^{\alpha_m} |m, r\rangle a_m \langle m, r|\phi\rangle \quad (72)$$

□

2. Cas de la dimension infinie

On considère un $\mathbf{K}$ -espace vectoriel $\mathcal{E}$, O un endomorphisme de $\mathcal{E}$ et $\lambda \in \mathbf{K}$.

Définition II.14. Si l'application $(\lambda \text{Id}_{\mathcal{E}} - O)$ n'est pas inversible, alors on dit que λ est une valeur spectrale de O . L'ensemble de ces valeurs spectrales forment le spectre de O . Au contraire, si l'application est inversible, on dit que λ est une valeur régulière de O . L'ensemble de ces valeurs régulières forment l'ensemble résolvant de O .

Définition II.15. Si l'application $(\lambda \text{Id}_{\mathcal{E}} - O)$ n'est pas injective, on dit que λ est une valeur propre de O (toute valeur propre est aussi une valeur spectrale). Un vecteur ϕ de $\mathcal{E}$ vérifiant $O(\phi) = \lambda\phi$ est un vecteur propre de O associé à la valeur propre λ , appartenant au spectre ponctuel. Au contraire, si l'application est injective et $\text{Im}(\lambda \text{Id}_{\mathcal{E}} - O)$ est distinct de $\mathcal{E}$, alors on dit que λ appartient au spectre continu.

Remarque II.16. Les deux notions de valeurs propres et valeurs spectrales sont identiques en dimension finie. En dimension infinie, le nombre de valeurs spectrales est supérieur ou égale à celui des valeurs propres.

Proposition II.17. Le spectre d'un opérateur auto-adjoint est réel.

L'équation aux valeurs propres n'admet pas toujours de solutions en dimension infinie.

III. DISTRIBUTIONS

A. Espace des fonctions tests et espace dual

Définition III.1. On appelle fonction test sur $\mathbf{R}$ toute fonction C^∞ de $\mathbf{R}$ dans $\mathbf{C}$ nulle en dehors d'un intervalle fermé borné de $\mathbf{R}$, appelé le support la fonction. L'ensemble des fonctions tests est noté $\mathcal{D}(\mathbf{R})$, ou plus simplement $\mathcal{D}$.

Proposition III.2. L'espace $\mathcal{D}$ est un $\mathbf{C}$ -espace vectoriel.

On considère un ensemble $\mathcal{K}$, un espace métrique $(\mathcal{E}, d)$ et une suite d'applications $(\phi_n)_n$ telle que $\forall n, \phi_n : \mathcal{K} \rightarrow \mathcal{E}$. La suite $(\phi_n)_n$ converge simplement vers ϕ sur $\mathcal{K}$ si $\forall x \in \mathcal{K}$, la suite $(\phi_n(x))_n$ converge vers $\phi(x)$ dans $\mathcal{E}$. Dit autrement, $(\phi_n)_n$ converge simplement vers ϕ si et seulement si :

$$\forall \varepsilon > 0, \forall x \in \mathcal{K}, \exists N_\varepsilon(x) \in \mathbf{N}, \forall n \geq N_\varepsilon(x), d[\phi_n(x), \phi(x)] \leq \varepsilon \quad (73)$$

La suite $(\phi_n)_n$ converge uniformément vers ϕ sur $\mathcal{K}$ si et seulement si :

$$\forall \varepsilon > 0, \exists N_\varepsilon \in \mathbf{N}, \forall n \geq N_\varepsilon, \forall x \in \mathcal{K}, d[\phi_n(x), \phi(x)] \leq \varepsilon \quad (74)$$

La convergence uniforme implique la convergence simple.

Définition III.3. On dit qu'une suite de fonctions tests ϕ_n converge dans $\mathcal{D}$ vers ϕ si :

- il existe un intervalle fermé $\mathcal{K}$ en dehors duquel toutes les fonctions ϕ_n et ϕ s'annulent ;
- la suite $(\phi_n)_n$ converge uniformément vers ϕ sur $\mathcal{K}$;
- toutes les suites des dérivées i -ième $(\phi_n^{(i)})_n$ convergent uniformément vers $\phi^{(i)}$ sur $\mathcal{K}$.

On note :

$$\phi_n \xrightarrow[n \rightarrow +\infty]{\mathcal{D}} \phi \quad (75)$$

Définition III.4. Un espace fonctionnel $\mathcal{F}$ est un ensemble de fonctions possédant une structure d'espace vectoriel.

Définition III.5. Une distribution T est une fonctionnelle linéaire et continue sur $\mathcal{D}$.

— Fonctionnelle linéaire veut dire qu'il s'agit d'une application T de $\mathcal{D}$ dans $\mathbf{K}$, qui à une fonction test $\phi \in \mathcal{D}$ fait correspondre un nombre noté $\langle T, \phi \rangle$ tel que :

$$\forall(a, b) \in \mathbf{K}^2, \forall(\phi, \psi) \in \mathcal{D}^2, \langle T, a\phi + b\psi \rangle = a \langle T, \phi \rangle + b \langle T, \psi \rangle \quad (76)$$

— continue veut dire que si une suite $(\phi_n)_n$ converge vers ϕ dans $\mathcal{D}$, alors la suite $(\langle T, \phi_n \rangle)_n$ converge vers $\langle T, \phi \rangle$;
— les distributions constituent donc le dual topologique de $\mathcal{D}$, noté $\mathcal{D}'(\mathbf{R})$ ou simplement $\mathcal{D}'$.

L'espace $\mathcal{D}'$ constitue un espace vectoriel avec les lois :

$$\begin{aligned} \forall(T, T_1, T_2) \in \mathcal{D}'^3, \forall\phi \in \mathcal{D}, \forall\lambda \in \mathbf{R} : \\ \langle T_1 + T_2, \phi \rangle &= \langle T_1, \phi \rangle + \langle T_2, \phi \rangle, \\ \langle \lambda T, \phi \rangle &= \lambda \langle T, \phi \rangle \end{aligned} \quad (77)$$

Remarque III.6. T est la distribution nulle si et seulement si $\forall\phi \in \mathcal{D}, \langle T, \phi \rangle = 0$.

Définition III.7. Une fonction est localement sommable sur $\mathbf{R}$ si $\int_a^b |f(x)| dx$ existe quelque soit l'intervalle fini $[a, b]$. On note $\mathcal{L}_{\text{loc}}^1$ l'ensemble des fonction localement sommables sur $\mathbf{R}$.

B. Distributions particulières

Définition III.8. Une distribution régulière associée à une fonction localement sommable f notée T_f ou $[f]$ est la distribution telle que :

$$\langle [f], \phi \rangle = \int_{-\infty}^{+\infty} f(x)\phi(x) dx \quad (78)$$

Définition III.9. Une distribution singulière est une distribution qui n'est pas régulière.

Définition III.10. La distribution de DIRAC est définie telle que :

$$\forall\phi \in \mathcal{D}, \langle \delta, \phi \rangle = \phi(0) \quad (79)$$

Cette distribution peut également être translatée en a , alors :

$$\forall\phi \in \mathcal{D}, \langle \delta_a, \phi \rangle = \phi(a) \quad (80)$$

C'est une distribution singulière.

Remarque III.11. La théorie des distributions sert donc à généraliser la notion de "fonction de carré intégrable" (les fonctions telles que l'intégrale (78) existe et constitue le produit scalaire) à des objets comme le delta de DIRAC. Par analogie avec ces fonctions, on invente le symbole delta de DIRAC $\delta(x)$, qu'on manipule comme une fonction et qui nous permet d'écrire :

$$\langle \delta_a, \phi \rangle = \int_{-\infty}^{+\infty} \delta(x)\phi(x) dx = \phi(a) \quad (81)$$

Ainsi, même si ce symbole ne correspond à aucune véritable fonction, son produit scalaire avec une fonction existe car il s'agit de l'application de la distribution de DIRAC à cette fonction, une opération bien définie.

C. Opérations sur les distributions

Définition III.12.

— On définit la translatée T_a d'une distribution T par :

$$\forall\phi \in \mathcal{D}, \langle T_a, \phi \rangle = \langle T, \phi_{-a} \rangle \quad (82)$$

où $\phi_{-a}(x) = \phi(x + a)$. Une distribution est dite périodique de période a si $T = T_a$.

— On définit la dilatée ${}_kT$ d'une distribution T par :

$$\forall \phi \in \mathcal{D}, \langle {}_kT, \phi \rangle = \frac{1}{|k|} \langle T, {}_{1/k}\phi \rangle \quad (83)$$

avec $k \neq 0$ et ${}_{1/k}\phi(x) = \phi(\frac{x}{k})$. Une distribution est dite paire si ${}_{-1}T = T$ et impaire si ${}_{-1}T = -T$.

— Le produit de deux distributions n'est pas défini de manière générale, mais on définit le produit d'une distribution par une fonction $g \in \mathcal{C}^\infty$ par :

$$\langle gT, \phi \rangle = \langle T, g\phi \rangle \quad (84)$$

Définition III.13. On définit la dérivée T' de la distribution T par :

$$\forall \phi \in \mathcal{D}, \langle T', \phi \rangle = -\langle T, \phi' \rangle \quad (85)$$

Toutes les distributions sont indéfiniment dérivables, et les dérivées n -ièmes sont données par :

$$\forall \phi \in \mathcal{D}, \langle T^{(n)}, \phi \rangle = (-1)^n \langle T, \phi^{(n)} \rangle \quad (86)$$

Remarque III.14. Cette définition fait sens dans la mesure où on voit $\langle T, \phi \rangle$ comme la généralisation du produit scalaire L^2 entre T et ϕ (cf. remarque III.11). Equation (85) est alors une formule d'intégration par parties, où le terme complètement intégré est nul car la fonction test ϕ est nulle à l'infini.

On a les propriétés suivantes :

- $\forall (T, S) \in \mathcal{D}'^2, (T + S)' = T' + S'$;
- $\forall \lambda \in \mathbf{R}, (\lambda T)' = \lambda T'$;
- $\forall g \in \mathcal{C}^\infty, (gT)' = g'T + gT'$;
- $(T_a)' = (T')_a$;
- $({}_aT)' = {}_a(T')$.

Définition III.15. On appelle $[f]'$ la dérivée de f au sens des distributions, la dérivée de la distribution régulière $[f]$ bâtie sur une fonction localement sommable f .

Cette définition n'est pas à confondre avec $[f']$, la distribution régulière de la dérivée de f . Si f est dérivable et f' est localement sommable, alors f' est aussi une distribution régulière et $[f]' = [f']$. Il est tout à fait possible que f ne soit pas dérivable au sens des fonctions mais le soit au sens des distributions.

Définition III.16. Une fonction f est de classe $\mathcal{C}^n$ sur un intervalle $I \subset \mathbf{R}$ (noté $f \in \mathcal{C}^n(I, \mathbf{R})$) si elle est n fois dérivable et sa dérivée n -ième est continue sur I .

Théorème III.17 (formule des sauts). Soit f définie sur $\mathbf{R}$ de classe $\mathcal{C}^1$ sauf en n points a_i où elle présente des discontinuités de première espèce (i.e. la fonction admet des limites finies à gauche et à droite de ces points), d'amplitude $\Delta f(a_i) = f(a_i^+) - f(a_i^-)$. Alors :

$$[f]' = [\tilde{f}'] + \sum_{i=1}^n \Delta f(a_i) \delta_{a_i} \quad (87)$$

où $\tilde{f}'$ représente la dérivée de f là où elle existe.

D. Suites et séries de distribution, convergence

Définition III.18. Une suite de distribution $(T_n)_n$ est dite converger vers la distribution $T \in \mathcal{D}'$ si, $\forall \phi \in \mathcal{D}$, la suite $(\langle T_n, \phi \rangle)_n$ converge vers $\langle T, \phi \rangle$, i.e.

$$T \xrightarrow[n \rightarrow +\infty]{\mathcal{D}'} T \Leftrightarrow \forall \phi \in \mathcal{D}, \langle T_n, \phi \rangle \xrightarrow[n \rightarrow +\infty]{} \langle T, \phi \rangle \quad (88)$$

Si une suite de distribution $(T_n)_n$ converge vers la distribution T dans $\mathcal{D}'$, alors la suite $(T'_n)_n$ (resp. $(T_n^{(p)})_n$) converge vers T' (resp. $T^{(p)}$) dans $\mathcal{D}'$. Attention cependant, si une suite de fonctions localement sommables $(f_n)_n$ converge vers une fonction localement sommable f , on ne peut pas conclure que la suite de distributions $([f_n])_n$ converge vers $[f]$. Cependant, on peut montrer ce résultat si il existe aussi g localement sommable telle que $\forall n \in \mathbf{N}, \forall x \in \mathbf{R}, f_n(x) \leq g(x)$, ou si la suite de fonction $(f_n)_n$ converge uniformément vers f .

Définition III.19. Soit $(T_n)_n$ une suite de distributions donnée. La série $\sum_n T_n$ est dite converger vers T si la suite des sommes partielles $(S_N)_N$, définie par la relation $\forall N \in \mathbf{N}, S_N = \sum_{n=0}^N T_n$, converge vers T dans $\mathcal{D}'$.

Toute série de distribution convergente peut être dérivée terme à terme.

Théorème III.20. Si f est une fonction localement sommable telle qu'elle vérifie la normalisation $\int_{-\infty}^{+\infty} f(x) dx = 1$, alors la suite de distributions $([f_n])_n$ (resp. $([\tilde{f}_\varepsilon])_\varepsilon$) associée aux fonctions $f_n(x) = nf(nx)$ (reps. $\tilde{f}_\varepsilon = \frac{1}{\varepsilon} f(\frac{x}{\varepsilon})$) converge vers la distribution de DIRAC lorsque $n \rightarrow \infty$ (resp. $\varepsilon \rightarrow 0^+$) :

$$\begin{aligned} \lim_{n \rightarrow +\infty} [f_n] &= \lim_{n \rightarrow +\infty} [nf(nx)] = \delta \\ \lim_{\varepsilon \rightarrow 0^+} [\tilde{f}_\varepsilon] &= \lim_{\varepsilon \rightarrow 0^+} \left[\frac{1}{\varepsilon} \tilde{f}\left(\frac{x}{\varepsilon}\right) \right] = \delta \end{aligned} \quad (89)$$

E. Produit de convolution de distributions

Définition III.21. On appelle produit de convolution de deux distributions $T = T_{(x)}$ et $S = S_{(y)}$ la distribution $T * S$ telle que, si elle existe :

$$\forall \phi \in \mathcal{D}, \langle T * S, \phi \rangle = \langle T_{(x)}, \langle S_{(y)}, \phi(x+y) \rangle \rangle \quad (90)$$

Les distributions T et S vérifient les propriétés suivantes :

- si $T * S$ existe, alors $T * S = S * T$;
- $\forall a \in \mathbf{R}, T * \delta_a = T_a$, en particulier $T * \delta = T$;
- $\forall (a, b) \in \mathbf{R}^2, \delta_a * \delta_b = \delta_{a+b}$;
- $\forall n \in \mathbf{N}, \delta^{(n)} * T = T^{(n)}$, en particulier $\delta' * T = T'$;
- si $T * S$ existe, alors $(T * S)' = T' * S = T * S'$;
- $\forall a \in \mathbf{R}, (T * S)_a = T * S_a$.

Le produit de convolution n'est pas associatif i.e. en général :

$$(T, S, R) \in \mathcal{D}'^3, T * (S * R) \neq (T * S) * R \quad (91)$$

F. Valeur principale de CAUCHY d'une intégrale

Définition III.22. On définit la valeur principale de CAUCHY de l'intégrale d'une fonction f présentant une discontinuité en $c \in [a, b] \subset \mathbf{R}$ par :

$$\text{VP} \int_a^b f(x) dx \equiv \lim_{\varepsilon \rightarrow 0} \left\{ \int_a^{c-\varepsilon} f(x) dx + \int_{c+\varepsilon}^b f(x) dx \right\} \quad (92)$$

Si la fonction f est continue sur $\mathbf{R}$ mais que son intégrale ne converge pas sur $\mathbf{R}$, on définit la valeur principale de CAUCHY par :

$$\text{VP} \int_{-\infty}^{+\infty} f(x) dx \equiv \lim_{t \rightarrow \infty} \int_{-t}^{+t} f(x) dx \quad (93)$$

Définition III.23. On définit la distribution « valeur principale de $\frac{1}{x}$ » $\text{VP} \left(\frac{1}{x} \right)$ par :

$$\forall \phi \in \mathcal{D}, \left\langle \text{VP} \left(\frac{1}{x} \right), \phi \right\rangle = \text{VP} \int_{-\infty}^{+\infty} \frac{\phi(x)}{x} dx = \lim_{\varepsilon \rightarrow 0} \left\{ \int_{-\infty}^{-\varepsilon} \frac{\phi}{x} dx + \int_{+\varepsilon}^{+\infty} \frac{\phi}{x} dx \right\} \quad (94)$$

On a $[\ln |x|]' = \text{VP} \left(\frac{1}{x} \right)$ (la démonstration de cette propriété n'est pas faite ici).

IV. SÉRIES DE FOURIER ET TRANSFORMÉES DE FOURIER

A. Séries de FOURIER

Définition IV.1. On note $\text{CM}_T(\mathbf{R}, \mathbf{K})$ l'ensemble des fonctions périodiques de période T continues par morceaux sur chaque période. L'intégrale sur une période de la fonction $f \in \text{CM}_T(\mathbf{R}, \mathbf{K})$ vérifie ($a \in \mathbf{R}$) :

$$\int_a^{a+T} f(t) dt = \int_0^T f(t) dt = \int_{-T/2}^{+T/2} f(t) dt \equiv \int_T f(t) dt \quad (95)$$

Définition IV.2. Soit $f \in \text{CM}_T(\mathbf{R}, \mathbf{K})$, on appelle coefficients de FOURIER trigonométriques de f les scalaires :

$$\begin{aligned} \forall n \in \mathbf{N}, a_n(f) &= \frac{2}{T} \int_T f(t) \cos(n\omega t) dt \\ b_n(f) &= \frac{2}{T} \int_T f(t) \sin(n\omega t) dt \end{aligned} \quad (96)$$

avec n'importe quelle intégrale sur une période et $\omega = \frac{2\pi}{T}$ la pulsation de la fonction périodique.

Si la fonction f est impaire, alors $\forall n \in \mathbf{N}, a_n(f) = 0$. Si cette fonction f est paire, alors $\forall n \in \mathbf{N}, b_n(f) = 0$.

Définition IV.3. Soit une fonction $f \in \text{CM}_T(\mathbf{R}, \mathbf{K})$, on appelle coefficients de FOURIER exponentiels de f les scalaires :

$$\forall n \in \mathbf{Z}, c_n(f) = \frac{1}{T} \int_T f(t) e^{-in\omega t} dt \quad (97)$$

On peut généralement noter sans ambiguïté $a_n = a_n(f)$, $b_n = b_n(f)$ et $c_n = c_n(f)$. On peut relier les deux catégories de coefficient de FOURIER avec les relations :

$$\begin{aligned} c_0 &= a_0 ; \forall n \geq 1, \begin{cases} c_n = \frac{a_n - ib_n}{2} \\ c_{-n} = \frac{a_n + ib_n}{2} \end{cases} \\ a_0 &= c_0 ; \forall n \geq 1, \begin{cases} a_n = c_n + c_{-n} \\ b_n = i(c_n - c_{-n}) \end{cases} \end{aligned} \quad (98)$$

Définition IV.4. Soit $f \in \text{CM}_T(\mathbf{R}, \mathbf{K})$:

— On appelle série de FOURIER de f au point t la série numérique :

$$\begin{aligned} S_F[f](x) &= a_0(f) + \sum_{n \geq 1} (a_n(f) \cos(n\omega t) + b_n(f) \sin(n\omega t)) \\ &= \sum_{n \in \mathbf{N}} (a_n(f) \cos(n\omega t) + b_n(f) \sin(n\omega t)) \\ &= \sum_{n \in \mathbf{Z}} c_n(f) e^{in\omega t} \end{aligned} \quad (99)$$

— On appelle polynôme de FOURIER de f au point t d'indice p la somme partielle d'ordre p de cette série :

$$\begin{aligned} S_p[f](t) &= a_0(f) + \sum_{n=1}^p (a_n(f) \cos(n\omega t) + b_n(f) \sin(n\omega t)) \\ &= \sum_{n=0}^p (a_n(f) \cos(n\omega t) + b_n(f) \sin(n\omega t)) \\ &= \sum_{n=-\infty}^p c_n(f) e^{in\omega t} \end{aligned} \quad (100)$$

Théorème IV.5 (formule de PARSEVAL⁴). Soit une fonction $f \in \text{CM}_T(\mathbf{R}, \mathbf{K})$, les séries $\sum_n |a_n|^2$, $\sum_n |b_n|^2$ et $\sum_n |c_n|^2$ convergent et :

$$\begin{aligned} \frac{1}{T} \int_T |f(t)|^2 dt &= |a_0|^2 + \frac{1}{2} \sum_{n=1}^{+\infty} (|a_n|^2 + |b_n|^2) \\ &= \sum_{n=-\infty}^{+\infty} |c_n|^2 \end{aligned} \quad (101)$$

La convergence des séries implique quelque soit la fonction considérée :

$$\lim_{n \rightarrow +\infty} a_n = 0 \quad ; \quad \lim_{n \rightarrow +\infty} b_n = 0 \quad ; \quad \lim_{n \rightarrow \pm\infty} c_n = 0 \quad (102)$$

B. Développements en séries de FOURIER

Soit $f \in \text{CM}_T(\mathbf{R}, \mathbf{K})$, on appelle régularisée de DIRICHLET de f la fonction $\tilde{f}$ définie par :

$$\begin{aligned} \tilde{f} : \mathbf{R} &\rightarrow \mathbf{K} \\ t &\mapsto \frac{f(t^+) + f(t^-)}{2} \end{aligned} \quad (103)$$

Théorème IV.6 (DIRICHLET). Soit f de classe $\mathcal{C}^1$ par morceaux et de période $T = \frac{2\pi}{\omega}$, la série de FOURIER de f converge en tout point t vers $\tilde{f}(t)$. On a alors la synthèse harmonique :

$$\begin{aligned} \forall t \in \mathbf{R}, \quad \tilde{f}(t) &= a_0 + \sum_{n=1}^{+\infty} (a_n \cos(n\omega t) + b_n \sin(n\omega t)) \\ &= \sum_{n=-\infty}^{+\infty} c_n e^{in\omega t} \end{aligned} \quad (104)$$

C. Transformées de FOURIER de fonctions

Définition IV.7. Soit une fonction f complexe de la variable réelle x . Si elle existe, on définit la transformée de FOURIER de f par la relation :

$$\text{TF}[f](k) = \mathcal{F}[f](k) = \hat{f}(k) \equiv \int_{-\infty}^{+\infty} f(x) e^{+ikx} dx \quad (105)$$

On note $\mathcal{L}^1(\mathbf{R})$ l'ensemble des fonctions définies sur $\mathbf{R}$ à valeurs dans $\mathbf{C}$ telles que l'intégrale $\int_{-\infty}^{+\infty} |f(x)| dx$ converge. Les fonctions de cet ensemble admettent une transformée de FOURIER. On définit de la même manière la transformée inverse de g :

$$\text{TF}^{-1}[g](x) = \mathcal{F}^{-1}[g](x) \equiv \frac{1}{2\pi} \int_{-\infty}^{+\infty} g(k) e^{-ikx} dk \quad (106)$$

On a alors la relation :

$$\text{TF}^{-1}[\text{TF}[f]] = f \quad (107)$$

Ces définitions peuvent varier selon le contexte, en particulier en mécanique quantique, où un facteur $\sqrt{2\pi\hbar}$ est en général ajouté.

4. Encore une fois, à ne pas confondre avec PERCEVAL.

Théorème IV.8 (RIEMAN-LEBESGUE). *Si une fonction $f \in \mathcal{L}^1(\mathbf{R})$, alors la fonction $\text{TF}[f]$ est continue et bornée sur $\mathbf{R}$, et $\lim_{k \rightarrow +\infty} \text{TF}[f](k) = 0$.*

Pour deux fonctions $(f, g) \in \mathcal{L}^1(\mathbf{R})$ et deux scalaires $(\lambda, \mu) \in \mathbf{C}$, on a :

$$\text{TF}[\lambda f + \mu g] = \lambda \text{TF}[f] + \mu \text{TF}[g] \quad ; \quad \text{TF}[\bar{f}(x)] = \overline{\text{TF}[f(-x)]} \quad (108)$$

où $\bar{z}$ désigne le conjugué complexe de z . On a aussi $\forall a \in \mathbf{R}$:

$$\text{TF}[f_a] = e^{+ika} \text{TF}[f] \quad ; \quad \text{TF}[{}_af](k) = \frac{1}{|a|} \text{TF}[f]\left(\frac{k}{a}\right) \quad (109)$$

Si $\forall x \in \mathbf{R}$, $g(x) = e^{-ik_0x} f(x)$, alors :

$$\text{TF}[g](k) = \text{TF}[f](k - k_0) \quad (110)$$

On a pour la dérivation :

$$\text{TF}\left[\frac{df}{dx}\right](k) = -ik \text{TF}[f] \quad (111)$$

et :

$$\frac{d^m \text{TF}[f]}{dk^m} = \text{TF}[(ix)^m f](k) \Rightarrow \frac{d \text{TF}[f]}{dk} = \text{TF}[ixf](k) \quad (112)$$

Définition IV.9. *Le produit de convolution de deux fonctions $(f, g) \in \mathcal{L}^1(\mathbf{R})$ est défini par :*

$$(f * g)(x) = \int_{-\infty}^{+\infty} f(t)g(x-t) dt \quad (113)$$

On a :

$$\text{TF}[f * g] = \text{TF}[f] \times \text{TF}[g] \quad ; \quad \text{TF}[f \times g] = \text{TF}[f] * \text{TF}[g] \quad (114)$$

On peut établir la formule de PARSEVAL-PLANCHEREL (si les intégrales existent) :

$$\int_{-\infty}^{+\infty} f(x)\bar{g}(x) dx = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \text{TF}[f](k)\overline{\text{TF}[g](k)} dk \quad (115)$$

Ainsi :

$$\int_{-\infty}^{+\infty} |f(x)|^2 dx = \frac{1}{2\pi} \int_{-\infty}^{+\infty} |\text{TF}[f](k)|^2 dk \quad (116)$$

D. Transformées de FOURIER de distributions

Définition IV.10. *On appelle fonction déclinante (ou fonction de SCHWARTZ) toute fonction f de classe $\mathcal{C}^\infty$ qui décroît plus vite que polynomialement à l'infini et dont toutes les dérivées décroissent plus vite que polynomialement, i.e. :*

$$\forall n \in \mathbf{N}, \forall p \in \mathbf{N}, \lim_{x \rightarrow \pm\infty} x^n f^{(p)}(x) = 0 \quad (117)$$

L'espace des fonctions déclinantes est appelé espace de SCHWARTZ et est noté $\mathcal{S}$.

- Si $f \in \mathcal{S}$, alors $\forall p \in \mathbf{N}$, $f^{(p)} \in \mathcal{S}$.
- L'ensemble $\mathcal{S}$ est stable par transformée de FOURIER, i.e. :

$$\forall f \in \mathcal{S}, \text{TF}[f] \in \mathcal{S} \quad (118)$$

et on a la formule d'inversion :

$$\forall f \in \mathcal{S}, \text{TF}[\text{TF}^{-1}[f]] = \text{TF}^{-1}[\text{TF}[f]] = f \quad (119)$$

En fait, la transformée de FOURIER est un autohoméomorphisme de $\mathcal{S}$ (une bijection continue de bijection réciproque continue de $\mathcal{S}$ dans lui-même).

— Si $(f, g) \in \mathcal{S}^2$, alors $fg \in \mathcal{S}$ et $f * g \in \mathcal{S}$ et :

$$\text{TF}[fg] = \text{TF}[f] * \text{TF}[g] \quad ; \quad \text{TF}[f * g] = \text{TF}[f]\text{TF}[g] \quad (120)$$

Définition IV.11. Une distribution tempérée T est une fonctionnelle linéaire continue sur $\mathcal{S}$. L'ensemble des distributions tempérées est donc le dual topologique $\mathcal{S}'$ de l'espace de SCHWARTZ $\mathcal{S}$ (c'est un $\mathbf{K}$ -espace vectoriel).

Remarque IV.12.

- Les fonctions tests sont naturellement des fonctions déclinantes ($\mathcal{D} \subset \mathcal{S}$) : toute fonction nulle à l'infini tend évidemment vers 0 plus vite que polynomialement... Mieux : $\mathcal{D}$ est un sous-espace vectoriel de $\mathcal{S}$.
- De ce fait, toute fonctionnelle (linéaire) sur $\mathcal{S}$ est en particulier une fonctionnelle (linéaire) sur $\mathcal{D}$. De plus, comme $\mathcal{D}$ et $\mathcal{S}$ ont la même topologie (cf. III.3), les fonctionnelles continues sur $\mathcal{S}$ sont en particulier continues sur $\mathcal{D}$. Ainsi, les distributions tempérées sont des distributions ($\mathcal{S}' \subset \mathcal{D}'$, notez le renversement de l'inclusion). $\mathcal{S}'$ est un sous-espace vectoriel de $\mathcal{D}'$.

Définition IV.13. La transformée de FOURIER d'une distribution tempérée $T \in \mathcal{S}'$ est définie par :

$$\forall \phi \in \mathcal{S}, \langle \text{TF}[T], \phi \rangle = \langle T, \text{TF}[\phi] \rangle \quad (121)$$

On définit de la même manière la transformée inverse :

$$\forall \phi \in \mathcal{S}, \langle \text{TF}^{-1}[T], \phi \rangle = \langle T, \text{TF}^{-1}[\phi] \rangle \quad (122)$$

On a par exemple $\text{TF}[\delta_a] = [e^{iax}]$ et $\text{TF}[[1]] = 2\pi\delta$.

Si f est une fonction sommable (i.e. $f \in \mathcal{L}^1(\mathbf{R})$), alors on a :

$$\text{TF}[[f]] = [\text{TF}[f]] \quad \text{ou avec une autre notation} \quad \text{TF}[T_f] = T_{\text{TF}[f]} \quad (123)$$

Théorème IV.14. Si (T_n) est une suite de distributions de $\mathcal{S}'$ et $T \in \mathcal{S}'$, alors :

$$\lim_{n \rightarrow +\infty} T_n = T \Rightarrow \lim_{n \rightarrow +\infty} \text{TF}[T_n] = \text{TF}[T] \quad (124)$$

Ce n'est pas vrai pour des suites de fonctions.

On a :

- $\forall T \in \mathcal{S}', \text{TF}\left[\frac{dx}{dT}\right] = -ik\text{TF}[T];$
- $\forall T \in \mathcal{S}', \frac{d\text{TF}[T]}{dk} = \text{TF}[ixT];$
- $\forall T \in \mathcal{S}', \text{TF}[T_a] = e^{+iax}\text{TF}[T];$
- $\forall T \in \mathcal{S}', \text{TF}[e^{-iax}T] = (\text{TF}[T])_a :$
- Quand cela a un sens :

$$\forall (T, U) \in \mathcal{S}', \text{TF}[T * U] = \text{TF}[T]\text{TF}[U] \quad (125)$$

V. PROBABILITÉ

Les probabilités sont au départ issues de la théorie des jeux, mais ont trouvé des applications fondamentales en physique, en particulier en mécanique quantique où la notion de densité de probabilité $\rho(x) = |\Psi(x)|^2$ est centrale.

A. Éléments mathématiques

On considère une urne remplie de boules numérotées et de différentes couleurs qu'on tire au hasard. Chaque point de l'univers (ou espace des épreuves) Ω représente une issue possible du tirage.

Une fois l'univers défini, il faut assigner une probabilité à chacun de ces éléments. Si aucune indication n'est fournie, on peut supposer l'équiprobabilité :

$$\forall \omega \in \Omega, \mathbb{P}(\omega) = \frac{1}{\text{card}(\Omega)} \quad (126)$$

En général, on s'intéresse plutôt à un ensemble d'issues qui possèdent une certaine propriété : une *partie* A de Ω . On désigne par ω une issue élémentaire et $A \in \mathcal{P}(\Omega)$ l'ensemble des ω qui réalise la propriété (on note $\mathcal{P}(\Omega)$ est l'ensemble des parties de Ω)

Si ω est un élément de A , on dit que l'événement A est réalisé.

Exemple V.1. On considère l'univers $\Omega = \{1, 2, 3, 4, 5, 6\}$, l'événement « pair » est $A = \{2, 4, 6\}$

D'une manière plus formelle, on désigne par (Ω, T) l'espace probabilisable, où Ω est l'univers et T est une tribu de Ω , i.e. un ensemble de parties de Ω ($T \subset \mathcal{P}(\Omega)$) tel que :

1. $\emptyset \in T$, $\Omega \in T$;
2. $\forall A \in T$, $\bar{A} \in T$;
3. T est stable par union dénombrable : $\cup_{n \in \mathbf{N}} A_n \in T$.

Exemple V.2 (pile ou face). On considère l'univers $\Omega = \{P, F\}$, avec la tribu naturelle $T = \{\emptyset, \{P\}, \{F\}, \Omega\}$. Soient A et B deux événements, on dit que A et B sont réalisés simultanément si l'issue ω de l'épreuve vérifie $\omega \in A \cap B$. Si $A \cap B = \emptyset$ alors les événements sont dits incompatibles.

Probabilité	Ensembles
issue	$\omega \in \Omega$
événement	$A \subset \Omega$
A est réalisé	$\omega \in A$
A implique B	$A \subset B$
A ou B	$A \cup B$
A et B	$A \cap B$

TABLE I – Correspondance entre les lexiques de la théorie des probabilités et de la théorie des ensembles.

B. Probabilités

1. Définition générale

Définition V.3. Soit (Ω, T) un espace probabilisable, une probabilité sur cet espace est une application $\mathbb{P} : T \rightarrow [0; 1]$ telle que $\mathbb{P}(\Omega) = 1$ et si $(A_n)_{n \in \mathbf{N}}$ est une famille dénombrable d'événements de T deux à deux disjoints, alors :

$$\mathbb{P}\left(\bigcup_{n \in \mathbf{N}} A_n\right) = \sum_{n \in \mathbf{N}} \mathbb{P}(A_n) \quad (127)$$

On dit que $(\Omega, T, \mathbb{P})$ est un espace probabilisé.

Proposition V.4. Soient A et B deux événements de T

- Si $A \subset B$, alors $\mathbb{P}(A) \leq \mathbb{P}(B)$;
- $\mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A)$;
- Si $A \in B$, alors $\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A)$
- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

2. Probabilités conditionnelles

Exemple V.5 (la roulette). On considère une roulette représentée en figure 1

On a :

- la probabilité d'avoir un chiffre impair : $\mathbb{P}(\text{impair}) = \frac{3}{7}$;
- la probabilité d'avoir un chiffre impair sachant qu'il est blanc :
 $\mathbb{P}(\text{impair}|\text{blanc}) = \frac{1/7}{3/7} = \frac{1}{3}$.

Définition V.6. On a la probabilité conditionnelle :

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} \quad (128)$$

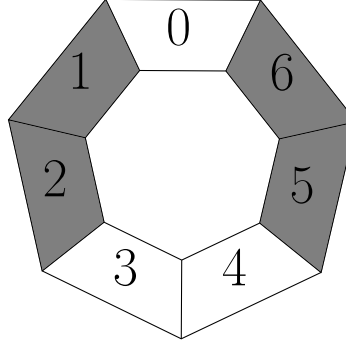


FIGURE 1 – Roulette.

3. Indépendance stochastique

Définition V.7. Deux événements A et B sont dits indépendants si $\mathbb{P}(A \cap B) = \mathbb{P}(A) \mathbb{P}(B)$.

L'indépendance stochastique est attachée à la probabilité associée à Ω et non à une propriété des événements.

Exemple V.8 (tirage dans une urne). On considère une urne remplie de 3 boules blanches B et deux boules noires N . On procède à deux tirages successifs.

$$\Omega = \{N_1 N_2, N_1 B_2, B_1 N_2, B_1 B_2\} \quad (129)$$

Calcul des probabilités (avec remise)

- $\mathbb{P}(N_1) = \mathbb{P}(N_2) = \frac{2}{5}$
- $\mathbb{P}(B_1) = \mathbb{P}(B_2) = \frac{3}{5}$
- $\mathbb{P}(N_1 N_2) = \mathbb{P}(N_1 \cap N_2) = \mathbb{P}(N_1) \mathbb{P}(N_2) = \frac{4}{25}$
- $\mathbb{P}(N_1 B_2) = \mathbb{P}(N_1 \cap B_2) = \mathbb{P}(N_1) \mathbb{P}(B_2) = \frac{6}{25}$
- $\mathbb{P}(B_1 N_2) = \mathbb{P}(B_1 \cap N_2) = \mathbb{P}(B_1) \mathbb{P}(N_2) = \frac{6}{25}$
- $\mathbb{P}(B_1 B_2) = \mathbb{P}(B_1 \cap B_2) = \mathbb{P}(B_1) \mathbb{P}(B_2) = \frac{9}{25}$

Calcul des probabilités (sans remise)

- $\mathbb{P}(N_1) = \mathbb{P}(N_2) = \frac{2}{5}$
- $\mathbb{P}(B_1) = \mathbb{P}(B_2) = \frac{3}{5}$
- $\mathbb{P}(N_1 N_2) = \mathbb{P}(N_1 \cap N_2) = \mathbb{P}(N_1) \mathbb{P}(N_2 | N_1) = \frac{2}{20}$
- $\mathbb{P}(N_1 B_2) = \mathbb{P}(N_1 \cap B_2) = \mathbb{P}(N_1) \mathbb{P}(B_2 | N_1) = \frac{6}{20}$
- $\mathbb{P}(B_1 N_2) = \mathbb{P}(B_1 \cap N_2) = \mathbb{P}(B_1) \mathbb{P}(N_2 | B_1) = \frac{6}{20}$
- $\mathbb{P}(B_1 B_2) = \mathbb{P}(B_1 \cap B_2) = \mathbb{P}(B_1) \mathbb{P}(B_2 | B_1) = \frac{6}{20}$

4. Loi binomiale

Une épreuve de BERNOULLI est une épreuve aléatoire n'admettant que deux issues (le « succès » et « l'échec »). On considère ici la répétition à l'identique d'une même épreuve de BERNOULLI.

Exemple V.9 (pile ou face). On considère un jeu de pile ou face avec les probabilités p et $q = 1 - p$. On réalise trois tirages. Les issues possibles sont résumées en fig. 2.

Vérification de la normalisation :

$$p^3 + 3p^2q + 3pq^2 + q^3 = (p + q)^3 = 1 \quad (130)$$

Définition V.10. La loi binomiale :

$$\mathbb{P}_N(n) = C_N^n p^n q^{N-n} \quad \text{avec} \quad C_N^n = \frac{N!}{n!(N-n)!} \quad (131)$$

est la probabilité de réaliser n succès lors de N répétition d'une épreuve de BERNOULLI de paramètre p .

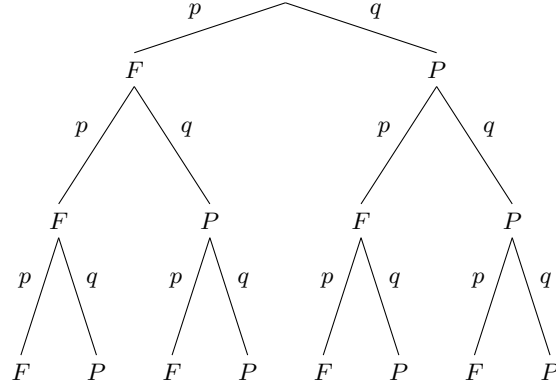
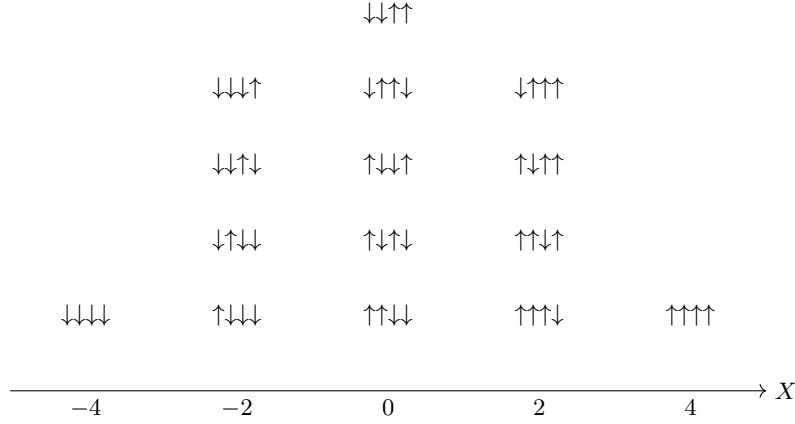


FIGURE 2 – Arbre de probabilité pour le jeu de pile ou face.

5. Formule de BAYES

Exemple V.11. On considère un ensemble de combinaisons de 4 spins.

FIGURE 3 – Ensemble des configurations de spins possibles. X désigne le spin total.

On considère les événements :

— $\{A\} = \uparrow\uparrow\downarrow\uparrow \Rightarrow \mathbb{P}(A) = \frac{1}{16}$;

— $\{B\} = \downarrow\downarrow\downarrow\uparrow \Rightarrow \mathbb{P}(B) = \frac{1}{16}$.

Si on fait une mesure tel que $X = 2$, alors $\mathbb{P}(A|X = 2) = \frac{1}{4}$ et $\mathbb{P}(B|X = 2) = 0$.

On a la formule de départ :

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \mathbb{P}(B|A) = \mathbb{P}(B) \mathbb{P}(A|B) \quad (132)$$

Soit R le résultat de la mesure ($(A_i)_i$ un partitionnement de l'univers), on a :

$$\mathbb{P}(A_i) \mathbb{P}(R|A_i) = \mathbb{P}(R) \mathbb{P}(A_i|R) \Rightarrow \mathbb{P}(A_i|R) = \frac{\mathbb{P}(A_i) \mathbb{P}(R|A_i)}{\mathbb{P}(R)} \quad (133)$$

donc :

$$\mathbb{P}(R) = \mathbb{P}\left(\bigcup_k (R \cap A_k)\right) = \sum_k \mathbb{P}(R \cap A_k) = \sum_k \mathbb{P}(A_k) \mathbb{P}(R|A_k) \quad (134)$$

On a alors :

Théorème V.12 (Formule de BAYES).

$$\mathbb{P}(A_i|R) = \frac{\mathbb{P}(A_i) \mathbb{P}(R|A_i)}{\sum_k \mathbb{P}(A_k) \mathbb{P}(R|A_k)} \quad (135)$$

Exemple V.13. On considère 6 tirage successifs avec remise dans une urne remplie de 6 boules (blanches et noires). Le but est d'estimer la composition de l'urne à partir de ces tirages. Avant tirage, il y a 7 configurations possibles équiprobables (table II). On peut alors estimer la probabilité de ces configurations en fonction du premier tirage (supposé ici B) (table III).

Noires	Blanche	Probabilité
0	6	1/7
1	5	1/7
2	4	1/7
3	3	1/7
4	2	1/7
5	1	1/7
6	0	1/7

TABLE II – Configurations possibles avant tirages.

A_k	$\mathbb{P}(A_k)$	$\mathbb{P}(B A_k)$	$\mathbb{P}(A_k B)$
(0B, 6N)	1/7	0	0
(1B, 5N)	1/7	$(1/6)^6$	10^{-5}
(2B, 4N)	1/7	$(2/6)^6$	10^{-3}
(3B, 3N)	1/7	$(3/6)^6$	0.01
(4B, 2N)	1/7	$(4/6)^6$	0.06
(5B, 1N)	1/7	$(5/6)^6$	0.23
(6B, 0N)	1/7	1	0.70

TABLE III – Tableau des probabilités du premier tirage en fonctions des configurations.

C. Variables aléatoires

On définit des lois de probabilité discrète ou continues. Pour les lois continues, on parle alors de densité de probabilité $\rho(x)$ telle que :

- $\forall x \in \mathbf{R}$, $\rho(x)$ est définie ;
- on a la normalisation $\int_{-\infty}^{+\infty} \rho(x) dx = 1$;
- $\rho(x < X < x + dx) = d\rho = \rho(x) dx$;
- $\rho(a < X < b) = \int_a^b \rho(x) dx$.

1. Caractéristiques numériques

X	$\mathbb{P}(X)$
x_1	p_1
x_2	p_2
$\vdots$	$\vdots$
x_n	p_n

TABLE IV – Probabilité de la variable aléatoire.

On définit la *moyenne* par :

$$\langle X \rangle = \bar{X} = \begin{cases} \sum_{i=1}^n x_i p_i & (\text{discret}) \\ \int_{-\infty}^{+\infty} x \rho(x) dx & (\text{continue}) \end{cases} \quad (136)$$

On définit la *variance* par :

$$\sigma^2(X) = V(X) = \begin{cases} \sum_{i=1}^n (x_i - \bar{X})^2 p_i & (\text{discret}) \\ \int_{-\infty}^{+\infty} (x - \bar{X})^2 \rho(x) dx & (\text{continue}) \end{cases} \quad (137)$$

D'une manière générale :

$$\overline{X^k} = \int_{-\infty}^{+\infty} x^k \rho(x) dx \quad ; \quad \overline{(X - \bar{X})^k} = \int_{-\infty}^{+\infty} (x - \bar{X})^k \rho(x) dx \quad (138)$$

Théorème V.14 (KÖNIG-HUYGENS). *On a :*

$$\overline{(X - \bar{X})^2} = \overline{X^2} - (\bar{X})^2 \quad (139)$$

D'autres moments peuvent être définis comme le coefficient d'asymétrie :

$$a = \frac{\overline{(X - \bar{X})^3}}{\sigma^3} \quad (140)$$

$a = 0$ si la répartition est symétrique, $a \neq 0$ sinon. On peut également définir de la même manière le coefficient d'aplatissement :

$$A = \frac{\overline{(X - \bar{X})^4}}{\sigma^4} \quad (141)$$

Exemple V.15. *On considère $\forall X \in [0; 2]$, $\rho(x) = C e^{-x}$ ($\rho(x) = 0$ en dehors de $[0; 2]$), où C est une constante à déterminer.*

Avec la normalisation, on a :

$$1 = \int_{-\infty}^{+\infty} \rho(x) dx = \int_0^2 C e^{-x} dx \Rightarrow C(1 - e^{-2}) = 1 \Rightarrow C = \frac{1}{1 - e^{-2}} \quad (142)$$

On cherche désormais la moyenne :

$$\bar{X} = \int_{-\infty}^{+\infty} x \rho(x) dx = \int_0^2 x C e^{-x} dx \Rightarrow [\dots] \Rightarrow \bar{X} = \frac{1 - 3e^{-2}}{1 - e^{-2}} \quad (143)$$

On calcule le moment d'ordre k , $\overline{X^k}$:

$$\overline{X^k} = \int_{-\infty}^{+\infty} x^k \rho(x) dx = \int_0^2 C x^k e^{-x} dx \quad (144)$$

On introduit un paramètre λ :

$$\begin{aligned} \overline{X^k} &= -C \int_0^2 \frac{\partial^k}{\partial \lambda^k} e^{-\lambda x} dx \Big|_{\lambda=1} \\ &= -C \frac{\partial^k}{\partial \lambda^k} \int_0^2 e^{-\lambda x} dx \Big|_{\lambda=1} \\ &= -C \frac{\partial^k}{\partial \lambda^k} \left[\frac{e^{-\lambda x}}{\lambda} \right]_0^2 \Big|_{\lambda=1} \\ &= -C \frac{\partial^k}{\partial \lambda^k} \frac{1}{\lambda} (1 - e^{2\lambda}) \Big|_{\lambda=1} \end{aligned} \quad (145)$$

Calcul de la moyenne à partir du moment :

$$\bar{X} = C \frac{\partial^k}{\partial \lambda^k} \left(\frac{1 - e^{-2\lambda}}{\lambda} \right) \Big|_{\lambda=1} = [\dots] = \frac{1 - 3e^{-2}}{1 - e^{-2}} \quad (146)$$

2. Les fonctions de répartition, génératrices et caractéristiques

On a la fonction de répartition :

$$F(x) = \mathbb{P}(X < x) = \int_{-\infty}^x \rho(y) dy \quad (147)$$

On peut alors écrire :

$$\mathbb{P}(x_2 < X < x_1) = \int_{x_2}^{x_1} \rho(x) dx = F(x_1) - F(x_2) \quad (148)$$

Cette fonction vérifie :

- $\forall x, F(x) \geq 0$;
- $\lim_{x \rightarrow -\infty} F(x) = 0$;
- $\lim_{x \rightarrow +\infty} F(x) = 1$;
- $\mathbb{P}(x < X < x + dx) = \rho(x) dx = dF$

On définit la **fonction génératrice des moments** par :

$$G(t) = \int_{-\infty}^{+\infty} e^{tx} \rho(x) dx = \overline{e^{tX}} \quad (149)$$

En calculant la dérivée k -ième :

$$\begin{aligned} \left. \frac{d^k G(t)}{dt^k} \right|_{t=0} &= \left. \frac{d^k}{dt^k} \int_{-\infty}^{+\infty} e^{tx} \rho(x) dx \right|_{t=0} \\ &= \int_{-\infty}^{+\infty} \frac{\partial^k}{\partial t^k} e^{tx} \rho(x) dx \Big|_{t=0} \\ &= \int_{-\infty}^{+\infty} x^k e^{tx} \rho(x) dx \Big|_{t=0} \\ &= \int_{-\infty}^{+\infty} x^k \rho(x) dx = \overline{X^k} \end{aligned} \quad (150)$$

On retrouve le moment d'ordre k .

$$G(t) = \overline{e^{tX}} = 1 + t\overline{X} + \frac{t^2 \overline{X^2}}{2!} + \dots = 1 + t\overline{X} + \frac{t^2}{2!} \overline{X^2} + \dots \quad (151)$$

Pour les moments centrés :

$$\tilde{G}(t) = \overline{e^{t(x-\overline{X})}} = \int_{-\infty}^{+\infty} e^{t(x-\overline{X})} \rho(x) dx \quad (152)$$

On définit la **fonction caractéristique** par :

$$\varphi(t) = \overline{e^{itX}} = \int_{-\infty}^{+\infty} e^{itx} \rho(x) dx = 1 + it\overline{X} - \frac{t^2}{2!} \overline{X^2} + \dots \quad (153)$$

3. Les lois classiques

On a la **loi binomiale** :

$$\mathbb{P}_N(n) = C_N^n p^n q^{N-n} \quad (154)$$

avec $p + q = 1$. On vérifie la normalisation :

$$\sum_{n=0}^N \mathbb{P}_N(n) = \sum_{n=0}^N C_N^n p^n q^{N-n} = (p + q)^N = 1 \quad (155)$$

Sa fonction génératrice est :

$$G(t) = \sum_{n=0}^N e^{tn} C_N^n p^n q^{N-n} = (pe^t + q)^N \quad (156)$$

On fait le développement limité en 0 :

$$G(t) \simeq \left[p \left(1 + t + \frac{t^2}{2} + \frac{t^3}{3} + \mathcal{O}(t^4) \right) + q \right]^N \simeq [p + q + \varepsilon]^N \quad (157)$$

où $\varepsilon \ll 1$.

$$\begin{aligned} G(t) &\simeq 1 + N\varepsilon + \frac{N(N-1)}{2!}\varepsilon^2 + \frac{N(N-1)(N-2)}{3!}\varepsilon^3 + \dots \\ &\simeq 1 + N \left(pt + \frac{pt^2}{2!} + \frac{pt^3}{3!} + \mathcal{O}(t^4) \right) + \frac{N(N-1)}{2!} \left(pt + \frac{pt^2}{2!} + \frac{pt^3}{3!} + \mathcal{O}(t^4) \right)^2 \\ &\quad + \frac{N(N-1)(N-2)}{3!} \left(pt + \frac{pt^2}{2!} + \frac{pt^3}{3!} + \mathcal{O}(t^4) \right)^3 + \dots \\ &\simeq 1 + t\bar{n} + \frac{t^2}{2!}\bar{n}^2 + \frac{t^3}{3!}\bar{n}^3 + \dots \end{aligned} \quad (158)$$

d'où :

$$\begin{cases} \bar{n} = Np \\ \bar{n}^2 = Np + N(N-1)p^2 \end{cases} \Rightarrow V(n) = Npq \quad (159)$$

On a la **loi de GAUSS** :

$$\mathbb{P}_N(n) = \frac{N!}{n!(N-n)!} p^n q^{N-n} \xrightarrow{\text{continu}} \mathbb{P}_N(x) = \frac{N!}{x!(N-x)!} p^x q^{N-x} \quad (160)$$

On développe autour de X^* (le mode, ici correspondant à la valeur moyenne) :

$$\begin{aligned} \ln[\mathbb{P}_N(X^* + \varepsilon)] &\simeq \ln[\mathbb{P}_N(X^*)] + \varepsilon \left. \frac{d}{dx} \{\ln[\mathbb{P}_N(x)]\} \right|_{X^*} \\ &\quad + \frac{\varepsilon^2}{2} \left. \frac{d^2}{dx^2} \{\ln[\mathbb{P}_N(x)]\} \right|_{X^*} \\ &\quad + \dots \\ &\simeq \ln[\mathbb{P}_N(X^*)] + \frac{\varepsilon^2}{2} \left. \frac{d^2}{dx^2} \{\ln[\mathbb{P}_N(x)]\} \right|_{X^*} + \dots \end{aligned} \quad (161)$$

Calcul du mode X^* . Avec la loi binomiale :

$$\ln[\mathbb{P}_N(x)] = \ln(N!) - \ln(x!) - \ln[(N-x)!] + x \ln(p) + (N-x) \ln(q) \quad (162)$$

Par dérivation :

$$\begin{aligned} \frac{d}{dx} \ln[\mathbb{P}_N(x)] &= -\ln(x) + \ln(N-x) + \ln(p) - \ln(q) \\ \frac{d}{dx} \ln[\mathbb{P}_N(X^*)] &= 0 \Rightarrow -\ln(X^*) + \ln(N-X^*) + \ln(p) - \ln(q) = 0 \\ &\Rightarrow X^* = Np \end{aligned} \quad (163)$$

Le calcul de la dérivée seconde donne :

$$\left. \frac{d^2}{dx^2} \ln[\mathbb{P}_N(x)] \right|_{x=X^*} = -\frac{1}{Npq} = \frac{1}{\sigma^2} < 0 \quad (164)$$

On a :

$$\ln[\mathbb{P}_N(x)] = \ln[\mathbb{P}_N(X^*)] + \frac{\varepsilon^2}{2} \frac{d^2}{dx^2} \{\ln[\mathbb{P}_N(x)]\} \Big|_{x=X^*} = C - \frac{\varepsilon^2}{2\sigma^2} \quad (165)$$

d'où :

$$\begin{aligned} \mathbb{P}_N(x) &= K e^{-\varepsilon^2/(2\sigma^2)} \quad \text{avec} \quad \varepsilon = X - X^* \quad ; \quad X^* = Npq = \bar{X} \\ \rho(x) &= K e^{-\frac{x-\bar{X}}{2\sigma^2}} \end{aligned} \quad (166)$$

Soit $I_\lambda(k) = \int_{-\infty}^{+\infty} u^k e^{-\lambda u^2} du$. Si k est impair, alors $I_\lambda(k) = 0$. Si k est pair, $k = 2p$:

$$\begin{aligned} I_\lambda(2p) &= \int_{-\infty}^{+\infty} u^{2p} e^{-\lambda u^2} du = \int_{-\infty}^{+\infty} \left(-\frac{\partial}{\partial \lambda}\right)^p e^{-\lambda u^2} du = \left(-\frac{\partial}{\partial \lambda}\right)^p \int_{-\infty}^{+\infty} e^{-\lambda u^2} du \\ &= \left(\frac{\partial}{\partial \lambda}\right) \frac{1}{\sqrt{\lambda}} \int_{-\infty}^{+\infty} e^{-x^2} dx = \left(-\frac{\partial}{\partial \lambda}\right)^p \sqrt{\frac{\pi}{\lambda}} \end{aligned} \quad (167)$$

Exemple V.16. On a par exemple $I_\lambda(0) = \sqrt{\frac{\pi}{\lambda}}$, $I_\lambda(2) = \frac{1}{2} \frac{\sqrt{\pi}}{\lambda^{3/2}}$.

Calcul de la normalisation :

$$\int_{-\infty}^{+\infty} \rho(x) dx = 1 \Leftrightarrow \bar{X}^0 = k I_{1/(2\sigma^2)}(0) = K \sqrt{2\pi\sigma^2} = 1 \quad (168)$$

On calcul $K = \frac{1}{\sqrt{2\pi\sigma^2}}$ avec la normalisation, donc :

$$\rho(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\bar{X})^2}{2\sigma^2}} \quad (169)$$

Calcul de la moyenne :

$$\bar{X} = \int_{-\infty}^{+\infty} x \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2} \frac{(x-\bar{X})^2}{\sigma^2}} \quad (170)$$

On pose $u = x - \bar{X}$:

$$\begin{aligned} \bar{X} &= \int_{-\infty}^{+\infty} (u + \bar{X}) \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{u^2}{2\sigma^2}} du = \int_{-\infty}^{+\infty} u \frac{1}{\sqrt{2\pi}} e^{-u^2} du + \bar{X} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{u^2}{2\sigma^2}} du \\ &\Rightarrow \bar{X} = X^* = Np \quad (\text{loi binômiale}) \end{aligned} \quad (171)$$

Calcul de $\bar{X}^2$:

$$\begin{aligned} \bar{X}^2 &= \int_{-\infty}^{+\infty} x^2 \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2} \frac{(x-\bar{X})^2}{\sigma^2}} dx = \int_{-\infty}^{+\infty} (u + \bar{X})^2 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2} \frac{u^2}{\sigma^2}} du \\ &= \int_{-\infty}^{+\infty} u^2 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{u^2}{2\sigma^2}} du + 2\bar{X} \int_{-\infty}^{+\infty} u \frac{1}{2\pi\sigma^2} e^{-\frac{u^2}{2\sigma^2}} du + \bar{X}^2 \int_{-\infty}^{+\infty} \frac{1}{2\pi\sigma^2} e^{-\frac{u^2}{2\sigma^2}} du \\ &\Rightarrow \bar{X}^2 = \sigma^2 + \bar{X}^2 \Rightarrow \sigma^2 = \bar{X}^2 - \bar{X}^2 \end{aligned} \quad (172)$$

On a la fonction génératrice :

$$G(t) = e^{t\bar{X}} = \int_{-\infty}^{+\infty} e^{tx} \underbrace{\frac{1}{\sqrt{2\pi\sigma^2}}}_{k} e^{-\frac{1}{2} \frac{(x-\bar{X})^2}{\sigma^2}} dx = k \int_{-\infty}^{+\infty} e^{-\frac{1}{2} \frac{(x-\bar{X})^2}{\sigma^2}} dx \quad (173)$$

Avec :

$$\begin{aligned} -\frac{1}{2} \frac{(x-\bar{X})^2}{\sigma^2} + tx &= -\frac{1}{2\sigma^2} [(x-\bar{X})^2 + 2\sigma^2 tx] \\ &= [\dots] \\ &= -\frac{1}{2\sigma^2} \left\{ [x - (\bar{X} + 2\sigma^2 t)]^2 + \bar{X}^2 - (\bar{X} + \sigma^2 t)^2 \right\} \\ &= -\frac{1}{2\sigma^2} [x - (\bar{X} + 2\sigma^2 t)]^2 + \frac{\sigma^2 t^2}{2} + t\bar{X} \end{aligned} \quad (174)$$

D'où l'expression de l'intégrale :

$$G(t) = e^{\frac{1}{2}(\sigma^2 t^2 + 2\bar{X}t)} \underbrace{\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}[x - (\bar{X} + \sigma^2 t)]^2} dx}_1$$

$$= e^{\frac{1}{2}(\sigma^2 t^2 + 2\bar{X}t)}$$
(175)

Calcul de $\mathbb{P}(a \leq X < b)$:

$$\mathbb{P}(a \leq X < b) = \int_a^b \rho(x) dx = F(b) - F(a)$$
(176)

Pour le calcul, on utilise l'expression de la gaussienne normalisée :

$$f(r) = \frac{1}{\sqrt{2\pi}} e^{-\frac{r^2}{2}}$$
(177)

On passe par le changement de variable $dp = \rho(x) dx = f(r) dr$:

$$r = \frac{x - \bar{X}}{\sigma} \Rightarrow f(r) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2} \frac{(x - \bar{X})^2}{\sigma^2}} \sigma = \frac{1}{\sqrt{2\pi}} e^{-\frac{r^2}{2}}$$
(178)

On a donc $r_a = \frac{a - \bar{X}}{\sigma}$ et $r_b = \frac{b - \bar{X}}{\sigma}$, alors :

$$\mathbb{P}(a \leq X < b) = \mathbb{P}(r_a \leq R < r_b) = \Pi(r_b) - \Pi(r_a)$$
(179)

où Π est la fonction de répartition de R (valeurs tabulée).

D. Systèmes de variables aléatoires

On considère :

$$F(x, y) = \mathbb{P}(X < x, Y < y)$$
(180)

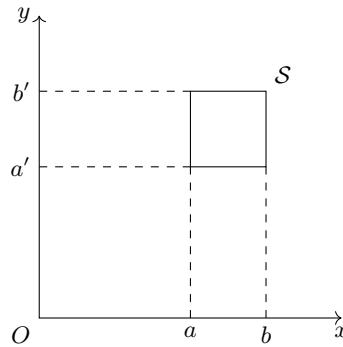


FIGURE 4 – Représentation graphique.

On a :

$$\mathbb{P}[(x, y) \in (\mathcal{S})] = F(b, b') - F(a, b') - F(b, a') + F(a, a')$$
(181)

On pose $a = x$, $b = x + dx$ et $a' = y$, $b' = y + dy$. On peut réaliser le développement limité :

$$\begin{aligned} \begin{cases} F(x + dx, y + dy) = F(x, y) + \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy + \frac{1}{2} \frac{\partial^2 F}{\partial x^2} dx^2 + \frac{1}{2} \frac{\partial^2 F}{\partial y^2} dy^2 + \frac{1}{2} \frac{\partial^2 F}{\partial x \partial y} dx dy \\ F(x + dx, y) = F(x, y) + \frac{1}{2} \frac{\partial^2 F}{\partial x^2} dx^2 \\ F(x + dx, y) = F(x, y) + \frac{1}{2} \frac{\partial^2 F}{\partial y^2} dy^2 \end{cases} \\ \Rightarrow F(x + dx, y + dy) = F(x, y) + \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy + \frac{1}{2} \frac{\partial^2 F}{\partial x^2} dx^2 + \frac{1}{2} \frac{\partial^2 F}{\partial y^2} dy^2 + \frac{\partial^2 F}{\partial x \partial y} dx dy \\ - F(x, y) - \frac{\partial F}{\partial x} dx - \frac{1}{2} \frac{\partial^2 F}{\partial x^2} dx^2 - F(x, y) - \frac{\partial F}{\partial y} dy - \frac{1}{2} \frac{\partial^2 F}{\partial y^2} dy^2 \\ = \frac{\partial^2 F}{\partial x \partial y} dx dy \end{aligned} \quad (182)$$

d'où :

$$\rho(x, y) = \frac{\partial^2 F}{\partial x \partial y} \quad ; \quad \mathbb{P}[(x, y) \in \mathcal{S}] = \iint_{\mathcal{S}} \rho(x, y) dx dy \quad (183)$$

Définition des densités marginales :

$$F_2(y) = F(+\infty, y) = \iint_{\substack{u=-\infty \\ v=-\infty}}^{u=+\infty \\ v=y} \rho(u, v) du dv \quad ; \quad \rho_2(x) = \frac{dF_2(y)}{dy} = \int_{-\infty}^{+\infty} \rho(x, y) dy \quad (184)$$

de même :

$$F_1(x) = F(x, +\infty) = \iint_{\substack{u=-\infty \\ v=-\infty}}^{u=x \\ v=+\infty} \rho(u, v) du dv \quad ; \quad \rho_1(x) = \frac{dF_1(x)}{dx} = \int_{-\infty}^{+\infty} \rho(x, y) dy \quad (185)$$

Densité conditionnelles.

$$\mathbb{P}(x \leq X < x + dx, y \leq Y < y + dy) = \rho(x, y) dx dy \quad (186)$$

Avec $\mathbb{P}(A \cap B) = \mathbb{P}(A) \mathbb{P}(B|A)$ et $\mathbb{P}(y \leq Y < y + dy | x \leq X < x + dx) = \rho(y|x) dy$:

$$\rho(x, y) dx dy = \rho_1(x) \rho(y|x) dx dy \Rightarrow \rho(y|x) = \frac{\rho(x, y)}{\rho_1(x)} = \frac{\rho(x, y)}{\int_{-\infty}^{+\infty} \rho(x, y) dy} \quad (187)$$

On a les conditions d'indépendance des variables aléatoires :

- $\mathbb{P}(A, B) = \mathbb{P}(A) \mathbb{P}(B)$;
- $\rho(x, y) = \rho_1(x) \rho_2(y)$;
- $\rho(y|x) = \rho_2(y)$.

Moment d'un système de variables aléatoires. Pour deux variables, on a :

$$\overline{X^k} = \iint_{-\infty}^{+\infty} x^k \rho(x, y) dx dy = \int_{-\infty}^{+\infty} x^k \left[\int_{-\infty}^{+\infty} \rho(x, y) dy \right] dx = \int_{-\infty}^{+\infty} x^k \rho_1(x) dx \quad (188)$$

On a :

$$V(X) = \overline{(X - \overline{X})^2} = \int_{-\infty}^{+\infty} (x - \overline{X})^2 \rho_1(x) dx \quad (189)$$

Les moments croisés sont :

$$\begin{aligned} M(X^k Y^l) &= \overline{X^k Y^l} = \iint_{-\infty}^{+\infty} x^k y^l \rho(x, y) dx dy \\ M(\dot{X}^k \dot{Y}^l) &= \overline{(X - \overline{X})^k (Y - \overline{Y})^l} = \iint_{-\infty}^{+\infty} (x - \overline{X})^k (y - \overline{Y})^l \rho(x, y) dx dy \end{aligned} \quad (190)$$

Remarque V.17. On a $\overline{(X - \bar{X})(Y - \bar{Y})} = \overline{XY} - \bar{X} \cdot \bar{Y}$. On définit le coefficient de corrélation :

$$\mathcal{C}_{\text{corr}}(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} \quad (191)$$

Pour un système de variables aléatoires, $F(x_1, \dots, x_n)$:

$$\begin{aligned} F_i(x_i) &= F(+\infty, \dots, +\infty, x_i, +\infty, \dots, +\infty) \\ \rho_i(x_i) &= \int \dots \int_{-\infty}^{+\infty} \rho(x_1, \dots, x_n) dx_1 dx_{i-1} dx_{i+1} dx_n \end{aligned} \quad (192)$$

E. Fonctions de variables aléatoires

Méthode. On procède suivant le protocole :

1. tracer la fonction ;
2. résoudre $y = f(x)$;
3. calcul de la densité ;
4. remplacer la variable x_i par y ;
5. vérifier la normalisation.

Soit une variable aléatoires $Y = g(X)$. On connaît la densité marginale $\rho_X(x)$. On cherche $\bar{Y}$ et σ_Y :

$$\bar{Y} = \int_{-\infty}^{+\infty} y \rho_Y(y) dy \quad (193)$$

Par conservation des probabilités $dp = \rho_X(x) dx = \rho_Y(y) dy$:

$$\bar{Y} = \pm \int_{-\infty}^{+\infty} y(x) \rho_X(x) dx \quad (194)$$

On a :

$$\rho_Y(y) = \sum_{i=1}^n \frac{\rho_X(x_i)}{|y'(x_i)|} \quad (195)$$

Exemple V.18. Soit $Y = (X + a)^2$, où X suit une loi normale centrée réduite, i.e. $\rho_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. On représente la fonction $f : y = (x + a)^2$.

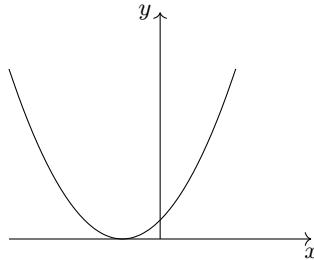


FIGURE 5 – Loi normale.

- si $y < 0$, alors $(x + a)^2 = y$ n'a pas de solution ;
- si $y > 0$, alors $(x + a)^2 = y$ a deux solutions : $x = -a \pm \sqrt{y}$.

On applique alors la formule :

$$\rho_Y(y) = \sum_{i=1}^N \frac{\rho_X(x_i)}{|y'(x_i)|} \quad (196)$$

avec :

$$|y'(x_i)| = |2(x+a)| = 2\sqrt{y} \quad (197)$$

on a alors :

$$\rho_Y(y) = \frac{1}{2\sqrt{y}} \left[\frac{e^{-\left|\frac{a+\sqrt{y}}{2}\right|^2}}{\sqrt{2\pi}} + \frac{e^{-\left|-a+\sqrt{y}\right|^2}}{\sqrt{2\pi}} \right] \quad (198)$$

d'où :

$$\int_0^{+\infty} \rho_Y(y) dy = \int_0^{+\infty} \frac{1}{2\sqrt{y}} \left[\frac{e^{-\left|\frac{a+\sqrt{y}}{2}\right|^2}}{\sqrt{2\pi}} + \frac{e^{-\left|-a+\sqrt{y}\right|^2}}{\sqrt{2\pi}} \right] dy \quad (199)$$

par un changement de variable, on obtient :

$$\int_{-\infty}^{+\infty} \rho_Y(y) dy = \int_{-\infty}^{+\infty} \frac{e^{-u^2/2}}{\sqrt{2\pi}} du = 1 \quad (200)$$

Calcul de la moyenne. On pose :

$$\begin{aligned} Z &= \lambda X + \mu Y \\ \Rightarrow \bar{Z} &= \int_{-\infty}^{+\infty} z f(z) dz \\ &= \iint_{-\infty}^{+\infty} (\lambda x + \mu y) f(x, y) dx dy \\ &= \lambda \iint_{-\infty}^{+\infty} x f(x, y) dx dy + \mu \iint_{-\infty}^{+\infty} y f(x, y) dx dy \\ &= \lambda \int_{-\infty}^{+\infty} x f_1(x) dx + \mu \int_{-\infty}^{+\infty} y f_2(y) dy \\ &= \lambda \bar{X} + \mu \bar{Y} \end{aligned} \quad (201)$$

Calcul de la variance. On pose :

$$\begin{aligned} &\begin{cases} Z_1 = \lambda_1 X + \mu_1 Y \\ Z_2 = \lambda_2 X + \mu_2 Y \end{cases} \\ \Rightarrow \text{cov}(Z_1, Z_2) &= \overline{(Z_1 - \bar{Z}_1)(Z_2 - \bar{Z}_2)} = \iint (z_1 - \bar{Z}_1)(z_2 - \bar{Z}_2) f(z_1, z_2) dz_1 dz_2 \end{aligned} \quad (202)$$

On calcul $\text{cov}(X, Z_2)$:

$$\begin{aligned} \text{cov}(X, Z_2) &= \text{cov}(X, \lambda_2 X + \mu_2 Y) \\ &= \iint (x - \bar{X}) [\lambda_2 (x - \bar{X}) + \mu_2 (y - \bar{Y})] f(x, y) dx dy \\ &= \lambda_2 \text{cov}(X, X) + \mu_2 \text{cov}(X, Y) \end{aligned} \quad (203)$$

La covariance est bilinéaire. On montre de même :

$$\text{cov}(Z_1, Z_2) = \lambda_1 \lambda_2 \text{cov}(X, X) + \lambda_1 \mu_2 \text{cov}(X, Y) + \mu_1 \lambda_2 \text{cov}(Y, X) + \mu_1 \mu_2 \text{cov}(Y, Y) \quad (204)$$

Remarque V.19. L'écart-type d'une somme $X + Y$ de variables aléatoires indépendantes est $\sigma_{X+Y} = \sqrt{\sigma_X^2 + \sigma_Y^2}$ (théorème de Pythagore). L'écart-type du produit d'une variable aléatoire par un scalaire λX est $\sigma_{\lambda X} = |\lambda| \sigma_X$. La variance de $\lambda X + Y$ est $\lambda^2 V(X) + 2\lambda \text{cov}(X, Y) + V(Y) \geq 0$. Comme le discriminant est positif, on a :

$$0 \leq \frac{\text{cov}(X, Y)}{V(X)V(Y)} \leq 1 \Rightarrow 0 \leq \mathcal{C}_{\text{corr}}(X, Y) \leq 1 \quad (205)$$

Loi de répartition. Pour une somme de variables aléatoires $Z = X + Y$:

$$F_Z(z) = \mathbb{P}(Z < z) \quad (206)$$

So $Z < z$ alors $X + Y < z$ donc $Y < -x + z$ d'où :

$$P(Z < z) = \iint_{\mathcal{D}} f(x, y) \, dx \, dy = \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{-x+z} f(x, y) \, dy \right] dx \quad (207)$$

or :

$$\begin{aligned} \rho_Z(z) &= \frac{dF(z)}{dz} \\ &= \frac{d}{dz} \left\{ \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{-x+z} f(x, y) \, dy \right] dx \right\} \\ &= \int_{-\infty}^{+\infty} \frac{d}{dz} \left[\int_{-\infty}^{-x+z} f(x, y) \, dy \right] dx \\ &= \int_{-\infty}^{+\infty} f(x, -x+z) \, dx \end{aligned} \quad (208)$$

Si X et Y sont deux variables aléatoires indépendantes :

$$f(x, y) = f_X(x)f_Y(y) \quad (209)$$

Dans ce cas :

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x)f_Y(z-x) \, dx = (f_X * f_Y)(z) \quad (210)$$

où $*$ désigne le produit de convolution.

Fonction caractéristique. On a :

$$\phi(t) = \overline{e^{itZ}} = \overline{e^{it(X+Y)}} = \overline{e^{itX}e^{itY}} \quad (211)$$

Si X et Y sont indépendantes :

$$\phi(t) = \overline{e^{itX}e^{itY}} \quad (212)$$

F. Limites de probabilités

Convergence en probabilité.

$$\begin{aligned} \forall \varepsilon > 0, \forall \delta > 0, \exists n_0 \in \mathbf{N}, \forall n > n_0, \mathbb{P}(|X_n - L| < \varepsilon) > 1 - \delta \\ \text{i.e.} \quad \mathbb{P}(|X_n - L| > \varepsilon) > \delta \end{aligned} \quad (213)$$

Inégalité de MARKOV. Soit Y une variable aléatoire, on a l'inégalité de MARKOV :

$$\forall a > 0, \mathbb{P}(Y \geq a) \leq \frac{\overline{Y}}{a} \quad (214)$$

Inégalité de BIENAYMÉ-TCHEBYCHEV. On pose dans l'inégalité de MARKOV $Y = (X - \bar{X})^2$, donc $\bar{Y} = \overline{(X - \bar{X})^2} = \sigma_X^2$. On pose $a = \varepsilon^2 > 0$ donc :

$$\mathbb{P}((X - \bar{X})^2 \geq \varepsilon^2) \leq \frac{\sigma_X^2}{\varepsilon^2} \Rightarrow \mathbb{P}(X - \bar{X} \geq \varepsilon) \leq \frac{\sigma_X^2}{\varepsilon^2} \Leftrightarrow \mathbb{P}(X - \bar{X} \leq \varepsilon) \geq 1 - \frac{\sigma_X^2}{\varepsilon^2} \quad (215)$$

Loi des grands nombres. Soit $(X_i)_{0 \leq i \leq n}$ une suite de variables aléatoires indépendantes et identiquement distribuées (*i.e.* qui suivent la même loi). Soit la variable aléatoires Y_n telle que :

$$Y_n = \frac{1}{n} \sum_{i=1}^n X_i$$

$$V(Y_n) = V\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n V(X_i) = \frac{1}{n^2} n \sigma_X^2 \Rightarrow \sigma_Y = \frac{\sigma_X}{\sqrt{n}} \quad (216)$$

où σ_X est l'écart-type de chaque variable X_i (étant identiquement distribuées, elles ont en particulier le même écart-type).

Théorème central limite. On conserve les notations précédentes et on considère une gaussienne $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. Soit $\phi(t) = \text{TF}[f](t) = e^{-t^2/2}$. La fonction caractéristique de Y_n est :

$$\phi_{Y_n}(t) = \overline{e^{itY_n}} = \overline{e^{it \sum_{i=1}^n X_i}} = \overline{\prod_{j=1}^n e^{itX_j}} = \prod_{j=1}^n \overline{e^{itX_j}} = [\phi_X(t)]^n \quad (217)$$

On passe au logarithme :

$$\ln[\phi_{Y_n}(t)] = n \ln[\phi_X(t)] \quad (218)$$

On suppose sans perte de généralité que $\bar{X} = 0$. Avec $\phi_X(t) \simeq 1 + i\bar{X} + \frac{t^2}{2} \overline{X^2} + \dots$, on a :

$$\ln[\phi_{Y_n}(t)] = n \ln[1 - \frac{t^2}{2} \overline{X^2} + \dots] \quad (219)$$

On réalise le changement de variable $Z_n = \frac{Y_n - \bar{Y}_n}{\sigma_Y \sqrt{n}}$. On a alors $\phi_{Z_n}(t) = \phi_{Y_n}\left(\frac{t}{\sigma_Y \sqrt{n}}\right)$ donc :

$$\ln[\phi_{Z_n}(t)] = \ln[1 - \frac{t^2}{2n} + \dots] = -\frac{t^2}{2} + f\left(\frac{1}{n}\right) \quad (220)$$

Ainsi, quand $n \rightarrow +\infty$:

$$\phi_{Z_n}(t) = e^{-\frac{t^2}{2}} \Rightarrow Z_n(t) = \frac{1}{\sqrt{2\pi}} e^{-z_n^2/2} \quad (221)$$