

Espaces de Hilbert et Analyse de Fourier  
(MHT 613), Printemps 2008

Alain Yger

3 mai 2009



# Table des matières

<b>1</b>	<b>Espaces de Hilbert</b>	<b>1</b>
1.1	Rappel : l'espace euclidien $\mathbb{R}^n$ et l'espace hilbertien $\mathbb{C}^n$ . . . . .	1
1.2	Le saut de la dimension finie à la dimension infinie . . . . .	5
1.3	Les notions d'espace de Hilbert (réel ou complexe) . . . . .	7
1.3.1	Le cadre Hilbert . . . . .	7
1.3.2	Un formulaire géométrique . . . . .	8
1.4	Les exemples majeurs . . . . .	9
1.4.1	Les exemples $\mathbb{R}^n$ et $\mathbb{C}^n$ . . . . .	9
1.4.2	Les espaces $l_{\mathbb{R}}^2(I)$ et $l_{\mathbb{C}}^2(I)$ . . . . .	10
1.4.3	Les espaces $L_{\mathbb{K}}^2(U, dx)$ , $U$ partie mesurable de $\mathbb{R}^n$ , $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ . . . . .	12
1.4.4	Les espaces $L_{\mathbb{K}}^2(\Omega, \mathcal{T}, P)$ . . . . .	15
1.5	Le théorème de projection orthogonale sur un convexe fermé . . . . .	17
1.6	Systèmes orthonormés, « <i>frames</i> » ; exemples . . . . .	20
1.7	Espaces de Hilbert séparables ; notion de base hilbertienne . . . . .	27
1.8	Le cas des espaces de Hilbert non séparables . . . . .	29
1.9	Exemples de bases hilbertiennes dans le cas séparable . . . . .	31
1.9.1	Densité des classes de fonctions polynômiales dans $L_{\mathbb{K}}^2(U, dx)$ pour $U$ mesurable borné . . . . .	31
1.9.2	Densité des classes de polynômes trigonométriques dans $L_{\mathbb{K}}^2(\mathbb{T}^n)$ . . . . .	32
1.9.3	A propos des polynômes de Hermite . . . . .	34
1.10	Le théorème de dualité . . . . .	35
1.10.1	Quelques préliminaires à propos des formes linéaires continues sur un Hilbert . . . . .	35
1.10.2	L'énoncé du théorème de dualité . . . . .	36
1.11	Opérateurs de $H$ dans $H$ ; notion d'adjoint . . . . .	37
1.11.1	Préliminaires sur les opérateurs linéaires continus de $H$ dans $H$ . . . . .	38
1.11.2	La notion d'opérateur adjoint . . . . .	40
1.11.3	L'intérêt de la prise d'adjoint ; deux exemples pratiques . . . . .	42
1.12	Un procédé algorithmique : le <i>matching</i> . . . . .	49
<b>2</b>	<b>L'analyse de Fourier et ses applications</b>	<b>53</b>
2.1	Introduction . . . . .	53
2.2	Transformation de Fourier dans le cadre discret . . . . .	54
2.2.1	La transformation de Fourier discrète . . . . .	54
2.2.2	La transformée de Fourier sur $l_{\mathbb{C}}^1(\mathbb{Z})$ . . . . .	59
2.2.3	La transformée de Fourier sur $l_{\mathbb{C}}^2(\mathbb{Z})$ . . . . .	62
2.3	Transformation de Fourier sur les espaces de fonctions $2\pi$ -périodiques . . . . .	66
2.3.1	La transformation de Fourier sur $L_{\mathbb{K}}^2(\mathbb{T})$ . . . . .	67
2.3.2	La transformation de Fourier sur $L_{\mathbb{K}}^1(\mathbb{T})$ . . . . .	69

2.3.3	Les théorèmes de Dirichlet et de Féjer . . . . .	75
2.4	La transformation de Fourier des classes de fonctions intégrables sur $\mathbb{R}^n$ . . . . .	82
2.4.1	Définition, exemples, et premières propriétés . . . . .	83
2.4.2	La relation avec l'optique . . . . .	89
2.4.3	Fourier et la convolution dans $L^1_{\mathbb{K}}(\mathbb{R}^n, dx)$ . . . . .	90
2.4.4	La formule d'inversion dans le cadre $L^1$ . . . . .	91
2.4.5	L'inversion locale et la formule sommatoire de Poisson . . . . .	95
2.5	La transformation de Fourier dans $L^2_{\mathbb{C}}(\mathbb{R}^n, dx)$ . . . . .	99
2.5.1	Une isométrie de $\mathcal{G}$ dans lui-même (pour la norme $L^2$ ) . . . . .	99
2.5.2	Extension à $L^2_{\mathbb{C}}(\mathbb{R}^n, dx)$ et formule de Plancherel . . . . .	100
2.5.3	Fourier <i>versus</i> la dérivation; le principe d'incertitude d'Heisenberg . . . . .	102
2.5.4	Fonctions $L^2$ à spectre borné; le théorème de Shannon . . . . .	106
2.6	L'outil Fourier couplé avec l'analyse hilbertienne; deux exemples . . . . .	108
2.6.1	Un premier exemple; l'inversion de transformations intégrales et le CAT-Scanner . . . . .	109
2.6.2	Un second exemple : l'extrapolation des signaux de spectre borné . . . . .	111

# Chapitre 1

## Espaces de Hilbert

### 1.1 Rappel : l'espace euclidien $\mathbb{R}^n$ et l'espace hilbertien $\mathbb{C}^n$

Le  $\mathbb{R}$ -espace vectoriel de dimension finie  $\mathbb{R}^n$  est le prototype d'un *espace euclidien* dans lequel on peut envisager la géométrie fondée sur le théorème de Pythagore. L'espace  $\mathbb{R}^n$  peut être en effet équipé de la forme quadratique définie positive<sup>1</sup>

$$x \mapsto \|x\|^2 := \sum_{j=1}^n x_j^2$$

dérivant du produit scalaire canonique

$$\langle x, y \rangle = \langle y, x \rangle = \sum_{j=1}^n x_j y_j,$$

où  $(x_1, \dots, x_n)$  (*resp.*  $(y_1, \dots, y_n)$ ) désignent les coordonnées de  $x$  (*resp.*  $y$ ) dans la base canonique de  $\mathbb{R}^n$ .

Le  $\mathbb{C}$ -espace vectoriel de dimension finie  $\mathbb{C}^n$  est, lui, le prototype d'un *espace hilbertien*<sup>2</sup> dans lequel on peut envisager également la géométrie fondée sur le théorème de Pythagore. L'espace  $\mathbb{C}^n$  peut être en effet lui aussi équipé de la forme hermitienne

$$z \mapsto \|z\|^2 := \sum_{j=1}^n |z_j|^2$$

définie positive<sup>3</sup> dérivant de la forme sesquilénaire canonique présentant la symétrie hermitienne  $\langle w, z \rangle = \overline{\langle z, w \rangle}$ , à savoir le produit scalaire canonique

$$\langle z, w \rangle = \sum_{j=1}^n z_j \overline{w}_j,$$

---

<sup>1</sup>On rappelle (voir le cours de MHT301) que dire qu'une forme quadratique  $Q : \mathbb{R}^n \rightarrow \mathbb{R}$  est *positive* signifie que  $Q(x) \geq 0$  pour tout  $x \in \mathbb{R}^n$  et que dire qu'elle est *définie* signifie  $Q(x) = 0 \iff x = 0$ .

<sup>2</sup>En référence au géomètre allemand David Hilbert, 1862-1943 ; on parle aussi, pour qualifier un tel espace, d'espace *hermitien*, en référence aux travaux du théoricien des nombres français Charles Hermite, 1822-1901, qui mania aussi l'outil « produit scalaire » dans le cadre complexe.

<sup>3</sup>On rappelle (voir le cours de MHT301) que dire qu'une forme hermitienne  $H : \mathbb{C}^n \rightarrow \mathbb{R}$  est *positive* signifie que  $H(z) \geq 0$  pour tout  $z \in \mathbb{C}^n$  et que dire qu'elle est *définie* signifie  $H(z) = 0 \iff z = 0$ .

où  $(z_1, \dots, z_n)$  (resp.  $(w_1, \dots, w_n)$ ) désignent les coordonnées de  $z$  (resp.  $w$ ) dans la base canonique de  $\mathbb{C}^n$ . Ce produit scalaire est linéaire en le premier bloc de variables ( $z$ ), antilinéaire en le second ( $w$ ), ce qui est précisément la définition de la propriété de *sesquilinearité*.

Ces deux espaces ( $\mathbb{R}^n$  ou  $\mathbb{C}^n$ ) peuvent être équipés de la topologie associée à la métrique euclidienne (resp. hermitienne) définie dans le cas de  $\mathbb{R}^n$  par

$$d(x, y) = \|x - y\|$$

(resp. dans le cas de  $\mathbb{C}^n$  par  $d(z, w) = \|z - w\|$ ). Munis respectivement de cette topologie d'espace métrique, ces deux espaces  $\mathbb{R}^n$  et  $\mathbb{C}^n$  sont des espaces métriques complets : toute suite de Cauchy (relativement à la métrique) est convergente<sup>4</sup>.

Les deux résultats majeurs inhérents au cadre de l'espace euclidien  $\mathbb{R}^n$  sont :

- le théorème de Pythagore : si  $x$  et  $y$  sont deux vecteurs de  $\mathbb{R}^n$ ,

$$\|x - y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y \rangle$$

(en particulier  $\|x \pm y\|^2 = \|x\|^2 + \|y\|^2$  si et seulement si  $x \perp y$ , l'orthogonalité  $\perp$  étant ici entendue au sens du produit scalaire canonique) ;

- l'inégalité de Cauchy-Schwarz : si  $x$  et  $y$  sont deux vecteurs de  $\mathbb{R}^n$ ,

$$|\langle x, y \rangle| \leq \|x\| \times \|y\|,$$

l'égalité ayant lieu si et seulement si  $x$  et  $y$  sont  $\mathbb{R}$ -colinéaires.

Ces deux résultats majeurs se transposent ainsi dans le cadre de l'espace hilbertien  $\mathbb{C}^n$  :

- le théorème de Pythagore devient : si  $z$  et  $w$  sont deux vecteurs de  $\mathbb{C}^n$ ,

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

(en particulier  $\|z \pm w\|^2 = \|z\|^2 + \|w\|^2$  si et seulement si  $z \perp w$ , l'orthogonalité étant encore ici entendue au sens du produit scalaire canonique) ;

- l'inégalité de Cauchy-Schwarz devient : si  $z$  et  $w$  sont deux vecteurs de  $\mathbb{C}^n$ ,

$$|\langle z, w \rangle| \leq \|z\| \times \|w\|,$$

l'égalité ayant lieu si et seulement si  $z$  et  $w$  sont  $\mathbb{C}$ -colinéaires.

Nous reviendrons dans la section suivante (dans un cadre élargi) sur ces deux propriétés (et d'autres) qui ne font intervenir que les notions algébriques liées à la définition d'une norme dérivant d'un produit scalaire.

On connaît<sup>5</sup> l'intérêt (dans le cadre tant de  $\mathbb{R}^n$  que de  $\mathbb{C}^n$ ) des outils que représentent le « produit scalaire », le théorème de Pythagore et l'inégalité de Cauchy-Schwarz. Ces outils permettent, par exemple dans le contexte euclidien de  $\mathbb{R}^n$ , de réaliser l'opération suivante : étant donné un  $\mathbb{R}$ -sous-espace affine  $D$  de  $\mathbb{R}^n$  (de sous-espace vectoriel sous-jacent  $F$ ), on peut définir un *opérateur de projection orthogonale sur  $D$* , schéma que le petit diagramme bien commode ci-dessous (figure 1.1) soutend (ici lorsque  $D$  est une droite du plan  $\mathbb{R}^2$ ) :

<sup>4</sup>D'ailleurs, le choix de la métrique, pourvu que celle-ci dérive d'une norme, est irrelevant puisqu'en dimension finie on sait que deux normes sont toujours équivalentes.

<sup>5</sup>Voir le cours d'algèbre de L2 (UE MHT301).

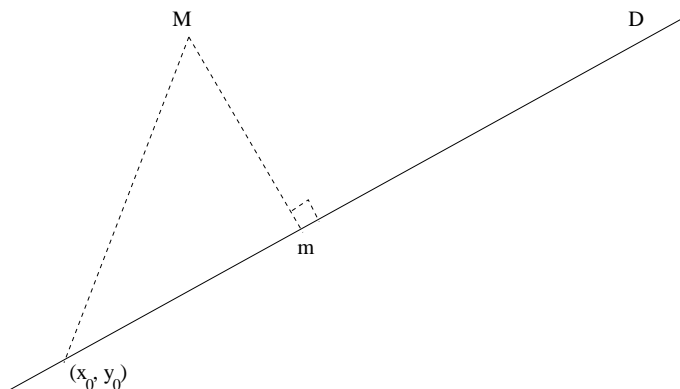


FIG. 1.1 – La projection orthogonale sur une droite  $D$  de  $\mathbb{R}^2$  passant par  $(x_0, y_0)$

On a utilisé ici le modèle de  $\mathbb{R}^2$  pour faire un dessin, mais l'on aurait pu tout aussi bien reproduire ce schéma dans le cadre abstrait. Si

$$D = M_0 + F,$$

où  $F$  (sous-espace vectoriel de  $\mathbb{R}^n$  sous-jacent à  $D$ ) est un sous-espace vectoriel de dimension  $k$  engendré par  $v_1, \dots, v_k$  et que  $(v_1, \dots, v_n)$  est une base orthonormée de  $\mathbb{R}^n$  pour le produit scalaire canonique, la projection orthogonale sur le sous-espace affine  $D$

$$M \longmapsto \text{Proj}_D(M)$$

est l'application qui à  $M$  associe le point

$$M_0 + \xi_1 v_1 + \dots + \xi_k v_k,$$

où  $(\xi_1, \dots, \xi_k, \xi_{k+1}, \dots, \xi_n)$  sont les coordonnées du vecteur  $\overrightarrow{M_0 M}$  dans la base orthonormée  $(v_1, \dots, v_n)$ . Cette opération de prise de « projection orthogonale » sur un sous-espace affine  $D$  est à la base d'une démarche capitale en mathématiques appliquées, la *méthode des moindres carrés* : étant donnée une liste de contraintes dans  $\mathbb{R}^n$  matérialisée par un jeu d'équations affines

$$\begin{aligned} a_{1,1}x_1 + \dots + a_{1,n}x_n &= b_1 \\ &\vdots \\ a_{m,1}x_1 + \dots + a_{m,n}x_n &= b_m, \end{aligned} \tag{1.1}$$

où les  $a_{j,k}$  et les  $b_j$ ,  $j = 1, \dots, m$ ,  $k = 1, \dots, n$  sont des scalaires réels, on peut calculer, étant donné un état  $x = (x_1, \dots, x_n)$  donné dans  $\mathbb{R}^n$ , l'unique point

$$\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_n)$$

le plus proche de  $x$  dans  $\mathbb{R}^n$  (pour la distance euclidienne) qui satisfasse au jeu de contraintes (1.1). Ce point répond donc à la double exigence :

- d'une part  $\tilde{x}$  satisfait les contraintes (1.1) ;
- d'autre part  $\|x - \tilde{x}\|$  est le minimum des nombres  $\|x - y\|$ , pris pour tous les  $y$  satisfaisant les contraintes (1.1).

Ce point  $\tilde{x}$  est exactement la projection orthogonale  $\text{Proj}_D(x)$  de  $x$  sur le  $\mathbb{R}$ -sous-espace affine  $D$  (de dimension au moins  $n - m$  car il y a  $m$  contraintes, certaines pouvant éventuellement être redondantes) défini par les équations affines (1.1). Le même raisonnement vaut dans le cadre complexe et cette méthode des moindres carrés dont nous serons souvent appelés à parler dans ce cours est l'illustration même de la force du théorème de Pythagore.

Les mêmes idées se transposent au cadre de l'espace hermitien  $\mathbb{C}^n$ .

Ce sont ces idées qui soutendent par exemple des démarches algorithmiques telles par exemple celle qu'illustrent les deux exemples suivants :

**Exemple 1.1.** Soient  $F_1$  et  $F_2$  deux sous-espaces vectoriels de  $\mathbb{R}^n$  (ou de  $\mathbb{C}^n$ ) tels que  $F_1 \cap F_2^\perp = \{0\}$  et  $s$  un élément inconnu du sous-espace  $F_1$  ; si l'on connaît la projection orthogonale  $s_0$  de  $s$  sur  $F_2$ , la démarche algorithmique « en zig-zag », basée sur l'utilisation d'un jeu de projections orthogonales itérées (alternativement sur  $F_1$  et sur le sous-espace affine  $s_0 + F_2^\perp$ ) conduit à une approximation de l'élément inconnu  $s$  (voir la figure 1.2 ci-dessous) <sup>6</sup>.

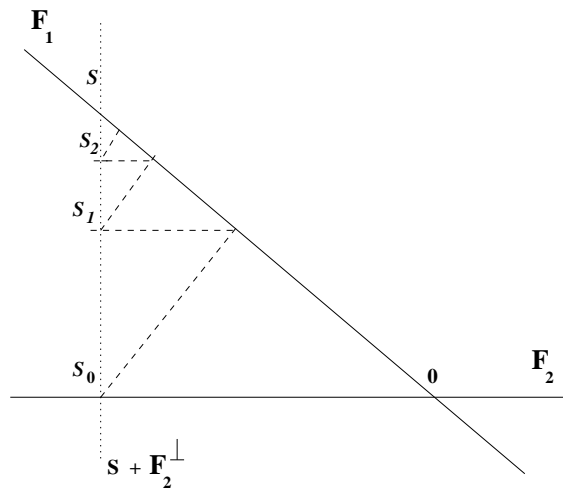


FIG. 1.2 – Itération de projections (exemple 1.1)

**Exemple 1.2.** Soient  $D_1, \dots, D_N$   $N$  sous-espaces affines de  $\mathbb{R}^n$  (ou  $\mathbb{C}^n$ ) s'intersectant en un unique point  $A$ . Le mécanisme qui consiste à partir d'un point  $M$  quelconque de l'espace et (comme sur la figure 1.3 ci-dessous) à projeter ce point orthogonalement sur  $F_1$ , puis à projeter le point obtenu orthogonalement sur  $F_2$ , etc., jusqu'à épuisement de toute la liste de sous-espaces, puis à recommencer, conduit (via une démarche « en escargot ») à une approximation du point  $A$  <sup>7</sup>.

<sup>6</sup>Nous reviendrons sur la preuve de ce résultat dans un contexte beaucoup plus général plus tard dans le cours.

<sup>7</sup>Ceci sera établi dans un contexte beaucoup plus général plus loin dans le cours.



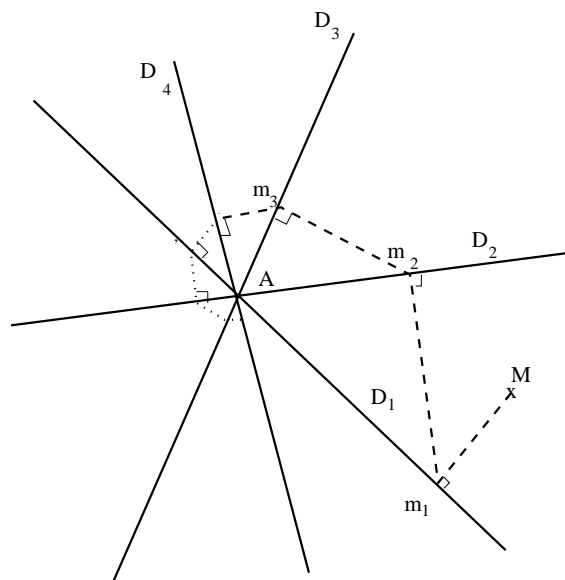


FIG. 1.3 – Itération de projections (exemple 1.2)

## 1.2 Le saut de la dimension finie à la dimension infinie

Autant le théorème de Pythagore et l'inégalité de Cauchy-Schwarz ne relevaient que des propriétés algébriques (seul le fait que la forme soit définie positive s'avère important pour assurer l'inégalité de Cauchy-Schwarz et la clause d'égalité dans cette inégalité), autant la démarche de recherche de projection orthogonale sur un sous-espace utilise (mais de manière détournée) le fait que  $(\mathbb{R}^n, d)$  et  $(\mathbb{C}^n, d)$  sont des espaces complets. Le cadre de la dimension finie est un peu trompeur car on se rend pas compte de ce phénomène !

Pour se convaincre de cela, plaçons nous dans le cadre plus troublant d'un espace de fonctions (de dimension cette fois infinie). Le  $\mathbb{R}$ -espace vectoriel  $E = \mathcal{C}([0, 1], \mathbb{R})$  des fonctions continues sur  $[0, 1]$ , à valeurs réelles, peut lui aussi être équipé d'un produit scalaire, à savoir

$$\langle f, g \rangle := \int_0^1 f(t)g(t) dt. \quad (1.2)$$

On retrouvera beaucoup ce produit scalaire par la suite ; il a une interprétation physique car on peut le relier à la notion d'énergie (l'énergie est toujours, pensez par exemple à l'énergie cinétique  $E = mv^2/2$ , une expression dépendant de manière non linéaire mais quadratique des phénomènes physiques qui la soutendent, ici en l'occurrence la vitesse  $v$ ). Dans ce  $\mathbb{R}$ -espace vectoriel  $E$ , on peut considérer le sous-espace affine

$$D := \left\{ f \in E, \int_0^{1/2} f(t)dt - \int_{1/2}^1 f(t)dt = 1 \right\}.$$

Si  $E$  est équipé de la topologie (d'espace métrique) associée à la distance

$$d_\infty(f, g) = \sup_{t \in [0, 1]} |f(t) - g(t)|$$

correspondant à la topologie de la convergence uniforme sur  $[0, 1]$ , on constate immédiatement que  $D$  est bien un sous-ensemble fermé de  $E$ . Equipé de cette distance  $d_\infty$ ,  $E$  est d'ailleurs un espace métrique complet. Il faut noter cependant que  $d_\infty$  n'est pas une distance associée à une norme dérivant d'un produit scalaire. La distance  $d_\infty$  ne nous intéresse donc pas autant que la distance euclidienne

$$d(f, g) = \|f - g\|_2 := \sqrt{\int_0^1 |f(t) - g(t)|^2 dt},$$

métrique dérivant associée au produit scalaire (1.2) pour laquelle on dispose certes du théorème de Pythagore, mais pour laquelle, malheureusement,  $E$  n'est plus un  $\mathbb{R}$ -espace vectoriel complet (ni d'ailleurs  $D$  un sous-espace affine fermé).

On peut naturellement se demander malgré tout s'il est concevable de définir un opérateur de projection orthogonale sur  $D$ . Il n'en est rien. En effet, on montre en exercice (voir le paragraphe ci-dessous) qu'il ne saurait exister aucun élément  $f_0 \in D$  en lequel la fonction

$$f \in D \longmapsto \|f\|_2^2 := \int_0^1 |f(t)|^2 dt$$

atteigne son minimum sur  $D$ ; comme l'existence d'une telle fonction  $f_0$  équivaut au fait que 0 admette une « projection orthogonale » sur  $D$  (il existe un point de  $D$  réalisant la distance euclidienne  $d$  entre 0 et  $D$ ), la non existence dans  $D$  d'un tel  $f_0$  rend donc bien compte du fait qu'il s'avère impossible de « projeter orthogonalement » l'origine 0 sur  $D$ .

**Exercice 1.1.** Montrons que  $f_0$  (élément de  $D$  tel que  $\|f_0\|_2$  soit minimal) ne peut exister. Soit  $h$  la fonction définie sur  $[0, 1]$  par

$$h(t) = \begin{cases} 1 & \text{si } t \in [0, 1/2[ \\ 0 & \text{si } t = 1/2 \\ -1 & \text{si } t \in ]1/2, 1]. \end{cases}$$

Même si

$$\int_0^{1/2} h(t) dt - \int_{1/2}^1 h(t) dt = 1,$$

la fonction  $h$  n'est pas dans  $D$  (ce n'est en effet pas une fonction continue!); néanmoins, on peut l'approcher (au sens de la distance  $d$ ) par la suite de fonctions  $(h_n)_{n \geq 3}$ , où

$$h_n(t) = \begin{cases} \frac{n}{n-1} & \text{si } t \in [0, 1/2 - 1/n[ \\ \frac{n}{n-1} \times (n(1/2 - t)) & \text{si } t \in [1/2 - 1/n, 1/2 + 1/n] \\ -\frac{n}{n-1} & \text{si } t \in ]1/2 + 1/n, 1]. \end{cases}$$

qui, elle, on le vérifie aisément, est une suite d'éléments de  $D$ . Si  $f_0$  existait, on aurait donc certainement

$$\|f_0\|_2^2 \leq \|h_n\|_2^2 \quad \forall n \geq 3,$$

et, par conséquent, en prenant la limite lorsque  $n$  tend vers  $+\infty$  (et en utilisant par exemple le théorème de convergence dominée)

$$\|f_0\|_2^2 \leq \lim_{n \rightarrow +\infty} \|h_n\|_2^2 = \|h\|_2^2 = \frac{1}{2} + \frac{1}{2} = 1.$$

On remarque que le trinôme

$$\lambda \longmapsto \int_0^1 (f_0(t) + \lambda h(t))^2 dt = \lambda^2 \|h\|^2 + 2\lambda \int_0^1 f_0(t)h(t) dt + \|f_0\|_2^2 = \lambda^2 + 2\lambda + \|f_0\|_2^2$$

a pour discriminant réduit  $1 - \|f_0\|_2^2 \geq 0$ , donc admet au moins une racine réelle  $\lambda_0$ ; ceci montre que

$$\int_0^1 (f_0(t) + \lambda_0 h(t))^2 dt = 0,$$

d'où l'on déduirait  $f_0 \equiv -\lambda_0 h$   $dt$ -presque partout sur  $[0, 1]$ , ce qui est impossible car  $f_0$  est continue tandis que  $-\lambda_0 h$  ne l'est pas! On conclut donc (par l'absurde) à l'impossibilité de l'existence de  $f_0$ , élément de  $D$  réalisant la distance (au sens de la métrique euclidienne  $d$ ) de l'origine à  $D$ .

Ce qui manque ici au  $\mathbb{R}$ -espace vectoriel  $E$  pour que l'on puisse pleinement profiter (comme dans le cas de  $\mathbb{R}^n$  et  $\mathbb{C}^n$ ) des atouts de l'algorithmique pythagoricienne est simplement le fait que  $E$  équipé de la métrique  $d$  n'est pas un espace métrique complet!

La leçon que nous retenons de l'exemple proposé dans ce paragraphe (exemple sortant du cadre de la dimension finie) est qu'il faut prendre garde (lorsque l'on sort du cadre des  $\mathbb{R}$  ou  $\mathbb{C}$ -espaces vectoriels de dimension finie tels  $\mathbb{R}^n$  et  $\mathbb{C}^n$ ) à coupler le modèle algébrique avec des contraintes relevant de l'analyse (en l'occurrence la complétude pour la distance euclidienne du  $\mathbb{R}$  ou  $\mathbb{C}$ -espace vectoriel sur lequel on travaille) si l'on espère profiter pleinement des outils que met à notre disposition l'introduction sur un  $\mathbb{R}$  ou  $\mathbb{C}$ -espace vectoriel, d'un produit scalaire (Pythagore, Cauchy-Schwarz, possibilité de projeter orthogonalement sur un sous-espace fermé, etc.).

## 1.3 Les notions d'espace de Hilbert (réel ou complexe)

### 1.3.1 Le cadre Hilbert

La section 1.1 nous guide vers les deux définitions suivantes, situant le cadre de travail dans lequel on entend mettre en pratique les idées inspirées du maniement de l'orthogonalité dans  $\mathbb{R}^n$  ou  $\mathbb{C}^n$  :

**Définition 1.1** *Un espace de Hilbert réel est un  $\mathbb{R}$ -espace vectoriel, couplé avec la donnée d'un produit scalaire*

$$(h_1, h_2) \in H \times H \longmapsto \langle h_1, h_2 \rangle \in \mathbb{R},$$

*c'est-à-dire une application bilinéaire à valeurs réelles*

- *symétrique, c'est-à-dire telle que  $\langle h_1, h_2 \rangle = \langle h_2, h_1 \rangle$  pour tous  $h_1, h_2$  dans  $H$  ;*
- *telle que  $\langle h, h \rangle \geq 0$  pour tout  $h \in H$  (clause de positivité) ;*
- *telle que  $\langle h, h \rangle = 0$  si et seulement si  $h = 0$  (clause de « définition »).*

*De plus, on exige que l'espace vectoriel normé  $(H, \|\cdot\|)$ , où*

$$\|h\| := \sqrt{\langle h, h \rangle}$$

*soit un espace normé complet (c'est-à-dire un espace de Banach <sup>8</sup>).*

---

<sup>8</sup>En référence au mathématicien polonais Stefan Banach (1892-1945) qui explora toutes les facettes et tout l'intérêt de ce concept, un *espace de Banach* est par définition un  $\mathbb{R}$  ou  $\mathbb{C}$ -espace vectoriel normé où toute suite de Cauchy (au sens de la norme) est convergente.

**Définition 1.2** *Un espace de Hilbert complexe est un  $\mathbb{C}$ -espace vectoriel, couplé avec la donnée d'un produit scalaire*

$$(h_1, h_2) \in H \times H \longmapsto \langle h_1, h_2 \rangle \in \mathbb{C},$$

*c'est-à-dire une application sesquilinéaire<sup>9</sup> à valeurs complexes*

- *présentant la symétrie hermitienne, c'est-à-dire :  $\langle h_1, h_2 \rangle = \overline{\langle h_2, h_1 \rangle}$  pour tout  $h_1, h_2$  dans  $H$  ;*
- *telle que  $\langle h, h \rangle \geq 0$  pour tout  $h \in H$  (clause de positivité) ;*
- *telle que  $\langle h, h \rangle = 0$  si et seulement si  $h = 0$  (clause de « définition »).*

*De plus, on exige que l'espace vectoriel normé  $(H, \|\cdot\|)$ , où*

$$\|h\| := \sqrt{\langle h, h \rangle}$$

*soit un espace normé complet (c'est-à-dire un espace de Banach).*

### 1.3.2 Un formulaire géométrique

En utilisant la bilinéarité et la symétrie du produit scalaire dans le cas réel, la sesquilinearité et la symétrie hermitienne du produit scalaire dans le cas complexe, on voit que l'on a l'identité importante

$$\|h_1 + h_2\|^2 + \|h_1 - h_2\|^2 = 2(\|h_1\|^2 + \|h_2\|^2) \quad \forall h_1, h_2 \in H \quad (1.3)$$

(dite *loi du parallélogramme*<sup>10</sup>) et, suivant que l'on est dans le cas réel ou dans le cas complexe, soit

$$\|h_1 + h_2\|^2 = \|h_1\|^2 + \|h_2\|^2 + 2\langle h_1, h_2 \rangle \quad \forall h_1, h_2 \in H, \quad (1.4)$$

soit

$$\|h_1 + h_2\|^2 = \|h_1\|^2 + \|h_2\|^2 + 2 \operatorname{Re} \langle h_1, h_2 \rangle \quad \forall h_1, h_2 \in H ; \quad (1.5)$$

les identités (1.4) et (1.5) sont dites *identités de Pythagore*. La loi du parallélogramme induit la *formule de la médiane* :

$$\left\| h_0 - \frac{h_1 + h_2}{2} \right\|^2 + \frac{\|h_1 - h_2\|^2}{4} = \frac{(\|h_0 - h_1\|^2 + \|h_0 - h_2\|^2)}{2} \quad \forall h_0, h_1, h_2 \in H ; \quad (1.6)$$

il suffit pour le voir d'appliquer (1.3) en remplaçant  $h_j$  par  $(h_0 - h_j)/2$ ,  $j = 1, 2$ . La positivité du produit scalaire implique dans le cas réel, si  $h_1$  et  $h_2$  sont deux éléments de  $H$ , que la fonction trinôme

$$\lambda \in \mathbb{R} \longmapsto \langle h_1 + \lambda h_2, h_1 + \lambda h_2 \rangle = \lambda^2 \|h_2\|^2 + 2\lambda \langle h_1, h_2 \rangle + \|h_1\|^2 \quad (1.7)$$

est positive ou nulle sur  $\mathbb{R}$ , dans le cas complexe que la fonction trinôme complexe

$$\lambda \in \mathbb{C} \longmapsto \langle h_1 + \lambda h_2, h_1 + \lambda h_2 \rangle = |\lambda|^2 \|h_2\|^2 + 2 \operatorname{Re} (\bar{\lambda} \langle h_1, h_2 \rangle) + \|h_1\|^2 \quad (1.8)$$

<sup>9</sup>Ceci signifie  $\mathbb{C}$ -linéaire en  $h_1$ ,  $\mathbb{C}$ -antilinéaire en  $h_2$ .

<sup>10</sup>Ainsi dénotée car elle s'appuie sur la figure du parallélogramme construit à partir des deux vecteurs  $h_1$  et  $h_2$  (on fera un petit dessin pour s'en convaincre).

prend ses valeurs dans  $[0, \infty[$ . Ceci implique dans le cas réel que le discriminant du trinôme (1.7) est négatif, d'où l'inégalité

$$|\langle h_1, h_2 \rangle| \leq \|h_1\| \times \|h_2\| \quad \forall h_1, h_2 \in H, \quad (1.9)$$

dite de Cauchy-Schwarz<sup>11</sup>. En spécifiant  $\lambda = te^{i\alpha}$ ,  $\alpha \in \mathbb{R}$  dans (1.8) et en faisant varier  $\alpha$ , on constate, dans le cas complexe, que l'on a la même inégalité de Cauchy-Schwarz (1.9). Il est important de noter que cette inégalité (1.9) est stricte dès que  $h_1$  et  $h_2$  ne sont pas colinéaires. Si  $h_1$  et  $h_2$  sont colinéaires, l'inégalité (1.9) devient une égalité.

Dans un espace de Hilbert, on peut définir, non pas l'angle de deux vecteurs non nuls, mais au moins leur cosinus, ce de la manière suivante :

$$\cos(\text{angle}(h_1, h_2)) := \frac{\operatorname{Re} \langle h_1, h_2 \rangle}{\|h_1\| \|h_2\|}.$$

Si  $F_1$  et  $F_2$  sont deux sous-espaces vectoriels de  $H$ , on peut définir le cosinus de l'angle entre  $F_1$  et  $F_2$  en posant

$$\cos(\text{angle}(F_1, F_2)) := \sup_{h_1 \in F_1 \setminus \{0\}, h_2 \in F_2 \setminus \{0\}} \frac{|\langle h_1, h_2 \rangle|}{\|h_1\| \|h_2\|} \in [0, 1]$$

(à cause de l'inégalité de Cauchy-Schwarz).

Rappelons, pour mémoire, les trois manières équivalentes de caractériser la complétude d'un  $\mathbb{R}$  (ou  $\mathbb{C}$ ) espace vectoriel normé  $(H, \|\cdot\|)$  :

- toute suite de Cauchy (pour la métrique  $d(x, y) = \|x - y\|$ ) est convergente ;
- si  $(A_k)_{k \in \mathbb{N}}$  est une suite décroissante (au sens de l'inclusion) de sous ensembles fermés de  $H$  dont le diamètre (au sens de la distance  $d$ ) tend vers 0, l'intersection de tous les  $A_k$  est non vide et réduite à un singleton ;
- toute série  $(\sum_n h_n)$ ,  $h_n \in H$ , telle que  $\sum_{n=0}^{\infty} \|h_n\| < \infty$ , converge dans  $H$ .

## 1.4 Les exemples majeurs

### 1.4.1 Les exemples $\mathbb{R}^n$ et $\mathbb{C}^n$

Le premier exemple d'espace de Hilbert réel est bien sûr  $\mathbb{R}^n$ , équipé de son produit scalaire canonique (voir la section 1.1). Le premier exemple d'espace de Hilbert complexe est  $\mathbb{C}^n$ , équipé de son produit scalaire canonique (voir aussi la section 1.1).

On peut penser ces deux espaces comme des espaces de fonctions, puisque l'on peut identifier le vecteur  $x$  avec la fonction  $X$  de  $\{1, \dots, n\}$  dans  $\mathbb{R}$  (*resp.*  $\mathbb{C}$ ) qui à l'entier  $k$  associe le scalaire  $X(k) = x_k$  ( $k$ -ème coordonnée de  $x$  dans la base canonique). Une telle fonction est un *signal digital réel* (*resp.* complexe).

Si  $n_1$  et  $n_2$  sont deux entiers strictement positifs, l'espace des matrices réelles (*resp.* complexes)  $I$  de type  $n_1 \times n_2$ , équipé du produit scalaire

$$\langle I_1, I_2 \rangle := \sum_{k_1=1}^{n_1} \sum_{k_2=1}^{n_2} I_1(k_1, k_2) I_2(k_1, k_2)$$

<sup>11</sup>Les noms de l'analyste français Augustin Cauchy (1789-1857) et du géomètre allemand Hermann Schwarz (1843-1921) sont attachés à cette inégalité, que dans l'ex culture soviétique, on connaît sous la terminologie d'inégalité de Cauchy-Bunyakovsky.

dans le cas réel ou

$$\sum_{k_1=1}^{n_1} \sum_{k_2=1}^{n_2} I_1(k_1, k_2) \overline{I_2(k_1, k_2)}$$

dans le cas complexe est un espace de Hilbert (ce n'est qu'une autre présentation de  $\mathbb{R}^{n_1 \times n_2}$  ou  $\mathbb{C}^{n_1 \times n_2}$  comme un espace de matrices); un élément de cet espace est une *image digitale* (réelle ou complexe suivant le contexte), de type  $(n_1, n_2)$ . Chaque point marqué  $I(k_1, k_2)$  d'un tableau  $[I(k_1, k_2)]_{1 \leq k_j \leq n_j}$  représente l'intensité ou encore la *brillance* (mesurée avec une unité convenue) à ce que l'on conviendra d'appeler le *pixel*  $(k_1, k_2)$ .

On peut aussi imaginer des espaces de tableaux de tableaux : une image couleur par exemple se code dans le système RGB<sup>12</sup> comme un tableau  $n_1 \times n_2$  de vecteurs  $(r, g, b)$  d'entiers entre 0 et 255 ; la configuration est bien toujours du type  $\mathbb{R}^n$ , avec  $n = 3n_1n_2$ , mais la présentation est différente.

Le produit scalaire entre deux vecteurs de  $\mathbb{R}^n$  (ou de  $\mathbb{C}^n$ ), vecteurs que l'on peut considérer comme des informations digitales de longueur  $n$ , est aussi appelé *corrélation* de ces deux informations ; si  $x$  et  $y$  ont même norme, plus  $\langle x, y \rangle$  est grand, plus les informations sont dites *corrélées* entre elles. Le carré de la norme euclidienne d'un tel vecteur  $x$  de  $\mathbb{R}^n$  ou  $z$  de  $\mathbb{C}^n$  est dit *énergie* du signal digital  $x$  (ou  $z$ ).

### 1.4.2 Les espaces $l_{\mathbb{R}}^2(I)$ et $l_{\mathbb{C}}^2(I)$

Si  $I$  est un ensemble d'indices (abstrait), on peut considérer dans le  $\mathbb{R}$ -espace vectoriel  $\mathbb{R}^I = \mathcal{F}(I, \mathbb{R})$  des fonctions de  $I$  dans  $\mathbb{R}$  le sous-espace  $l_{\mathbb{R}}^2(I)$  constitué des éléments  $(x_i)_{i \in I}$  tels que

$$\sup_{\{J \subset I; \#J < \infty\}} \left( \sum_{i \in J} |x_i|^2 \right) := \|(x_i)_{i \in I}\|^2 < \infty.$$

On constate alors (en utilisant la définition de la borne supérieure et l'inégalité de Cauchy-Schwarz dans  $\mathbb{R}$ ) que si  $(x_i)_{i \in I}$  et  $(y_i)_{i \in I}$  sont deux éléments de  $l_{\mathbb{R}}^2(I)$ , alors pour tout  $\epsilon > 0$ , il existe une partie finie  $J(\epsilon) \subset I$  telle que pour toute partie finie  $K$  de  $I$  d'intersection vide avec  $J(\epsilon)$ , on ait

$$\left| \sum_{i \in K} x_i y_i \right| \leq \sqrt{\sum_{i \in K} |x_i|^2} \sqrt{\sum_{i \in K} |y_i|^2} < \epsilon. \quad (1.10)$$

Soient donc  $(x_i)_{i \in I}$  et  $(y_i)_{i \in I}$  deux éléments du  $\mathbb{R}$ -espace vectoriel  $l_{\mathbb{R}}^2(I)$ . Pour  $n \in \mathbb{N}^*$ , notons  $C_n$  l'adhérence dans  $\mathbb{R}$  de l'ensemble des nombres

$$\sum_{i \in J} x_i y_i,$$

où  $J$  est une partie finie de  $I$  contenant  $J(1) \cup J(1/2) \cup \dots \cup J(1/n)$ . A cause de (1.10) avec  $\epsilon = 1/n$  (pour toute partie finie  $K$  de  $I$  ne rencontrant pas  $\bigcup_{k=1}^n J(1/k)$ ), on constate que le diamètre de  $C_n$  (pour la distance associée à la valeur absolue sur  $\mathbb{R}$ ) tend vers 0. Comme  $\mathbb{R}$  est complet, l'intersection des fermés  $C_n$  est non vide et

<sup>12</sup> « Red-Green-Blue ».

réduite à un singleton ; on définit ainsi le nombre réel (comme étant précisément ce singleton)

$$\sum_{\iota \in I} x_{\iota} y_{\iota}$$

et on montre que le nombre réel ainsi défini est l'unique nombre réel  $S$  caractérisé par la propriété suivante : *pour tout  $\epsilon > 0$ , il existe une partie finie  $\tilde{J}(\epsilon)$  telle pour toute partie finie  $J$  de  $I$  contenant  $\tilde{J}(\epsilon)$ ,*

$$\left| S - \sum_{\iota \in J} x_{\iota} y_{\iota} \right| < \epsilon.$$

Ceci nous amène naturellement à la notion de *famille sommable* de nombres réels ou complexes.

**Définition 1.3** Une famille  $(u_{\iota})$  de nombres réels ou complexes indexée par un ensemble d'indices  $I$  est dite *sommable* et de somme  $S$  si et seulement si, pour tout  $\epsilon > 0$ , il existe une partie finie  $\tilde{J}(\epsilon)$  telle que pour toute partie finie  $J$  de  $I$  contenant  $\tilde{J}(\epsilon)$ ,

$$\left| S - \sum_{\iota \in J} u_{\iota} \right| < \epsilon.$$

Ici la famille  $(x_{\iota} y_{\iota})_{\iota \in I}$  est sommable dès que  $(x_{\iota})_{\iota \in I}$  et  $(y_{\iota})_{\iota \in I}$  sont deux éléments de  $l_{\mathbb{R}}^2(I)$  et l'on définit un produit scalaire sur  $l_{\mathbb{R}}^2(I)$  en posant précisément :

$$\left\langle (x_{\iota})_{\iota \in I}, (y_{\iota})_{\iota \in I} \right\rangle := \sum_{\iota} x_{\iota} y_{\iota}.$$

**Proposition 1.1** Le  $\mathbb{R}$ -espace vectoriel  $l_{\mathbb{R}}^2(I)$  équipé du produit scalaire

$$\left\langle (x_{\iota})_{\iota \in I}, (y_{\iota})_{\iota \in I} \right\rangle := \sum_{\iota} x_{\iota} y_{\iota}.$$

est un espace de Hilbert.

**Preuve.** Si  $\left( (x_{\iota}^{(n)})_{\iota \in I} \right)_n$  est une suite de Cauchy dans  $l_{\mathbb{R}}^2(I)$ , on a

$$\forall \epsilon > 0, \exists N(\epsilon), \forall p, q \geq N(\epsilon), \sup_{\{J \subset I; \#J < \infty\}} \left( \sum_{\iota \in J} |x_{\iota}^{(p)} - x_{\iota}^{(q)}|^2 \right) < \epsilon. \quad (1.11)$$

Pour chaque  $\iota$  dans  $I$ , la suite de scalaires  $(x_{\iota}^{(n)})_n$  est donc de Cauchy, donc convergente dans  $\mathbb{R}$  puisque  $\mathbb{R}$  est complet vers une limite  $x_{\iota}^{(\infty)}$ . En « gelant »  $p \geq N(\epsilon)$  et en faisant courir  $q$  vers l'infini dans (1.11), on constate que

$$\forall \epsilon > 0, \exists N(\epsilon), \forall p \geq N(\epsilon), \sup_{\{J \subset I; \#J < \infty\}} \left( \sum_{\iota \in J} |x_{\iota}^{(p)} - x_{\iota}^{(\infty)}|^2 \right) < \epsilon.$$

Il résulte de cela que  $(x_{\iota}^{(\infty)})_{\iota \in I}$  est bien un élément de  $l_{\mathbb{R}}^2(I)$  et que cet élément est la limite dans  $l_{\mathbb{R}}^2(I)$  (lorsque  $n$  tend vers  $+\infty$  et pour la norme  $\| \cdot \|$ ) de la suite  $\left( (x_{\iota}^{(n)})_{\iota \in I} \right)_n$ . La complétude de  $l_{\mathbb{R}}^2(I)$  pour la norme  $\| \cdot \|$  est donc assurée.  $\diamond$

Toujours à partir du même ensemble d'indices  $I$ , on peut considérer dans le  $\mathbb{C}$ -espace vectoriel  $\mathbb{C}^I = \mathcal{F}(I, \mathbb{C})$  des fonctions de  $I$  dans  $\mathbb{C}$  le sous-espace  $l_{\mathbb{C}}^2(I)$  constitué des éléments  $(z_\iota)_{\iota \in I}$  tels que

$$\sup_{\{J \subset I; \#J < \infty\}} \left( \sum_{\iota \in J} |z_\iota|^2 \right) := \|(z_\iota)_{\iota \in I}\|^2 < \infty.$$

Exactement comme dans le cadre réel (toujours en utilisant l'inégalité de Cauchy-Schwarz), on constate que si  $(z_\iota)_{\iota \in I}$  et  $(w_\iota)_{\iota \in I}$  sont deux éléments de  $l_{\mathbb{C}}^2(I)$ , la famille  $(z_\iota \bar{w}_\iota)_{\iota \in I}$  est sommable dans  $\mathbb{C}$ . On définit donc un produit scalaire dans  $l_{\mathbb{C}}^2(I)$  en posant

$$\left\langle (z_\iota)_{\iota \in I}, (w_\iota)_{\iota \in I} \right\rangle := \sum_{\iota} z_\iota \bar{w}_\iota.$$

La proposition suivante se prouve exactement comme la proposition 1.1 (mais en utilisant la complétude de  $\mathbb{C}$  cette fois).

**Proposition 1.2** *Le  $\mathbb{C}$ -espace vectoriel  $l_{\mathbb{C}}^2(I)$  équipé du produit scalaire*

$$\left\langle (z_\iota)_{\iota \in I}, (w_\iota)_{\iota \in I} \right\rangle := \sum_{\iota} z_\iota \bar{w}_\iota.$$

*est un espace de Hilbert.*

Les deux espaces de Hilbert  $l_{\mathbb{R}}^2(I)$  (dans le cadre réel) et  $l_{\mathbb{C}}^2(I)$  (dans le cadre complexe) seront, on le verra plus tard, des modèles pour tout espace de Hilbert réel (dans ce cas le modèle est  $l_{\mathbb{R}}^2(I)$ ) ou complexe (auquel cas le modèle est  $l_{\mathbb{C}}^2(I)$ ). Seront plus intéressants pour nous les modèles correspondant au cas où  $I$  est dénombrable<sup>13</sup>. Les cas  $I = \mathbb{Z}$  ou  $I = \mathbb{N}$  seront des cas d'exemples de Hilbert très importants au niveau des applications.

### 1.4.3 Les espaces $L_{\mathbb{K}}^2(U, dx)$ , $U$ partie mesurable de $\mathbb{R}^n$ , $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$

Soit  $\mathcal{L}(\mathbb{R}^n)$  la tribu de Lebesgue sur  $\mathbb{R}^n$  (plus petite tribu contenant la tribu borélienne et tous les ensembles de mesure de Lebesgue nulle)<sup>14</sup>.

Si  $U$  est un sous-ensemble mesurable de  $\mathbb{R}^n$  (par exemple un sous-ensemble ouvert, un sous-ensemble fermé,  $\mathbb{R}^n$  tout entier, ...), l'espace  $\mathcal{L}_{\mathbb{R}}^2(U, \mathcal{L}(\mathbb{R}^n)|_U, dx)$  des fonctions  $f : U \rightarrow \mathbb{R}$  mesurables (relativement à la tribu  $\mathcal{L}(\mathbb{R}^n)$  restreinte à  $U$ ) telles que

$$\int_U |f(x)|^2 dx < \infty$$

( $dx$  est ici la mesure de Lebesgue) peut être équipé d'une forme bilinéaire symétrique positive

$$(f, g) \mapsto \langle f, g \rangle := \int_U f(x)g(x) dx.$$

<sup>13</sup>Remarquons toutefois que si  $(x_\iota)_{\iota \in I} \in l_{\mathbb{R}}^2(I)$ , l'ensemble  $\{\iota \in I; |x_\iota| \geq 1/p\}$  (pour  $p \in \mathbb{N}^*$  fixé) est fini, de cardinal au plus  $p^2 \|(x_\iota)_{\iota \in I}\|^2$ ; ceci montre que l'ensemble des indices  $\iota$  tels que  $x_\iota \neq 0$  est forcément au plus fini ou dénombrable. Ceci vaut aussi pour un élément donné  $(z_\iota)_{\iota \in I}$  de  $l_{\mathbb{C}}^2(I)$  : l'ensemble des indices  $\iota$  tels que  $z_\iota \neq 0$  est au plus fini ou dénombrable.

<sup>14</sup>On pourrait d'ailleurs tout aussi bien se contenter de la tribu borélienne  $\mathcal{B}(\mathbb{R}^n)$  à la place de la tribu de Lebesgue  $\mathcal{L}(\mathbb{R}^n)$ ; l'avantage que présente la tribu de Lebesgue sur la tribu borélienne est la propriété de complétude (section 1.6 du cours de MHT512).



En effet, on sait (inégalité de Cauchy-Schwarz pour les fonctions mesurables) que

$$\int_U |f(x)g(x)| dx \leq \sqrt{\int_U f^2(x) dx} \sqrt{\int_U g^2(x) dx} < +\infty.$$

Cette forme bilinéaire symétrique positive sur le  $\mathbb{R}$ -espace vectoriel

$$\mathcal{L}_{\mathbb{R}}^2(U, \mathcal{L}(\mathbb{R}^n)|_U, dx)$$

est dite *forme bilinéaire d'énergie*; cette terminologie tient au fait que la quantité positive

$$\|f\|_2^2 := \int_U |f(x)|^2 dx$$

est (interprétée du point de vue physique) l'indicateur d'une énergie (l'énergie en physique repose, on l'a rappelé, sur une expression non pas linéaire, mais quadratique). Le  $\mathbb{R}$ -espace vectoriel quotient

$$L_{\mathbb{R}}^2(U, \mathcal{L}(\mathbb{R}^n)|_U, dx) := \frac{\mathcal{L}^2(U, \mathcal{L}(\mathbb{R}^n)|_U, dx)}{\mathcal{R}}$$

où  $\mathcal{R}$  est la relation d'équivalence

$$f \mathcal{R} g \iff f - g = 0 \text{ dx - presque partout sur } U, \quad (1.12)$$

est un  $\mathbb{R}$ -espace normé avec une norme quotient

$$\|\dot{f}\|_2 := \inf_{f \in \dot{f}} \|f\|_2$$

qui dérive d'un produit scalaire<sup>15</sup>, à savoir la forme bilinéaire symétrique

$$(\dot{f}, \dot{g}) \longmapsto \langle f, g \rangle := \int_U f(x)g(x) dx.$$

qui est en effet non seulement symétrique et positive, mais aussi définie (du fait du passage au quotient modulo la relation  $\mathcal{R}$  introduite en (1.12)). On a la proposition très importante, avatar de la théorie de l'intégration de Lebesgue :

**Proposition 1.3** *Si  $U$  est un sous-ensemble mesurable de  $\mathbb{R}^n$ , le  $\mathbb{R}$ -espace vectoriel*

$$L_{\mathbb{R}}^2(U, dx),$$

*équipé du produit scalaire*

$$(\dot{f}, \dot{g}) \longmapsto \langle f, g \rangle := \int_U f(x)g(x) dx,$$

*est un espace de Hilbert réel.*

<sup>15</sup>En fait, il s'agit là le l'espace de Minkowski  $L_{\mathbb{R}}^2(\Omega, \mathcal{T}, \mu)$  défini dans le cours de MHT512 (section 4.3) avec dans ce cas particulier ici  $\Omega := U$ ,  $\mathcal{T} := \mathcal{L}(\mathbb{R}^n)|_U$ ,  $\mu := dx$  étant la mesure de Lebesgue sur  $\mathbb{R}^n$  restreinte à  $U$ .

**Preuve.** Voir le cours de théorie de l'intégration (UE MHT512) ; il s'agit là d'un cas particulier du théorème de Riesz-Fischer (théorème 4.1 du cours de MHT512)<sup>16</sup>.  $\diamond$

L'espace  $\mathcal{L}_{\mathbb{C}}^2(U, dx) := \mathcal{L}_{\mathbb{C}}^2(U, \mathcal{L}(\mathbb{R}^n)|_U, dx)$  des fonctions mesurables sur  $U$ , à valeurs complexes, telles que

$$\|f\|_2^2 := \int_U |f(x)|^2 dx < \infty$$

se traite de manière identique ; en quotientant par la relation d'équivalence  $\mathcal{R}$ , on obtient encore un espace  $L_{\mathbb{C}}^2(U, \mathcal{L}(\mathbb{R}^n)|_U, dx)$  équipé d'un produit scalaire, qui n'est rien d'autre que le  $\mathbb{C}$ -espace normé  $L_{\mathbb{C}}^2(U, \mathcal{L}(\mathbb{R}^n), dx)$  de Minkowski ( $p = 2$ ) rencontré, lui aussi, dans la section 4.3 du cours d'intégration (MHT512).

**Proposition 1.4** *Si  $U$  est un sous-ensemble mesurable de  $\mathbb{R}^n$ , le  $\mathbb{C}$ -espace vectoriel*

$$L_{\mathbb{C}}^2(U, dx),$$

*équipé du produit scalaire*

$$(\dot{f}, \dot{g}) \longmapsto \langle f, g \rangle := \int_U f(x) \overline{g(x)} dx,$$

*est un espace de Hilbert complexe.*

**Preuve.** Voir le cours de théorie de l'intégration (UE MHT512, théorème 4.1 de Riesz-Fischer dans le cas  $p = 2$ ).  $\diamond$

Seront particulièrement intéressants pour nous les exemples correspondant au cas  $U = \mathbb{R}^n$ , espaces que nous noterons  $L_{\mathbb{R}}^2(\mathbb{R}^n)$  (espace de Hilbert réel) et  $L_{\mathbb{C}}^2(\mathbb{R}^n)$  (espace de Hilbert complexe).

### Le cas périodique

Seront intéressants aussi pour nous plus tard les espaces

$$L_{\mathbb{R}}^2\left(\frac{\mathbb{R}^n}{(2\pi\mathbb{Z})^n}\right) = L_{\mathbb{R}}^2(\mathbb{T}^n) \quad (\text{resp. } L_{\mathbb{C}}^2\left(\frac{\mathbb{R}^n}{(2\pi\mathbb{Z})^n}\right) = L_{\mathbb{C}}^2(\mathbb{T}^n))$$

des classes de fonctions mesurables sur  $\mathbb{R}^n$ , à valeurs réelles (*resp.* complexes), qui sont périodiques de période  $2\pi$  en chacune des variables réelles  $\theta_1, \dots, \theta_n$ , et telles que

$$\int_{[0, 2\pi]^n} |f(\theta)|^2 d\theta < \infty.$$

Le produit scalaire sur ces espaces est

$$(\dot{f}, \dot{g}) \longmapsto \langle \dot{f}, \dot{g} \rangle_{\text{per}} := \frac{1}{(2\pi)^n} \int_{[0, 2\pi]^n} f(\theta) g(\theta) d\theta$$

(dans le cas de  $L_{\mathbb{R}}^2(\mathbb{T}^n)$ ) ou

$$(\dot{f}, \dot{g}) \longmapsto \langle \dot{f}, \dot{g} \rangle_{\text{per}} := \frac{1}{(2\pi)^n} \int_{[0, 2\pi]^n} f(\theta) \overline{g(\theta)} d\theta$$

(dans le cas de  $L_{\mathbb{C}}^2(\mathbb{T}^n)$ ) ; le premier est un espace de Hilbert réel, le second un espace de Hilbert complexe.

<sup>16</sup>C'est d'ailleurs dans ce cas particulier ( $p = 2$ ) que ce théorème fut formulé en 1907 (et exploité dans le cadre de la théorie des séries de Fourier que nous retrouverons au chapitre 2) par le mathématicien hongrois Frigyes Riesz (1880-1956) et le mathématicien autrichien Ernst Fischer (1875-1954).

### Le cas des espaces $L^2$ « à poids »

Si  $U$  est un sous-ensemble mesurable de  $\mathbb{R}^n$  et  $\omega$  une classe de fonctions mesurables sur  $U$  (modulo la relation  $\mathcal{R}$  introduite en (1.12)), admettant un représentant strictement positif (*i.e*  $\omega > 0$  presque partout sur  $U$  au sens de Lebesgue) et que l'on appelle un « poids », on peut définir les espaces

$$L_{\mathbb{R}}^2(U, \omega dx) := \left\{ \dot{f}; f \text{ mesurable } U \rightarrow \mathbb{R}, \int_U |f(x)|^2 \omega(x) dx < \infty \right\}$$

et

$$L_{\mathbb{C}}^2(U, \omega dx) := \left\{ \dot{f}; f \text{ mesurable } U \rightarrow \mathbb{C}, \int_U |f(x)|^2 \omega(x) dx < \infty \right\}$$

(suivant que les classes de fonctions envisagées sont des classes de fonctions à valeurs réelles ou complexes). Ce sont des espaces de Hilbert, le produit scalaire étant défini par

$$\langle \dot{f}, \dot{g} \rangle := \int_U f(x) g(x) \omega(x) dx$$

(dans le cas réel) ou par

$$\langle \dot{f}, \dot{g} \rangle := \int_U f(x) \overline{g(x)} \omega(x) dx$$

(dans le cas complexe)<sup>17</sup>.

#### 1.4.4 Les espaces $L_{\mathbb{R}}^2(\Omega, \mathcal{T}, P)$

Soit  $(\Omega, \mathcal{T}, P)$  un espace probabilisé ( $\Omega$  est un ensemble d'« événements »,  $\mathcal{T}$  une tribu sur  $\Omega$ ,  $P$  une distribution de probabilité sur  $\mathcal{T}$ <sup>18</sup>).

On peut considérer le  $\mathbb{R}$  espace vectoriel  $\mathcal{L}_{\mathbb{R}}^2(\Omega, \mathcal{T}, P)$  des variables aléatoires réelles sur l'espace probabilisé  $(\Omega, \mathcal{T}, P)$ <sup>19</sup> ayant de plus un moment d'ordre 2 fini, c'est-à-dire telles que

$$\int_{\Omega} X(\omega)^2 dP = \text{Espérance}[X^2] < \infty;$$

l'espace  $L_{\mathbb{R}}^2(\Omega, \mathcal{T}, P)$  est obtenu en quotient l'espace  $\mathcal{L}_{\mathbb{R}}^2(\Omega, \mathcal{T}, P)$  par la relation d'équivalence

$$X \sim Y \iff P(\{\omega \in \Omega; X(\omega) \neq Y(\omega)\}) = 0. \quad (1.13)$$

Cet espace  $L_{\mathbb{R}}^2(\Omega, \mathcal{T}, P)$  peut être équipé du produit scalaire

$$(\dot{X}, \dot{Y}) \longmapsto E[XY] := \int_{\Omega} X(\omega)Y(\omega) dP;$$

ce produit scalaire est intimement lié à la notion de *covariance* puisque par définition la *covariance* des deux variables aléatoires  $X$  et  $Y$  est définie par :

$$\begin{aligned} \text{cov}(X, Y) &:= E[(X - EX)(Y - EY)] = E[XY] - 2E[X]E[Y] + E[X]E[Y] \\ &= E[XY] - E[X]E[Y] \end{aligned}$$

<sup>17</sup>Il s'agit là encore des espaces de Minkowski  $L_{\mathbb{K}}^2(\Omega, \mathcal{T}, \mu)$  de la section 4.3 du cours de MHT512 (ici avec  $p = 2$ ), où  $\Omega := U$ ,  $\mathcal{T} = \mathcal{L}(\mathbb{R}^n)$  et  $\mu = \omega dx$  est la mesure à densité  $\omega$  par rapport à la mesure de Lebesgue restreinte à  $U$ .

<sup>18</sup>C'est-à-dire une mesure positive  $P : \mathcal{T} \rightarrow [0, \infty]$  telle que  $P(\Omega) = 1$ .

<sup>19</sup>Rappelons que ce sont les applications  $X$  de  $\Omega$  dans  $\mathbb{R}$  telle que  $X^{-1}(\mathcal{B}(\mathbb{R})) \subset \mathcal{T}$ .

(la covariance entre  $X$  et  $X$  étant par définition la *variance* de  $X$ ). On rappelle dans ce nouveau le résultat suivant, avatar du théorème de Riesz-Fischer (théorème 4.1 du cours de MHT512) :

**Proposition 1.5** *Le  $\mathbb{R}$ -espace vectoriel  $L_{\mathbb{R}}^2(\Omega, \mathcal{T}, P)$ , équipé du produit scalaire*

$$(\dot{X}, \dot{Y}) \longmapsto E[XY] := \int_{\Omega} X(\omega)Y(\omega) dP,$$

*est un espace de Hilbert réel.*

**Preuve.** Voir le cours de théorie de l'intégration MHT512 (théorème 4.1).  $\diamond$

On peut aussi envisager le  $\mathbb{C}$ -espace vectoriel  $L_{\mathbb{C}}^2(\Omega, \mathcal{T}, P)$  des variables aléatoires  $X$  à valeurs complexes définies sur l'espace probalilisé  $(\Omega, \mathcal{T}, P)$  et telles que

$$E[|X|^2] < \infty.$$

En quotientant cet espace par la relation d'équivalence  $\sim$  introduite en (1.13), on trouve un  $\mathbb{C}$ -espace  $L_{\mathbb{C}}^2(\Omega, \mathcal{T}, P)$  que l'on peut naturellement équiper du produit scalaire

$$(\dot{X}, \dot{Y}) \longmapsto E[X\bar{Y}] := \int_{\Omega} X(\omega)\overline{Y(\omega)} dP;$$

on définit de même la covariance de  $X$  avec  $Y$  par

$$\text{cov}(X, Y) = E[(X - E[X])(\overline{Y - E[Y]})] = E[X\bar{Y}] - E[X] \times \overline{E[Y]}$$

et la *variance* de  $X$  par

$$V(X) := \text{cov}(X, X) = E[|X - E[X]|^2] = E[|X|^2] - |E[X]|^2.$$

**Proposition 1.6** *Le  $\mathbb{C}$ -espace vectoriel  $L_{\mathbb{C}}^2(\Omega, \mathcal{T}, P)$ , équipé du produit scalaire*

$$(\dot{X}, \dot{Y}) \longmapsto E[X\bar{Y}] := \int_{\Omega} X(\omega)\overline{Y(\omega)} dP,$$

*est un espace de Hilbert complexe.*

**Preuve.** Voir le cours de théorie de l'intégration MHT512 (théorème 4.1 de Riesz-Fischer).  $\diamond$

Signalons que si  $X$  est une variable aléatoire réelle représentant d'un élément de  $L_{\mathbb{R}}^2(\Omega, \mathcal{T}, P)$ , la tribu  $\mathcal{T}_X$  engendrée par les ensembles  $X^{-1}(B)$ , où  $B \in \mathcal{B}(\mathbb{R})$ , est une sous-tribu de la tribu  $\mathcal{T}$ ; c'est la tribu des évènements « dépendant de  $X$  ». On peut travailler cette fois sur l'espace probalilisé  $(\Omega, \mathcal{T}_X, P|_{\mathcal{T}_X})$  et introduire le  $\mathbb{R}$ -espace de Hilbert  $L_{\mathbb{R}}^2(\Omega, \mathcal{T}_X, P|_{\mathcal{T}_X})$ . Il existe une injection naturelle  $i_X$  (isométrie au niveau des normes) de  $L_{\mathbb{R}}^2(\Omega, \mathcal{T}_X, P|_{\mathcal{T}_X})$  dans  $L^2(\Omega, \mathcal{T}, P)$  qui consiste juste à associer à la classe de  $Y$  ( $Y$  étant considérée comme variable aléatoire sur l'espace probalilisé  $(\Omega, \mathcal{T}_X, P|_{\mathcal{T}_X})$ ) la classe de  $Y$  (cette fois considérée comme variable aléatoire sur l'espace probalilisé  $(\Omega, \mathcal{T}, P)$ ). Le  $\mathbb{R}$ -espace de Hilbert  $L_{\mathbb{R}}^2(\Omega, \mathcal{T}_X, P|_{\mathcal{T}_X})$  est ainsi un exemple de sous-espace vectoriel fermé du  $\mathbb{R}$ -espace de Hilbert  $L_{\mathbb{R}}^2(\Omega, \mathcal{T}, P)$ .

La même remarque vaut pour le cas complexe. Si  $X$  est un représentant d'un élément  $\dot{X}$  de  $L_{\mathbb{C}}^2(\Omega, \mathcal{T}, P)$ , le  $\mathbb{C}$ -espace de Hilbert  $L_{\mathbb{C}}^2(\Omega, \mathcal{T}_X, P|_{\mathcal{T}_X})$  est ainsi un exemple particulier de sous-espace vectoriel fermé du  $\mathbb{C}$ -espace de Hilbert  $L_{\mathbb{C}}^2(\Omega, \mathcal{T}, P)$ .

On profitera de cette configuration ultérieurement pour y faire entrer le mécanisme de projection orthogonale ; ce sera la base de la définition du « conditionnement par la variable aléatoire  $X$  »<sup>20</sup>.

## 1.5 Le théorème de projection orthogonale sur un convexe fermé

Soit  $H$  un espace de Hilbert (réel ou complexe).

**Définition 1.4** *Un sous ensemble  $C$  de  $H$  est dit convexe si et seulement si*

$$\forall h_1 \in C, \forall h_2 \in C, \forall t \in [0, 1], (1 - t)h_1 + th_2 \in C.$$

*L'ensemble  $C$  est dit strictement convexe si et seulement si*

$$\forall h_1 \in C, \forall h_2 \in C, \forall t \in ]0, 1[, (1 - t)h_1 + th_2 \in \text{intérieur}(C).$$

Voici le premier théorème majeur concernant la géométrie « à la Pythagore » dans un espace de Hilbert ; c'est ce résultat fondamental qui soutend toutes les figures inspirées du concept d'orthogonalité et qui plus tard nous servira de guide et d'auxiliaire constant dans les preuves ou les constructions algorithmiques que nous serons amenés à faire dans un espace de Hilbert (réel ou complexe), telles celles proposées dans les exemples 1.1 et 1.2 évoqués dans la section 1.1.

**Théorème 1.1** *Soit  $C$  un sous ensemble fermé, convexe et non vide d'un espace de Hilbert  $H$  réel ou complexe (la norme étant notée  $\| \cdot \|$ ). Soit  $h$  un élément de  $H$ . Il existe un et un seul élément  $a = \text{Proj}_C(h)$  (dit projection orthogonale de  $h$  sur  $C$ ), tel que*

$$\|h - a\| = \min_{x \in C} \|h - x\|. \quad (1.14)$$

*Cet élément  $a = \text{Proj}_C(h)$  est caractérisé par le jeu d'inégalités*

$$\forall x \in C, \langle h - a, x - a \rangle \leq 0 \quad (1.15)$$

*dans le cas réel et par le jeu d'inégalités*

$$\forall x \in C, \text{Re} \langle h - a, x - a \rangle \leq 0 \quad (1.16)$$

*dans le cas complexe.*

**Remarque 1.1.** Les jeux d'inégalités (1.15) ou (1.16) s'interprètent géométriquement et de manière intuitive en disant que pour tout  $x$  de  $C$ , les vecteurs  $x - a$  et  $h - a$  font un angle « obtus ». On se souvient en effet d'avoir (dans la sous-section 1.3.2) défini le cosinus de l'angle de deux vecteurs  $h_1$  et  $h_2$

$$\cos(\text{angle}(h_1, h_2)) := \frac{\text{Re} \langle h_1, h_2 \rangle}{\|h_1\| \|h_2\|}.$$

Les conditions (1.15) et (1.16) se lisent donc

$$\forall x \in C, \cos(\text{angle}(h - a, x - a)) \leq 0.$$

<sup>20</sup>Voir aussi le cours de probabilités MHT601.

**Preuve.** Soit  $d$  la distance de  $h$  à  $C$ ;  $d = \inf_{x \in C} \|x - h\| > 0$  puisque  $C$  est fermé. Soit  $(x_n)_{n \geq 1}$  une suite de points de  $C$  telle que  $d^2 \leq \|h - x_n\|^2 \leq d^2 + 1/n$  (une telle suite existe du fait de la définition de la borne inférieure). En utilisant la formule de la médiane (1.6), il vient

$$\begin{aligned} \forall n, m \in \mathbb{N}^*, \|x_n - x_m\|^2 &= 2(\|h - x_n\|^2 + \|h - x_m\|^2) - 4\left\|h - \frac{x_n + x_m}{2}\right\|^2 \\ &\leq 4d^2 + \frac{1}{n^2} + \frac{2}{m^2} - 4\left\|h - \frac{x_n + x_m}{2}\right\|^2 \\ &\leq \frac{2}{n^2} + \frac{2}{m^2} \end{aligned}$$

du fait que  $(x_n + x_m)/2$  est dans  $C$  ( $C$  est convexe) et que par conséquent

$$d = d(h, C) \leq d\left(h, \frac{x_n + x_m}{2}\right).$$

La suite  $(x_n)_{n \geq 1}$  est donc une suite de Cauchy qui converge dans  $H$  (puisque  $H$  est complet), ce vers un point de  $C$  puisque  $C$  est fermé. L'élément  $a = \lim_n x_n$  vérifie (1.14). Un tel  $a$  s'avère unique puisque la formule de la médiane (1.6) implique aussi, si l'on dispose de deux candidats  $a_1$  et  $a_2$  pour (1.14)

$$\|a_1 - a_2\|^2 + 4\left\|h - \frac{a_1 + a_2}{2}\right\|^2 = 4d^2 \geq 4d^2 + \|a_1 - a_2\|^2$$

puisque  $\|h - a_1\|^2 = \|h - a_2\|^2 = d^2$  et que  $(a_1 + a_2)/2 \in C$ . Le premier volet de la proposition est acquis. Voyons maintenant ce qui concerne (1.15) (ou (1.16)). Pour simplifier, nous nous placerons dans le cas complexe (c'est le plus général). Si  $a$  réalise la distance de  $h$  à  $C$ , on a

$$\forall t \in [0, 1], \forall x \in C, \|h - a(1 - t) - tx\|^2 = \|h - a - t(x - a)\|^2 \geq \|h - a\|^2 \quad (1.17)$$

(puisque  $C$  est convexe et que  $ta + (1 - t)x$  est dans  $C$  quand  $x$  est dans  $C$ ). Ceci donne en développant

$$\forall t \in [0, 1], t^2\|x - a\|^2 - 2t\operatorname{Re} \langle h - a, x - a \rangle \geq 0, \quad (1.18)$$

et donc, après division par  $t$ ,

$$\forall t \in ]0, 1], t\|x - a\|^2 - 2\operatorname{Re} \langle h - a, x - a \rangle \geq 0$$

ce qui prouve (en faisant tendre  $t$  vers 0) que (1.16) est remplie si  $a$  vérifie (1.14). La réciproque est immédiate : si (1.16) est remplie, alors on a (1.18) et par conséquent (1.17) pour tout  $x \in C$ . En prenant  $t = 1$ , on a  $\|h - x\|^2 \geq \|h - a\|^2$  pour tout  $x$  dans  $C$ , ce qui montre que  $a$  réalise bien la distance de  $z$  à  $C$ . La preuve du théorème est achevée.  $\diamond$

**Remarque 1.2.** Quand  $C = F$  est un sous espace fermé de  $H$ , on peut écrire les conditions (1.15) ou (1.16) (suivant le cas, réel ou complexe) sous la forme

$$\langle h - a, x \rangle = 0 \quad \forall x \in F, \quad (1.19)$$

ce que l'on traduit en disant que  $h - a$  est orthogonal à  $F$  (en abrégé  $h - a \in F^\perp$ ).

La remarque 1.2 ci-dessus nous amène naturellement à la définition suivante :

**Définition 1.5** *Etant donné un sous espace  $F$  d'un espace de Hilbert  $H$ , on appelle orthogonal de  $F$  et on note  $F^\perp$  le sous espace vectoriel (nécessairement fermé cette fois) défini comme*

$$F^\perp := \{h \in H; \forall x \in F, \langle x, h \rangle = 0\}.$$

*Si  $F$  est un sous-espace fermé, la projection orthogonale  $\text{Proj}_F(h)$  de  $h$  sur  $F$  est par définition l'unique point  $a$  de  $F$  tel que  $h - a \in F^\perp$ .*

Il est aussi très important d'étudier dans un Hilbert le comportement de la suite des projections d'un point sur des convexes fermés  $C_1, \dots, C_n, \dots$  de  $H$ . Voici un résultat lors que les suites de convexes fermés  $(C_n)_n$  sont des suites de fermés emboîtés.

**Proposition 1.7** *Soit  $C_1, C_2, \dots, C_n, \dots$  une suite de sous ensembles non vides, convexes et fermés dans un espace de Hilbert  $H$ . On note  $P_1, \dots, P_n, \dots$ , les projections orthogonales sur les convexes respectifs  $C_1, C_2, \dots, C_n, \dots$*

1. *Si  $C_{k+1} \subset C_k$  pour tout  $k \in \mathbb{N}^*$  et si*

$$\bigcap_{k=1}^{\infty} C_k = C_\infty \neq \emptyset,$$

*alors  $C_\infty$  est aussi un sous ensemble convexe fermé de  $H$  et, pour tout  $h \in H$ , la suite  $(a_k)_k$ , où*

$$a_k := P_k(h), \quad k \in \mathbb{N}^*,$$

*converge vers  $P_\infty(z) = \text{Proj}_{C_\infty}(h)$ .*

2. *Si  $C_k \subset C_{k+1}$  pour tout  $k \in \mathbb{N}^*$  et*

$$C_{[\infty]} := \overline{\bigcup_{k=1}^{\infty} C_k},$$

*alors  $C_{[\infty]}$  est aussi un sous ensemble convexe fermé de  $H$  et, pour tout  $h \in H$ , la suite  $(a_k)_{k \geq 1}$ , où*

$$a_k = P_k(h), \quad k \in \mathbb{N}^*,$$

*converge vers  $P_{[\infty]}(z) = \text{Proj}_{C_{[\infty]}}(h)$ .*

### Preuve.

*Preuve du point 1.* Que  $C_\infty$  soit aussi un sous ensemble convexe fermé est immédiat (faire l'exercice, une intersection de convexes est convexe). On a, pour tout  $k \in \mathbb{N}^*$ ,

$$\|h - a_k\|^2 \leq \|h - a_{k+1}\|^2 \leq \|h - P_\infty(h)\|^2$$

puisque  $C_\infty \subset C_{k+1} \subset C_k$ . La suite  $(\|h - a_k\|^2)_{k \geq 1}$  est donc une suite croissante majorée convergent vers un nombre  $d > 0$ . La formule de la médiane (1.6) nous assure, comme dans la preuve du théorème 1.1, que la suite  $a_k$  est de Cauchy dans  $H$  pour la distance associée à la norme hilbertienne; elle converge donc vers un point  $a_\infty$  de  $H$  (nécessairement dans  $C_\infty$ ). Mais on peut vérifier que pour tout  $u \in C_\infty$ ,

$$\text{Re} \langle h - a_\infty, u - a_\infty \rangle = \lim_{k \rightarrow \infty} \text{Re} \langle h - a_k, u - a_k \rangle \leq 0$$

car  $u \in C_k$  pour tout  $k$  ; mais, d'après le théorème 1.1, la clause ci-dessus est remplie si et seulement si  $a_\infty = P_\infty(h)$ .

*Preuve du point 2.* Que  $C_{[\infty]}$  est un sous ensemble convexe fermé est immédiat (faire l'exercice, une union croissante de convexes est convexe). Pour tout  $k \in \mathbb{N}^*$ , on a cette fois

$$\|h - a_k\|^2 \geq \|h - a_{k+1}\|^2$$

puisque  $C_k \subset C_{k+1}$ . La suite  $(\|h - a_k\|^2)_{k \geq 1}$  (suite décroissante minorée par 0) converge vers un nombre  $d \geq 0$ . La formule de la médiane (1.6) nous assure encore, comme dans la preuve du théorème 1.1, que la suite  $(a_k)_{k \geq 1}$  est de Cauchy ; elle converge donc vers un élément  $a_{[\infty]}$  (appartenant nécessairement à  $C_{[\infty]}$ ). Mais on peut vérifier à nouveau que, si  $k \in \mathbb{N}^*$ , pour tout  $u \in C_k$ ,

$$\operatorname{Re} \langle h - a_\infty, u - a_\infty \rangle = \lim_{k \rightarrow \infty} \operatorname{Re} \langle h - a_k, u - a_k \rangle \leq 0 ;$$

Cette propriété est satisfaite pour tout  $u$  dans

$$\bigcup_{k \in \mathbb{N}^*} C_k,$$

donc pour tout  $u$  dans

$$\overline{\bigcup_{k \in \mathbb{N}^*} C_k} = C_{[\infty]}.$$

Comme  $C_{[\infty]}$  est convexe et fermé,  $a_{[\infty]} = P_{[\infty]}(h)$  par définition de  $P_{[\infty]}(h)$  (théorème 1.1).  $\diamond$

## 1.6 Systèmes orthonormés, « frames » ; exemples

Soit  $(H, \langle \cdot, \cdot \rangle)$  un espace de Hilbert (réel ou complexe). Il sera important (pour permettre la représentation des éléments de  $H$  de manière non redondante) de disposer de systèmes dits *orthonormés*.

**Définition 1.6** Soit  $(H, \langle \cdot, \cdot \rangle)$  un espace de Hilbert (réel ou complexe). Un système de vecteurs  $(e_\iota)_{\iota \in I}$  indexé par un ensemble d'indices  $I$  est dit *orthonormé* si et seulement si, pour tout  $\iota, \iota'$  dans  $I$ ,

$$\langle e_\iota, e_{\iota'} \rangle = \begin{cases} 1 & \text{si } \iota = \iota' \\ 0 & \text{si } \iota \neq \iota' \end{cases}$$

Le système est dit *orthogonal* si les vecteurs  $e_\iota$  sont tous non nuls et si  $\langle e_\iota, e_{\iota'} \rangle = 0$  pour tout couple d'éléments distincts  $\iota$  et  $\iota'$  de l'ensemble d'indices  $I$ .

On remarque immédiatement qu'un système orthogonal est toujours libre (au sens algébrique), ce qui signifie que pour toute partie finie  $J \subset I$ , la famille  $(e_\iota)_{\iota \in J}$  est libre.

**Exemple 1.3.** Un exemple qui sera pour nous fondamental sera celui emprunté au cadre des espaces de Hilbert  $l_{\mathbb{R}}^2(I)$  ou  $l_{\mathbb{C}}^2(I)$  (par exemple pour  $I = \mathbb{N}$  ou  $I = \mathbb{Z}$  qui seront pour nous les deux cas importants ultérieurement). Dans ce cas, le système de vecteurs  $(e_\iota)_{\iota \in I}$  définis par  $e_\iota = (e_{\iota, \iota'})_{\iota' \in I}$ , où

$$e_{\iota, \iota'} = \begin{cases} 1 & \text{si } \iota' = \iota \\ 0 & \text{sinon} \end{cases}$$



est un système orthonormé, le produit scalaire étant, rappelons le, défini par

$$\langle (x_\iota)_{\iota \in I}, (y_\iota)_{\iota \in I} \rangle := \sum_{\iota \in I} x_\iota \overline{y_\iota}$$

(ici dans le cas complexe ; dans le cas réel, on ôte la conjugaison complexe).

La notion de système orthonormé étant une notion *idéale*, souvent difficile à réaliser dans la pratique (et on le verra, particulièrement instable, l'orthogonalité étant une propriété très fragile), on introduira aussi une notion plus « molle », mais souvent bien utile dans la pratique, celle de *frame*<sup>21</sup>. C'est cette dernière notion qui nous permettra une représentation certes redondante (au contraire de la représentation dans un système orthonormé), mais stable (ce qui est important) des éléments de l'espace de Hilbert dans lequel on est amené à travailler.

**Définition 1.7** Soit  $(H, \langle \cdot, \cdot \rangle)$  un espace de Hilbert (réel ou complexe). Un système de vecteurs  $(e_\iota)_{\iota \in I}$  indexé par un ensemble d'indices  $I$  est un « frame » si et seulement si il existe deux constantes strictement positives  $C_1$  et  $C_2$  telles que, pour toute partie finie  $J$  de  $I$ , pour toute collection de scalaires  $(\lambda_\iota)_{\iota \in J}$ ,

$$C_1 \left( \sum_{\iota \in J} |\lambda_\iota|^2 \right) \leq \left\| \sum_{\iota \in J} \lambda_\iota e_\iota \right\|^2 \leq C_2 \left( \sum_{\iota \in J} |\lambda_\iota|^2 \right).$$

Un système orthonormé est bien sûr un « frame ».

**Exemple 1.4.** Soit  $H = L^2_{\mathbb{C}}(\mathbb{T})$  le  $\mathbb{C}$ -espace de Hilbert des classes de fonctions mesurables de  $\mathbb{R}$  dans  $\mathbb{C}$ ,  $2\pi$ -périodiques, d'énergie finie sur  $[0, 2\pi]$ , équipé du produit scalaire (dit produit scalaire associé à l'énergie)

$$\langle \dot{f}, \dot{g} \rangle := \frac{1}{2\pi} \int_0^{2\pi} f(\theta) \overline{g(\theta)} d\theta.$$

Les « harmoniques fondamentales de période  $2\pi$  », c'est-à-dire les classes  $(\dot{e}_n)_{n \in \mathbb{Z}}$  des fonctions

$$e_n : \theta \in \mathbb{R} \mapsto \exp(in\theta) = \cos(n\theta) + i \sin(n\theta)$$

forment un système orthonormé dans  $H$ . Ce système jouera plus tard un rôle essentiel dans l'analyse de Fourier des signaux périodiques de période  $2\pi$ <sup>22</sup>.

**Exemple 1.5.** Soit  $H = L^2_{\mathbb{C}}([-1, 1], dt)$  l'espace de Hilbert des classes de fonctions mesurables de  $[-1, 1]$  dans  $\mathbb{C}$ , équipé du produit scalaire

$$\langle \dot{f}, \dot{g} \rangle := \int_{-1}^1 f(t) \overline{g(t)} dt.$$

<sup>21</sup>Difficile ici de donner une traduction française de cette terminologie anglo-saxonne ; le concept le plus proche est sans doute celui de « trame » ou de « bâti » (dans une structure).

<sup>22</sup>Notons que  $2\pi$  ne joue aucun rôle particulier ici et que bien sûr, on peut remplacer  $2\pi$  par  $T > 0$ , le produit scalaire (sur le  $\mathbb{C}$ -espace des classes de fonctions mesurables  $T$ -périodiques de carré intégrable sur  $[0, T]$ ) devenant

$$\langle \dot{f}, \dot{g} \rangle := \frac{1}{T} \int_0^T f(\theta) \overline{g(\theta)} d\theta$$

et les fonctions  $e_n$  les fonctions  $e_{T,n}$  :

$$e_{T,n} : \theta \in \mathbb{R} \mapsto \exp(2i\pi n\theta/T) = \cos(2\pi n\theta/T) + i \sin(2\pi n\theta/T)$$

(ce sont les harmoniques fondamentales de période  $T$ ).

On vérifie immédiatement que, pour tout entier positif  $n$ , la fonction

$$t \mapsto \frac{1}{2^n n!} \sqrt{n + \frac{1}{2}} \times \frac{d^n}{dt^n} [(1 - t^2)^n] \quad (1.20)$$

est une fonction polynômiale  $L_n$  de degré exactement  $n$  et que le système  $(L_n)_{n \in \mathbb{N}}$  est un système orthonormé de l'espace de Hilbert  $H^2$ . Le polynôme  $L_n$  est appelé  $n$ -ème polynôme de Legendre<sup>24</sup> et il est intéressant de constater que plus  $n$  augmente, plus le graphe de  $L_n$  sur  $[-1, 1]$  présente des oscillations de plus en plus accentuées (au niveau de l'amplitude) lorsque l'on s'approche des bornes de l'intervalle  $[-1, 1]$ . Voici par exemple, sur la figure 1.4 ci-dessous, le graphe de  $L_{30}$  :

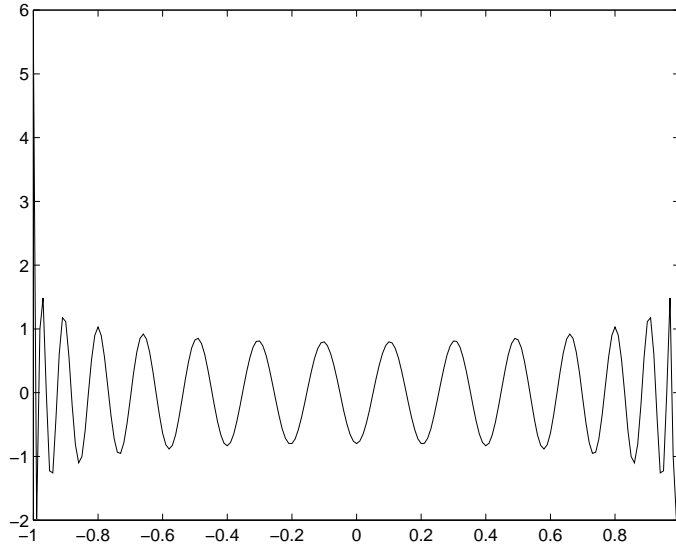


FIG. 1.4 – Le graphe de la fonction polynômiale de Legendre  $L_{30}$  sur  $[-1, 1]$

Ce constat sera important lorsque l'on sera amené à utiliser le système  $(L_n)_{n \in \mathbb{N}}$  à des fins pratiques (en relation en particulier avec les méthodes numériques d'intégration par quadrature proposées par Gauss). On retrouve (avec le souci causé par l'amplification des amplitudes des oscillations lorsque l'on s'approche du bord) un problème aussi présent dans l'interpolation de Lagrange faite sur un intervalle  $[a, b]$  avec un pas de discrétisation régulier<sup>25</sup>.

**Exemple 1.6.** Considérons l'espace  $L^2_{\mathbb{C}}([-1, 1], \omega(t) dt)$ , où

$$\omega(t) := \frac{1}{\sqrt{1 - t^2}}, \quad \forall t \in ]-1, 1[.$$

Il s'agit, rappelons le, de l'espace de Hilbert des classes de fonctions  $f$  mesurables (pour la mesure de Lebesgue) sur  $[-1, 1]$ , telles que

$$\int_{-1}^1 \frac{|f(t)|^2}{\sqrt{1 - t^2}} dt < +\infty,$$

équipé du produit scalaire

$$\langle f, g \rangle := \int_{-1}^1 f(t) \overline{g(t)} \frac{dt}{\sqrt{1 - t^2}}.$$

<sup>23</sup>Vérifier ce dernier point constitue un exercice classique de  $L^2$  souvent utilisé pour illustrer l'utilisation de la formule d'intégration par parties.

<sup>24</sup>Adrien-Marie Legendre (1752-1833) introduit ces polynômes en 1782 dans son traité *Figure des planètes*.

<sup>25</sup>Pour ces questions relevant plus de l'analyse numérique, voir le polycopié du cours de MHT304, <http://www.math.u-bordeaux1.fr/~yger/mht304.pdf>, chapitre 4.

C'est un exercice immédiat que de constater que les classes des fonctions polynômiales  $(\Theta_n)_{n \in \mathbb{N}}$  définies par

$$\begin{aligned}\Theta_0(t) &= \frac{1}{\sqrt{\pi}} \\ \Theta_n(t) &= \sqrt{\frac{2}{\pi}} \cos(n \arccos(t)), \quad n \geq 1,\end{aligned}\tag{1.21}$$

forment un système orthonormé de  $H$  (faire l'exercice en faisant le changement de variables  $t = \cos u$  dans les intégrales). Cette fois, l'amplitude des oscillations des fonctions  $\Theta_n$  n'augmente plus (lorsque  $n$  croît) lorsque l'on s'approche des bords de  $[-1, 1]$ , mais par contre la fréquence de ces oscillations augmente très fortement au fur et à mesure que l'on s'approche du bord, ce qui, on s'en doute, posera des problèmes aigûs au niveau numérique si l'on entend utiliser ces fonctions polynômiales (dites fonctions polynômiales de Tchebychev<sup>26</sup>) pour coder une classe de fonction mesurable sur  $[-1, 1]$ . Voici par exemple (sur la figure 1.5 ci-dessous) le graphe de  $\Theta_{100}$  sur  $[-1, 1]$  :

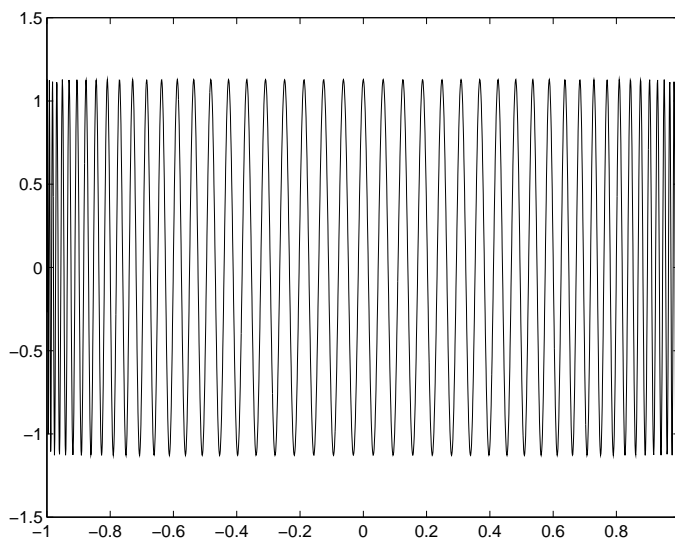


FIG. 1.5 – Le graphe de la fonction polynômiale de Tchebychev  $\Theta_{100}$  sur  $[-1, 1]$

**Exemple 1.7.** Un autre exemple important (du point de vue de la physique) sera l'espace de Hilbert  $L^2_{\mathbb{C}}(\mathbb{R}, e^{-t^2} dt)$  des classes de fonctions mesurables sur  $\mathbb{R}$ , à valeurs complexes, telles que

$$\int_{\mathbb{R}} |f(t)|^2 e^{-t^2} dt < \infty,$$

équipé du produit scalaire

$$\langle \dot{f}, \dot{g} \rangle := \int_{\mathbb{R}} f(t) \overline{g(t)} e^{-t^2} dt.$$

Un système orthonormé de cet espace est le système des classes de fonctions polynômiales  $(\dot{H}_n)_{n \geq 0}$  ( $H_n$  est une fonction polynômiale de degré exactement  $n$ , dite fonction polynômiale de Hermite) donné par

$$\begin{aligned}H_0(t) &:= \frac{1}{\pi^{1/4}} \\ H_n(t) &:= \frac{(-1)^n}{2^{n/2} \pi^{1/4} \sqrt{n!}} \times e^{t^2} \times \frac{d^n}{dt^n} [e^{-t^2}], \quad n \geq 1.\end{aligned}\tag{1.22}$$

L'importance de cet espace de Hilbert (et du système orthonormé de Hermite) tient au fait que la gaussienne  $t \mapsto \exp(-t^2)$  sera amenée à jouer (on le verra) un rôle très important vis-à-vis

<sup>26</sup>On adopte cette terminologie en l'honneur du probabiliste et statisticien russe Pafnouti Tchebychev (1821-1894).

du principe d'incertitude en physique ou dans les théorèmes limites du calcul des probabilités (le théorème « *limite centrale* » en particulier<sup>27</sup>).

On retrouvera plus loin une manière pratique de retrouver tous ces exemples à partir du système des fonctions monômes  $t \mapsto t^n$ ,  $n = 0, 1, \dots$ , qui lui n'est pas orthonormé (dans l'espace de Hilbert de classes de fonctions d'une variable où l'on travaille). La construction se fera de proche en proche suivant le procédé d'orthonormalisation algébrique bien connu de Gram-Schmidt. Voici un autre exemple important :

**Exemple 1.8.** Dans l'espace  $L_{\mathbb{R}}^2([a, b], dt)$ , les classes de fonctions

$$t \in [a, b] \mapsto \sqrt{\frac{2}{b-a}} \sin \left[ \pi \left( n + \frac{1}{2} \right) \frac{t-a}{b-a} \right], \quad n \in \mathbb{N},$$

forment un système orthonormé ; ce sont les *harmoniques locales* subordonnées à l'intervalle  $[a, b]$ . Ce système orthonormé (pensez que l'intervalle  $[a, b]$  peut se dilater ou se contracter tel un accordéon) jouera un rôle primordial pour le codage de la parole ou de la musique.

Voici enfin, pour conclure cette galerie d'exemples, un exemple de *frame*.

**Exemple 1.9 (un exemple de « frame »).** Voici un exemple de « frame » important : considérons l'espace de Hilbert (réel)  $H$  des fonctions continues, affines par morceaux sur  $\mathbb{R}$ , à valeurs réelles, avec « noeuds » aux points d'un maillage  $\tau\mathbb{Z}$  ( $\tau > 0$ ), telles que

$$\int_{\mathbb{R}} |f(t)|^2 dt < +\infty,$$

ce qui revient en fait à dire

$$\sum_{n \in \mathbb{Z}} |f(n\tau)|^2 < \infty,$$

équipé du produit scalaire

$$\int_{\mathbb{R}} f(t) g(t) dt.$$

Pour simplifier, on prendra  $\tau = 1$ . Les fonctions triangle (ou « *splines* » dans la terminologie des mathématiques appliquées)

$$\Delta_n : t \mapsto \max(0, 1 - |t - n|), \quad n \in \mathbb{Z},$$

ne forment pas un système orthonormé de  $H$ , ni même orthogonal (le produit scalaire de  $\Delta_n$  et  $\Delta_{n+1}$  vaut  $1/6$ ) ; par contre, le système  $(\Delta_n)_{n \in \mathbb{Z}}$  est, on le verra plus loin, un *frame*. Ces fonctions (que l'on a représenté sur la figure 1.6 ci-dessous) forment un *frame* important car c'est souvent comme éléments de cet espace de Hilbert que sont représentés visuellement (sur un écran d'ordinateur) les signaux digitaux  $(x_k)_{k \in \mathbb{Z}}$  éléments de l'espace de Hilbert  $l_{\mathbb{R}}^2(\mathbb{Z})$ .

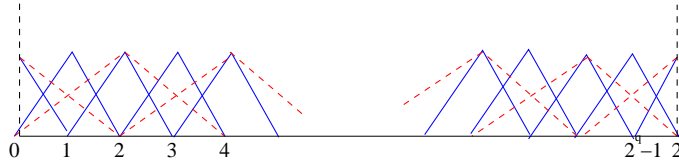


FIG. 1.6 – Les éléments du *frame*  $(\Delta_n)_{n \in \mathbb{Z}}$  avec  $\tau = 1$  et  $\tau = 2$

Supposons maintenant que l'espace de Hilbert ne soit pas de dimension finie et admette au moins un système orthonormé dénombrable. La proposition 1.7 induit la proposition facile (mais très importante) suivante :

<sup>27</sup>Voir le cours de l'UE MHT601.

**Proposition 1.8** Soit  $H$  un  $\mathbb{K}$ -espace de Hilbert (réel ou complexe :  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$ ), la norme étant notée  $\| \cdot \|$  et le produit scalaire  $\langle \cdot, \cdot \rangle$ , et  $(e_n)_{n \in \mathbb{N}^*}$  un système orthonormé dénombrable de  $H$ . Soit  $V$  le sous-espace fermé de  $H$  défini comme l'adhérence (dans  $H$ ) de l'ensemble constitué de toutes les combinaisons linéaires finies

$$\sum_{k=1}^N \lambda_k e_k, \quad \lambda_k \in \mathbb{K}, \quad N \in \mathbb{N}^*.$$

Alors

1. pour tout  $h \in H$ , on a l'inégalité suivante, dite de Bessel<sup>28</sup> :

$$\sum_{k=1}^{\infty} |\langle h, e_k \rangle|^2 \leq \|h\|^2; \quad (1.23)$$

2. pour tout  $h \in H$ , la série de terme général  $\langle h, e_n \rangle e_n \in H$  est convergente (dans  $H$ ) et l'on a

$$\text{Proj}_V(h) = \sum_{k=1}^{\infty} \langle h, e_k \rangle e_k; \quad (1.24)$$

3. si  $(\lambda_n)_{n \in \mathbb{N}^*}$  est un élément de  $l_{\mathbb{K}}^2(\mathbb{N}^*)$ , il existe un unique  $h_\lambda$  dans  $V$  tel que

$$\langle h_\lambda, e_k \rangle = \lambda_k, \quad \forall k \in \mathbb{N}^*;$$

cet élément  $h_\lambda$  est la projection orthogonale sur  $V$  de tout autre vecteur  $h$  de  $H$  tel que

$$\langle h, e_k \rangle = \lambda_k, \quad \forall k \in \mathbb{N}^*.$$

**Preuve.** Pour prouver les points 1 et 2, on introduit la projection orthogonale  $\text{Proj}_V(h)$  et on écrit, du fait du théorème de Pythagore

$$\begin{aligned} \|h\|^2 &= \|h - \text{Proj}_V(h) + \text{Proj}_V(h)\|^2 = \|h - \text{Proj}_V(h)\|^2 + \|\text{Proj}_V(h)\|^2 \\ &\geq \|\text{Proj}_V(h)\|^2. \end{aligned} \quad (1.25)$$

Si  $V_n$  désigne le sous-espace vectoriel (de dimension finie) engendré par  $e_1, \dots, e_n$ , le vecteur

$$\sum_{k=1}^n \langle h, e_k \rangle e_k$$

représente la projection orthogonale de  $h$  sur  $V_n$ . D'après le volet 2 de la proposition 1.7, la suite des projections orthogonales  $\text{Proj}_{V_n}(h)$  converge vers la projection orthogonale sur le sous-espace fermé

$$V = \overline{\bigcup_{k=1}^{\infty} V_k}.$$

On a donc bien

$$\text{Proj}_V(h) = \sum_{k=1}^{\infty} \langle h, e_k \rangle e_k$$

<sup>28</sup>On l'attribue au mathématicien allemand Friedrich Wilhelm Bessel, 1784-1846.

et la formule (1.24) du point 2 est ainsi démontrée. En utilisant le théorème de Pythagore, on a

$$\|\text{Proj}_{V_n}(h)\|^2 = \left\| \sum_{k=1}^n \langle h, e_k \rangle e_k \right\|^2 = \sum_{k=1}^n |\langle h, e_k \rangle|^2$$

puisque le système  $(e_n)_{n \in \mathbb{N}^*}$  est orthonormé. On a donc, par continuité de la norme

$$\lim_{n \rightarrow \infty} \left( \sum_{k=1}^n |\langle h, e_k \rangle|^2 \right) = \|\text{Proj}_V(h)\|^2,$$

et en utilisant (1.25),

$$\lim_{n \rightarrow \infty} \left( \sum_{k=1}^n |\langle h, e_k \rangle|^2 \right) \leq \|h\|^2,$$

ce qui prouve l'inégalité de Bessel (1.23).

Reste à prouver le point 3. Si  $(\lambda_n)_{n \in \mathbb{N}}$  est un élément de  $l_{\mathbb{K}}^2(\mathbb{N}^*)$ , la suite

$$S_\lambda := \left( \sum_{l=1}^n \lambda_l e_l \right)_{n \in \mathbb{N}^*}$$

est de Cauchy dans  $H$  car, pour tout  $p > n \geq 1$ ,

$$\left\| \sum_{l=1}^p \lambda_l e_l - \sum_{l=1}^n \lambda_l e_l \right\|^2 = \left\| \sum_{l=n+1}^p \lambda_l e_l \right\|^2 = \sum_{l=n+1}^p |\lambda_l|^2$$

en utilisant le théorème de Pythagore et le fait que le système  $(e_n)_{n \in \mathbb{N}^*}$  est orthonormé. Cette suite  $S_\lambda$  converge donc (puisque  $H$  est complet) vers un élément  $h_\lambda$  de  $H$ ; en fait  $h_\lambda \in V$  car le terme général  $S_{\lambda,n}$  de la suite  $S_\lambda$  est dans  $V_n$ . On a bien, par continuité du produit scalaire

$$\langle h_\lambda, e_k \rangle := \lim_{n \rightarrow \infty} \left\langle \sum_{l=1}^n \lambda_l e_l, e_k \right\rangle = \lambda_k.$$

Si  $h$  est un autre élément de  $V$  tel que

$$\langle h_\lambda, e_k \rangle = \lambda_k, \quad \forall k \in \mathbb{N}^*,$$

alors  $h - h_\lambda$  est orthogonal à  $V$ , ce qui prouve que  $h_\lambda$  est bien la projection orthogonale de  $h$  sur  $V$ . Le point 3 est ainsi complètement démontré.  $\diamond$

Si  $(v_n)_{n \in \mathbb{N}^*}$  est un système libre dans un espace de Hilbert, il existe un processus algorithmique bien connu des algébristes pour construire à partir de  $v_1, v_2, \dots$  (pris dans cet ordre) un système orthonormé  $(e_n)_{n \in \mathbb{N}^*}$ , ce en faisant en sorte qu'à chaque cran  $n$  de l'algorithme, les sous-espaces vectoriels de dimension finie  $\text{Vec}(v_1, \dots, v_n)$  et  $\text{Vec}(e_1, \dots, e_n)$  coïncident.

Voici ce procédé (dit *d'orthonormalisation de Gram-Schmidt*<sup>29</sup>), décrit à l'étape  $n$ , en supposant que  $e_1, \dots, e_{n-1}$  ont été construits (après avoir démarré avec  $e_1 =$

<sup>29</sup>Si on l'attribue au mathématicien allemand Erhard Schmidt (1876-1959), spécialiste d'analyse fonctionnelle, et au mathématicien danois Jørgen Pedersen Gram (1853-1916), qui s'est penché sur la méthode des moindres carrés, ce procédé algorithmique est certainement bien plus ancien et connu de Cauchy et Laplace dès le début du XIX-ème siècle (Cauchy le manipula expressément vers 1830).

$v_1/\|v_1\|$ ) : on calcule la projection orthogonale de  $v_n$  sur le sous-espace vectoriel  $V_{n-1}$  engendré par  $v_1, \dots, v_{n-1}$  (qui est aussi le sous-espace engendré par  $(e_1, \dots, e_{n-1})$ , cela pouvant se faire de deux façons :

- soit on utilise la formule

$$\text{Proj}_{V_{n-1}}(v_n) = \sum_{k=1}^{n-1} \langle v_n, e_k \rangle e_k ;$$

- soit on calcule la matrice de Gram

$$G_{n-1} := \left[ \langle v_i, v_j \rangle \right]_{1 \leq i, j \leq n-1}$$

(qui est symétrique ou hermitienne), puis on cherche les scalaires  $\lambda_1, \dots, \lambda_{n-1}$  tels que

$$G_{n-1} \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_{n-1} \end{pmatrix} = \begin{pmatrix} \langle v_n, v_1 \rangle \\ \vdots \\ \langle v_n, v_{n-1} \rangle \end{pmatrix} ;$$

on a alors

$$\text{Proj}_{V_{n-1}}(v_n) = \sum_{k=1}^{n-1} \lambda_k v_k .$$

On pose ensuite, le calcul précédent ayant été fait d'une manière ou d'une autre,

$$e_n := \frac{v_n - \text{Proj}_{V_{n-1}}(v_n)}{\|v_n - \text{Proj}_{V_{n-1}}(v_n)\|} .$$

C'est ce procédé qui a été utilisé dans les exemples 1.5 et 1.7 pour construire les fonctions polynômiales de Legendre ou de Hermite à partir du système libre constitué des fonctions monômes  $1, t, t^2, \dots$

## 1.7 Espaces de Hilbert séparables ; notion de base hilbertienne

**Définition 1.8** *Un espace de Hilbert  $H$  (réel ou complexe) est dit séparable si et seulement si s'il existe dans  $H$  un système de vecteurs  $(v_n)_{n \in \mathbb{N}^*}$  tel que l'on ait*

$$\overline{\bigcup_{n=1}^{\infty} \text{Vec}(v_1, \dots, v_n)} = H$$

(on dit aussi que le système  $(v_n)_{n \in \mathbb{N}^*}$  est un système total).

Si  $H$  est un espace de Hilbert séparable, deux configurations sont envisageables :

- soit  $H$  est de dimension finie ;
- soit il existe dans  $H$  un système libre et total de vecteurs  $(v_n)_{n \in \mathbb{N}^*}$ .

Dans l'un ou l'autre des cas, on peut construire (en utilisant le procédé d'orthonormalisation de Gram-Schmidt) soit une base orthonormée  $(e_1, \dots, e_n)$  dans le premier cas, soit un système orthonormé total  $(e_n)_{n \in \mathbb{N}^*}$  dans le second cas. Cela nous conduit naturellement à la définition suivante :

**Définition 1.9** Soit  $H$  un espace de Hilbert séparable. On appelle base hilbertienne de  $H$  tout système orthonormé fini (si  $H$  est de dimension finie) ou total (sinon). Tout espace de Hilbert séparable admet toujours au moins une base hilbertienne.

Le cas des espaces hilbertiens de dimension finie ( $\mathbb{R}^n$  ou  $\mathbb{C}^n$ ) relevant totalement de l'algèbre, nous nous intéresserons surtout ici au cadre des espaces de Hilbert séparables mais non de dimension finie. On a alors la proposition très importante suivante, conséquence immédiate de la proposition 1.8 <sup>30</sup> :

**Proposition 1.9** Soit  $H$  un  $\mathbb{K}$  espace de Hilbert séparable et de dimension infinie ( $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$ ),  $(e_n)_{n \in \mathbb{N}^*}$  une base hilbertienne de  $H$  (il en existe toujours). Tout élément  $h$  de  $H$  se représente comme la somme de la série convergente

$$h = \sum_{k=1}^{\infty} \langle h, e_k \rangle e_k$$

avec

$$\sum_{k=1}^{\infty} |\langle h, e_k \rangle|^2 = \|h\|^2.$$

De plus, l'application

$$\text{inj} : h \in H \longmapsto \left( \langle h, e_k \rangle \right)_{k \in \mathbb{N}^*}$$

est un  $\mathbb{K}$ -isomorphisme isométrique (conservant la norme) entre  $H$  et l'espace de Hilbert  $l_{\mathbb{K}}^2(\mathbb{N}^*)$ . Les scalaires  $\langle h, e_k \rangle$ ,  $k \in \mathbb{N}^*$ , sont dits coordonnées de l'élément  $h$  exprimé dans la base hilbertienne  $(e_n)_{n \in \mathbb{N}^*}$ .

**Preuve.** On applique la proposition 1.8, mais en remarquant cette fois que  $V$  (qui est par définition le plus petit sous-espace fermé contenant tous les vecteurs  $e_n$ ) est tel que  $V = H$ . La première affirmation résulte donc du point 2 de la proposition 1.8. La seconde affirmation tient au fait (voir la preuve de la proposition 1.8) que

$$\|\text{Proj}_V(h)\|^2 = \sum_{k=1}^{\infty} |\langle h, e_k \rangle|^2 = \|h\|^2.$$

La surjectivité de  $\text{inj}$  est une conséquence du point 3 de la proposition 1.8. Le fait que  $\text{inj}$  préserve la norme implique bien sûr l'injectivité.  $\diamond$

Ainsi, à isométrie près, il n'y a qu'un seul  $\mathbb{K}$ -espace de Hilbert séparable (si  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$ , à savoir  $l_{\mathbb{K}}^2(\mathbb{N}^*)$ ). Même si cela est vrai, chaque « incarnation » de cet espace  $l_{\mathbb{K}}^2(\mathbb{N}^*)$  aura son intérêt propre et sera étudiée séparément : le lien en effet entre  $L_{\mathbb{K}}^2(U, dx)$ , (où  $U$  désigne un sous-ensemble mesurable de  $\mathbb{R}^n$ ),  $L^2(\Omega, \mathcal{T}, P)$  (où  $(\Omega, \mathcal{T}, P)$  désigne un espace probabilisé<sup>31</sup>), qui pourtant sont, on le verra, tous les deux des espaces de Hilbert séparables, et  $l_{\mathbb{K}}^2(\mathbb{N}^*)$ ,  $l_{\mathbb{K}}^2(\mathbb{Z})$  (qui en sont d'autres bien sûr) n'est nullement « intuitif » ! Les premiers relèvent plutôt du cadre de l'information continue (à moins que  $P$  ne corresponde à une distribution de probabilités discrète), les deux autres relèvent du cadre des mathématiques discrètes !

<sup>30</sup>Le résultat établi dans cette proposition est indubitablement, avec le théorème 1.1, un des outils majeurs de l'analyse hilbertienne appliquée.

<sup>31</sup>Pourvu que la tribu  $\mathcal{T}$  soit *séparable*, c'est-à-dire puisse être engendrée par une famille dénombrable d'éléments pris dans  $\mathcal{T}$ .



## 1.8 Le cas des espaces de Hilbert non séparables

Dans un espace de Hilbert quelconque (non réduit à  $\{0\}$ , réel ou complexe) il existe bien sûr toujours des systèmes orthonormés. L'ensemble  $\mathcal{E}$  des systèmes orthonormés (dans un espace de Hilbert donné) peut être équipé naturellement d'un ordre : on dira que le système orthonormé  $(e_\mu)_{\mu \in M}$  « précède » le système  $(f_\nu)_{\nu \in N}$  si pour chaque  $\mu$  dans le premier ensemble d'indices  $M$ , il existe  $\nu = \nu(\mu)$  (forcément unique car tout système orthonormé est automatiquement libre) tel que  $e_\mu = f_{\nu(\mu)}$ . On notera ceci

$$(e_\mu)_{\mu \in M} \leq (f_\nu)_{\nu \in N}.$$

Considérons maintenant une famille totalement ordonnée de systèmes orthonormés

$$(e_{\alpha, \mu_\alpha})_{\mu_\alpha \in M_\alpha}, \quad \alpha \in A,$$

de l'espace de Hilbert  $H$  ; ceci signifie que si l'on prend deux systèmes dans cette famille, par exemple

$$(e_{\alpha_1, \mu_{\alpha_1}})_{\mu_{\alpha_1} \in M_{\alpha_1}}, \quad (e_{\alpha_2, \mu_{\alpha_2}})_{\mu_{\alpha_2} \in M_{\alpha_2}},$$

alors soit le premier précède le second pour l'ordre, soit l'inverse. On constate que, si l'on dispose d'une telle famille, on construit encore un système orthonormé en « concaténant » toutes ces familles : en effet si  $e$  et  $e'$  sont deux éléments du système concaténé, ils sont tous les deux dans le même système

$$(e_{\alpha_j, \mu_{\alpha_j}})_{\mu_{\alpha_j} \in M_{\alpha_j}}$$

(soit pour  $j = 1$ , soit pour  $j = 2$ , cela dépend de l'ordre entre les systèmes correspondant aux indices  $\alpha_1$  et  $\alpha_2$ ) et on peut donc calculer leur produit scalaire (qui vaut 0 si ces vecteurs sont distincts, 1 s'ils sont égaux) entre ces deux vecteurs en se positionnant dans le système adéquat. Ce système orthonormé « majorant » réalise en fait la borne supérieure de la famille totalement ordonnée dont on est parti.

Ainsi, toute famille totalement ordonnée de systèmes orthonormés de  $H$  admet toujours une borne supérieure (on dit aussi que l'ordre est *inductif*). L'axiome de Zorn (équivalent de fait à l'axiome du choix) nous assure alors que  $\mathcal{E}$  admet au moins un élément maximal (pour l'ordre que nous venons de définir). Nous remarquons alors :

**Proposition 1.10** *Si  $(e_\iota)_{\iota \in I}$  est un système orthonormé maximal pour l'ordre entre systèmes orthonormés de  $H$  décrit ci-dessus, alors le plus petit sous-espace vectoriel fermé contenant tous les  $e_\iota$ ,  $\iota \in I$ , est  $H$  lui-même ; on dit que le système  $(e_\iota)_{\iota \in I}$  est orthonormé et total, ou encore que c'est une base hilbertienne de l'espace de Hilbert  $H$ .*

**Preuve.** On utilise un raisonnement par l'absurde. Soit  $V$  le plus petit sous-espace vectoriel fermé de  $H$  contenant tous les  $e_\iota$ ,  $\iota \in I$ , et  $h$  un élément de  $H$  n'appartenant pas à  $V$ . Le vecteur

$$\frac{h - \text{Proj}_V(h)}{\|h - \text{Proj}_V(h)\|}$$

formerait, avec les  $e_\iota$ ,  $\iota \in I$ , un système orthonormé strictement plus grand (pour l'ordre mentionné sur les systèmes) que le système pourtant maximal  $(e_\iota)_{\iota \in I}$ . Ceci

est absurde, ce qui montre que l'existence d'un tel  $h$  est impossible, donc que l'on a bien  $V = H$ .  $\diamond$

Voici maintenant la raison pour laquelle les exemples des  $\mathbb{K}$ -espaces de Hilbert  $l_{\mathbb{K}}^2(I)$ ,  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$ ) sont si importants :

**Proposition 1.11** *Soit  $H$  un  $\mathbb{K}$ -espace de Hilbert et  $(e_\iota)_{\iota \in I}$  une base hilbertienne de  $H$  (il en existe d'après l'axiome de Zorn). L'application  $\mathbb{K}$ -linéaire*

$$h \in H \longmapsto \left( \langle h, e_\iota \rangle \right)_{\iota \in I} \in l_{\mathbb{K}}^2(I)$$

*réalise un  $\mathbb{K}$ -isomorphisme isométrique entre  $H$  et  $l_{\mathbb{K}}^2(I)$ .*

**Remarque 1.3.** La proposition ci-dessus nous assure donc qu'à un  $\mathbb{K}$ -isomorphisme isométrique près, les seuls  $\mathbb{K}$ -espaces de Hilbert sont les espaces du type  $l_{\mathbb{K}}^2(I)$  pour un certain ensemble d'indices  $I$ . Lorsque  $I$  est dénombrable, on récupère ainsi tous les espaces de Hilbert séparables.

**Preuve.** Elle est immédiate ; si  $h$  est dans  $H$ , alors, pour tout  $\epsilon > 0$ , il existe une partie  $J(\epsilon)$  de  $I$  telle que, si  $V_{J(\epsilon)}$  désigne le  $\mathbb{K}$ -sous-espace vectoriel engendré par les  $e_\iota$ ,  $\iota \in J(\epsilon)$ ,

$$\|h - \text{Proj}_{V_{J(\epsilon)}}(h)\|^2 = \min_{x \in V_{J(\epsilon)}} \|h - x\|^2 \leq \epsilon.$$

On voit tout de suite que si  $J$  est une partie finie contenant  $J(\epsilon)$

$$\|h - \text{Proj}_{V_J}(h)\|^2 = \min_{x \in V_J} \|h - x\|^2 \leq \min_{x \in V_{J(\epsilon)}} \|h - x\|^2 \leq \epsilon$$

puisque  $V_J$  contient  $V_{J(\epsilon)}$ . La famille  $(\langle h, e_\iota \rangle e_\iota)_{\iota \in I}$  est donc sommable et de somme  $h$ . Ceci induit (avec le théorème de Pythagore) l'égalité

$$\begin{aligned} \|h\|^2 &= \sup_{J \subset I, \#J < \infty} \sum_{\iota \in J} \|\text{Proj}_{V_J}(h)\|^2 \\ &= \sup_{J \subset I, \#J < \infty} \sum_{\iota \in J} |\langle h, e_\iota \rangle|^2. \end{aligned}$$

On utilise ici le fait que si  $K$  est une partie finie de  $I$  ne rencontrant pas  $J(\epsilon)$ , on a

$$\sum_{\iota \in K} |\langle h, e_\iota \rangle|^2 \leq \|h - \text{Proj}_{V_{J(\epsilon)}} h\|^2 \leq \epsilon.$$

Il reste à montrer que si  $(\lambda_\iota)_{\iota \in I}$  est dans  $l_{\mathbb{K}}^2(I)$ , alors la famille

$$\left( \lambda_\iota e_\iota \right)_{\iota \in I}$$

est sommable, de somme un élément  $h_\lambda$  de  $H$  tel que

$$\lambda_\iota = \langle h_\lambda, e_\iota \rangle, \quad \forall \iota \in I.$$

La preuve du premier point repose sur le fait que pour toute partie finie  $K$  de  $I$ , on a

$$\left| \sum_{\iota \in K} \lambda_\iota e_\iota \right|^2 = \sum_{\iota \in K} |\lambda_\iota|^2$$

(cela résulte du théorème de Pythagore) ; on reprend ensuite le raisonnement fait pour montrer la sommabilité de  $(z_\iota \bar{w}_\iota)_{\iota \in I}$  lorsque l'on a deux éléments  $z$  et  $w$  de  $l_{\mathbb{K}}^2(I)$ . Pour le second point, on utilise simplement la continuité de la projection orthogonale sur chaque droite vectorielle  $\mathbb{K}e_\iota$ ,  $\iota \in I$ . La proposition est ainsi démontrée.  $\diamond$

## 1.9 Exemples de bases hilbertiennes dans le cas séparable

### 1.9.1 Densité des classes de fonctions polynômiales dans $L_{\mathbb{K}}^2(U, dx)$ pour $U$ mesurable borné

On rappelle<sup>32</sup> que si  $(\epsilon_k)_{k \geq 1}$  et si  $g$  désigne la gaussienne

$$(x_1, \dots, x_n) \longmapsto \frac{1}{(2\pi)^{n/2}} \exp(-\|x\|^2/2),$$

la suite  $(\dot{g}_{\epsilon_k})_{k \geq 1}$ , où  $g_{\epsilon}(x) := \epsilon^{-n} g(x/\epsilon)$  constitue une approximation de la masse de Dirac dans  $L_{\mathbb{R}}^1(\mathbb{R}^n, \mathcal{L}, dx)$  (au sens de la définition 4.3 du cours de théorie de l'intégration). On en déduit, en utilisant la proposition 4.10 du cours de théorie de l'intégration que, si  $f$  est un élément de  $\mathcal{L}_{\mathbb{K}}^2(\mathbb{R}^n, dx)$ , la suite  $(\dot{f} * \dot{g}_{\epsilon_k})_{k \geq 1}$ , qui est une suite d'éléments de  $L_{\mathbb{K}}^2(\mathbb{R}^n, dx)$  d'après le théorème de Young<sup>33</sup>, converge vers  $\dot{f}$  dans l'espace de Hilbert  $L_{\mathbb{K}}^2(\mathbb{R}^n, dx)$ .

Si  $\dot{f}$  a un représentant identiquement nul  $dx$ -presque partout hors d'un compact  $K$  de  $\mathbb{R}^n$ , un représentant de  $\dot{f} * \dot{g}_{\epsilon}$  est la fonction

$$\begin{aligned} x \longmapsto \int_K g_{\epsilon}(x-y) f(y) dy &= \frac{1}{\epsilon^n (2\pi)^{n/2}} \int_K \left( \sum_{k=0}^{\infty} \frac{(-1)^k}{2^k k! \epsilon^{2k}} \|x-y\|^{2k} \right) f(y) dy \\ &= \frac{1}{\epsilon^n (2\pi)^{n/2}} \sum_{k=0}^{\infty} \frac{(-1)^k}{2^k k! \epsilon^{2k}} \int_K \|x-y\|^{2k} f(y) dy \end{aligned}$$

qui se présente donc comme la limite uniforme sur  $K$  de la suite de fonctions polynômiales  $(P_N)_N$ , où

$$P_N(x) := \frac{1}{\epsilon^n (2\pi)^{n/2}} \sum_{k=0}^N \frac{(-1)^k}{2^k k! \epsilon^{2k}} \int_K \|x-y\|^{2k} f(y) dy.$$

On déduit de toutes ses remarques la proposition suivante :

**Proposition 1.12** *Si  $U$  est un sous-ensemble mesurable borné de  $\mathbb{R}^n$ , les classes de fonctions polynômiales à coefficients dans le corps  $\mathbb{K}$  ( $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$ ) engendrent un  $\mathbb{K}$ -sous espace dense dans le  $\mathbb{K}$ -espace de Hilbert  $L_{\mathbb{K}}^2(U, dx)$ .*

**Exemple 1.5 revisité.** Suite à cette proposition, nous pouvons donc affirmer que les classes  $(\dot{L}_n)_n$  des polynômes de Legendre introduits dans l'exemple 1.5 forment une base Hilbertienne du  $\mathbb{K}$ -espace de Hilbert  $L_{\mathbb{K}}^2([-1, 1], dt)$ . Ce système se construit pas-à-pas en utilisant le procédé d'orthonormalisation de Gram-Schmidt à partir du système libre constitué des fonctions monômes  $t \longmapsto t^n$ ,  $n = 0, 1, \dots$

<sup>32</sup>Voir l'exemple 4.2 du cours de théorie de l'intégration (UE MHT512), polycopié en ligne : <http://www.math.u-bordeaux1.fr/~yger/mht512.pdf>

<sup>33</sup>Théorème 4.2 du cours de MHT512.

### 1.9.2 Densité des classes de polynômes trigonométriques dans $L_{\mathbb{K}}^2(\mathbb{T}^n)$

Le but de cette sous-section est de montrer dans un premier temps que les classes de monômes trigonométriques complexes

$$e_k := (\theta_1, \dots, \theta_n) \mapsto \exp(i(k_1\theta_1 + \dots + k_n\theta_n)) = \exp(i\langle k, \theta \rangle), \quad k \in \mathbb{Z}^n,$$

engendrent un  $\mathbb{C}$ -sous-espace dense dans le  $\mathbb{C}$ -espace de Hilbert  $L_{\mathbb{C}}^2(\mathbb{T}^n)$ . On en déduira immédiatement que les classes des monômes trigonométriques réels

$$(\theta_1, \dots, \theta_n) \mapsto \prod_{j=1}^n (\cos \theta_j)^{k_j} \times \prod_{j=1}^n (\sin \theta_j)^{l_j}, \quad k, l \in \mathbb{N}^n,$$

engendrent un  $\mathbb{K}$ -sous-espace vectoriel dense dans  $L_{\mathbb{K}}^2(\mathbb{T}^n)$  ( $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$ ). Le matériel introduit dans cette sous-section sera repris au chapitre 2.

Nous introduisons pour cela le concept d'approximation de l'unité dans  $L_{\mathbb{R}}^2(\mathbb{T}^n)$ .

**Définition 1.10** Une approximation de la masse de Dirac dans  $L_{\mathbb{R}}^1(\mathbb{T}^n)$  est une suite  $(\dot{\psi}_k)_{k \in \mathbb{N}^*}$  d'éléments de  $L_{\mathbb{R}}^1(\mathbb{T}^n)$  telle que :

- pour tout  $k \in \mathbb{N}^*$ ,  $\dot{\psi}_k$  a un représentant  $(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n))$ - ( $\mathbb{R}, \mathcal{L}$ )-mesurable,  $2\pi$ -périodique en chaque variable  $\theta_1, \dots, \theta_n$ , tel que  $\psi_k \geq 0$  sur  $[0, 2\pi[^n$  (donc sur  $\mathbb{R}^n$ );
- $\forall k \in \mathbb{N}^*$ ,  $\|\dot{\psi}_k\|_{\mathbb{T}, 1} := \frac{1}{(2\pi)^n} \int_{[0, 2\pi[^n} |\dot{\psi}_k(\theta)| d\theta = \frac{1}{(2\pi)^n} \int_{[0, 2\pi[^n} \psi_k(\theta) d\theta = 1$ ;
- pour tout  $\delta \in [0, \pi[$ ,

$$\lim_{k \rightarrow +\infty} \left( \int_{\substack{\delta < |\theta_j| < \pi \\ j=1, \dots, n}} \psi_k(\theta) d\theta \right) = 0, \quad (1.26)$$

conditions traduisant bien le fait que toute la « masse » (égale à 1) de  $\dot{\psi}_k$  se concentre sur l'origine  $\dot{0}$  de  $\mathbb{T}^n$  lorsque  $k$  tend vers  $+\infty$ .

**Exemple 1.10.** Soit, pour  $k \geq 1$ ,

$$\Psi_k(\theta) := (2\pi)^n \frac{\prod_{j=1}^n \left( \frac{1 + \cos \theta_j}{2} \right)^k}{\int_{]-\pi, \pi[^n} \prod_{j=1}^n \left( \frac{1 + \cos u_j}{2} \right)^k du}. \quad (1.27)$$

Chaque fonction  $\Psi_k$ ,  $k \in \mathbb{N}^*$ , est une fonction  $2\pi$ -périodique positive ou nulle en les  $n$  variables (c'est même un polynôme trigonométrique) et on a par construction même  $\|\dot{\Psi}_k\|_{\mathbb{T}, 1} = 1$ ; de plus, si  $\delta \in ]0, \pi[$ , la convergence uniforme vers 0 sur  $[-\pi, -\delta] \cup [\delta, \pi]$  de la suite de fonctions

$$\theta \mapsto \frac{\left( \frac{1 + \cos \theta}{2} \right)^k}{4 \int_0^{\pi/2} (\sin u)^{2k} du} \leq \frac{\left( \frac{1 + \cos \theta}{2} \right)^k}{4 \int_0^{\pi/2} (2u/\pi)^{2k} du} \leq \frac{(2k+1)(\cos(\theta/2))^{2k}}{2\pi}$$

nous assure que la condition (1.26) est satisfaite et donc que la suite  $(\dot{\Psi}_k)_{k \geq 1}$  réalise une approximation de la masse de Dirac dans  $L_{\mathbb{R}}^1(\mathbb{T}^n, \mathcal{B}(\mathbb{T}^n), d\dot{\theta})$ , approximation qui, il faut le souligner, est donnée par des classes de polynômes trigonométriques.

On introduit la convolution d'un  $\dot{g} \overset{\text{per}}{*} \dot{f}$  d'un élément  $\dot{g}$  de  $L_{\mathbb{K}}^p(\mathbb{T}^n)$  et d'un élément  $\dot{f}$  de  $L_{\mathbb{K}}^q(\mathbb{T}^n)$  lorsque  $p$  et  $q$  sont deux éléments de  $[1, +\infty]$  tels que  $1/p + 1/q = 1 + 1/r$ ,  $r \in [1, \infty]$  comme étant l'élément

$$\dot{g} \overset{\text{per}}{*} \dot{f}$$

de  $L^r(\mathbb{T}^n)$  dont un représentant est la fonction mesurable  $2\pi$ -périodique en les  $n$  variables  $\theta_1, \dots, \theta_n$

$$\theta \longmapsto \frac{1}{(2\pi)^n} \int_{[0, 2\pi[^n} g(\theta - u) f(u) du = \frac{1}{(2\pi)^n} \int_{[0, 2\pi[^n} g(u) f(\theta - u) du. \quad (1.28)$$

On a d'ailleurs, d'après les inégalités de Young<sup>34</sup> transposables à ce cadre périodique,

$$\|\dot{g} \overset{\text{per}}{*} \dot{f}\|_r \leq (2\pi)^n \|\dot{g}\|_p \times \|\dot{f}\|_q \quad (1.29)$$

si

$$\|\dot{g}\|_p := \left( \frac{1}{(2\pi)^n} \int_{[0, 2\pi[^n} |g(\theta)|^p d\theta \right)^{1/p}$$

et

$$\|\dot{f}\|_q := \left( \frac{1}{(2\pi)^n} \int_{[0, 2\pi[^n} |f(\theta)|^q d\theta \right)^{1/q}.$$

On peut en particulier appliquer cette opération à partir d'un élément  $\dot{g}$  de  $L_{\mathbb{K}}^1(\mathbb{T}^n)$  et d'un élément  $\dot{f}$  de  $L_{\mathbb{K}}^2(\mathbb{T}^n)$ , ce qui nous donne un élément

$$\dot{g} \overset{\text{per}}{*} \dot{f} \in L_{\mathbb{K}}^2(\mathbb{T}^n).$$

Dans exactement la même veine que la proposition 4.10 du cours de théorie de l'intégration, nous avons (la preuve est quasiment identique) la proposition suivante :

**Proposition 1.13** *Soit  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$  et  $\dot{f} \in L_{\mathbb{K}}^2(\mathbb{T}^n)$ . Si  $(\dot{\psi}_k)_{k \geq 1}$  est une approximation de la masse de Dirac dans  $L_{\mathbb{R}}^1(\mathbb{T}^n)$ , la suite*

$$(\dot{\psi}_k \overset{\text{per}}{*} \dot{f})_{k \geq 1}$$

*est une suite d'éléments de  $L_{\mathbb{K}}^2(\mathbb{T}^n)$  convergeant vers  $\dot{f}$  dans le  $\mathbb{K}$ -espace de Hilbert  $L_{\mathbb{K}}^2(\mathbb{T}^n)$ .*

Le corollaire suivant de cette proposition 1.13 sera particulièrement important en ce qui concerne ultérieurement l'analyse de Fourier des signaux ou des images périodiques :

**Corollaire 1.1** *Les classes de monômes trigonométriques complexes*

$$e_k : \theta \longmapsto \exp(i(k_1\theta_1 + \dots + k_n\theta_n)) = \exp(i\langle k, \theta \rangle), \quad (k_1, \dots, k_n) \in \mathbb{Z}^n,$$

*engendrent un  $\mathbb{C}$ -sous-espace vectoriel dense dans le  $\mathbb{K}$ -espace de Hilbert  $L_{\mathbb{C}}^2(\mathbb{T}^n)$ . Les classes des monômes trigonométriques réels*

$$(\theta_1, \dots, \theta_n) \longmapsto \prod_{j=1}^n (\cos \theta_j)^{k_j} \times \prod_{j=1}^n (\sin \theta_j)^{l_j}, \quad k, l \in \mathbb{N}^n$$

*engendrent un  $\mathbb{K}$ -sous-espace dense dans  $L_{\mathbb{K}}^2(\mathbb{T}^n)$  ( $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$ ).*

<sup>34</sup>Voir le théorème 4.2 du cours de MHT512 et sa transposition au cadre périodique dans la définition 4.5 de ce même cours.

**Preuve.** Il suffit de se souvenir que  $(\dot{\Psi}_k)_{k \geq 1}$  (où  $\Psi_k$  est le polynôme trigonométrique défini par (1.27)) est une approximation de la masse de Dirac dans  $L^1_{\mathbb{R}}(\mathbb{T}^n)$  (voir l'exemple 1.10), d'appliquer la proposition 1.13, et de remarquer dans un second temps que si  $f \in L^2_{\mathbb{K}}(\mathbb{T}^n)$ , la fonction

$$\theta \longmapsto \int_{[0, 2\pi[^n} f(\theta - u) \Psi_k(u) du = \int_{[0, 2\pi[^n} f(u) \Psi_k(\theta - u) du$$

est un polynôme trigonométrique complexe car, pour tout  $k = (k_1, \dots, k_n) \in \mathbb{Z}^n$ , pour tout  $\theta \in \mathbb{R}^n$

$$\int_{[0, 2\pi[^n} f(u) e^{i\langle k, \theta - u \rangle} du = e^{i\langle k, \theta \rangle} \times \int_{[0, 2\pi[^n} f(u) e^{-i\langle k, u \rangle} du,$$

où  $\langle \cdot, \cdot \rangle$  désigne le produit scalaire canonique dans  $\mathbb{R}^n$ . Le passage des monômes trigonométriques complexes aux monômes trigonométriques réels se fait à l'aide des formules de Moivre et d'Euler.  $\diamond$

**Exemple 1.4 revisité.** Il résulte du résultat établi dans l'exemple 1.4 (transposé du cadre monodimensionnel au cadre multidimensionnel, ce qui se fait immédiatement) que les classes des monômes trigonométriques complexes

$$e_k : (\theta_1, \dots, \theta_n) \longmapsto \exp(i(k_1\theta_1 + \dots + k_n\theta_n)), \quad k \in \mathbb{Z}^n,$$

forment un système orthonormé dans  $L^2_{\mathbb{C}}(\mathbb{T}^n)$ . Du fait du corollaire 1.1, ces classes  $e_k$ ,  $k \in \mathbb{Z}^n$ , forment une base hilbertienne du  $\mathbb{C}$ -espace de Hilbert  $L^2_{\mathbb{C}}(\mathbb{T}^n)$ . Cette base hilbertienne jouera un rôle majeur dans l'analyse de Fourier des phénomènes périodiques ; on la retrouvera au chapitre 2.

**Exemple 1.6 revisité.** Si  $\hat{f}$  est un élément de  $L^2_{\mathbb{C}}([-1, 1], (1 - t^2)^{-1/2} dt)$  et que  $f$  soit un représentant de  $\hat{f}$ , la formule de changement de variables dans les intégrales<sup>35</sup> assure que

$$\int_{[-1, 1]} |f(t)|^2 \frac{dt}{\sqrt{1 - t^2}} = \int_0^\pi |f(\cos \theta)|^2 \frac{\sin \theta d\theta}{\sqrt{1 - \cos^2 \theta}} = \int_0^\pi |f(\cos \theta)|^2 d\theta < +\infty.$$

Or, d'après le corollaire 1.1, les monômes trigonométriques réels pairs

$$\theta \longmapsto \cos(n\theta), \quad n \in \mathbb{N},$$

engendrent un  $\mathbb{K}$ -sous-espace dense dans le  $\mathbb{K}$ -sous-espace de  $L^2_{\mathbb{K}}(\mathbb{T})$  constitué des classes de fonctions  $2\pi$ -périodiques paires à valeurs dans  $\mathbb{K}$  (ici  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{K} = \mathbb{C}$ ). On en déduit que les classes des polynômes de Tchebychev  $\Theta_n$  définis précisément (voir (1.21)) par

$$\Theta_n(\cos \theta) \equiv \cos(n\theta), \quad n \in \mathbb{N},$$

engendrent un  $\mathbb{K}$ -sous-espace dense de  $L^2_{\mathbb{K}}([-1, 1], (1 - t^2)^{-1/2} dt)$  et par conséquent, compte tenu de leur orthogonalité deux à deux (voir l'exemple 1.6) forment une base hilbertienne de ce  $\mathbb{K}$ -espace de Hilbert (toujours si  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{K} = \mathbb{C}$ ). Ce système peut se reconstruire *via* le procédé d'orthonormalisation de Gram-Schmidt à partir du système libre constitué des classes des monômes  $t^n$ ,  $n = 0, 1, \dots$

### 1.9.3 A propos des polynômes de Hermite

Nous verrons au chapitre 2 que les classes de fonctions polynômiales (à coefficients dans  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$ ) engendrent un  $\mathbb{K}$ -sous-espace dense dans le  $\mathbb{K}$ -espace de

<sup>35</sup>Formule (3.11) dans l'énoncé du théorème 3.5 du cours de MHT512.

Hilbert  $L_{\mathbb{K}}^2(\mathbb{R}, e^{-t^2} dt)$  introduit dans l'exemple 1.7. Il en résulte que les classes  $\dot{H}_n$ ,  $n \in \mathbb{N}$  des fonctions polynômiales de Hermite définies par

$$H_n(t) := \frac{(-1)^n}{2^{n/2} \pi^{1/4} \sqrt{n!}} \times e^{t^2} \times \frac{d^n}{dt^n} [e^{-t^2}], \quad n \in \mathbb{N}, \quad (1.30)$$

forment une base hilbertienne de  $L_{\mathbb{K}}^2(\mathbb{R}, e^{-t^2} dt)$ . Ce système peut ici encore, se reconstruire *via* le procédé d'orthonormalisation de Gram-Schmidt à partir du système libre constitué des classes des monômes  $t^n$ ,  $n = 0, 1, \dots$

## 1.10 Le théorème de dualité

Nous revenons maintenant à l'étude générale des propriétés des espaces de Hilbert. Nous donnons dans cette section une application clef du théorème de projection, qui nous dit que le dual d'un espace de Hilbert (au sens topologique) s'identifie isométriquement à l'espace lui-même. Rappelons que le dual (au sens topologique) du  $\mathbb{K}$ -espace de Hilbert  $H$  est par définition le  $\mathbb{K}$ -espace des formes  $\mathbb{K}$ -linéaires continues de  $H$  dans  $\mathbb{K}$  (« continues » s'entendant naturellement au sens de la topologie définie par la norme euclidienne  $\| \cdot \|$ ).

### 1.10.1 Quelques préliminaires à propos des formes linéaires continues sur un Hilbert

Le fait qu'une forme  $\mathbb{K}$ -linéaire  $L : H \rightarrow K$  soit continue équivaut au fait qu'il existe une constante  $C$  telle que

$$\forall x \in H, \|x\| \leq 1 \implies |L(x)| \leq C.$$

Si tel est le cas, on définit la *norme* de  $L$  par

$$\|L\|^* := \sup_{\{x; \|x\| \leq 1\}} |L(x)| = \sup_{\{x; \|x\|=1\}} |L(x)|.$$

Cette norme  $\| \cdot \|^*$  est dite norme de l'espace normé dual du  $\mathbb{K}$ -espace de Hilbert  $(H, \| \cdot \|)$  (la norme  $\| \cdot \|$  étant la norme dérivant du produit scalaire sur  $H$ )<sup>36</sup>.

Dans la pratique, on ne dispose bien souvent de la description de l'action d'une forme linéaire  $L$  que sur un sous-espace vectoriel dense dans  $H$  (et non explicitement sur  $H$ ); un exemple important est celui d'un espace de Hilbert séparable où l'on dispose d'une base hilbertienne  $(e_k)_{k \in \mathbb{Z}}$  (ou d'un « *frame* » engendrant un sous-espace dense  $V$ ) et où l'on ne connaît l'action de  $L$  que par le biais des  $L(e_k)$ ,  $k \in \mathbb{Z}$ . On ne connaît donc *a priori* que l'action de  $L$  sur les combinaisons linéaires finies des vecteurs  $(e_k)_{k \in \mathbb{Z}}$  :

$$L\left(\sum_{-N}^N \lambda_k e_k\right) := \sum_{-N}^N \lambda_k L(e_k), \quad N \in \mathbb{N}^*.$$

Une question naturelle se pose : celle de « prolonger » sans ambiguïté  $L$  à l'espace de Hilbert tout entier (en une forme linéaire continue bien sûr) de manière à pouvoir vraiment travailler avec un élément du dual de  $H$ . Le petit lemme de prolongement suivant, anodin à première vue, est en fait très important du point de vue pratique :

<sup>36</sup>Pour ne pas alourdir les notations, on omettra par la suite de faire figurer l'exposant  $*$  dans l'expression de la norme duale; on la notera aussi  $\| \cdot \|$  par souci de simplicité, ce lorsque la confusion avec la norme  $\| \cdot \|$  sur  $H$  s'avèrera impossible.

**Lemme 1.1 (Lemme de prolongement des formes linéaires)** *Soit  $V$  un  $\mathbb{K}$ -sous-espace dense (mais attention, pas nécessairement fermé, c'est le problème!) de  $H$  et  $L$  une forme linéaire de  $V$  dans  $H$ . Une condition nécessaire et suffisante pour que l'on puisse prolonger  $L$  en un élément du dual topologique de  $H$  est qu'il existe une constante positive  $C$  telle que*

$$\forall v \in V, |L(v)| \leq C\|v\|.$$

**Preuve.** Si  $L$  se prolonge de  $V$  à  $H$  en une forme linéaire continue, on peut prendre  $C = \|L\|$  et l'inégalité  $|L(h)| \leq C\|h\|$  est satisfaite pour tout  $h$  dans  $H$ , donc *a fortiori* pour tout  $v$  dans  $V$ . Pour la réciproque, on approche un élément arbitraire  $h$  de  $H$  par une suite  $(v_n)_{n \geq 0}$  d'éléments de  $V$  ( $V$  est dense, c'est donc possible). On remarque que, puisque

$$|L(v_n) - L(v_m)| = |L(v_n - v_m)| \leq C\|v_n - v_m\|, \quad \forall m, n \in \mathbb{N},$$

la suite  $(L(v_n))_{n \geq 0}$  est de Cauchy dans le corps  $\mathbb{K}$  comme la suite convergente  $(v_n)_{n \geq 0}$ ; elle converge donc et l'on peut poser

$$L(h) := \lim_{n \rightarrow +\infty} L(v_n)$$

après avoir remarqué que cette limite ne dépendait pas en fait de la suite  $(v_n)_{n \geq 0}$  choisie pour approcher  $h$ . L'action de  $L$  est ainsi prolongée à  $H$  tout entier et la forme linéaire  $L : H \rightarrow \mathbb{K}$  définie ainsi prolonge l'action de  $L$  sur  $V$ , est unique et continue car on a bien sûr

$$\forall h \in H, |L(h)| \leq C\|h\|.$$

Le résultat repose donc essentiellement sur la complétude de  $\mathbb{K}$ .  $\diamond$

### 1.10.2 L'énoncé du théorème de dualité

**Théorème 1.2** *Soit  $L$  une forme linéaire continue sur un  $\mathbb{K}$ -espace de Hilbert  $H$  (le corps de base  $\mathbb{K}$  étant  $\mathbb{R}$  ou  $\mathbb{C}$ ); il existe un et un seul  $h \in H$  tel que*

$$\forall x \in H, L(x) = \langle x, h \rangle.$$

*De plus, l'application*

$$L \longmapsto h = h(L)$$

*est une isométrie anti-linéaire<sup>37</sup> entre  $H^*$  (muni de la norme  $\|\cdot\|^*$  d'espace normé dual définie ci-dessus) et  $H$ ; le dual  $H^*$  hérite donc naturellement ainsi d'une structure de  $\mathbb{K}$ -espace de Hilbert avec un produit scalaire défini par*

$$\langle L_1, L_2 \rangle = \langle h(L_2), h(L_1) \rangle$$

<sup>37</sup>L'anti-linéarité signifie que pour toute paire de formes  $\mathbb{K}$ -linéaires  $L_1, L_2$ , pour toute paire de scalaires  $\lambda_1, \lambda_2$ ,

$$h(\lambda_1 L_1 + \lambda_2 L_2) = \overline{\lambda_1} h(L_1) + \overline{\lambda_2} h(L_2).$$



**Preuve.** Le sous espace  $F = \text{Ker } L$  est un sous espace fermé  $F$  de  $H$ . A moins que  $L$  ne soit l'application nulle (auquel cas  $h = 0$  convient dans l'énoncé du théorème),  $F$  est strictement inclus dans  $H$ . De plus, d'après le théorème de projection orthogonale 1.1,  $H$  est somme directe orthogonale de  $F$  et  $F^\perp$  puisque tout  $x \in H$  s'écrit

$$x = \text{Proj}_F(x) + \left(x - \text{Proj}_F(x)\right) \in F \overset{\perp}{\oplus} F^\perp,$$

la somme étant directe. On peut par conséquent toujours trouver (lorsque  $L \neq 0$ ) au moins un vecteur  $u$  non nul dans  $F^\perp$ .

Remarquons aussi (toujours si  $L \neq 0$ ) qu'il est impossible de trouver dans  $F^\perp$  deux vecteurs indépendants  $u_1$  et  $u_2$  : en effet, si tel était le cas, on aurait  $L(u_1) = \lambda L(u_2)$  pour un certain  $\lambda \in \mathbb{C}^*$  et par conséquent  $u_1 - \lambda u_2 \in F \cap F^\perp = \{0\}$ , ce qui contredirait l'indépendance de  $u_1$  et  $u_2$ . On a donc

$$H = F \overset{\perp}{\oplus} \mathbb{C}u.$$

Lorsque  $L \neq 0$ , choisissons un vecteur non nul  $u$  de  $F^\perp$ , c'est-à-dire, d'après ce qui précède, une base de  $F^\perp$ . Pour tout  $x \in F$ , on a  $L(x) = \langle x, u \rangle = 0$ . Posons

$$h := \frac{\overline{L(u)}}{\|u\|^2} u.$$

On a, avec ce choix,

$$\forall x \in H, L(x) = \langle x, h \rangle$$

et la première partie du théorème est prouvée. L'unicité de  $h$  est immédiate, car

$$\langle x, h_1 \rangle = \langle x, h_2 \rangle, \quad \forall x \in H$$

implique en particulier  $\langle h_1 - h_2, h_1 - h_2 \rangle = 0$ , soit  $h_1 = h_2$ . Enfin, on a, du fait de l'inégalité de Cauchy-Schwarz (1.9) (avec son cas d'égalité),

$$\|L\| = \sup_{\|x\|=1} |\langle x, h \rangle| = \|h\|,$$

ce qui montre bien que

$$L \longmapsto h = h(L)$$

est une isométrie. L'antilinearité de cette application est évidente.  $\diamond$

## 1.11 Opérateurs de $H$ dans $H$ ; notion d'adjoint

Après nous être intéressés à décrire le  $\mathbb{K}$ -espace vectoriel des formes linéaires continues sur un  $\mathbb{K}$ -espace de Hilbert  $H$ , nous allons dans cette section nous intéresser à l'étude (très succincte) des opérateurs  $\mathbb{K}$ -linéaires continus d'un  $\mathbb{K}$ -espace de Hilbert dans lui-même.

### 1.11.1 Préliminaires sur les opérateurs linéaires continus de $H$ dans $H$

Soit  $T : H \longrightarrow H$  une application  $\mathbb{K}$ -linéaire du  $\mathbb{K}$ -espace de Hilbert  $H$  dans lui-même<sup>38</sup>. Dire que  $T$  est continue (pour la topologie d'espace normé définie par la norme  $\| \cdot \|$  dérivant du produit scalaire) équivaut à dire qu'il existe une constante  $C$  telle que

$$\forall h \in H, \|h\| \leq 1 \implies \|T(h)\| \leq C.$$

On pose dans ce cas

$$\|T\| := \sup_{\{h \in H; \|h\| \leq 1\}} \|T(h)\| = \sup_{\{h \in H; \|h\|=1\}} \|T(h)\|. \quad (1.31)$$

Comme pour les formes linéaires, dans de nombreux problèmes pratiques, on ne dispose de la description de l'action d'un opérateur que sur un sous-espace vectoriel dense dans  $H$  (et non explicitement sur  $H$ ); un exemple important est celui d'un espace de Hilbert séparable où l'on dispose d'une base hilbertienne  $(e_k)_{k \in \mathbb{Z}}$  (ou d'un « *frame* » engendrant un sous-espace dense  $V$ ) et où l'on ne connaît l'action de  $T$  que par le biais des  $T(e_k)$ ,  $k \in \mathbb{Z}$ . On ne connaît donc *a priori* que l'action de  $T$  sur les combinaisons linéaires finies des vecteurs  $(e_k)_{k \in \mathbb{Z}}$  :

$$T\left(\sum_{k=-N}^N \lambda_k e_k\right) := \sum_{k=-N}^N \lambda_k T(e_k), \quad N \in \mathbb{N}^*.$$

Une question naturelle se pose : celle de « prolonger » sans ambiguïté  $T$  à l'espace de Hilbert tout entier (en un opérateur  $\mathbb{K}$ -linéaire continu bien sûr) de manière à pouvoir vraiment travailler avec un opérateur  $\mathbb{K}$  linéaire continu de  $H$  dans lui-même. Le petit lemme de prolongement suivant, anodin à première vue, est en fait très important du point de vue pratique :

**Lemme 1.2 (Lemme de prolongement des opérateurs)** *Soit  $V$  un  $\mathbb{K}$ -sous-espace dense (mais attention, pas nécessairement fermé, c'est tout le problème!) de  $H$  et  $T$  une application  $\mathbb{K}$ -linéaire de  $V$  dans  $H$ . Une condition nécessaire et suffisante pour que l'on puisse prolonger  $T$  de manière unique en un opérateur  $\mathbb{K}$ -linéaire continu de  $H$  dans lui-même est qu'il existe une constante positive  $C$  telle que*

$$\forall v \in V, \|T(v)\| \leq C\|v\|.$$

**Preuve.** Si  $T$  se prolonge de  $V$  à  $H$  en un opérateur linéaire continu, on peut prendre  $C = \|T\|$  et l'inégalité  $\|T(h)\| \leq C\|h\|$  est satisfaite pour tout  $h$  dans  $H$ , donc *a fortiori* pour tout  $v$  dans  $V$ . Pour la réciproque, on approche un élément arbitraire  $h$  de  $H$  par une suite  $(v_n)_{n \geq 0}$  d'éléments de  $V$  ( $V$  est dense, c'est donc possible). On remarque que, puisque

$$\|T(v_n) - T(v_m)\| = \|T(v_n - v_m)\| \leq C\|v_n - v_m\|, \quad \forall m, n \in \mathbb{N},$$

la suite  $(T(v_n))_{n \geq 0}$  est de Cauchy comme la suite convergente  $(v_n)_{n \geq 0}$ ; elle converge donc et l'on peut poser

$$T(h) := \lim_{n \rightarrow +\infty} T(v_n)$$

<sup>38</sup>On dit aussi un *opérateur* linéaire continu de  $H$  dans  $H$ .

après avoir remarqué que cette limite ne dépendait pas en fait de la suite  $(v_n)_{n \geq 0}$  choisie pour approcher  $h$ . L'action de  $T$  est ainsi prolongée à  $H$  tout entier et le nouvel opérateur  $T : H \rightarrow H$  défini ainsi prolonge l'action de  $T$  sur  $V$ , est unique et continu car on a bien sûr

$$\forall h \in H, \|T(h)\| \leq C\|h\|.$$

Le résultat repose donc sur la complétude de  $H$  en place de celle de  $\mathbb{K}$  utilisée pour le lemme 1.1.  $\diamond$

**Exemple d'application 1.11 (thème d'exercice).** Voici un exemple d'application important de ce lemme de prolongement dans un  $\mathbb{K}$ -espace de Hilbert séparable où l'on dispose d'une base hilbertienne  $(e_k)_{k \in \mathbb{Z}}$ . On se donne l'action d'un opérateur linéaire  $T$  en se la donnant uniquement sur les vecteurs  $e_k$  de la base hilbertienne. On veille simplement à ce que l'action de  $T$  ne détériore pas trop l'orthogonalité, au sens suivant : il existe une suite  $(\tau_k)_{k \in \mathbb{Z}}$  de nombres positifs telle que

$$\|\tau\|_1 := \sum_{k \in \mathbb{Z}} \tau_k < \infty$$

et

$$\forall k, l \in \mathbb{Z}, |\langle T(e_k), T(e_l) \rangle| \leq \tau_{k-l}$$

(on parle alors pour la famille  $(T(e_k))_k$  de famille « presque orthogonale »). Alors, si  $N$  est un entier positif non nul et  $(\lambda_k)_{k=-N}^N$  des scalaires, on a, du fait du théorème de Pythagore,

$$\begin{aligned} \left\| T \left( \sum_{k=-N}^N \lambda_k e_k \right) \right\|^2 &= \sum_{k=-N}^N \sum_{l=-N}^N \lambda_k \bar{\lambda}_l \langle T(e_k), T(e_l) \rangle \\ &\leq \sum_{k=-N}^N \sum_{l=-N}^N |\lambda_k| |\lambda_l| \tau_{k-l} \\ &\leq \sum_{l=-N}^N \sum_{p=-l-N}^{-l+N} |\lambda_l| |\lambda_{l+p}| \tau_p \\ &\leq \sum_{p=-2N}^{2N} \tau_p \left( \sum_l |\lambda_l| |\lambda_{l+p}| \right) \\ &\leq \sum_{p \in \mathbb{Z}} \tau_p \left( \sqrt{\sum_l |\lambda_l|^2} \sqrt{\sum_l |\lambda_{l+p}|^2} \right) \\ &\leq \|\tau\|_1 \times \sum_{k=-N}^N |\lambda_k|^2 \end{aligned}$$

si l'on convient de prolonger la collection  $(\lambda_k)_k$  par des zéros hors de  $\{-N, \dots, N\}$ , que l'on applique l'inégalité de Cauchy-Schwarz (1.9) et l'invariance par translation de la mesure de comptage sur  $\mathbb{Z}$ . On voit qu'il existe une constante  $C = \sqrt{\|\tau\|_1}$  telle que, pour tout  $h$  dans le sous-espace vectoriel engendré par les  $e_k$ ,  $k \in \mathbb{Z}$ , on ait  $\|T(h)\| \leq C\|h\|$ . Le lemme de prolongement s'applique donc et l'action de  $T$  se prolonge bien en celle d'un opérateur linéaire continu de  $H$  dans lui-même.

Par exemple, si  $H = L^2_{\mathbb{C}}(\mathbb{T})$  et si

$$e_k : \theta \mapsto e^{ik\theta}, \quad k \in \mathbb{Z},$$

on définit un opérateur linéaire continu de  $H$  dans lui-même en posant

$$T(e_k) : \theta \mapsto a_k(\theta) e^{ik\theta}, \quad \forall k \in \mathbb{Z},$$

où  $a_k$  désigne (pour tout  $k \in \mathbb{Z}$ ) une fonction  $C^\infty$   $2\pi$ -périodique<sup>39</sup>, avec la contrainte qu'il existe une constante absolue  $K \geq 0$  telle que

$$\forall k \in \mathbb{Z}, \|a_k\|_\infty + \|a'_k\|_\infty + \|a''_k\|_\infty \leq K.$$

<sup>39</sup>Il s'agit concrètement d'un mécanisme de *modulation de l'amplitude*, dans la mesure où l'on remplace l'amplitude constante égale à 1 de  $\theta \mapsto e^{ik\theta}$  par l'amplitude « modulée » (de façon douce)  $\theta \mapsto a_k(\theta)$ .

On voit en effet (en faisant deux intégrations par parties successives) que, si tel est le cas, il existe bien une constante  $C$  telle que, pour tout couple d'entiers  $(m, n)$  avec  $m \neq n$ ,

$$|\langle T(e_k), T(e_l) \rangle| = \left| \frac{1}{2\pi} \int_0^{2\pi} a_k(\theta) \overline{a_l(\theta)} e^{-i(k-l)\theta} d\theta \right| \leq \frac{C}{|k-l|^2},$$

ce qui assure l'existence d'une suite  $(\tau_k)_{k \in \mathbb{Z}}$  telle que

$$|\langle T(e_k), T(e_l) \rangle| \leq \tau_{k-l}, \quad \forall k, l \in \mathbb{Z}$$

(la série de Riemann  $[1/k^2]_{k \geq 1}$  étant convergente). La possibilité de prolonger  $T$  en un opérateur linéaire continu à l'espace  $L^2_{\mathbb{C}}(\mathbb{R}/2\pi\mathbb{Z})$  tout entier entre donc bien dans le cadre de l'exemple.

### 1.11.2 La notion d'opérateur adjoint

La notion importante que nous attacherons dans ce cours à un opérateur  $\mathbb{K}$ -linéaire continu  $T : H \longrightarrow H$  sera celle d'*opérateur adjoint*.

Soit  $T$  un tel opérateur  $\mathbb{K}$ -linéaire continu  $T : H \longrightarrow H$ . Pour tout  $h$  fixé dans  $H$ , l'application

$$x \in H \longmapsto \langle T(x), h \rangle$$

est une  $\mathbb{K}$ -forme linéaire continue sur  $H$ . Il existe donc (d'après le théorème de dualité 1.2), un unique élément  $T^*(h)$  tel que

$$\langle T(x), h \rangle = \langle x, T^*(h) \rangle, \quad \forall x \in H. \quad (1.32)$$

Si l'on remplace  $h$  par  $\lambda_1 h_1 + \lambda_2 h_2$  avec  $h_1, h_2 \in H$ ,  $\lambda_1, \lambda_2 \in \mathbb{K}$ , on constate que

$$\begin{aligned} \langle T(x), \lambda_1 h_1 + \lambda_2 h_2 \rangle &= \overline{\lambda_1} \langle T(x), h_1 \rangle + \overline{\lambda_2} \langle T(x), h_2 \rangle \\ &= \overline{\lambda_1} \langle x, T^*(h_1) \rangle + \overline{\lambda_2} \langle x, T^*(h_2) \rangle \\ &= \left\langle x, \lambda_1 T^*(h_1) + \lambda_2 T^*(h_2) \right\rangle, \end{aligned}$$

ce qui montre que l'application

$$h \longmapsto T^*(h)$$

définie suivant (1.32) est bien  $\mathbb{K}$ -linéaire. La relation définissant complètement cette application  $\mathbb{K}$ -linéaire  $T^* : H \longrightarrow H$ , soit

$$\forall x \in H, \forall h \in H, \langle T(x), h \rangle = \langle x, T^*(h) \rangle \quad (1.33)$$

est dite *formule d'adjonction*. Cette formule d'adjonction se lit aussi

$$\begin{aligned} \forall x \in H, \forall h \in H, \langle T^*(x), h \rangle &= \overline{\langle h, T^*(x) \rangle} \\ &= \overline{\langle T(h), x \rangle} \\ &= \langle x, T(h) \rangle, \end{aligned}$$

on constate que  $(T^*)^* = T$ .

Si  $h$  est de norme 1, on a, pour tout  $x$  dans  $H$ ,

$$|\langle x, T^*(h) \rangle| = |\langle T(x), h \rangle| \leq \|T(x)\| \times \|h\| \leq \|T\| \times \|x\|,$$

ce qui montre (en prenant  $x = T^*(h)$ ) que  $\|T^*(h)\| \leq \|T\|$ , donc que  $T^*$  est une application  $\mathbb{K}$ -linéaire continue de  $H$  dans  $H$  (comme  $T$ ), avec  $\|T^*\| \leq \|T\|$ . Comme  $(T^*)^* = T$ , on a aussi l'inégalité inverse

$$\|(T^*)^*\| = \|T\| \leq \|T^*\|.$$

Tout ce que nous venons de faire se résume en la proposition-définition suivante :

**Proposition 1.14** Soit  $H$  un  $\mathbb{K}$ -espace de Hilbert ( $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$ ) et  $T : H \longrightarrow H$  une application  $\mathbb{K}$ -linéaire continue de  $H$  dans lui-même (on dit aussi un opérateur  $\mathbb{K}$ -linéaire continu de  $H$  dans  $H$ ). La formule d'adjonction (1.33)

$$\forall x \in H, \forall h \in H, \langle T(x), h \rangle = \langle x, T^*(h) \rangle$$

permet de définir sans équivoque un autre opérateur  $\mathbb{K}$ -linéaire continu de  $H$  dans lui-même, dit adjoint de  $T$ . Les opérateurs  $T$  et  $T^*$  sont tels que  $\|T\| = \|T^*\|$  et l'opération de prise d'adjoint  $T \longmapsto T^*$  est une involution isométrique antilinéaire du  $\mathbb{K}$ -espace vectoriel  $\mathcal{L}_{\mathbb{K}}(H, H)$  des opérateurs  $\mathbb{K}$ -linéaires continus de  $H$  dans  $H$  (équipé de la norme d'opérateur continu définie par (1.31) à partir de la norme  $\|\cdot\|$  dérivant du produit scalaire sur  $H$ ).

On remarque de plus (toujours grâce à la caractérisation de l'adjoint *via* la formule d'adjonction (1.33)) que si  $T_1$  et  $T_2$  sont deux opérateurs  $\mathbb{K}$ -linéaires continus de  $H$  dans lui-même,

$$(T_2 \circ T_1)^* = T_1^* \circ T_2^*.$$

En particulier, si un opérateur linéaire continu  $T : H \longrightarrow H$  admet un inverse aussi linéaire continu  $T^{-1} : H \longrightarrow H$ , on a nécessairement  $(T^{-1})^* = (T^*)^{-1}$ .

**Exemple 1.12 : le cas de la dimension finie.** Si  $H = \mathbb{R}^n$  ou  $H = \mathbb{C}^n$  sont équipés d'un produit scalaire  $\langle \cdot, \cdot \rangle$  (pas nécessairement le produit scalaire canonique), la notion d'adjoint rejoint celle que l'on connaît en algèbre. Si  $M$  est la matrice de l'opérateur  $T$  exprimée dans une base orthonormée de  $\mathbb{R}^n$  ou  $\mathbb{C}^n$  (relativement au produit scalaire  $\langle \cdot, \cdot \rangle$ ), la matrice de l'adjoint  $T^*$  (relativement à la structure hilbertienne attachée au choix du produit scalaire) exprimée dans cette même base est exactement la matrice  $M^*$ , c'est-à-dire la transconjugée de la matrice  $M$  (on transpose et on conjugue les coefficients).

**Exemple 1.13 : les « shifts ».** Si  $H = l_{\mathbb{K}}^2(\mathbb{Z})$ , l'opérateur  $(x_k)_{k \in \mathbb{Z}} \longmapsto (x_{k-1})_{k \in \mathbb{Z}}$  (dit *shift* d'un pas vers la droite) a pour adjoint l'opérateur  $(x_k)_{k \in \mathbb{Z}} \longmapsto (x_{k+1})_{k \in \mathbb{Z}}$  (dit *shift* d'un pas vers la gauche).

**Exemple 1.14 : le cas d'une projection orthogonale sur un sous-espace fermé.** Si  $V$  est un sous-espace fermé de  $H$ , l'opérateur  $\text{Proj}_V$  (qui est continu de norme 1 puisque  $\|\text{Proj}_V x\| \leq \|x\|$  avec égalité si et seulement si  $x \in V$ ) a pour adjoint

$$(\text{Proj}_V)^* = \text{Proj}_V$$

car on a, pour tout  $x, h$  dans  $H$ ,

$$\begin{aligned} \langle x, \text{Proj}_V(h) \rangle &= \langle \text{Proj}_V(x), \text{Proj}_V(h) \rangle \\ &= \langle \text{Proj}_V(x), h \rangle. \end{aligned}$$

L'opérateur  $\text{Proj}_V$  est un exemple d'opérateur *auto-adjoint*. La norme d'opérateur d'une projection orthogonale  $P_V$  sur un sous-espace fermé  $V$  est égale à 1 car  $P_V(x) = x$  pour  $x \in V$  et d'autre part  $\|P_V(x)\| \leq \|x\|$  pour tout  $x$  dans  $H$ .

La proposition suivante nous sera utile car elle relie noyau et image d'un opérateur linéaire continu  $T$  au noyau et image de son adjoint  $T^*$ .

**Proposition 1.15** Soit  $H$  un  $\mathbb{K}$ -espace de Hilbert et  $T$  un opérateur  $\mathbb{K}$ -linéaire continu de  $H$  dans lui-même. On a les relations :

$$\text{Ker } T = (\overline{\text{Im } T^*})^{\perp}$$

$$\begin{aligned}
\text{Ker } T^* &= (\overline{\text{Im } T})^\perp \\
H &= \text{Ker } T \overset{\perp}{\oplus} \overline{\text{Im } T^*} \\
&= \text{Ker } T^* \overset{\perp}{\oplus} \overline{\text{Im } T}
\end{aligned} \tag{1.34}$$

**Preuve.** Si  $h$  est dans le noyau de  $T$ , on a

$$\forall x \in H, \langle T(h), x \rangle = 0.$$

En utilisant la formule d'adjonction (1.33), il vient donc

$$\forall x \in H, \langle h, T^*(x) \rangle = 0,$$

ce qui signifie que  $h$  est orthogonal au sous-espace  $\text{Im } T^*$ . Comme le produit scalaire est continu sur  $H \times H$ ,  $h$  est aussi orthogonal au sous-espace fermé  $\overline{\text{Im } T^*}$ , ce qui prouve bien l'inclusion

$$\text{Ker } T \subset (\overline{\text{Im } T^*})^\perp.$$

Prenons maintenant  $h$  orthogonal à  $\text{Im } T^*$ ; pour montrer que  $T(h) = 0$ , il suffit de montrer que pour tout  $x \in H$ ,

$$\langle T(h), x \rangle = 0.$$

Mais, toujours d'après la formule d'adjonction (1.33)

$$\langle T(h), x \rangle = \langle h, T^*(x) \rangle = 0$$

puisque  $h \perp T^*(x)$ . On a donc bien aussi l'autre inclusion

$$(\overline{\text{Im } T^*})^\perp \subset \text{Ker } T$$

(on peut passer à gauche à l'adhérence car l'on sait que  $\text{Ker } T$  est fermé). Les formules des première et troisième lignes de (1.34) sont donc prouvées (pour celle de la troisième ligne, on applique le théorème de projection orthogonale). En remplaçant  $T$  par  $T^*$  et en utilisant l'involutivité de la prise d'adjoint  $(T^*)^* = T$ , on en déduit les formules des seconde et quatrième lignes.  $\diamond$

### 1.11.3 L'intérêt de la prise d'adjoint; deux exemples pratiques

Nous donnons dans ce paragraphe deux exemples de raisonnements où l'opération de prise d'adjoint facilite la tâche. Ces raisonnements conduisent chacun à une application pratique importante : la première concerne la « poursuite » d'une information inconnue par le biais de projections orthogonales itérées judicieuses, la seconde concerne la possibilité, dans un espace de Hilbert séparable, de « perturber » une base hilbertienne tout en conservant certaines des propriétés opérationnelles qui lui sont attachées. Cette sous-section peut être pensée comme deux thèmes de problèmes à regarder à tête reposée. On peut la lire à titre d'exercice (ces exercices seront d'ailleurs partiellement repris dans les guides d'activités proposés sous Ulysse).

### Itération de projections

Le résultat de ce paragraphe est celui que soutend la figure illustrant l'exemple 1.2 de la section 1.1.

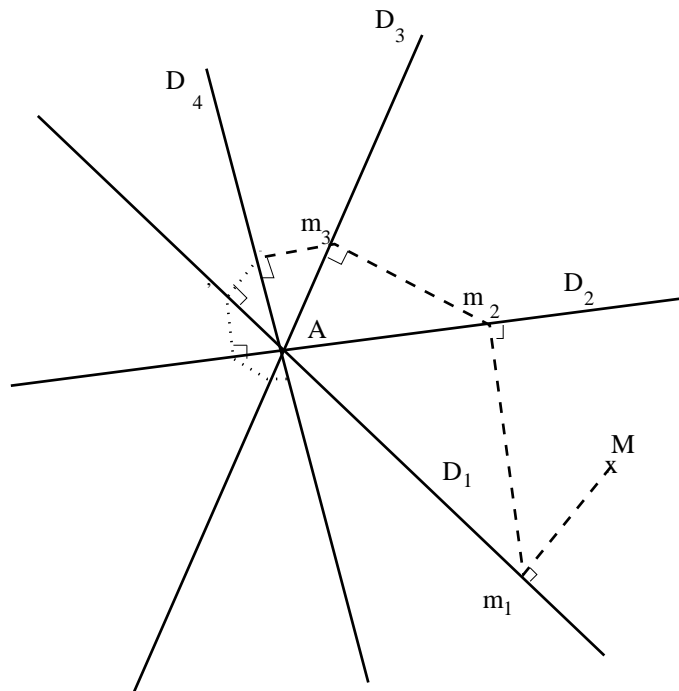


FIG. 1.7 – Itération de projections (rappel de l'exemple 1.2)

**Proposition 1.16** Soient  $V_1, V_2, \dots, V_N$ ,  $N$  sous espaces vectoriels fermés d'un espace de Hilbert  $H$ ,  $h$  un élément de  $H$ , et  $P_1, \dots, P_N$ , les projections orthogonales sur les sous-espaces affines respectifs  $D_1 := h + V_1, \dots, D_N := h + V_N$ . Soit  $V := V_1 \cap \dots \cap V_N$  et  $P$  la projection orthogonale sur le sous espace affine  $A = h + V$ . Alors, si  $Q$  est l'application affine  $Q = P_N \circ \dots \circ P_1$ , on a

$$Q^k(x) \longrightarrow P(x) \quad (1.35)$$

pour tout  $x$  dans  $H$  lorsque  $k$  tend vers l'infini. De plus, lorsque  $N = 2$  et

$$\cos \theta := \sup_{\substack{h_1 \in V_1^* \cap V^\perp \\ h_2 \in V_2^* \cap V^\perp}} \frac{|\langle h_1, h_2 \rangle|}{\|h_1\| \|h_2\|} < 1,$$

alors, on a l'information quantitative permettant l'estimation d'erreur dans cet algorithme itératif :

$$\|Q^k(x) - P(x)\| \leq (\cos \theta)^k \|x - P(x)\|. \quad (1.36)$$

**Preuve.** On se ramène au cas  $h = 0$ , ce qui nous permet de poser le problème au niveau des sous-espaces vectoriels et non plus affines. Les projections orthogonales sur des sous-espaces vectoriels fermés étant des opérateurs de norme égale à 1, on voit que si  $x$  est dans le noyau de  $Q - \text{Id}$ , on a

$$\|x\| \geq \|P_1(x)\| \geq \dots \geq \|Q(x)\| = \|x\|,$$

et donc  $\|P_j(x)\| = \|x\|$  pour  $j = 1, \dots, N$ . Ceci implique (*via* le théorème de Pythagore), que  $P_j(x) = x$  pour tout  $j = 1, \dots, N$ , soit  $x \in V$ . On a donc

$$\text{Ker}(I - Q) = V.$$

Introduisons ici la prise d'adjoint. Comme  $\text{Ker}(I - Q^*) = \text{Ker}(I - Q) = V$  (il suffit d'inverser l'ordre des projections orthogonales  $P_k$ , qui sont toutes, on le rappelle, des opérateurs autoadjoints, *cf.* exemple 1.14 ci-dessus), on peut d'après la proposition 1.15 décomposer l'espace de Hilbert  $H$  sous la forme

$$H = \text{Ker}(I - Q^*) \overset{\perp}{\oplus} \overline{\text{Im}(I - Q)} = V \overset{\perp}{\oplus} \overline{\text{Im}(I - Q)}.$$

Prouver le premier volet de la proposition revient donc à prouver (1.35) pour tout  $x$  dans  $\overline{\text{Im}(I - Q)}$ , ou encore, ce qui est suffisant par continuité de  $I - Q$ , pour tout  $x$  dans  $\text{Im}(I - Q)$ . Mais si  $x = (I - Q)(y)$ , posons

$$a := \lim_{k \rightarrow \infty} \|Q^k(y)\|.$$

( $a$  existe bien car on a affaire à une suite décroissante de nombres positifs). Distinguons maintenant deux cas.

– Si  $a > 0$ , on considère la suite  $(u_k)_{k \in \mathbb{N}}$ , où

$$u_k = \frac{Q^k(y)}{\|Q^k(y)\|};$$

on a  $\|Q(u_k)\| \rightarrow 1$ ; ceci implique

$$(I - Q)(u_k) \rightarrow 0,$$

comme on le voit en utilisant une fois de plus le théorème de Pythagore et en raisonnant par récurrence sur le nombre  $N$  de projecteurs impliqués dans la factorisation de  $Q$ ; comme  $a > 0$ , il en résulte

$$Q^k \circ (I - Q)(y) = (I - Q) \circ Q^k(y) \rightarrow 0.$$

– Si maintenant  $a = 0$ , on a également

$$(I - Q) \circ Q^k(y) \mapsto 0.$$

Ainsi, pour tout  $x$  dans  $\text{Im}(I - Q)$ , on a  $Q^k(x) \rightarrow 0 = P(x)$ , ce qui, avec le fait que  $Q^k(x)$  converge vers  $x = P(x)$  lorsque  $x \in V$ , conclut la preuve du premier volet de la proposition.

En ce qui concerne le second volet de cette même proposition, il suffit de remarquer que, compte tenu de la définition de  $\cos \theta$ ,

$$\|Q(x)\| \leq \cos \theta \|x\|$$

pour tout  $x$  dans  $V^\perp$  et d'appliquer cette inégalité de manière itérative au vecteur  $x - P(x)$ .  $\diamond$

**Application.** Une application importante de ce résultat est la suivante : c'est celle qu'illustre la figure déjà introduite dans l'exemple 1.1 de la section 1.1 et que pour mémoire nous redonnons ci-dessous.



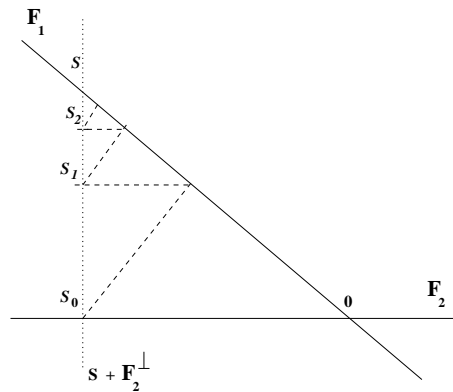


FIG. 1.8 – Itération de projections (rappel de l'exemple 1.1)

On se donne deux sous-espaces  $F_1$  et  $F_2$  de  $H$  tels que  $F_1 \cap F_2^\perp = \{0\}$ . On note  $P_1$  (resp.  $P_2$ ) la projection orthogonale sur  $F_1$  (resp.  $F_2$ ) et  $Q_1$  (resp.  $Q_2$ ) la projection orthogonale sur  $F_1^\perp$  (resp.  $F_2^\perp$ ). On suppose que  $s$  est un élément de  $F_1$  a priori inconnu, la seule information connue étant  $s_0 = P_2(s)$ . Alors, l'algorithme itératif

$$\begin{aligned} s_0 &= P_2(s) \\ \vdots & \\ s_{k+1} &= s_0 + Q_2 \circ P_1(s_k) \\ &= (s - Q_2(s)) + Q_2 \circ P_1(s_k) \\ &= s + Q_2(P_1(s_k) - s) \\ &= s + Q_2 \circ P_1(s_k - s), \quad \forall k \in \mathbb{N}, \end{aligned}$$

(c'est plus parlant sur la figure) converge vers  $s$ . Il suffit d'utiliser la proposition 1.16. De plus, si  $F_1$  et  $F_2^\perp$  sont tels que

$$\sup_{\substack{h_1 \in F_1 \setminus \{0\} \\ h_2 \in F_2^\perp \setminus \{0\}}} \frac{|\langle h_1, h_2 \rangle|}{\|h_1\| \|h_2\|} = \cos \theta < 1, \quad (1.37)$$

on a une estimation de la vitesse de convergence de  $s_k$  vers  $s$  en

$$\|s_k - s\| \leq (\cos \theta)^k \|s_0 - s\| = (\cos \theta)^k \|Q_2(s)\| \leq (\cos \theta)^k \|s\|;$$

la décroissance de l'erreur d'approximation est exponentielle. Souvent, pareille idée est exploitée, même dans le cas où l'« angle » des deux sous-espaces  $F_1$  et  $F_2^\perp$  est égal à 0 ( $\cos \theta = 1$ ); l'algorithme n'a plus alors, on s'en doute, la robustesse que l'on pourrait escompter.

### Systèmes bi-orthonormés dans un espace de Hilbert séparable

Souvent les bases orthonormées des espaces de Hilbert sont amenées à se présenter comme « perturbées », l'orthogonalité n'étant alors plus assurée. Il est néanmoins possible de pallier à ce problème grâce à la proposition suivante<sup>40</sup>.

**Proposition 1.17** Soit  $(e_k)_{k \geq 1}$  une base hilbertienne d'un  $\mathbb{K}$ -espace de Hilbert séparable et  $(f_k)_{k \in \mathbb{N}^*}$  un système de vecteurs tel qu'il existe une constante absolue  $\tau \in ]0, 1[$

<sup>40</sup>Que l'on peut voir comme un résultat de stabilité concernant les bases hilbertiennes, même si l'orthogonalité reste une notion « fragile » comparée à la redondance qui elle est une notion beaucoup plus robuste! Nous y reviendrons dans la section 1.12 suivante.

avec

$$\left\| \sum_{k=1}^N \lambda_k (e_k - f_k) \right\|^2 \leq \tau^2 \sum_{k=1}^N |\lambda_k|^2 \quad (1.38)$$

pour tout  $N \in \mathbb{N}^*$ , pour tout choix de scalaires  $\lambda_1, \dots, \lambda_N$  dans le corps<sup>41</sup>. Il existe alors un système de vecteurs  $(g_k)_{k \in \mathbb{N}^*}$  tel que

$$\langle f_k, g_l \rangle = \begin{cases} 1 & \text{si } k = l \\ 0 & \text{si } k \neq l \end{cases}$$

et que tout élément  $h$  de  $H$  puisse s'écrire

$$h = \sum_{k=1}^{\infty} \langle h, g_k \rangle f_k = \sum_{k=1}^{\infty} \langle h, f_k \rangle g_k$$

avec

$$\begin{aligned} \frac{\|h\|}{1 + \tau} &\leq \left( \sum_{k=1}^{\infty} |\langle h, g_k \rangle|^2 \right)^{1/2} \leq \frac{\|h\|}{1 - \tau} \\ (1 - \tau)\|h\| &\leq \left( \sum_{k=1}^{\infty} |\langle h, f_k \rangle|^2 \right)^{1/2} \leq (1 + \tau)\|h\|. \end{aligned}$$

**Preuve.** On définit tout d'abord un opérateur linéaire continu  $T : H \longrightarrow H$  en définissant dans un premier temps son action sur les éléments de la base hilbertienne  $(e_k)_{k \geq 1}$  par

$$T(e_k) = e_k - f_k, \quad \forall n \in \mathbb{Z}.$$

Ainsi, pour tout  $N \in \mathbb{N}^*$ , pour tout choix de scalaires  $\lambda_1, \dots, \lambda_N$ ,

$$T\left(\sum_{k=1}^N \lambda_k e_k\right) = \sum_{k=1}^N \lambda_k (e_k - f_k).$$

Pour tout élément  $h$  dans le sous-espace vectoriel engendré par les  $e_k$ ,  $k \geq 1$ , on a

$$\|T(h)\| \leq \tau \|h\| \quad (1.39)$$

d'après l'hypothèse (1.38), ce qui prouve (si l'on invoque le lemme 1.2 de prolongement des opérateurs) que  $T$  se prolonge bien en un opérateur linéaire continu de  $H$  dans lui-même puisque le sous-espace vectoriel engendré par les  $e_k$ ,  $k \geq 1$  est dense. L'opérateur  $T : H \longrightarrow H$  ainsi défini est de norme majorée par  $\tau$  du fait de l'inégalité (1.39) (valable par prolongement pour tout  $h$  dans  $H$ ). L'opérateur  $T$  étant de norme  $\|T\| \leq \tau < 1$ , l'opérateur  $S = \text{Id}_H - T = I - T$  admet un inverse continu  $S^{-1}$  dont l'action est définie par

$$S^{-1}(h) := \sum_{k=0}^{\infty} T^k(h), \quad \forall h \in H, \quad (1.40)$$

<sup>41</sup>Ceci signifie que le système  $(f_k)_{k \geq 1}$  est en quelque sorte une version « perturbée » du système orthonormé idéal qu'est la base hilbertienne  $(e_k)_{k \geq 1}$ .

où  $T^k$  désigne, pour  $k \in \mathbb{N}$ , l'itéré  $k$  fois de l'opérateur  $T$  ( $T^0 = \text{Id}$ ) ; la série au second membre de (1.40) est en effet absolument convergente car

$$\sum_{k=0}^{\infty} \|T^k(h)\| \leq \sum_{k=0}^{\infty} \|T\|^k \|h\| \leq \|h\| \sum_{k=0}^{\infty} \tau^k \leq \frac{\|h\|}{1-\tau},$$

ce qui prouve non seulement que  $S^{-1} : H \rightarrow H$  est bien défini (comme opérateur linéaire continu de  $H$  dans lui-même), mais encore que

$$\|S^{-1}\| \leq \frac{1}{1-\tau}.$$

Notons que l'on a, pour tout  $h$  dans  $H$ , les inégalités :

$$\begin{aligned} (1-\tau)\|h\| &= (1-\tau)\|S^{-1}(S(h))\| \leq \|S(h)\| \leq \|h\| + \|T(h)\| \leq (1+\tau)\|h\| \\ (1-\tau)\|S^{-1}(h)\| &\leq \|h\| = \|S(S^{-1}(h))\| \leq (1+\tau)\|S^{-1}(h)\|. \end{aligned} \quad (1.41)$$

Ici intervient l'opération de prise d'adjoint : nous introduisons en effet l'opérateur  $S^*$  adjoint de  $S$  ; la norme de cet opérateur  $S^*$  est égale à la norme de  $S$ , donc majorée par  $1+\tau$ . Du fait de la remarque suivant la proposition 1.14, l'opérateur  $S^*$  est un opérateur inversible, d'inverse l'opérateur continu  $(S^{-1})^*$  (de norme égale à  $\|S^{-1}\|$ ) ; on a donc aussi, pour tout  $h$  dans  $H$ , l'encadrement

$$(1-\tau)\|h\| = (1-\tau)\|(S^*)^{-1}(S^*(h))\| \leq \|S^*(h)\| \leq (1+\tau)\|h\|. \quad (1.42)$$

Posons, pour  $k \in \mathbb{N}^*$ ,

$$g_k := (S^{-1})^*(e_k) = (S^*)^{-1}(e_k).$$

D'après la formule d'adjonction (1.33), on a bien, pour tout couple  $(k, l)$  d'entiers positifs,

$$\begin{aligned} \langle f_k, g_l \rangle &= \langle S(e_k), (S^{-1})^*(e_l) \rangle = \langle S^{-1} \circ S(e_k), e_l \rangle \\ &= \langle e_k, e_l \rangle = \begin{cases} 1 & \text{si } k = l \\ 0 & \text{si } k \neq l. \end{cases} \end{aligned}$$

On a aussi, pour tout  $h$  dans  $H$ ,

$$\begin{aligned} h = S(S^{-1}(h)) &= S\left(\sum_{k=1}^{\infty} \langle S^{-1}(h), e_k \rangle e_k\right) \\ &= \sum_{k=1}^{\infty} \langle S^{-1}(h), e_k \rangle S(e_k) \\ &= \sum_{k=1}^{\infty} \langle h, (S^{-1})^*(e_k) \rangle S(e_k) = \sum_{k=1}^{\infty} \langle h, g_k \rangle f_k \end{aligned}$$

et aussi

$$h = (S^*)^{-1} \circ S^*(h) = (S^*)^{-1}\left(\sum_{k=1}^{\infty} \langle S^*(h), e_k \rangle e_k\right)$$

$$\begin{aligned}
&= \sum_{k=1}^{\infty} \langle S^*(h), e_k \rangle (S^*)^{-1}(e_k) \\
&= \sum_{k=1}^{\infty} \langle h, (S^*)^*(e_k) \rangle g_k \\
&= \sum_{k=1}^{\infty} \langle h, S(e_k) \rangle g_k = \sum_{k=1}^{\infty} \langle h, f_k \rangle g_k.
\end{aligned}$$

Comme de plus (toujours en utilisant la formule d'adjonction (1.33))

$$\begin{aligned}
\sum_{k=1}^{\infty} |\langle h, g_k \rangle|^2 &= \sum_{k=1}^{\infty} |\langle h, (S^{-1})^*(e_k) \rangle|^2 = \sum_{k=1}^{\infty} |\langle S^{-1}(h), e_k \rangle|^2 = \|S^{-1}(h)\|^2 \\
\sum_{k=1}^{\infty} |\langle h, f_k \rangle|^2 &= \sum_{k=1}^{\infty} |\langle h, S(e_k) \rangle|^2 = \sum_{k=1}^{\infty} |\langle S^*(h), e_k \rangle|^2 = \|S^*(h)\|^2,
\end{aligned}$$

on conclut à toutes les assertions de la proposition en invoquant les doubles encadrements (1.41) et (1.42).  $\diamond$

**Exemple 1.15.** Dans l'espace de Hilbert  $L_{\mathbb{C}}^2(\mathbb{T})$  où les classes de fonctions

$$e_k : \theta \mapsto e^{ik\theta}, \quad k \in \mathbb{Z},$$

(correspondant aux harmoniques fondamentales) forment une base hilbertienne, on peut considérer les classes « perturbées » des fonctions obtenues en  $2\pi$ -périodisant les

$$f_k : \theta \in [-\pi, \pi[ \mapsto e^{i\omega_k \theta}, \quad k \in \mathbb{Z},$$

où les  $\omega_k$  sont des nombres réels avec

$$\mu := \max_k |\omega_k - k| < \frac{\log 2}{\pi}.$$

Comme on le vérifie aisément, pour tout  $N \in \mathbb{N}^*$  et pour tout choix de scalaires complexes  $(\lambda_k)_{k=-N}^N$ ,

$$\begin{aligned}
\left\| \sum_{k=-N}^N \lambda_k (f_k - e_k) \right\| &= \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \sum_{k=-N}^N \lambda_k (e^{i(\omega_k - k)\theta} - 1) e^{ik\theta} \right|^2 d\theta \right)^{1/2} \\
&= \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \sum_{k=-N}^N \lambda_k \left( \sum_{l=1}^{\infty} \frac{(i(\omega_k - k))^l \theta^l}{l!} \right) e^{ik\theta} \right|^2 d\theta \right)^{1/2} \\
&\leq \sum_{l=1}^{\infty} \frac{\pi^l}{l!} \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \sum_{k=-N}^N \lambda_k (i(\omega_k - k))^l e^{ik\theta} \right|^2 d\theta \right)^{1/2} \\
&\leq \sum_{l=1}^{\infty} \frac{\pi^l \mu^l}{l!} \left( \sum_{k=-N}^N |\lambda_k|^2 \right)^{1/2} \\
&\leq (e^{\mu\pi} - 1) \left( \sum_{k=-N}^N |\lambda_k|^2 \right)^{1/2} \\
&\leq \tau \left( \sum_{k=-N}^N |\lambda_k|^2 \right)^{1/2}
\end{aligned}$$

avec  $\tau < 1$  si  $\mu < \log 2/\pi$ . La proposition s'applique dans ce contexte et l'on peut associer un système  $(g_k)_{k \in \mathbb{Z}}$  au système « perturbé »  $(f_k)_{k \in \mathbb{Z}}$ . Notons que ceci nous autorise à travailler dans l'espace  $L_{\mathbb{C}}^2(\mathbb{T})$  avec un système de fonctions de référence qui correspondent à des classes de fonctions oscillantes, mais cette fois aperiodiques, puisque les fréquences  $\omega_k$ ,  $k \in \mathbb{Z}$ , ne sont plus entières. Ceci est souvent très intéressant dans les problèmes pratiques liés à l'analyse de Fourier (sur laquelle on se penchera au chapitre 2 de ce cours).

## 1.12 Un procédé algorithmique : le *matching*

Comme on l'a déjà mentionné dans ce cours, les notions de système orthonormé et de base hilbertienne, si elles constituent des modèles mathématiques « idéaux » pour décomposer un élément dans un espace de Hilbert, ne sont pas des objets « robustes » lorsqu'on les perturbe. Il est souvent préférable du point de vue pratique de travailler avec des « *frames* » ou même des « dictionnaires » souvent très redondants : au contraire de ce qui se passe pour les systèmes orthonormés, les éléments distincts du dictionnaire sont autorisés cette fois à être corrélés entre eux, c'est-à-dire à ne plus être nécessairement orthogonaux relativement au produit scalaire avec lequel on travaille.

On utilise souvent, tant en mathématiques appliquées qu'en informatique, des algorithmes « gloutons » pour « pister <sup>42</sup> » un élément d'un espace de Hilbert contre un dictionnaire (en général redondant) d'éléments « test » pris dans ce même espace de Hilbert (en utilisant toute la force de l'algorithme de projection orthogonale).

Au titre d'annexe à ce cours (et de thème de problème de révision), voici un exemple d'un tel procédé algorithmique ; on peut ici le voir comme un thème d'exercice guidé, mais c'est une méthode très couramment utilisée dans la pratique, notamment en informatique (par exemple en imagerie), dans toutes les questions relatives à l'analyse, au stockage, au classement, à la compression (souvent aux fins de synthèse ultérieure), de données expérimentales. La « traque » d'un visage ou d'un objet dans un film vidéo, la confrontation d'une image médicale pathologique avec un dictionnaire d'images « type », constituent des exemples illustrant parfaitement ce type de démarche algorithmique d'inspiration naturellement hilbertienne.

**Proposition 1.18** *Soit  $H$  un  $\mathbb{K}$ -espace de Hilbert,  $h$  un élément de  $H$  et  $D$  un dictionnaire d'éléments (tous de norme 1) que l'on suppose dénombrable et engendrant un sous-espace dense ; on suppose que l'on peut construire de manière itérative une suite d'éléments  $d_1, \dots, d_n, \dots$  du dictionnaire  $D$  (les  $d_j$  dépendent de  $h$ ), une suite de scalaires  $\lambda_1, \dots, \lambda_n, \dots$  (dépendant aussi de  $h$ ) tels que*

$$\begin{aligned} |\langle h, d_1 \rangle| &= \max_{d \in D} |\langle h, d \rangle| \\ \lambda_1 d_1 &= \text{Proj}_{\mathbb{K}d_1}(h), \end{aligned}$$

et, pour tout  $n \geq 1$ ,

$$\begin{aligned} \left| \left\langle h - \sum_{k=1}^n \lambda_k d_k, d_{n+1} \right\rangle \right| &= \max_{d \in D} \left| \left\langle h - \sum_{k=1}^n \lambda_k d_k, d \right\rangle \right| \\ \lambda_{n+1} d_{n+1} &= \text{Proj}_{\mathbb{K}d_{n+1}} \left( h - \sum_{k=1}^n \lambda_k d_k \right). \end{aligned}$$

Alors, la suite

$$(h_n)_{n \geq 1} := \left( \sum_{k=1}^n \lambda_k d_k \right)_{n \geq 1}$$

---

<sup>42</sup>C'est la l'origine de la terminologie anglo-saxonne de « *matching* ».

converge vers  $h$  dans  $H$  ; on dit qu'elle réalise le *matching*<sup>43</sup> du vecteur  $h$  contre le dictionnaire  $D$ .

**Preuve.** Posons  $v_k := \lambda_k d_k$  pour  $k \in \mathbb{N}^*$  et notons  $r_n$  le « reste » de la décomposition de  $h$  une fois l'étape  $n$  de l'algorithme effectuée, soit

$$r_n := h - \sum_{k=1}^n v_k$$

partant de  $r_0 = h$ . Si  $1 \leq n < m$ , on a

$$r_n = r_m + \sum_{k=n+1}^m v_k.$$

Comme on effectue à chaque cran de l'algorithme une projection orthogonale unidimensionnelle sur la droite vectorielle engendrée par le nouvel élément du dictionnaire trouvé (le procédé est très simple et très naïf!), on a, pour tout  $k \in \mathbb{N}^*$ ,

$$\|r_k\|^2 = \|r_{k-1}\|^2 - |\langle r_{k-1}, d_k \rangle|^2.$$

La suite  $(\|r_k\|^2)_{k \geq 0}$  est donc une suite décroissante minorée, donc convergente. On en déduit aussi

$$\sum_{k=1}^{\infty} \|v_k\|^2 < \infty. \quad (1.43)$$

Pour montrer dans un premier temps que la suite  $(r_k)_{k \geq 0}$  converge, nous allons prouver qu'elle est de Cauchy. Fixons donc  $\epsilon > 0$ . Comme on le voit aisément, on a, si  $1 \leq n < m$ ,

$$\|r_n - r_m\|^2 = \|r_n\|^2 + \|r_m\|^2 - 2\|r_m\|^2 - 2 \sum_{k=n+1}^m \operatorname{Re} \langle r_m, v_k \rangle.$$

Mais on a, pour tout  $k$  entre  $n+1$  et  $m$ ,

$$|\langle r_m, v_k \rangle| \leq \|v_k\| \times \|v_{m+1}\| ; \quad (1.44)$$

en effet, toujours pour  $k$  entre  $n+1$  et  $m$ ,

$$\begin{aligned} |\langle r_m, v_k \rangle| &= \left| \left\langle r_m, \langle r_{k-1}, d_k \rangle d_k \right\rangle \right| \\ &= |\langle r_m, d_k \rangle| \times |\langle r_{k-1}, d_k \rangle| \\ &= \|v_k\| |\langle r_m, d_k \rangle| \\ &\leq \|v_k\| \times \|v_{m+1}\| \end{aligned}$$

du fait même du principe de l'algorithme, puisque  $|\langle r_m, d_{m+1} \rangle|$  (qui est par définition la norme de  $v_{m+1}$ ) maximise tous les  $|\langle r_m, d \rangle|$  avec  $d \in D$ . On a donc

$$\|r_n - r_m\|^2 \leq \|r_n\|^2 - \|r_m\|^2 + 2\|v_{m+1}\| \sum_{k=n+1}^m \|v_k\| \quad (1.45)$$

---

<sup>43</sup>Comme le dictionnaire  $D$  peut fort bien être redondant (on a d'ailleurs tout intérêt à ce qu'il en soit ainsi), la décomposition de  $h$  suivant  $D$  n'a rien d'unique ; la méthode proposée dans ce scénario algorithmique fournit simplement une décomposition efficace en « pistant » (de manière « hiérarchique ») les éléments tests qui comptent le plus dans la représentation de  $h$ .

Or

$$\|v_{m+1}\| \sum_{k=n+1}^m \|v_k\| \leq \|v_{m+1}\| \sum_{k=1}^{m+1} \|v_k\|.$$

Mais la clause (1.43) implique

$$\lim_{p \rightarrow \infty} \left( \inf_{q \geq p} \left( \|v_q\| \sum_{k=1}^q \|v_k\| \right) \right) = 0. \quad (1.46)$$

Si  $n$  est assez grand, on est donc certain que

$$\|r_n\|^2 - \|r_m\|^2 \leq \epsilon.$$

On choisira  $n$  assez grand pour qu'il en soit ainsi. Il résulte de (1.46) qu'il existe  $q > m$  tel que

$$\|v_q\| \sum_{k=1}^q \|v_k\| \leq \epsilon.$$

Mais on a

$$\|r_n - r_m\| \leq \|r_n - r_q\| + \|r_m - r_q\|.$$

En appliquant (1.45) (mais cette fois avec les couples  $(n, q)$  et  $(m, q)$  au lieu de  $(m, n)$ ), on voit que

$$\max(\|r_n - r_q\|^2, \|r_m - r_q\|^2) \leq 3\epsilon.$$

En mettant tout ensemble, on voit donc que

$$\|r_n - r_m\| \leq 2\sqrt{3\epsilon}$$

pour ce choix de  $n$ , ce qui implique bien que la suite  $(r_k)_{k \geq 0}$  est de Cauchy, donc convergente. Mais on a aussi la convergence vers 0 de  $\|v_k\|$  lorsque  $k$  tend vers l'infini.

On a donc

$$\lim_{k \rightarrow \infty} |\langle r_k, d_{k+1} \rangle| = 0.$$

Le principe de l'algorithme (appliqué une fois de plus) nous assure que pour tout élément  $d$  du dictionnaire

$$\lim_{k \rightarrow \infty} |\langle r_k, d \rangle| = 0.$$

Comme le dictionnaire forme une partie totale, la limite de la suite  $(r_k)_{k \geq 0}$  doit être orthogonale à tous les atomes du dictionnaire, est par conséquent nulle, et la proposition est ainsi démontrée.  $\diamond$

**Remarque 1.4.** On peut raffiner de beaucoup de manières cet algorithme. Une manière consiste par exemple (en utilisant le procédé de Gram-Schmidt) à faire en sorte qu'à chaque étape le reste  $r_n$  soit orthogonal aux éléments tests atomes  $d_1, \dots, d_n$  du dictionnaire précédemment sélectionnés. Ceci évite de ré-utiliser le dictionnaire complet à chaque cran du processus. Le risque est de réintroduire la fragilité liée au souci d'orthogonalité.





# Chapitre 2

## L'analyse de Fourier et ses applications

### 2.1 Introduction

Si l'analyse de Fourier puise ses sources dans le travail de pionnier de Jean-Baptiste Joseph Fourier (1768-1830)<sup>1</sup> et en particulier dans son mémoire fondateur « *Sur la propagation de la chaleur dans les corps solides* » de 1807 où il introduit la théorie des séries trigonométriques, elle a fédéré tout au long du XIX-ème et du XX-ème siècles tant des mathématiciens (on retrouvera dans ce cours les noms de Friedrich Bessel, Karl Weierstrass, Gustav Lejeune Dirichlet, Lipót Féjer, Alfred Tauber, Johannes Radon, Michaël Plancherel,..., théoriciens des nombres ou analystes), que des physiciens (on y croiera Werner Heisenberg avec le célèbre principe d'incertitude, Denis Gabor, Josiah Gibbs, des opticiens comme Joseph von Fraunhofer) et des ingénieurs mathématiciens ou informaticiens (Claude Shannon, James W. Cooley et John W. Tukey,...). À travers ce second chapitre, où l'on présentera l'analyse de Fourier tant en mathématiques discrètes qu'en mathématiques « continues », on tentera de mettre en lumière non seulement les idées et résultats mathématiques, mais aussi la manière dont ces idées et ces résultats s'insèrent aujourd'hui de manière essentielle (et surtout depuis la « révolution numérique » survenue à l'aube des années 1970 avec le célèbre algorithme proposé par les ingénieurs américains James Cooley et John Tukey autour de 1965) dans de nombreuses branches relevant des mathématiques, de la physique, de l'ingénierie ou de l'informatique : traitement des signaux et des images, théorie de l'information, cryptologie et cryptanalyse, optique, physique quantique, tomographie en instrumentation médicale (CAT-Scanner, IRM), analyse et synthèse du son et de la parole, téléphonie mobile et télécommunications,...

Si les mathématiques (avec leur présentation pouvant parfois paraître « académique ») constituent le fil directeur des notes de ce chapitre, il est évident qu'il convient de prendre ici un certain recul qui permette aussi de voir ces mathématiques « en situation » dans le monde qui nous entoure ; cela était déjà vrai avec l'analyse hilbertienne et l'algorithme « à la Pythagore » présentée au premier chapitre,

---

<sup>1</sup>Mathématicien, mais aussi conseiller scientifique et compagnon de Napoléon qu'il accompagna avec Champollion durant la campagne d'Égypte, homme politique aussi (il fut préfet de l'Isère et l'Université Grenoble 1 porte aujourd'hui son nom), Joseph Fourier a marqué profondément l'essor conjoint des mathématiques et de la physique à l'aube du XIX-ème siècle. Il faut veiller à ne pas le confondre avec son homonyme le philosophe François-Marie Charles Fourier (1772-1837), le père du « Fourierisme » et fondateur de l'École Sociétaire.

c'est bien sûr encore plus criant pour cette seconde partie du cours, évidemment très en prise avec la première.

Nous suivrons l'analyse de Fourier depuis le cadre de la dimension finie (avec les modèles  $\mathbb{R}^N$  ou  $\mathbb{C}^N$ ), relevant de l'algèbre et de l'arithmétique, jusqu'aux cadres discret dénombrable ( $l_{\mathbb{K}}^p(\mathbb{Z}^n)$ ) et continu ( $L_{\mathbb{K}}^p(\mathbb{R}^n)$  ou  $L_{\mathbb{K}}^p(\mathbb{T}^n)$ ), relevant, eux, plus de l'analyse car la notion de dimension infinie y est présente ( $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$ ). On insistera en particulier sur le cadre  $p = 2$  présentant l'intérêt majeur d'être propice à l'algorithmique hilbertienne « à la Pythagore » étudiée au chapitre 1.

## 2.2 Transformation de Fourier dans le cadre discret

Cette section sera dévolue au cadre des mathématiques discrètes : notre premier modèle sera celui des espaces (de Hilbert)  $l_{\mathbb{K}}^2(\{0, \dots, N-1\})$  (espace des signaux digitaux de longueur  $N$  réels ou complexes) ou  $l_{\mathbb{K}}^2(\{0, \dots, N_1-1\} \times \{0, \dots, N_2-1\})$ , (espace des images digitales de taille  $N_1 \times N_2$ ),  $\mathbb{K}$  désignant le corps des réels ou celui des complexes. Ensuite, notre modèle sera celui des espaces  $l_{\mathbb{K}}^p(\mathbb{Z})$ ,  $p \geq 1$ , parmi lesquels les cas  $p = 1$  et bien sûr  $p = 2$  (c'est le modèle, on l'a vu, de  $\mathbb{K}$ -espace de Hilbert séparable <sup>2</sup>) retiendront surtout notre attention.

Notre souci est d'introduire la transformation de Fourier avec les deux motivations qui l'accompagnent :

- fournir un outil puissant de calcul pour transformer une opération (la « convolution ») omniprésente mais difficile à gérer en une opération autrement plus maniable (la multiplication) ;
- permettre la décomposition (dite *harmonique*) des phénomènes discrets en phénomènes périodiques (toujours discrets).

### 2.2.1 La transformation de Fourier discrète

Soit  $N \in \mathbb{N}^*$  un entier strictement positif. Un système orthonormé du  $\mathbb{C}$ -espace de Hilbert  $l_{\mathbb{C}}^2(\{0, \dots, N-1\})$  des vecteurs  $(s(k))_{k=0, \dots, N-1}$  de nombres complexes, équipé du produit scalaire

$$\langle s_1, s_2 \rangle = \sum_{k=0}^{N-1} s_1(k) \overline{s_2(k)},$$

est donné par les suites  $(e_j(k))_{k=0, \dots, N-1}$ ,  $j = 0, \dots, N-1$ , où  $e_j(k) = 1$  si  $j = k$ , 0 sinon ; c'est la classique base canonique de  $\mathbb{C}^n$ . Mais il existe d'autres systèmes orthogonaux particulièrement intéressants ; l'un d'eux est constitué des suites

$$\left( \exp(2i\pi jk/N) \right)_{k=0, \dots, N-1} = \left( \cos(2jk\pi/N) + i \sin(2jk\pi/N) \right)_{k=0, \dots, N-1},$$

$j = 0, \dots, N-1$ . Le signal digital  $f_j := (\exp(2i\pi jk/N))_{k=0, \dots, N-1}$  correspond à la suite des valeurs prises successivement aux points  $t = 0, t = 1, \dots, t = N-1$ , par la

<sup>2</sup>On choisit ici l'indexation par  $\mathbb{Z}$  plutôt que par  $\mathbb{N}$  ou  $\mathbb{N}^*$  parce que  $\mathbb{Z}$  est équipé avec l'addition d'une structure de groupe commutatif, ce qui jouera un rôle précieux à la fois dans la définition et le maniement de la transformation de Fourier.

fonction

$$t \in \mathbb{R} \longmapsto \exp(2i\pi jt/N)$$

qui est une fonction du type  $t \longmapsto \exp(i\omega_j t)$  avec  $\omega_j = 2\pi j/N$ , donc une fonction périodique de fréquence  $2\pi j/N$ . Le système orthogonal constitué des  $N$  signaux digitaux  $f_0, \dots, f_{N-1}$  constitue un système totalement différent du système orthonormé constitué par les  $e_j = (e_j(k))_{k=0, \dots, N-1}$ ; il est constitué de signaux discrets « oscillants » et non de fonctions « pic »; en divisant ces vecteurs par  $\sqrt{N}$ , on obtient donc un nouveau système orthonormé  $(f_j)_{j=0, \dots, N-1}$  de  $\mathbb{C}^N$  constitué de vecteurs correspondant à des signaux digitaux par essence très différents des signaux « pic »  $e_j$ ,  $j = 0, \dots, N-1$ , correspondant aux vecteurs de la base canonique.

Changer de base orthonormée (ici remplacer  $(e_j)_{j=0, \dots, N-1}$  par  $(f_j)_{j=0, \dots, N-1}$ ) pour exprimer un vecteur de  $\mathbb{C}^N$  (que l'on pensera ici comme un signal digital à valeurs complexes défini sur  $\{0, \dots, N-1\}$ ) ne change bien sûr en rien le carré de la norme  $l^2$  du vecteur des coordonnées; si

$$s = \sum_{j=0}^{N-1} \xi_j e_j = \sum_{j=0}^{N-1} \tilde{\xi}_j f_j,$$

on a bien sûr

$$\|\xi\|_2^2 = \sum_{j=0}^{N-1} |\xi_j|^2 = \sum_{j=0}^{N-1} |\tilde{\xi}_j|^2 = \|\tilde{\xi}\|_2^2 = \|s\|^2.$$

Mais il est une quantité qui, elle, peut changer de manière drastique, c'est l'entropie du signal lorsqu'il est exprimé dans la base  $\mathcal{B} = (e_j)_{j=0, \dots, N-1}$  (resp. dans la base  $\tilde{\mathcal{B}} = (f_j)_{j=0, \dots, N-1}$ ), définie par

$$\begin{aligned} \text{Entr}[s; \mathcal{B}] &:= - \sum_{j=0}^{N-1} |\xi_j|^2 \log_2 |\xi_j|^2 \\ (\text{resp. } \text{Entr}[s; \tilde{\mathcal{B}}] &:= - \sum_{j=0}^{N-1} |\tilde{\xi}_j|^2 \log_2 |\tilde{\xi}_j|^2). \end{aligned}$$

Cette notion d'entropie se révèle être (pensez à son incarnation en thermodynamique) un indicateur du chaos; plus l'entropie d'un signal digital exprimé dans une base est petite, mieux ce signal digital se présente comme « organisé » dans la base considérée, ce qui se lit aussi de la manière suivante: plus forte sera la pente du graphe de la fonction  $j \longmapsto |\xi_j|$  obtenue en réorganisant les  $|\xi_j|$  suivant l'indice  $j$  de manière à ce que cette fonction soit décroissante. Trouver une base orthonormée de  $\mathbb{C}^N$  dans laquelle l'entropie soit minimale est évidemment un challenge très intéressant aux fins d'applications futures (compression, transmission, stockage, etc.). Il est intuitivement évident que la base canonique  $(e_j)_{j=0, \dots, N-1}$  ne saurait être la base orthonormée de ce point de vue la plus adéquate pour encoder un signal digital « oscillant » de même que bien sûr la base  $(f_j)_{j=0, \dots, N-1}$  est sûrement bien moins intéressante que ne l'est la base canonique pour encoder une combinaison linéaire de fonctions « pic ». C'est face à ce type de problématique que se situe l'intérêt de l'analyse de Fourier discrète.

En effet, prendre la transformée de Fourier discrète d'un signal digital

$$s = (s(k))_{k=0, \dots, N-1},$$

ce n'est rien d'autre que précisément calculer (au facteur  $\sqrt{N}$  près) les coefficients de  $s$  dans cette nouvelle base orthonormée  $(f_j)_{j=0,\dots,N-1}$  alors que les nombres  $s(k)$ ,  $k = 0, \dots, N-1$  sont les coordonnées de  $s$  dans la base canonique.

**Proposition 2.1** *La  $N$ -transformée de Fourier discrète d'un élément*

$$(s(0), \dots, s(N-1))$$

de  $l_{\mathbb{C}}^2(\{0, \dots, N-1\})$  est par définition l'élément  $(\widehat{s}(0), \dots, \widehat{s}(N-1))$  défini par

$$\begin{aligned} \widehat{s}(j) &:= \left\langle s, (\exp(2i\pi jk/N))_{k=0,\dots,N-1} \right\rangle := \sum_{k=0}^{N-1} s(k) \exp\left(-\frac{2i\pi jk}{N}\right) \\ &= \sum_{k=0}^{N-1} s(k) \left[ \exp\left(-i\frac{2\pi j}{N}t\right) \right]_{t=k}, \quad j = 0, \dots, N \end{aligned} \quad (2.1)$$

Il s'agit d'une transformation linéaire inversible de  $\mathbb{C}^N$  dans  $\mathbb{C}^N$ , dont l'application inverse est donnée par

$$s(k) = \frac{1}{N} \sum_{j=0}^{N-1} \widehat{s}(j) \exp \frac{2i\pi jk}{N}$$

car la matrice symétrique  $M_N = \left[ \exp(-2i\pi jk/N) \right]_{0 \leq j, k \leq N-1}$  vérifie

$$M_N \cdot (M_N)^* = N I_N.$$

**Preuve.** Pour voir que  $M_N \cdot (M_N)^*$  est  $N$  fois la matrice identité, on utilise simplement le fait que  $W_N = \exp(-2i\pi/N)$  est une racine de  $X^N = 1$  et que toutes les racines  $\xi$  de cette équation autres que  $X = 1$  (c'est à dire tous les  $W_N^k$ ,  $k = 1, \dots, N-1$ ) vérifient

$$1 + \xi + \xi^2 + \dots + \xi^{N-1} = 0.$$

Le calcul de  $M_N \cdot (M_N)^*$  est alors considérablement simplifié.  $\diamond$

Dans le cas  $\mathbb{K} = \mathbb{R}$ , un système orthogonal intéressant de  $l_{\mathbb{R}}^2(\{0, \dots, N-1\})$  est celui qui est donné par les suites

$$\left( \alpha_j \cos \frac{\pi j(2k+1)}{N} \right)_{k=0,\dots,N-1}$$

avec  $\alpha_0 = \sqrt{1/N}$  et  $\alpha_j = \sqrt{2/N}$  si  $j > 0$  (ces suites forment, elles, un système orthonormé de  $l_{\mathbb{R}}^2(\{0, \dots, N-1\})$ ) et la transformation qui associe au vecteur

$$(s(0), \dots, s(N-1)) \in \mathbb{R}^n$$

le vecteur de coordonnées

$$\begin{aligned} s^{\cos}(j) &:= \left\langle s, \left( \alpha_j \cos \frac{\pi j(2k+1)}{2N} \right)_{k=0,\dots,N-1} \right\rangle \\ &= \alpha_j \sum_{k=0}^{N-1} s(k) \cos \frac{\pi j(2k+1)}{2N} \\ &= \alpha_j \sum_{k=0}^{N-1} s(k) \left[ \cos \frac{\pi j(2t+1)}{2N} \right]_{t=k}, \quad j = 0, \dots, N-1, \end{aligned}$$

est appelée *N-transformation en cosinus discrète*. La transformation inverse est donnée par les formules

$$s(k) = \sum_{j=0}^{N-1} \alpha_j s^{\cos}(j) \cos \frac{\pi j(2k+1)}{2N}, \quad k = 0, \dots, N-1.$$

Il faut donc interpréter ces transformations (*N-transformation de Fourier discrète* dans le cas complexe ou *N-transformation en cosinus discrète* dans le cas réel) comme l'effet sur la suite des coordonnées de changements de base ; dans les deux cas, on cherche à décomposer le vecteur  $(s(0), \dots, s(N-1))$  de  $\mathbb{K}^N$  (exprimé ainsi dans la base canonique de  $\mathbb{K}^N$ ) dans une nouvelle base constituée de fonctions oscillantes discrétisées aux points  $0, \dots, N-1$ , la gamme de fréquences étant  $[0, 2\pi[$  (les fréquences étant  $0, 2\pi/N, \dots, 2(N-1)\pi/N$ )<sup>3</sup>. Ces transformations se transposent au cadre des espaces de Hilbert

$$l_{\mathbb{K}}^2(\{0, \dots, N_1-1\} \times \{0, \dots, N_2-1\}).$$

Par exemple, la  $(N_1, N_2)$ -transformée de Fourier discrète d'une matrice

$$[I(k_1, k_2)]_{\substack{0 \leq k_1 \leq N_1-1 \\ 0 \leq k_2 \leq N_2-1}}$$

est la matrice  $[\widehat{I}(j_1, j_2)]_{\substack{0 \leq j_1 \leq N_1-1 \\ 0 \leq j_2 \leq N_2-1}}$  dont les entrées sont donnés par

$$\widehat{I}(j_1, j_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} I(k_1, k_2) \exp\left(-\frac{2i\pi k_1 j_1}{N_1}\right) \exp\left(-\frac{2i\pi k_2 j_2}{N_2}\right)$$

pour  $j_1 = 0, \dots, N_1-1$ ,  $j_2 = 0, \dots, N_2-1$ . Les formules inverses sont

$$I(k_1, k_2) = \frac{1}{N_1 N_2} \sum_{j_1=0}^{N_1-1} \sum_{j_2=0}^{N_2-1} \widehat{I}(j_1, j_2) \exp\frac{2i\pi k_1 j_1}{N_1} \exp\frac{2i\pi k_2 j_2}{N_2}$$

pour  $k_1 = 0, \dots, N_1-1$ ,  $k_2 = 0, \dots, N_2-1$ . Dans les versions bi-dimensionnelles, on privilégie souvent pour des raisons pratiques de symétrie les transformations réelles, donc plutôt la  $(N_1, N_2)$ -transformation en cosinus discrète plutôt que la  $(N_1, N_2)$ -transformation de Fourier discrète<sup>4</sup>.

Pour comprendre l'importance de la transformation de Fourier du point de vue opérationnel (on retrouvera toujours cette propriété par la suite), prenons deux suites  $(s_1(0), \dots, s_1(N-1))$  et  $(s_2(0), \dots, s_2(N-1))$  de  $l_{\mathbb{K}}^2(\{0, \dots, N-1\})$  et associons

<sup>3</sup>En fait, cela n'aurait pas grand sens d'envisager des fréquences au delà de ce seuil  $2\pi$  car la discrétisation aux entiers  $0, 1, \dots, N-1$  ne rendrait absolument plus compte des oscillations ! Notons que  $\exp(-2i\pi jk/N) = \exp(2i\pi j(N-k)/N)$  pour  $j = [N/2] + 1, \dots, N$ , ce qui fait que l'on peut aussi considérer que la gamme des fréquences explorées (si l'on entend prendre un intervalle symétrique autour de l'origine) est  $[-\pi, \pi]$ .

<sup>4</sup>C'est par exemple la version bi-dimensionnelle de la transformation en cosinus discrète qui est l'un des ingrédients majeurs des algorithmes de traitement ou de compression d'image tels JPEG (la compression se faisant au niveau de la transformée en cosinus discrète de blocs  $8 \times 8$  en lesquels le tableau  $(N_1, N_2)$  a été préalablement décomposé).

leur les deux polynômes

$$P_1(X) := s_1(0) + s_1(1)X + \cdots + s_1(N-1)X^{N-1} = \sum_{k=0}^{N-1} s_1(k)X^k$$

$$P_2(X) := s_2(0) + s_2(1)X + \cdots + s_2(N-1)X^{N-1} = \sum_{k=0}^{N-1} s_2(k)X^k.$$

Le reste de  $P_1(X)P_2(X)$  après division euclidienne par  $X^N - 1$  est le polynôme

$$\sum_{k=0}^{N-1} (s_1 *_{N} s_2(k)) X^k,$$

où

$$s_1 *_{N} s_2(k) := \sum_{l=0}^{N-1} \tilde{s}_1(k-l)\tilde{s}_2(l) = \sum_{l=0}^{N-1} \tilde{s}_2(k-l)\tilde{s}_1(l), \quad k = 0, \dots, N-1, \quad (2.2)$$

$\tilde{s}_1$  et  $\tilde{s}_2$  désignant les fonctions définies sur  $\mathbb{Z}$  et à valeurs dans  $\mathbb{K}$  prolongeant  $s_1$  et  $s_2$  par 1-périodicité. On a alors la proposition (et définition) immédiate suivante.

**Proposition 2.2** *L'opération  $(s_1, s_2) \mapsto s_1 *_{N} s_2$  de  $(l_{\mathbb{K}}^2(\{0, \dots, N-1\}))^2$  dans  $l_{\mathbb{K}}^2(\{0, \dots, N-1\})$  est appelée  $N$ -convolution cyclique. On a*

$$\widehat{s_1 *_{N} s_2}(j) = \widehat{s}_1(j) \times \widehat{s}_2(j), \quad \forall j \in \{0, \dots, N-1\}, \quad (2.3)$$

ce qui fait de la  $N$ -transformée de Fourier discrète un isomorphisme entre le groupe  $(l_{\mathbb{K}}^2(\{0, \dots, N-1\}), *_{N})$  et  $l_{\mathbb{K}}^2(\{0, \dots, N-1\})$  équipé de la multiplication « terme à terme ».

**Preuve.** On écrit

$$P_1(X)P_2(X) \equiv \sum_{k=0}^{N-1} (s_1 *_{N} s_2(k)) X^k \pmod{X^N - 1},$$

puis on pose  $X = \exp(-2i\pi j/N)$  ; on obtient ainsi

$$\widehat{s}_1(j)\widehat{s}_2(j) = \sum_{k=0}^{N-1} (s_1 *_{N} s_2(k)) \exp(-2i\pi jk/N) = \widehat{s_1 *_{N} s_2}(j),$$

ce qui donne le résultat voulu.  $\diamond$

Ce résultat très simple (mais d'un intérêt majeur<sup>5</sup>) soutend l'opération algorithmique de *multiplication rapide* de polynômes. Multiplier deux polynômes de degré au plus  $N-1$  est une opération qui « consomme »  $N^2$  multiplications, comme il est immédiat de s'en convaincre si l'on utilise aucune astuce particulière. Il en est de même du calcul de la  $N$ -transformée de Fourier discrète d'un signal digital de longueur  $N$  (on doit multiplier un vecteur par une matrice carrée de taille  $N \times N$ ), sauf dans le cas certes très particulier (mais pourtant essentiel)  $N = 2$  où aucune multiplication

<sup>5</sup>On en retrouvera de nombreuses versions tout au long de ce cours, transposées du cadre fini au cadre discret, du cadre discret au cadre continu,...

n'entre en jeu (seulement une addition et une soustraction <sup>6</sup>). On comprend aisément que si  $N$  est très grand (pensez à  $N = 10^6$ ), le coût d'un tel calcul numérique ( $N^2 = 10^{12}$ ) échappe très vite aux normes raisonnables !

Mais il y eut dans cette histoire un miracle ! Il fallut attendre 1965 pour que deux ingénieurs informaticiens américains, James William Cooley et John Wilder Tukey, mettent au point un algorithme récursif très simple<sup>7</sup> exploitant la simplicité du cas  $N = 2$  et permettant, lorsque  $N = 2^p$ , d'effectuer le calcul de la  $N$ -transformée de Fourier discrète d'un signal digital avec  $p \times N/2 = p \times 2^{p-1}$  et non plus  $N^2 = 2^{2p}$  multiplications ! Le gain est énorme et déclencha à l'aube des années 1970 ce qui allait être la « révolution numérique ». Jusque là resté instrument théorique au service du mathématicien ou du physicien, la transformation de Fourier allait devenir le formidable outil technologique qu'elle n'a cessé d'être depuis dans tous les domaines (télécommunications, imagerie, équations aux dérivées partielles, ...). La transformation de Fourier lorsque  $N$  est une puissance de 2 devient la *Fast Fourier Transform* (FFT) et l'on comprend pourquoi les écrans de télévision ou les consoles d'ordinateur sont souvent des configurations à 1024 sur 1024, 512 sur 256, ..., pixels ; on comprend aussi le regain d'intérêt pour les grands nombres de Fermat premiers  $M = 2^{2^m} + 1$  ou de Mersenne premiers  $M = 2^p - 1$ , pour envisager la multiplication « rapide » de polynômes de degrés strictement inférieur à  $[M/2]$  en calculant leur produit dans  $\mathbb{K}[X]/(X^{2^{2^m}} - 1)$  ou dans  $\mathbb{K}[X]/(X^{2^p} - 1)$  via la proposition 2.2, donc grâce à la  $2^{2^m}$  ou  $2^p$ -FFT et à son inversion (et en profitant dans ce cas des propriétés de l'arithmétique modulaire).

### 2.2.2 La transformée de Fourier sur $l_{\mathbb{C}}^1(\mathbb{Z})$

Comme on l'a vu dans la section précédente 2.2.1, le calcul de la transformée de Fourier d'un élément  $s = (s(0), \dots, s(N-1))$  de  $l_{\mathbb{C}}^2(\{0, \dots, N-1\})$  revient à calculer le produit scalaire de  $s$  avec les versions échantillonnées aux points  $0, \dots, N-1$  des fonction oscillantes

$$t \longmapsto \exp(i\omega_j t), \quad \omega_j = \frac{2\pi j}{N}, \quad j = 0, \dots, N-1.$$

Plus  $N$  augmente, plus l'éventail de fréquences  $\{2j\pi/N; j = 0, \dots, N-1\}$  remplit l'intervalle  $[0, 2\pi[$ . Le fait maintenant d'imaginer que  $N$  tende vers  $+\infty$  nous incite à définir la transformée de Fourier d'une suite  $(s_k)_{k \in \mathbb{Z}}$  de  $l_{\mathbb{C}}^1(\mathbb{Z})$  (en calquant les formules 2.1) comme suit :

**Définition 2.1** Soit  $s = (s_k)_{k \in \mathbb{Z}}$  un élément de  $l_{\mathbb{C}}^1(\mathbb{Z})$ . La transformée de Fourier (ou spectre) de  $s$  est la fonction continue  $2\pi$ -périodique sur  $\mathbb{R}$ , à valeurs en général

<sup>6</sup>La 2-transformation de Fourier discrète se résume aux deux formules  $\widehat{s}(0) = s(0) + s(1)$  et  $\widehat{s}(1) = s(0) - s(1)$ ; c'est d'ailleurs le seul cas où la matrice de la  $N$ -transformation de Fourier discrète (dite dans ce cas « matrice papillon » car c'est sous la forme de ce type de diagramme que l'on schématise cette transformation) est une matrice réelle.

<sup>7</sup>On trouvera par exemple une présentation succincte de l'algorithme de Cooley-Tukey dans mes notes du cours de l'UE MHT304 : « Calcul scientifique et symbolique et initiation aux logiciels », en ligne sur <http://www.math.u-bordeaux1.fr/~yger/mht304.pdf>, section 4.6. La clef de cet algorithme tient simplement dans la remarque  $W_{2^p}^{2^{p-1}} = -1$  si  $W_N := \exp(-2i\pi/N)$ , les entrées du vecteur colonne  $s$  étant préalablement triées entre entrées d'indice pair et entrées d'indice impair.

complexes<sup>8</sup> :

$$\widehat{s} : \omega \in \mathbb{R} \mapsto \sum_{k \in \mathbb{Z}} s_k e^{-ik\omega}. \quad (2.4)$$

L'application  $\mathcal{F} : s \mapsto \widehat{s}$  ainsi définie est une application linéaire continue injective du  $\mathbb{C}$ -espace de Banach  $l_{\mathbb{C}}^1(\mathbb{Z})$  dans le  $\mathbb{C}$ -espace de Banach  $C^0(\mathbb{T}, \mathbb{C})$  des fonctions continues  $2\pi$ -périodiques de  $\mathbb{R}$  dans  $\mathbb{C}$ , équipé de la norme  $\|\cdot\|_{\infty}$ .

**Preuve des affirmations incluses dans la définition.** La convergence des deux séries de fonctions  $(\omega \mapsto s_k e^{-ik\omega})_{k \in \mathbb{N}}$  et  $(\omega \mapsto s_{-k} e^{ik\omega})_{k \in \mathbb{N}^*}$  étant une convergence normale puisque

$$|s_{\pm k} e^{-\pm ik\omega}| = |s_{\pm k}|$$

et

$$\sum_{k \in \mathbb{Z}} |s_k| = \|s\|_1 < +\infty$$

par hypothèses, le spectre  $\widehat{s}$  d'un élément  $s \in l_{\mathbb{C}}^1(\mathbb{Z})$  est une fonction continue (en général à valeurs complexes, sauf dans le cas particulier où  $s_{-k} = \overline{s_k}$  pour tout  $k \in \mathbb{Z}$ ). De plus, on a

$$\|\widehat{s}\|_{\infty} := \sup_{[0, 2\pi]} |\widehat{s}| \leq \|s\|_1,$$

ce qui prouve la continuité de  $\mathcal{F}$ .

Il est immédiat de retrouver  $s$  à partir de son spectre (cela prouvera donc l'injectivité de  $\mathcal{F}$ ) *via* les relations

$$\begin{aligned} \frac{1}{2\pi} \int_0^{2\pi} \widehat{s}(\omega) e^{ik\omega} d\omega &= \frac{1}{2\pi} \int_0^{2\pi} \left( \sum_{l=-\infty}^{\infty} s_l e^{-il\omega} \right) e^{ik\omega} d\omega \\ &= \sum_{l \in \mathbb{Z}} s_l \left( \frac{1}{2\pi} \int_0^{2\pi} e^{i\omega(k-l)} d\omega \right) = s_k \end{aligned}$$

pour tout  $k \in \mathbb{Z}$  (c'est la convergence normale qui permet l'interversion série/intégrale à la seconde ligne du calcul ci-dessus).  $\diamond$

**Remarque 2.1.** Attention ! Il est faux que  $\mathcal{F} : s \mapsto \widehat{s}$ , considérée comme une application injective de  $l_{\mathbb{C}}^1(\mathbb{Z})$  dans l'espace  $C^0(\mathbb{T}, \mathbb{C})$ , soit surjective. Par exemple, si  $(\epsilon_k)_{k \geq 1}$  est une suite décroissante de nombres positifs tendant vers 0, alors le lemme d'Abel (uniforme) montre que la suite de fonctions

$$\left( \omega \in \mathbb{R} \mapsto \sum_{l=1}^k \epsilon_l \frac{\sin l\omega}{l} \right)_{k \geq 1}$$

converge uniformément sur  $\mathbb{R}$  vers une fonction continue  $2\pi$ -périodique ; pour le voir, on utilise le fait qu'il existe une constante absolue  $C$  telle que

$$\left\| \omega \mapsto \sum_{k=1}^N \frac{\sin k\omega}{k} \right\|_{\infty} \leq C \quad \forall N \in \mathbb{N}^* ;$$

en effet<sup>9</sup> :

$$\sum_{k=1}^N \frac{\sin k\omega}{k} = \int_0^{\omega} \left( \sum_{k=1}^N \cos ku \right) du$$

<sup>8</sup>Le spectre de  $(s_k)_{k \in \mathbb{Z}}$  est à valeurs réelles si et seulement si  $s_{-k} = \overline{s_k}$ ,  $\forall k \in \mathbb{Z}$ .

<sup>9</sup>Le calcul sera conduit plus en détails plus loin, voir la suite d'égalités 2.21 dans la section 2.3.3.



$$\begin{aligned}
&= \frac{1}{2} \int_0^\omega \left( \sum_{k=-N}^N e^{iku} - 1 \right) du \\
&= \frac{1}{2} \int_0^\omega \left( \frac{\sin(N+1/2)u}{\sin u/2} - 1 \right) du,
\end{aligned}$$

et l'on a à la fois

$$\left| \int_0^\omega \frac{\sin Nu}{u} du \right| \leq K_1, \quad \forall N \in \mathbb{N}^*, \quad \forall \omega \in [0, \pi]$$

et

$$\int_0^\omega \left| \frac{\sin(N+1/2)u}{2 \sin u/2} - \frac{\sin Nu}{u} \right| du \leq K_2, \quad \forall N \in \mathbb{N}^*, \quad \forall \omega \in [0, \pi]$$

pour deux constantes absolues  $K_1$  et  $K_2$ . Si l'on choisit par exemple comme suite  $(\epsilon_k)_{k \geq 1}$  la suite  $((\log(k+1))^{-1})_{k \geq 1}$ , on constate que la fonction

$$\omega \in \mathbb{R} \mapsto \sum_{k=1}^{\infty} \frac{\sin k\omega}{k \log(k+1)}$$

est une fonction continue  $2\pi$ -périodique qui ne peut pas être la transformée de Fourier d'une suite de  $l_{\mathbb{C}}^1(\mathbb{Z})$  puisque la série de Bertrand  $[1/(p \log p)]_{p \geq 2}$  est divergente!

On a vu dans le cours de théorie de l'intégration (proposition 4.9 du polycopié de l'UE MHT512) le rôle essentiel que revêtait l'opération interne de *convolution discrète* entre éléments de  $l_{\mathbb{K}}^1(\mathbb{Z})$  ( $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$ ), opération qui à deux suites  $(s_{1,k})_{k \in \mathbb{Z}}$  et  $(s_{2,k})_{k \in \mathbb{Z}}$  de  $l_{\mathbb{K}}^1(\mathbb{Z})$  associe la suite de terme général

$$(s_1 * s_2)_k := \sum_{l \in \mathbb{Z}} s_{1,l} s_{2,k-l} = \sum_{l \in \mathbb{Z}} s_{1,k-l} s_{2,l} = (s_2 * s_1)_k.$$

L'importance de cette opération tient surtout au fait que tout opérateur linéaire continu  $\mathcal{L}$  de  $l_{\mathbb{K}}^1(\mathbb{Z})$  dans lui-même invariant par translation (c'est-à-dire tel que  $\mathcal{L}[(s_{k-k_0})_k] = (\mathcal{L}[s]_{k-k_0})_k$  pour tout  $s = (s_k)_k$  dans  $l_{\mathbb{K}}^1(\mathbb{Z})$  et pour tout  $k_0$  dans  $\mathbb{Z}$ <sup>10</sup>) se représente précisément comme un tel opérateur de convolution

$$\mathcal{L} : s \mapsto h * s,$$

où  $h = \mathcal{L}[e_0]$ ,  $e_0$  désignant la suite  $(\delta_{0,k})_{k \in \mathbb{Z}}$ , est un élément particulier de  $l_{\mathbb{K}}^1(\mathbb{Z})$  dit *réponse impulsionnelle* du filtre digital  $\mathcal{L}$ .

La proposition 2.2 devient dans ce nouveau contexte l'énoncé suivant, confirmant le rôle « opérationnel » de la transformation de Fourier, rôle que l'on retrouvera dans le contexte continu plus tard :

**Proposition 2.3** *Si  $s_1$  et  $s_2$  sont deux éléments de  $l_{\mathbb{C}}^1(\mathbb{Z})$ , on a*

$$\widehat{s_1 * s_2}(\omega) = \widehat{s_1}(\omega) \times \widehat{s_2}(\omega), \quad \forall \omega \in \mathbb{R}.$$

**Preuve.** C'est une application du théorème de Fubini et de l'invariance par translation de la mesure de décompte sur  $\mathbb{Z}$  : on écrit

$$\begin{aligned}
\widehat{s_1 * s_2}(\omega) &= \sum_{k \in \mathbb{Z}} \left( \sum_{l \in \mathbb{Z}} s_{1,l} s_{2,k-l} \right) e^{-i\omega k} \\
&= \sum_{k \in \mathbb{Z}} \sum_{l \in \mathbb{Z}} s_{1,l} s_{2,k-l} e^{-i\omega k}
\end{aligned}$$

<sup>10</sup>Un tel opérateur est aussi appelé *boîte noire* ou *filtre digital*.

$$\begin{aligned}
&= \sum_{k \in \mathbb{Z}} \sum_{l \in \mathbb{Z}} s_{1,l} s_{2,k-l} e^{-i\omega(k-l)} e^{-i\omega l} \\
&= \sum_{l \in \mathbb{Z}} s_{1,l} e^{-i\omega l} \left( \sum_{k \in \mathbb{Z}} s_{2,k-l} e^{-i\omega(k-l)} \right) \\
&= \sum_{l \in \mathbb{Z}} s_{1,l} e^{-i\omega l} \left( \sum_{k \in \mathbb{Z}} s_{2,k} e^{-i\omega k} \right) \\
&= \widehat{s}_1(\omega) \times \widehat{s}_2(\omega).
\end{aligned}$$

On a appliqué Fubini deux fois (après avoir vérifié qu'en mettant des modules partout l'intégrale double est convergente, ce qui est vrai car les exponentielles introduites sont de module 1, donc invisibles lorsque l'on prend les modules des termes) : de la ligne 1 à la ligne 2, puis de la ligne 3 à la ligne 4. De la ligne 4 à la ligne 5, c'est l'invariance par translation de la mesure de comptage que l'on a utilisé. Mais il est clair que la formule  $\exp(z_1 + z_2) = \exp(z_1) \times \exp(z_2)$  (pour passer de la ligne 2 à la ligne 3) a joué un rôle décisif (c'est en fait la clef du problème).  $\diamond$

Tout ce que nous avons fait dans cette sous-section s'adapte pour définir naturellement la transformée de Fourier sur  $l_{\mathbb{C}}^1(\mathbb{Z}^n)$ . Par exemple, dans le cas  $n = 2$  (correspondant au cadre des images digitales), si  $[I(k_1, k_2)]_{k_1, k_2 \in \mathbb{Z}}$  est un tableau (indexé par  $\mathbb{Z}^2$ ) d'entrées complexes tel que

$$\sum_{k_1 \in \mathbb{Z}} \sum_{k_2 \in \mathbb{Z}} |I(k_1, k_2)| < \infty,$$

le spectre de  $I$  est la fonction continue ( $2\pi$ -périodique en les deux variables) :

$$(\omega_1, \omega_2) \in \mathbb{R}^2 \longmapsto \widehat{I}(\omega_1, \omega_2) := \sum_{k_1 \in \mathbb{Z}} \sum_{k_2 \in \mathbb{Z}} I(k_1, k_2) e^{-i(k_1 \omega_1 + k_2 \omega_2)}. \quad (2.5)$$

La convolution bidimensionnelle<sup>11</sup> de deux tels tableaux  $I_1$  et  $I_2$  est le tableau  $I_1 * I_2$  défini par

$$\begin{aligned}
\forall (k_1, k_2) \in \mathbb{Z}^2, (I_1 * I_2)(k_1, k_2) &= \sum_{l_1 \in \mathbb{Z}} \sum_{l_2 \in \mathbb{Z}} I_1(l_1, l_2) I_2(k_1 - l_1, k_2 - l_2) \\
&= \sum_{l_1 \in \mathbb{Z}} \sum_{l_2 \in \mathbb{Z}} I_1(k_1 - l_1, k_2 - l_2) I_2(l_1, l_2)
\end{aligned} \quad (2.6)$$

et on a la formule

$$\forall (\omega_1, \omega_2) \in \mathbb{R}^2, \widehat{I_1 * I_2}(\omega_1, \omega_2) = \widehat{I_1}(\omega_1, \omega_2) \times \widehat{I_2}(\omega_1, \omega_2). \quad (2.7)$$

### 2.2.3 La transformée de Fourier sur $l_{\mathbb{C}}^2(\mathbb{Z})$

L'espace  $l_{\mathbb{C}}^2(\mathbb{Z})$  (contenant strictement  $l_{\mathbb{C}}^1(\mathbb{Z})$ <sup>12</sup>) est un cadre de travail plus propice que celui de l'espace  $l_{\mathbb{C}}^1(\mathbb{Z})$ , ce pour deux raisons : d'une part parce que c'est

<sup>11</sup>On parle aussi de « masque » à propos d'une image  $I$  agissant par convolution sur toutes les autres.

<sup>12</sup>Si  $(s_k)_{k \in \mathbb{Z}}$  est dans  $l_{\mathbb{C}}^1(\mathbb{Z})$  avec  $\|s\|_1 \leq 1$ , alors on a  $|s_k| \leq 1$  pour tout  $k \in \mathbb{Z}$  et par conséquent  $\sum_k |s_k|^2 \leq \sum_k |s_k| = \|s\|_1 \leq 1$ .

un  $\mathbb{C}$ -espace de Hilbert (et l'on peut y faire de la géométrie « à la Pythagore »), d'autre part parce que la norme  $l^2$  correspond à une entité « physique », la notion d'énergie. Nous allons donc dans ce cadre élargi (par rapport au cadre de  $l^1_{\mathbb{C}}(\mathbb{Z})$ ) définir la notion de spectre.

On a vu au chapitre 1 (exemple 1.4) que le système constitué des fonctions

$$\omega \longmapsto \exp(ik\omega), \quad k \in \mathbb{Z},$$

constitue une base hilbertienne du  $\mathbb{C}$ -espace de Hilbert  $L^2_{\mathbb{C}}(\mathbb{T})$ . Par conséquent, si  $(s_k)_{k \in \mathbb{Z}}$  est un élément de  $l^2_{\mathbb{C}}(\mathbb{Z})$ , la série bilatère des classes de fonctions

$$\omega \longmapsto \sum_{k \in \mathbb{Z}} s_k e^{-ik\omega}, \quad k \in \mathbb{Z},$$

converge (au niveau des classes dans  $L^2_{\mathbb{C}}(\mathbb{T})$ ) vers un élément  $\widehat{s}$  de  $L^2_{\mathbb{C}}(\mathbb{T})$ . Cela conduit à la définition suivante :

**Définition 2.2** On appelle spectre  $\widehat{s}$  d'un élément  $s = (s_k)_{k \in \mathbb{Z}}$  de  $l^2_{\mathbb{C}}(\mathbb{Z})$  l'élément  $\widehat{s}$  de  $L^2_{\mathbb{C}}(\mathbb{T})$  (c'est donc une classe de fonction mesurable de  $\mathbb{R}$  dans  $\mathbb{C}$ ,  $2\pi$ -périodique et d'énergie finie sur  $[0, 2\pi]$ ) tel que

$$\lim_{N \rightarrow +\infty} \frac{1}{2\pi} \int_0^{2\pi} \left| \widehat{s}(\omega) - \sum_{k=-N}^N s_k e^{-ik\omega} \right|^2 d\omega = 0.$$

L'application  $\mathcal{F} : s \in l^2_{\mathbb{C}}(\mathbb{Z}) \longmapsto \widehat{s} \in L^2_{\mathbb{C}}(\mathbb{T})$  réalise une isométrie linéaire bijective entre ces deux espaces, qui coïncide sur  $l^1_{\mathbb{C}}(\mathbb{Z})$  avec la transformation  $\mathcal{F}$  de la définition 2.1.

**Remarque 2.2.** Le fait que l'application  $\mathcal{F}$  de la définition 2.2 réalise une isométrie correspond à un principe physique, le principe de *conservation d'énergie*; nous verrons en effet ultérieurement que la transformation de Fourier  $\mathcal{F}$  (ou encore la prise de spectre), pensée dans le cadre continu et non plus comme ici discret, a une incarnation concrète en deux dimensions ou trois (on le verra dans une section ultérieure lorsque nous envisagerons la transformation de Fourier agissant sur  $L^1_{\mathbb{K}}(\mathbb{R}^n)$ ), à savoir le processus de diffraction à travers une lentille ou un prisme en optique ou en cristallographie. La formule qui rend compte de ce principe (suivant le premier volet de la proposition 1.9), à savoir ici

$$\sum_{k \in \mathbb{Z}} |s_k|^2 = \frac{1}{2\pi} \int_0^{2\pi} |\widehat{s}(\omega)|^2 d\omega, \quad \forall s = (s_k)_{k \in \mathbb{Z}} \in l^2_{\mathbb{C}}(\mathbb{Z}),$$

est, dans ce cadre discret, connue comme la formule de Parseval<sup>13</sup>. Cette formule se « duplique » de la manière suivante : si  $s_1 = (s_{1,k})_{k \in \mathbb{Z}}$  et  $s_2 = (s_{2,k})_{k \in \mathbb{Z}}$  sont deux éléments de  $l^2_{\mathbb{C}}(\mathbb{Z})$ ,

$$\frac{1}{2\pi} \int_0^{2\pi} \widehat{s}_1(\omega) \overline{\widehat{s}_2(\omega)} d\omega = \sum_{k \in \mathbb{Z}} s_{1,k} \overline{s_{2,k}}. \quad (2.8)$$

**Remarque 2.3.** Attention! Au contraire de ce qui se passe concernant le spectre  $\widehat{s}$  d'une suite  $s \in l^1_{\mathbb{C}}(\mathbb{Z})$ , le spectre d'une suite  $s$  de  $l^2_{\mathbb{C}}(\mathbb{Z})$  est une classe de fonctions, non une fonction, donc n'a pas de définition ponctuelle ( $\widehat{s}(\omega_0)$  pour une valeur spécifique de  $\omega_0$  n'a aucun sens!).

<sup>13</sup>C'est au mathématicien français Marc Antoine Parseval des Chênes (1755-1836), dont d'ailleurs on connaît très peu de la vie, que revient en 1799 l'intuition de la formule qui porte son nom; la formule de Plancherel concerne plutôt, elle, la théorie relative à la transformation intégrale de Fourier (point de vue continu, que l'on verra lorsque nous envisagerons la transformation de Fourier agissant sur  $L^2_{\mathbb{C}}(\mathbb{R}^n)$ , et non, comme ici, dans un cadre discret.

Contrairement à ce qui se passait dans le cadre de  $l_{\mathbb{K}}^1(\mathbb{Z})$ , la convolution n'est plus une opération interne sur  $l_{\mathbb{K}}^2(\mathbb{Z})$  ( $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$ ). Cependant, du fait du théorème de Young dans le cadre discret (théorème 4.3 du cours de théorie de l'intégration de l'UE MHT512), on a le résultat suivant. Si  $s_1 = (s_{1,k})_{k \in \mathbb{Z}}$  et  $s_2 = (s_{2,k})_{k \in \mathbb{Z}}$  sont deux éléments de  $l_{\mathbb{K}}^2(\mathbb{Z})$ , on peut encore définir leur convolution (sur le modèle de ce que l'on a fait pour définir la convolution entre éléments de  $l_{\mathbb{K}}^1(\mathbb{Z})$ ) ; en effet, pour chaque  $k \in \mathbb{Z}$ ,

$$\sum_{l \in \mathbb{Z}} |s_{1,l}| |s_{2,k-l}| \leq \sqrt{\sum_{l \in \mathbb{Z}} |s_{1,l}|^2} \times \sqrt{\sum_{l \in \mathbb{Z}} |s_{2,k-l}|^2} = \|s_1\|_2 \times \|s_2\|_2,$$

ce qui prouve que

$$s_1 * s_2 := \left( \sum_{l \in \mathbb{Z}} s_{1,l} s_{2,k-l} \right)_{k \in \mathbb{Z}}$$

est bien défini : c'est un élément de  $l_{\mathbb{K}}^\infty(\mathbb{Z})$  avec

$$\|s_1 * s_2\|_\infty \leq \|s_1\|_2 \|s_2\|_2.$$

On a toujours commutativité et associativité de cette opération,  $e_0$  est élément neutre, et l'opération coïncide (si on la restreint à  $l_{\mathbb{K}}^1(\mathbb{Z})$ ) avec celle qui a été définie dans le cadre  $l^1$ .

Dans ce contexte  $l^2$ , nous avons donc la proposition suivante :

**Proposition 2.4** *Soit  $\mathcal{L}$  un opérateur linéaire continu de  $l_{\mathbb{C}}^2(\mathbb{Z})$  dans lui-même tel que, pour tout  $s = (s_k)_{k \in \mathbb{Z}} \in l_{\mathbb{C}}^2(\mathbb{Z})$ , on ait<sup>14</sup>*

$$\mathcal{L}[(s_{k-k_0})_{k \in \mathbb{Z}}] = \left( \mathcal{L}[s]_{k-k_0} \right)_{k \in \mathbb{Z}}, \quad \forall k_0 \in \mathbb{Z}. \quad (2.9)$$

Si  $h = (h_k)_{k \in \mathbb{Z}} = \mathcal{L}[e_0]$ , on a  $\widehat{h} \in L_{\mathbb{C}}^\infty(\mathbb{T})$ . De plus, pour tout  $s \in l_{\mathbb{C}}^2(\mathbb{Z})$ , on a

$$\mathcal{L}[s] = s * h \in l_{\mathbb{C}}^2(\mathbb{Z})$$

et

$$\widehat{s * h} = \widehat{s} \times \widehat{h} \quad (2.10)$$

(la dernière égalité étant entendue au sens des fonctions  $2\pi$ -périodiques à valeurs complexes définies presque partout). Réciproquement, si  $h \in l_{\mathbb{C}}^2(\mathbb{Z})$  est tel que l'on ait  $\widehat{h} \in L_{\mathbb{C}}^\infty(\mathbb{T})$ , l'opérateur

$$s \longmapsto \mathcal{F}^{-1}[\widehat{s} \cdot \widehat{h}]$$

est un opérateur linéaire continu de  $l_{\mathbb{C}}^2(\mathbb{Z})$  dans lui-même, tel que  $\mathcal{L}[e_0] = h$ , se pliant à la clause (2.9).

<sup>14</sup>Cette dernière condition (2.9) signifie que les paramètres de l'appareil réalisant l'action  $\mathcal{L}$  sont immuables dans le temps. On dit qu'un tel opérateur linéaire continu  $\mathcal{L}$  est un *filtre discret stationnaire*, ce dernier qualificatif rendant précisément compte de l'immuabilité dans le temps des paramètres de l'appareil réalisant l'action du filtre.

**Preuve.** Prouvons d'abord l'assertion directe. En utilisant la linéarité et la continuité de  $\mathcal{L}$ , couplées avec la clause (2.9), on voit, si  $h = \mathcal{L}[e_0]$ , que

$$\forall k \in \mathbb{Z}, \quad \mathcal{L}[s]_k = \sum_{l \in \mathbb{Z}} s_l h_{k-l} = \sum_{l \in \mathbb{Z}} h_l s_{k-l}.$$

On a donc bien  $\mathcal{L}[s] = s * h$ ; on a affaire avec  $\mathcal{L}[s]$  *a priori* à un élément de  $l_{\mathbb{C}}^{\infty}(\mathbb{Z})$ , mais de fait cet élément est dans  $l_{\mathbb{C}}^2(\mathbb{Z})$  puisque  $\mathcal{L}$  est supposé être un opérateur de  $l_{\mathbb{C}}^2(\mathbb{Z})$  dans lui-même. La propriété (2.10) se montre en observant tout d'abord que la continuité de  $\mathcal{L}$ , combinée avec la continuité de  $\mathcal{F}$ , permettent de se ramener au cas où les  $s_k$  sont tous nuls pour  $|k| > M$ ,  $M$  étant un entier positif (on approche ensuite  $s$  dans  $l_{\mathbb{C}}^2(\mathbb{Z})$  par ses versions « tronquées »  $s^{(n)}$  où  $s_k^{(n)} = s_k$  si  $|k| \leq n$ ,  $s_k^{(n)} = 0$  sinon). Pour un tel  $s$  et pour tout  $N \in \mathbb{N}$ , on a

$$\begin{aligned} \sum_{k=-N}^N (s * h)_k e^{-ik\omega} &= \sum_{k=-N}^N \left( \sum_{l=-M}^M s_l h_{k-l} \right) e^{-ik\omega} \\ &= \sum_{l=-M}^M s_l \left( \sum_{k=-N}^N h_{k-l} e^{-ik\omega} \right) \\ &= \sum_{l=-M}^M s_l \left( \sum_{k=-N}^N h_{k-l} e^{-i(k-l)\omega} e^{-il\omega} \right) \\ &= \sum_{l=-M}^M s_l e^{-il\omega} \left( \sum_{k=-N-l}^{N-l} h_k e^{-ik\omega} \right) \end{aligned} \quad (2.11)$$

(grâce au théorème de Fubini applicable ici car les ensembles d'intégration sont finis et à l'invariance par translation de la mesure de comptage). Comme

$$\lim_{\substack{N \rightarrow \infty \\ L_{\mathbb{C}}^2(\mathbb{T})}} \left( \omega \mapsto \sum_{k=-N-l}^{N-l} h_k e^{-ik\omega} \right) = \widehat{h}$$

pour tout  $l \in \{-M, \dots, M\}$  (par continuité de  $\mathcal{F}$  sur  $l_{\mathbb{C}}^2(\mathbb{Z})$ ) et que

$$\lim_{\substack{N \rightarrow \infty \\ L_{\mathbb{C}}^2(\mathbb{T})}} \left( \omega \mapsto \sum_{k=-N}^N (s * h)_k e^{-ik\omega} \right) = \widehat{s * h},$$

on déduit bien de (2.11) en faisant tendre  $N$  vers l'infini et en prenant les limites dans  $L_{\mathbb{C}}^2(\mathbb{T})$  la formule (2.10). Comme  $\|\mathcal{L}[s]\|_2 \leq C\|s\|_2$  pour tout  $s \in l_{\mathbb{C}}^2(\mathbb{Z})$  (pour une certaine constante  $C \geq 0$ ) parce que  $\mathcal{L}$  est continu de  $l_{\mathbb{C}}^2(\mathbb{Z})$  dans lui-même, on a

$$\begin{aligned} \|\mathcal{L}[s]\|_2^2 &= \frac{1}{2\pi} \int_0^{2\pi} |\widehat{s * h}(\omega)|^2 d\omega = \frac{1}{2\pi} \int_0^{2\pi} |\widehat{s}(\omega)|^2 |\widehat{h}(\omega)|^2 d\omega \\ &\leq C^2 \|s\|_2^2 = C^2 \int_0^{2\pi} |\widehat{s}(\omega)|^2 d\omega \end{aligned}$$

pour tout  $s$  dans  $l_{\mathbb{C}}^2(\mathbb{Z})$ , donc pour tout  $\widehat{s}$  dans  $L_{\mathbb{C}}^2(\mathbb{T})$ ; on en déduit

$$|\widehat{h}(\omega)|^2 \leq C^2 \text{ pour presque tout } \omega,$$

donc que  $\widehat{h}$  est dans  $L^\infty(\mathbb{T})$ , de norme  $\|\widehat{h}\|_\infty \leq C$ .

L'assertion réciproque résulte de la continuité de  $\mathcal{F}^{-1}$  et du fait que l'opérateur de  $L^2_{\mathbb{C}}(\mathbb{T})$  dans lui-même défini par multiplication par une fonction  $2\pi$ -périodique mesurable essentiellement bornée par une constante  $C$  est un opérateur continu de norme inférieure ou égale à  $C$ . Le fait que l'opérateur ainsi construit se plie à la clause (2.9) résulte du fait que la spectre de  $(s_{k-k_0})_k$  s'obtient en multipliant  $\widehat{s}$  par la classe du caractère  $\omega \mapsto e^{-ik_0\omega}$ .  $\diamond$

**Remarque 2.4 : le vocabulaire des ingénieurs.** Étant donné un tel filtre discret stationnaire  $\mathcal{L}$ , la classe de la fonction  $2\pi$ -périodique  $\widehat{h}$  (où  $h = \mathcal{L}[e_0]$ ) est dite *transformée de Fourier* du filtre, tandis que la classe de la fonction  $2\pi$ -périodique

$$\omega \in \mathbb{R} \mapsto \sum_{k \in \mathbb{Z}} h_k e^{ik\omega}$$

est dite *fonction de transfert* du filtre  $\mathcal{L}$ . La classe  $\dot{D}_{\mathcal{L}}$  de  $|\widehat{h}|^2$  dans  $L^\infty(\mathbb{T})$  est appelée *densité spectrale de puissance* du filtre  $\mathcal{L}$ . L'énergie de  $\mathcal{L}[s]$  s'obtient donc (via la formule de Parseval) par

$$\|\mathcal{L}[s]\|_2^2 = \frac{1}{2\pi} \int_{[0, 2\pi[} D_{\mathcal{L}}(\omega) |\widehat{s}(\omega)|^2 d\omega.$$

**Exemple 2.1.** Un exemple important de cette proposition est celui des opérateurs de la forme

$$s \mapsto s * h$$

où  $h = (h_k)_{k \in \mathbb{Z}}$  est un élément de  $l^1_{\mathbb{C}}(\mathbb{Z})$ . On remarque en effet dans ce cas que l'opérateur

$$\mathcal{L}_h : s \in l^2_{\mathbb{C}}(\mathbb{Z}) \mapsto \mathcal{F}^{-1}[\widehat{h} \cdot \widehat{s}]$$

est bien continu de  $l^2_{\mathbb{C}}(\mathbb{Z})$  dans lui-même et que  $\mathcal{L}_h[e_0] = \mathcal{F}^{-1}[\widehat{h}] = h$ ; la convolée de  $s$  et de  $h$  (considéré comme élément de  $l^2_{\mathbb{C}}(\mathbb{Z})$ ) est donc bien dans  $l^2_{\mathbb{C}}(\mathbb{Z})$  et on définit ainsi un opérateur continu de  $l^2_{\mathbb{C}}(\mathbb{Z})$  dans lui-même.

Tous les résultats établis dans cette section en dimension 1 se transposent naturellement, exactement comme nous l'avons vu à la fin de la section 2.2.2 (par exemple, dans le cas  $n = 2$ , mêmes définitions pour le spectre et pour la convolutions de deux tableaux dont l'un des deux a un spectre dans  $L^\infty(\mathbb{T}^2)$  que les définitions (2.5), (2.6) et même formule (2.7)).

## 2.3 Transformation de Fourier sur les espaces de fonctions $2\pi$ -périodiques

On note toujours  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$ . Les deux espaces qui vont nous intéresser dans cette section sont les  $\mathbb{K}$ -espaces de Hilbert  $L^2_{\mathbb{K}}(\mathbb{T})$  et les  $\mathbb{K}$ -espaces de Banach  $L^1_{\mathbb{K}}(\mathbb{T})$  (le premier  $\mathbb{K}$ -espace vectoriel s'injecte ici continuellement dans le second, alors que l'on avait  $l^1_{\mathbb{K}}(\mathbb{Z}) \subset l^2_{\mathbb{K}}(\mathbb{Z})$ ). On rappelle que, si  $p \in [1, \infty]$ , la norme  $\|\cdot\|_{\mathbb{T}, p}$  est définie par

$$\|f\|_{\mathbb{T}, p}^p := \frac{1}{2\pi} \int_{[0, 2\pi[} |f(\theta)|^p d\theta.$$

Ces définitions peuvent bien sûr être étendues au cadre multidimensionnel, avec la définition des  $\mathbb{K}$ -espaces de Banach  $L^p_{\mathbb{K}}(\mathbb{T}^n)$ , avec leur norme définie par

$$\|f\|_{\mathbb{T}, p}^p := \frac{1}{(2\pi)^n} \int_{([0, 2\pi[)^n} |f(\theta)|^p d\theta,$$

mais, pour simplifier les choses ici, nous nous limiterons essentiellement dans cette section au cas  $n = 1$ .

### 2.3.1 La transformation de Fourier sur $L_{\mathbb{K}}^2(\mathbb{T})$

On notera à partir de maintenant

$$\|\dot{f}\|_{\mathbb{T},2}^2 := \frac{1}{2\pi} \int_0^{2\pi} |f(\theta)|^2 d\theta$$

le carré de la norme hilbertienne d'un élément  $\dot{f}$  du  $\mathbb{K}$ -espace de Hilbert  $L_{\mathbb{K}}^2(\mathbb{T})$ ,  $\dot{f}$  étant un représentant de la classe  $f$ .

On a vu dans la section 2.2.3 que l'application  $\mathcal{F}$  (prise de spectre ou transformation de Fourier) transformait les suites  $(s_k)_{k \in \mathbb{Z}}$  de  $l_{\mathbb{C}}^2(\mathbb{Z})$  de manière isométrique en les éléments  $\hat{s}$  de  $L_{\mathbb{C}}^2(\mathbb{T})$ . L'application inverse  $\mathcal{F}^{-1}$  est l'application

$$\mathcal{F}^{-1} : \dot{f} \in L_{\mathbb{C}}^2(\mathbb{T}) \longmapsto \left( \frac{1}{2\pi} \int_0^{2\pi} f(\theta) e^{ik\theta} d\theta \right)_{k \in \mathbb{Z}}. \quad (2.12)$$

Comme on l'a vu dès l'introduction de la transformation discrète (proposition 2.1), le principe de l'inversion de la transformée de Fourier est, du point de vue algébrique, celui qui consiste à remplacer un certain opérateur (de matrice  $M$  dans le cas de la transformation de Fourier discrète) par l'opérateur « conjugué » (de matrice  $\overline{M}$  à une constante près dans le cas de la transformation de Fourier discrète). Suivant ce principe, puisque  $\mathcal{F}^{-1} : L_{\mathbb{C}}^2(\mathbb{T}) \longmapsto l_{\mathbb{C}}^2(\mathbb{Z})$  est définie suivant le mécanisme (2.12), il est naturel de définir la transformation de Fourier  $\mathcal{TF} : L_{\mathbb{C}}^2(\mathbb{T}) \longmapsto l_{\mathbb{C}}^2(\mathbb{Z})$  suivant le mécanisme

$$\mathcal{TF} = \overline{\mathcal{F}^{-1}} : \dot{f} \in L_{\mathbb{C}}^2(\mathbb{T}) \longmapsto \left( \frac{1}{2\pi} \int_0^{2\pi} f(\theta) e^{-ik\theta} d\theta \right)_{k \in \mathbb{Z}}. \quad (2.13)$$

Le nombre

$$c_k(\dot{f}) := \frac{1}{2\pi} \int_0^{2\pi} f(\theta) e^{-ik\theta} d\theta, \quad k \in \mathbb{Z},$$

(qui est appelé *k-ème coefficient de Fourier complexe* de  $\dot{f}$ ) représente le produit scalaire (ou plus concrètement la « corrélation ») entre  $\dot{f}$  et l'harmonique fondamentale  $\dot{e}_k$  de période  $2\pi$ , classe de :

$$\theta \longmapsto e^{ik\theta}, \quad k \in \mathbb{Z}.$$

Remarquons que les seules fonctions oscillantes élémentaires sur  $\mathbb{R}$  du type

$$t \longmapsto e^{i\omega t}$$

avec  $\omega \in \mathbb{R}$  qui soient  $2\pi$ -périodiques sont précisément les fonctions

$$\theta \longmapsto e^{ik\theta}, \quad k \in \mathbb{Z},$$

et uniquement celles-ci. Le dictionnaire (dénombrable)  $\mathcal{D}$  constitué de classes de ces fonctions est bien LE dictionnaire de toutes les classes de fonctions oscillantes élémentaires de la forme  $t \longmapsto e^{i\omega t}$  qui soient  $2\pi$ -périodiques. Calculer le spectre

d'une classe de fonction  $2\pi$ -périodique  $\dot{f} \in L^2_{\mathbb{C}}(\mathbb{R}/2\pi\mathbb{Z})$ , c'est chercher à exprimer  $\dot{f}$  suivant précisément ce dictionnaire, ce aux fins de reconnaître en  $\dot{f}$  une certaine combinaison de fonctions oscillantes élémentaires<sup>15</sup>. Ceci revient à calculer la liste des coordonnées  $(\langle \dot{f}, \dot{e}_k \rangle)_{k \in \mathbb{Z}}$  de  $\dot{f}$  dans la base hilbertienne  $(\dot{e}_k)_{k \in \mathbb{Z}}$  de  $L^2_{\mathbb{C}}(\mathbb{T})$  (voir l'exemple 1.4).

**Définition 2.3** On appelle *spectre (complexe) d'un élément  $\dot{f} \in L^2_{\mathbb{C}}(\mathbb{T})$*  la suite

$$(c_k(\dot{f}))_{k \in \mathbb{Z}} = (\langle \dot{f}, \dot{e}_k \rangle)_{k \in \mathbb{Z}} \in l^2_{\mathbb{C}}(\mathbb{Z})$$

de ses coefficients de Fourier. L'application  $\mathcal{TF}$  de prise de spectre est une isométrie bijective entre ces deux espaces, l'inverse étant l'application

$$(\mathcal{TF})^{-1} = \overline{\mathcal{F}} : (s_k)_{k \in \mathbb{Z}} \mapsto \left( \theta \mapsto \sum_{k \in \mathbb{Z}} s_k e^{ik\theta} \right),$$

où la convergence de la série  $\sum_k s_k e^{ik(\cdot)}$  est entendue au sens de la norme  $\| \cdot \|_{\mathbb{T},2}$ , soit

$$\dot{f} = \lim_{\substack{N \rightarrow \infty \\ L^2_{\mathbb{R}}(\mathbb{T})}} \left( \theta \mapsto \sum_{k=-N}^N c_k(\dot{f}) e^{ik\theta} \right). \quad (2.14)$$

**Remarque 2.5.** Le fait que  $\mathcal{TF}$  soit une isométrie (ce qui correspond toujours à la conservation d'énergie après prise de spectre) nous assure que si  $\dot{f}_1$  et  $\dot{f}_2$  sont deux éléments de  $L^2_{\mathbb{C}}(\mathbb{T})$ , alors :

$$\frac{1}{2\pi} \int_0^{2\pi} f_1(\theta) \overline{f_2(\theta)} d\theta = \sum_{k \in \mathbb{Z}} c_k(\dot{f}_1) \overline{c_k(\dot{f}_2)}. \quad (2.15)$$

C'est la version « dupliquée » de la formule de Parseval (que nous avons en fait déjà rencontré en 2.8).

Pour  $\dot{f} \in L^2_{\mathbb{C}}(\mathbb{T})$  et  $N \in \mathbb{N}$ , le polynôme trigonométrique

$$S_N[\dot{f}] : \theta \mapsto \sum_{k=-N}^N c_k(\dot{f}) e^{ik\theta}$$

est appelé *N-ième somme partielle de Fourier* de  $\dot{f}$ . Du point de vue de la géométrie hilbertienne, sa classe est la projection orthogonale de  $\dot{f}$  sur le sous-espace vectoriel engendré par les harmoniques fondamentales  $\dot{e}_k$ ,  $-N \leq k \leq N$ , de représentants

$$e_k : \theta \mapsto e^{ik\theta}, \quad k = -N, \dots, N.$$

On remarque, si l'on pose

$$\begin{aligned} a_k(\dot{f}) &:= \frac{1}{\pi} \int_0^{2\pi} f(\theta) \cos(k\theta) d\theta, \quad k \in \mathbb{N} \\ b_k(\dot{f}) &:= \frac{1}{\pi} \int_0^{2\pi} f(\theta) \sin(k\theta) d\theta, \quad k \in \mathbb{N}^*, \end{aligned}$$

<sup>15</sup>Il est intuitivement évident que cette prise de spectre n'est sans doute pas la chose la plus intelligente à tenter lorsqu'il est manifeste que  $\dot{f}$  ne s'exprime pas naturellement dans le dictionnaire  $\mathcal{D}$ ; pensez par exemple à la fonction obtenue en  $2\pi$ -périodisant une fonction pic telle la fonction  $\tilde{g}_\epsilon$ , avec  $\epsilon$  petit, définie par

$$g_\epsilon(t) := \frac{1}{\sqrt{2\pi\epsilon}} \exp\left(-\frac{t^2}{2\epsilon^2}\right).$$



on peut réécrire

$$S_N[\dot{f}](\theta) = \frac{a_0(\dot{f})}{2} + \sum_{k=1}^N \left( a_k(\dot{f}) \cos(k\theta) + b_k(\dot{f}) \sin(k\theta) \right), \quad \forall \theta \in \mathbb{R}. \quad (2.16)$$

Ces remarques nous conduisent à la définition suivante (dans le cadre réel cette fois) :

**Définition 2.4** On appelle spectre (réel) d'un élément  $\dot{f} \in L^2_{\mathbb{R}}(\mathbb{T})$  la liste des scalaires  $(a_0(\dot{f}), (a_k(\dot{f}), b_k(\dot{f}))_{k \geq 1})$ . Pour tout  $\dot{f} \in L^2_{\mathbb{R}}(\mathbb{T})$ , on a

$$\dot{f} = \lim_{N \rightarrow \infty} \left( \theta \mapsto \frac{a_0(\dot{f})}{2} + \sum_{k=1}^N \left( a_k(\dot{f}) \cos(k\theta) + b_k(\dot{f}) \sin(k\theta) \right) \right); \quad (2.17)$$

de plus

$$\|\dot{f}\|_{\mathbb{T},2}^2 = \frac{(a_0(\dot{f}))^2}{4} + \frac{1}{2} \sum_{k=1}^{\infty} \|(a_k(\dot{f}), b_k(\dot{f}))\|^2.$$

**Preuve.** Les assertions complétant cette définition sont de simples transcriptions des assertions de la définition 2.3 via la réécriture 2.16 et les relations  $a_0(\dot{f}) = 2c_0(\dot{f})$  et  $a_k(\dot{f}) + ib_k(\dot{f}) = 2c_k(\dot{f})$ ,  $a_k(\dot{f}) - ib_k(\dot{f}) = 2c_{-k}(\dot{f})$  si  $k \geq 1$ .  $\diamond$

Les formules d'inversion (2.14) et (2.17) sont à la base même de l'analyse de Fourier des signaux périodiques<sup>16</sup>. Tout ceci se transpose bien sûr au cadre multidimensionnel.

### 2.3.2 La transformation de Fourier sur $L^1_{\mathbb{K}}(\mathbb{T})$

Si  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$ , le  $\mathbb{K}$ -espace de Banach  $L^1_{\mathbb{K}}(\mathbb{T})$  contient strictement (comme  $\mathbb{K}$ -sous-espace fermé) le  $\mathbb{K}$ -sous-espace  $L^2_{\mathbb{K}}(\mathbb{T})$ ; l'injection du  $\mathbb{K}$ -espace de Hilbert  $L^2_{\mathbb{K}}(\mathbb{T})$  (équipé de sa norme  $\|\cdot\|_{\mathbb{T},2}$ ) dans  $L^1_{\mathbb{K}}(\mathbb{T})$  (équipé de sa norme  $\|\cdot\|_{\mathbb{T},1}$ ) est d'ailleurs une injection linéaire continue. L'action de la transformation de Fourier, définie à la section précédente sur  $L^2_{\mathbb{K}}(\mathbb{T})$ , s'étend de fait au cadre élargi qu'est celui du  $\mathbb{K}$ -espace de Banach  $L^1_{\mathbb{K}}(\mathbb{T})$ . L'aspect hilbertien sera toutefois perdu et l'on ne pourra donc plus en profiter dans ce cadre élargi.

Si  $\dot{f}$  est un élément de  $L^1_{\mathbb{C}}(\mathbb{T})$ , on peut toujours définir la suite  $(c_k(\dot{f}))_{k \in \mathbb{Z}}$  de ses coefficients de Fourier complexes :

$$c_k(\dot{f}) := \frac{1}{2\pi} \int_0^{2\pi} f(\theta) e^{-ik\theta} d\theta, \quad k \in \mathbb{Z},$$

quand bien même on ne peut plus interpréter ces nombres comme des produits scalaires entre deux éléments d'un  $\mathbb{C}$ -espace de Hilbert.

<sup>16</sup>C'est l'idée de base du mathématicien français Jean-Baptiste Joseph Fourier (1768-1830), sur laquelle repose l'étude des phénomènes physiques oscillants, qu'en un sens à préciser, tout phénomène physique 1-dimensionnel  $T$ -périodique se réalise comme un empilement de  $T$ -harmoniques fondamentales complexes (*resp.* réelles), affectées de coefficients correspondant précisément aux coefficients de Fourier complexes (*resp.* réels).

**Définition 2.5** On appelle spectre (complexe) d'un élément  $\dot{f} \in L^1_{\mathbb{C}}(\mathbb{T})$  la suite  $(c_k(\dot{f}))_{k \in \mathbb{Z}}$  de ses coefficients de Fourier. L'application  $\mathcal{TF}$  de prise de spectre est une application linéaire continue de  $L^1_{\mathbb{C}}(\mathbb{T})$  dans l'espace de Banach  $l^{\infty,0}_{\mathbb{C}}(\mathbb{Z})$  des suites de nombres complexes tendant vers 0 à l'infini<sup>17</sup>, équipé de la norme uniforme. Cette application prolonge l'application  $\mathcal{TF} : L^2_{\mathbb{C}}(\mathbb{T}) \mapsto l^2_{\mathbb{C}}(\mathbb{Z})$  de la définition 2.3.

**Preuve.** On a immédiatement

$$\sup_k |c_k(\dot{f})| \leq \|\dot{f}\|_{\mathbb{T},1}. \quad (2.18)$$

Pour tout  $k \in \mathbb{N}^*$ , on introduit (comme dans l'exemple 1.10) la classe  $\dot{\Psi}_k$  de  $L^1_{\mathbb{R}}(\mathbb{T})$ , de représentant la fonction polynômiale trigonométrique

$$\theta \mapsto 2\pi \frac{\left(\frac{1+\cos\theta}{2}\right)^k}{\int_{[-\pi,\pi[} \left(\frac{1+\cos u}{2}\right)^k du}.$$

D'après la proposition 1.13 appliquée avec l'exemple 1.10, on peut, étant donné  $\epsilon > 0$ , choisir  $k_{\epsilon} \in \mathbb{N}^*$  assez grand pour que

$$\|\dot{f} - \dot{f} \overset{\text{per}}{*} \dot{\Psi}_{k_{\epsilon}}\| < \epsilon.$$

Or, si  $k_{\epsilon}$  est ainsi fixé, la classe

$$\dot{f} \overset{\text{per}}{*} \dot{\Psi}_{k_{\epsilon}}$$

a pour représentant une fonction de la forme

$$\theta \mapsto \sum_{-k_{\eta}}^{k_{\epsilon}} c_{k_{\epsilon},l}(\dot{f}) e^{il\theta}$$

et l'on vérifie immédiatement que l'on a

$$c_k(\dot{f} \overset{\text{per}}{*} \dot{\Psi}_{k_{\epsilon}}) = 0$$

pour  $|k| > k_{\epsilon}$ . On a donc

$$\forall k \in \mathbb{Z}, |c_k(\dot{f})| = |c_k(\dot{f} - \dot{f} \overset{\text{per}}{*} \dot{\Psi}_{k_{\epsilon}})| \leq \epsilon.$$

Ceci prouve bien

$$\lim_{|k| \rightarrow +\infty} |c_k(\dot{f})| = 0$$

et donc le fait que l'opération  $\mathcal{TF}$  de prise de spectre réalise une application  $\mathbb{C}$ -linéaire de  $L^1_{\mathbb{C}}(\mathbb{T})$  dans l'espace de Banach  $l^{\infty,0}_{\mathbb{C}}(\mathbb{Z})$ , sous-espace fermé de  $l^{\infty}_{\mathbb{C}}(\mathbb{Z})$  constitué des suites tendant vers 0 à l'infini. La continuité de la prise de spectre complexe résulte de la majoration 2.18.  $\diamond$

<sup>17</sup>Cette propriété, pour la suite  $(c_k(\dot{f}))_{k \in \mathbb{Z}}$  des coefficients de Fourier d'un élément de  $L^1_{\mathbb{C}}(\mathbb{T})$ , de tendre vers 0 à l'infini, est la version discrète du lemme de Riemann-Lebesgue.

**Remarque 2.6.** Attention ! La transformation de Fourier  $\mathcal{TF} : L_{\mathbb{C}}^1(\mathbb{T}) \longrightarrow l_{\mathbb{C}}^{\infty,0}(\mathbb{Z})$  n'est malheureusement pas surjective : pour le voir, anticipons un peu le cours et prenons par exemple la suite  $(u_k)_{k \in \mathbb{Z}}$  de  $l_{\mathbb{C}}^{\infty,0}(\mathbb{Z})$  définie par

$$u_k = \begin{cases} \frac{1}{i \log k}, & \text{si } k \geq 2 \\ 0 & \text{si } k = -1, 0, 1 \\ -\frac{1}{i \log |k|} & \text{si } k \leq -2. \end{cases}$$

S'il existait  $\dot{f}$  dans  $L_{\mathbb{C}}^1(\mathbb{T})$  tel que  $u_k = c_k(\dot{f})$  et que l'on introduise la fonction (nécessairement continue et  $2\pi$ -périodique puisque  $\dot{f} \in L_{\mathbb{C}}^1(\mathbb{T})$  et que  $c_0(\dot{f}) = 0$ ) définie par

$$F(\theta) = \int_0^\theta f(u) du, \quad \theta \in \mathbb{R},$$

on vérifierait par intégration par parties que, pour tout  $k$  tel que  $|k| \geq 2$ ,

$$c_{|k|}(\dot{F}) = \frac{1}{2\pi} \int_0^{2\pi} F(\theta) e^{-ik\theta} d\theta = \frac{1}{2\pi} \left[ F(\theta) \frac{e^{-ik\theta}}{-ik} \right]_0^{2\pi} + \frac{1}{ik} c_k(\dot{f}) = -\frac{1}{|k| \log |k|}.$$

Mais la série de Bertrand  $[1/(p \log p)]_{p \geq 2}$  est une série divergente<sup>18</sup>, alors que le théorème de Féjér version locale (que nous verrons plus loin, théorème 2.3) nous indique que

$$\lim_{N \rightarrow +\infty} \left( c_0(\dot{F}) - \frac{1}{N} \sum_{n=0}^{N-1} \sum_{\substack{k=-n \\ k \neq -1, 0, 1}}^n \frac{1}{|k| \log |k|} \right) = F(0) = 0.$$

Or, comme

$$\sum_{\substack{k=-n \\ k \neq -1, 0, 1}}^n \frac{1}{|k| \log |k|} \sim 2 \log(\log n) \xrightarrow{n \rightarrow +\infty} +\infty,$$

on a

$$\lim_{N \rightarrow +\infty} \left( -\frac{1}{N} \sum_{n=0}^{N-1} \sum_{\substack{k=-n \\ k \neq -1, 0, 1}}^n \frac{1}{|k| \log |k|} \right) = -\infty,$$

ce qui est contradictoire.

Comme dans le contexte discret (propositions 2.2, 2.3, 2.4), la transformation  $\mathcal{TF}$  transforme l'opération de convolution  $2\pi$ -périodique entre éléments de  $L_{\mathbb{C}}^1(\mathbb{T})$ , définie en (1.28) par le fait que  $\dot{f} \overset{\text{per}}{*} \dot{g}$  a pour représentant

$$\theta \longmapsto \frac{1}{2\pi} \int_{[0, 2\pi[} f(\theta - u) g(u) du = \frac{1}{2\pi} \int_{[0, 2\pi[} f(u) g(\theta - u) du \quad \text{p.p.},$$

en multiplication terme à terme au niveau des spectres. On a le résultat suivant :

**Proposition 2.5** Soit  $\dot{f}_1$  et  $\dot{f}_2$  deux éléments de  $L_{\mathbb{C}}^1(\mathbb{T})$ . Pour tout  $k \in \mathbb{Z}$ , on a

$$c_k(\dot{f}_1 \overset{\text{per}}{*} \dot{f}_2) = c_k(\dot{f}_1) \times c_k(\dot{f}_2).$$

**Preuve.** On a

$$c_k(\dot{f}_1 \overset{\text{per}}{*} \dot{f}_2) = \frac{1}{2\pi} \int_0^{2\pi} \left( \frac{1}{2\pi} \int_0^{2\pi} f_1(u) f_2(\theta - u) du \right) e^{-ik\theta} d\theta$$

<sup>18</sup>Voir par exemple le polycopié du cours de MHT401.

$$\begin{aligned}
&= \frac{1}{4\pi^2} \int_0^{2\pi} f_1(u) \left( \int_0^{2\pi} f_2(\theta - u) e^{-ik(\theta - u)} d\theta \right) e^{-iku} du \\
&= \frac{1}{4\pi^2} \int_0^{2\pi} f_1(u) e^{-iku} \left( \int_0^{2\pi} f_2(\theta) e^{-ik\theta} d\theta \right) du \\
&= c_k(\dot{f}_1) \times c_k(\dot{f}_2)
\end{aligned}$$

(le théorème de Fubini permet de passer de la ligne 1 à la ligne 2, puis vient le fait que l'exponentielle échange addition et multiplication, enfin pour finir on utilise l'invariance de la mesure de décompte par translation).  $\diamond$

Pour envisager la restitution de  $\dot{f}$  à partir de la connaissance des coefficients de Fourier  $c_k(\dot{f})$ ,  $k \in \mathbb{Z}$ , on peut par exemple utiliser la propriété (proposition 1.13 et exemple 1.10) :

$$\dot{f} = \lim_{k \rightarrow \infty} (\dot{f} \underset{*}{\overset{\text{per}}{\Psi}}_k),$$

où

$$\Psi_k : \theta \longmapsto 2\pi \frac{\left(\frac{1+\cos\theta}{2}\right)^k}{\int_{[-\pi, \pi[} \left(\frac{1+\cos u}{2}\right)^k du} = \sum_{l=-k}^{l=k} \gamma_{k,l} e^{il\theta}, \quad k \in \mathbb{N}^*.$$

On remarque en effet, grâce à la proposition 2.5, que l'on a

$$\dot{f} \underset{*}{\overset{\text{per}}{\Psi}}_k = \sum_{l=-k}^{l=k} \gamma_{k,l} c_l(\dot{f}) \dot{e}_l,$$

ce qui nous permet d'affirmer

$$\lim_{k \rightarrow +\infty} \left\| \dot{f} - \sum_{l=-k}^k \gamma_{k,l} c_l(\dot{f}) \dot{e}_l \right\|_{\mathbb{T},1} = 0. \quad (2.19)$$

On dispose ainsi d'un procédé « en deux temps » pour « reconstituer »  $\dot{f}$  à partir de la connaissance de son spectre complexe : pour rendre l'erreur d'approximation assez petite, il faut commencer par fixer  $k$  assez grand, puis, ce choix étant fait, former la classe du polynôme trigonométrique

$$\theta \longmapsto \sum_{l=-k}^k \gamma_{k,l} c_l(\dot{f}) e^{il\theta}.$$

Un tel procédé s'appelle un procédé *taubérien*<sup>19</sup>.

Bien sûr un résultat idéal au niveau de l'inversion de la transformation  $\mathcal{TF} : L_{\mathbb{C}}^1(\mathbb{T}) \longmapsto l_{\mathbb{C}}^{\infty,0}(\mathbb{Z})$  aurait été

$$\lim_{N \rightarrow +\infty} \left\| \dot{f} - \sum_{k=-N}^N c_k(\dot{f}) e^{-ik\theta} \right\|_{\mathbb{T},1} = 0.$$

<sup>19</sup>La terminologie fait référence au mathématicien austro-hongrois Alfred Tauber, né en 1866 et mort en déportation autour de 1942, qui a introduit ce type de procédé pour « resommer » des séries divergentes.

Mais il se trouve que ce résultat est FAUX en général pour  $\dot{f} \in L^1_{\mathbb{C}}(\mathbb{T})$ !<sup>20</sup>. Nous y reviendrons dans la section suivante et verrons comment rectifier le tir.

Il est un cas particulier important où l'inversion de la transformation de Fourier est immédiate (vous avez sûrement croisé l'exemple 2.2 qui l'illustre dans le cours de L2 sur les séries de Fourier) :

**Proposition 2.6** *Si  $\dot{f}$  est un élément de  $L^1_{\mathbb{C}}(\mathbb{T})$  tel que la suite  $(c_k(\dot{f}))_{k \in \mathbb{Z}}$  soit dans  $l^1_{\mathbb{C}}(\mathbb{Z})$ , un représentant de  $\dot{f}$  est donné par la fonction continue  $2\pi$ -périodique*

$$\theta \longmapsto \sum_{k=-\infty}^{\infty} c_k(\dot{f}) e^{ik\theta}, \quad (2.20)$$

la série ci-dessus étant normalement convergente sur  $\mathbb{R}$ .

**Preuve.** La convergence normale de la série (2.20) est immédiate ; la suite  $(c_k)_{k \in \mathbb{Z}}$  est aussi un élément de  $l^2_{\mathbb{C}}(\mathbb{Z})$  et  $\dot{f}$  est aussi la limite dans  $L^2_{\mathbb{C}}(\mathbb{T})$  de la suite des classes

$$\left( \sum_{k=-N}^N c_k(\dot{f}) e_k \right)_{N \in \mathbb{N}^*}$$

(puisque l'injection de  $L^2_{\mathbb{C}}(\mathbb{T})$  dans  $L^1_{\mathbb{C}}(\mathbb{T})$  est continue). La classe  $\dot{f}$  est dans  $L^2_{\mathbb{C}}(\mathbb{T})$  et, d'après le fait que l'on puisse extraire d'une suite  $(\dot{S}_N)_{N \in \mathbb{N}}$  convergeant dans  $L^2_{\mathbb{C}}(\mathbb{T})$  une sous-suite dont les représentants convergent presque partout (avatar du théorème de Riesz-Fischer, voir la proposition 4.6 du cours de théorie de l'intégration), un représentant de  $\dot{f}$  est bien la fonction  $2\pi$ -périodique continue

$$\theta \longmapsto \sum_{k=-\infty}^{\infty} c_k(\dot{f}) e^{ik\theta}.$$

La proposition est démontrée.  $\diamond$

**Exemple 2.2.** Un exemple majeur d'application de ce résultat est celui où  $\dot{f}$  est la classe d'une fonction  $2\pi$ -périodique continue et  $C^1$  par morceaux, c'est-à-dire telle qu'il existe une subdivision

$$a_0 = 0 < a_1 < a_2 < \dots < a_N = 2\pi$$

de manière à ce que la restriction de  $f$  à chaque segment  $[a_j, a_{j+1}]$ ,  $j = 0, \dots, N-1$ , soit de classe  $C^1$ . Alors, on vérifie facilement en écrivant (grâce à des intégrations par parties sur chaque  $[a_j, a_{j+1}]$ , pour  $k \in \mathbb{Z}^*$ ,

$$\begin{aligned} 2\pi c_k(\dot{f}) &= \sum_{j=0}^{N-1} \int_{a_j}^{a_{j+1}} f(\theta) e^{-ik\theta} d\theta \\ &= \sum_{j=0}^{N-1} \left( \left[ f(\theta) \frac{e^{-ik\theta}}{-ik} \right]_{a_j}^{a_{j+1}} + \frac{1}{ik} \int_{a_j}^{a_{j+1}} f'(\theta) e^{-ik\theta} d\theta \right), \end{aligned}$$

<sup>20</sup>Ceci est cependant vrai dès que  $\dot{f}$  admet un représentant dans  $L^2_{\mathbb{C}}(\mathbb{T})$ , en vertu de la formule (2.14) établie à la section 2.3.1 et du fait que la convergence dans  $L^2_{\mathbb{C}}(\mathbb{T})$  implique la convergence dans  $L^1_{\mathbb{C}}(\mathbb{T})$ , ce qui complique singulièrement le problème pour exhiber un contre-exemple. En fait, on peut exhiber un élément  $\dot{f}$  de  $L^1_{\mathbb{C}}(\mathbb{T})$  tel que la suite de fonctions  $(S_N[\dot{f}])_{N \in \mathbb{Z}}$  diverge presque partout sur  $\mathbb{R}$  ; on ne peut donc extraire de sous-suite convergente presque partout, ce qui aurait été le cas si la suite  $(S_N[\dot{f}])_{N \geq 0}$  avait convergé vers  $\dot{f}$  dans  $L^1_{\mathbb{C}}(\mathbb{T})$  (proposition 4.6 du cours de théorie de l'intégration MHT512). Pour voir comment effectuer cette construction (subtile), voir par exemple le livre d'Y. Katznelson, *An introduction to harmonic analysis*, John Wiley and Sons, 1968, pp. 59-61 (voir aussi les exercices p. 50 du même ouvrage). Cette construction sort du cadre du cours, mais il était utile de mentionner le résultat.

que

$$(ik)c_k(\dot{f}) = c_k(\dot{g}),$$

où  $g$  est définie par  $g(\theta) = f'(\theta)$  sur  $\mathbb{R}$  privé des  $a_j + 2\pi\mathbb{Z}$ ,  $j = 0, \dots, N$ . Comme  $\dot{g}$  est dans  $L^2_{\mathbb{C}}(\mathbb{T})$ , la suite  $(c_k(\dot{g}))_{k \in \mathbb{Z}}$  est dans  $l^2_{\mathbb{C}}(\mathbb{Z})$  et la suite  $(c_k(\dot{f}))_{k \in \mathbb{Z}}$  est dans  $l^1_{\mathbb{C}}(\mathbb{Z})$  grâce à l'inégalité de Cauchy-Schwarz (comme produit terme à terme de deux éléments de  $l^2_{\mathbb{C}}(\mathbb{Z})$ ). La proposition 2.6 s'applique ici et l'on peut affirmer que

$$f(\theta) = \sum_{k \in \mathbb{Z}} c_k(\dot{f}) e^{ik\theta} \quad \forall \theta \in \mathbb{R},$$

la série de fonctions au second membre étant normalement convergente sur  $\mathbb{R}$ .

On termine cette section en disant deux mots du cas réel, où la notion de spectre est plus mal commode à gérer (on préférera de loin à travailler, même lorsque l'on a affaire à des éléments de  $L^1_{\mathbb{R}}(\mathbb{T})$ , avec le spectre complexe en pensant ces éléments comme des éléments de  $L^2_{\mathbb{C}}(\mathbb{T})$ ).

**Définition 2.6** On appelle spectre (réel) d'un élément  $\dot{f} \in L^1_{\mathbb{R}}(\mathbb{T})$  la collection  $(a_0(\dot{f}), (a_k(\dot{f}), b_k(\dot{f}))_{k \in \mathbb{N}^*})$  de ses coefficients réels :

$$\begin{aligned} a_k(\dot{f}) &:= \frac{1}{\pi} \int_0^{2\pi} f(\theta) \cos(k\theta) d\theta, \quad k \in \mathbb{N} \\ b_k(\dot{f}) &:= \frac{1}{\pi} \int_0^{2\pi} f(\theta) \sin(k\theta) d\theta, \quad k \in \mathbb{N}^*. \end{aligned}$$

On a  $\lim_{k \rightarrow +\infty} \|(a_k(\dot{f}), b_k(\dot{f}))\| = 0$ . Cette application prolonge l'application  $\mathcal{TF} : L^2_{\mathbb{R}}(\mathbb{T}) \mapsto l^2_{\mathbb{R}}(\mathbb{Z})$  de la définition 2.4.

Le point important à retenir concernant la transformation de Fourier réelle et sa relation avec la transformée de Fourier complexe est la double écriture des sommes partielles de Fourier

$$S_N[\dot{f}](\theta) = \sum_{k=-N}^N c_k(\dot{f}) e^{ik\theta} = \frac{a_0(\dot{f})}{2} + \sum_{k=1}^N (a_k(\dot{f}) \cos(k\theta) + b_k(\dot{f}) \sin(k\theta)), \quad \forall N \geq 1.$$

et toujours les relations  $a_k(\dot{f}) + ib_k(\dot{f}) = 2c_k(\dot{f})$ ,  $a_k(\dot{f}) - ib_k(\dot{f}) = 2c_{-k}(\dot{f})$  si  $k \geq 1$ .

### 2.3.3 Les théorèmes de Dirichlet et de Féjer

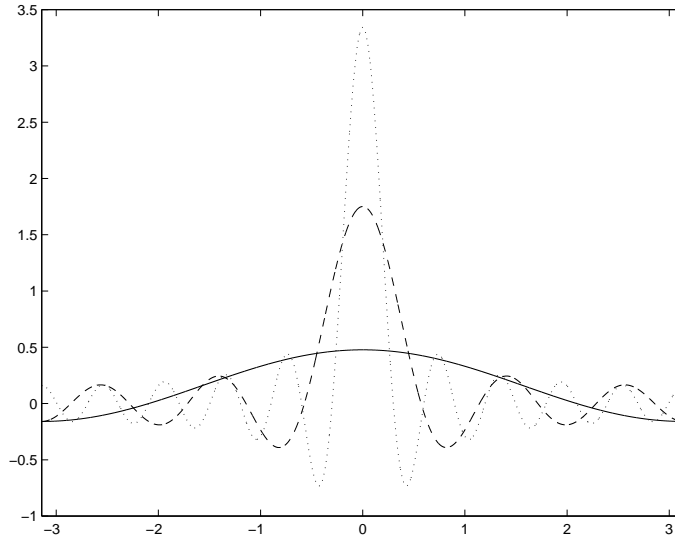


FIG. 2.1 – Graphes sur  $[-\pi, \pi]$  de  $D_N/2\pi$ ,  $N = 1, 5, 10$

Si  $\dot{f}$  est un élément de  $L^1_{\mathbb{C}}(\mathbb{T})$ , il est facile d'exprimer  $S_N[f]$  comme la convolée  $2\pi$ -périodique de  $\dot{f}$  et d'un certain noyau (dépendant de  $N$ ); on a en effet :

$$\begin{aligned}
 \forall \theta \in \mathbb{R}, \quad S_N[\dot{f}](\theta) &= \sum_{k=-N}^{k=N} \left( \frac{1}{2\pi} \int_0^{2\pi} f(u) e^{-iku} du \right) e^{ik\theta} \\
 &= \int_0^{2\pi} f(u) \left[ \frac{1}{2\pi} \sum_{k=-N}^{k=N} e^{ik(\theta-u)} \right] du \\
 &= \int_0^{2\pi} f(u) \left[ \frac{1}{2\pi} \left( 2\operatorname{Re} \left( \sum_{k=0}^N e^{ik(\theta-u)} \right) - 1 \right) \right] du \\
 &= \int_0^{2\pi} f(u) \left[ \frac{1}{2\pi} \left( 2\operatorname{Re} \left[ \frac{1 - e^{i(N+1)(\theta-u)}}{1 - e^{i(\theta-u)}} \right] - 1 \right) \right] du \\
 &= \int_0^{2\pi} f(u) \left[ \frac{1}{2\pi} \left( 2\operatorname{Re} \left[ e^{-i\frac{N(\theta-u)}{2}} \frac{\sin \frac{(N+1)(\theta-u)}{2}}{\sin \frac{\theta-u}{2}} \right] - 1 \right) \right] du \\
 &= \int_0^{2\pi} f(u) \left[ \frac{1}{2\pi} \left( \frac{2 \cos \frac{N(\theta-u)}{2} \sin \frac{(N+1)(\theta-u)}{2}}{\sin \frac{\theta-u}{2}} - 1 \right) \right] du \\
 &= \frac{1}{2\pi} \int_0^{2\pi} f(u) D_N(\theta - u) du, \tag{2.21}
 \end{aligned}$$

où  $D_N$  est la fonction continue  $2\pi$ -périodique continue définie par

$$D_N(\theta) := \sum_{k=-N}^{k=N} e^{ik\theta} = \begin{cases} \frac{\sin(N + \frac{1}{2})\theta}{\sin \frac{\theta}{2}} & \text{si } \theta \neq 0 \text{ mod. } 2\pi \\ 2N + 1 & \text{si } \theta = 0 \text{ mod. } 2\pi. \end{cases}$$

Cette fonction  $D_N$ , dont nous avons représenté le graphe pour diverses valeurs de  $N$  ( $N = 1, 5, 10$ ) sur la figure 2.1, est appelée *noyau de Dirichlet* (d'ordre  $N$ )<sup>21</sup>. Ce graphe sur  $[-\pi, \pi]$  présente un lobe central et des lobes latéraux ; on remarque que l'intégrale sur  $[-\pi, \pi]$  de la fonction  $D_N$  vaut  $2\pi$ , mais que cette fonction  $D_N$  n'est pas positive, ce qui représentera, on le verra plus loin, un handicap important : la norme  $\|\dot{D}_N\|_{\mathbb{T},1}$  tend vers  $+\infty$  lorsque  $N$  tend vers l'infini !

Reprenons maintenant  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$ . On peut donc écrire, pour tout  $\dot{f} \in L^1_{\mathbb{K}}(\mathbb{T})$  :

$$\dot{S}_N[\dot{f}] = \dot{f} \overset{\text{per}}{*} \dot{D}_N. \quad (2.22)$$

Nous pouvons alors énoncer le résultat suivant dans le cadre  $L^1_{\mathbb{K}}(\mathbb{T})$  :

**Théorème 2.1 (Théorème de Jordan-Dirichlet)** *Soit  $\dot{f} \in L^1_{\mathbb{K}}(\mathbb{T})$  et  $\theta_0 \in \mathbb{R}, l \in \mathbb{K}$  deux scalaires tels que*

$$\int_0^\pi \frac{|f(\theta_0 + u) + f(\theta_0 - u) - 2l|}{|u|} du < +\infty. \quad (2.23)$$

alors

$$\lim_{N \rightarrow +\infty} S_N[\dot{f}](\theta_0) = l.$$

**Remarque 2.7 et rappels de L2.** Il s'agit ici d'un résultat de nature très différente de ce que nous avons fait jusque là, à savoir un résultat local. On en connaît depuis le cours de L2<sup>22</sup> l'application majeure : si  $\dot{f}$  a un représentant  $f$  présentant une limite à gauche ( $f(\theta_0 - 0)$ ) en  $\theta_0$  ainsi qu'une limite à droite ( $f(\theta_0 + 0)$ ) en ce même point, alors on peut affirmer que

$$\lim_{N \rightarrow +\infty} S_N[\dot{f}](\theta_0) = \frac{f(\theta_0 - 0) + f(\theta_0 + 0)}{2}$$

à condition toutefois que soit remplie la clause de sécurité suivante :

Le graphe de  $f$  présente des demi-tangentes à gauche et à droite respectivement aux points  $(\theta_0, f(\theta_0 - 0))$  et  $(\theta_0, f(\theta_0 + 0))$ .

Si tel est le cas, on peut écrire en effet, dans  $]0, \epsilon]$  avec  $\epsilon > 0$  assez petit

$$\begin{aligned} f(\theta_0 + u) &= f(\theta_0 + 0) + u f'(\theta_0 + 0) + u \varphi(u) \\ f(\theta_0 - u) &= f(\theta_0 - 0) - u f'(\theta_0 - 0) + u \psi(u) \end{aligned}$$

où  $\varphi$  et  $\psi$  sont des fonctions bornées sur  $]0, \epsilon]$ , et par conséquent

$$\frac{|f(\theta_0 + u) + f(\theta_0 - u) - (f(\theta_0 - 0) + f(\theta_0 + 0))|}{|u|} = |f'(\theta_0 + 0) - f'(\theta_0 - 0) + o(1)| \leq c$$

pour une certaine constante  $c > 0$  ; la clause (2.23) est donc bien remplie et l'on a dans ce cas

$$\lim_{N \rightarrow +\infty} S_N[\dot{f}](\theta_0) = \frac{f(\theta_0 + 0) + f(\theta_0 - 0)}{2}.$$

<sup>21</sup>La terminologie fait référence à l'analyste et théoricien des nombres allemand Johann Peter Gustav Lejeune Dirichlet (1805-1859) qui l'introduisit et le manipula en 1828.

<sup>22</sup>Voir le polycopié du cours de l'UE MHT401, en ligne sur : <http://www.math.u-bordeaux1.fr/~yger/mat401.pdf>



**Preuve.** On écrit, en utilisant la formule (2.22) en  $\theta_0$ , la parité de  $D_N$  et le fait que  $D_N$  soit d'intégrale  $2\pi$  sur  $[\theta_0 - \pi, \theta_0 + \pi]$ ,

$$\begin{aligned} S_N[f](\theta_0) - l &= \frac{1}{2\pi} \int_{\theta_0 - \pi}^{\theta_0 + \pi} (f(\theta_0 - u) - l) D_N(u) du \\ &= \frac{1}{2\pi} \int_0^\pi (f(\theta_0 + u) + f(\theta_0 - u) - 2l) D_N(u) du \\ &= \frac{1}{2\pi} \int_0^\pi \frac{f(\theta_0 + u) + f(\theta_0 - u) - 2l}{u} \frac{u}{\sin(u/2)} \sin \left[ \left( N + \frac{1}{2} \right) u \right] du. \end{aligned}$$

La fonction

$$u \longmapsto \chi_{[0, \pi]}(u) \frac{f(\theta_0 + u) + f(\theta_0 - u) - 2l}{u} \frac{u}{\sin(u/2)},$$

prolongée par  $2\pi$ -périodicité, définit (du fait de l'hypothèse (2.23)) une classe de  $L^1_{\mathbb{R}}(\mathbb{T})$  et l'on conclut avec le lemme de Riemann-Lebesgue mentionné dans les définitions 2.5 et 2.6.  $\diamond$

Ce résultat est vraiment un résultat local ; pour nous en convaincre, voici un exemple troublant.

On considère la fonction  $2\pi$ -périodique  $\varphi$  définie par

$$\varphi(\theta) = \begin{cases} 1 & \text{si } \theta \in [-\pi, 0[ \\ 0 & \text{si } \theta \in [0, \pi[. \end{cases} \quad (2.24)$$

Cette fonction définit un élément de  $L^1_{\mathbb{C}}(\mathbb{T})$ .

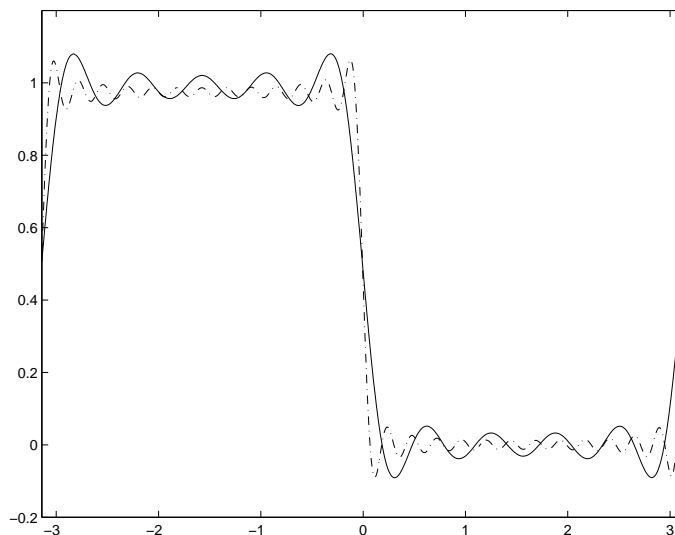


FIG. 2.2 – Phénomène de Gibbs pour  $\varphi$  avec  $N = 10$  et  $N = 25$

Sur la figure 2.2, voici ce que l'on observe, en traçant les graphes de  $S_{10}[\varphi]$  (en trait plein) et  $S_{25}[\varphi]$  (en pointillés) : lorsque  $N$  augmente, il y a un phénomène de « rehaussement » de la fonction  $\varphi$  (lorsqu'on tente de l'approcher par  $S_N[\varphi]$ ) avant la singularité en 0 et d'« affaissement » passée cette singularité ; l'amplitude de ce phénomène ne fléchit pas lorsque  $N$  augmente, c'est ce que l'on appelle un effet

d'*aliasing*; c'est le phénomène de Gibbs<sup>23</sup>, qui refroidit ici nos espoirs d'approcher dans  $L^1_{\mathbb{K}}(\mathbb{T})$  la classe  $\dot{f}$  par la suite  $(S_N[\dot{f}])_{N \geq 0}$ .

Pour tenter une parade, nous allons « couper en douceur » les composantes « hautes-fréquences », c'est-à-dire les contributions à  $\dot{f}$  des

$$\theta \longmapsto e^{ik\theta}, \quad |k| > N.$$

On envisage pour cela

$$T_N[\dot{f}] : \theta \longmapsto \sum_{k=-N}^N \left(1 - \frac{|k|}{N}\right) e^{ik\theta}$$

que l'on appelle la *N-ième somme de Féjer*<sup>24</sup> de  $\dot{f}$ . Reprenons ici les calculs de trigonométrie que nous avons fait pour exprimer  $S_N[\dot{f}]$  en terme d'une convolution  $2\pi$ -périodique avec le noyau de Dirichlet (formule (2.22)). On a :

$$\begin{aligned} \forall \theta \in \mathbb{R}, \quad T_N[\dot{f}](\theta) &= \sum_{k=-(N-1)}^{k=N-1} \frac{N - |k|}{N} \left( \frac{1}{2\pi} \int_0^{2\pi} f(u) e^{-iku} du \right) e^{ik\theta} \\ &= \frac{1}{N} \int_0^{2\pi} f(u) \left[ \frac{1}{2\pi} \sum_{k=-N-1}^{k=N-1} (N - |k|) e^{ik(\theta-u)} \right] du \\ &= \frac{1}{N} \int_0^{2\pi} f(u) \left[ \frac{1}{2\pi} \left( 2\operatorname{Re} \left( \sum_{k=0}^{N-1} (N - k) e^{ik(\theta-u)} \right) - N \right) \right] du \end{aligned}$$

Or, pour  $\theta$  réel et non congru à 0 (modulo  $2\pi$ )

$$2\operatorname{Re} \left( \sum_{k=0}^{N-1} (N - k) e^{ik\theta} \right) - N = N\Phi_{N-1}(\theta) - \frac{d}{d\theta} [\Psi_{N-1}(\theta)]$$

avec

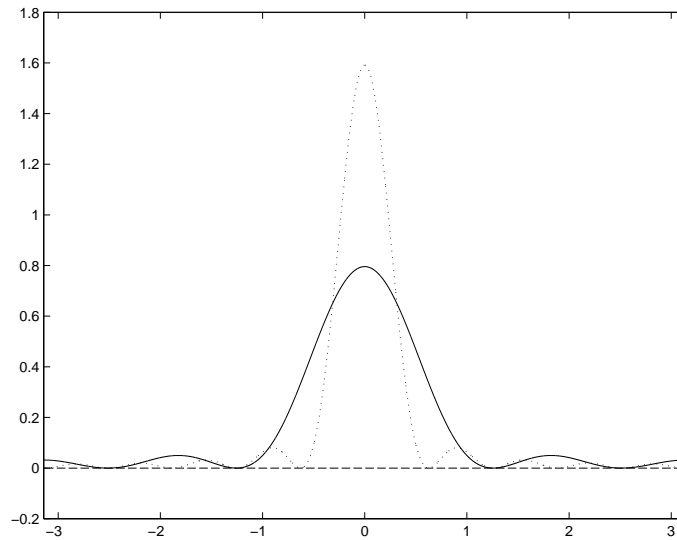
$$\Phi_{N-1}(\theta) := 2\operatorname{Re} \left[ \sum_{k=0}^{N-1} e^{ik\theta} \right] - 1 = \frac{\sin \left[ \left( N - \frac{1}{2} \right) \theta \right]}{\sin \frac{\theta}{2}}$$

(voir le calcul du noyau de Dirichlet  $D_{N-1}$  fait précédemment) et

$$\begin{aligned} \Psi_{N-1}(\theta) &:= 2\operatorname{Im} \left[ \sum_{k=0}^{N-1} e^{ik\theta} \right] = 2\operatorname{Im} \left[ \frac{1 - e^{iN\theta}}{1 - e^{i\theta}} \right] \\ &= 2\operatorname{Im} \left( e^{i\frac{(N-1)\theta}{2}} \frac{\sin \frac{N\theta}{2}}{\sin \frac{\theta}{2}} \right) \\ &= \frac{2 \sin \frac{(N-1)\theta}{2} \sin \frac{N\theta}{2}}{\sin \frac{\theta}{2}} \\ &= - \frac{\cos \left[ \left( N - \frac{1}{2} \right) \theta \right] - \cos \left( \frac{\theta}{2} \right)}{\sin \frac{\theta}{2}}. \end{aligned}$$

<sup>23</sup>H. Wilbraham l'observa en 1848 et c'est Josiah W. Gibbs (1839-1903) qui l'analysa ; ce n'est pas réellement une surprise, mais c'est, on le verra, un défaut qui entache l'opération consistant à « couper » les hautes fréquences (par exemple lors du « repiquage » de vieux enregistrements).

<sup>24</sup>Cette idée a été introduite en 1900 par le mathématicien hongrois Lipót Féjer (1880-1959).

FIG. 2.3 – Graphes sur  $[-\pi, \pi]$  de  $K_N/2\pi$ ,  $N = 5, 10$ 

On a donc, toujours pour  $\theta$  réel et non congru à 0 (modulo  $2\pi$ )

$$\begin{aligned} \frac{d}{d\theta}[\Psi_{N-1}(\theta)] &= N\Phi_{N-1}(\theta) - \frac{\sin\left(N - \frac{1}{2}\right) + \sin\frac{\theta}{2}}{2\sin\frac{\theta}{2}} \\ &\quad - \frac{\cos\frac{\theta}{2}\left(\cos\frac{\theta}{2} - \cos\left[\left(N - \frac{1}{2}\right)\theta\right]\right)}{2\sin^2\frac{\theta}{2}} \\ &= N\Phi_{N-1}(\theta) + \frac{\cos(N\theta) - 1}{2\sin^2\frac{\theta}{2}} \\ &= N\Phi_{N-1}(\theta) - \frac{\sin^2\frac{N\theta}{2}}{\sin^2\frac{\theta}{2}}. \end{aligned}$$

Finalement, tous calculs faits, on trouve

$$\dot{T}_N[f] = \dot{f} \overset{\text{per}}{*} \dot{K}_N \quad (2.25)$$

avec

$$K_N(\theta) := \sum_{k=-N}^{k=N} \left(1 - \frac{|k|}{N}\right) e^{ik\theta} = \begin{cases} \frac{1}{N} \left(\frac{\sin\frac{N\theta}{2}}{\sin\frac{\theta}{2}}\right)^2 & \text{si } \theta \not\equiv 0 \pmod{2\pi} \\ N & \text{si } \theta \equiv 0 \pmod{2\pi}. \end{cases}$$

Ce nouveau noyau  $K_N$ , dit aussi *noyau de Féjer* est toujours d'intégrale  $2\pi$  sur  $[0, 2\pi]$ , mais a cette particularité essentielle qui le différencie du noyau de Dirichlet qui est le fait d'être positif. Pour les valeurs de  $N = 5, 10$ , on a représenté sur la figure 2.3 les graphes des fonctions  $K_N/2\pi$ ; on remarque que, à valeurs de  $N$  égales, le lobe central est plus « enflé » qu'il ne l'est pour le noyau  $D_N$  de Dirichlet. Mais encore une fois, le phénomène le plus frappant est la positivité du noyau  $K_N$ . Si l'on utilise la suite  $(\dot{T}_N[f])_{N \geq 1}$  pour approcher un élément de  $L^1_{\mathbb{K}}(\mathbb{T})$  (comme l'élément  $\dot{\phi}$  défini en (2.24)), on voit cette fois que l'approximation, même si elle est plus lente, n'est

plus cette fois entachée du phénomène de Gibbs ; c'est ce que l'on voit par exemple sur la figure 2.4, où nous avons approché par la suite des classes  $(\dot{T}_N[\dot{\varphi}])_{N \geq 1}$  la classe  $\dot{\varphi}$ .

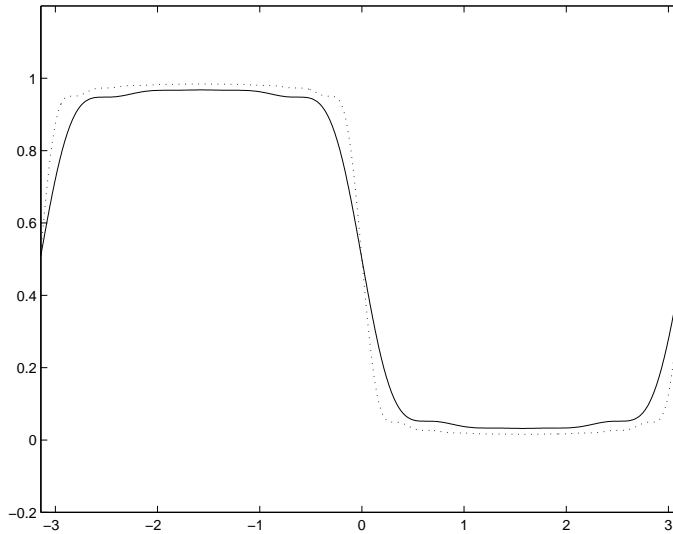


FIG. 2.4 – Graphes sur  $[-\pi, \pi]$  de  $T_N[\dot{\varphi}]$ ,  $N = 5, 10$

**Théorème 2.2 (théorème de Féjer, version  $L^1$ )** Pour tout  $\dot{f} \in L^1_{\mathbb{K}}(\mathbb{T})$ , on a

$$\lim_{N \rightarrow +\infty} \|\dot{f} - \dot{T}_N[\dot{f}]\|_{\mathbb{T},1} = 0,$$

soit encore

$$\dot{f} = \lim_{\substack{N \rightarrow +\infty \\ L^1_{\mathbb{K}}(\mathbb{T})}} \left( \theta \mapsto \sum_{k=-N}^N c_k(\dot{f}) \left(1 - \frac{|k|}{N}\right) e^{ik\theta} \right). \quad (2.26)$$

**Remarque 2.8.** Il est clair que la formule (2.26) constitue la bonne formulation pour l'inversion de  $\mathcal{TF} : L^1_{\mathbb{K}}(\mathbb{T}) \rightarrow l^{\infty,0}_{\mathbb{K}}(\mathbb{Z})$ . Contrairement au procédé taubérien en deux temps (2.19), il s'agit là d'un procédé en un temps. Pas tout à fait cependant si l'on pense à un procédé depuis les sommes  $S_N[\dot{f}]$ , puisqu'il est immédiat de constater que la  $N$ -ième somme de Féjer

$$T_N[\dot{f}] = \frac{S_0[\dot{f}] + \cdots + S_{N-1}[\dot{f}]}{N}$$

est obtenue à partir de la suite des sommes partielles de Fourier  $(S_k[\dot{f}])_{k \geq 0}$  par le procédé (lui aussi taubérien) de Césaro.

**Preuve.** On écrit, en utilisant le fait que  $K_N$  est d'intégrale  $2\pi$  sur  $[-\pi, \pi]$  et la formule (2.25), puis le théorème de Fubini-Tonelli :

$$\begin{aligned} \|\dot{f} - \dot{T}_N[\dot{f}]\|_{\mathbb{T},1} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} (f(\theta) - f(\theta - u)) K_N[\dot{f}](u) du \right| d\theta \\ &\leq \frac{1}{2\pi} \int_{-\pi}^{\pi} \|\dot{f} - \dot{f}(\cdot - u)\|_{\mathbb{T},1} |K_N(u)| du \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \|\dot{f} - \dot{f}(\cdot - u)\|_{\mathbb{T},1} K_N(u) du. \end{aligned}$$

C'est au passage de la ligne 2 à la ligne 3 qu'intervient de manière cruciale la positivité de  $K_N$  (ceci ne marcherait pas avec  $D_N$  en place de  $K_N$ ). Comme on l'a vu dans le cours de théorie de l'intégration (lemme 4.1 du polycopié de l'UE MHT512, transposé ici du cadre de  $\mathbb{R}^n$  au cadre périodique de  $\mathbb{T}^n$  avec  $n = 1$ , ce qui ne pose pas de difficulté), quelque soit  $\epsilon > 0$ , il existe  $\eta > 0$  tel que

$$\|u\| < \eta \implies \|\dot{f} - \dot{f}(\cdot - u)\|_{\mathbb{T},1} < \epsilon/2.$$

On peut donc poursuivre nos majorations :

$$\begin{aligned} \|\dot{f} - \dot{T}_N[\dot{f}]\|_{\mathbb{T},1} &\leq \frac{\epsilon}{4\pi} \int_{-\eta}^{\eta} K_N(u) du + \frac{\|\dot{f}\|_{\mathbb{T},1}}{\pi N} \int_{\eta < |u| \leq \pi} \left( \frac{\sin(Nu/2)}{\sin u/2} \right)^2 du \\ &\leq \frac{\epsilon}{4\pi} \int_{-\pi}^{\pi} K_N(u) du + \frac{\|\dot{f}\|_{\mathbb{T},1}}{N} \frac{1}{(\sin(\eta/2))^2} \\ &\leq \epsilon/2 + \frac{\|\dot{f}\|_{\mathbb{T},1}}{N} \frac{1}{(\sin(\eta/2))^2} \leq \epsilon/2 + \epsilon/2 = \epsilon \end{aligned}$$

dès que  $N$  est assez grand. Le résultat est démontré.  $\diamond$

**Remarque 2.9.** Du point de vue pratique, l'approximation de  $\dot{f}$  par les sommes de Féjer (lorsque l'on se soucie de « couper » les composantes « hautes-fréquences ») a le défaut d'être trop lente (cela tient au fait que le lobe central du graphe du noyau de Féjer  $K_N$  est trop « enflé »). Si l'on veut profiter du fait que celui du noyau de Dirichlet l'est beaucoup moins (mais que l'on veuille éviter l'écueil du phénomène de Gibbs), on imagine un noyau construit à partir de celui de Dirichlet de manière à récupérer presque la positivité sans perdre trop relativement à l'étroussure du lobe central. Les solutions proposées :

$$H_{\alpha,N}(\theta) = \alpha D_N(\theta) + \frac{1-\alpha}{2} \left( D_N\left(\theta - \frac{2\pi}{2N+1}\right) + D_N\left(\theta + \frac{2\pi}{2N+1}\right) \right)$$

avec  $\alpha = 0.54$  (Hamming<sup>25</sup>) ou  $\alpha = 0.5$  (Hanning<sup>26</sup>) sont les plus utilisées en pratique. On approche  $\dot{f}$  par les  $\dot{f} \overset{\text{per}}{*} H_{\alpha,N}$ ,  $N$  tendant vers l'infini.

Signalons au passage qu'une preuve tout à fait identique a déjà été conduite dans le cours de L2<sup>27</sup> pour obtenir le résultat suivant, dans la lignée du théorème de Jordan-Dirichlet vu précédemment (théorème 2.1) :

**Théorème 2.3 (Théorème de Féjer, version locale)** Soit  $\dot{f} \in L^1_{\mathbb{K}}(\mathbb{T})$  et  $\theta_0 \in \mathbb{R}, l \in \mathbb{K}$  deux scalaires tels que, pour un certain représentant  $f$  de  $\dot{f}$ , on ait

$$\lim_{u \rightarrow 0_+} (f(\theta_0 + u) + f(\theta_0 - u)) = 2l.$$

Alors

$$\lim_{N \rightarrow +\infty} T_N[\dot{f}](\theta_0) = l.$$

**Remarque 2.10 et rappels de L2.** Dans le cadre élémentaire du cours de L2 où l'on ne travaille que dans le contexte de l'intégration au sens de Riemann, ce résultat se répercute au niveau des fonctions  $2\pi$ -périodiques (définies cette fois partout, il ne s'agit plus de classes de fonctions) bornées et réglées, c'est à dire ayant une limite à gauche et à droite en tout point. Comme limite uniforme

<sup>25</sup>Les travaux du mathématicien américain Richard Hamming (1915-1998) portaient principalement sur l'informatique et les télécommunications.

<sup>26</sup>En référence à Julius Ferdinand von Hann (1839-1921), météorologiste autrichien, l'un des pionniers de la météorologie moderne.

<sup>27</sup>Voir le cours et le polycopié en ligne de l'UE MHT401.

de fonctions en escalier, une telle fonction est intégrable au sens de Riemann. On peut donc définir le spectre complexe  $(c_k(f))_{k \in \mathbb{Z}}$  et les sommes partielles de Fourier  $S_N[f]$  et de Féjer  $T_N(f)$ ,  $N \in \mathbb{N}$ . On a, pour tout  $\theta_0 \in \mathbb{R}$ , en adaptant la preuve du théorème 2.3 (la preuve est quasiment identique) :

$$\lim_{N \rightarrow +\infty} T_N[f](\theta_0) = \frac{f(\theta_0 + 0) + f(\theta_0 - 0)}{2}.$$

**Preuve.** On écrit, en utilisant la formule (2.25) en  $\theta_0$ , la parité de  $K_N$  et le fait que  $K_N$  soit d'intégrale  $2\pi$  sur  $[\theta_0 - \pi, \theta_0 + \pi]$ ,

$$\begin{aligned} |T_N[f](\theta_0) - l| &= \frac{1}{2\pi} \left| \int_{\theta_0 - \pi}^{\theta_0 + \pi} (f(\theta_0 - u) - l) K_N(u) du \right| \\ &= \frac{1}{2\pi} \left| \int_0^\pi (f(\theta_0 + u) + f(\theta_0 - u) - 2l) K_N(u) du \right| \\ &\leq \frac{1}{2\pi} \int_0^\eta |f(\theta_0 + u) + f(\theta_0 - u) - 2l| K_N(u) du \\ &\quad + \frac{1}{2\pi N} \int_\eta^\pi |f(\theta_0 + u) + f(\theta_0 - u) - 2l| \left( \frac{\sin(Nu/2)}{\sin(u/2)} \right)^2 du \\ &\leq \sup_{u \in [0, \eta]} |f(\theta_0 + u) + f(\theta_0 - u) - 2l| + \frac{\|f\|_{\mathbb{T}, 1} + 2|l|\pi}{2\pi \sin^2(\eta/2)} \times \frac{1}{N}. \end{aligned}$$

On conclut en fixant  $\eta$  assez petit pour que

$$\sup_{u \in [0, \eta]} |f(\theta_0 + u) + f(\theta_0 - u) - 2l| < \epsilon/2$$

( $\epsilon$  étant arbitraire), puis on choisit ensuite  $N$ , tenant compte du choix de  $\eta$ , pour que

$$\frac{\|f\|_{\mathbb{T}, 1} + 2|l|\pi}{2\pi \sin^2(\eta/2)} \times \frac{1}{N} < \epsilon/2.$$

Le théorème est ainsi démontré.  $\diamond$ .

**Remarque 2.11.** Si  $f$  est une fonction  $2\pi$ -périodique continue, la preuve ci-dessus, couplée avec le théorème de Heine<sup>28</sup> montre que

$$\sup_{\mathbb{R}} |f - T_N[f]| \mapsto 0$$

lorsque  $N$  tend vers l'infini, autrement dit, la convergence des sommes de Féjer vers  $f$  est uniforme.

## 2.4 La transformation de Fourier des classes de fonctions intégrables sur $\mathbb{R}^n$

Nous allons dans cette section définir la transformation de Fourier sur  $L^1_{\mathbb{C}}(\mathbb{R}^n, dx)$ ,  $dx$  désignant la mesure de Lebesgue. L'idée de base est toujours la même : tenter de « mesurer » quantitativement comment une classe  $\dot{f}$  de fonctions intégrables est globalement corrélée sur  $\mathbb{R}^n$  aux classes d'harmoniques spatiales

$$(x_1, \dots, x_n) \mapsto e^{i(\omega_1 x_1 + \dots + \omega_n x_n)} = e^{i\langle x, \omega \rangle}, \quad \omega \in \mathbb{R}^n$$

<sup>28</sup>Toute fonction  $2\pi$ -périodique et continue de  $\mathbb{R}$  dans  $\mathbb{C}$  est uniformément continue car on peut la voir comme une fonction continue sur le compact  $\mathbb{R}/(2\pi\mathbb{Z}) = \mathbb{T}$  (avec la métrique quotient) qui s'identifie topologiquement (et métriquement) au cercle unité de  $\mathbb{R}^2$ .

(dites harmoniques temporelles si  $n = 1$ ,  $t$  désignant dans ce cas une variable modélisant le temps). Bien sûr, ces harmoniques ne définissent pas des éléments de  $\mathcal{L}_{\mathbb{C}}^1(\mathbb{R}^n, dx)$  puisque ce sont des fonctions de module 1 qui ne sauraient donc être intégrables sur  $\mathbb{R}^n$ . Néanmoins, leur « corrélation » avec un élément  $\dot{f}$  de  $L_{\mathbb{C}}^1(\mathbb{R}^n, dx)$  (inspirée de la démarche hilbertienne que l'on a conduit par exemple dans la section 2.3) a un sens puisque le produit d'un élément  $f$  de  $\mathcal{L}_{\mathbb{C}}^1(\mathbb{R}^n, dx)$  par un élément  $e$  de  $\mathcal{L}_{\mathbb{C}}^{\infty}(\mathbb{R}^n, dx)$  est une fonction appartenant à  $\mathcal{L}_{\mathbb{C}}^1(\mathbb{R}^n, dx)$ , avec

$$\|\dot{f}e\|_1 \leq \|f\|_1 \times \|e\|_{\infty}.$$

### 2.4.1 Définition, exemples, et premières propriétés

**Définition 2.7** Soit  $\dot{f}$  une classe de fonction mesurable et intégrable sur  $\mathbb{R}^n$ , à valeurs complexes ; la transformée de Fourier (ou spectre) de  $\dot{f}$  est par définition la fonction  $\widehat{f} : \omega \in \mathbb{R}^n \mapsto \mathbb{C}$  définie ponctuellement par

$$\widehat{f}(\omega_1, \dots, \omega_n) := \int_{\mathbb{R}^n} f(x) e^{-i(x_1\omega_1 + \dots + x_n\omega_n)} dx_1 \cdots dx_n,$$

$f$  étant un représentant arbitraire de  $\dot{f}$ .

**Remarque importante 2.12.** Si la transformée de Fourier d'un élément de  $L_{\mathbb{C}}^1(\mathbb{R}_x^n, dx)$  est bien définie sur  $\mathbb{R}^n = \mathbb{R}_x^n$ , il est important de se faire à l'idée que la copie  $\mathbb{R}_\omega^n$  de  $\mathbb{R}^n$  sur laquelle est défini le spectre de  $\dot{f}$  (copie dans laquelle nous conviendrons de noter  $\omega = (\omega_1, \dots, \omega_n)$  le point courant) n'est pas la même que celle ( $\mathbb{R}_x^n$ ) sur laquelle sont définis les représentants de  $\dot{f}$  (copie où l'on note, pour bien marquer cette distinction,  $x = (x_1, \dots, x_n)$  le point courant<sup>29</sup>). Cette dernière copie  $\mathbb{R}_x^n$  est l'univers spatial sur lequel vivent les fonctions que l'on prétend traiter (on parle de « plan des images » si  $n = 2$ , d'« espace des temps » si  $n = 1$ ), tandis que la copie  $\mathbb{R}_\omega^n$  sur laquelle vivent les spectres des objets est appelée, elle, « espace des fréquences ». On verra plus loin qu'il est naturel de considérer l'espace  $\mathbb{R}_\omega^n$  des fréquences comme le dual algébrique  $(\mathbb{R}_x^n)^*$  de l'univers spatial  $\mathbb{R}_x^n$ . On verra aussi (dans la section 2.4.2) comment il est physiquement concevable de « séparer » ces deux copies dans le cas  $n = 2$  en intercalant entre elles un mécanisme optique diffractant. La dernière assertion (formule 2.28 de la proposition 2.7 à venir) constituera une première indication de la relation entre transformation de Fourier et *dualité*<sup>30</sup> en algèbre linéaire.

**Exemple 2.3 : un premier exemple basique ( $n = 1$ ).** Le spectre de la fonction caractéristique  $\chi_{[a,b]}$  d'un segment fermé borné de  $\mathbb{R}$  est la fonction

$$\omega \in \mathbb{R} \mapsto \int_a^b e^{-i\omega t} dt = \frac{1}{i\omega} (e^{-i\omega a} - e^{-i\omega b}).$$

On remarque ici que

$$\lim_{|\omega| \rightarrow 0} |\widehat{\chi}_{[a,b]}(\omega)| = 0,$$

ce qui, on le verra, s'avèrera une propriété clef du spectre d'un élément  $\dot{f}$  de  $L_{\mathbb{C}}^1(\mathbb{R}^n, dx)$  en général (voir la proposition 2.7 à venir). Lorsque  $[a, b] = [-T, T]$ , on trouve en particulier

$$\widehat{\chi}_{[-T,T]}(\omega) = \frac{2}{\omega} \sin(\omega T),$$

<sup>29</sup>Ou  $t$  lorsque  $n = 1$ .

<sup>30</sup>Voir le polycopié du cours de MIAS301, partie algèbre, pages 31-36, disponible sur le site <http://www.math.u-bordeaux1.fr/~yger/coursalg.pdf>

et l'on voit apparaître ici une fonction jouant un rôle très important en analyse de Fourier, la fonction *sinus cardinal* définie par

$$\text{sinc} : u \in \mathbb{R} \mapsto \frac{\sin(\pi u/2)}{\pi u/2}$$

(c'est en général ainsi qu'elle est normalisée, ce de manière à ce que son intégrale semi-convergente sur  $[0, +\infty[$  vaille 1). Le spectre de toutes ces fonctions  $\chi_{[a,b]}$ , notons le, n'est pas intégrable sur  $\mathbb{R}_\omega$ , ce qui s'avèrera un handicap (en relation avec la transposition au cadre continu du phénomène de Gibbs, comme on le verra dans la section 2.4.5).

**Exemple 2.4 : un second exemple très important.** Lorsque  $n = 1$ , le spectre de la gaussienne

$$g : t \mapsto \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$$

(qui est la densité de la loi normale  $\mathcal{N}(0, 1)$ ) est la gaussienne  $\omega \mapsto e^{-\omega^2/2}$ , résultat qui se répercute en  $n$  variables (*via* le théorème de Fubini) sous la forme

$$\left(x \mapsto \frac{1}{(2\pi)^{n/2}} e^{-\|x\|^2/2}\right)^\wedge = \left[\omega \mapsto e^{-\|\omega\|^2/2}\right], \quad (2.27)$$

autrement dit, si l'on pouvait superposer les deux copies de  $\mathbb{R}^n$  que sont  $\mathbb{R}_x^n$  et  $\mathbb{R}_\omega^n$ , on dirait que la gaussienne convenablement normalisée

$$g : x \in \mathbb{R}^n \mapsto (2\pi)^{-n/2} e^{-\|x\|^2/2}$$

est un vecteur propre de la prise de spectre<sup>31</sup>, ce avec la valeur propre  $(2\pi)^{n/2}$ . Faisons le calcul pour  $n = 1$ <sup>32</sup>; on remarque, en appliquant le théorème de dérivation de Lebesgue (théorème 3.2 du cours de théorie de l'intégration, MHT512), que

$$\omega \in \mathbb{R} \mapsto \widehat{g}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-t^2/2} e^{-i\omega t} dt$$

est de classe  $C^1$ , de dérivée la fonction continue

$$\begin{aligned} \omega \in \mathbb{R} \mapsto \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-t^2/2} \frac{d}{d\omega} [e^{-i\omega t}] dt &= -\frac{i}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-i\omega t} e^{-t^2/2} t dt \\ &= -\frac{i}{\sqrt{2\pi}} \left[ -e^{-t^2/2} e^{-i\omega t} \right]_{-\infty}^{\infty} - \frac{\omega}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-t^2/2} e^{-i\omega t} dt \\ &= -\omega \widehat{g}(\omega) \end{aligned}$$

(on a utilisé une intégration par parties pour passer de la ligne 1 à la ligne 2). La fonction  $\widehat{g}$  étant donc solution de l'équation différentielle linéaire homogène du premier ordre  $y'(\omega) = -\omega y(\omega)$ , on constate, en résolvant cette équation, que  $\widehat{g}(\omega) = \widehat{g}(0) e^{-\omega^2/2} = e^{-\omega^2/2}$ . On rappelle que la vérification du fait que  $\widehat{g}(0) = 1$  se fait en remarquant (grâce au théorème de Fubini) que

$$\begin{aligned} \left( \int_{\mathbb{R}} e^{-t^2/2} dt \right)^2 &= \int_{\mathbb{R}} e^{-x^2/2} dx \times \int_{\mathbb{R}} e^{-y^2/2} dy = \int \int_{\mathbb{R}^2} e^{-(x^2+y^2)/2} dx dy \\ &= 2\pi \int_0^\infty e^{-r^2/2} r dr = 2\pi \end{aligned}$$

une fois l'intégrale double exprimée en coordonnées polaires. Le résultat 2.27 est donc ainsi démontré.

<sup>31</sup>On verra plus loin, dans la section 2.5.3, l'importance de ce résultat qui explique, entre autres, pourquoi les gaussiennes sont de bons candidats mathématiques pour modéliser les particules en mécanique quantique.

<sup>32</sup>Ce calcul a déjà été fait dans le cours d'intégration et dans les guides d'activités de ce cours pour illustrer l'application du théorème de dérivation des intégrales fonction d'un paramètre, théorème 3.2 du cours de l'UE MHT 512.



**Exemple 2.5 : la distribution exponentielle symétrique.** La fonction

$$f : t \in \mathbb{R} \mapsto \frac{1}{2^n} e^{-|x_1| - \dots - |x_n|}$$

définit une densité de probabilité intéressante sur  $\mathbb{R}^n$ , la loi associée est appelée loi exponentielle symétrique. On la retrouve par exemple (lorsque  $n > 1$ ) dans les processus d'évolution spaciaux. On vérifie immédiatement (en commençant par traiter le cas  $n = 1$ ) que

$$\widehat{f}(\omega) = \prod_{j=1}^n \frac{1}{(1 + \omega_j^2)}.$$

En effet

$$\frac{1}{2} \int_{\mathbb{R}} e^{-|t|} e^{-i\omega t} dt = \frac{1}{2} \int_0^{+\infty} (e^{-(1+i\omega)t} + e^{-(1-i\omega)t}) dt = \frac{1}{2} \left( \frac{1}{1+i\omega} + \frac{1}{1-i\omega} \right) = \frac{1}{1+\omega^2}.$$

La densité de probabilités sur l'espace des fréquences  $\mathbb{R}_\omega^n$  définie par

$$\omega \mapsto \left(\frac{1}{\pi}\right)^n \prod_{j=1}^n \frac{1}{(1 + \omega_j^2)}$$

(et qui est à une constante près le spectre de la densité de probabilités d'un vecteur aléatoire suivant une loi exponentielle dans  $\mathbb{R}^n$ ) induit la loi de probabilité dite *loi de Poisson*<sup>33</sup>.

**Exemple 2.6.** Dans le cas  $n = 1$ , une loi tout aussi importante est la loi exponentielle sur  $[0, \infty[$  de paramètre  $\lambda > 0$  : il s'agit de la loi de probabilité de densité la fonction définie presque partout par

$$f_\lambda(t) := \lambda \chi_{[0, \infty[}(t) e^{-\lambda t}.$$

On la rencontre par exemple dans les processus de désintégration atomique. Le spectre de  $f_\lambda$  a pour représentant dans l'espace des fréquences la fonction

$$\omega \mapsto \lambda \int_0^{+\infty} e^{-i\omega t - \lambda t} dt = \frac{\lambda}{\lambda + i\omega};$$

c'est un spectre complexe de par la dissymétrie de  $f_\lambda$  comparée à la symétrie de la densité  $f$  introduite à l'exemple précédent.

**Proposition 2.7** *Le spectre  $\widehat{f}$  d'un élément  $f$  de  $L^1_{\mathbb{C}}(\mathbb{R}^n, dx)$  est une fonction continue sur  $\mathbb{R}^n$ , tendant vers 0 à l'infini<sup>34</sup>. L'application linéaire  $f \mapsto \widehat{f}$  est donc une application linéaire continue (de norme inférieure à 1) du  $\mathbb{C}$ -espace de Banach  $L^1_{\mathbb{C}}(\mathbb{R}^n, dx)$  dans le  $\mathbb{C}$ -espace de Banach  $\mathcal{C}(\mathbb{R}^n, \mathbb{C})_0$  des fonctions continues de  $\mathbb{R}^n$  dans  $\mathbb{C}$  tendant vers 0 à l'infini, autrement dit, on a  $\|\widehat{f}\|_{\infty} \leq \|f\|_1$ . De plus, si  $A$  est un  $\mathbb{R}$ -isomorphisme algébrique de  $\mathbb{R}^n$ , on a*

$$f(\widehat{A \cdot (\cdot)}) = \frac{1}{|\det A|} \widehat{f}(A' \cdot (\cdot)), \tag{2.28}$$

où  $A' = (A^{-1})^t$  désigne la transposée de l'inverse de  $A$ , soit la matrice duale<sup>35</sup> de  $A$ .

<sup>33</sup>Mathématicien français, Siméon-Denis Poisson (1781-1840), fut l'un des pionniers de la théorie des séries et des intégrales de Fourier. Attention toutefois à ne pas confondre la loi de Poisson à densité sur  $\mathbb{R}$  introduite ici avec la loi discrète sur  $\mathbb{N}^*$  (dite aussi *loi de Poisson de paramètre*  $\lambda > 0$ ), définie par  $P(X = n) = \lambda^n e^{-\lambda} / n!$  pour  $n \in \mathbb{N}$ .

<sup>34</sup>Cette dernière assertion est connue sous la terminologie de *lemme de Riemann-Lebesgue*; c'est la version continue du résultat établi lors de l'étude de la transformation de Fourier sur  $l^1_{\mathbb{C}}(\mathbb{Z})$  (voir la définition 2.5).

<sup>35</sup>Si  $v_1, \dots, v_n$  sont les vecteurs colonnes de  $A$  (exprimés dans la base canonique  $e_1, \dots, e_n$ ) et  $(v_1^*, \dots, v_n^*)$  la base duale de la base  $(v_1, \dots, v_n)$ , la  $j$ -ème colonne de  $A'$  représente la liste des coordonnées de  $v_j^*$  exprimé dans la base  $(e_1^*, \dots, e_n^*)$ , base duale de la base canonique  $(e_1, \dots, e_n)$  de  $\mathbb{R}^n$ .

**Preuve.** La continuité de  $\widehat{f}$  est une conséquence immédiate du théorème de Lebesgue (de continuité des intégrales dépendant d'un paramètre) vu dans le cours d'intégration (UE MHT512, théorème 3.1) puisque, si  $f$  est un représentant de  $\dot{f}$ ,

$$|f(x)e^{-i\langle x, \omega \rangle}| = |f(x)|$$

et que  $|f| \in \mathcal{L}^1_{\mathbb{R}}(\mathbb{R}^n, dx)$  définit par conséquent un chapeau majorant intégrable. L'inégalité

$$\|\widehat{f}\|_{\infty} \leq \|\dot{f}\|_1$$

est immédiate (le module de l'intégrale d'une fonction est majoré par l'intégrale du module de cette fonction). La formule (2.28) résulte de l'application du théorème de changement de variables (théorème 3.5 du cours de théorie de l'intégration) :

$$\begin{aligned} \int_{\mathbb{R}^n} f(A.x)e^{-i\langle x, \omega \rangle} dx &= \frac{1}{|\det A|} \int_{\mathbb{R}^n} f(y)e^{-i\langle A^{-1}.y, \omega \rangle} dy \\ &= \frac{1}{|\det A|} \int_{\mathbb{R}^n} f(y)e^{-i\langle y, (A^{-1})^t.\omega \rangle} dy = \frac{1}{|\det A|} \widehat{f}(A'.\omega). \end{aligned}$$

En particulier, si  $\epsilon > 0$ , la transformée de Fourier de la fonction

$$x \mapsto g_{\epsilon}(x) = \frac{1}{\sqrt{(2\pi)^n \epsilon^n}} e^{-\frac{\|x\|^2}{2\epsilon^2}}$$

est la fonction

$$\omega \mapsto \exp(-\epsilon^2\|\omega\|^2/2).$$

Cette gaussienne tend uniformément vers 0 (ainsi d'ailleurs que chacune de ses dérivées partielles à tout ordre) lorsque  $\|\omega\|$  tend vers l'infini.

Il nous reste à prouver que  $\widehat{f}(\omega)$  tend vers 0 lorsque  $\|\omega\|$  tend vers  $+\infty$  (propriété dite de *Riemann-Lebesgue*). Fixons  $\eta > 0$  et une suite  $(\epsilon_k)_{k \geq 1}$  de nombres strictement positifs tendant vers 0. Puisque la famille  $(g_{\epsilon_k})_{k \geq 1}$  réalise une approximation de la masse de Dirac dans  $L^1_{\mathbb{R}}(\mathbb{R}^n, dx)$ , on sait<sup>36</sup> que, pour  $k$  assez grand

$$\|\widehat{f} - \widehat{f}_{\epsilon_k}\|_{\infty} \leq \|\dot{f} - \dot{f}_{\epsilon_k}\|_1 \leq \eta/2, \quad (2.29)$$

où  $\dot{f}_{\epsilon_k}$  a pour représentant la fonction

$$x \mapsto \int_{\mathbb{R}^n} f(x-y)g_{\epsilon_k}(y) dy.$$

Fixons  $k \in \mathbb{N}^*$  pour qu'il en soit ainsi. On peut calculer le spectre de  $\dot{f}_{\epsilon_k}$  en remarquant

$$\begin{aligned} &\int_{\mathbb{R}^n} \left( \int_{\mathbb{R}^n} g_{\epsilon_k}(t)f(x-t) dt \right) e^{-i\langle x, \omega \rangle} dx = \iint_{\mathbb{R}^n \times \mathbb{R}^n} g_{\epsilon_k}(t)f(x-t)e^{-i\langle x, \omega \rangle} dx dt \\ &= \iint_{\mathbb{R}^n \times \mathbb{R}^n} g_{\epsilon_k}(t)f(x-t)e^{-i\langle t, \omega \rangle} e^{-i\langle x-t, \omega \rangle} dx dt \\ &= \int_{\mathbb{R}^n} g_{\epsilon_k}(t)e^{-i\langle t, \omega \rangle} \left( \int_{\mathbb{R}^n} f(x-t)e^{-i\langle x-t, \omega \rangle} dx \right) dt \\ &= \int_{\mathbb{R}^n} g_{\epsilon_k}(t)e^{-i\langle t, \omega \rangle} \left( \int_{\mathbb{R}^n} f(x)e^{-i\langle x, \omega \rangle} dx \right) dt = \widehat{f}(\omega) \widehat{g_{\epsilon_k}}(\omega) = \widehat{f}(\omega) e^{-\epsilon_k^2\|\omega\|^2/2} \end{aligned} \quad (2.30)$$

<sup>36</sup>Voir la proposition 4.10 du cours de théorie de l'intégration, UE MHT512; ce résultat a d'ailleurs été rappelé (dans le contexte  $L^2$ , ici nous l'invoquons dans le contexte  $L^1$ ) dans la section 1.9.1 du chapitre 1 du présent cours.

(la première égalité résulte de l'application de Fubini, la seconde de la propriété clef de l'exponentielle échangeant addition et multiplication, puis vient finalement Fubini une nouvelle fois et l'invariance de la mesure de Lebesgue par translation sur  $\mathbb{R}^n$ )<sup>37</sup>. On a donc

$$\|\widehat{f_\epsilon}(\omega)\| \leq \|f\|_1 e^{-\epsilon^2 \|\omega\|^2/2},$$

d'où il résulte qu'il existe,  $k$  étant ainsi choisi, un seuil  $\Omega(\eta, k)$  tel que

$$\|\omega\| \geq \Omega(\eta, k) \implies |\widehat{f_{\epsilon_k}}(\omega)| \leq \eta/2.$$

En combinant avec (2.29), on a donc

$$\|\omega\| \geq \Omega(\eta, k) \implies |\widehat{f}(\omega)| \leq \|\dot{f} - \dot{f_{\epsilon_k}}\|_1 + |\widehat{f_{\epsilon_k}}(\omega)| \leq \eta/2 + \eta/2 = \eta.$$

Comme  $\eta$  avait été choisi arbitrairement (puis  $k$  fonction de  $\eta$ ), on a bien

$$\lim_{\|\omega\| \rightarrow +\infty} |\widehat{f}(\omega)| = 0$$

et la proposition est ainsi complètement démontrée.  $\diamond$

**Remarque 2.13.** Nous avons choisi ici, par souci de cohérence avec la trame du cours, où les gaussiennes sont appelées à jouer un rôle important, de démontrer la propriété de Riemann Lebesgue en utilisant la régularisation par une approximation de la masse de Dirac du type  $(g_{\epsilon_k})_{k \geq 1}$ . Nous aurions pu aussi remarquer que la propriété de Riemann-Lebesgue (le fait que le spectre de  $\dot{f}$  tende vers 0 à l'infini) est satisfaite lorsque  $f$  est la fonction indicatrice<sup>38</sup> d'un pavé fermé borné  $[a_1, b_1] \times \dots \times [a_n, b_n]$  (voir l'exemple 2.3). On sait que les classes des combinaisons linéaires finies à coefficients complexes de telles fonctions indicatrices constituent un sous-espace dense  $\mathcal{E}$  dans  $L^1_{\mathbb{C}}(\mathbb{R}^n, dx)$ . En utilisant le même schéma de preuve que ci-dessus ( $\dot{f_{\epsilon_k}}$  dans (2.29) étant remplacé par une approximation  $\dot{f}_\eta$  précisément dans  $\mathcal{E}$ ), on prouverait également le résultat.

**Exemple d'application 4.7.** En appliquant la formule (2.28), on voit que si  $m \in \mathbb{R}^n$  et  $\sigma \in ]0, +\infty[^n$  le spectre de

$$(x_1, \dots, x_n) \mapsto \frac{1}{(2\pi)^{n/2}} \frac{1}{\sigma_1 \dots \sigma_n} \exp\left(-\frac{1}{2} \sum_{j=1}^n \frac{|x_j - m_j|^2}{\sigma_j^2}\right) \quad (2.31)$$

(loi d'un vecteur de variables aléatoires indépendantes  $X_j$ , avec  $X_j$  de loi  $\mathcal{N}(m_j, \sigma_j)$ ,  $j = 1, \dots, n$ ) est

$$(\omega_1, \dots, \omega_n) \mapsto \exp\left(i\langle m, \omega \rangle - \frac{1}{2} \sum_{j=1}^n \sigma_j^2 \omega_j^2\right). \quad (2.32)$$

On constate d'ailleurs sur l'exemple précédent un fait tout à fait général : si  $\dot{f}$  est un élément de  $L^1_{\mathbb{C}}(\mathbb{R}^n, dx)$  et  $m$  un vecteur de  $\mathbb{R}^n$ , alors le spectre de la classe translatée  $\dot{f}(\cdot - m)$  est

$$\left[\dot{f}(\cdot - m)\right]^{\wedge}(\omega) = e^{i\langle m, \omega \rangle} \widehat{f}(\omega), \quad \forall \omega \in \mathbb{R}^n. \quad (2.33)$$

<sup>37</sup>Le calcul fait ici est un cas particulier d'un calcul plus général que nous ferons une fois plus loin, en nous souvenant que  $\dot{f_{\epsilon_k}} = \dot{f} * \dot{g_{\epsilon_k}}$ .

<sup>38</sup>On préfère utiliser ici l'épithète « *indicatrice* » plutôt que « *caractéristique* » (souvent utilisée dans le cours d'intégration) pour éviter la confusion avec la notion de *fonction caractéristique* d'un vecteur de variables aléatoires réelles (voir plus loin).

L'effet d'une translation de  $m$  dans l'espace  $\mathbb{R}_x^n$  se traduit donc (dans l'espace  $\mathbb{R}_\omega^n$ ) par une modulation du spectre  $\widehat{f}$  par la fonction oscillante

$$\omega \longmapsto e^{i\langle m, \omega \rangle}.$$

Si maintenant inversement  $\dot{f}$  est modulée par multiplication par la fonction oscillante

$$x \in \mathbb{R}^n \longmapsto e^{i\langle x, m \rangle}$$

alors le spectre de  $\dot{f}$  est translaté de  $m$  dans l'espace  $\mathbb{R}_\omega^n$ , ce que l'on résume en écrivant

$$\left[ \dot{f} e^{i\langle x, m \rangle} \right]^\wedge(\omega) = \widehat{f}(\omega - m), \quad \forall \omega \in \mathbb{R}^n. \quad (2.34)$$

**Une incursion du côté des probabilités.** Puisque nous avons ici (en particulier dans notre galerie d'exemples 2.3 à 2.7) fait souvent référence aux probabilités, il est important de faire ici le lien avec une notion introduite dans le cours de théorie de l'intégration (section 3.2.3 du cours de MHT512) ainsi que dans le cours de MHT601 : celle de *fonction caractéristique* d'un vecteur de variables aléatoires réelles  $X = (X_1, \dots, X_n)$  définie sur un espace probabilisé  $(\Omega, \mathcal{T}, P)$  ; il s'agit, rappelons le, de la fonction continue  $\Phi_X : \mathbb{R}^n \longrightarrow \mathbb{C}$  définie par <sup>39</sup>

$$\Phi_X(\tau_1, \dots, \tau_n) := \int_{\Omega} e^{i\langle \tau, X(\omega) \rangle} dP(\omega) = E[e^{i\langle \tau, X \rangle}].$$

On remarque que, si la loi de  $X$  est une loi à densité  $\dot{f} \in L^1_{\mathbb{R}}(\mathbb{R}^n, dx)$ , on a

$$\overline{\Phi_X(\tau)} = \int_{\Omega} e^{-i\langle \tau, X(\omega) \rangle} dP(\omega) = \int_{\mathbb{R}^n} \dot{f}(x) e^{-i\langle \tau, x \rangle} dx = \widehat{\dot{f}}(\tau),$$

d'où l'intime relation entre les notions de transformation de Fourier sur  $L^1$  et de prise de fonction caractéristique d'un vecteur de variables aléatoires suivant une loi à densité sur  $\mathbb{R}^n$ . Le concept de fonction caractéristique s'étend cependant aux vecteurs de variables aléatoires de loi quelconque et non plus à densité, ce qui élargit le cadre de la transformation de Fourier (même si ici elle n'est envisagée que pour des classes  $\dot{f}$  de fonctions positives d'intégrale 1, comme c'est le cas pour toute densité de probabilité). Notons deux règles de calcul importantes : si  $X_1, \dots, X_n$  sont des variables aléatoires réelles mutuellement indépendantes, alors

$$\forall \tau \in \mathbb{R}^n, \quad \Phi_X(\tau_1, \dots, \tau_n) = \prod_{j=1}^n \Phi_{X_j}(\tau_j) \quad (2.35)$$

et

$$\forall \tau \in \mathbb{R}, \forall \lambda \in \mathbb{R}^n, \quad \Phi_{\lambda_1 X_1 + \dots + \lambda_n X_n}(\tau) = \prod_{j=1}^n \Phi_{X_j}(\lambda_j \tau), \quad (2.36)$$

toutes les deux conséquences de la mutuelle indépendance des  $X_j$  et du fait que l'exponentielle complexe réalise un homomorphisme entre  $(\mathbb{C}, +)$  et  $(\mathbb{C}^*, \times)$  :

$$\forall z, w \in \mathbb{C}, \quad \exp(z + w) = \exp z \times \exp w,$$

propriété algébrique dont on a très souvent dans ce chapitre déjà eu l'occasion de souligner l'intérêt majeur.

<sup>39</sup>Attention ici ! on ne peut plus utiliser  $\omega$  comme variable sur l'espace des « fréquences » puisque  $\omega$  est ici réservé à l'indexation des événements dans l'ensemble abstrait  $\Omega$ .

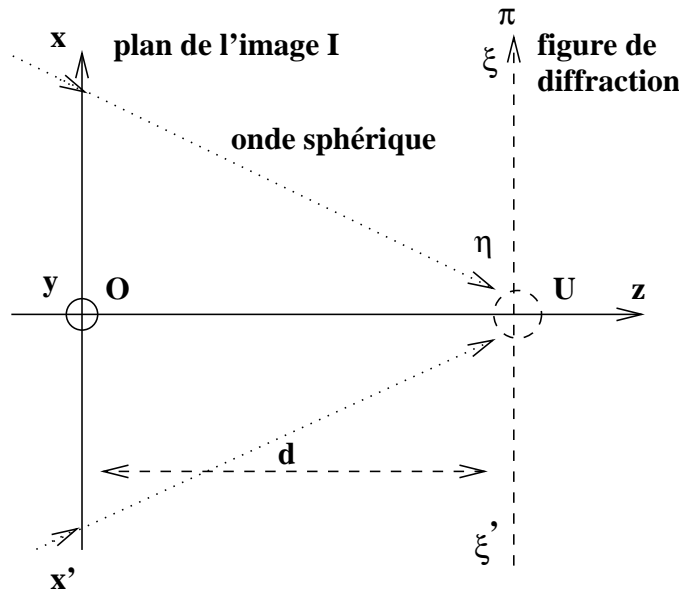


FIG. 2.5 – Réalisation de la transformation de Fourier optique

### 2.4.2 La relation avec l'optique

C'est un dispositif optique qui (lorsque  $n = 2$ ) permet de réaliser comme une transformation physique la transformation de Fourier ; c'est en effet la *diffraction de Fraunhofer*<sup>40</sup> qui réalise de manière optique la transformation d'une image 2D en son spectre. Utilisons donc ici un instant le langage des physiciens. Supposons qu'une onde sphérique monochromatique ( $\lambda$ ),

$$(t, x, y, z) \mapsto a \exp \left( -i\omega t - ik \frac{x^2 + y^2}{2d} \right)$$

(avec *nombre d'onde*  $k = 2\pi/\lambda$ ), convergeant à la distance  $d$  du plan diffractant  $\pi = xOy = \{z = 0\}$  (voir la figure 2.5), éclaire un objet qui se trouve dans ce plan diffractant et dont la transmittance en amplitude réalise une distribution d'image  $(x, y) \mapsto I(x, y)$  ; alors (sous réserve que l'on puisse se placer dans le contexte où est valide la règle d'approximation dite « des petits angles »), l'amplitude diffractée dans le plan de convergence  $\xi U \eta$  vaut

$$a_{\text{diffract}}(\xi, \eta) = \frac{ka}{2\pi id} \exp \left( ikd - i\omega t + ik \frac{\xi^2 + \eta^2}{2d} \right) \widehat{I}(\xi k/d, \eta k/d). \quad (2.37)$$

Cette formule montre donc que le mécanisme de diffraction de Fraunhofer réalise optiquement la prise de spectre d'une image, à savoir ici l'image  $I$  dont le spectre se lit à l'aide de la formule (2.37) à partir de l'amplitude lumineuse diffractée  $a_{\text{diffract}}$  (on constate en particulier que le module de  $\widehat{I}$  correspond en particulier au module de l'amplitude diffractée). Cette interprétation « physique » de la transformation de Fourier expliquera le principe de conservation d'énergie (que matérialisera la formule de Plancherel) que nous énoncerons dans la section 2.5.2.

<sup>40</sup>Mise en évidence et étudiée par Joseph von Fraunhofer, opticien allemand (1787-1828).

### 2.4.3 Fourier et la convolution dans $L_{\mathbb{K}}^1(\mathbb{R}^n, dx)$

On rappelle<sup>41</sup> que la convolée  $\dot{f} * \dot{g}$  de deux éléments de  $L_{\mathbb{K}}^1(\mathbb{R}^n, dx)$  ( $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$ ) est définie comme l'élément de  $L_{\mathbb{K}}^1(\mathbb{R}^n)$  dont un représentant est la fonction définie  $dx$ -presque partout par

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

l'opération de convolution définie ainsi est une opération interne sur  $L_{\mathbb{K}}^1(\mathbb{R}^n, dx)$ , associative, commutative, mais sans élément neutre<sup>42</sup>.

Certes aucun élément de  $L_{\mathbb{K}}^1(\mathbb{R}^n, dx)$  ne vient remplacer l'élément  $e_{0, \dots, 0}$  de la base canonique de  $l_{\mathbb{K}}^1(\mathbb{Z}^n)$  et l'on ne peut (comme on aurait envie de le faire) illustrer l'omniprésence de l'opération de convolution dans le cadre continu aussi bien que comme on l'a fait dans le cadre discret<sup>43</sup> : la translation  $\dot{f} \rightarrow \dot{f}(\cdot - x_0)$  est par exemple un opérateur continu de  $L_{\mathbb{K}}^1(\mathbb{R}^n, dx)$  dans lui-même, invariant par translation, mais ne se représente pourtant pas comme un opérateur de la forme

$$\dot{f} \mapsto \dot{f} * \dot{h}$$

avec  $\dot{h} \in L_{\mathbb{K}}^1(\mathbb{R}^n, dx)$ . Le cadre  $L^1$  (ou même  $L^2$ , le problème est en fait le même), lorsque l'on travaille en continu et non plus en discret, s'avère en effet être un cadre trop restreint pour que l'opération de convolution puisse réellement pendre toute sa signification. C'est la théorie des *distributions* (voir les cours ultérieurs de master 1) qui permettra d'élargir ce cadre, de donner à cette opération tout son sens et de comprendre par là même la raison de son omniprésence en ingénierie (lorsque l'on travaille avec des appareils agissant de manière linéaire, continue en un certain sens, et surtout présentant des caractéristiques immuables en temps ou en espace). Il est cependant bien utile à ce niveau de déjà suggérer cet aspect « pratique » que revêt l'opération de convolution en ingénierie.

La prise de spectre persiste dans le cadre  $L^1$  (comme elle le faisait dans le cadre fini, discret ou périodique, voir les propositions 2.2, 2.3, 2.4, 2.5) à échanger les opérations de convolution et de multiplication; on a en effet la :

**Proposition 2.8** Soit  $\dot{f}_1$  et  $\dot{f}_2$  deux éléments de  $L_{\mathbb{C}}^1(\mathbb{R}^n, dx)$ ; on a, pour tout  $\omega$  dans  $\mathbb{R}^n$ , la formule

$$[\dot{f}_1 * \dot{f}_2]^{\wedge}(\omega) = \hat{f}_1(\omega) \times \hat{f}_2(\omega).$$

**Preuve.** Grâce au théorème de Fubini (on vérifie la clause de sécurité car les représentants  $f_1$  et  $f_2$  choisis sont intégrables sur  $\mathbb{R}^n$ ) et à l'invariance de la mesure de Lebesgue par translation, on a, pour  $\omega$  fixé dans  $\mathbb{R}^n$  :

$$\begin{aligned} & \int_{\mathbb{R}^n} \left( \int_{\mathbb{R}^n} f_1(y)f_2(x-y) dy \right) e^{-i\langle x, \omega \rangle} dx = \int \int_{\mathbb{R}^n \times \mathbb{R}^n} f_1(y)f_2(x-y)e^{-i\langle x, \omega \rangle} dx dy \\ & = \int \int_{\mathbb{R}^n \times \mathbb{R}^n} f_1(y)f_2(x-y)e^{-i\langle y, \omega \rangle} e^{-i\langle x-y, \omega \rangle} dx dy \end{aligned}$$

<sup>41</sup>Voir l'énoncé du théorème de Young (théorème 4.2 du cours de théorie de l'intégration MHT 512).

<sup>42</sup>Voir la troisième séquence du guide d'activité 10 proposé sous Ulysse pour le cours de MHT512.

<sup>43</sup>Voir la section 4.5.1 du cours de théorie de l'intégration.

$$\begin{aligned} &= \int_{\mathbb{R}^n} f_1(y) e^{-i\langle y, \omega \rangle} \left( \int_{\mathbb{R}^n} f_2(x-y) e^{-i\langle x-y, \omega \rangle} dx \right) dy \\ &= \int_{\mathbb{R}^n} f_1(y) e^{-i\langle y, \omega \rangle} \left( \int_{\mathbb{R}^n} f_2(x) e^{-i\langle x, \omega \rangle} dx \right) dy = \widehat{f}_1(\omega) \widehat{f}_2(\omega). \end{aligned}$$

La proposition est donc ainsi démontrée.  $\diamond$ .

**Exemple d'application 2.8.** Considérons deux fonctions de la forme

$$\begin{aligned} \varphi &: x \in \mathbb{R}^n \mapsto \frac{1}{(2\pi)^{n/2}} \frac{1}{\sigma_1 \cdots \sigma_n} \exp\left(-\frac{1}{2} \sum_{j=1}^n \frac{|x_j - m_j|^2}{\sigma_j^2}\right) \\ \tilde{\varphi} &: x \in \mathbb{R}^n \mapsto \frac{1}{(2\pi)^{n/2}} \frac{1}{\tilde{\sigma}_1 \cdots \tilde{\sigma}_n} \exp\left(-\frac{1}{2} \sum_{j=1}^n \frac{|x_j - \tilde{m}_j|^2}{\tilde{\sigma}_j^2}\right), \end{aligned}$$

où  $m, \tilde{m}$  sont des éléments de  $\mathbb{R}^n$  et  $\sigma, \tilde{\sigma}$  des éléments de  $]0, +\infty[^n$ . En passant (aller-retour) par la transformation de Fourier (et en utilisant les calculs faits pour passer de (2.31) à (2.32) *via* Fourier), on observe que

$$\varphi * \tilde{\varphi} : x \in \mathbb{R}^n \mapsto \frac{1}{(2\pi)^{n/2}} \frac{1}{\Sigma_1 \cdots \Sigma_n} \exp\left(-\frac{1}{2} \sum_{j=1}^n \frac{|x_j - M_j|^2}{\Sigma_j^2}\right)$$

avec

$$\begin{aligned} M &= m + \tilde{m} \\ \Sigma_j &= \sqrt{\sigma_j^2 + \tilde{\sigma}_j^2}, \quad j = 1, \dots, n. \end{aligned}$$

Ceci montre en particulier que, si  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$ , le  $\mathbb{K}$ -sous-espace vectoriel  $G$ <sup>44</sup> de  $L^1_{\mathbb{K}}(\mathbb{R}^n, dx)$  engendré (au sens algébrique) par les classes des gaussiennes  $g_\epsilon(\cdot - x_0)$ ,  $\epsilon > 0$ ,  $x_0 \in \mathbb{R}^n$ , est stable par convolution. Remarquons aussi que, lorsque  $\mathbb{K} = \mathbb{C}$ , l'image de  $G$  par la transformation de Fourier est un  $\mathbb{C}$ -sous-espace  $\widehat{G}$  de  $L^1_{\mathbb{C}}(\mathbb{R}^n_\omega, d\omega)$ . Le sous-espace  $\mathcal{G}$ <sup>45</sup> engendré par les fonctions

$$x \in \mathbb{R}^n \mapsto e^{i\langle x, \omega \rangle} g_\epsilon(x - x_0), \quad x_0, \omega \in \mathbb{R}^n, \quad \epsilon > 0$$

(et contenant donc  $G$ ) est, lui, stable par prise de spectre (et aussi en fait par convolution, vérifiez le en exercice) ; il contient à la fois  $G$  et  $\widehat{G}$  si l'on décide d'identifier les deux copies de  $\mathbb{R}^n$  en jeu dans la transformation de Fourier. C'est d'ailleurs le plus petit sous-espace de  $L^1_{\mathbb{C}}(\mathbb{R}^n_x \simeq \mathbb{R}^n_\omega, dx \simeq d\omega)$  contenant à la fois  $G$  et  $\widehat{G}$ .

### 2.4.4 La formule d'inversion dans le cadre $L^1$

L'un des résultats majeurs (et assez inattendu) de la théorie de Fourier au niveau  $L^1$  est la *formule d'inversion* qui nous assure que tout élément de  $L^1_{\mathbb{C}}(\mathbb{R}^n, dx)$  peut s'écrire comme un « empilement » de fonctions oscillantes

$$x \mapsto e^{i\langle x, \omega \rangle}, \quad \omega \in \mathbb{R}^n.$$

Pourquoi inattendu ? Aucune de ces fonctions oscillantes (sauf la fonction nulle), ne définit en effet une classe de fonctions intégrable ; comment, en « empilant »

<sup>44</sup>La terminologie que l'on a choisi ici pour qualifier ce sous-espace particulièrement intéressant (car engendré par les classes des versions translatées, comprimées ou dilatées, de la gaussienne de référence  $g$ ) fait référence (ce n'est pas une surprise) à Carl Friedrich Gauss (1777-1855).

<sup>45</sup>La terminologie choisie ici pour qualifier ce sous-espace (engendré cette fois par les classes des versions translatées, comprimées ou dilatées, mais cette fois aussi modulées, de la gaussienne de référence  $g$ ) fait référence aux travaux du physicien d'origine hongroise Dennis Gabor (1900-1979), prix Nobel de Physique en 1971, à qui l'on doit l'introduction d'un tel « dictionnaire » d'atomes.

astucieusement de telles fonctions, espérer récupérer toutes les fonctions intégrables ? Tout est bien sûr de fait caché dans l'« empilement » qui permet, lui, les phénomènes de cancellation (ou de *battements* dans le langage des physiciens, employé surtout en dimension  $n = 1$ ). On retrouve dans cette formule d'inversion la grande idée de Fourier, suivant laquelle tout phénomène physique s'écrirait comme un empilement de phénomènes ondulatoires élémentaires.

**Théorème 2.4 (Formule d'inversion dans le cadre  $L^1$ )** Soit  $\dot{f}$  un élément de  $L^1_{\mathbb{C}}(\mathbb{R}^n, dx)$ , tel que  $\widehat{f}$  soit une fonction intégrable sur l'espace des fréquences  $\mathbb{R}^n_{\omega}$ . Alors  $\dot{f}$  admet pour représentant la fonction continue :

$$x \in \mathbb{R}^n \longmapsto \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \widehat{f}(\omega) e^{i\langle x, \omega \rangle} d\omega.$$

**Preuve.** Le point de départ est la proposition 4.10 du cours d'intégration MHT 512 qui assure que si  $(\epsilon_k)_{k \geq 1}$  est une suite de nombres strictement positifs tendant vers 0, alors

$$\lim_{k \rightarrow \infty} \dot{f} * \dot{g}_{\epsilon_k} = \dot{f}.$$

Confondons un instant les deux copies de  $\mathbb{R}^n$  que sont  $\mathbb{R}^n_x$  et  $\mathbb{R}^n_{\omega}$ . Pour  $\epsilon > 0$ , la fonction

$$x \simeq \omega \longmapsto g_{\epsilon}(x)$$

est en fait le spectre de la classe de fonction

$$x \simeq \omega \longmapsto \frac{1}{(2\pi)^n} e^{-\epsilon^2 \|x\|^2/2} = h_{\epsilon}(x)$$

(il suffit de considérer cette classe sur le modèle (2.31) et de calculer son spectre sur le modèle (2.32)). Fixons  $k \in \mathbb{N}^*$ . On peut donc écrire, en utilisant Fubini (la clause de sécurité est remplie car  $\dot{f}$  et  $\dot{h}_{\epsilon_k}$  sont dans  $L^1_{\mathbb{R}}(\mathbb{R}^n, dx)$ ) et le fait que  $g_{\epsilon}(x - y) = g_{\epsilon}(y - x)$  :

$$\begin{aligned} f * g_{\epsilon_k}(x) &= \int_{\mathbb{R}^n} f(y) \left( \int_{\mathbb{R}^n} h_{\epsilon_k}(\omega) e^{-i\langle \omega, y-x \rangle} d\omega \right) dy \\ &= \iint_{\mathbb{R}^n \times \mathbb{R}^n} f(y) h_{\epsilon_k}(\omega) e^{-i\langle \omega, y-x \rangle} dy d\omega \\ &= \int_{\mathbb{R}^n} e^{i\langle x, \omega \rangle} h_{\epsilon_k}(\omega) \left( \int_{\mathbb{R}^n} f(y) e^{-i\langle y, \omega \rangle} dy \right) d\omega \\ &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i\langle x, \omega \rangle} e^{-\epsilon_k^2 \|\omega\|^2/2} \widehat{f}(\omega) d\omega. \end{aligned}$$

Si  $k$  tend vers l'infini, le théorème de convergence dominée de Lebesgue (applicable ici puisque  $\widehat{f}$  est supposée intégrable sur  $\mathbb{R}^n_{\omega}$ ) implique

$$\lim_{k \rightarrow +\infty} \left( \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i\langle x, \omega \rangle} e^{-\epsilon_k^2 \|\omega\|^2/2} \widehat{f}(\omega) d\omega \right) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \widehat{f}(\omega) e^{i\langle x, \omega \rangle} d\omega.$$

Comme l'on sait d'autre part que  $\dot{f} * \dot{g}_{\epsilon_k}$  converge vers  $\dot{f}$  dans  $L^1_{\mathbb{R}}(\mathbb{R}^n, dx)$  et que ceci implique donc<sup>46</sup> l'existence d'une suite  $(f * g_{\epsilon_{k_p}})_{p \geq 0}$  convergeant en presque tout  $x$

<sup>46</sup>Voir la proposition 4.6 du cours de théorie de l'intégration MHT 512, avatar essentiel de la démonstration du théorème de Riesz-Fischer.



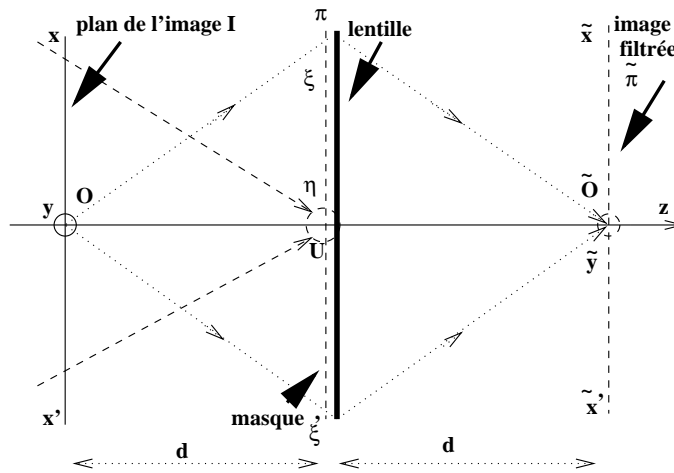


FIG. 2.6 – Expérience de Abbe

vers  $f(x)$ , où  $f$  est un représentant de  $\hat{f}$ , la conclusion du théorème est donc bien démontrée.  $\diamond$

### Une nouvelle incursion du côté de l'optique : l'expérience de Abbe.

Voici une nouvelle incursion du côté de l'optique, à la lumière de la formule d'inversion du théorème 2.4. C'est l'expérience de la *double diffraction* (voir la figure 2.6), dite aussi *expérience de Abbe*<sup>47</sup> qui permet elle la réalisation optique de la convolution comme opération optique. On suppose que le plan  $\pi = \xi U \eta$  dans lequel se forme la première figure de diffraction de l'image  $I$  ( $I$  étant une image vivante dans le plan des images  $xOy$ ) est le plan d'un masque de filtrage dont la transmittance en amplitude est  $\hat{h}(k\xi/d, k\eta/d)$ , où  $h$  est une autre distribution d'image dans le plan  $\pi$ . L'effet du masque est de transformer l'amplitude lumineuse *diffractée*

$$a_{\text{diffract}}(\xi, \eta) = \frac{ka}{2\pi id} \exp\left(ikd - i\omega t + ik\frac{\xi^2 + \eta^2}{2d}\right) \hat{I}(\xi k/d, \eta k/d)$$

en l'amplitude *diffractée-masquée*

$$a_{\text{diffract}}(\xi, \eta) \times \hat{h}(k\xi/d, k\eta/d).$$

Comme le spectre de la distribution d'image

$$(\xi, \eta) \mapsto \hat{I}(k\xi/d, k\eta/d) \hat{h}(k\xi/d, k\eta/d)$$

est formellement (d'après le théorème 2.4) donné en les variables  $(x, y)$  par

$$(x, y) \mapsto \frac{4\pi^2 d^2}{k^2} [I * h](-dx/k, -dy/k),$$

l'image obtenue dans le plan  $\tilde{x}\tilde{O}\tilde{y}$  de la figure 2.6 après la seconde diffraction est

$$(\tilde{x}, \tilde{y}) \mapsto a \exp\left(2ikd - i\omega t + ik\frac{\tilde{x}^2 + \tilde{y}^2}{2d}\right) [I * h](\tilde{x}, \tilde{y})$$

<sup>47</sup>Ernst Abbe (1840-1905) est un mathématicien et physicien allemand à qui l'on doit le schéma de cette expérience.

et l'on réalise bien via ce procédé de double diffraction la convolution de manière optique. L'idée est toujours de prendre les spectres des deux êtres  $I$  et  $h$  à convoler (c'est le but de la première diffraction), de multiplier terme à terme ces spectres (c'est l'effet du « masque » correspondant à  $\widehat{h}$  dans le plan  $\pi$ ), puis d'utiliser la formule d'inversion de Fourier, ce qui revient à effectuer une seconde transformée de Fourier, donc une seconde diffraction, pour retrouver l'antécédent, en l'occurrence  $I * h$ . La proposition 2.8 soutend bien sûr cette démarche.

### Un avatar en probabilités : le théorème « limite central ».

Concluons cette section par une nouvelle incursion vers la théorie des probabilités. Supposons que  $(X_k)_{k \in \mathbb{N}^*}$  soit une suite de variables aléatoires réelles, mutuellement indépendantes, toutes définies sur le même espace probabilisé  $(\Omega, \mathcal{T}, P)$ , toutes de même loi, avec  $E[X^2] = 1$  et  $E[X] = 0$  si  $X := X_1$ . Si l'on calcule la fonction caractéristique de

$$Y_k := \frac{X_1 + \cdots + X_k}{\sqrt{k}}$$

(notons que l'écart type de  $Y_k$  est ainsi normalisé égal à 1), on constate, du fait de l'indépendance mutuelle des  $X_k$  et donc de la formule (2.36), que

$$\Phi_{Y_k}(\tau) = [\Phi_{X/\sqrt{k}}(\tau)]^k.$$

Or

$$\Phi_{X/\sqrt{k}}(\tau) = \int_{\Omega} e^{i\tau X(\omega)/\sqrt{k}} dP(\omega) = \Phi_X(\tau/\sqrt{k}).$$

D'après le théorème de dérivation des intégrales dépendant d'un paramètre (appliqué deux fois), le fait que le moment d'ordre deux  $E[X^2]$  soit fini implique que la fonction  $\Phi_X$  est deux fois différentiable en  $\tau = 0$  et que l'on peut écrire, grâce à la formule de Taylor-Young,

$$\Phi_X(\tau/\sqrt{k}) = 1 - \frac{t^2}{2k} E[X^2] + o((\tau/\sqrt{k})^2) = 1 - \frac{t^2}{2k} + o((\tau/\sqrt{k})^2),$$

d'où l'on déduit immédiatement

$$\lim_{k \rightarrow +\infty} \Phi_{Y_k}(\tau) = \lim_{k \rightarrow +\infty} \left( 1 - \frac{t^2}{2k} + o((\tau/\sqrt{k})^2) \right)^k = \exp(-\tau^2/2) = \Phi_{\mathcal{N}(0,1)}(\tau),$$

où  $\mathcal{N}(0,1)$  désigne n'importe quelle variable aléatoire réelle sur  $(\Omega, \mathcal{T}, P)$  suivant une loi normale (de densité la gaussienne  $g$ ). Si maintenant  $\varphi$  est un élément de  $\mathcal{L}_{\mathbb{R}}^1(\mathbb{R}, dt)$  tel que le spectre de  $\varphi$  soit intégrable<sup>48</sup>, on peut écrire, d'après le théorème 2.4 et le théorème de Fubini, que pour tout  $k \in \mathbb{N}$  ( $P_{Y_k}$  désignant la distribution de probabilité de la variable  $Y_k$  sur  $\mathbb{R}$ ),

$$\begin{aligned} \int_{\Omega} \varphi(Y_k) dP &= \int_{\mathbb{R}} \varphi(t) dP_{Y_k}(t) = \int_{\mathbb{R}} \left[ \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{\varphi}(\tau) e^{i\tau t} d\tau \right] dP_{Y_k}(t) \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{\varphi}(\tau) \left[ \int_{\Omega} e^{i\tau t} dP_{Y_k}(t) \right] d\tau \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{\varphi}(\tau) \Phi_{Y_k}(\tau) d\tau \end{aligned}$$

<sup>48</sup>D'après la formule d'inversion du théorème 2.4, on peut donc considérer  $\varphi$  comme une fonction continue tendant vers 0 à l'infini (mais attention, une fonction continue tendant vers 0 à l'infini n'est pas forcément du type envisagé ici, les hypothèses faites ici sur  $\varphi$  étant plus restrictives).

Comme  $\|\Phi_{Y_k}\|_\infty \leq 1$ , on déduit du théorème de convergence dominée de Lebesgue et du fait que  $\Phi_{Y_k}$  converge ponctuellement vers la gaussienne  $g$  que

$$\lim_{k \rightarrow +\infty} \int_{\Omega} \varphi(Y_k) dP = \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{\varphi}(\tau) g(\tau) d\tau = \frac{1}{2\pi} \int_{\mathbb{R}} \varphi(t) \widehat{g}(t) dt = \int_{\mathbb{R}} \varphi(t) g(t) dt,$$

soit, utilisant le principe de transport des mesures de probabilités,

$$\lim_{k \rightarrow +\infty} \int_{\Omega} \varphi(Y_k) dP = \int_{\Omega} \varphi(\mathcal{N}(0, 1)) dP. \quad (2.38)$$

De fait, ce résultat est en fait valable pour toute fonction  $\varphi$  continue bornée sur  $\mathbb{R}$  (on l'admettra ici, c'est la conséquence d'un célèbre théorème de Paul Lévy<sup>49</sup>, et le résultat (partiellement) prouvé ici n'est rien d'autre que le célèbre *théorème « limite central »*<sup>50</sup> (voir le cours de MHT 601). Ce théorème (dont on vient de voir, ne serait-ce que par la preuve évoquée ici, qu'il n'était pas étranger à la transformation de Fourier) est l'une des raisons expliquant le rôle crucial joué par la gaussienne (on en verra une autre raison lorsque nous évoquerons dans la section 2.5.3 le principe d'incertitude de Werner Heisenberg, théorème 2.7). La célèbre *expérience de Galton* consistant à laisser tomber un stock de billes au dessus d'une grille avec des plots (la bille pouvant tomber indifféremment à gauche ou à droite lorsqu'elle rencontre un plot, ce suivant une loi binomiale) montre au final que le tas de billes se répartit dans les canaux du réceptacle suivant une distribution de Gauss, comme sur la figure 2.7 ci-dessous ; c'est l'illustration même du théorème limite central.

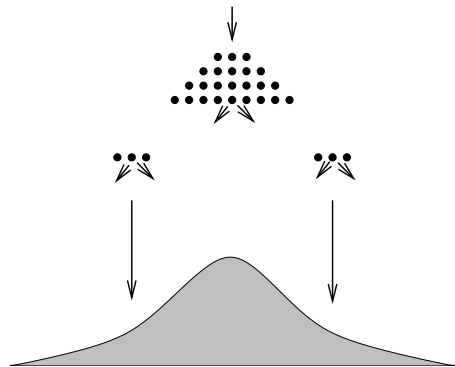


FIG. 2.7 – Le triangle de Galton

### 2.4.5 L'inversion locale et la formule sommatoire de Poisson

Commençons dans cette section par énoncer un analogue continu du théorème de Jordan-Dirichlet (théorème 2.1) concernant la restitution « locale » d'un élément  $f \in \mathcal{L}_C^1(\mathbb{R}, dt)$  à partir de la connaissance de son spectre. On oubliera dans cette

<sup>49</sup> « Si une suite  $(Y_k)_{k \geq 1}$  de variables aléatoires réelles toutes définies sur un même espace de probabilités est telle que la suite des fonctions caractéristiques  $(\Phi_{Y_k})_{k \geq 1}$  converge ponctuellement vers une fonction  $\Phi$  continue en  $\tau = 0$ , alors  $\Phi$  est la fonction caractéristique d'une distribution de probabilité  $P_Y$  et la suite des lois des variables aléatoires  $(Y_k)_{k \geq 1}$  converge en loi vers la distribution de probabilité  $P_Y$  ». Le probabiliste français Paul Lévy (1886-1971), certainement l'un des pionniers de la théorie des probabilités modernes, publia ce résultat en 1924 et son application au théorème limite central en est une des applications les plus célèbres.

<sup>50</sup> Il s'agit d'un théorème limite, central dans la théorie.

section les classes de fonctions  $\hat{f}$  pour nous intéresser plus spécifiquement (parce que les valeurs ponctuelles des fonctions seront en jeu) à des représentants particuliers  $f$  de ces classes. Nous nous placerons aussi en dimension  $n = 1$  (la variable temporelle sera donc le temps  $t$ ).

**Théorème 2.5 (Théorème de Jordan-Dirichlet, version continue)** *Soit  $f$  un élément de  $\mathcal{L}^1_{\mathbb{C}}(\mathbb{R}, dt)$  et  $t_0, l$  deux nombres réels tels que*

$$\int_0^1 \frac{|f(t_0 + u) + f(t_0 - u) - 2l|}{|u|} < +\infty \quad (2.39)$$

alors

$$\lim_{\Omega \rightarrow +\infty} \frac{1}{2\pi} \int_{-\Omega}^{\Omega} \hat{f}(\omega) e^{it_0\omega} d\omega = l.$$

**Preuve.** Grâce au théorème de Fubini

$$\begin{aligned} \frac{1}{2\pi} \int_{-\Omega}^{\Omega} \hat{f}(\omega) e^{i(t_0, \omega)} d\omega &= \frac{1}{2\pi} \int_{-\Omega}^{\Omega} \left( \int_{\mathbb{R}} f(t) e^{-i\omega t} dt \right) e^{it_0\omega} d\omega \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} f(t) \left( \int_{-\Omega}^{\Omega} e^{-i(t-t_0)\omega} d\omega \right) dt \\ &= \frac{1}{\pi} \int_{\mathbb{R}} f(t) \frac{\sin \Omega(t-t_0)}{t-t_0} dt \\ &= \frac{1}{\pi} \int_0^{\infty} (f(t_0+u) + f(t_0-u)) \frac{\sin \Omega u}{u} du. \end{aligned}$$

On a, en utilisant le fait que

$$\forall \Omega > 0, \int_0^{\infty} \frac{\sin \Omega u}{u} du = \int_0^{\infty} \frac{\sin u}{u} du = \frac{\pi}{2}$$

(il s'agit d'une intégrale semi convergente, voir par exemple le cours de L2 et le polycopié du cours de MHT401 <sup>51</sup>),

$$\begin{aligned} &\frac{1}{\pi} \int_0^{\infty} (f(t_0+u) + f(t_0-u)) \frac{\sin \Omega u}{u} du - l \\ &= \frac{1}{\pi} \int_0^1 (f(t_0+u) + f(t_0-u) - 2l) \frac{\sin \Omega u}{u} du \\ &\quad + \frac{1}{\pi} \int_1^{\infty} (f(t_0+u) + f(t_0-u)) \frac{\sin \Omega u}{u} du - \frac{1}{\pi} \int_1^{\infty} \frac{\sin \Omega u}{u} du. \end{aligned} \quad (2.40)$$

Les deux premières intégrales dans le membre de droite de (2.40) tendent vers 0 grâce au lemme de Riemann-Lebesgue (proposition 2.7) puisque les fonctions mesurables

$$\begin{aligned} u &\longmapsto \chi_{[0,1]}(u) \frac{f(t_0+u) + f(t_0-u) - 2l}{u} \\ u &\longmapsto \chi_{[1,\infty[}(u) \frac{f(t_0+u) + f(t_0-u)}{u} \end{aligned}$$

<sup>51</sup><http://www.math.u-bordeaux1.fr/~yger/mat401.pdf>

sont dans  $\mathcal{L}_{\mathbb{C}}^1(\mathbb{R}, dt)$ . Enfin

$$\lim_{\Omega \rightarrow +\infty} \frac{1}{\pi} \int_1^{\infty} \frac{\sin \Omega u}{u} du = \lim_{\Omega \rightarrow +\infty} \frac{1}{\pi} \int_{\Omega}^{\infty} \frac{\sin \Omega u}{u} du = 0.$$

La proposition est donc démontrée.  $\diamond$

**Exemple d'application 2.9.** Voici l'application majeure de ce résultat, comme dans le cadre périodique (comparez avec l'application du théorème 2.1) : si  $f$  présente une limite à gauche ( $f(t_0 - 0)$ ) en  $t_0$  ainsi qu'une limite à droite ( $f(t_0 + 0)$ ) en ce même point, alors on peut affirmer que

$$\lim_{\Omega \rightarrow +\infty} \frac{1}{2\pi} \int_{-\Omega}^{\Omega} \widehat{f}(\omega) e^{it_0\omega} d\omega = \frac{f(t_0 - 0) + f(t_0 + 0)}{2}$$

à condition toutefois que soit remplie la clause de sécurité suivante : le graphe de  $f$  présente des demi-tangentes à gauche et à droite respectivement aux points  $(t_0, f(t_0 - 0))$  et  $(t_0, f(t_0 + 0))$ . Il est important de souligner cependant que ce phénomène de restitution locale de  $f$  à partir de son spectre est, comme dans le cadre périodique, entaché du phénomène d'*aliasing* de Gibbs (figure 2.2 de la section 2.3.3).

Une autre application importante est la formule sommatoire de Poisson qui permet de « resommer » une série numérique convergente de manière à pouvoir améliorer la vitesse de convergence des sommes partielles de la série.

**Proposition 2.9 (Formule sommatoire de Poisson<sup>52</sup>)** Soit  $\dot{f}$  un élément de  $L_{\mathbb{C}}^1(\mathbb{R}, dt)$  tel que  $\widehat{f} \in L_{\mathbb{C}}^1(\mathbb{R}, d\omega)$  et  $f$  son unique représentant continu. La série bilatère

$$[\widehat{f}(\cdot - 2k\pi)]_{k \in \mathbb{Z}}$$

est donc absolument convergente dans  $L_{\mathbb{C}}^1([-\pi, \pi], d\theta)$  et sa somme (dans cet espace de Banach) définit un élément  $\psi \in L_{\mathbb{C}}^1([-\pi, \pi], d\theta)$  que l'on suppose de plus avoir un représentant continu en  $\theta = 0$  et présentant des dérivées à gauche et à droite en ce point. Alors

$$\lim_{N \rightarrow +\infty} \sum_{k=-N}^N f(k) = \lim_{N \rightarrow +\infty} \sum_{k=-N}^N \widehat{f}(2k\pi) \in \mathbb{C}.$$

**Preuve.** Comme

$$\begin{aligned} \int_{\mathbb{R}} |\widehat{f}(\omega)| d\omega &= \sum_{k=-\infty}^{\infty} \int_{(2k-1)\pi}^{(2k+1)\pi} |\widehat{f}(\omega)| d\omega = \sum_{k \in \mathbb{Z}} \int_{-\pi}^{\pi} |\widehat{f}(\theta + 2k\pi)| d\theta \\ &= \int_{-\pi}^{\pi} \left( \sum_{k \in \mathbb{Z}} |\widehat{f}(\theta + 2k\pi)| \right) d\theta < +\infty, \end{aligned}$$

on a bien convergence absolue de la série bilatère  $[\widehat{f}(\cdot - 2k\pi)]_{k \in \mathbb{Z}}$  dans  $L_{\mathbb{C}}^1(\mathbb{R}, d\theta)$  vers un élément  $\psi$  de cet espace. Grâce à la formule d'inversion de Fourier (théorème 2.4), on sait que

$$f(t) = \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{f}(\omega) e^{i\omega t} d\omega \quad \forall t \in \mathbb{R}$$

---

<sup>52</sup>On l'attribue au mathématicien français Siméon-Denis Poisson (1781-1840) que nous avons déjà évoqué à propos de la *loi de Poisson*.

et que par conséquent

$$\begin{aligned} \sum_{k=-N}^N f(k) &= \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{f}(\omega) \left( \sum_{k=-N}^N e^{ik\omega} \right) d\omega \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{f}(\omega) D_N(\omega) d\omega \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \psi(u) D_N(u) du. \end{aligned}$$

Mais nous nous trouvons exactement dans conditions d'application du théorème de Jordan-Dirichlet (théorème 2.1 et son application) en  $t_0 = 0$  à cause des hypothèses faites sur  $\psi$ . On en déduit

$$\lim_{N \rightarrow +\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} \psi(u) D_N(u) du = \psi(0),$$

d'où le résultat voulu.  $\diamond$

**Exemple 2.10.** On a vu (exemple 2.5 de la section 2.4.1) que la fonction

$$t \mapsto \frac{e^{-|t|}}{2}$$

avait pour spectre la fonction

$$\omega \mapsto \frac{1}{1 + \omega^2}.$$

D'après la formule d'inversion de Fourier (théorème 2.4), la fonction

$$f : t \mapsto \frac{1}{1 + t^2}$$

a pour spectre

$$\widehat{f} : \omega \mapsto \pi e^{-|\omega|}.$$

On vérifie que les conditions d'application de la proposition 2.9 sont ici remplies et que par conséquent, pour chaque  $\tau > 0$ ,

$$\sum_{k \in \mathbb{Z}} \frac{1}{1 + \tau^2 k^2} = \frac{\pi}{\tau} \sum_{k \in \mathbb{Z}} e^{-2\pi|k|/\tau}$$

(on travaille avec  $t \mapsto f(\tau t)$  au lieu de  $f$  pour appliquer la formule). On obtient ainsi

$$\begin{aligned} 2 \sum_{k=1}^{\infty} \frac{1}{1 + k^2 \tau^2} + 1 &= \frac{\pi}{\tau} \left( 2 \sum_{k=0}^{\infty} e^{-2k\pi/\tau} - 1 \right) \\ &= \frac{\pi}{\tau} \frac{1 + e^{-2\pi/\tau}}{1 - e^{-2\pi/\tau}}. \end{aligned}$$

C'est un exemple où la série de droite (série géométrique) converge beaucoup plus vite que la série de gauche (série équivalente à une série de Riemann)<sup>53</sup>. Si l'on prend  $\tau \gg 1$ , on écrit ceci

$$\frac{2}{\tau^2} \sum_{k=1}^{\infty} \frac{1}{k^2 + \frac{1}{\tau^2}} + 1 = \frac{\pi}{\tau} \times \frac{\tau}{2\pi} \left( 2 + \frac{1}{6} (2\pi/\tau)^2 + \dots \right)$$

et l'on retrouve en mettant les deux membres sous la forme

$$1 + \frac{\alpha_1}{\tau^2} + o(\tau^{-2}) = 1 + \frac{\pi^2}{3} \frac{1}{\tau^2} + o(\tau^{-2})$$

<sup>53</sup>Ceci sera un fait général dans l'utilisation de la formule de Poisson (c'est d'ailleurs une raison majeure de son intérêt) car on verra plus loin avec le principe d'incertitude d'Heisenberg (théorème 2.7) que mieux  $f$  est localisée en temps, plus  $\widehat{f}$  est diffus en fréquences (et *vice-versa*).

la valeur de

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\alpha_1}{2} = \frac{\pi^2}{6}.$$

**Exemple 2.11.** Si  $\tau > 0$ , la gaussienne  $t \mapsto e^{-t^2\tau^2/2}$  a pour transformée de Fourier la gaussienne  $\omega \mapsto \frac{\sqrt{2\pi}}{\tau} e^{-\omega^2/(2\tau^2)}$ ; on déduit par conséquent de la formule sommatoire de Poisson la formule

$$\sum_{k \in \mathbb{Z}} e^{-k^2\tau^2/2} = \frac{\sqrt{2\pi}}{\tau} \sum_{k \in \mathbb{Z}} e^{-\frac{2\pi^2 k^2}{\tau^2}};$$

si  $\tau$  est très petit, disons par exemple  $\tau \simeq 10^{-3}$ ,  $\tau^2 = 10^{-6}$  et il faut un nombre impressionnant de termes pour calculer la somme

$$\sum_{k \in \mathbb{Z}} e^{-k^2\tau^2/2}$$

avec une erreur de  $10^{-3}$  (entraînez vous à calculer combien!); en revanche, le calcul approché de la somme

$$\sum_{k \in \mathbb{Z}} e^{-\frac{2\pi^2 k^2}{\tau^2}}$$

se fait lui avec très peu de termes du fait que  $1/\tau^2 \simeq 10^6$ ! C'est là toute la force de la formule sommatoire de Poisson!

## 2.5 La transformation de Fourier dans $L^2_{\mathbb{C}}(\mathbb{R}^n, dx)$

### 2.5.1 Une isométrie de $\mathcal{G}$ dans lui-même (pour la norme $L^2$ )

On rappelle la définition du  $\mathbb{C}$ -sous-espace vectoriel  $\mathcal{G}$  de  $L^1_{\mathbb{C}}(\mathbb{R}^n, dx)$  constitué des classes de fonctions intégrables de la forme

$$f : x \in \mathbb{R}^n \longmapsto \sum_{j=1}^N \lambda_j e^{i(\omega^{(j)}, x)} g_{\epsilon_j}(x - x^{(j)}),$$

où  $x^{(j)}, \omega^{(j)}$ ,  $j = 1, \dots, N$ , désignent des points respectivement de  $\mathbb{R}_x^n$  et  $\mathbb{R}_\omega^n$ ,  $\epsilon_j$ ,  $j = 1, \dots, N$ , des nombres strictement positifs,  $\lambda_1, \dots, \lambda_N$  des coefficients complexes.

Nous avons la proposition suivante :

**Proposition 2.10** *Pour tout  $\dot{f} \in \mathcal{G}$ , on a, pour tout représentant  $f$  de  $\dot{f}$  :*

$$\int_{\mathbb{R}^n} |f(x)|^2 dx = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} |\widehat{f}(\omega)|^2 d\omega.$$

**Preuve.** On considère un représentant  $f$  de  $\dot{f}$  et la fonction

$$F : x \in \mathbb{R}^n \longmapsto \dot{f} * \overline{\dot{f}(-\cdot)}(x) = \int_{\mathbb{R}^n} f(y) \overline{f(y-x)} dy \quad (2.41)$$

(il se trouve que l'intégrale ci dessus est bien définie pour tout  $x \in \mathbb{R}^n$ , la fonction de  $x$  ainsi définie est donc bien un représentant de l'élément  $\dot{f} * \overline{\dot{f}(-\cdot)}$  de  $L^1_{\mathbb{C}}(\mathbb{R}^n, dx)$ ). On remarque que la fonction  $F$  définie par (2.41) est toujours un élément de  $\mathcal{G}$  (car  $\mathcal{G}$  est stable par convolution). Il résulte de la proposition 2.8 que l'on a

$$\left[ \dot{f} * \overline{\dot{f}(-\cdot)} \right]^{\wedge} = \widehat{f} \times \left[ \overline{\dot{f}(-\cdot)} \right]^{\wedge},$$

ce qui donne, pour tout  $\omega \in \mathbb{R}_\omega^n$  :

$$\begin{aligned} \left[ \dot{f} * \overline{\dot{f}(-\cdot)} \right]^\wedge(\omega) &= \widehat{f}(\omega) \times \int_{\mathbb{R}^n} \overline{f(-x)} e^{-i\langle x, \omega \rangle} dx \\ &= \widehat{f}(\omega) \times \int_{\mathbb{R}^n} \overline{f(x)} e^{i\langle x, \omega \rangle} dx \\ &= \widehat{f}(\omega) \times \overline{\int_{\mathbb{R}^n} f(x) e^{-i\langle x, \omega \rangle} dx} \\ &= \widehat{f}(\omega) \times \widehat{\widehat{f}}(\omega) = |\widehat{f}(\omega)|^2. \end{aligned}$$

Or la formule d'inversion de Fourier (théorème 2.4) nous assure qu'un autre représentant (lui aussi continu, comme  $F$ ) de la classe  $\dot{f} * \overline{\dot{f}(-\cdot)}$  est

$$\widetilde{F} : x \in \mathbb{R}^n \longmapsto \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} |\widehat{f}(\omega)|^2 e^{i\langle x, \omega \rangle} d\omega.$$

Comme  $F$  et  $\widetilde{F}$  sont deux représentants continus du même élément de  $L_{\mathbb{R}}^1(\mathbb{R}^n, dx)$ , ces deux représentants coïncident et on a, en prenant leur valeur commune en  $x = 0$ ,

$$F(0) = \int_{\mathbb{R}^n} |f(x)|^2 dx = \widetilde{F}(0) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} |\widehat{f}(\omega)|^2 d\omega,$$

ce qui est la formule escomptée.  $\diamond$ .

Les fonctions  $f$  représentant les éléments de  $\mathcal{G}$  étant clairement des fonctions de  $\mathcal{L}_{\mathbb{C}}^2(\mathbb{R}^n, dx)$ , on peut prendre leurs classes dans  $L_{\mathbb{C}}^2(\mathbb{R}^n, dx)$  et noter également  $\mathcal{G}$  le  $\mathbb{C}$ -sous-espace que ces classes constituent dans  $L_{\mathbb{C}}^2(\mathbb{R}^n, dx)$  cette fois. On a alors le résultat suivant :

**Proposition 2.11** *L'application qui à  $\dot{f} \in \mathcal{G} \subset L_{\mathbb{C}}^2(\mathbb{R}_x^n, dx)$  associe la classe de  $\widehat{f}$  dans  $\mathcal{G} \subset L_{\mathbb{C}}^2(\mathbb{R}_\omega^n, d\omega)$  est un isomorphisme entre ces deux copies de  $\mathcal{G}$  (l'une sur  $\mathbb{R}_x^n$ , l'autre sur  $\mathbb{R}_\omega^n$ ) et l'on a de plus*

$$\|\dot{f}\|_2^2 = \frac{1}{(2\pi)^n} \|\widehat{f}\|_2^2. \quad (2.42)$$

*L'isomorphisme inverse est l'application qui à  $\dot{f} \in \mathcal{G} \subset L_{\mathbb{C}}^2(\mathbb{R}_\omega^n, d\omega)$  associe la classe (dans  $L_{\mathbb{C}}^2(\mathbb{R}_x^n, dx)$ ) de la fonction*

$$x \in \mathbb{R}^n \longmapsto \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} f(\omega) e^{i\langle x, \omega \rangle} d\omega.$$

**Preuve.** C'est une application immédiate de la proposition 2.10 et de la formule d'inversion de Fourier (théorème 2.4) applicable ici puisque  $\mathcal{G}$  (qui est invariant sous la transformation de Fourier si l'on décide d'identifier les deux copies  $\mathbb{R}_x^n$  et  $\mathbb{R}_\omega^n$  en jeu dans cette transformation) est dans  $L_{\mathbb{C}}^1(\mathbb{R}_x^n \simeq \mathbb{R}_\omega^n, dx \simeq d\omega)$ .  $\diamond$

## 2.5.2 Extension à $L_{\mathbb{C}}^2(\mathbb{R}^n, dx)$ et formule de Plancherel

Le  $\mathbb{C}$ -sous-espace vectoriel  $\mathcal{G} \subset L_{\mathbb{C}}^2(\mathbb{R}_x^n \simeq \mathbb{R}_\omega^n, dx \simeq d\omega)$  est invariant par prise de transformation de Fourier et est d'autre part un  $\mathbb{C}$ -sous-espace dense dans l'espace



de Hilbert  $L^2_{\mathbb{C}}(\mathbb{R}^n, dx)$ . Si en effet  $\dot{f}$  est un élément de  $L^1_{\mathbb{C}}(\mathbb{R}^n, dx) \cap L^2_{\mathbb{C}}(\mathbb{R}^n, dx)$  orthogonal à  $\mathcal{G}$ , on a

$$\int_{\mathbb{R}^n} f(y) e^{-i\langle y, \omega \rangle} g_{\epsilon}(y-x) dy = \int_{\mathbb{R}^n} f(y) e^{-i\langle y, \omega \rangle} g_{\epsilon}(x-y) dy = 0$$

pour tout  $x, \omega$  dans  $\mathbb{R}^n$  et pour tout  $\epsilon > 0$  (compte tenu de la définition de  $\mathcal{G}$ ). En particulier, on a  $\dot{f} * \dot{g}_{\epsilon} = 0$  pour tout  $\epsilon > 0$ , ce qui implique  $\dot{f} = 0$ , puisque l'on sait que la suite  $(\dot{f} * \dot{g}_{\epsilon_k})_{k \geq 1}$  converge vers  $\dot{f}$  dans  $L^1_{\mathbb{C}}(\mathbb{R}^n, dx)$  lorsque  $(\epsilon_k)_{k \geq 1}$  est une suite de nombres strictement positifs tendant vers 0<sup>54</sup>. Comme le sous-espace  $L^1_{\mathbb{C}}(\mathbb{R}^n, dx) \cap L^2_{\mathbb{C}}(\mathbb{R}^n, dx)$  est un sous-espace dense de  $L^2_{\mathbb{C}}(\mathbb{R}^n, dx)$ <sup>55</sup>, on a bien densité dans  $L^2_{\mathbb{C}}(\mathbb{R}^n_x \simeq \mathbb{R}^n_x, dx \simeq dx)$  du sous-espace  $\mathcal{G}$ .

Le lemme 1.2 établi dans la section 1.11.1 du chapitre 1, combiné avec la formule d'isométrie (2.42) (valable pour tout  $\dot{g} \in \mathcal{G}$ ) établie dans la proposition 2.11, nous autorise à étendre (de manière unique) l'action de l'opération

$$\dot{f} \in \mathcal{G} \subset L^2_{\mathbb{C}}(\mathbb{R}^n_x, dx) \longmapsto \text{classe}(\widehat{f}) \in \mathcal{G} \subset L^2_{\mathbb{C}}(\mathbb{R}^n_{\omega}, d\omega)$$

en un opérateur linéaire continu

$$\mathcal{F} : \dot{f} \in L^2_{\mathbb{C}}(\mathbb{R}^n_x, dx) \longmapsto \mathcal{F}(\dot{f}) \in L^2_{\mathbb{C}}(\mathbb{R}^n_{\omega}, d\omega).$$

L'opération inverse

$$\text{classe}(\widehat{f}) \in \mathcal{G} \subset L^2_{\mathbb{C}}(\mathbb{R}^n_{\omega}, d\omega) \longmapsto \dot{f} \in \mathcal{G} \subset L^2_{\mathbb{C}}(\mathbb{R}^n_x, dx)$$

peut également être étendue en un opérateur linéaire continu

$$\mathcal{F}^{-1} : \tilde{f} \in L^2_{\mathbb{C}}(\mathbb{R}^n_{\omega}, d\omega) \longmapsto \mathcal{F}^{-1}(\tilde{f}) \in L^2_{\mathbb{C}}(\mathbb{R}^n_x, dx).$$

D'une part la formule (2.42) demeure valide, d'autre part l'opérateur  $\mathcal{F}$  coïncide sur le sous-espace  $L^1_{\mathbb{C}}(\mathbb{R}^n_x, dx) \cap L^2_{\mathbb{C}}(\mathbb{R}^n_x, dx)$  avec l'opérateur consistant à associer à  $\dot{f}$  la classe de  $\widehat{f}$ ,  $\dot{f}$  étant cette fois considéré comme élément de  $L^1_{\mathbb{C}}(\mathbb{R}^n_x, dx)$ .

Tout cela peut être résumé en le théorème suivant :

**Théorème 2.6** *Soit  $\dot{f} \in L^2_{\mathbb{C}}(\mathbb{R}^n_x, dx)$  et  $\tilde{f} \in L^2_{\mathbb{C}}(\mathbb{R}^n_{\omega}, d\omega)$ . Les limites*

$$\lim_{\substack{T \rightarrow +\infty \\ L^2_{\mathbb{C}}(\mathbb{R}^n_{\omega}, d\omega)}} \left[ \omega \longmapsto \int_{[-T, T]^n} f(x) e^{-i\langle x, \omega \rangle} dx \right] := \mathcal{F}(\dot{f}) \in L^2_{\mathbb{C}}(\mathbb{R}^n_{\omega}, d\omega)$$

$$\lim_{\substack{\Omega \rightarrow +\infty \\ L^2_{\mathbb{C}}(\mathbb{R}^n_x, dx)}} \left[ x \longmapsto \frac{1}{(2\pi)^n} \int_{[-\Omega, \Omega]^n} \tilde{f}(\omega) e^{i\langle x, \omega \rangle} d\omega \right] := \mathcal{F}^{-1}(\tilde{f}) \in L^2_{\mathbb{C}}(\mathbb{R}^n_x, dx)$$

*et définissent l'action de deux opérateurs inverses l'un de l'autre que l'on appelle respectivement transformation de Fourier et transformation de Fourier inverse. De plus, on a la formule, dite formule de Plancherel<sup>56</sup>*

$$\forall \dot{f} \in L^2_{\mathbb{C}}(\mathbb{R}^n_x, dx), \quad \|\dot{f}\|_2^2 = \frac{1}{(2\pi)^n} \|\mathcal{F}(\dot{f})\|_2^2, \quad (2.43)$$

<sup>54</sup>Toujours la proposition 4.10 du cours de théorie de l'intégration MHT 512.

<sup>55</sup>Voir le cours de théorie de l'intégration : ceci résulte du fait que ce  $\mathbb{C}$ -sous-espace contient les classes des fonctions indicatrices des pavés  $[a_1, b_1] \times \dots \times [a_n, b_n]$ .

<sup>56</sup>On fait référence ici aux travaux du mathématicien suisse Michaël Plancherel, 1885-1967.

formule que l'on peut dédoubler en

$$\forall \dot{f}_1, \dot{f}_2 \in L^2_{\mathbb{C}}(\mathbb{R}^n_x, dx), \quad \int_{\mathbb{R}^n} f_1(x) \overline{f_2(x)} dx = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \mathcal{F}(\dot{f}_1)[\omega] \overline{\mathcal{F}(\dot{f}_2)[\omega]} d\omega. \quad (2.44)$$

**Attention !** Contrairement à ce qui se passe pour  $\widehat{f}$  lorsque  $\dot{f} \in L^1_{\mathbb{C}}(\mathbb{R}^n_x, dx)$ ,  $\mathcal{F}(\dot{f})$  est un élément de  $L^2_{\mathbb{C}}(\mathbb{R}^n_\omega, d\omega)$  (défini comme une limite dans cet espace de classes de fonctions elles ponctuellement définies, à savoir les classes des spectres des  $f\chi_{[-T, T]^n}$ ,  $T \rightarrow +\infty$ , qui, eux, sont des spectres d'éléments de  $L^1_{\mathbb{C}}(\mathbb{R}^n_x, dx)$ ) et n'a donc pas de sens ponctuel. On ne saurait parler de  $\mathcal{F}(\dot{f})[\omega_0]$  en un point spécifique  $\omega_0$  de l'espace  $\mathbb{R}^n_\omega$ . Il faut prendre garde à ce point subtil mais essentiel. C'est la même chose en ce qui concerne  $\mathcal{F}^{-1}(\tilde{f})$  lorsque  $\tilde{f}$  est un élément de  $L^2_{\mathbb{C}}(\mathbb{R}^n_\omega, d\omega)$ .

**Preuve.** L'énoncé est juste un résumé de ce qui précède dans cette section 2.5.2. Tout a été dit, si ce n'est le fait bien connu qu'un élément de  $L^2_{\mathbb{C}}(\mathbb{R}^n_x, dx)$  (*resp.*  $L^2_{\mathbb{C}}(\mathbb{R}^n_\omega, d\omega)$ ) s'approche dans cet espace par les classes des fonctions « tronquées »

$$f\chi_{[-T, T]^n}, \quad T > 0,$$

en faisant tendre  $T$  vers l'infini.  $\diamond$

**Exemple 2.12.** On a vu (exemple 2.3 dans la section 2.4.1) que le spectre de la fonction caractéristique  $\chi_{[-T, T]}$  était la fonction

$$\omega \mapsto 2 \frac{\sin \omega T}{\omega}.$$

Grâce au théorème 2.6, la transformée de de Fourier

$$\mathcal{F}\left(t \mapsto 2 \frac{\sin Tt}{t}\right)$$

a pour représentant dans  $L^2_{\mathbb{C}}(\mathbb{R}_\omega, d\omega)$  la fonction

$$\omega \mapsto 2\pi \times \chi_{[-T, T]}.$$

La transformée de Fourier de la fonction

$$\text{sinc} : t \in \mathbb{R} \mapsto \frac{\sin(\pi t/2)}{\pi t/2}$$

a pour représentant

$$\omega \mapsto 2\chi_{[-\pi/2, \pi/2]}.$$

### 2.5.3 Fourier versus la dérivation ; le principe d'incertitude d'Heisenberg

Plus le spectre d'un élément  $\dot{f}$  de  $L^1_{\mathbb{C}}(\mathbb{R}^n_x, dx)$  décroît rapidement lorsque  $\|\omega\|$  tend vers l'infini, plus l'on peut affirmer que  $\dot{f}$  est régulier, c'est à dire admet un représentant de classe  $C^m$ , avec  $m$  lié à la vitesse de convergence du spectre. En effet, on a la

**Proposition 2.12** Si  $\dot{f} \in L^1_{\mathbb{C}}(\mathbb{R}^n_x, dx) \cap L^2_{\mathbb{C}}(\mathbb{R}^n_x, dx)$  et si  $\omega \mapsto \|\omega\|^m |\widehat{\dot{f}}(\omega)|$  définit un élément de  $L^1_{\mathbb{C}}(\mathbb{R}^n_\omega, d\omega) \cap L^2_{\mathbb{C}}(\mathbb{R}^n_\omega, d\omega)$  pour un certain entier  $m \in \mathbb{N}$ ,  $\dot{f}$  admet un représentant  $f$  de classe  $C^m$  et pour tout opérateur différentiel  $P(\partial/\partial x_1, \dots, \partial/\partial x_n)$

d'ordre total inférieur ou égal à  $m$ ,  $P[D](f)$  définit un élément de  $L^2_{\mathbb{C}}(\mathbb{R}^n_x, dx)$  dont la transformée de Fourier a pour représentant

$$\omega \in \mathbb{R}^n \longmapsto P(i\omega)\widehat{f}(\omega) ;$$

de plus, on a

$$\int_{\mathbb{R}^n} |P[D](f)(x)|^2 dx = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} |P(i\omega)|^2 |\widehat{f}(\omega)|^2 d\omega .$$

**Preuve.** On utilise tout d'abord la formule d'inversion de Fourier (théorème 2.4) qui nous assure que  $\widehat{f}$  admet pour représentant

$$f : x \in \mathbb{R}^n \longmapsto \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \widehat{f}(\omega) e^{i\langle x, \omega \rangle} d\omega .$$

En appliquant à répétition le théorème de dérivation de Lebesgue des intégrales à paramètres <sup>57</sup>, on voit que le fait que

$$\omega \longmapsto \|\omega\|^m |\widehat{f}(\omega)|$$

soit dans  $L^1_{\mathbb{C}}(\mathbb{R}^n_{\omega}, d\omega)$  implique que  $f$  est de classe  $C^m$  et que l'on a, pour tout opérateur différentiel  $P(\partial/\partial x_1, \dots, \partial/\partial x_n)$  d'ordre au plus  $m$ ,

$$P(D)[f](\omega) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} P(i\omega) \widehat{f}(\omega) e^{i\langle x, \omega \rangle} d\omega .$$

Le théorème 2.6 nous assure que  $P(D)[f]$  définit bien un élément de  $L^2_{\mathbb{C}}(\mathbb{R}^n_x, dx)$  puisque

$$\omega \longmapsto \|\omega\|^m |\widehat{f}(\omega)|$$

est aussi dans  $L^2_{\mathbb{C}}(\mathbb{R}^n_{\omega}, d\omega)$  et que le degré total de  $P$  est au plus  $m$ . De plus, ce même théorème nous assure que  $\mathcal{F}(P[D](f))$  a pour représentant

$$\omega \in \mathbb{R}^n \longmapsto P(i\omega)\widehat{f}(\omega) .$$

Enfin, on a

$$\int_{\mathbb{R}^n} |P[D](f)(x)|^2 dx = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} |P(i\omega)|^2 |\widehat{f}(\omega)|^2 d\omega$$

d'après la formule de Plancherel (2.43). La proposition est ainsi démontrée.  $\diamond$

La proposition 2.12 peut aussi se lire en échangeant les deux mondes  $\mathbb{R}^n_x$  et  $\mathbb{R}^n_{\omega}$ ; plus  $\widehat{f}$  a un représentant décroissant vers 0 rapidement, plus le spectre  $\widehat{f}$  est régulier. On a la proposition :

**Proposition 2.13** *Si  $f$  est une fonction mesurable sur  $\mathbb{R}^n_x$  telle que*

$$x \longmapsto \|x\|^m |f(x)|$$

définisse un élément de  $L^1_{\mathbb{C}}(\mathbb{R}^n_x, dx) \cap L^2_{\mathbb{C}}(\mathbb{R}^n_x, dx)$  pour un certain entier  $m \in \mathbb{N}$ , alors  $\widehat{f}$  est de classe  $C^m$  et pour tout opérateur différentiel  $P(\partial/\partial x_1, \dots, \partial/\partial x_n)$  d'ordre total inférieur ou égal à  $m$ ,  $P[D](\widehat{f})$  définit un élément de  $L^2_{\mathbb{C}}(\mathbb{R}^n_{\omega}, d\omega)$ , avec

$$P[D](\widehat{f})(\omega) = \left[ x \longmapsto P(-ix)f(x) \right]^{\wedge}(\omega) \quad \forall \omega \in \mathbb{R}^n_{\omega} ;$$

<sup>57</sup>Théorème 3.3 du cours de théorie de l'intégration MHT 512.

de plus, on a

$$\int_{\mathbb{R}^n} |P[D](\widehat{f})(\omega)|^2 d\omega = (2\pi)^n \int_{\mathbb{R}^n} |f(x)|^2 |P(-ix)|^2 dx.$$

**Preuve.** Les hypothèses permettent d'appliquer à répétition (pas plus de  $m$  fois cependant) le théorème de dérivation de Lebesgue des intégrales dépendant d'un paramètre à

$$\omega \in \mathbb{R}^n \longmapsto \widehat{f}(\omega) = \int_{\mathbb{R}^n} f(x) e^{-i\langle x, \omega \rangle} dx.$$

Cette fonction est donc de classe  $C^m$  et on a

$$P[D](\widehat{f})(\omega) = \int_{\mathbb{R}^n} P(-ix) f(x) e^{-i\langle x, \omega \rangle} dx \quad \forall \omega \in \mathbb{R}^n.$$

Le théorème 2.6 nous permet d'affirmer que  $P[D](\widehat{f})$  définit un élément de l'espace  $L^2_{\mathbb{C}}(\mathbb{R}^n, d\omega)$  et que

$$\int_{\mathbb{R}^n} |P[D](\widehat{f})(\omega)|^2 d\omega = (2\pi)^n \int_{\mathbb{R}^n} |f(x)|^2 |P(-ix)|^2 dx$$

d'après la formule de Plancherel (2.43). La proposition est démontrée.  $\diamond$

Derrière les énoncés des propositions 2.12 et 2.13, on voit poindre la construction d'une hiérarchie d'espaces de Hilbert, tous sous-espaces de  $L^2_{\mathbb{C}}(\mathbb{R}^n, dx)$ , introduits par le mathématicien russe Sergei Sobolev (1908-1989) vers 1935<sup>58</sup>.

Si nous ne pouvons aller plus avant en direction des contributions de Sobolev, nous pouvons cependant formuler un principe essentiel en physique, soutenu par l'analyse de Fourier, principe mettant une nouvelle fois en lumière le rôle privilégié de la gaussienne dans cette analyse. Pour simplifier, nous nous limiterons à la dimension  $n = 1$ .

**Théorème 2.7 (Principe d'incertitude d'Heisenberg<sup>59</sup>)** Soit  $f$  une fonction mesurable sur  $\mathbb{R}$  telle que  $t \longmapsto |t| |f(t)|$  définisse un élément de  $L^1_{\mathbb{C}}(\mathbb{R}_t, dt) \cap L^2_{\mathbb{C}}(\mathbb{R}_t, dt)$ , que  $\omega \longmapsto |\omega| |\widehat{f}(\omega)|$  définisse un élément de  $L^1_{\mathbb{C}}(\mathbb{R}_\omega, d\omega) \cap L^2_{\mathbb{C}}(\mathbb{R}_\omega, d\omega)$ , et que de plus  $\|f\|_2 = 1$ <sup>60</sup>. Alors

$$\sqrt{\int_{\mathbb{R}} |t|^2 |f(t)|^2 dt} \times \sqrt{\int_{\mathbb{R}} |\omega|^2 |\widehat{f}(\omega)|^2 d\omega} \geq \sqrt{\frac{\pi}{2}}, \quad (2.45)$$

l'inégalité ne devenant une égalité que si  $f$  est une gaussienne centrée de la forme

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

avec  $\operatorname{Re} \lambda > 0$ .

<sup>58</sup>C'est à Sergei Sobolev en Union Soviétique et à Laurent Schwartz (1915-2002) en France que l'on doit les fondements de la théorie des fonctions généralisées (ou distributions), formidable outil de modélisation dans la théorie des équations aux dérivées partielles; le formalisme de la théorie des distributions, enseigné en master 1, est malheureusement nécessaire pour asseoir la définition mathématique des espaces de Sobolev. Les deux propositions 2.12 et 2.13 ne font ici que suggérer cette définition.

<sup>59</sup>Il fut formulé par le physicien allemand Werner Heisenberg au printemps 1927, à l'aube des développements ultérieurs de la mécanique quantique.

<sup>60</sup>Cette dernière condition est juste une condition de normalisation.

**Preuve.** D'après la proposition 2.12,  $\hat{f}$  admet un représentant dérivable (c'est lui que nous noterons maintenant  $f$ ), tel que  $f'$  définisse un élément de  $L^2_{\mathbb{C}}(\mathbb{R}_t, dt)$  et que

$$\int_{\mathbb{R}} |f'(t)|^2 dt = \frac{1}{2\pi} \int_{\mathbb{R}} |\omega|^2 |\hat{f}(\omega)|^2 d\omega.$$

De plus, si  $u$  et  $v$  sont deux réels strictement positifs, alors, une intégration par parties donne

$$\operatorname{Re} \left( \int_{-u}^v t f(t) \overline{f'(t)} dt \right) = \frac{1}{2} \left[ t |f(t)|^2 \right]_{-u}^v - \frac{1}{2} \int_{-u}^v |f(t)|^2 dt. \quad (2.46)$$

Comme

$$\int_{\mathbb{R}} |t| |f(t)| dt < +\infty,$$

il existe deux suites strictement croissantes et tendant vers  $+\infty$  de nombres réels strictement positifs  $(u_k)_{k \geq 0}$  et  $(v_k)_{k \geq 0}$  telles que

$$\lim_{k \rightarrow +\infty} (-u_k) |f(-u_k)|^2 = \lim_{k \rightarrow +\infty} v_k |f(v_k)|^2 = 0.$$

De plus,  $tf$  et  $f'$  définissent des éléments de  $L^2_{\mathbb{C}}(\mathbb{R}_t, dt)$ , ce qui implique la convergence (absolue) de l'intégrale impropre

$$\int_{\mathbb{R}} t f(t) \overline{f'(t)} dt.$$

Si l'on choisit  $u = u_k$ ,  $v = v_k$  dans (2.46) et que l'on fasse tendre  $k$  vers l'infini, on trouve donc

$$\operatorname{Re} \left( \int_{-u}^v t f(t) \overline{f'(t)} dt \right) = -\frac{1}{2} \int_{\mathbb{R}} |f(t)|^2 dt = -\frac{1}{2}.$$

Il vient donc, en utilisant l'inégalité de Cauchy-Schwarz et la proposition 2.12 (pour exprimer  $\|f'\|_2$ ),

$$\begin{aligned} \frac{1}{2} &\leq \left| \int_{\mathbb{R}} t f(t) \overline{f'(t)} dt \right| \leq \sqrt{\int_{\mathbb{R}} t^2 |f(t)|^2 dt} \sqrt{\int_{\mathbb{R}} |f'(t)|^2 dt} \\ &\leq \sqrt{\int_{\mathbb{R}} t^2 |f(t)|^2 dt} \times \frac{1}{\sqrt{2\pi}} \sqrt{\int_{\mathbb{R}} |\omega|^2 |\hat{f}(\omega)|^2 d\omega}, \end{aligned} \quad (2.47)$$

ce qui fournit l'inégalité voulue. Pour que cette inégalité ne soit pas stricte, il faut que l'on se trouve dans le cas d'égalité lors de notre utilisation de l'inégalité de Cauchy-Schwarz dans  $L^2_{\mathbb{C}}(\mathbb{R}_t, dt)$  avec les deux éléments que sont les classes de  $t \mapsto tf(t)$  et de  $t \mapsto f'(t)$  (première étape de la chaîne d'inégalités (2.47)). L'égalité n'est possible que s'il existe une constante  $\lambda \in \mathbb{C}$  telle que

$$\forall t \in \mathbb{R}, f'(t) + 2\lambda t f(t) = 0.$$

Cette équation différentielle se résout immédiatement et l'on trouve bien

$$f : t \mapsto \frac{1}{2} \sqrt{\frac{\lambda}{\pi}} e^{-\lambda t^2}$$

avec  $\operatorname{Re} \lambda > 0$  puisque  $f$  doit être intégrable ; il y a deux solutions possibles (puisque un nombre complexe non nul a deux racines carrées), la constante multiplicative tenant compte de la normalisation  $\|\dot{f}\|_2 = 1$ .  $\diamond$

Voici l'interprétation du théorème 2.7 ; les deux facteurs figurant au membre de gauche dans l'inégalité (2.45) peuvent s'interpréter comme les racines carrées des moments d'inertie (par rapport à l'origine) des répartitions d'énergie  $|f(t)|^2 dt$  et  $|\widehat{f}(\omega)|^2 d\omega$ . Plus le premier facteur est petit, meilleure est la concentration de l'énergie de  $f$  au voisinage de l'origine dans l'espace  $\mathbb{R}_t$  des temps ; plus le second facteur est petit, meilleure est la concentration de l'énergie du spectre de  $f$  au voisinage de l'origine dans l'espace  $\mathbb{R}_\omega$  des fréquences. Ce que dit l'inégalité (2.45), c'est que l'on est dans l'incapacité de rendre ces deux facteurs petits, autrement dit, on ne peut bien localiser à la fois l'énergie de  $f$  et l'énergie de son spectre. Le moins pire « compromis » s'avère être le choix des gaussiennes, d'où leur rôle pour modéliser les particules en mécanique quantique. Ce principe crucial conditionne autant en mathématiques qu'en physique l'outil qu'est la transformation de Fourier. Il faut tenter de « faire avec ». Il est important de bien l'avoir compris.

### 2.5.4 Fonctions $L^2$ à spectre borné ; le théorème de Shannon

Les deux dernières sections de ce chapitre (et du cours) doivent être vues comme des digressions (pouvant servir de thème d'exercice ou de problème de révision) autour des deux chapitres. La présente section est dévolue à la notion de « spectre borné » et surtout au résultat majeur qu'est le *théorème d'échantillonnage* de Shannon-Nyquist.

**Définition 2.8** Un élément  $f \in L^2_{\mathbb{C}}(\mathbb{R}^n, dx)$  est dit à spectre borné s'il existe  $\Omega > 0$  tel que

$$\mathcal{F}(f) = 0 \quad \text{presque partout hors de } [-\Omega, \Omega]^n.$$

Du fait du théorème 2.6, un tel élément  $f$  admet toujours un représentant continu, à savoir<sup>61</sup>

$$f : x \in \mathbb{R}^n \longmapsto \frac{1}{(2\pi)^n} \int_{[-\Omega, \Omega]^n} \mathcal{F}(f)(\omega) e^{i\langle x, \omega \rangle} d\omega.$$

Nous allons montrer que dans cette section qu'en fait, si  $\tau_1, \dots, \tau_n$  sont des nombres strictement positifs tels que  $\tau_j \in ]0, \pi/\Omega]$ ,  $j = 1, \dots, n$ , alors cette fonction continue est entièrement déterminée par la donnée de tous les échantillons

$$f(k_1\tau_1, \dots, k_n\tau_n), \quad k_1, \dots, k_n \in \mathbb{Z}^n.$$

L'hypothèse portant sur les  $\tau_j$  ( $\tau_j \leq \pi/\Omega$  pour  $j = 1, \dots, n$ ) est essentielle : par exemple, si  $\epsilon \in ]0, \Omega[$ , la fonction

$$f : t \in \mathbb{R} \longmapsto \frac{\sin(\Omega - \epsilon)t}{t} \times \sin(\epsilon t),$$

<sup>61</sup>En fait, par exemple lorsque  $n = 1$ , ce représentant est même une fonction réelle analytique, c'est à dire une fonction s'exprimant sur  $\mathbb{R}$  comme la somme d'une série entière de rayon de convergence  $+\infty$  (il suffit de développer l'exponentielle en série sous l'intégrale, puis d'intervertir, ce qui licite ici d'après le théorème de Lebesgue, sommation et intégration) ; c'est donc en particulier une fonction  $C^\infty$  qui, à moins d'être identiquement nulle, ne saurait présenter que des zéros isolés.

qui définit bien un élément  $\dot{f}$  de  $L^2_{\mathbb{C}}(\mathbb{R}, dt)$  tel que  $\mathcal{F}(\dot{f}) = 0$  presque partout hors de  $[-\Omega, \Omega]$ , ne saurait être déterminée à partir de ses échantillons  $f(\pi k/(\Omega - \epsilon))$ ,  $k \in \mathbb{Z}$ , car ces nombres sont tous nuls !

Nous énonçons pour simplifier le résultat dans le cas  $n = 1$  mais il est immédiatement transposable au cadre multidimensionnel. Il s'agit, on s'en doute, d'un résultat très important du point de vue pratique puisqu'il valide la possibilité d'échantillonner une fonction continue d'énergie finie et de spectre vivant dans  $[-\Omega, \Omega]$  sans courir le risque d'une perte d'informations (pourvu que le pas d'échantillonnage soit plus petit que le seuil  $\pi/\Omega$  que les ingénieurs qualifient de *seuil de Nyquist*) ; c'est le théorème de Shannon-Nyquist<sup>62</sup> :

**Théorème 2.8** *Soit  $\dot{f} \in L^2_{\mathbb{C}}(\mathbb{R}_t, dt)$  une fonction à spectre borné, inclus dans l'intervalle  $[-\Omega, \Omega]$  avec  $\Omega > 0$ , et  $f$  un représentant continu de  $\dot{f}$ . Alors, pour tout  $\tau \in ]0, \pi/\Omega]$ , la suite de fonctions  $(f_N)_{N \in \mathbb{N}}$ , où*

$$f_N : t \in \mathbb{R} \longmapsto \sum_{k=-N}^N f(k\tau) \operatorname{sinc}\left(\frac{2(t - k\tau)}{\tau}\right), \quad N \in \mathbb{N},$$

*converge uniformément vers  $f$  sur  $\mathbb{R}$  <sup>63</sup>.*

**Preuve.** D'après le théorème 2.6, on sait que l'unique représentant continu de  $\dot{f}$  est la fonction

$$\begin{aligned} f : t \in \mathbb{R} \longmapsto \frac{1}{2\pi} \int_{\mathbb{R}} \mathcal{F}(\dot{f})(\omega) e^{i\omega t} d\omega &= \frac{1}{2\pi} \int_{-\Omega}^{\Omega} \mathcal{F}(\dot{f})(\omega) e^{i\omega t} d\omega \\ &= \frac{1}{2\pi} \int_{-\pi/\tau}^{\pi/\tau} \mathcal{F}(\dot{f})(\omega) e^{i\omega t} d\omega \\ &= \frac{1}{2\pi} \frac{1}{\tau} \int_{-\pi}^{\pi} \mathcal{F}(\dot{f})(\omega/\tau) e^{i\omega t/\tau} d\omega \\ &= \frac{1}{2\pi} \frac{1}{\tau} \int_{-\pi}^{\pi} \mathcal{F}(\dot{f})(\theta/\tau) \overline{e^{-i\theta t/\tau}} d\theta \end{aligned}$$

On fixe  $t \in \mathbb{R}$  et on utilise maintenant la formule (2.15) (formule de Parseval) introduite dans la section 2.3.1 avec les deux fonctions

$$\begin{aligned} f_1 : \theta \in [-\pi, +\pi[ &\longmapsto \mathcal{F}(\dot{f})(\theta/\tau) \\ f_2 : \theta \in [-\pi, \pi[ &\longmapsto e^{-i\theta t/\tau} \end{aligned}$$

(ces deux fonctions sont ensuite prolongées par  $2\pi$ -périodicité en des fonctions définissant des classes  $\dot{f}_1$  et  $\dot{f}_2$  dans  $L^2_{\mathbb{C}}(\mathbb{T})$ ). Les coefficients de Fourier complexes de

<sup>62</sup>Ingénieur électrique américain, Claude Elwood Shannon (1916-2001) fut un des pionniers de la théorie de l'information ; on lui doit l'introduction de la notion d'entropie et la mise en évidence de son rôle majeur dans les problèmes de gain ou de compression d'information, les bases de la théorie du codage, *etc.*. Le résultat que nous citons ici, attribué à Shannon et à Harry Nyquist (1889-1976), physicien suédois collaborateur de Shannon aux Bell Labs, est un principe majeur en théorie de la communication et dans le traitement des signaux ou des images.

<sup>63</sup>On rappelle la définition de la fonction *sinus cardinal* sur  $\mathbb{R}$  par  $\operatorname{sinc}(u) = \frac{\sin(\pi u/2)}{\pi u/2}$ .

$\dot{f}_1$  sont donnés par

$$\begin{aligned} c_k(\dot{f}_1) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathcal{F}(\dot{f})(\theta/\tau) e^{-ik\theta} d\theta = \frac{\tau}{2\pi} \int_{-\pi/\tau}^{\pi/\tau} \mathcal{F}(\dot{f})(-\omega) e^{ik\tau\omega} d\omega \\ &= \frac{\tau}{2\pi} \int_{-\Omega}^{\Omega} \mathcal{F}(\dot{f})(-\omega) e^{ik\tau\omega} d\omega = \tau f(-k\tau) \quad \forall k \in \mathbb{Z}. \end{aligned}$$

Les coefficients de Fourier complexes de  $\dot{f}_2$  sont donnés, eux, par

$$\begin{aligned} c_k(\dot{f}_2) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i\theta t/\tau} e^{-ik\theta} d\theta = \frac{1}{2\pi} \frac{e^{i(k+t/\tau)\pi} - e^{-i(k+t/\tau)\pi}}{i(k+t/\tau)} \\ &= \frac{\sin(k+t/\tau)\pi}{(k+t/\tau)\pi} = \operatorname{sinc}\left(\frac{2(t - (-k\tau))}{\tau}\right) \quad \forall k \in \mathbb{Z}. \end{aligned}$$

En appliquant la formule (2.15), on trouve

$$\begin{aligned} f(t) &= \sum_{k=-\infty}^{k=+\infty} f(-k\tau) \operatorname{sinc}\left(\frac{2(t - (-k\tau))}{\tau}\right) = \sum_{k=-\infty}^{k=+\infty} f(k\tau) \operatorname{sinc}\left(\frac{2(t - k\tau)}{\tau}\right) \\ &= \lim_{N \rightarrow +\infty} f_N(t) \end{aligned}$$

(la convergence étant ici une convergence ponctuelle). Le fait que la convergence soit uniforme résulte du fait que la fonction

$$t \in \mathbb{R} \mapsto \sum_{k \in \mathbb{Z}} \operatorname{sinc}^2\left(\frac{2(t - k\tau)}{\tau}\right)$$

soit une fonction  $\tau$ -périodique continue, donc bornée<sup>64</sup> par une constante  $C > 0$ , et que  $(f(k\tau))_{k \in \mathbb{Z}} \in l^2_{\mathbb{C}}(\mathbb{Z})$ . Pour  $N_2 > N_1$ , on a, d'après l'inégalité de Cauchy-Schwarz,

$$\|f_{N_2} - f_{N_1}\|_{\infty} \leq \sqrt{C} \left( \sum_{N_1 < |k| \leq N_2} |f(k\tau)|^2 \right)^{1/2}.$$

Le critère de Cauchy uniforme est donc rempli car on peut réaliser  $\|f_{N_2} - f_{N_1}\|_{\infty} < \epsilon$  pour  $N_2 \geq N_1 \geq N(\epsilon)$  avec  $N(\epsilon)$  assez grand ; on en déduit la convergence uniforme de  $(f_N)_{N \in \mathbb{N}}$  vers  $f$  et donc la conclusion du théorème de Shannon.  $\diamond$

## 2.6 L'outil Fourier couplé avec l'analyse hilbertienne ; deux exemples

Pour faire le pont entre les deux chapitres de ce cours, nous avons choisi deux exemples où analyse hilbertienne et analyse de Fourier se croisent et se complètent : le principe mathématique du CAT-Scanner et le problème de l'extrapolation des informations.

<sup>64</sup>Il se trouve en fait que cette fonction est en fait constante et égale à 1, ce que l'on voit en travaillant un peu plus, mais nous n'en avons pas besoin ici.



### 2.6.1 Un premier exemple ; l'inversion de transformations intégrales et le CAT-Scanner

Soit  $\dot{f} \in L^2_{\mathbb{R}}(\mathbb{R}^n, dx)$ , admettant un représentant nul presque partout hors de la boule euclidienne fermée de  $\mathbb{R}^n$  de rayon  $R$  (avec  $R > 0$ )<sup>65</sup>. Pour tout  $\xi \in \mathbb{S}^{n-1}(\mathbb{R})$  (où  $\mathbb{S}^{n-1}(\mathbb{R})$  désigne la sphère unité de  $\mathbb{R}^n$ ), il résulte du théorème de Fubini que la fonction

$$p \in \mathbb{R} \longmapsto \int_{y \in \xi^\perp} f(p\xi + y) d_{\xi^\perp} y$$

(où  $\xi^\perp$  désigne l'hyperplan orthogonal à  $\mathbb{R}\xi$  pour le produit scalaire usuel,  $d_{\xi^\perp} y$  désignant la mesure de Lebesgue sur  $\mathbb{R}^n$  restreinte à cet hyperplan) est bien définie presque partout (elle est nulle presque partout hors de  $[-R, R]$ ) et induit (comme fonction de  $p$ ) une classe  $R_\theta[\dot{f}]$  de  $L^2_{\mathbb{R}}(\mathbb{R}_p, dp)$ . L'application qui à une telle classe  $\dot{f} \in L^2_{\mathbb{R}}(\mathbb{R}^n, dx)$  associe la collection d'éléments

$$\left( R_\theta[\dot{f}] \right)_{\xi \in \mathbb{S}^{n-1}(\mathbb{R})}$$

est appelée *transformation de Radon par les hyperplans*. Être capable de restaurer  $\dot{f}$  à partir de la connaissance de toutes les « coupes »  $R_\xi[\dot{f}]$ ,  $\xi \in \mathbb{S}^{n-1}(\mathbb{R})$ , est un problème d'un intérêt pratique majeur, on le verra plus loin. Nous allons ici proposer deux solutions pour ce faire, l'une reposant sur l'utilisation de la transformation de Fourier, l'autre sur le maniement de l'analyse hilbertienne.

Pour mettre en œuvre la première solution, fixons  $\xi \in \mathbb{S}^{n-1}(\mathbb{R})$  et calculons la transformée de Fourier (au sens  $L^2$ ) de  $R_\xi[\dot{f}]$ . Un représentant de  $\mathcal{F}(R_\xi[\dot{f}])$  est la fonction

$$\begin{aligned} \tau \in \mathbb{R} &\longmapsto \int_{-R}^R \left( \int_{\xi^\perp} f(p\vec{\theta} + y) d_{\xi^\perp}(y) \right) e^{-i p \tau} dp \\ &= \int_{\mathbb{R}^n} f(x) e^{-i \langle x, \tau \xi \rangle} dx = \widehat{f}(\tau \xi). \end{aligned}$$

Or, d'après le théorème 2.6, on a, en utilisant les coordonnées polaires,

$$\begin{aligned} f &= \lim_{\substack{\Omega \rightarrow +\infty \\ L^2(\mathbb{R}^n_x, dx)}} \frac{1}{(2\pi)^n} \left[ x \longmapsto \int_{\|\omega\| \leq \Omega} \widehat{f}(\omega) e^{i \langle x, \omega \rangle} d\omega \right] \\ &= \lim_{\substack{\Omega \rightarrow +\infty \\ L^2(\mathbb{R}^n_x, dx)}} \frac{1}{(2\pi)^n} \left[ x \longmapsto \int_{\xi \in \mathbb{S}^{n-1}(\mathbb{R})} \left( \int_0^\Omega \widehat{f}(\tau \xi) e^{i \langle x, \tau \xi \rangle} \tau^{n-1} d\tau \right) d_{\mathbb{S}^{n-1}}(\xi) \right] \\ &= \lim_{\substack{\Omega \rightarrow +\infty \\ L^2(\mathbb{R}^n_x, dx)}} \frac{1}{(2\pi)^n} \left[ x \longmapsto \int_{\xi \in \mathbb{S}^{n-1}} \left( \int_0^\Omega (R_\xi[\dot{f}])^\wedge(\tau) e^{i \langle x, \tau \xi \rangle} \tau^{n-1} d\tau \right) d_{\mathbb{S}^{n-1}}(\xi) \right] \end{aligned} \tag{2.48}$$

où nous avons convenu de noter  $\xi \longmapsto d_{\mathbb{S}^{n-1}}(\xi)$  la mesure de Lebesgue sur la sphère unité de  $\mathbb{R}^n$ . Le procédé d'inversion décrit dans (2.48) est un premier procédé, basé sur l'utilisation de la transformée de Fourier, pour recomposer la classe inconnue  $\dot{f}$  à partir de toutes les « coupes »  $R_\xi[\dot{f}]$ ,  $\xi \in \mathbb{S}^{n-1}(\mathbb{R})$ .

<sup>65</sup>La classe  $\dot{f}$  admet donc un représentant intégrable.

La seconde solution passe, elle, par l'utilisation de l'algorithmique hilbertienne, plus précisément de la proposition 1.16 de la section 1.11.3 du chapitre 1. Soit  $H$  le  $\mathbb{R}$ -espace des classes de fonctions mesurables sur la boule euclidienne  $B := B_n(0, R)$ , telles

$$\int_B |f(x)|^2 dx < +\infty$$

et, pour tout  $\xi \in \mathbb{S}^{n-1}(\mathbb{R})$ , soit

$$R_\xi : \dot{f} \in H \longmapsto R_\xi[\dot{f}] \in L^2_{\mathbb{R}}([-R, R], \frac{dp}{(1-p^2)^{(n-1)/2}})$$

qui à  $\dot{f}$  associe la classe de fonction

$$p \longmapsto \int_{\xi^\perp} f(p\xi + y) d_{\xi^\perp} y.$$

Il s'agit d'un opérateur continu entre les deux espaces de Hilbert réels que sont  $H$  et  $L^2([-R, R], dp)$ . Le noyau  $F_\xi$  de  $R_\xi$  est donc un  $\mathbb{R}$ -sous-espace vectoriel fermé de  $H$ . La projection d'une classe  $\dot{g}$  de  $H$  sur l'orthogonal  $F_\xi^\perp$  de ce noyau est la classe de représentant (défini presque partout)

$$x \in B_n(0, R) \longmapsto \frac{R_\xi[\dot{g}]}{R_\xi[\chi_{B_n(0, R)}]} (\langle x, \xi \rangle).$$

En effet on vérifie immédiatement que, si  $g$  est un représentant de  $\dot{g}$ ,

$$x \in B_n(0, R) \longmapsto g(x) - \frac{R_\xi[\dot{g}]}{R_\xi[\chi_{B_n(0, R)}]} (\langle x, \xi \rangle)$$

est le représentant d'une classe appartenant au noyau  $F_\xi$  de  $R_\xi$ ; d'autre part, pour toute classe  $\dot{h}$  (de représentant  $h$ ) dans ce noyau,

$$\begin{aligned} & \int_B \frac{R_\xi[\dot{g}]}{R_\xi[\chi_{B_n(0, R)}]} (\langle x, \xi \rangle) h(x) dx = \int_{\substack{|p|^2 + \|y\|^2 \leq R^2 \\ y \in \xi^\perp}} \frac{R_\xi[\dot{g}]}{R_\xi[\chi_{B_n(0, R)}]} (p) h(p\xi + y) dp d_{\xi^\perp} y \\ &= \int_{\substack{|p|^2 + \|y\|^2 \leq R^2 \\ y \in \xi^\perp}} h(p\xi + y) \left( \int_{\substack{|p|^2 + \|u\|^2 \leq R^2 \\ u \in \xi^\perp}} \frac{g(p\xi + u)}{R_\xi[\chi_{B_n(0, R)}](p)} d_{\xi^\perp} u \right) dp d_{\xi^\perp} y \\ &= \int_{\substack{|p|^2 + \|u\|^2 \leq R^2 \\ u \in \xi^\perp}} \frac{g(p\xi + u)}{R_\xi[\chi_{B_n(0, R)}](p)} \left( \int_{\substack{|p|^2 + \|y\|^2 \leq R^2 \\ y \in \xi^\perp}} h(p\xi + y) d_{\xi^\perp} y \right) dp d_{\xi^\perp} u \\ &= \int_{\substack{|p|^2 + \|u\|^2 \leq R^2 \\ u \in \xi^\perp}} \frac{g(p\xi + u)}{R_\xi[\chi_{B_n(0, R)}](p)} R_\xi[\dot{h}](p) dp d_{\xi^\perp} u = 0. \end{aligned}$$

Si l'on choisit une collection finie  $\{\xi_1, \dots, \xi_M\}$  de directions dans  $\mathbb{S}^{n-1}(\mathbb{R})$  et si  $P_1, \dots, P_M$  désignent les projections orthogonales respectivement sur les sous-espaces affines fermés  $F_1, \dots, F_M$ , où  $F_j := f + F_{\xi_j}$ ,  $j = 1, \dots, M$ ,  $f$  désignant un élément *a priori* inconnu de  $H$ , alors, pour tout  $\dot{g}$  dans  $H$ , la suite

$$\left( (P_M \circ P_{M-1} \circ \dots \circ P_1)^k(\dot{g}) \right)_{k \geq 1}$$

converge lorsque  $k$  tend vers  $+\infty$  vers la projection de  $\dot{g}$  sur le sous espace affine

$$f + \bigcap_{j=1}^M F_{\xi_j}.$$

En particulier, si l'on prend  $\dot{g} = \dot{0}$ , on a

$$\lim_{k \rightarrow +\infty} (P_M \circ P_{M-1} \circ \dots \circ P_1)^k (\dot{0}) \longrightarrow \text{Proj}_{f+(F_{\xi_1} \cap \dots \cap F_{\xi_M})} (\dot{0}),$$

ce qui fournit un moyen (itératif) d'approcher  $\dot{f}$  (modulo les éléments du sous-espace vectoriel  $F_{\xi_1} \cap \dots \cap F_{\xi_M}$ ) lorsque l'on dispose de la connaissance de représentants pour les classes  $R_{\xi_j}[\dot{f}]$ ,  $j = 1, \dots, M$  (la connaissance de toutes ces classes fournit le moyen de calculer de proche en proche tous les  $(P_M \circ P_{M-1} \circ \dots \circ P_1)^k (0)$ , on pourra le vérifier en exercice dans le cas  $n = 2$ ). Plus  $M$  est grand, plus le  $\mathbb{R}$ -sous-espace vectoriel  $F_{\xi_1} \cap \dots \cap F_{\xi_M}$  est petit <sup>66</sup> et par conséquent, meilleure est l'approximation de  $\dot{f}$  ainsi réalisée.

Il a fallu attendre les années 1970 <sup>67</sup> pour que les formules établies en 1911 par le mathématicien autrichien Johannes Radon pour inverser la transformée qui porte aujourd'hui son nom trouvent, avec les progrès de l'informatique et de l'algorithmique de la transformation de Fourier (et l'irruption de l'algorithme de *Fast Fourier Transform* proposé par J.W. Cooley et J.W. Tukey en 1965 et dont nous avons parlé dans la section 2.2.1), leur pleine réalisation au service de la tomographie médicale et de son instrument clef, le CAT-scanner<sup>68</sup>.

## 2.6.2 Un second exemple : l'extrapolation des signaux de spectre borné

Pour cette seconde illustration (qui peut encore être vue comme un thème de problème dirigé), on se place dans le cadre  $n = 1$ . Le  $\mathbb{C}$ -espace de Hilbert  $H = L^2_{\mathbb{C}}(\mathbb{R}_t, dt)$  contient deux types de sous-espaces fermés intéressants :

- d'une part, si  $T > 0$ , le  $\mathbb{C}$ -sous espace  $H^T$  constitué des classes  $\dot{f}$  ayant un représentant presque partout nul hors de  $[-T, T]$  ;
- d'autre part, si  $\Omega > 0$ , le  $\mathbb{C}$ -sous espace vectoriel  $H_{\Omega}$  constitué des classes  $\dot{f}$  telles que  $\mathcal{F}[\dot{f}]$  admette un représentant nul presque partout hors de  $[-\Omega, \Omega]$ .

Si  $\dot{f} \in H$ , la projection orthogonale de  $\dot{f}$  sur  $H^T$  a pour représentant la fonction  $f\chi_{[-T, T]}$ , où  $f$  est un représentant de  $\dot{f}$ .

Si  $\dot{f} \in H$ , la projection orthogonale de  $\dot{f}$  sur  $H_{\Omega}$  est

$$\mathcal{F}^{-1}[\widehat{f}\chi_{[-\Omega, \Omega]}],$$

où  $\widehat{f}$  est un représentant de  $\mathcal{F}[\dot{f}]$ . Il résulte du théorème 2.6 et de l'étude de l'exemple 2.12 qui l'accompagne qu'un représentant de cette projection est la fonction

$$t \in \mathbb{R} \longmapsto \frac{1}{\pi} \int_{\mathbb{R}} f(u) \frac{\sin(\Omega(t-u))}{t-u} du,$$

<sup>66</sup>Notons que les éléments de ce sous-espace sont les classes  $\dot{h}$  telles que  $R_{\xi_j}[\dot{h}] = \dot{0}$ ,  $j = 1, \dots, M$ .

<sup>67</sup>Avec l'attribution en 1979 du prix Nobel de médecine au physicien (et mathématicien) américain Allan M. Cormack et à l'informaticien et ingénieur électronicien anglais Godfrey N. Hounsfield pour leurs travaux en CAT-scanner tomographie.

<sup>68</sup>CAT pour Computed Axial Tomography.

où  $f$  est un représentant de  $\dot{f}$ .

Comme un élément  $\dot{f}$  de  $H_\Omega$  admet pour représentant la fonction

$$t \longmapsto \frac{1}{2\pi} \int_{-\Omega}^{\Omega} \widehat{f}(\omega) e^{i\omega t} d\omega$$

qui est une fonction réelle analytique et ne donc peut avoir (à moins d'être identiquement nulle) que des zéros isolés, on a  $H_\Omega \cap (H^T)^\perp = \{0\}$  puisque tout représentant d'un élément de  $(H^T)^\perp$  doit être presque partout nul hors de  $[-T, T]$ .

Nous sommes donc en situation de mettre en œuvre l'application de la proposition 1.16 de la section 1.11.3 du chapitre 1 : si  $\dot{f}_0 := \text{Proj}_{H^T}(\dot{f})$ , alors, par le jeu du mécanisme itératif

$$\dot{f}_{k+1} = \dot{f}_0 + (1 - \chi_{[-T, T]}) \mathcal{F}^{-1} \left[ \chi_{[-\Omega, \Omega]} \mathcal{F}[\dot{f}_k] \right], \quad k = 0, \dots,$$

on approche  $\dot{f}$  dans  $H$ , ce qui donne ainsi un procédé (théorique !) pour « extrapoler » le représentant continu de  $\dot{f}$  (supposé ici seulement connu sur  $[-T, T]$  puisque l'on ne dispose que de sa projection orthogonale sur  $H^T$ ) à  $\mathbb{R}$  tout entier. Ce procédé algorithmique, introduit dans les années 1970, nommé du fait de ses inventeurs procédé de Gerchberg-Papoulis<sup>69</sup>, même si bien sûr il est très difficile à réaliser de manière stable numériquement (les signaux à spectre borné relevant de l'utopie car tout signal est inévitablement bruité et le bruit ne saurait être à spectre borné), s'avère une bonne illustration du profit que l'on a à coupler les outils Fourier du chapitre 2 avec les notions d'analyse hilbertienne présentées au chapitre 1.

La présentation de ce dernier exemple, fédérant les deux chapitres de ce cours, le conclut tout naturellement. Elle souligne aussi combien les outils de l'algorithmique pythagoricienne et ceux de l'analyse de Fourier sont intimement complémentaires, le meilleur cadre pour travailler étant sans aucun doute celui des espaces de Hilbert, c'est-à-dire le cadre  $l^2$  ou  $L^2$ .

**FIN**

---

<sup>69</sup>A. Papoulis (1921-2002) est un théoricien du signal américain d'origine grecque contemporain ; ses ouvrages d'enseignement en analyse de Fourier, traitement du signal et processus stochastiques sont largement utilisés tant par la communauté mathématique que celle des ingénieurs ; ils ont inspiré la présentation du chapitre 2 de ce cours.