Des espaces Euclidiens aux espaces de Hilbert

Une introduction à l'analyse fonctionnelle et à ses applications

Edoardo Provenzi

Table des matières

Préambule **5**

1 Les espaces vectoriels avec produit scalaire (ou pre-Hilbertiens) **6**
1.1 Le produit scalaire réel et complexe 6
1.2 La norme associée à un produit scalaire (hilbertienne) et les espaces vectoriels normés 11
1.2.1 La formule du parallélogramme et de polarisation 13
1.3 Familles orthogonales et orthonormales dans des espaces vectoriels avec produit scalaire 14
1.4 Le théorème de Pythagore généralisé 15
1.5 Orthogonalité et indépendance linéaire 16
1.6 La projection orthogonale dans les espaces vectoriels avec produit scalaire . . . 18
1.7 Existence d'une base orthonormale : l'algorithme itératif de Gram-Schmidt . . 23
1.8 Les propriétés fondamentales d'une base orthonormale et orthogonale 24
1.9 Résumé du chapitre 1 30

2 La transformée de Fourier discrète et ses applications à la théorie des signaux et des images **32**
2.1 L'espace $\ell^2(\mathbb{Z}_N)$ et sa base canonique 32
2.1.1 La base orthogonale des exponentiels complexes de $\ell^2(\mathbb{Z}_N)$ 35
2.2 La base orthonormale de Fourier de $\ell^2(\mathbb{Z}_N)$ 39
2.3 La base orthogonale de Fourier de $\ell^2(\mathbb{Z}_N)$ 40
2.4 Les coefficients de Fourier et la transformée de Fourier discrète (DFT) 42
2.4.1 La transformée de Fourier inverse (IDFT) 44
2.4.2 Définition de DFT et IDFT avec la base orthonormale de Fourier 46
2.4.3 La base (orthonormale) de Fourier réelle 47
2.5 Interprétation matricielle de la DFT et IDFT 48
2.5.1 « FFT » : Fast Fourier Transform 50
2.6 La transformée de Fourier dans le traitement des signaux 51
2.6.1 La formule de synthèse des signaux 1D : décomposition sur la base des harmoniques 51
2.6.2 Signification des coefficients de Fourier et spectres d'un signal 1D 52
2.6.3 La formule de synthèse et les coefficients de Fourier de l'impulsion unitaire δ 53
2.6.4 Hautes et basses fréquences m dans la formule de synthèse 55
2.6.5 Filtrage de signaux dans la représentation fréquentielle 57

2.6.6 L'opérateur de multiplication et sa représentation matricielle diagonale . 58
2.6.7 Le multiplicateur de Fourier 59
2.7 Propriétés de la DFT 60
2.7.1 La périodicité de $\hat{z}$ et $\check{z}$ 60
2.7.2 DFT et translation (invariance du spectre par translations) 61
2.7.3 DFT et conjugaison 64
2.7.4 DFT et convolution 66
2.8 La DFT et les opérateurs stationnaires 70
2.8.1 La DFT et la diagonalisation des opérateurs stationnaires 71
2.8.2 Matrices circulantes 73
2.8.3 La caractérisation exhaustive des opérateurs stationnaires 74
2.8.4 Filtres passe-haut, passe-bas, passe-bande 78
2.8.5 La caractérisation des opérateurs stationnaires via les opérateurs de translation 79
2.8.6 Analyse fréquentielle des opérateurs de dérivation (discrète) première et seconde 80
2.9 La transformée de Fourier bidimensionnelle (DFT 2D) 83
2.9.1 Représentation matricielle de la DFT 2D : produit de Kronecker vs. itération de deux DFT 1D 86
2.9.2 Les propriétés de la DFT 2D 87
2.9.3 La DFT 2D et les opérateurs stationnaires 89
2.9.4 Les opérateurs gradient et Laplacien et leurs action sur les images numériques 91
2.9.5 Visualisation du spectre d'amplitude en 2D 92
2.9.6 Un exemple remarquable de filtrage d'une image numérique dans l'espace de Fourier : le floutage 94
2.10 Résumé du chapitre 2 95

3 Rappel sur la théorie de la mesure et de l'intégration à la Lebesgue 97
3.1 Riemann vs Lebesgue 97
3.2 Tribu (σ-algèbre), espace mesurable, mesure et espace mesuré 98
3.3 Fonctions mesurables et propriétés possédées presque partout (p.p) 100
3.4 Fonctions intégrables, intégrales à la Lebesgue 100
3.5 La caractérisation de la mesure de Lebesgue sur $\mathbb{R}$ et les ensembles de mesure de Lebesgue nulle 102
3.6 Les trois théorèmes d'opération de limite dans la théorie de l'intégration 104
3.7 Résumé du chapitre 3 105

4 Espaces de Banach et de Hilbert 106
4.1 La topologie métrique des espaces vectoriels avec produit scalaire 106
4.2 La continuité des opérations fondamentales des espaces vectoriels avec produit scalaire 111
4.2.1 L'équivalence des topologies séparées des espaces vectoriels de dimension finie 116
4.3 Suites de Cauchy et complétude : espaces de Banach et de Hilbert 117
4.3.1 La complétude des espaces vectoriel 121
4.3.2 Une caractérisation de la complétude d'espaces vectoriels normés via séries 122

4.3.3 Le théorème de point fixe de Banach . . . 126
4.4 Exemples remarquables d'espaces de Banach et de Hilbert . . . 130
4.4.1 Les espaces L^p et ℓ^p et leur complétude . . . 131
4.4.2 Les espaces L^∞ et ℓ^∞ et leur complétude . . . 139
4.4.3 Relations d'inclusion parmi les espaces ℓ^p . . . 143
4.4.4 Relations d'inclusion dans les espaces L^p . . . 145
4.4.5 Théorèmes de densité dans L^p(X,A,μ) . . . 147
4.5 Résumé du chapitre 4 . . . 151

5 La structure géométrique des espaces de Hilbert 152
5.1 Le complément orthogonal dans un espace de Hilbert et ses propriétés . . . 152
5.2 Le théorème de projection sur un convexe fermé et ses conséquences . . . 155
5.2.1 Une caractérisation des sous-espaces vectoriels fermés dans les espaces de Hilbert . . . 160
5.3 Sous-ensemble polaire et bipolaire d'un espace de Hilbert . . . 162
5.4 Le théorème de la projection (orthogonale) dans un espace de Hilbert . . . 164
5.5 Systèmes orthonormés et bases hilbertiennes . . . 166
5.5.1 L'inégalité de Bessel et les coefficients de Fourier . . . 167
5.5.2 Le théorème de Fischer - Riesz . . . 169
5.5.3 Les caractérisations d'une base hilbertienne . . . 171
5.5.4 Isomorphismes entre espaces de Hilbert . . . 175
5.5.5 $\ell^2(\mathbb{N},\mathbb{K})$ comme prototype des espaces de Hilbert séparables de dimension infinie . . . 177
5.6 La base hilbertienne de Fourier en L^2 . . . 177
5.6.1 $L^2[-\pi,\pi]$ ou $L^2[0,2\pi]$ . . . 178
5.6.2 $L^2(\mathbb{T})$ . . . 179
5.6.3 $L^2[a,b]$ . . . 180
5.6.4 Série de Fourier réelle . . . 181
5.6.5 Convergence ponctuelle de la série de Fourier réelle : le théorème de Dirichlet . . . 186
5.6.6 Le phénomène de Gibbs et les sommes de Cesàro . . . 187
5.6.7 La vitesse de convergence à 0 des coefficients de Fourier . . . 187
5.6.8 Transformée de Fourier en $L^2(\mathbb{T})$ et translation . . . 190
5.7 Résumé du chapitre 5 . . . 192

6 Les opérateurs linéaires bornés dans les espaces de Hilbert 194
6.1 Les propriétés fondamentales des opérateurs linéaires bornés entre espaces vectoriels normés . . . 196
6.1.1 La continuité des opérateurs linéaires définis sur un espace vectoriel normé de dimension finie . . . 199
6.2 La norme opératorielle, la convergence des suites d'opérateurs et les algèbres de Banach . . . 199
6.2.1 Un exemple (classique) d'opérateur linéaire non borné sur un espace vectoriel de dimension infinie . . . 208
6.3 Inversibilité des opérateurs linéaires . . . 209
6.4 L'espace dual d'un espace de Hilbert et le théorème de représentation de Riesz 213
6.4.1 Le produit scalaire induit sur le dual d'un espace de Hilbert . . . 217

6.5 Formes bilinéaires, sesquilinéaires et formes quadratiques associées 218
6.5.1 Le théorème de Lax-Milgram et ses conséquences 224
6.6 L'opérateur adjoint et ses propriétés 227
6.7 Les opérateurs de projection orthogonale dans un espace de Hilbert 234
6.7.1 Les opérateurs de multiplication bornés et leur relation avec les projecteurs orthogonaux 242
6.7.2 La réalisation géométrique des opérateurs de projection orthogonale via systèmes orthonormés 243
6.8 Les opérateurs isométriques et unitaires 249
6.8.1 Les caractérisations des opérateurs isométriques et unitaires 250
6.8.2 Relation entre opérateurs isométriques et unitaires et systèmes orthonormés 254
6.9 La transformée de Fourier sur $\mathcal{S}(\mathbb{R}^n)$, $L^1(\mathbb{R}^n)$ et $L^2(\mathbb{R}^n)$ 256
6.9.1 Un espace invariante par la transformée de Fourier : l'espace de Schwarz 257
6.9.2 Extension de la transformée de Fourier de $\mathcal{S}(\mathbb{R}^n)$ à $L^1(\mathbb{R}^n)$: le théorème de Riemann-Lebesgue 261
6.9.3 Extension de la transformée de Fourier à un opérateur unitaire sur $L^2(\mathbb{R}^n)$: la transformée de Fourier Plancherel 262
6.9.4 Relation entre transformée de Fourier-Plancherel et base hilbertienne de Hermite 264
6.9.5 La transformée de Fourier et la convolution 265
6.9.6 Convolution et transformée de Fourier en L^2 : le problème de la localisation de la transformée de Fourier 268
6.10 Le théorème d'échantillonnage de Shannon, Nyquist et Whittaker 269
6.10.1 La fréquence de Nyquist : « aliasing » et « oversampling » 271
6.11 Application de la transformée de Fourier à la résolution d'équations différentielles en dérivées ordinaires et partielles 271
6.11.1 Un exemple de EDO résolue avec la transformée de Fourier 271
6.11.2 Transformée de Fourier et EDPs 273
6.11.3 Résolution de l'EDP de la propagation de la chaleur avec la transformée de Fourier 273
6.12 Résumé du chapitre 6 276

Appendices **279**

A L'espace vectoriel quotient **280**

B L'opérateur transposé (ou dual) d'un opérateur linéaire **285**

C La convergence uniforme, forte et faible **287**
C.1 Convergence forte et faible dans un espace de Banach 288
C.2 Convergence forte et faible dans un espace de Hilbert 288
C.3 Convergence uniforme, forte et faible dans l'algèbre de Banach $\mathcal{B}(\mathcal{H})$ 289

Index **290**

Bibliographie **293**

Préambule

Le but de ce cours est d'introduire les concepts les plus importants de la théorie des espaces de Hilbert et des opérateurs définis sur ces espaces.

La raison pour laquelle on dédie un cours entier aux espaces de Hilbert est simple à comprendre : parmi les espaces vectoriels de dimension infinie, les espaces de Hilbert constituent l'objet mathématique le plus proche des espaces Euclidiens de dimension finie, i.e. $\mathbb{R}^n$ ou $\mathbb{C}^n$, où on développe l'analyse et l'algèbre linéaire classiques.

Lorsque l'on passe en dimension infinie, des subtilités de nature topologique imposent de rajouter des conditions (qui en dimension finie sont toujours vérifiées) pour pouvoir assurer la validité des résultats que l'on connaît dans les espaces Euclidiens. Pour les espaces de Hilbert, une de ces conditions topologiques est la complétude, i.e. le fait que toute suite de Cauchy soit convergente dans l'espace où elle est définie.

Cette considération montre que la théorie des espaces de Hilbert peut être pensée comme un mélange très élégant d'algèbre, analyse et topologie : ceci est l'héritage de grands mathématiciens du début du vingtième siècle, entre autres Riesz, Banach et, évidemment, Hilbert, qui ont mis en lumière l'exigence de ce mélange pour étendre les résultats classiques de l'algèbre et de l'analyse à la dimension infinie.

Une transformation linéaire particulièrement importante sera le fil rouge de ce cours : la transformée de Fourier. On examinera d'abord ses propriétés en dimension finie, avec ce que l'on appelle transformée de Fourier discrète, et on passera après à l'extension en dimension infinie, en considérant différents domaines pour cette transformation.

Bien assimiler les concepts introduits dans ce cours est essentiel pour l'évolution de la carrière d'un mathématicien dans toute direction, soit elle théorique ou appliquée, et permet de faire connaissance avec des techniques et des outils mathématiques développés dans une période particulièrement fertile et créative de l'histoire de la science. Ces techniques et ces outils restent toujours actuelles dans la recherche mathématique pure et appliquée.

L'auteur.

Chapitre 1

Les espaces vectoriels avec produit scalaire (ou pre-Hilbertiens)

Dans ce chapitre on va se concentrer sur les espaces vectoriels avec un produit scalaire, en particulier ceux de *dimension finie*.

1.1 Le produit scalaire réel et complexe

On commence par rappeler que, dans les espace Euclidiens réels $\mathbb{R}^2$ et $\mathbb{R}^3$, le produit scalaire entre deux vecteurs v, w est défini comme le nombre réel

$$v \bullet w = \langle v, w \rangle = \|v\| \|w\| \cos(\vartheta),$$

où ϑ est l'angle le plus petit entre v et w et $\| \ \|$ représente la norme (ou le module, en $\mathbb{R}^2$ et $\mathbb{R}^3$) des vecteurs.

On rappelle aussi que, à travers du produit scalaire, on peut définir la projection orthogonale du vecteur v dans la direction définie par le vecteur w. Il faut distinguer entre :

— la *projection scalaire* de v en direction w : $\|v\| \cos(\theta) = \frac{\langle v,w \rangle}{\|w\|}$;

— la *projection vectorielle* de v en direction w : $\|v\| \cos(\theta) \frac{w}{\|w\|} = \frac{\langle v,w \rangle}{\|w\|^2} w$,

où $\frac{w}{\|w\|}$ est le vecteur unitaire en direction de w. Le rôle de v et w peut, bien sûr, être échangé.

La valeur absolue de la projection scalaire *mesure la « similarité » entre la direction de deux vecteurs*. Pour comprendre ce concept, allons considérer deux positions relatives remarquables entre v et w :

— si v et w ont la même direction, alors l'angle ϑ entre eux est nul ou ça vaut π, et donc $\cos(\vartheta) = \pm 1$, i.e. la valeur absolue de la projection scalaire de v en direction w est maximale et ça vaut $\|v\|$;

— au contraire, si v et w sont perpendiculaires, alors $\vartheta = \frac{\pi}{2}$ ou $\vartheta = \frac{3}{2}\pi$ et donc $\cos(\vartheta) = 0$, ce qui montre que la projection scalaire de v en direction w est nulle.

Lorsque v a une position relative a w intermédiaire entre les deux ci-avant, la valeur absolue de la projection scalaire de v en direction de w bascule entre la norme de v et 0, ce qui explique pourquoi sa valeur est considérée comme une mesure de similarité entre direction de vecteurs.

Dans le cours, on considérera des espaces vectoriels bien plus compliqués que $\mathbb{R}^2$ et $\mathbb{R}^3$ et la mesure de similarité entre vecteurs donnée par la projection nous donnera des informations encore plus importantes de la cohérence entre directions.

Avant d'arriver à obtenir ces informations, la première chose que nous devons faire est de passer des espaces Euclidiens $\mathbb{R}^2$ et $\mathbb{R}^3$ à des espaces vectoriels abstraits. La définition générale de produit scalaire et de projection orthogonale dans ces espaces peut être pensée comme une extension des définitions précédentes pour des espaces où notre capacité de représentation des vecteurs devient impossible.

Les propriétés géométriques (dont l'intuition et, surtout, la visualisation, est possible seulement en dimension 2 ou 3) seront remplacées par un ensemble de propriétés algébriques qui ont l'avantage de pouvoir être utilisées en dimension quelconque sans difficulté.

Ces propriétés algébriques doivent, bien sûr, être celles nécessaires et suffisantes pour caractériser le produit scalaire entre vecteurs du plan ou de l'espace réel. Cette forme de généraliser des concepts « intuitifs » en dimension 2 et 3 est classique dans les mathématiques.

Dans ce chapitre, le symbole V sera utilisé pour décrire un espace vectoriel défini sur le corps $\mathbb{K}$, où $\mathbb{K}$ sera soit $\mathbb{R}$, soit $\mathbb{C}$, et de dimension finie $n < +\infty$. Le corps $\mathbb{K}$ contient les *scalaires* avec lesquels on construit les combinaisons linéaires entre vecteurs de V. On rappelle que deux espaces vectoriel de dimension finie sont isomorphes si et seulement s'ils ont la même dimension. De plus, si on fixe une base $B = (b_1, \ldots, b_n)$ de V, alors on peut construire un *isomorphisme* entre V et $\mathbb{K}^n$ comme ça :

$$\begin{array}{rccc} I: & V & \longrightarrow & \mathbb{K}^n \\ & v = [v]_B = \sum_{i=1}^{n} v_i b_i & \longmapsto & \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \end{array}$$

i.e. I associe à chaque $v \in V$ le vecteur de $\mathbb{K}^n$ donné par les composantes scalaires de v par rapport à la bases fixée B. Comme I est un isomorphisme, il suit que $\mathbb{K}^n$ **est le prototype de tous les espaces vectoriels de dimension** n **sur le corps** $\mathbb{K}$.

Déf. 1.1.1 ($\mathbb{K}$-forme) *Soit V un espace vectoriel défini sur un corps $\mathbb{K}$.*

Une $\mathbb{K}$-forme sur V est une application définie sur $V \times V$ à valeurs en $\mathbb{K}$, i.e.

$$\begin{array}{rccc} \phi: & V \times V & \longrightarrow & \mathbb{K} \\ & (v, w) & \longmapsto & \phi(v, w). \end{array}$$

Déf. 1.1.2 *Soit V un espace vectoriel <u>réel</u>. On dit que le couple $(V, \langle , \rangle)$ est un* ***espace vectoriel réel avec produit scalaire*** *(ou un* ***espace pré-Hilbertien réel****) si $\langle , \rangle$ est une forme :*

1. ***bilinéaire****, i.e.*[1] *linéaire relativement à chaque argument (l'autre étant fixé) :*

$$\langle v_1 + v_2, w_1 + w_2 \rangle = \langle v_1, w_1 \rangle + \langle v_1, w_2 \rangle + \langle v_2, w_1 \rangle + \langle v_2, w_2 \rangle, \quad \forall\, v_1, v_2, w_1, w_2 \in V$$

et

$$\langle \alpha v, \beta w \rangle = \alpha \langle v, \beta w \rangle = \beta \langle \alpha v, w \rangle = \alpha\beta \langle v, w \rangle, \quad \forall \alpha, \beta \in \mathbb{R},\ v, w \in V;$$

1. i.e. est l'abréviation de « id est », l'expression latine équivalente à « c'est-à-dire ». Cette écriture est très utilisée dans les livres de mathématiques.

2. ***symétrique*** : $\langle v, w\rangle = \langle w, v\rangle$, $\forall v, w \in V$;
3. ***définie*** : $\langle v, v\rangle = 0 \iff v = 0_V$, *le vecteur nul de l'espace vectoriel* V ;
4. ***positive*** : $\langle v, v\rangle > 0$ $\forall v \in V$, $v \neq 0_V$.

Certains auteurs utilisent la définition **strictement positive** pour une forme définie positive.

Avec un moment de réflexion, nous pouvons comprendre que, pour une forme réelle sur V, *demander la symétrie et la bilinéarité équivaut à demander la symétrie et la linéarité à gauche*, i.e. :

$$\langle v_1 + v_2, w\rangle = \langle v_1, w\rangle + \langle v_2, w\rangle, \quad \langle \alpha v, w\rangle = \alpha\langle v, w\rangle, \quad \forall v, w \in V,\ \forall \alpha \in \mathbb{R}.$$

L'exemple le plus simple et important de produit scalaire réel est le **produit scalaire Euclidien réel**, qui est défini ainsi : soient $v = (v_1, v_2, \ldots, v_n)$, $w = (w_1, w_2, \ldots, w_n)$ deux vecteurs de $\mathbb{R}^n$ écrits avec leurs composantes par rapport à une base $(b_i)_{i=1}^n$ de $\mathbb{R}^n$ quelconque mais fixée, alors le produit scalaire Euclidien réel entre v et w est :

$$\langle v, w\rangle \equiv \sum_{i=1}^{n} v_i w_i = v^t \cdot w = v \cdot w^t,$$

où, dans les deux dernières équations, on a écrit avec v^t et w^t les vecteurs transposés de v et w, et donc on revient à un produit matriciel entre un vecteur ligne (à interpréter comme une matrice $1 \times n$) et un vecteur colonne (à interpréter comme une matrice $n \times 1$).

L'extension de ces définitions à des espaces vectoriels complexes n'est pas triviale. On commence par observer que, si V est un espace vectoriel *complexe*, alors *il ne peut pas exister une transformation bilinéaire et définie positive* su $V \times V$. En fait, dans ce cas, on aurait, pour tout vecteur $v \in V$,

$$\langle iv, iv\rangle = i^2\langle v, v\rangle = -\langle v, v\rangle \leqslant 0 \ \text{ car } \langle v, v\rangle \geqslant 0 \ \text{ par positivité.}$$

On verra que, pour définir une norme (et, ainsi, une distance, et donc une topologie) à partir d'un produit scalaire complexe, la propriété de positivité est indispensable. Donc, si on veut une structure algébrique de produit scalaire complexe compatible avec une structure topologique, il faut renoncer à la bilinéarité.

Afin de chercher une alternative à la bilinéarité, rappelons que, si V et W sont deux espaces vectoriels sur le même corps $\mathbb{K}$ et $\tau, \sigma : \mathbb{K} \to \mathbb{K}$ est un automorphisme de $\mathbb{K}$, i.e. une application bijective de $\mathbb{K}$ en $\mathbb{K}$, alors une forme $\phi : V \times W \to \mathbb{K}$ est dite (τ, σ)-linéaire si ϕ est additive et si $\phi(\alpha v, \beta w) = \tau(\alpha)\sigma(\beta)\phi(v, w)$.

$\mathbb{R}$ a seulement un automorphisme, l'identité $id_{\mathbb{R}}$ et l'$id_{\mathbb{R}}$-linéarité par rapport aux deux arguments de ϕ est la bilinéarité usuelle. Par contre, $\mathbb{C}$ a une infinité d'automorphismes, néanmoins, si on demande des conditions minimales de régularité sur ces automorphismes, par exemple la continuité ou simplement la mesurabilité, ça restent que deux automorphismes de $\mathbb{C}$: l'identité $id_{\mathbb{C}}$ et la conjugaison complexe $\tau(z) = \overline{z}$. L' $id_{\mathbb{C}}$-linéarité dans les deux arguments de ϕ est la bilinéarité usuelle, tandis que la τ-linéarité par rapport aux deux arguments de ϕ est dite antilinéarité [2], i.e. $\phi(\alpha v, \beta w) = \bar{\alpha}\bar{\beta}\phi(v, w)$, ou, si on utilise la notation du produit scalaire :

$$\langle \alpha v, \beta w\rangle = \bar{\alpha}\bar{\beta}\langle v, w\rangle.$$

2. Les symboles z^* et $\bar{z}$ représentent la **conjugaison complexe**, i.e. si $z \in \mathbb{C}$, $z = a + ib$, $a, b \in \mathbb{R}$, alors $z^* = \bar{z} = a - ib$. On rappelle que $\sum_{k=1}^n \overline{z_k} = \overline{\sum_{k=1}^n z_k}$, $\prod_{k=1}^n \overline{z_k} = \overline{\prod_{k=1}^n z_k}$, $|z|^2 = z\bar{z}$ et $z = \bar{z}$ si et seulement si $z \in \mathbb{R}$.

Il est clair que l'antilinéarité a le même problème d'incompatibilité avec la définie positivité de la bilinéarité, en fait $\langle iv, iv\rangle = (-i)(-i)\langle v, v\rangle = i^2\langle v, v\rangle = -\langle v, v\rangle^2 \leqslant 0$.

La dernière option qui reste est dite **sesquilinéarité** [3], i.e. le mélange des deux automorphismes de $\mathbb{C}$:

$$\begin{cases} \phi(\alpha v, \beta w) = id_{\mathbb{C}}(\alpha)\tau(\beta)\phi(v, w) = \alpha\bar{\beta}\phi(v, w) \\ \text{ ou} \\ \phi(\alpha v, \beta w) = \tau(\alpha)id_{\mathbb{C}}(\beta)\phi(v, w) = \bar{\alpha}\beta\phi(v, w). \end{cases}$$

Les deux choix sont compatibles avec la définie positivité, par conséquent, la sesquilinéaire est la seule alternative (régulière) à la bilinéarité pour une forme sur deux espaces vectoriels complexes compatible avec la définie positivité.

Le choix de la variable linéaire et antilinéaire est totalement arbitraire. L'antilinéarité à droite est la *convention* utilisée parmi les *mathématiciens*, tandis que l'antilinéarité à gauche est la *convention* typiquement adoptée parmi les *physiciens*.

Pour uniformiser la notation, on considèrera dorénavant toujours la convention utilisée par les mathématiciens, i.e., si on utilise la notation du produit scalaire : $\langle \alpha v, \beta w\rangle = \alpha\bar{\beta}\langle v, w\rangle$.

Maintenant il est important d'observer que *la sesquilinéarité n'est pas compatible avec la symétrie*, en fait, si les deux propriétés sont vérifiées, alors $\langle v, \alpha w\rangle = \bar{\alpha}\langle v, w\rangle$, mais aussi $\langle v, \alpha w\rangle = \langle \alpha w, v\rangle = \alpha\langle w, v\rangle = \alpha\langle v, w\rangle$. Donc $\langle v, \alpha w\rangle = \bar{\alpha}\langle v, w\rangle = \alpha\langle v, w\rangle$ et cela peut être vrai si et seulement si $\alpha \in \mathbb{R}$.

Par conséquent, $\langle , \rangle$ ne peut pas être simultanément sesquilinéaire et symétrique si on veut travailler avec des vecteurs qui appartiennent à un espace vectoriel complexe.

L'exemple qu'on vient d'examiner montre que, au lieu de la symétrie, on doit demander la propriété suivante pour toute couple de vecteurs v, w : $\langle v, w\rangle = \overline{\langle w, v\rangle}$, i.e. que l'échange de l'ordre des vecteurs dans $\langle , \rangle$ soit équivalent à la conjugaison complexe. Une transformation qui a cette propriété est dite *hermitienne* [4].

Ces observations justifient complètement la définition suivante.

Déf. 1.1.3 *Soit V un espace vectoriel complexe. On dit que le couple $(V, \langle , \rangle)$ est un* ***espace vectoriel complexe avec produit scalaire*** *(ou un* ***espace pré-Hilbertien complexe****) si $\langle , \rangle$ est une forme complexe :*

1. ***sesquilinéaire*** *:*

$$\langle v_1 + v_2, w_1 + w_2\rangle = \langle v_1, w_1\rangle + \langle v_1, w_2\rangle + \langle v_2, w_1\rangle + \langle v_2, w_2\rangle$$

$\forall\ v_1, v_2, w_1, w_2 \in V$*, et*

$$\xrightarrow[\text{Linéarité à gauche}]{\text{Anti-linéarité à droite}} \langle \alpha v, \beta w\rangle = \alpha\langle v, \beta w\rangle = \bar{\beta}\langle \alpha v, w\rangle = \alpha\bar{\beta}\langle v, w\rangle$$

$\forall\ \alpha, \beta \in \mathbb{C}$, $\forall\ v, w \in V$ *;*

2. ***hermitienne*** *:* $\langle v, w\rangle = \overline{\langle w, v\rangle}$, $\forall v, w \in V$ *;*
3. ***définie*** *:* $\langle v, v\rangle = 0 \iff v = 0_V$*, le vecteur nul de l'espace vectoriel V ;*
4. ***positive*** *:* $\langle v, v\rangle > 0\ \forall v \in V$, $v \neq 0_V$.

3. Sesqui vient du latin *semisque*, qui veut dire *une fois et demi*, on utilise cette nomenclature pour souligner qu'on n'a pas deux fois la linéarité, mais une fois « et demi », dû à la présence de la conjugaison complexe.

4. Charles Hermite, mathématicien français, 1822 Dieuze - 1901 Paris.

Comme dans le cas du produit scalaire réel, pour une forme complexe sur V, *demander la symétrie et la sesquilinéairité équivaut à demander la propriété hermitienne et la linéarité à gauche*, en fait, si ces propriétés sont valides, alors :

$$\langle v, \alpha w\rangle = \overline{\langle \alpha w, v\rangle} = \overline{\alpha\langle w, v\rangle} = \bar{\alpha}\overline{\langle w, v\rangle} = \bar{\alpha}\overline{\overline{\langle v, w\rangle}} = \bar{\alpha}\langle v, w\rangle, \quad \forall \alpha \in \mathbb{C}.$$

Si on considère la somme de n vecteurs au lieu de deux, la sesquilinéarité est représentée par les formules suivantes :

$$\boxed{\langle \sum_{i=1}^{n} \alpha_i v_i \, , \, w\rangle = \sum_{i=1}^{n} \alpha_i \langle v_i, w\rangle}\,; \tag{1.1}$$

$$\boxed{\langle v \, , \, \sum_{i=1}^{n} \alpha_i w_i\rangle = \sum_{i=1}^{n} \overline{\alpha_i}\langle v, w_i\rangle}\,. \tag{1.2}$$

En $\mathbb{C}^n$, le **produit scalaire Euclidien complexe** est défini ainsi :

$$\langle v, w\rangle \equiv \sum_{i=1}^{n} v_i \overline{w_i} = v \cdot (\overline{w})^t = v^t \cdot \overline{w},$$

où $v = (v_1, v_2, \ldots, v_n), w = (w_1, w_2, \ldots, w_n) \in \mathbb{C}^n$ sont écrits avec leurs composantes par rapport à une base $(b_i)_{i=1}^n$ quelconque mais fixée de $\mathbb{C}^n$.

Dorénavant, le symbole $\mathbb{K}$ sera utilisé pour représenter soit $\mathbb{R}$ soit $\mathbb{C}$ quand on veut discuter des propriétés valides indépendamment de la réalité ou complexité du produit scalaire.

Théorème 1.1.1 *Soit $(V, \langle\ ,\ \rangle)$ un espace vectoriel avec produit scalaire. Alors :*

1. $\langle v, 0_V\rangle = 0\ \forall v \in V$;
2. *Si $\langle u, w\rangle = \langle v, w\rangle\ \forall w \in V$, nécessairement u et v coïncident.*
3. $\langle v, w\rangle = 0\ \forall v \in V \iff w = 0_V$, *i.e. le seul vecteur orthogonal à tous les autres vecteurs est le vecteur nul.*

Preuve.

1. $\langle v, 0_V\rangle = \langle v, 0_V + 0_V\rangle = \langle v, 0_V\rangle + \langle v, 0_V\rangle$ par linéarité, i.e. $\langle v, 0_V\rangle - \langle v, 0_V\rangle = 0 = \langle v, 0_V\rangle$.

2. $\langle u, w\rangle = \langle v, w\rangle\ \forall w \in V$ implique, par linéarité, que $\langle u - v, w\rangle = 0\ \forall w \in V$ et donc, en particulier, si on considère $w = u - v$, on obtient $\langle u - v, u - v\rangle = 0$, ce qui implique, par la définie positivité du produit scalaire, que $u - v = 0_V$, i.e. $u = v$.

3. Si $w = 0_V$, alors $\langle v, w\rangle = 0\ \forall v \in V$ grâce à **1.** Inversement, par hypothèse ça vaut que $\langle v, w\rangle = 0 = \langle v, 0_V\rangle\ \forall v \in V$, mais alors la propriété **2.** implique que $w = 0_V$. □

Terminons avec une propriété typique du produit scalaire complexe, qui descend directement d'une propriété des nombre complexes.

Théorème 1.1.2 *Soit $(V, \langle\ ,\ \rangle)$ un espace vectoriel avec produit scalaire complexe. Alors :*

$$\Im(\langle v, w\rangle) = \Re(\langle v, iw\rangle) \qquad \forall v, w \in V.$$

Preuve. Considérons un nombre complexe quelconque $z = a + ib$, alors $-iz = b - ia$, donc $b = \Im(z) = \Re(-iz)$. Si, maintenant, $z = \langle v, w\rangle$, alors $\Im(\langle v, w\rangle) = \Re(-i\langle v, w\rangle) = \Re(\langle v, iw\rangle)$ par sesquilinéarité. □

1.2 La norme associée à un produit scalaire (hilbertienne) et les espaces vectoriels normés

Si $(V, \langle , \rangle)$ est un espace vectoriel avec produit scalaire sur $\mathbb{K}$, alors on peut définir une norme sur V comme ça :

$$\begin{array}{rccl} \| \ \| : & V & \to & \mathbb{R}_0^+ = [0, +\infty) \\ & v & \to & \|v\| = \sqrt{\langle v, v \rangle}. \end{array}$$

On note que $\|v\|$ est bien définie car $\langle v, v \rangle \geqslant 0 \ \forall v \in V$. Une fois qu'on a une norme, on peut toujours définir une *distance* entre deux vecteurs v, w de V : $d(v, w) = \|v - w\|$.

On appelle un vecteur $v \in V$ tel que $\|v\| = 1$ un **vecteur unitaire**. Chaque vecteur $v \in V$ peut être **normalisé** pour devenir un vecteur unitaire simplement via la division par sa norme. Exemples remarquables :

$$(\mathbb{R}^n, \langle , \rangle) : \|v\| = \sqrt{\sum_{i=1}^{n} v_i^2} \ ;$$

$$(\mathbb{C}^n, \langle , \rangle) : \|v\| = \sqrt{\sum_{i=1}^{n} v_i \overline{v_i}} = \sqrt{\sum_{i=1}^{n} |v_i|^2}.$$

On résume dans la liste ci-dessous trois propriétés de la norme qui devraient déjà être connues. $\forall v, w \in V$, $\forall \alpha \in \mathbb{K}$ ça vaut :

1. $\|v\| \geqslant 0, \ \|v\| = 0 \iff v = 0_V$
2. $\|\alpha v\| = |\alpha| \|v\|$ (Homogénéité)
3. $\|v + w\| \leqslant \|v\| + \|w\|$ (Inégalité Triangulaire).

Déf. 1.2.1 (Espace vectoriel normé) *On appelle espace vectoriel normé le couple $(V, \| \ \|)$ donné par un espace vectoriel V et une fonction, dite norme $\| \ \| : V \to \mathbb{R}_0^+$, qui satisfait les trois propriétés ci-dessus.*

La norme $\| \ \|$ est une norme hilbertienne s'il existe un produit scalaire $\langle \ , \ \rangle$ sur V tel que $\|v\| = \sqrt{\langle v, v \rangle} \ \forall v \in V$.

Un espace vectoriel avec produit scalaire est donc, canoniquement, un espace vectoriel normé. Des contre-exemples montrent que le vice-versa n'est pas vrai en général.

On observe que, par définition, $\langle v, v \rangle = \|v\| \, \|v\|$, mais, en général, le module du produit scalaire est dominé par le produit des normes. Cela est assuré par la célèbre inégalité suivante.

Théorème 1.2.1 (Inégalité de Cauchy-Schwarz) *Pour tout $v, w \in (V, \langle \ , \ \rangle)$ ça vaut :*

$$\boxed{|\langle v, w \rangle| \leqslant \|v\| \|w\|}.$$

Preuve. Il y a des dizaines de preuves de l'inégalité de Cauchy-Schwarz. Nous allons détailler ci-dessous l'une des plus élégantes et, après, la plus simple qui existe.

Première preuve :

Si $w = 0_V$, alors l'inégalité est trivialement vérifiée avec $0 = 0$. Si $w \neq 0_V$, alors on peut définir $z = v - \frac{\langle v,w\rangle}{\|w\|^2} w$, i.e. $v = \frac{\langle v,w\rangle}{\|w\|^2} w + z$, et observer que

$$\langle z, w\rangle = \langle v - \frac{\langle v, w\rangle}{\|w\|^2} w, w\rangle = \langle v, w\rangle - \frac{\langle v, w\rangle}{\cancel{\|w\|^2}} \cancel{\langle w, w\rangle} = 0.$$

Donc,

$$\begin{aligned}
\|v\|^2 = \langle v, v\rangle &= \left\langle \frac{\langle v, w\rangle}{\|w\|^2} w + z, \frac{\langle v, w\rangle}{\|w\|^2} w + z \right\rangle \\
&= \frac{\langle v, w\rangle}{\|w\|^2} \left\langle w, \frac{\langle v, w\rangle}{\|w\|^2} w + z \right\rangle + \left\langle z, \frac{\langle v, w\rangle}{\|w\|^2} w + z \right\rangle \\
&= \frac{\langle v, w\rangle}{\|w\|^2} \frac{\overline{\langle v, w\rangle}}{\cancel{\|w\|^2}} \cancel{\langle w, w\rangle} + \frac{\langle v, w\rangle}{\|w\|^2} \langle w, z\rangle + \frac{\overline{\langle v, w\rangle}}{\|w\|^2} \langle z, w\rangle + \langle z, z\rangle \\
&= \frac{|\langle v, w\rangle|^2}{\|w\|^2} + \|z\|^2,
\end{aligned}$$

car les deux termes intermédiaires dans l'avant dernière étape sont nuls vu que $\langle z, w\rangle = \overline{\langle w, z\rangle} = 0$.

Comme $\|z\|^2 \geqslant 0$, on a que

$$\|v\|^2 = \frac{|\langle v, w\rangle|^2}{\|w\|^2} + \|z\|^2 \geqslant \frac{|\langle v, w\rangle|^2}{\|w\|^2},$$

i.e. $|\langle v, w\rangle|^2 \leqslant \|v\|^2 \|w\|^2$, et alors $|\langle v, w\rangle| \leqslant \|v\| \|w\|$.

Deuxième preuve (1 ligne !) : $\forall t \in \mathbb{R}$ ça vaut

$$0 \leqslant \|tv - w\|^2 = \langle tv - w, tv - w\rangle = t^2 \|v\|^2 - 2t\langle v, w\rangle + \|w\|^2 \iff \Delta = 4\langle v, w\rangle^2 - 4\|v\|^2 \|w\|^2 \leqslant 0.$$

□

L'inégalité de Cauchy-Schwarz permet de généraliser aux espaces vectoriels abstraits le concept d'angle entre deux vecteurs, en fait elle implique qu'il existe un coefficient k entre -1 et $+1$ tel que $\langle v, w\rangle = \|v\| \|w\| k$, mais alors, vu que la restriction du cos à $[0, \pi]$ réalise une bijection avec $[-1, 1]$, cela veut dire qu'il existe un seul $\vartheta \in [0, \pi]$ tel que $\langle v, w\rangle = \|v\| \|w\| \cos \vartheta$. $\vartheta \in [0, \pi]$ est dit **l'angle entre les deux vecteurs**.

Une autre propriété très importante de la norme est la suivante.

Théorème 1.2.2 *Soit $(V, \| \ \|)$ un espace vectoriel normé quelconque et $v, w \in V$, alors :*

$$\boxed{|\|v\| - \|w\|| \leqslant \|v - w\|}. \tag{1.3}$$

Preuve. D'un côté :

$$\|v\| = \|v - w + w\| = \|(v - w) + w\| \leqslant \|v - w\| + \|w\|,$$

par l'inégalité triangulaire, et alors $\|v\| - \|w\| \leqslant \|v - w\|$. De l'autre côté :

$$\|w\| = \|w - v + v\| = \|(w - v) + v\| \leqslant \|w - v\| + \|w\|,$$

et alors $\|w\| - \|v\| \leqslant \|v - w\|$, i.e. $\|v\| - \|w\| \geqslant -\|v - w\|$.
Donc, $-\|v - w\| \leqslant \|v\| - \|w\| \leqslant \|v - w\|$, i.e. $|\|v\| - \|w\|| \leqslant \|v - w\|$. □

Allons maintenant présenter une formule très utile.

Théorème 1.2.3 (Théorème de Carnot) *Soient $v, w \in (V, \langle\ ,\ \rangle)$, alors :*

$$\boxed{\|v \pm w\|^2 = \|v\|^2 + \|w\|^2 \pm 2\langle v, w\rangle, \quad (\mathbb{K} = \mathbb{R})}, \tag{1.4}$$

et

$$\boxed{\|v \pm w\|^2 = \|v\|^2 + \|w\|^2 \pm \langle v, w\rangle \pm \langle w, v\rangle, \quad (\mathbb{K} = \mathbb{C})}. \tag{1.5}$$

Preuve. Calcul direct :

$$\|v \pm w\|^2 = \langle v \pm w, v \pm w\rangle = \langle v, v\rangle \pm \langle v, w\rangle \pm \langle w, v\rangle + \langle w, w\rangle = \|v\|^2 + \|w\|^2 \pm \langle v, w\rangle \pm \langle w, v\rangle.$$

□

Si $\mathbb{K} = \mathbb{C}$, comme $\langle w, v\rangle = \overline{\langle v, w\rangle}$, et comme, si $z = a + ib = \Re(z) + i\Im(z)$, alors $z + \bar{z} = 2a = 2\Re(z)$, on peut réécrire (1.5) comme ça :

$$\boxed{\|v \pm w\|^2 = \|v\|^2 + \|w\|^2 \pm 2\Re(\langle v, w\rangle)}. \tag{1.6}$$

Les formules qu'on a examiné dans cette section ont des conséquences immédiates, qu'on va mettre en avant dans la sous-section suivante à cause de leur importance.

1.2.1 La formule du parallélogramme et de polarisation

En $\mathbb{R}^2$ c'est bien connue la formule du parallélogramme représentée dans la figure 1.1. On peut généraliser cette formule dans un espace vectoriel avec produit scalaire quelconque.

Théorème 1.2.4 (Formule du parallélogramme) *Soit $(V, \langle, \rangle)$ un espace vectoriel sur $\mathbb{K}$ avec produit scalaire. Alors, $\forall v, w \in V$:*

$$\|v + w\|^2 + \|v - w\|^2 = 2(\|v\|^2 + \|w\|^2).$$

Preuve. Conséquence directe de la formule 1.4 ou (1.5) en considérant $\|v + w\|^2$ et après $\|v - w\|^2$. □

Nous avons vu qu'un produit scalaire induit une norme. L'importante formule de polarisation permet d'« inverser » les rôles pour écrire le produit scalaire à travers de la norme.

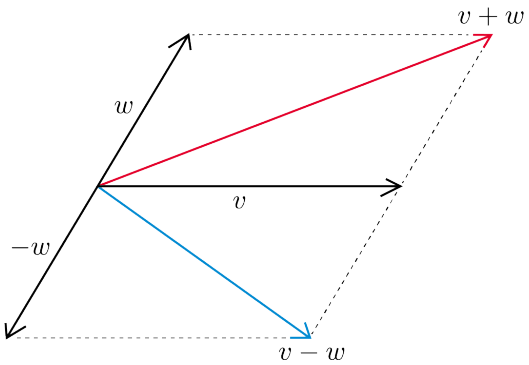


FIGURE 1.1 – La formule du parallélogramme en $\mathbb{R}^2$: la somme des carrés des normes des diagonales indiquées est égale à deux fois la somme des normes de v et w.

Théorème 1.2.5 (Formule de polarisation) *Soit $(V, \langle , \rangle)$ un espace vectoriel sur $\mathbb{K}$ avec produit scalaire. Alors, $\forall v, w \in V$:*

$$\langle v, w\rangle = \frac{1}{4}\left(\|v+w\|^2 - \|v-w\|^2\right), \quad (\mathbb{K} = \mathbb{R})$$

et

$$\langle v, w\rangle = \frac{1}{4}\left[\|v+w\|^2 - \|v-w\|^2 + i\left(\|v+iw\|^2 - \|v-iw\|^2\right)\right], \quad (\mathbb{K} = \mathbb{C}).$$

Preuve. La formule est une conséquence directe de la formule (1.4) dans le cas réel. Pour le cas complexe il faut remplacer w par iw dans la formule (1.5), par sesquilinéairité on obtient :

$$\|v \pm iw\|^2 = \|v\|^2 + \|w\|^2 \mp i\langle v, w\rangle \pm i\langle w, v\rangle.$$

Par calcul direct, on vérifie que $\|v+w\|^2 - \|v-w\|^2 + i\,\|v+iw\|^2 - i\,\|v-iw\|^2 = 4\langle v, w\rangle$. □

Observation : la formule (1.4) et la preuve de la formule de polarisation utilisent seulement la bilinéarité ou la sesquilinéarité de la forme $\langle\ ,\ \rangle$, la symétrie et la définie positivité n'interviennent nulle part. Cette observation sera utilisée dans la section 6.5.

Il est surprenant qu'une formule si simple comme celle du parallélogramme permet de donner une condition nécessaire et suffisante pour garantir qu'une norme sur un espace vectoriel est induite par un produit scalaire, i.e. elle est une norme hilbertienne. Nous formaliserons ça dans le chapitre 4.

1.3 Familles orthogonales et orthonormales dans des espaces vectoriels avec produit scalaire

La définition « géométrique » de produit scalaire en $\mathbb{R}^2$ et $\mathbb{R}^3$ montre qu'il est nul si et seulement si ϑ, l'angle entre les vecteur, est $\pi/2$, car dans ce deux cas $\cos(\theta) = 0$.
Dans les espaces vectoriels plus compliqués (par exemple les espaces de polynômes) ou même dans les espaces vectoriels Euclidiens de dimension supérieure à 3, on ne peut plus visualiser les vecteurs et donc on doit *axiomatiser* leur orthogonalité à travers de la nullité de leur produit scalaire.

Déf. 1.3.1 *Soit $(V, \langle , \rangle)$ un espace vectoriel avec produit scalaire réel ou complexe de dimension finie n. Soit $F = \{v_1, \cdots, v_n\}$ une famille de vecteurs de V, alors :*

- *F est une **famille orthogonale** de vecteurs si chaque couple de vecteurs différents a pour produit scalaire 0 : $\langle v_i, v_j \rangle = 0$;*
- *F est une **famille orthonormale** si elle est orthogonale et, en plus $\|v_i\| = 1 \ \forall i$. Donc, si $\{v_i\}_{i=1}^n$ est une famille orthogonale, $\{u_i = \|v_i\|^{-1} v_i\}_{i=1}^n$ est une famille orthonormale.*

Une famille orthonormale (vecteurs unitaires et orthogonaux entre eux) peut être caractérisé comme ça :

$$\boxed{\langle v_i, v_j \rangle = \delta_{i,j} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \qquad \textbf{Famille orthonormale}}$$

$\delta_{i,j}$ est dit **symbole de Kronecker** [5]**.**

1.4 Le théorème de Pythagore généralisé

On peut déterminer une généralisation très utile du théorème de Pythagore dans les espaces vectoriels abstraits avec produit scalaire. Pour arriver à la formulation générale, on nécessite une lemme.

Lemme 1.4.1 *Soit $(V, \langle , \rangle)$ un espace vectoriel avec produit scalaire réel ou complexe. Soit $u \in V$ orthogonal à tous les vecteurs $v_1, \ldots, v_n \in V$. Alors u est orthogonal aussi à tous les vecteurs de V obtenus comme combinaison linéaire de $v_1, \ldots, v_n$.*

Preuve. Soit $w = \sum_{i=1}^{n} \alpha_i v_i$, $\alpha_i \in \mathbb{K} \ \forall i = 1, \ldots, n$, une combinaison linéaire quelconque des vecteurs $v_1, \ldots, v_n$. Par calcul direct :

$$\langle u, w \rangle = \langle u, \sum_{i=1}^{n} \alpha_i v_i \rangle \underset{\text{(sesquilinéarité)}}{=} \sum_{i=1}^{n} \overline{\alpha_i} \langle u, v_i \rangle \underset{u \perp v_i}{=} \sum_{i=1}^{n} \overline{\alpha_i} 0 = 0.$$

□

Théorème 1.4.1 (Théorème de Pythagore généralisé) *Soit $(V, \langle , \rangle)$ un espace vectoriel sur $\mathbb{K}$ avec produit scalaire. Soient $u, v \in V$ orthogonaux entre eux, alors :*

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

Plus en général, si les vecteurs $v_1, \ldots, v_n \in V$ sont orthogonaux, alors :

$$\left\| \sum_{i=1}^{n} v_i \right\|^2 = \sum_{i=1}^{n} \|v_i\|^2 \quad \Longleftrightarrow \quad \|v_1 + \ldots + v_n\|^2 = \|v_1\|^2 + \ldots + \|v_n\|^2 .$$

5. Leopold Kronecker (Liegnitz 1823 – Berlin 1891).

Preuve. Par le théorème de Carnot et par orthogonalité, on a :

$$\begin{aligned}\|u+v\|^2 &= \langle u,u\rangle + \cancelto{0}{\langle u,v\rangle} + \cancelto{0}{\langle v,u\rangle} + \langle v,v\rangle \\ &= \|u\|^2 + \|v\|^2.\end{aligned}$$

La preuve dans le cas de n vecteurs est obtenue par récurrence :

— Le cas $n = 2$ vient d'être démontré

— On suppose que $\left\|\sum_{i=1}^{n-1} v_i\right\|^2 = \sum_{i=1}^{n-1} \|v_i\|^2$ (hypothèse de récurrence)

— On écrit maintenant $u = v_n$ et $z = \sum_{i=1}^{n-1} v_i$, alors $u \perp z$ grâce au Lemme 1.4.1, donc, grâce au cas $n = 2$: $\|u+z\|^2 = \|u\|^2 + \|z\|^2$, mais

$$u + z = v_n + \sum_{i=1}^{n-1} v_i = \sum_{i=1}^{n} v_i,$$

donc

$$\|u+z\|^2 = \left\|\sum_{i=1}^{n} v_i\right\|^2,$$

et

$$\|u\|^2 + \|z\|^2 = \|v_n\|^2 + \left\|\sum_{i=1}^{n-1} v_i\right\|^2 \underset{\text{(Hypothèse de récurrence)}}{=} \|v_n\|^2 + \sum_{i=1}^{n-1} \|v_i\|^2 = \sum_{i=1}^{n} \|v_i\|^2,$$

d'ici la thèse.

□

On remarque que *la thèse du théorème de Pythagore est une double implication si V est réel*, tandis que *si V est complexe, alors elle est une implication simple.* En fait, grâce à la formule (1.6) ça vaut $\|u+v\|^2 = \|u\|^2 + \|v\|^2$ si et seulement si $\Re(\langle u,v\rangle) = 0$, ce qui est équivalent à l'orthogonalité seulement si V est réel.

Le résultat suivant donne une information sur la distance entre deux vecteurs quelconques d'une famille orthonormale.

Théorème 1.4.2 *Soit $(V, \langle,\rangle)$ un espace vectoriel sur $\mathbb{K}$ avec produit scalaire et soit F une famille orthonormale en V. Alors la distance entre deux éléments quelconques de F est constante et ça vaut $\sqrt{2}$.*

Preuve. Par le théorème de Pythagore : $\|u+(-v)\|^2 = \|u\|^2 + \|v\|^2 = 2$, grâce au fait que $u \perp v$. □

1.5 Orthogonalité et indépendance linéaire

La condition d'orthogonalité est plus forte que l'indépendance linéaire : en fait, **toute famille orthogonale est libre**.

Théorème 1.5.1 *Soit F une famille orthogonale de $(V, \langle , \rangle)$, $F = \{v_1, \cdots, v_n\}$, $\underline{v_i \neq 0}$ $\forall i$, alors F est libre.*

Preuve. On doit démontrer l'indépendance linéaire des v_i, i.e. : $\sum_{i=1}^{n} a_i v_i = 0 \Rightarrow a_i = 0 \ \forall i$. Pour le prouver, on va calculer le produit scalaire entre la combinaison linéaire $\sum_{i=1}^{n} a_i v_i$ et un vecteur quelconque v_j avec $j \in \{1, \cdots, n\}$:

$$\langle \sum_{i=1}^{n} a_i v_i, v_j \rangle \underset{(1.1)}{=} \sum_{i=1}^{n} a_i \langle v_i, v_j \rangle \underset{(\langle v_i, v_j \rangle \neq 0 \Leftrightarrow i=j)}{=} a_j \langle v_j, v_j \rangle = a_j \|v_j\|^2.$$

Par hypothèse, aucun vecteur de F est nul, donc, l'hypothèse que $\sum_{i=1}^{n} a_i v_i = 0$ implique que :

$$\underbrace{\langle 0, v_j \rangle}_{\shortparallel} = a_j \underbrace{\|v_j\|^2}_{\neq} \Rightarrow a_j = 0.$$
$$0 \qquad\qquad 0$$

Comme cela vaut pour tout $j \in \{1, \ldots, n\}$, la famille orthogonale F est libre. □

Grâce à la théorie générale des espaces vectoriels en dimension finie, on peut dériver un corollaire immédiat du théorème précédent.

Corollaire 1.5.1 *Une famille orthogonale de n vecteurs non nuls dans un espace $(V, \langle , \rangle)$ de dimension n est une base de V.*

Déf. 1.5.1 *On appelle une famille de n vecteurs orthogonaux non nuls d'un espace vectoriel $(V, \langle , \rangle)$ de dimension n, une **base orthogonale** de V. Si, en plus, la famille est orthonormale, on l'appelle une **base orthonormale** de V.*

Dans le chapitre 5 on verra comment étendre le concept de base orthogonale à un espace vectoriel avec produit scalaire de dimension infinie. Pour cela, il est important de souligner qu'*une base orthogonale est caractérisée par le fait d'être composée par le nombre maximal de vecteurs qui peuvent être orthogonaux entre eux dans un espace vectoriel* : en fait, dite n la dimension de l'espace V, si, par l'absurde, il existe un autre vecteur $u^* \in V$, $u \neq 0$, orthogonal à tous les vecteurs d'une base orthogonale $(u_i)_{i=1}^n$, alors l'ensemble $(u^*, u_i)_{i=1}^n$ serait libre, car les vecteurs orthogonaux sont linéairement indépendants, et la dimension de V serait $n+1$ et non pas n ! Il est habituel de résumer cette propriété en disant qu'une famille orthogonale est une base si elle n'est pas un sous ensemble propre d'une autre famille orthogonale de vecteurs de V.

Nous observons que, pour déterminer les composantes d'un vecteur par rapport à une base quelconque on doit résoudre un système linéaire de n équations avec n inconnues.

Par contre, si on a une base orthogonale ou orthonormale, les composantes sont déterminées par des produits scalaires, comme le montre le théorème suivant.

On observe que la résolution d'un système linéaire de n équations avec n inconnues nécessite, en général, beaucoup plus d'opérations que le calcul de produits scalaires, ceci montre déjà un des avantages de connaître une base orthogonale d'un espace vectoriel.

Théorème 1.5.2 *Soit $B = \{u_1, \ldots, u_n\}$ une base* ***orthogonale*** *de $(V, \langle , \rangle)$, alors :*

$$\boxed{v = \sum_{i=1}^{n} \frac{\langle v, u_i \rangle}{\|u_i\|^2} u_i}$$

En particulier, si B est une base ***orthonormale****, alors :*

$$\boxed{v = \sum_{i=1}^{n} \langle v, u_i \rangle u_i}.$$

Preuve. B est une base, donc il y a un ensemble des scalaires $\alpha_1, \ldots, \alpha_n$ tels que $v = \sum_{j=1}^{n} \alpha_j u_j$. On considère le produit scalaire de cette expression de v avec un vecteur fixé u_i, $i \in \{1, \ldots, n\}$:

$$\langle v, u_i \rangle = \langle \sum_{j=1}^{n} \alpha_j u_j, u_i \rangle = \sum_{j=1}^{n} \alpha_j \langle u_j, u_i \rangle \underset{(u_i \perp u_j \ \forall i \neq j)}{=} \alpha_i \langle u_i, u_i \rangle = \alpha_i \|u_i\|^2$$

donc $\alpha_i = \frac{\langle v, u_i \rangle}{\|u_i\|^2}$ $\forall i = 1, \cdots, n$, et alors $v = \sum_{i=1}^{n} \frac{\langle v, u_i \rangle}{\|u_i\|^2} u_i$. Si B est une base orthonormale, alors $\|u_i\| = 1$ et donc on a la deuxième formule du théorème. □

1.6 La projection orthogonale dans les espaces vectoriels avec produit scalaire

Comme vu pour le produit scalaire, nous allons étendre la définition de projection orthogonale en examinant les propriétés géométriques et algébriques de cette opération en $\mathbb{R}^2$ et en $\mathbb{R}^3$. Commençons par $\mathbb{R}^2$.

Dans l'espace Euclidien $\mathbb{R}^2$, il est clair que le produit scalaire d'un vecteur v avec un vecteur *unitaire* réalise la projection orthogonale de v dans la direction donnée par ce vecteur, comme le montre la figure 1.2, où la projection orthogonale est faite sur l'axe x.

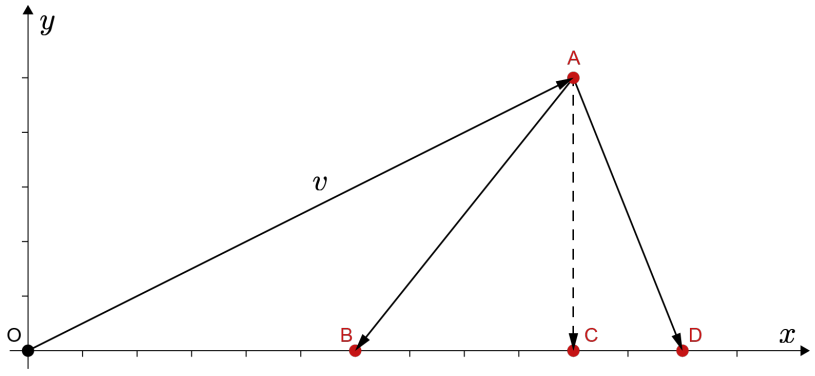


FIGURE 1.2 – Projection orthogonale $P_x v = \vec{OC}$ et projections diagonales $\vec{OB}$ et $\vec{OD}$ d'un vecteur de $v \in \mathbb{R}^2$ sur l'axe x.

Les propriétés vérifiés par la projection dans ce cas sont listées ci-dessous.

1. Si on projette une deuxième fois sur l'axe x, évidemment le vecteur $P_x v$ reste inaltéré, en étant déjà sur l'axe x, i.e. $P_x^2(v) := P_x(P_x v) = P_x v \ \forall v \in V$. Une autre manière de dire ça est que l'opérateur P_x restreint à l'axe x est l'identité de cet axe ;

2. Le vecteur différence entre v et sa projection $v - P_x v$ est orthogonale à l'axe x, comme le montre la figure 1.3 ;

3. $P_x v$ minimise la distance entre l'extrémité de v et l'axe x, en fait, dans la figure 1.2, $\vec{AB}$ et $\vec{AD}$ sont les hypoténuses des triangles rectangles ABC et ACD, par contre $\vec{AC}$ est un côté de ces triangles et, par conséquent, sa taille est inférieure à celle de $\vec{AB}$ et $\vec{AD}$. $\vec{AC}$ est la distance de l'extrémité de v à l'extrémité de $P_x v$, tandis que $\vec{AB}$ et $\vec{AD}$ sont les distances entre l'extrémité de v et les projections diagonales de v sur x avec pies en B et D, respectivement.

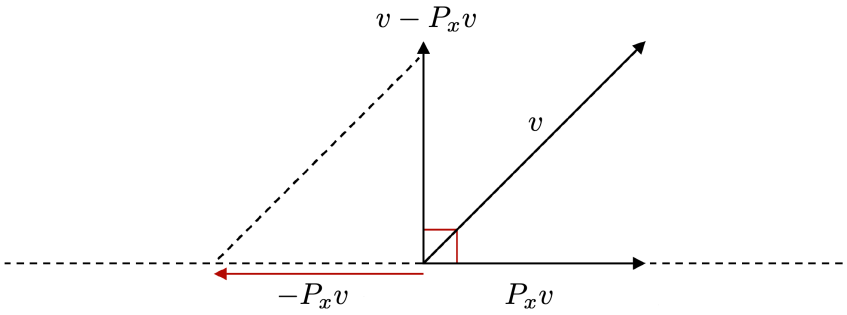


FIGURE 1.3 – Visualisation de la propriété 2) en $\mathbb{R}^2$.

Nous voulons définir une opération de projection orthogonale dans un espace vectoriel abstrait avec produit scalaire de dimension n qui préserve ces propriétés géométriques.

Une idée pour définir algébriquement cette opération est donnée grâce à l'analyse de la projection orthogonale en $\mathbb{R}^3$. Dans la figure 1.4 nous pouvons voir un vecteur $v \in \mathbb{R}^3$ et le plan engendré par les vecteurs *orthogonaux* u_1 et u_2. La projection p de v sur ce plan est la somme vectorielle des projections orthogonales $p_1 = \frac{\langle v, u_1 \rangle}{\|u_1\|^2} u_1$ et $p_2 = \frac{\langle v, u_2 \rangle}{\|u_2\|^2} u_2$ sur les deux vecteurs u_1 et u_2 considérés séparément, i.e. $p = p_1 + p_2 = \sum_{i=1}^{2} \frac{\langle v, u_i \rangle}{\|u_i\|^2} u_i$.

La généralisation maintenant devrait être claire : considérons un espace vectoriel de dimension n avec produit scalaire $(V, \langle , \rangle)$ et une **famille orthogonale** de vecteurs non nuls $F = \{u_1, \ldots, u_m\}$, $m \leqslant n$, $u_i \neq 0_V \ \forall i = 1, \ldots, m$.

On écrira le sous-espace vectoriel de V engendré par toutes les combinaisons linéaires des vecteurs de F comme $\text{span}(F)$:

$$\text{span}(F) \equiv S = \left\{ s \in V \ : \ \exists \alpha_1, \ldots, \alpha_m \in \mathbb{K} \text{ tels que } s = \sum_{j=1}^{m} \alpha_j u_j \right\}.$$

On définit l'**opérateur de projection orthogonale**, aussi dit **projecteur orthogonal**,

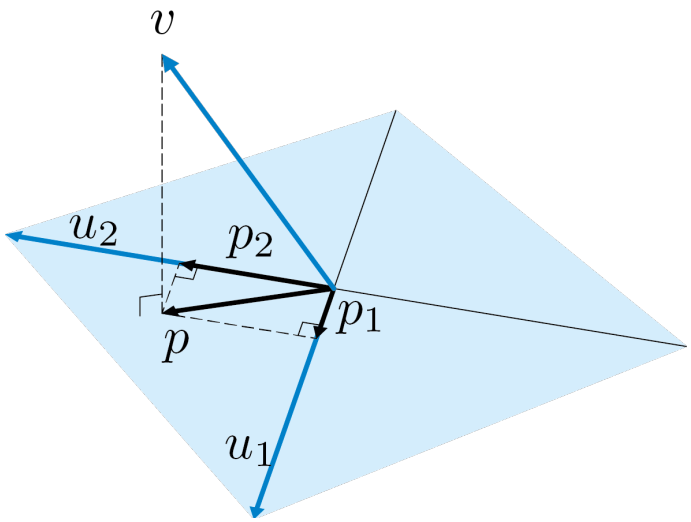


FIGURE 1.4 – Projection orthogonale p d'un vecteur de $\mathbb{R}^3$ sur le plan engendré par deux vecteurs unitaires.

d'un vecteur $v \in V$ sur S comme l'application (évidemment) linéaire :

$$\begin{array}{rccc} P_S : & V & \longrightarrow & S \subseteq V \\ & v & \longmapsto & \boxed{P_S(v) = \sum_{i=1}^{m} \frac{\langle v, u_i \rangle}{\|u_i\|^2} u_i} \end{array}.$$

Le théorème suivant montre que la projection orthogonale définie ci-dessus préserve toutes les propriétés de la projection orthogonale qu'on a mis en évidence en $\mathbb{R}^2$.

Théorème 1.6.1 *Avec les notations ci-dessus, ça vaut que :*

1) si $s \in S$ alors $P_S(s) = s$, i.e. l'action de P_S sur les vecteurs de S est l'identité ;

*2) $\forall v \in V$ et $s \in S$, le vecteur **résidu** de la projection, i.e. $v - P_S(v)$, est $\perp$ à S :*

$$\langle v - P_S(v), s \rangle = 0 \quad \Longleftrightarrow \quad v - P_S(v) \perp s \, ;$$

3) $\forall v \in V$ et $s \in S$: $\|v - P_S(v)\| \leqslant \|v - s\|$ et l'égalité vaut si et seulement si $s = P_S(v)$, on écrit

$$\boxed{P_S(v) = \operatorname*{argmin}_{s \in S} \|v - s\|}.$$

Preuve.

1) Soit $s \in S$, i.e. $s = \sum_{j=1}^{m} \alpha_j u_j$, alors :

$$P_S(s) = \sum_{i=1}^{m} \frac{\langle \sum_{j=1}^{m} \alpha_j u_j, u_i \rangle}{\|u_i\|^2} u_i = \sum_{i=1}^{m} \frac{\sum_{j=1}^{m} \alpha_j \langle u_j, u_i \rangle}{\|u_i\|^2} u_i \underset{(u_i \perp u_j \ \forall i \neq j)}{=} \sum_{i=1}^{m} \frac{\alpha_i \langle u_i, u_i \rangle}{\|u_i\|^2} u_i = \sum_{i=1}^{m} \alpha_i u_i = s.$$

2) On considère le produit scalaire de $P_S(v)$ avec un vecteur u_j, $j \in \{1, \ldots, m\}$ fixé :

$$\langle P_S(v), u_j \rangle = \langle \sum_{i=1}^{m} \frac{\langle v, u_i \rangle}{\|u_i\|^2} u_i, u_j \rangle \underset{\text{(linéarité)}}{=} \sum_{i=1}^{m} \frac{\langle v, u_i \rangle}{\|u_i\|^2} \langle u_i, u_j \rangle \underset{(u_i \perp u_j \ \forall i \neq j)}{=} \frac{\langle v, u_j \rangle}{\|u_j\|^2} \langle u_j, u_j \rangle = \langle v, u_j \rangle,$$

d'où

$$\langle v, u_j \rangle - \langle P_S(v), u_j \rangle = 0 \underset{\text{linéarité de } \langle\ ,\ \rangle}{\iff} \langle v - P_S(v), u_j \rangle = 0 \quad \forall j \in \{1, ..., m\}.$$

Le lemme 1.4.1 garantit que $\langle v - P_S(v), s \rangle = 0 \ \forall s \in S$.

3) Il est utile d'écrire la différence $v - s$ comme ceci : $v - P_S(v) + P_S(v) - s$. La propriété 2) nous dit que $v - P_S(v) \perp S$, par contre, $P_S(v), s \in S$ donc $P_S(v) - s \in S$. Par conséquent : $(v - P_S(v)) \perp (P_S(v) - s)$. Le théorème de Pythagore implique que :

$$\|v - s\|^2 = \|v - P_S(v) + P_S(v) - s\|^2 = \|v - P_S(v)\|^2 + \underbrace{\|P_S(v) - s\|^2}_{\geqslant 0} \geqslant \|v - P_S(v)\|^2,$$

d'où : $\|v - s\| \geqslant \|v - P_S(v)\| \ \forall v \in V, \ s \in S$.

Bien évidemment, $\|P_S(v) - s\| = 0$ si et seulement si $s = P_S(v)$, et dans ce cas on a $\|v - s\| = \|v - P_S(v)\|$. □

Le théorème qu'on vient de démontrer nous dit que, parmi tous les vecteurs du sous-espace vectoriel $S \subseteq V$, le plus « similaire à $v \in V$ » (dans le sens de la norme induite par le produit scalaire) est donné par la projection orthogonale. Dans le chapitre 5 nous allons voir comment on peut généraliser ce résultat aux espaces de Hilbert de dimension infinie.

Comme déjà souligné dans le cas de la projection sur $\mathbb{R}^2$ et $\mathbb{R}^3$, la quantité scalaire non-négative $\frac{|\langle v, u_i \rangle|}{\|u_i\|}$ donne une mesure de l'importance de $\frac{u_i}{\|u_i\|}$ dans la reconstruction de la meilleure approximation de v en S via la formule $P_S(v) = \sum_{i=1}^{m} \frac{\langle v, u_i \rangle}{\|u_i\|^2} u_i$: si cette quantité est grande, alors $\frac{u_i}{\|u_i\|}$ est extrêmement important pour reconstruire $P_S(v)$, sinon, dans certaines conditions, il peut être ignoré. En fait, dans les applications à la compression des signaux, une stratégie courante consiste en réordonner la somme qui définit $P_S(v)$ en ordre décroissant des quantités $\frac{|\langle v, u_i \rangle|}{\|u_i\|}$ et d'essayer d'éliminer la plus grande quantité de termes qui ont ces valeurs petits tout en préservant la qualité du signal.

Cette observation aura une importance fondamentale pour comprendre la signification de la décomposition de Fourier que nous allons examiner dans le cas discret et continu dans les chapitres qui suivent.

Pour terminer, nous observons que la formule $v = v - s + s$, apparemment triviale, quand on sait que $s \in S$ a une signification bien plus profonde : en fait, dans ce cas, on sait que $v - s$ et s sont orthogonaux.

La décomposition d'un vecteur comme la somme d'une composante qui appartient au sous-espace S et une qui appartient à son orthogonale est dite *théorème de la projection orthogonale* .

Cette décomposition est unique, et, dans le chapitre 5, nous allons voir comment elle peut être généralisée en dimension infinie et quelles sont ses conséquences sur la structure géométrique des espaces de Hilbert.

1.7 Existence d'une base orthonormale : l'algorithme itératif de Gram-Schmidt

On a vu que quand on a à disposition une base orthonormale, les formules de projection et de décomposition sont plus simples. Le théorème suivant dit que, dans un espace vectoriel de dimension finie avec produit scalaire, on peut toujours construire une base orthonormale à partir d'une famille libre de générateurs.

Théorème 1.7.1 (Procédure itérative de Gram-Schmidt [6]) *Si* $(v_1, \ldots, v_n)$, $n \leqslant \infty$, *est une base de* $(V, \langle , \rangle)$ *alors on peut trouver une base orthonormale de* $(V, \langle , \rangle)$ *à partir de* $(v_1, \ldots, v_n)$.

Preuve. La preuve est constructive, dans le sens qu'elle donne la méthode pour construire une base orthonormale à partir d'une base quelconque.

— 1ère étape, normalisation de v_1 :

$$u_1 = \frac{v_1}{\|v_1\|}.$$

— 2ème étape, comme montré dans la figure 1.5, on projette v_2 en direction de u_1, i.e. on considère $\langle v_2, u_1 \rangle u_1$. Grâce au théorème 1.6.1 on sait que la différence vectorielle $v_2 - \langle v_2, u_1 \rangle u_1$ est orthogonale à u_1, enfin on normalise le résultat :

$$u_2 = \frac{v_2 - \langle v_2, u_1 \rangle u_1}{\|v_2 - \langle v_2, u_1 \rangle u_1\|}.$$

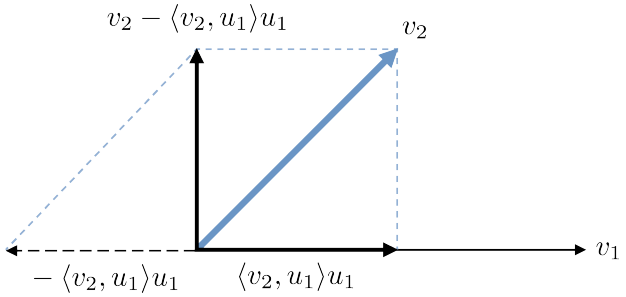


FIGURE 1.5 – Représentation graphique de la deuxième étape de la procédure d'orthonormalisation de Gram-Schmidt.

— n-ième étape, par itération :

$$u_n = \frac{v_n - (\langle v_n, u_{n-1} \rangle u_{n-1} + \ldots + \langle v_n, u_1 \rangle u_1)}{\|v_n - (\langle v_n, u_{n-1} \rangle u_{n-1} + \ldots + \langle v_n, u_1 \rangle u_1)\|}.$$

□

6. Jørgen Pedersen Gram (Nustrup 1850 - Copenhagen 1916), Erhard Schmidt (Tatu 1876 - Berlin 1959).

1.8 Les propriétés fondamentales d'une base orthonormale et orthogonale

On résume les propriétés les plus importantes d'une base orthonormale dans le théorème suivant :

Théorème 1.8.1 *Soit $(u_1, \ldots, u_n)$ une base orthonormale de $(V, \langle , \rangle)$, dim$(V) = n$. Alors, $\forall v, w \in V$:*

1. ***Théorème de décomposition sur une base orthonormale :***

$$\boxed{v = \sum_{i=1}^{n} \langle v, u_i \rangle u_i}\,; \tag{1.7}$$

2. ***Identité de Parseval***[7] ***:***

$$\boxed{\langle v, w \rangle = \sum_{i=1}^{n} \langle v, u_i \rangle \langle u_i, w \rangle}\,; \tag{1.8}$$

3. ***Identité de Plancherel***[8] ***:***

$$\boxed{\|v\|^2 = \sum_{i=1}^{n} |\langle v, u_i \rangle|^2}\,. \tag{1.9}$$

Preuve de 1) : c'est une conséquence immédiate du théorème 1.6.1. En fait, comme $(u_1, \ldots, u_n)$ est une base, $v \in \text{Span}(u_1, \ldots, u_n)$, en plus $(u_1, \ldots, u_n)$ est orthonormale, donc $v = P_S(v) = \sum_{i=1}^{n} \langle v, u_i \rangle u_i$. La division par $\|u_i\|^2$ dans la sommation n'est pas nécessaire car $\|u_i\| = 1 \ \forall i$.

Preuve de 2) : en utilisant 1) on peut écrire $v = \sum_{i=1}^{n} \langle v, u_i \rangle u_i$, et, si on calcule le produit scalaire de v, écrit comme ceci, avec w, grâce à la formule (1.1), on a :

$$\langle v, w \rangle = \langle \sum_{i=1}^{n} \langle v, u_i \rangle u_i, w \rangle = \sum_{i=1}^{n} \langle v, u_i \rangle \langle u_i, w \rangle.$$

Preuve de 3) : si on écrit $w = v$ dans le côté gauche de l'identité de Parseval on a $\langle v, v \rangle = \|v\|^2$, par contre, à droite on a :

$$\sum_{i=1}^{n} \langle v, u_i \rangle \langle u_i, v \rangle = \sum_{i=1}^{n} \langle v, u_i \rangle \overline{\langle v, u_i \rangle} = \sum_{i=1}^{n} |\langle v, u_i \rangle|^2$$

donc $\|v\|^2 = \sum_{i=1}^{n} |\langle v, u_i \rangle|^2$. □

7. Marc-Antoine de Parseval des Chênes (Rosières-aux-Salines 1755, Paris 1836).
8. Michel Plancherel (Bussy 1885 - Zurich 1967).

Observations :

1. L'**interprétation physique de l'identité de Plancherel** est la suivante : l'énergie de v, mesurée par sa norme carrée, peut être décomposée par la somme des modules carrés de chaque projection de v sur les n directions de la base orthonormale $(u_1, ..., u_n)$. Dans la théorie de Fourier les directions de la base orthonormale seront les harmoniques fondamentales (sinus et cosinus avec des fréquences bien définies), c'est pourquoi on appelle l'analyse de Fourier *analyse harmonique*.

2. Si, au lieu d'être une base orthonormale, $(u_1, \ldots, u_n)$ est une base orthogonale, alors, grâce à la formule du projecteur et au théorème 1.6.1, les résultats du théorème 1.8.1 peuvent être écrits comme ceci :

   (a) **Décomposition de $v \in V$ sur une base orthogonale** :

   $$v = \sum_{i=1}^{n} \frac{\langle v, u_i \rangle}{\|u_i\|^2} u_i \,; \tag{1.10}$$

   (b) **Identité de Parseval pour une base orthogonale** :

   $$\langle v, w \rangle = \sum_{i=1}^{n} \frac{\langle v, u_i \rangle \langle u_i, w \rangle}{\|u_i\|^2} \,; \tag{1.11}$$

   (c) **Identité de Plancherel pour une base orthogonale** :

   $$\|v\|^2 = \sum_{i=1}^{n} \frac{|\langle v, u_i \rangle|^2}{\|u_i\|^2} \,. \tag{1.12}$$

Le théorème suivant résume toutes les propriétés des bases orthonormales d'un espace vectoriel avec produit scalaire de dimension finie. Il sera généralisé au cas des espaces de Hilbert dans le chapitre 5 (voir théorème 5.5.4).

Théorème 1.8.2 *Soit V un espace vectoriel de dimension n sur $\mathbb{K}$. Les affirmations suivantes sont équivalentes :*

1. $(u_i)_{i=1}^n$ *est une base orthonormée de* V

2. $\langle v, u_i \rangle = 0 \ \forall i = 1, \ldots, n \iff v = 0_V$, *i.e.* 0_V *est le seul vecteur orthogonal à tous les vecteurs d'une base orthonormée*

3. $\operatorname{span}((u_i)_{i=1,\ldots,n}) = V$

4. $\forall v \in V \ : v = \sum_{i=1}^{n} \langle x, u_i \rangle u_i$

5. $\forall v, w \in V \ : \langle v, w \rangle = \sum_{i=1}^{n} \langle v, u_i \rangle \langle u_i, w \rangle$

6. $\forall v \in V \ : \quad \|v\|^2 = \sum_{i=1}^{n} |\langle v, u_i \rangle|^2$.

Nous terminons ce chapitre avec un exercice qui a le but de tester la connaissance de la théorie des espaces vectoriels avec produit scalaire de dimension finie et avec deux exercices qui montre explicitement un produit scalaire différent de celui Euclidien.

Exercice 1.1

Nous considérons l'espace vectoriel $\mathbb{C}^3$ avec produit scalaire Euclidien complexe et les trois vecteurs :

$$u = (0, i, 2i), \quad v = (2i, 0, -i), \quad w = \left(0, i, \frac{1}{2}e^{-\frac{\pi}{2}i}\right).$$

1) Déterminer les relations d'orthogonalité entre les vecteurs u, v, w ;

2) Calculer la norme de u, v, w et les distances Euclidiennes entre eux ;

3) Vérifier que (u, v, w) est une base (non orthogonale) de $\mathbb{C}^3$;

4) Soit S le sous-espace vectoriel de $\mathbb{C}^3$ engendré par u et w. Calculer $P_S v$, la projection orthogonale de v sur S. Calculer $d(v, P_S v)$, i.e. la distance Euclidienne entre v et sa projection sur S, et vérifier qu'elle minimise la distance entre v et les vecteurs de S (*suggestion* : considérer la distance au carrée) ;

5) En utilisant ce que nous avons obtenu dans les points précédents, déterminer une base orthogonale et une base orthonormale de $\mathbb{C}^3$ sans passer par la procédure d'orthonormalisation de Gram-Schmidt (*suggestion* : se rappeler de la relation géométrique du vecteur résidu r par rapport au sous-espace S) ;

6) Étant donné le vecteur $a = (2i, -1, 0)$, écrire la décomposition de a et l'identité de Plancherel par rapport à la base orthonormale déterminé dans le point 5). En déduire le vecteur de la base orthonormale qui a le poids le plus importante dans la décomposition de a (et qui donne la meilleure « reconstruction grossière » de a). À l'aide d'un logiciel graphique, dessiner la somme vectorielle progressive de a en commençant par la représentation grossière et en rajoutant les détails fins donnés par les deux autres vecteurs.

Solution de l'exercice 1.1

1) Évidemment $e^{-\frac{\pi}{2}i} = -i$, donc, par calcul direct des produits scalaires : $\langle u, v\rangle = 2$, $\langle u, w\rangle = 0$ et $\langle v, w\rangle = \frac{1}{2}$.

2) Par calcul direct : $\|u\|^2 = 5$, $\|v\|^2 = 5$, $\|w\|^2 = \frac{5}{4}$. Après avoir calculé les vecteurs différences, on obtient : $d(u, v) = \|u - v\| = \sqrt{14}$, $d(u, w) = \|u - w\| = \frac{5}{2}$, $d(v, w) = \|v - w\| = \frac{\sqrt{21}}{2}$.

3) Les trois vecteurs u, v, w sont linéairement indépendants, donc ils forment une base de $\mathbb{C}^3$, non orthogonale car seulement u et w sont orthogonaux.

4) $S = \text{span}(u, w)$. Comme (u, w) est une base orthogonale de S, nous pouvons écrire :

$$P_S(v) = \frac{\langle v, u\rangle}{\|u\|^2}u + \frac{\langle v, w\rangle}{\|w\|^2}w = (0, 0, -i).$$

Le vecteur résidu de la projection de v sur S est $r = v - P_S v = (2i, 0, 0)$ et donc $d(v, P_S v)^2 = \|r\|^2 = 4$. Par contre, le vecteur le plus général de S est $s = \alpha u + \beta w = (0, (\alpha+\beta)i, (2\alpha - \frac{\beta}{2})i)$ et $d(v,s)^2 = \|v-s\|^2 = 4 + (\alpha+\beta)^2 + (2\alpha - \frac{\beta}{2} + 1)^2 \geqslant 4 = d(v, P_S v)^2$. Ceci confirme que $P_S v$ est le vecteur de S qui a distance minimale à v par rapport à la norme Euclidienne.

5) r est orthogonal à S, qui est engendré par u et w, donc (u, w, r) est un ensemble de vecteurs orthogonaux en $\mathbb{C}^3$, i.e. une base orthogonale de $\mathbb{C}^3$. Pour obtenir une base orthonormale il suffit de diviser chaque vecteur par sa norme :

$$(\hat{u}, \hat{w}, \hat{r}) \equiv \left(\frac{u}{\|u\|}, \frac{w}{\|w\|}, \frac{r}{\|r\|} \right) = \left(\left(0, \frac{i}{\sqrt{5}}, \frac{2i}{\sqrt{5}} \right), \left(0, \frac{2i}{\sqrt{5}}, -\frac{i}{\sqrt{5}} \right), (i, 0, 0) \right)$$

6) Décomposition : $a = \langle a, \hat{u} \rangle \hat{u} + \langle a, \hat{w} \rangle \hat{w} + \langle a, \hat{r} \rangle \hat{r} = \frac{i}{\sqrt{5}} \hat{u} + \frac{2i}{\sqrt{5}} \hat{w} + 2\hat{r}$.
Identité de Plancherel : $\|a\|^2 = 5 = |\langle a, \hat{u} \rangle|^2 + |\langle a, \hat{w} \rangle|^2 + |\langle a, \hat{r} \rangle|^2 (= \frac{1}{5} + \frac{4}{5} + 4 = 5)$.
Le vecteur qui a le poids le plus important dans la reconstruction de a est alors $\hat{r}$, c'est donc celui-ci qui donne la meilleure reconstruction grossière de a. Si on fait la somme vectorielle de cette représentation grossière avec les autres deux vecteurs, nous allons reconstruire les « détails fins » de a, avant avec $\hat{w}$ et, en dernier, avec $\hat{u}$. □

Exercice 1.2

Soit $\mathcal{M}_n(\mathbb{C})$ l'espace des matrices complexes $n \times n$. On définit l'application $\phi : M(n, \mathbb{C}) \times M(n, \mathbb{C}) \to \mathbb{C}$ par

$$\phi(A, B) = \mathrm{tr}(B^\dagger A),$$

où $B^\dagger := \overline{B}^t$ désigne la matrice adjointe de B et tr est la trace matricielle. Montrer que ϕ est un produit scalaire.

Solution de l'exercice 1.2

La distributivité de la multiplication matricielle sur l'addition et la linéarité de la trace donne la linéarité de ϕ par rapport à la première variable.

Montrons maintenant que ϕ est hermitienne. Soient $A = (a_{i,j})_{1 \leqslant i,j \leqslant n}$ et $B = (b_{i,j})_{1 \leqslant i,j \leqslant n}$ deux matrices de $M(n, \mathbb{C})$. Notons $(c_{i,j})_{1 \leqslant i,j \leqslant n} = (\overline{b_{j,i}})_{1 \leqslant i,j \leqslant n}$ les coefficients de la matrice $B^\dagger$ et $(d_{i,j})_{1 \leqslant i,j \leqslant n} = (\overline{a_{j,i}})_{1 \leqslant i,j \leqslant n}$ ceux de $A^\dagger$.

On a alors

$$\begin{aligned}
\overline{\phi(A,B)} = \overline{\mathrm{tr}(B^\dagger A)} &= \overline{\mathrm{tr}\left[\left(\sum_{k=1}^{n} c_{i,k} a_{k,j} \right)_{i,j} \right]} = \overline{\mathrm{tr}\left[\left(\sum_{k=1}^{n} \overline{b_{k,i}} a_{k,j} \right)_{i,j} \right]} \\
&= \sum_{i,k=0}^{n} b_{k,i} \overline{a_{k,i}} = \sum_{i,k=0}^{n} d_{i,k} b_{k,i} = \mathrm{tr}(A^\dagger B) \\
&= \phi(B, A).
\end{aligned}$$

Donc ϕ est une forme sesquilinéaire hermitienne. D'autre part, ϕ est positive :

$$\forall A \in M(n, \mathbb{C}),\ \phi(A, A) = \sum_{i,k=0}^{n} |a_{k,i}|^2 \geqslant 0$$

Et elle est définie :

$$\begin{aligned}\phi(A,A)=0 &\iff \sum_{i,k=0}^{n}|a_{k,i}|^2=0\\ &\iff \forall 1\leqslant k,i\leqslant n,\ a_{k,i}=0\\ &\iff A=0\end{aligned}$$

Ainsi ϕ est un produit scalaire. □

Exercice 1.3

Soit $E=\mathbb{R}[X]$, l'espace vectoriel des polynômes en une variable, avec coefficients réels. Pour $P,Q\in E$, on pose

$$\Phi(P,Q)=\int_{-1}^{1}\frac{P(t)Q(t)}{\sqrt{1-t^2}}dt.$$

1. On rappelle que l'écriture $f(t)\underset{t\to t_0}{=}O(g(t))$ veut dire que $\exists\ a,C>0$ telles que $|t-t_0|<a\implies|f(t)|\leqslant C\,|g(t)|$. Montrer que, pour tout $P,Q\in E$, ça vaut :

$$\frac{P(t)Q(t)}{\sqrt{1-t^2}}\underset{t\to 1}{=}O\left(\frac{1}{\sqrt{1-t}}\right),$$

et

$$\frac{P(t)Q(t)}{\sqrt{1-t^2}}\underset{t\to -1}{=}O\left(\frac{1}{\sqrt{1+t}}\right).$$

En déduire que Φ est bien définie sur $E\times E$.

2. Montrer que Φ est un produit scalaire sur E que l'on notera $\langle\ ,\ \rangle$.
3. Pour $n\in\mathbb{N}$, on note T_n le n-ième **polynôme de Tchebychev**, c'est-à-dire l'unique polynôme tel que $\forall\theta\in\mathbb{R},\ T_n(\cos\theta)=\cos(n\theta)$. En effectuant le changement de variable $t=\cos\theta$, montrer que $(T_n)_{n\in\mathbb{N}}$ est une famille orthogonale de E. Suggestion : se rappeler de la formule trigonométrique suivante

$$\frac{1}{2}(\cos((n+m)\theta)+\cos((n-m)\theta))=\cos(n\theta)\cos(m\theta)\qquad\forall n,m\in\mathbb{N}.\tag{1.13}$$

4. Montrer que pour tout $n\in\mathbb{N}$, $(T_0,\dots,T_n)$ est une base orthogonale de $\mathbb{R}_n[X]$, l'espace vectoriel des polynômes de $\mathbb{R}[X]$ de degré inférieur ou égal à n. En déduire que $(T_n)_{n\in\mathbb{N}}$ est une base orthogonale (au sens algébrique : tout élément de E est une combinaison linéaire *finie* d'éléments de la base) de E.
5. Calculer la norme de T_n pour tout n et en déduire une base orthonormée (au sens algébrique) de E.

Solution de l'exercice 1.3

1. Notons $f(t)=\frac{P(t)Q(t)}{\sqrt{1-t^2}}=\frac{P(t)Q(t)}{\sqrt{1-t}\sqrt{1+t}}$. Puisque P et Q sont des polynômes, la fonction $t\mapsto\frac{P(t)Q(t)}{\sqrt{1+t}}$ est continue en un voisinage $V_1(1)$ et donc, par le théorème de Weierstrass, elle est bornée dans ce voisinage, i.e. $\exists\ C_1>0$ tel que $t\in V_1(1)\implies\left|\frac{P(t)Q(t)}{\sqrt{1+t}}\right|\leqslant C_1$.

De la même manière, la fonction $t \mapsto \frac{P(t)Q(t)}{\sqrt{1-t}}$ est continue en un voisinage $V_2(-1)$ et donc, $\exists\ C_2 > 0$ tel que $t \in V_2(-1) \implies \left|\frac{P(t)Q(t)}{\sqrt{1-t}}\right| \leqslant C_2$. Donc, on a bien :

$$t \in V_1(1) \implies |f(t)| \leqslant C_1 \frac{1}{\sqrt{1-t}} \iff f(t) \underset{t\to 1}{=} O\left(\frac{1}{\sqrt{1-t}}\right)$$

et

$$t \in V_2(-1) \implies |f(t)| \leqslant C_2 \frac{1}{\sqrt{1+t}} \iff f(t) \underset{t\to -1}{=} O\left(\frac{1}{\sqrt{1+t}}\right).$$

Ceci implique que l'intégrale qui définit Φ est bien définit, en fait $f(t)$ est continue sur $]-1,1[$ et donc intégrable. Le résultat que l'on vient de prouver montre que $f(t)$ est intégrable dans un voisinage droit de -1 et un voisinage gauche de 1, car l'intégrale de sa valeur absolue est majorée par une fonction intégrable dans les deux cas.

2. La bilinéarité de Φ descend de la linéarité de l'intégrale par calcul direct. La symétrie est une conséquence de la symétrie du produit ponctuel entre fonctions. La seule propriété qui n'est pas évidente est la définie positivité. On commence par la positivité :

$$\Phi(P,P) = \int_{-1}^{1} \frac{P^2(t)}{\sqrt{1-t^2}} dt \geqslant 0$$

et [9]

$$\Phi(P,P) = 0 \iff \frac{P^2(t)}{\sqrt{1-t^2}} dt = 0 \text{ p.p. sur }]-1,1[\iff P(t) = 0 \text{ p.p. sur }]-1,1[,$$

mais le seul polynôme qui a un nombre infini de racines est le polynôme nul $\mathbf{0}(t) \equiv 0$, alors $P = \mathbf{0}$. Donc, Φ est un produit scalaire sur E.

3. Pour tout $n, m \in \mathbb{N}$,

$$\begin{aligned}
\langle T_n, T_m \rangle &= \int_{-1}^{1} \frac{T_n(t)T_m(t)}{\sqrt{1-t^2}} dt \qquad (t = \cos\theta,\ dt = -\sin\theta d\theta) \\
&\qquad t = \cos\theta = -1 \iff \theta = \pi,\ t = \cos\theta = 1 \iff \theta = 0 \\
&= \int_{\pi}^{0} -\sin\theta \frac{T_n(\cos\theta)T_m(\cos\theta)}{\sqrt{1-\cos^2(\theta)}} d\theta \\
&= \int_{0}^{\pi} \sin\theta \frac{\cos(n\theta)\cos(m\theta)}{|\sin\theta|} d\theta \\
&= \int_{0}^{\pi} \cancel{\sin\theta} \frac{\cos(n\theta)\cos(m\theta)}{\cancel{\sin\theta}} d\theta \quad (\sin\theta \geqslant 0 \text{ sur } [0,\pi]) \\
&= \int_{0}^{\pi} \cos(n\theta)\cos(m\theta) d\theta \\
&= \frac{1}{2}\left(\int_{0}^{\pi} \cos((n+m)\theta) d\theta + \int_{0}^{\pi} \cos((n-m)\theta) d\theta\right) \quad (\text{grâce à } (1.13)).
\end{aligned}$$

9. p.p. : presque partout (voir le chapitre 3).

Alors, pour tout $n \neq m$, ça vaut :

$$\langle T_n, T_m \rangle = \frac{1}{2} \left(\left[\frac{\sin((n+m)\theta)}{n+m} \right]_0^\pi + \left[\frac{\sin((n-m)\theta)}{n-m} \right]_0^\pi \right) = 0,$$

i.e. les polynômes de Tchebychev forment une famille orthogonale de polynômes par rapport au produit scalaire que nous avons défini ci-dessus.

4. La famille $(T_0, T_1, \ldots, T_n)$ est une famille orthogonale (donc libre) de $n+1$ éléments de $\mathbb{R}[X]$, qui a dimension $n+1$, par conséquent elle est une base orthogonale de $\mathbb{R}_n[X]$. Pour montrer que $(T_n)_{n\in\mathbb{N}}$ est une base dans le sens algébrique de E, on considère un polynôme $P \in E$ de degré arbitraire $d \in \mathbb{N}$, i.e. $P \in \mathbb{R}_d[X]$ et nous observons que $(T_0, T_1, \ldots, T_d)$ est une famille orthogonale (libre) de générateurs de $\mathbb{R}_d[X]$, i.e. une base dans le sens algébrique.

5. Pour calculer la norme de T_n nous utilisons l'égalité

$$\langle T_n, T_m \rangle = \frac{1}{2} \left(\int_0^\pi \cos((n+m)\theta) d\theta + \int_0^\pi \cos((n-m)\theta) d\theta \right)$$

qu'on a montré dans le point 3. En prenant $n = m$ on a :

$$\|T_n\|^2 = \langle T_n, T_n \rangle = \frac{1}{2} \left(\int_0^\pi \cos(2n\theta) d\theta + \int_0^\pi d\theta \right) = \frac{1}{2} \left(\left[\frac{\sin 2n\theta}{2n} \right]_0^\pi + \pi \right) =$$

$$\|T_n\|^2 = \langle T_n, T_n \rangle = \frac{1}{2} \left(\int_0^\pi \cos(2n\theta) d\theta + \int_0^\pi d\theta \right) =$$

$$= \begin{cases} \frac{1}{2} \left(\int_0^\pi d\theta + \int_0^\pi d\theta \right) = \pi & \text{si } n = 0 \\ \frac{1}{2} \left(\left[\frac{\sin 2n\theta}{2n} \right]_0^\pi + \pi \right) = \frac{\pi}{2} & \text{si } n \geqslant 1, \end{cases}$$

d'où $\|T_0\| = \sqrt{\pi}$ et $\|T_n\| = \sqrt{\pi/2}$ pour $n \geqslant 1$. Pour terminer, la famille

$$\left\{ \frac{T_0}{\sqrt{\pi}} \right\} \cup \left\{ \sqrt{\frac{2}{\pi}} T_n \right\}_{n \geqslant 1}$$

est une base orthonormale de l'espace vectoriel des polynômes en une variable à coefficients réels E.

□

1.9 Résumé du chapitre 1

- Nous avons examiné les propriétés du produit scalaire réel et complexe, en soulignant leur différence. En particulier, la symétrie et la bilinéarité du produit scalaire réel doivent être remplacées par la symétrie conjuguée et par la sesquilinéarité pour pouvoir obtenir un ensemble de propriétés compatibles avec la définie positivité. Cette dernière propriété est essentielle pour pouvoir engendrer une norme à partir du produit scalaire.

- Nous avons rappelé que le prototype de tous les espaces vectoriels avec produit scalaire, aussi dits pré-hilbertiens, de dimension finie n est l'espace Euclidien $\mathbb{K}^n$, où $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$.

- Grâce au produit scalaire on peut étendre le concept d'orthogonalité entre vecteurs à n'importe quel espace vectoriel muni d'un produit scalaire. Deux vecteurs sont orthogonaux quand leur produit scalaire est nul. Le vecteur nul est le seul vecteur qui est orthogonal à tous les autres vecteurs et, grâce à la définie positivité, il est le seul vecteur orthogonal à lui-même. Si deux vecteurs ont le même produit scalaire avec tous les autres vecteurs, i.e. la même projection dans n'importe quelle direction, alors ils coïncident.

- Une norme sur un espace vectoriel est dite hilbertienne s'il est possible de munir l'espace d'un produit scalaire qui génère la norme de la manière canonique. Remarquablement, une norme est hilbertienne si et seulement si elle satisfait la règle du parallélogramme, ceci vaut en dimension finie et infinie. À partir d'une norme hilbertienne on peut définir un produit scalaire compatible avec elle grâce à la formule de polarisation.

- L'orthogonalité entre vecteurs implique leur indépendance linéaire, ce qui assure qu'un ensemble de n vecteurs orthogonaux dans un espace vectoriel de dimension n est une base. L'expansion d'un vecteur sur une base orthonormale est très simple : les composantes par rapport à cette base sont les produits scalaires du vecteur avec les vecteurs de la base. Ceci simplifie énormément le calcul des composantes, qui, sans l'orthonormalité de la base, nécessite la résolution d'un système linéaire.

- Nous avons aussi rappelé le concept de projection orthogonale sur un sous-espace vectoriel S : si on connaît une base orthogonale de cet espace, on peut représenter la projection comme une expansion sur les vecteurs de la base, avec des coefficients donnés encore par des produits scalaires (normalisés si la base n'est pas orthonormée). On a vu aussi que la différence entre un vecteur et sa projection orthogonale, qu'on appelle vecteur résidu, est orthogonale au sous-espace de projection S. Pour terminer, on a démontré que la projection orthogonale est le vecteur de S qui minimise la distance (par rapport à la norme de l'espace vectoriel) entre le vecteur et les vecteurs de S.

- Étant donné un espace vectoriel avec produit scalaire, de dimension finie ou infinie, une base orthonormale existe toujours grâce à l'algorithme d'orthonormalisation de Gram-Schmidt.

- Pour terminer, nous avons démontré les importantes identités de Parseval et de Plancherel relatives à une base orthonormale ou orthogonale. Dans le chapitre 5 nous allons voir comment étendre ces propriétés en dimension infinie.

Chapitre 2

La transformée de Fourier discrète et ses applications à la théorie des signaux et des images

Avec le rappel qu'on a fait sur les espaces vectoriels complexes avec produit scalaire et leurs propriétés, on peut introduire la transformée de Fourier discrète d'une façon très simple.

Il s'agit tout simplement de prouver que certaines fonctions des exponentiels complexes sont une base orthogonale pour un espace vectoriel complexe avec produit scalaire de dimension finie.

D'un point de vue mathématique la transformée de Fourier discrète est un simple changement de base dans un espace vectoriel, néanmoins, son interprétation a une importance et une utilité énorme dans les applications, en particulier dans le cadre de la théorie des signaux, comme on le verra dans la section 2.6.

Cette section est inspiré par l'excellent livre de M. Frazier [9].

2.1 L'espace $\ell^2(\mathbb{Z}_N)$ et sa base canonique

Pour introduire l'espace vectoriel où on travaillera pour construire la DFT, on va faire des petits changements dans la notation qu'on a utilisée jusqu'à maintenant.

On considèrera encore des vecteurs complexes avec un nombre N, $1 < N < +\infty$, de composantes, mais on va interpréter un vecteur de $\mathbb{C}^N$ comme une suite finie.

On commence par définir $\mathbb{Z}_N$.

Déf. 2.1.1 *On dit que deux entiers $i, j \in \mathbb{Z}$ sont **congrus modulo** N si leur différence est divisible par N, i.e.*

$$\frac{a-b}{N} = m \in \mathbb{Z},$$

cela implique qu'on peut écrire $a = b + mN$. La notation (dite de Gauss) pour deux entiers congrus modulo N est la suivante :

$$a \equiv b \pmod{N}.$$

Il est possible de démontrer que être congrus modulo N est une relation d'équivalence en $\mathbb{Z}$. Comme toute relation d'équivalence, elle crée une partition de $\mathbb{Z}$ en classes d'équivalence

distinctes. On écrit l'ensemble de ces classes d'équivalences comme

$$\mathbb{Z}_N = \mathbb{Z}\,(\mathrm{mod}\ N).$$

Une suite (finie) sur $\mathbb{Z}_N$ à valeurs complexes est une fonction :

$$\begin{array}{rccc} z & : & \mathbb{Z}_N & \to & \mathbb{C} \\ & & j & \mapsto & z(j). \end{array}$$

Dans la pratique, on va appliquer l'arithmétique qu'on appelle circulaire ou de l'horloge, qui consiste en identifier une suite définie sur $\mathbb{Z}_N$ comme une suite définie sur $\{0, 1, \ldots, N-1\}$ et étendue à $\mathbb{Z}$ par N-périodicité :

$$\boxed{z(j+mN) = z(j)} \quad \forall j, m \in \mathbb{Z},$$

i.e., si on connaît la définition de $z(j)$ quand $j \in \{0, 1, \ldots, N-1\}$, pour déterminer $z(j)$ quand $j \notin \{0, \ldots, N-1\}$, on doit ajouter à j un multiple entier de N, qu'on écrit mN, $m \in \mathbb{Z}$, tel que $\bar{\jmath} = j + mN \in \{0, 1, \ldots, N-1\}$ et après on définit $z(j+mN) = z(\bar{\jmath})$. Voyons un exemple.

Exemple : $N = 12$, $z = (1, i, i, \sqrt{2}i, 0, 0, 0, -1, 0, 0, 0, 2)$, i.e.

$$\begin{cases} z(0) = 1 \\ z(1) = z(2) = i \\ z(3) = \sqrt{2}i \\ z(4) = z(5) = z(6) = 0 \\ z(7) = -1 \\ z(8) = z(9) = z(10) = 0 \\ z(11) = 2. \end{cases}$$

étendue par 12-périodicité à $\mathbb{Z}$, déterminer $z(-21)$. Comme $N = 12$, on doit chercher l'entier $m \neq 0$ tel que $-21 + 12m \in \{0, 1, \ldots, 11\}$:

$$-21 + 12m = \begin{cases} -21 + 12 = -9 & m = 1 \\ -21 + 24 = 3 & m = 2 \\ -21 + 36 = 15 & m = 3 \\ \ldots \end{cases}$$

La valeur de m qui fait que $-21 + 12m$ soit inclus dans $\{0, \ldots, 11\}$ est $m = 2$ et dans ce cas on a $-21 + 2 \cdot 12 = 3$, ce qui implique $z(-21) = z(3) = \sqrt{2}i$.

Malgré le fait que l'on identifiera souvent $\mathbb{Z}_N$ avec l'ensemble des représentatifs canoniques $\{0, 1, \ldots, N-1\}$, **on peut penser z défini sur n'importe quel sous-ensemble de $\mathbb{Z}$ donné par N entiers consécutifs, et pas forcément sur** $\{0, \ldots, N-1\}$, on utilisera cette convention pendant tout le reste des polycopiés sans le spécifier ultérieurement.
L'espace vectoriel complexe dans lequel on va travailler est l'ensemble de toutes les suites sur $\mathbb{Z}_N$ à valeurs complexes :

$$\boxed{\ell^2(\mathbb{Z}_N) = \{z : \mathbb{Z}_N \to \mathbb{C}\}}.$$

La raison de cette notation particulière sera plus claire dans la suite du cours, pour le moment il faudra l'accepter.

$\ell^2(\mathbb{Z}_N)$ est un espace vectoriel complexe avec les opérations de somme et de multiplication scalaire usuelles, i.e. donnés $z, w \in \ell^2(\mathbb{Z}_N)$, $\alpha \in \mathbb{C}$, la somme et la multiplication par un scalaire complexe sont définies comme ceci :

$$\begin{array}{rccl} z+w & : & \mathbb{Z}_N & \to & \mathbb{C} \\ & & j & \mapsto & (z+w)(j) = z(j) + w(j), \end{array}$$

$$\begin{array}{rccl} \alpha z & : & \mathbb{Z}_N & \to & \mathbb{C} \\ & & j & \mapsto & (\alpha z)(j) = \alpha z(j). \end{array}$$

$\ell^2(\mathbb{Z}_N)$ a dimension N, en fait, l'application qui associe à chaque suite $z \in \ell^2(\mathbb{Z}_N)$ ses images $(z(0), z(1), \ldots, z(N-1))$:

$$\begin{array}{ccc} \ell^2(\mathbb{Z}_N) & \longleftrightarrow & \mathbb{C}^N \\ z & \longleftrightarrow & (z(0), z(1), \ldots, z(N-1)) = \begin{pmatrix} z(0) \\ z(1) \\ \vdots \\ z(N-1) \end{pmatrix}, \end{array}$$

est un *isomorphisme linéaire* (la preuve est laissée comme exercice). On utilisera la représentation de z comme vecteur ligne ou comme vecteur colonne selon nos besoins.

Grâce à l'isomorphisme ci-dessus, on peut définir la **base canonique** B **de** $\ell^2(\mathbb{Z}_N)$ comme l'ensemble des N suites suivantes :

$$B = (e_0, e_1, \ldots, e_{N-1}), \quad e_j(k) = \delta_{j,k} = \begin{cases} 1 & k = j \\ 0 & k \neq j. \end{cases}$$

On peut aussi introduire un **produit scalaire dans** $\ell^2(\mathbb{Z}_N)$ avec :

$$\langle z, w \rangle = \sum_{k=0}^{N-1} z(k)\overline{w(k)},$$

donc $z, w \in \ell^2(\mathbb{Z}_N)$ sont orthogonales si et seulement si $\langle z, w \rangle = \sum_{k=0}^{N-1} z(k)\overline{w(k)} = 0$.

La norme induite par ce produit scalaire est :

$$\|z\| = \left(\sum_{k=0}^{N-1} |z(k)|^2 \right)^{\frac{1}{2}},$$

que l'on appellera « **norme** $\ell^2(\mathbb{Z}_N)$ ».

2.1.1 La base orthogonale des exponentiels complexes de $\ell^2(\mathbb{Z}_N)$

Le système de fonctions que l'on va définir sera essentiel pour le développement de l'analyse de Fourier.

Avant de définir ce système, on rappelle que :

1. $z \in \mathbb{C}$ quelconque, $z = \rho[\cos\alpha + i\sin\alpha] = \rho e^{i\alpha}$, $\rho, \alpha \in \mathbb{R}$, $\rho \geqslant 0$;
2. Formules d'Euler : $\cos\alpha = \frac{1}{2}(e^{i\alpha} + e^{-i\alpha})$, $\sin\alpha = \frac{1}{2i}(e^{i\alpha} - e^{-i\alpha})$;
3. $|z| = 1 \Leftrightarrow z = e^{i\alpha}$;
4. $e^{i\alpha} = e^{i(\alpha+2\pi k)}$, $k \in \mathbb{Z}$;
5. Comme cas particulier du point ci-dessus, si $\alpha = 0$ on trouve :
$$\boxed{e^{2\pi ik} = 1} \qquad \forall k \in \mathbb{Z} ;$$
6. $e^{i\alpha}e^{i\beta} = e^{i(\alpha+\beta)}$;
7. $(e^{i\alpha})^n = e^{in\alpha}$;
8. $(e^{i\alpha})^* = e^{-i\alpha}$;
9. Étant donné $z = \rho e^{i\alpha}$, les solutions de l'équation $w^N = z$ sont les N racines complexes données par la formule : $w_m = \sqrt[N]{\rho} e^{i\frac{2\pi m+\alpha}{N}}$, $m = 0, \ldots, N-1$;
10. En particulier :
$$\text{Racines } N\text{-ièmes de l'unité } : \boxed{\omega_m = e^{2\pi i\frac{m}{N}}}, \quad m = 0, \ldots, N-1.$$

La dernière information qu'on a besoin de rappeler est la formule de la sommation géométrique, définie par

$$S_k = 1 + z + z^2 + \ldots + z^{k-1} + z^k = \sum_{j=0}^{k} z^j.$$

Si $z = 1$, alors $S_k = k + 1$. Si $z \neq 1$, on observe que :

$$(1-z)S_k = 1 + z + z^2 + \ldots + z^k - (z + z^2 + \ldots + z^k + z^{k+1}) = 1 - z^{k+1},$$

et donc :

$$\sum_{j=0}^{k} z^j = \begin{cases} \frac{1-z^{k+1}}{1-z} & \text{si } z \in \mathbb{C}\backslash\{1\} \\ k+1 & \text{si } z = 1. \end{cases}$$

On considère maintenant les suites de $\ell^2(\mathbb{Z}_N)$ définies par des exponentiels complexes, comme ceci :

$$\begin{array}{rccc} \mathcal{E}_m & : & \mathbb{Z}_N & \longrightarrow & \mathbb{C} \\ & & n & \longmapsto & \mathcal{E}_m(n), \end{array}$$

où :

$$\begin{cases} \mathcal{E}_0(n) = 1 \\ \mathcal{E}_1(n) = e^{2\pi i \frac{n}{N}} \\ \mathcal{E}_2(n) = e^{2\pi i \frac{2n}{N}} \\ \quad \vdots \\ \mathcal{E}_{N-1}(n) = e^{2\pi i \frac{(N-1)n}{N}}. \end{cases}$$

Donc :

— $\mathcal{E}_0$ est la suite constante $\mathcal{E}_0(n) \equiv 1 \ \forall n \in \mathbb{Z}_N$;

— $\mathcal{E}_1$ est la suite $\mathcal{E}_1 = \left(1, e^{2\pi i \frac{1}{N}}, e^{2\pi i \frac{2}{N}}, \ldots, e^{2\pi i \frac{(N-1)}{N}}\right)$;

— $\mathcal{E}_2$ est la suite $\mathcal{E}_2 = \left(1, e^{2\pi i \frac{2}{N}}, e^{2\pi i \frac{4}{N}}, \ldots, e^{2\pi i \frac{2(N-1)}{N}}\right)$;

— $\mathcal{E}_{N-1}$ est la suite $\mathcal{E}_{N-1} = \left(1, e^{2\pi i \frac{N-1}{N}}, e^{2\pi i \frac{2(N-1)}{N}}, \ldots, e^{2\pi i \frac{(N-1)^2}{N}}\right)$.

La suite générale de l'ensemble est :

$$\boxed{\mathcal{E}_m(n) = e^{2\pi i \frac{mn}{N}} = (\omega_m)^n} \qquad \forall m, n = 0, \ldots, N-1,$$

où $(\omega_m)^n$ est la puissance n-ième des racines N-èmes de l'unité, $\forall n \in \{0, ..., N-1\}$, en fait :

$$(\omega_m)^n = \left(e^{2\pi i \frac{m}{N}}\right)^n = e^{2\pi i \frac{mn}{N}}.$$

Grâce à la formule $z = e^{i\alpha} = [\cos\alpha + i \sin\alpha]$, on sait que le système qu'on vient de définir est **un ensemble de suites de valeurs qui oscillent avec des fréquences différentes**, car les arguments des fonctions cos et sin changent avec les coefficients m et n. Comme on le verra bientôt, la signification de ces fréquences a une importance centrale dans l'analyse de Fourier.

Pour le moment, il est plus important de démontrer que le système des exponentiels qu'on vient de définir est une base orthogonale de $\ell^2(\mathbb{Z}_N)$. La preuve nécessite un lemme préliminaire.

Lemme 2.1.1 *Pour tout $j, k \in \{0, 1, \ldots, N-1\}$ ça vaut la formule :*

$$\sum_{n=0}^{N-1} e^{2\pi i n \frac{j-k}{N}} = \sum_{n=0}^{N-1} e^{-2\pi i n \frac{j-k}{N}} = N\delta_{j,k} = \begin{cases} N & j = k \\ 0 & j \neq k. \end{cases} \tag{2.1}$$

On reviendra après sur l'interprétation physique de cette formule très importante. Avant de faire la preuve on observe que dans le cas $j, k \in \mathbb{Z}_N$, $j \neq k$, on a que $j - k \in \{1, 2, \ldots, N-1\}$, donc $\frac{j-k}{N} = -\frac{k-j}{N} \notin \mathbb{Z}$.

Preuve. On fait la preuve pour la première sommation, il sera clair que la même démonstration vaut aussi pour la deuxième. On commence par utiliser les propriétés des exponentiels complexes pour réécrire la formule comme ceci :

$$\sum_{n=0}^{N-1} e^{2\pi i n \frac{j-k}{N}} = \sum_{n=0}^{N-1} \left(e^{2\pi i \frac{j-k}{N}}\right)^n.$$

On analyse les deux cas suivants :

— Si $j = k$, les exponentiels dans la somme sont égaux à 1, car $e^{2\pi i \frac{0}{N}} = 1$, et donc

$$\sum_{n=0}^{N-1} e^{2\pi i n \frac{j-j}{N}} = \sum_{n=0}^{N-1} 1 = N.$$

— Si $j \neq k$, les exponentiels sont $\neq 1$, alors, grâce à la formule de la sommation géométrique :

$$\begin{aligned}\sum_{n=0}^{N-1} \left(e^{2\pi i \frac{j-k}{N}}\right)^n &= \frac{1 - \left(e^{2\pi i \frac{j-k}{N}}\right)^{N-1+1}}{1 - e^{2\pi i \frac{j-k}{N}}} \\ &= \frac{1 - e^{2\pi i \frac{(j-k)N}{N}}}{1 - e^{2\pi i \frac{j-k}{N}}} \\ &= \frac{1 - e^{2\pi i (j-k)}}{1 - e^{2\pi i \frac{j-k}{N}}}.\end{aligned}$$

Comme $j - k = m \in \mathbb{Z}$, $e^{2\pi i (j-k)} = 1$, alors le numérateur de la dernière formule est 0 toutefois que $j \neq k$. Par contre, le dénominateur ne s'annule jamais, car, comme observé avant de la preuve, si $j \neq k$, alors $\frac{j-k}{N} \notin \mathbb{Z}$. Donc, dans ce cas la sommation vaut 0. □

Maintenant, on peut démontrer facilement que $\mathcal{E}$ est une base orthogonale de $\ell^2(\mathbb{Z}_N)$.

Théorème 2.1.1 *$\mathcal{E} = (\mathcal{E}_0, \ldots, \mathcal{E}_{N-1})$ est une base orthogonale de $\ell^2(\mathbb{Z}_N)$.*

Preuve. $\mathcal{E}$ est donnée par N éléments d'un espace vectoriel N dimensionnel avec produit scalaire, donc si on prouve que $\mathcal{E}$ est une famille orthogonale, alors on démontre le théorème. En fait, on sait qu'une famille orthogonale est libre, et une famille libre de N vecteurs dans un espace vectoriel de dimension N est une base.

On va donc calculer les produits scalaires $\langle \mathcal{E}_j, \mathcal{E}_k \rangle$, $\forall j, k \in \{0, \ldots, N-1\}$:

$$\langle \mathcal{E}_j, \mathcal{E}_k \rangle = \sum_{n=0}^{N-1} \mathcal{E}_j(n)\overline{\mathcal{E}_k(n)} = \sum_{n=0}^{N-1} e^{2\pi i \frac{jn}{N}} e^{-2\pi i \frac{kn}{N}} = \sum_{n=0}^{N-1} e^{2\pi i \frac{(j-k)n}{N}} = N\delta_{j,k},$$

où on a utilisé le Lemme 2.1 pour arriver à la dernière égalité, qui montre que $\langle \mathcal{E}_j, \mathcal{E}_k \rangle = N\delta_{j,k}$, i.e. les éléments de la base sont orthogonaux entre eux. □

Si dans l'équation $\langle \mathcal{E}_j, \mathcal{E}_k \rangle = N\delta_{j,k}$ on considère $j = k = m$, alors $\langle \mathcal{E}_m, \mathcal{E}_m \rangle = N\delta_{m,m} = N$, donc :

$$\boxed{\|\mathcal{E}_m\|^2 = N}, \quad \boxed{\|\mathcal{E}_m\| = \sqrt{N}}, \qquad \forall m \in \{0, 1, \ldots, N-1\}.$$

On va examiner deux exemples dans lesquels les exponentiels complexes ont une expression particulièrement simple : $N = 2$ et $N = 4$ ($N = 3$ ne donne pas une expression de la base orthonormale de Fourier si simple).

— $N = 2$. $\ell^2(\mathbb{Z}_2) = \{z = (z(0), z(1)) \in \mathbb{C}^2\}$, dans ce cas $\mathcal{E}_m(n) = e^{2\pi i \frac{mn}{2}} = e^{\pi i mn}$ et donc :

$$m = 0 : \qquad \mathcal{E}_0 = \left(e^{\pi i 0\cdot 0}, e^{\pi i 0\cdot 1}\right) = (1, 1)$$

$$m = 1 : \qquad \mathcal{E}_1 = \left(e^{\pi i 1\cdot 0}, e^{\pi i 1\cdot 1}\right) = \left(1, e^{\pi i}\right)$$

Mais $e^{\pi i} = \cos(\pi) + i\sin(\pi) = -1$, donc $\mathcal{E}_1 = (1, -1)$. Alors,

$$\mathcal{E} = ((1, 1), (1, -1)), \tag{2.2}$$

est la base des exponentiels complexes de $\ell^2(\mathbb{Z}_2)$. Noter la présence d'une suite constante, la première, et d'une suite oscillante, la deuxième. On reviendra après sur cette forme particulière de la base.

— $N = 4$. $\ell^2(\mathbb{Z}_4) = \{z = (z(0), z(1), z(2), z(3)) \in \mathbb{C}^4\}$, la base de Fourier sera donnée par quatre suites complexes de quatre composantes. C'est laissé comme simple exercice de vérifier que la base des exponentiels complexes de $\ell^2(\mathbb{Z}_4)$ est :

$$\mathcal{E} = ((1, 1, 1, 1), (1, i, -1, -i), (1, -1, 1, -1), (1, -i, -1, i)). \tag{2.3}$$

On peut utiliser les résultats (1.10), (1.11), (1.12) de la section 1.8 pour écrire les formules suivantes, qui sont valides pour deux éléments $z, w \in \ell^2(\mathbb{Z}_N)$ quelconques :

— Décomposition sur la base orthogonale $\mathcal{E}$:

$$z = \sum_{m=0}^{N-1} \frac{\langle z, \mathcal{E}_m \rangle}{N} \mathcal{E}_m \tag{2.4}$$

— Identité de Parseval pour la base orthogonale $\mathcal{E}$:

$$\langle z, w \rangle = \sum_{m=0}^{N-1} \frac{\langle z, \mathcal{E}_m \rangle \langle \mathcal{E}_m, w \rangle}{N} \tag{2.5}$$

— Identité de Plancherel pour $\mathcal{E}$:

$$\|z\|^2 = \sum_{m=0}^{N-1} \frac{|\langle z, \mathcal{E}_m \rangle|^2}{N}. \tag{2.6}$$

On calculera explicitement les expressions ci-dessus dans la section 2.3.

On a la possibilité de renormaliser la base $\mathcal{E}$ de plusieurs manières, dans les deux sections suivantes on montrera les deux plus répandues, qui permettront aussi de définir la transformée de Fourier discrète.

2.2 La base orthonormale de Fourier de $\ell^2(\mathbb{Z}_N)$

Dans la section précédente on a vu que la norme $\ell^2(\mathbb{Z}_N)$ de toutes les suites $\mathcal{E}_m$ est $\sqrt{N}$, donc il est évident que l'on peut obtenir une base orthonormale si on divise par $\sqrt{N}$. Ceci justifie la définition suivante.

Déf. 2.2.1 *On appelle base orthonormale de Fourier de $\ell^2(\mathbb{Z}_N)$ l'ensemble*

$$E = (E_0, E_1, E_2, \ldots, E_{N-1})$$

des N suites $E_m \in \ell^2(\mathbb{Z}_N)$

$$\begin{array}{rccc} E_m & : & \mathbb{Z}_N & \longrightarrow & \mathbb{C} \\ & & n & \longmapsto & E_m(n), \end{array}$$

où :

$$\begin{cases} E_0(n) = \frac{1}{\sqrt{N}} \\ E_1(n) = \frac{1}{\sqrt{N}} e^{2\pi i \frac{n}{N}} \\ \\ E_2(n) = \frac{1}{\sqrt{N}} e^{2\pi i \frac{2n}{N}} \\ \\ \vdots \\ \\ E_{N-1}(n) = \frac{1}{\sqrt{N}} e^{2\pi i \frac{(N-1)n}{N}}. \end{cases}$$

La suite générale de la base orthonormale de Fourier est :

$$\boxed{E_m(n) = \frac{1}{\sqrt{N}} e^{2\pi i \frac{mn}{N}} = \frac{1}{\sqrt{N}} (\omega_m)^n} \qquad \forall m, n = 0, \ldots, N-1,$$

et la formule d'orthonormalité : $\langle E_j, E_k \rangle = \delta_{j,k}$ est valide.

Grâce aux formules (2.2) et (2.3) on peut dire que :

$$E = \frac{1}{\sqrt{2}} \left((1,1), (1,-1) \right) \tag{2.7}$$

est la base orthonormale de Fourier de $\ell^2(\mathbb{Z}_2)$ et

$$E = \frac{1}{2} \left((1,1,1,1), (1,i,-1,-i), (1,-1,1,-1), (1,-i,-1,i) \right). \tag{2.8}$$

est la base orthonormale de Fourier de $\ell^2(\mathbb{Z}_4)$.

La traduction du théorème 1.8.1 dans $\ell^2(\mathbb{Z}_N)$ avec la base orthonormale de Fourier est la suivante. Étant donnés les éléments $z, w \in \ell^2(\mathbb{Z}_N)$ quelconques, on a :

— Décomposition sur la base orthonormale de Fourier :

$$z = \sum_{m=0}^{N-1} \langle z, E_m \rangle E_m \tag{2.9}$$

— Identité de Parseval :

$$\langle z, w\rangle = \sum_{m=0}^{N-1} \langle z, E_m\rangle\langle E_m, w\rangle \tag{2.10}$$

— Identité de Plancherel :

$$\|z\|^2 = \sum_{m=0}^{N-1} |\langle z, E_m\rangle|^2 . \tag{2.11}$$

2.3 La base orthogonale de Fourier de $\ell^2(\mathbb{Z}_N)$

Malgré la constante de normalisation $1/\sqrt{N}$, qui apparaît dans la définition de la base orthonormale de Fourier, est le choix le plus naturel pour la normalisation de la base $\mathcal{E}$ de $\ell^2(\mathbb{Z}_N)$, dans les applications pratiques une autre normalisation est plus souvent utilisée. La raison est la suivante : avec le choix que l'on montrera tout de suite, beaucoup de formules peuvent être écrites d'une façon plus simple.

Déf. 2.3.1 *On appelle base orthogonale de Fourier de $\ell^2(\mathbb{Z}_N)$ l'ensemble*

$$F = (F_0, F_1, F_2, \ldots, F_{N-1})$$

des N suites $F_m \in \ell^2(\mathbb{Z}_N)$

$$\begin{array}{rccc} F_m & : & \mathbb{Z}_N & \longrightarrow & \mathbb{C} \\ & & n & \longmapsto & F_m(n), \end{array}$$

où :

$$\begin{cases} F_0(n) = \frac{1}{N} \\ F_1(n) = \frac{1}{N} e^{2\pi i \frac{n}{N}} \\ \\ F_2(n) = \frac{1}{N} e^{2\pi i \frac{2n}{N}} \\ \\ \vdots \\ \\ F_{N-1}(n) = \frac{1}{N} e^{2\pi i \frac{(N-1)n}{N}}. \end{cases}$$

La suite générale de la base orthogonale de Fourier est :

$$\boxed{F_m(n) = \frac{1}{N} e^{2\pi i \frac{mn}{N}} = \frac{1}{N} (\omega_m)^n} \qquad \forall m, n = 0, \ldots, N-1.$$

Les relations parmi les trois bases $\mathcal{E}$, E et F sont les suivantes :

$$E_m = \frac{\mathcal{E}_m}{\sqrt{N}}, \quad F_m = \frac{\mathcal{E}_m}{N}, \quad F_m = \frac{E_m}{\sqrt{N}} \qquad \forall m \in \{0, 1, \ldots, N\}. \tag{2.12}$$

Grâce aux formules ci-dessus on peut calculer facilement les bases orthogonales de Fourier de $\ell^2(\mathbb{Z}_2)$ et de $\ell^2(\mathbb{Z}_4)$:

— Base orthogonale de Fourier de $\ell^2(\mathbb{Z}_2)$:

$$F = \frac{1}{2}\left((1,1),(1,-1)\right) \tag{2.13}$$

— Base orthogonale de Fourier de $\ell^2(\mathbb{Z}_4)$:

$$F = \frac{1}{4}\left((1,1,1,1),(1,i,-1,-i),(1,-1,1,-1),(1,-i,-1,i)\right). \tag{2.14}$$

Toujours grâce à (2.12) on peut déterminer l'équivalent des formules (2.9) (ou 2.4), (2.10) (ou 2.5) et (2.11) (ou 2.6) pour deux éléments $z, w \in \ell^2(\mathbb{Z}_N)$ quelconques :

— Décomposition sur la base orthogonale de Fourier :

$$z = N \sum_{m=0}^{N-1} \langle z, F_m\rangle F_m$$

— Identité de Parseval pour la base orthogonale de Fourier :

$$\langle z, w\rangle = N \sum_{m=0}^{N-1} \langle z, F_m\rangle\langle F_m, w\rangle$$

— Identité de Plancherel pour la base orthogonale de Fourier :

$$\|z\|^2 = N \sum_{m=0}^{N-1} |\langle z, F_m\rangle|^2 .$$

Il est utile de résumer dans la table suivante la différence parmi les bases et les formules :

$$\mathcal{E}_m(n) = e^{2\pi i \frac{mn}{N}},\ E_m(n) = \frac{1}{\sqrt{N}} e^{2\pi i \frac{mn}{N}},\ F_m(n) = \frac{1}{N} e^{2\pi i \frac{mn}{N}}.$$

Base	Décomposition	Identité de Parseval	Identité de Plancherel
$\mathcal{E}$	$z = \sum_{m=0}^{N-1} \frac{\langle z,\mathcal{E}_m\rangle}{N}\mathcal{E}_m$	$\langle z, w\rangle = \sum_{m=0}^{N-1} \frac{\langle z,\mathcal{E}_m\rangle\langle\mathcal{E}_m,w\rangle}{N}$	$\|z\|^2 = \sum_{m=0}^{N-1} \frac{\|\langle z,\mathcal{E}_m\rangle\|^2}{N}$
E	$z = \sum_{m=0}^{N-1} \langle z, E_m\rangle E_m$	$\langle z, w\rangle = \sum_{m=0}^{N-1} \langle z, E_m\rangle\langle E_m, w\rangle$	$\|z\|^2 = \sum_{m=0}^{N-1} \|\langle z, E_m\rangle\|^2$
F	$z = N \sum_{m=0}^{N-1} \langle z, F_m\rangle F_m$	$\langle z, w\rangle = N \sum_{m=0}^{N-1} \langle z, F_m\rangle\langle F_m, w\rangle$	$\|z\|^2 = N \sum_{m=0}^{N-1} \|\langle z, F_m\rangle\|^2 .$

TABLE 2.1 – Différentes normalisations des bases de Fourier, et formules relatives.

2.4 Les coefficients de Fourier et la transformée de Fourier discrète (DFT)

La définition de la transformée de Fourier discrète change selon les auteurs et les applications. Les deux définitions les plus répandues utilisent la base orthonormale E et un mélange des bases orthogonales $\mathcal{E}$ et F.

Les deux choix sont utiles pour des raisons différentes :

— Utiliser la base orthonormale E permet d'obtenir des opérateurs unitaires ;

— Utiliser un mélange des bases orthogonales $\mathcal{E}$ et F permet de simplifier beaucoup de formules, notamment la formule de la convolution, que l'on verra plus tard dans ce chapitre, qui est très utilisée dans les applications.

On travaillera avec les formules qui viennent des bases orthogonales $\mathcal{E}$ et F. Cette décision a été prise principalement par cohérence avec les formules de beaucoup de logiciels, notamment MATLAB.

On commence par reconsidérer la décomposition

$$z = \sum_{m=0}^{N-1} \frac{\langle z, \mathcal{E}_m \rangle}{N} \mathcal{E}_m = \sum_{m=0}^{N-1} \langle z, \mathcal{E}_m \rangle \frac{\mathcal{E}_m}{N},$$

mais $\mathcal{E}_m/N = F_m$, donc :

$$z = \sum_{m=0}^{N-1} \langle z, \mathcal{E}_m \rangle F_m,$$

i.e. on peut décomposer un élément $z \in \ell^2(\mathbb{Z}_N)$ quelconque sur la base orthogonale de Fourier F avec des composantes données par les produits scalaires de z avec les éléments de la base $\mathcal{E}$.

En rappelant la définition du produit scalaire de $\ell^2(\mathbb{Z}_N)$, on peut écrire :

$$\begin{aligned} \langle z, \mathcal{E}_m \rangle = \sum_{n=0}^{N-1} z(n) \overline{\mathcal{E}_m(n)} &= \sum_{n=0}^{N-1} z(n) \overline{e^{2\pi i \frac{mn}{N}}} \\ &= \sum_{n=0}^{N-1} z(n) e^{-2\pi i \frac{mn}{N}}. \end{aligned}$$

Déf. 2.4.1 *Étant donné $z \in \ell^2(\mathbb{Z}_N)$ quelconque, on appelle les valeurs complexes $\langle z, \mathcal{E}_m \rangle$, $m \in \{0, 1, \ldots, N-1\}$, les coefficients de Fourier de z, qu'on écrit avec $\hat{z}(m)$. Explicitement :*

$$\boxed{\hat{z}(m) = \sum_{n=0}^{N-1} z(n) e^{-2\pi i \frac{mn}{N}}} \quad \ll \textbf{\textit{Coefficients de Fourier de z}} \gg . \tag{2.15}$$

On écrit avec $\hat{z} \in \ell^2(\mathbb{Z}_N)$ la suite des coefficients de Fourier de z :

$$\hat{z} = (\hat{z}(0), \hat{z}(1), \hat{z}(2), \ldots, \hat{z}(N-1)). \tag{2.16}$$

On appelle l'opérateur linéaire qui transforme $z \in \ell^2(\mathbb{Z}_N)$ en la suite $\hat{z} \in \ell^2(\mathbb{Z}_N)$ de ses coefficients de Fourier, i.e.

$$\begin{array}{rccc} \mathrm{DFT} \equiv \hat{} \equiv \mathcal{F} \ : & \ell^2(\mathbb{Z}_N) & \longrightarrow & \ell^2(\mathbb{Z}_N) \\ & z & \longmapsto & \mathrm{DFT}(z) \equiv \hat{z} \equiv \mathcal{F}(z), \end{array}$$

$$\hat{z}(m) = \sum_{n=0}^{N-1} z(n) e^{-2\pi i \frac{mn}{N}} \qquad \forall m \in \{0, 1, \ldots, N-1\},$$

la ***transformée de Fourier discrète****, qu'on écrira* DFT « *Discrete Fourier Transform* »*, à partir de maintenant.*

Il faut observer **la variable de z est n, tandis que la variable de $\hat{z}$ est m**. Dans la section 2.6 on donnera l'interprétation de n et m dans la théorie des signaux : n est la valeur discrète d'un instant temporel (ou d'une position spatiale), où on mesure un signal z, par contre m est proportionnel à la fréquence d'oscillation d'une onde (dite harmonique), multiple d'une fréquence fondamentale. Donc, la DFT permet de passer d'une description en termes d'échantillons temporels (ou spatiales) d'un signal, à une description en termes de fréquences du signal même. Dans la section 2.6 on formalisera cette affirmation.

Avec les définitions ci-dessus, on peut écrire la décomposition de z comme ceci :

$$z = \sum_{m=0}^{N-1} \hat{z}(m) F_m, \tag{2.17}$$

i.e. les coefficients de Fourier de z sont les composantes de z dans la base orthogonale de Fourier F :

$$\hat{z} = [z]_F. \tag{2.18}$$

Si on utilise la notation que l'on vient d'introduire, on peut réécrire le théorème de décomposition sur la base orthonormale de Fourier et les identités de Parseval et Plancherel comme cela :

— **Décomposition de z sur la base orthogonale de Fourier** :

$$\boxed{z(n) = \frac{1}{N} \sum_{m=0}^{N-1} \hat{z}(m) e^{2\pi i \frac{mn}{N}}} \qquad \forall n = 0, 1, \ldots, N-1 \tag{2.19}$$

— **Identité de Parseval** :

$$\boxed{\langle z, w \rangle = \frac{1}{N} \sum_{m=0}^{N-1} \hat{z}(m) \overline{\hat{w}(m)} = \frac{1}{N} \langle \hat{z}, \hat{w} \rangle} \tag{2.20}$$

— **Identité de Plancherel** :

$$\boxed{\|z\|^2 = \frac{1}{N} \sum_{m=0}^{N-1} |\hat{z}(m)|^2 = \frac{1}{N} \|\hat{z}\|^2}. \tag{2.21}$$

2.4.1 La transformée de Fourier inverse (IDFT)

Il est intéressant de comparer les formules (2.15) et (2.19) :

$$\hat{z}(m) = \sum_{n=0}^{N-1} z(n) e^{-2\pi i \frac{mn}{N}}, \quad z(n) = \frac{1}{N} \sum_{m=0}^{N-1} \hat{z}(m) e^{2\pi i \frac{mn}{N}}, \quad \forall n, m \in \{0, 1, \ldots, N-1\}.$$

La première relation dit que, si on connait les valeurs $z(n)$, alors on peut reconstruire les valeurs $\hat{z}(m)$ grâce à la formule (2.15).

La deuxième relation dit que, si on connait les valeurs $\hat{z}(m)$, alors on peut reconstruire les valeurs $z(n)$ grâce à la formule (2.19).

Ceci montre une « dualité » entre les deux formules : on peut passer de la suite z à la suite $\hat{z}$ et vice-versa via les relations (2.15) et (2.19).

On va formaliser cette dualité avec la définition et le théorème qui suivent.

Déf. 2.4.2 *On appelle l'opérateur linéaire :*

$$\begin{array}{rccccc} \text{IDFT} \equiv \check{} \equiv \mathcal{F}^{-1} & : & \ell^2(\mathbb{Z}_N) & \longrightarrow & \ell^2(\mathbb{Z}_N) \\ & & z & \longmapsto & \text{IDFT}(z) \equiv \check{z} \equiv \mathcal{F}^{-1}(z), \end{array}$$

$$\boxed{\check{z}(n) = \frac{1}{N} \sum_{m=0}^{N-1} z(m) e^{2\pi i \frac{mn}{N}}} \qquad \forall n \in \{0, 1, \ldots, N-1\},$$

la ***transformée de Fourier discrète inverse****, que l'on écrira* IDFT *« Inverse Discrete Fourier Transform », à partir de maintenant.*

Théorème 2.4.1 *La* IDFT *est l'opérateur linéaire inverse de la* DFT *et vice-versa :*

$$\text{IDFT} = \text{DFT}^{-1}, \quad \text{DFT} = \text{IDFT}^{-1},$$

autrement dit :

$$\check{\hat{z}} = z, \quad \hat{\check{z}} = z \qquad \forall z \in \ell^2(\mathbb{Z}_N).$$

Preuve. On doit démontrer que la composition entre DFT et IDFT et entre IDFT et DFT donne l'opérateur identité id : DFT$\circ$IDFT=IDFT$\circ$DFT= id, $id(z) = z$, $\forall z \in \ell^2(\mathbb{Z}_N)$.

On commence par vérifier que, si on a une suite $z \in \ell^2(\mathbb{Z}_N)$ quelconque, et on applique la DFT pour obtenir la suite des coefficients de Fourier $\hat{z} \in \ell^2(\mathbb{Z}_N)$, alors on peut revenir à la suite initiale via l'application de la IDFT :

$$\begin{array}{ccccccc} \ell^2(\mathbb{Z}_N) & \underset{\text{DFT}}{\longrightarrow} & \ell^2(\mathbb{Z}_N) & \underset{\text{IDFT}}{\longrightarrow} & \ell^2(\mathbb{Z}_N) \\ z & \longmapsto & \hat{z} & \longmapsto & \check{\hat{z}} = z. \end{array}$$

Avant d'écrire la composition, on souligne qu'il ne faut pas confondre l'indice de sommation, dont le symbole n'a aucune importance, avec les variables fixées n, m de $\check{z}(n)$ et $\hat{z}(m)$. Pour éviter ce problème, on utilise j comme symbole de sommation de la première transformation, qui est celle écrite à l'intérieur de l'expression composée.

$$
\begin{aligned}
\check{\hat{z}}(n) &= \frac{1}{N}\sum_{m=0}^{N-1}\hat{z}(m)e^{2\pi i\frac{mn}{N}} = \frac{1}{N}\sum_{m=0}^{N-1}\left(\sum_{j=0}^{N-1}z(j)e^{-2\pi i\frac{mj}{N}}\right)e^{2\pi i\frac{mn}{N}} \\
&= \frac{1}{N}\sum_{m=0}^{N-1}\sum_{j=0}^{N-1}z(j)e^{2\pi im\frac{n-j}{N}} \\
&= \frac{1}{N}\sum_{j=0}^{N-1}z(j)\left(\sum_{m=0}^{N-1}e^{2\pi im\frac{n-j}{N}}\right) \\
&\underset{\text{(Lemme 2.1)}}{=} \frac{1}{N}\sum_{j=0}^{N-1}z(j)N\delta_{j,n} \\
&= z(n) \quad \forall n\in\{0,1,\ldots,N-1\}.
\end{aligned}
$$

Maintenant on vérifie que la composition inverse donne encore l'identité :

$$
\begin{array}{ccccccc}
\ell^2(\mathbb{Z}_N) & \underset{\text{IDFT}}{\longrightarrow} & \ell^2(\mathbb{Z}_N) & \underset{\text{DFT}}{\longrightarrow} & \ell^2(\mathbb{Z}_N) \\
z & \longmapsto & \check{z} & \longmapsto & \hat{\check{z}} = z.
\end{array}
$$

$$
\begin{aligned}
\hat{\check{z}}(m) &= \sum_{n=0}^{N-1}\check{z}(n)e^{-2\pi i\frac{mn}{N}} = \sum_{n=0}^{N-1}\left(\frac{1}{N}\sum_{j=0}^{N-1}z(j)e^{2\pi i\frac{jn}{N}}\right)e^{-2\pi i\frac{mn}{N}} \\
&= \frac{1}{N}\sum_{n=0}^{N-1}\sum_{j=0}^{N-1}z(j)e^{2\pi in\frac{j-m}{N}} \\
&= \frac{1}{N}\sum_{j=0}^{N-1}z(j)\left(\sum_{n=0}^{N-1}e^{2\pi in\frac{j-m}{N}}\right) \\
&\underset{\text{(Lemme 2.1)}}{=} \frac{1}{N}\sum_{j=0}^{N-1}z(j)N\delta_{j,m} \\
&= z(m) \quad \forall m\in\{0,1,\ldots,N-1\}.
\end{aligned}
$$

Donc, $\check{\hat{z}}(n) = z(n)$ et $\hat{\check{z}}(m) = z(m)$, $\forall n,m\in\{0,1,\ldots,N-1\}$ et le théorème est prouvé. □

On note la grande similarité entre DFT et IDFT : seulement le coefficient $1/N$ et le signe de l'exponentiel complexe changent.

C'est utile de souligner les formules qu'on vient de démontrer :

$$
\boxed{\check{\hat{z}}(n) = \frac{1}{N}\sum_{m=0}^{N-1}\hat{z}(m)e^{2\pi i\frac{mn}{N}} = z(n)} \qquad \forall n\in\mathbb{Z}_N,
$$

$$
\boxed{\hat{\check{z}}(m) = \sum_{n=0}^{N-1}\check{z}(n)e^{-2\pi i\frac{mn}{N}} = z(m)} \qquad \forall m\in\mathbb{Z}_N.
$$

Déf. 2.4.3 *On appelle le couple $(z,\hat{z})\in\ell^2(\mathbb{Z}_N)\times\ell^2(\mathbb{Z}_N)$ un* ***couple de Fourier****.*

2.4.2 Définition de DFT et IDFT avec la base orthonormale de Fourier

Une définition alternative des coefficients de Fourier, DFT et IDFT, que l'on trouve surtout dans la littérature mathématique plus théorique, utilise la base orthonormale de Fourier E :

Analyse de Fourier discrète orthonormale :

— $z, w \in \ell^2(\mathbb{Z}_N)$;

— Coefficients de Fourier :

$$\hat{z}(m) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} z(n) e^{-2\pi i \frac{mn}{N}}, \tag{2.22}$$

l'écriture $\hat{z}$ dans les formules suivantes de cette liste (et seulement celles-ci) fait référence aux coefficients de Fourier ci-dessus.

— Décomposition sur la base orthonormale de Fourier :

$$z(m) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \hat{z}(n) e^{2\pi i \frac{mn}{N}}.$$

— DFT :

$$\hat{z}(m) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} z(n) e^{-2\pi i \frac{mn}{N}} \qquad \forall m \in \{0, 1, \ldots, N-1\}.$$

— IDFT :

$$\check{z}(n) = \frac{1}{\sqrt{N}} \sum_{m=0}^{N-1} z(m) e^{2\pi i \frac{mn}{N}} \qquad \forall n \in \{0, 1, \ldots, N-1\}.$$

— Identité de Parseval :

$$\langle z, w \rangle = \sum_{m=0}^{N-1} \hat{z}(m) \overline{\hat{w}(m)} = \langle \hat{z}, \hat{w} \rangle.$$

— Identité de Plancherel :

$$\|z\|^2 = \sum_{m=0}^{N-1} |\hat{z}(m)|^2 = \|\hat{z}\|^2.$$

On voit que l'avantage le plus important d'utiliser la base orthonormale de Fourier pour définir les objets de l'analyse de Fourier discrète est que la DFT et la IDFT sont des opérateurs qui conservent le produit scalaire, et donc aussi la norme, par conséquent, ils sont représentés avec des matrices unitaires.

On observe aussi que, indépendamment de la définition que l'on utilise, *le produit des coefficients de $\hat{z}$ et $\check{z}$ doit toujours être égal à $1/N$ pour garantir que* IDFT=DFT^{-1}.

2.4.3 La base (orthonormale) de Fourier réelle

On peut écrire la base de Fourier et la DFT avec une notation réelle. L'avantage d'avoir une base de Fourier réelle est que, si z est réel, alors on peut éviter l'introduction de composantes imaginaires. On considère simplement la base de Fourier orthonormale pour simplicité.

Il faut distinguer si N est paire ou impaire. On commence par le cas N paire : $N = 2M$, $M \in \mathbb{N}$, $M \geqslant 1$. Alors, $\forall n = 0, 1, \ldots, N-1$, on écrit :

$$\begin{cases} c_0(n) = \frac{1}{\sqrt{N}} \\ c_m(n) = \sqrt{\frac{2}{N}} \cos\left(\frac{2\pi mn}{N}\right) \qquad m = 1, 2, ..., M-1 \\ c_M(n) = \frac{1}{\sqrt{N}} \cos\left(\frac{2\pi \frac{N}{2} n}{N}\right) = \frac{(-1)^n}{\sqrt{N}} \\ s_m(n) = \sqrt{\frac{2}{N}} \sin\left(\frac{2\pi mn}{N}\right) \qquad m = 1, 2, \ldots, M-1. \end{cases}$$

Si $N = 2M+1$ est impaire, alors on définit encore c_0, c_m et s_m comme ci-dessus, mais bien sûr le cas $m = N/2$ ne doit pas être considéré car $N/2$ dans ce cas n'est pas un nombre entier.

Théorème 2.4.2 *L'ensemble $\{c_0, c_1, \ldots, c_{M-1}, c_M, s_1, \ldots, s_{M-1}\}$ quand $N = 2M$ ou l'ensemble $\{c_0, c_1, \ldots, c_{M-1}, s_1, \ldots, s_{M-1}\}$ quand $N = 2M+1$ est une base orthonormale de $\ell^2(\mathbb{Z}_N)$. Donc, pour tout $z \in \ell^2(\mathbb{Z}_N)$:*

$$z = \sum_{m=0}^{M} \langle z, c_m \rangle c_m + \sum_{m=1}^{M-1} \langle z, s_m \rangle s_m \quad (N = 2M), \quad z = \sum_{m=0}^{M-1} \langle z, c_m \rangle c_m + \sum_{m=1}^{M-1} \langle z, s_m \rangle s_m \quad (N = 2M+1).$$

Déf. 2.4.4 *On appelle base orthonormale réelle de Fourier de $\ell^2(\mathbb{Z}_N)$ l'ensemble de suites de $\ell^2(\mathbb{Z}_N)$ $\{c_0, c_1, \ldots, c_{M-1}, c_M, s_1, \ldots, s_{M-1}\}$ quand $N = 2M$, ou l'ensemble de suites de $\ell^2(\mathbb{Z}_N)$ $\{c_0, c_1, \ldots, c_{M-1}, s_1, \ldots, s_{M-1}\}$ quand $N = 2M+1$.*

La relation avec les coefficients de Fourier est donnée par les formules suivantes.

$$\begin{cases} \langle z, c_0 \rangle = \frac{\hat{z}(0)}{\sqrt{N}} \\ \langle z, c_M \rangle = \frac{\hat{z}(M)}{\sqrt{N}} \\ \langle z, c_m \rangle = \frac{1}{\sqrt{2N}} (\hat{z}(m) + \hat{z}(N-m)), \quad m = 1, 2, \ldots, M-1 \\ \langle z, s_m \rangle = \frac{-i}{\sqrt{2N}} (\hat{z}(m) - \hat{z}(N-m)), \quad m = 1, 2, \ldots, M-1 \\ \hat{z}(0) = \sqrt{N} \langle z, c_0 \rangle \\ \hat{z}(M) = \sqrt{N} \langle z, c_M \rangle \\ \hat{z}(m) = \sqrt{N/2} (\langle z, c_m \rangle - i \langle z, s_m \rangle), \quad m = 1, 2, \ldots, M-1 \\ \hat{z}(m) = \sqrt{N/2} (\langle z, c_{N-m} \rangle + i \langle z, s_{N-m} \rangle), \quad m = M+1, M+2, \ldots, N-1. \end{cases}$$

2.5 Interprétation matricielle de la DFT et IDFT

Par définition, la DFT transforme les suites de $z \in \ell^2(\mathbb{Z}_N)$ représentées dans la base canonique B de $\ell^2(\mathbb{Z}_N)$ en suites de $\ell^2(\mathbb{Z}_N)$ représentées dans la base orthogonale F de Fourier de $\ell^2(\mathbb{Z}_N)$ (2.17) :

$$\begin{array}{rccc} \text{DFT}: & \ell^2(\mathbb{Z}_N) & \longrightarrow & \ell^2(\mathbb{Z}_N) \\ & z = [z]_B & \longmapsto & \text{DFT}(z) = \hat{z} = [z]_F \end{array}$$

Donc, la DFT est l'opérateur de passage de la base canonique B de $\ell^2(\mathbb{Z}_N)$ à la base de Fourier F de $\ell^2(\mathbb{Z}_N)$ et, par conséquent, la IDFT est l'opérateur de passage inverse, de F à B.

On veut trouver la représentation matricielle des opérateurs linéaires DFT et IDFT. Pour cela, on introduit une notation habituelle dans la littérature relative à la DFT : $\omega_N = e^{-\frac{2\pi i}{N}}$. Grâce aux propriétés des exponentiels complexes on peut écrire :

$$\omega_N^{mn} = e^{-2\pi i \frac{mn}{N}}$$

et donc les coefficients de Fourier peuvent être réécrits comme ceci :

$$\hat{z}(m) = \sum_{n=0}^{N-1} z(n) e^{-2\pi i \frac{mn}{N}} = \sum_{n=0}^{N-1} z(n) \omega_N^{mn}.$$

On définit la matrice W_N qui a comme éléments ω_N^{mn} :

$$w_{mn} = \omega_N^{mn},$$

i.e. explicitement :

$$W_N = \begin{pmatrix} 1 & 1 & 1 & 1 & \dots & 1 \\ 1 & \omega_N & \omega_N^2 & \omega_N^3 & \dots & \omega_N^{N-1} \\ 1 & \omega_N^2 & \omega_N^4 & \omega_N^6 & \dots & \omega_N^{2(N-1)} \\ 1 & \omega_N^3 & \omega_N^6 & \omega_N^9 & \dots & \omega_N^{3(N-1)} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega_N^{N-1} & \omega_N^{2(N-1)} & \omega_N^{3(N-1)} & \dots & \omega_N^{(N-1)(N-1)} \end{pmatrix}.$$

Cette matrice $N \times N$ est dite **matrice de Sylvester** [1]. Elle est **symétrique** : $W_N = W_N^t$, i.e. $w_{mn} = w_{nm}$ (ce qui est une conséquence évidente de la définition de w_{mn} car $\omega_N^{mn} = \omega_N^{nm}$) et chaque ligne ou colonne est donnée par la progression géométrique [2] d'une puissance de ω_N. Une telle matrice est dite de Vandermonde [3].

La convention intrinsèque quand on considère W_N consiste à considérer la variabilité de l'indice des lignes et des colonnes entre 0 et $N-1$ (au lieu de la variabilité canonique de 1 jusqu'à N), c'est grâce à cette convention qu'on a tous les éléments de la première ligne ($m = 0$) et de la première colonne ($n = 0$) égaux à 1.

1. James Joseph Sylvester (1814 London - 1897 London).

2. On rappelle qu'une progression géométrique de raison r est la suite de puissances $1 = r^0, r = r^1, r^2, r^3, \dots, r^n$.

3. Alexandre-Théophile Vandermonde (1735 Paris - 1796 Paris).

Si on applique W_N à z *identifié avec un vecteur colonne* de $\mathbb{C}^N$, alors, par définition de produit matriciel, on obtient un vecteur $W_N z$ dont la m-ième composante $(W_N z)(m)$ est donnée par [4] :

$$(W_N z)(m) = \sum_{n=0}^{N-1} w_{mn} z(n) = \sum_{n=0}^{N-1} z(n) e^{-2\pi i \frac{mn}{N}} = \hat{z}(m),$$

ça vaut $\forall m \in \mathbb{Z}_N$, donc :

$$\boxed{\hat{z} = W_N z} \qquad \forall z \in \ell^2(\mathbb{Z}_N).$$

Avec les mêmes considérations, on peut vérifier que la IDFT est implémentée via la matrice conjuguée de W_N normalisée par le coefficient $1/N$ (la transposition n'est pas nécessaire car W_N est symétrique) :

$$\boxed{W_N^{-1} = \frac{1}{N}\overline{W_N}}, \quad \boxed{\check{z} = W_N^{-1} z} \qquad \forall z \in \ell^2(\mathbb{Z}_N).$$

W_N est la *matrice de passage de base de B à F*, et $W_N^{-1} = \frac{1}{N}\overline{W_N}$ est la *matrice de passage de base de F à B*.

Observation : si on utilise la définition de DFT correspondant à l'eq. (2.22), i.e. via la base orthonormale de Fourier, alors la matrice associée devient $\tilde{W}_N = W_N/\sqrt{N}$, qui est une matrice unitaire, et donc sa matrice inverse est $\overline{\tilde{W}_N}$.

Exemples :

— $N = 2 : \omega_2 = e^{-2\pi i/2} = e^{-i\pi} = \cos(\pi) - i\sin(\pi) = -1$, donc

$$W_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

d'où :

$$W_2^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

— $N = 4 : \omega_4 = e^{-2\pi i/4} = e^{-i\pi/2} = \cos(\pi/2) - i\sin(\pi/2) = -i$, donc

$$W_4 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & (-i)^2 & (-i)^3 \\ 1 & (-i)^2 & (-i)^4 & (-i)^6 \\ 1 & (-i)^3 & (-i)^6 & (-i)^9 \end{pmatrix},$$

d'où :

$$W_4 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{pmatrix}. \tag{2.23}$$

4. C'est le produit Euclidien réel de la m-ième ligne de W_N, i.e. $(w_{m0}, w_{m1}, \ldots, w_{m(N-1)})$ fois les composantes $(z(0), z(1), \ldots, z(n))$ de z.

La matrice inverse est :

$$W_4^{-1} = \frac{1}{4}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{pmatrix}. \tag{2.24}$$

On observe que la matrice W_4^{-1} a comme colonnes la base *orthogonale* F de $\ell^2(\mathbb{Z}_4)$ comme vu dans la formule eq. (2.14), ce qui est cohérent avec le fait qu'elle est la matrice de passage de la base *orthogonale* F à la base canonique de $\ell^2(\mathbb{Z}_4)$.

2.5.1 « FFT » : Fast Fourier Transform

On vient de voir que l'action de la DFT sur un signal $z \in \ell^2(\mathbb{Z}_N)$ peut être représentée comme un produit matriciel. Par conséquent, on a besoin de calculer N multiplications pour chaque élément $\hat{z}(m)$ de la suite $\hat{z} \in \ell^2(\mathbb{Z}_N)$. Comme $\hat{z}$ a N composantes, l'algorithme de calcul de la DFT a une complexité de $\mathcal{O}(N^2)$.

Pour des signaux de grande dimension, cette complexité implique que la DFT est très lente, c'est pour cela que la transformée de Fourier a été utilisée presque seulement dans un cadre théoriques, plutôt que dans les applications, jusqu'aux années 60 du XX siècle.

Heureusement, **Cooley et Tukey**, en 1965, ont utilisé des symétries cachées dans la DFT pour construire un algorithme rapide pour le calcul de la DFT, ils ont appelé l'algorithme Fast Fourier Transform « FFT ».

La FFT a une complexité de l'ordre de $\mathcal{O}(N \log N)$ et elle permet de calculer la transformée de Fourier d'un signal de grande dimension dans l'ordre d'une fraction de secondes avec les ordinateurs modernes.

En particulier, la FFT est très efficace quand la dimension des signaux est une puissance de 2. Cela explique pourquoi le format typique des images numériques est de 512 ou 1024, comme ceci on peut manipuler ces images efficacement avec la FFT.

Le développement de la FFT est considéré comme *une des plus grandes avancées scientifique du XXème siècle*, car il a permis d'utiliser dans une quantité énorme d'applications pratiques la transformée de Fourier.

2.6 La transformée de Fourier dans le traitement des signaux

La théorie de Fourier est appliquée dans beaucoup de domaines, par exemple la résolution des équations différentielles ordinaires et partielles, la physique classique et quantique, les statistiques et probabilités et le traitement des signaux.

C'est surtout dans le premier et le dernier domaine qu'on appliquera les résultats de l'analyse de Fourier. Dans cette section on commencera à voir l'importance de la théorie de Fourier dans le traitement des signaux en dimension 1 (1D).

2.6.1 La formule de synthèse des signaux 1D : décomposition sur la base des harmoniques

Un signal discret 1D de dimension N peut être défini comme l'ensemble des N échantillons d'une variable, qui peut être dépendant du temps, d'une dimension spatiale (x,y ou z), où d'un autre paramètre avec un seul degré de liberté.

Deux exemples remarquables de signaux discret 1D qui dépendent du temps ou d'une dimension spatiale sont :

— l'ensemble de valeurs d'intensité d'un morceau de musique échantillonné en correspondance de N instants temporels différents ;

— l'ensemble de valeurs de niveau de gris d'une ligne ou d'une colonne d'une image échantillonnée en correspondance de N positions différentes.

On peut traiter un signal discret 1D dans le cadre de la théorie de Fourier avec les identifications basiques suivantes :

— la représentation mathématique abstraite d'un signal discret 1D est donnée par une suite $z \in \ell^2(\mathbb{Z}_N)$;

— $n \in \mathbb{Z}_N = \{0, 1, \ldots, N-1\}$ représente la valeur du paramètre (temps, dimension spatiale, etc.) où on mesure les échantillons du signal. L'unité de mesure de n est typiquement le second ou le mètre ;

— l'énergie du signal z est associée à la norme carrée $\|z\|^2$.

On veut maintenant interpréter la formule de décomposition sur la base de Fourier, les coefficients de Fourier, la DFT et IDFT et l'identité de Plancherel dans le cadre du traitement des signaux.

L'interprétation de l'identité de Plancherel est la plus simple : l'énergie du signal z est décomposée dans la somme des modules carrés des coefficients de Fourier.

La formule de décomposition sur la base de Fourier, eq. (2.19), dans le domaine du traitement des signaux est dite

$$\textbf{Formule de synthèse} \ : z(n) = \frac{1}{N} \sum_{m=0}^{N-1} \hat{z}(m) e^{2\pi i \frac{mn}{N}} \qquad \forall n \in \mathbb{Z}_N.$$

Avec cette formule on peut reconstruire (ou synthétiser, d'où le nom) le signal z grâce à la connaissance des coefficients de Fourier $\hat{z}(m)$ et aux fonctions oscillantes $e^{2\pi i \frac{mn}{N}}$.

Les fonctions qu'on utilise dans l'opération de synthèse du signal sont :

$$e^{2\pi i \frac{m}{N} n} = \cos\left(2\pi \frac{m}{N} n\right) + i \sin\left(2\pi \frac{m}{N} n\right). \tag{2.25}$$

Quand $m = 0$ il n'y a pas d'oscillations, mais à partir de $m = 1$ jusqu'à $m = N - 1$ les fonctions $e^{2\pi i \frac{m}{N} n}$ oscillent avec une certaine fréquence (donc l'unité de mesure m est le *hertz* ou le rad/s) qu'on va discuter en détaille dans la section 2.6.4. Ces fonctions ont un nom spéciale, dérivé de la musique, comme le montre la définition suivante.

Déf. 2.6.1 (Harmoniques) *La fonction $n \mapsto e^{2\pi i \frac{1}{N} n}$ est dite harmonique (discrète*[5]*) fondamentale et les fonctions $n \mapsto e^{2\pi i \frac{m}{N} n}$ pour $m = 2, \ldots, N-1$ sont dites harmoniques (discrètes) d'ordre (supérieur) m.*

2.6.2 Signification des coefficients de Fourier et spectres d'un signal 1D

La formule de synthèse montre que le signal z dans la valeur n de son paramètre est reconstruit par une combinaison linéaire d'ondes harmoniques avec fréquences multiples de $1/N$ via le coefficient de multiplication $m : \{0, 1/N, 2/N, \ldots, (N-1)/N\}$. Les scalaires complexes de la combinaison linéaire sont les coefficients de Fourier $\hat{z}(m)$.

Chaque coefficient de Fourier $\hat{z}(m) \in \mathbb{C}$ peut être écrit ainsi :

$$\hat{z}(m) = a(m) + ib(m) = |\hat{z}(m)| e^{i \mathrm{Arg}(\hat{z}(m))}$$

où $|\hat{z}(m)| = \sqrt{a(m)^2 + b(m)^2}$ est le module du coefficient de Fourier $\hat{z}(m)$ et $\mathrm{Arg}(\hat{z}(m)) = \arctan\left(\frac{b(m)}{a(m)}\right)$ est son argument.

Il est donc évident que le « poids » qui mesure l'importance de chaque harmonique $e^{2\pi i \frac{mn}{N}}$ dans la reconstruction du signal z est le module[6] du coefficient de Fourier $\hat{z}(m)$:

$|\hat{z}(m)|$: **mesure de l'importance de l'harmonique $e^{2\pi i \frac{mn}{N}}$ dans la reconstruction de z.**

Dû à cela, dans le cadre du traitement des signaux, on appelle la formule des coefficients de Fourier :

$$\textbf{Formule d'analyse} \ : \hat{z}(m) = \sum_{n=0}^{N-1} z(n) e^{-2\pi i \frac{mn}{N}} \qquad \forall m \in \mathbb{Z}_N,$$

car avec $\hat{z}$ on peut *analyser les composantes fréquentielles d'un signal.*

Si le signal discret z dépend du temps t (ou d'une dimension spatiale x), alors on dit que la transformation $z \to \hat{z}$ réalisée par la DFT permet de passer de la **représentation temporelle (ou spatiale)** du signal à la **représentation fréquentielle**, ou à l'**espace de Fourier**.

La transformée de Fourier est souvent définie comme *l'équivalent du prisme de Newton pour les mathématiques.* Le prisme de Newton permet de décomposer la lumière dans les composantes fréquentielles « cachées » qui correspondent aux couleurs spectrales. La transformée de Fourier nous permet de mettre en évidence les composantes fréquentielles « cachées » dans un signal quelconque.

Cette analogie a amené à donner les définitions suivantes.

---

5. On spécifie le fait que les harmoniques sont discrètes, car les harmonique continues sont données par les fonctions $t \mapsto e^{2\pi i m \nu t} = e^{im\omega t}$, où ν est la fréquence et $\omega = 2\pi\nu$ est la pulsation.

6. Il faut prendre le module car les nombres complexes ne sont pas ordonnés.

Déf. 2.6.2 *Étant donné $z \in \ell^2(\mathbb{Z}_N)$, on appelle :*

- $\{|\hat{z}(m)|,\ m \in \mathbb{Z}_N\}$ *:* ***spectre d'amplitude*** *de z, ou plus simplement* ***spectre*** *de z ;*
- $\{|\hat{z}(m)|^2,\ m \in \mathbb{Z}_N\}$ *:* ***spectre de puissance*** *de z ;*
- $\{Arg(\hat{z}(m)),\ m \in \mathbb{Z}_N\}$ *:* ***spectre de phases*** *de z.*

On reviendra sur la signification de ces spectres plus tard.

Maintenant, on veut observer que, parmi les coefficients de Fourier il y en a un spécial : $\hat{z}(0)$, qui donne une information sur la valeur moyenne de z, en fait :

$$\hat{z}(0) = \sum_{n=0}^{N-1} z(n) e^{2\pi i \frac{0n}{N}} = \sum_{n=0}^{N-1} z(n) = N\langle z \rangle \quad \Longrightarrow \quad \boxed{\hat{z}(0) = N\langle z \rangle},$$

où $\langle z \rangle = \frac{1}{N} \sum_{n=0}^{N-1} z(n)$ est la valeur moyenne du signal z.

Si on introduit cette expression de $\hat{z}(0)$ dans la formule de synthèse et on sépare le premier terme du reste de la sommation on obtient :

$$z(n) = \frac{1}{N} N\langle z \rangle + \frac{1}{N} \sum_{m=1}^{N-1} \hat{z}(m) e^{2\pi i \frac{mn}{N}}, \quad \text{i.e.} \quad \boxed{z(n) = \langle z \rangle + \frac{1}{N} \sum_{m=1}^{N-1} \hat{z}(m) e^{2\pi i \frac{mn}{N}}},$$

on dit que le coefficient de Fourier $\hat{z}(0)$ est la composante « DC » de la formule de synthèse, tandis que les autres termes constituent la composante « AC ». Cette nomenclature vient de l'électrotechnique, où on appelle le courant continue DC « Direct Current » (courant de fréquence nulle) et le courant alternatif AC « Alternating Current ».

On peut interpréter la formule qu'on vient de déterminer en disant que *z est décomposée dans la somme de sa valeur moyenne et des détails plus fins reconstruits par les harmoniques d'ordre supérieur avec des poids donnés par les modules des coefficients de Fourier de z.*

2.6.3 La formule de synthèse et les coefficients de Fourier de l'impulsion unitaire δ

C'est instructif de comparer la formule de synthèse avec la formule (2.1), i.e.

$$\sum_{n=0}^{N-1} e^{\pm 2\pi i n \frac{j-k}{N}} = N\delta_{j,k} = \begin{cases} N & j = k \\ 0 & j \neq k, \end{cases}$$

si on réécrit $j - k = m \in \mathbb{Z}_N$, on échange m avec n (ce qui est possible car sont deux valeurs quelconques de $\mathbb{Z}_N$) et on normalise par N, alors la formule devient

$$\frac{1}{N} \sum_{m=0}^{N-1} e^{\pm 2\pi i n \frac{m}{N}} = \begin{cases} 1 & n = 0 \\ 0 & n = 1, \ldots, N-1 \end{cases} = e_0(n) \equiv \delta_0(n) \equiv \delta(n).$$

On appelle δ l'*impulsion unitaire*. La formule précédente est donc la formule de synthèse de l'impulsion unitaire, dans laquelle on voit que tous les coefficients de Fourier sont unitaires :

$$\widehat{\delta_0}(m) = 1 = \left|\widehat{\delta_0}(m)\right| \qquad \forall m \in \mathbb{Z}_N.$$

Ce résultat est très profond : *la DFT transforme un signal complètement « localisé » dans une valeur de son paramètre en un signal complètement « délocalisé » dans son spectre : les harmoniques relatives à toutes les fréquences interviennent avec le même poids dans la reconstruction du signal.*

La généralisation de ce comportement à espaces plus compliqués que $\ell^2(\mathbb{Z}_N)$, notamment $L^2(\Omega)$, $\Omega \subseteq \mathbb{R}^n$ qu'on examinera plus tard, est à la base de la compréhension de ce qu'on appelle le *principe d'indétermination de Heisenberg*, un des noyaux conceptuels de la mécanique quantique.

Si on calcule les coefficients de Fourier du signal constant $z(n) = \frac{1}{N}$, $\forall n \in \mathbb{Z}_N$, on obtient :

$$\hat{z}(m) = \sum_{n=0}^{N-1} \frac{1}{N} e^{-2\pi i \frac{mn}{N}} = \frac{1}{N} \sum_{n=0}^{N-1} e^{-2\pi i \frac{mn}{N}} = \delta_0(m) = \begin{cases} 1 & m = 0 \\ 0 & m = 1, \ldots, N-1, \end{cases}$$

donc, la DFT d'un signal constant (et donc complètement délocalisé par rapport à son paramètre) est une impulsion unitaire dans le domaine de Fourier, et donc, complètement localisé dans les fréquences.

On peut donner une interprétation physique de la formule du lemme (2.1) : $e^{2\pi i n \frac{m}{N}} = \cos\left(2\pi n \frac{m}{N}\right) + i \sin\left(2\pi n \frac{m}{N}\right) = 0$ si et seulement si $\cos\left(2\pi n \frac{m}{N}\right) = 0$ et $\sin\left(2\pi n \frac{m}{N}\right) = 0$ quand $m \in \{1, 2, \ldots, N-1\}$. Donc, le résultat qu'on vient de voir affirme que la superposition de $N-1$ fonctions harmoniques avec des fréquences multiples entières entre elles est nulle partout, i.e. elles sont sujettes à un phénomène d'interférence destructive, sauf en une valeur, où les harmoniques s'accumulent.

La formule de synthèse dit qu'il faut donner des poids différents aux harmoniques pour pouvoir reconstruire un signal différent d'une impulsion.

2.6.4 Hautes et basses fréquences m dans la formule de synthèse

On veut comprendre plus en profondeur la signification des coefficients fréquentiels m dans l'ensemble :

$$\left\{e^{2\pi i\frac{mn}{N}} = \cos\left(2\pi\frac{mn}{N}\right) + i\sin\left(2\pi\frac{mn}{N}\right),\ n = 0, 1, \ldots, N-1\right\},$$

qui représente les valeurs des harmoniques dans chacun des N paramétrés n.

Pour simplifier l'analyse, on considère seulement la partie réelle de l'ensemble ci-dessus, i.e.

$$H_m = \left\{\cos\left(2\pi\frac{mn}{N}\right),\ n = 0, 1, \ldots, N-1\right\},$$

les considérations qu'on fera sur le cosinus peuvent être répétées pour le sinus.

On discute du comportement de $\cos\left(2\pi\frac{mn}{N}\right)$ quand m prend les valeurs de 0 jusqu'à $N-1$, N *paire* (on discutera plus tard du cas N impaire) :

— $m = 0$: comme on l'a déjà vu, dans ce cas *on a pas d'oscillations*, mais une suite de valeurs constantes, $\cos\left(2\pi\frac{0n}{N}\right) = 1$, donc

$$H_0 = \{1, 1, \ldots, 1\};$$

— $m = 1$:

$$H_1 = \left\{1, \cos\left(2\pi\frac{1}{N}\right), \cos\left(2\pi\frac{2}{N}\right), \ldots, \cos\left(2\pi\frac{N-1}{N}\right)\right\},$$

les valeurs de H_1 représentent N échantillons d'*une oscillation du cosinus*. Le cycle n'est pas terminé car on ne considère pas la valeur $n = N$, qui permettrait d'obtenir $\cos\left(2\pi\frac{N}{N}\right) = \cos(2\pi) = 1$. Dans la figure 2.1 on montre le graphe de H_m pour $m = 1, N = 16$.

— $m = 2$:

$$H_2 = \left\{1, \cos\left(2\pi\frac{2}{N}\right), \cos\left(2\pi\frac{4}{N}\right), \ldots, \cos\left(2\pi\frac{2\frac{N}{2}}{N}\right) = 1, \ldots, \cos\left(2\pi\frac{2(N-1)}{N}\right)\right\},$$

les valeurs de H_2 représentent N échantillons de *deux oscillations du cosinus*. En fait, en correspondance de $n = N/2$, un cycle du cosinus est terminé. Dans la figure 2.2 on montre le graphe de H_m pour $m = 2, N = 16$. On voit que pour $n = 8 = 16/2$ la valeur du cosinus est 1.

— Si on augmente m jusqu'à $N/2$, on augmente la fréquence des oscillations du cosinus. La fréquence maximale est obtenue quand $m = N/2$, en fait, dans ce cas, $\cos\left(2\pi\frac{\frac{N}{2}n}{N}\right) = \cos(\pi n)$, donc

$$H_{\frac{N}{2}} = \{(-1)^n,\ n = 0, 1, \ldots, N-1\}.$$

— On pourrait penser que la fréquence des oscillations du cosinus augmente jusqu'à $N-1$, mais ceci n'est pas vrai. En fait, *à partir de $m = N/2 + 1$, la fréquence des oscillations du cosinus décroît.*

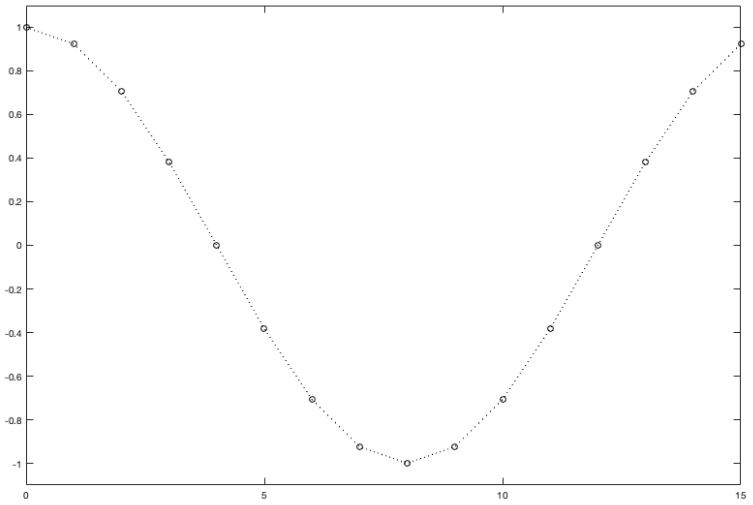


FIGURE 2.1 – H_m pour $m = 1, N = 16$.

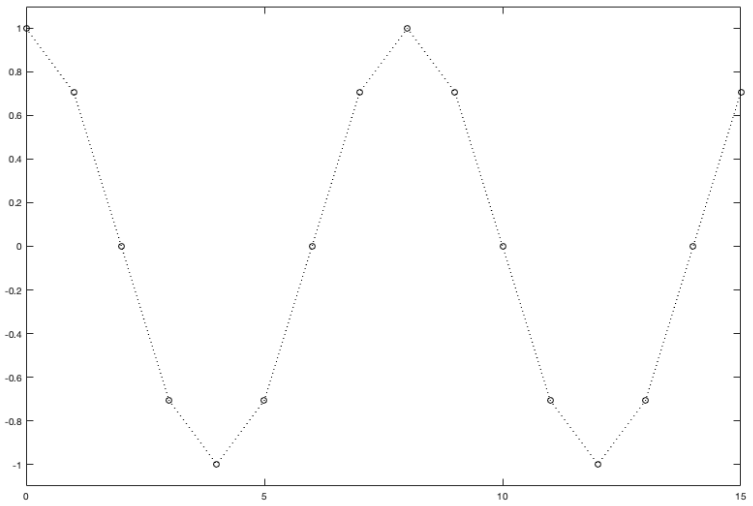


FIGURE 2.2 – H_m pour $m = 2, N = 16$.

Pour comprendre ce comportement, on considère $\cos\left(2\pi\frac{nm}{N}\right)$ quand m appartient à l'ensemble $\left\{\frac{N}{2}+1, \frac{N}{2}+2, \ldots, N-1\right\}$ et on fait un changement de variable :

$$k = N-m \;\Leftrightarrow\; m = N-k, \quad m \in \left\{\frac{N}{2}+1, \frac{N}{2}+2, \ldots, N-1\right\} \;\Leftrightarrow\; k \in \left\{\frac{N}{2}-1, \ldots, 2, 1\right\},$$

donc quand m croît de $\frac{N}{2}+1$ jusqu'à $N-1$, k décroît de $\frac{N}{2}-1$ jusqu'à 1. Si on introduit

le changement de variable dans le cosinus on obtient :

$$\cos\left(2\pi\frac{n(N-k)}{N}\right) = \cos\left(2\pi n - 2\pi\frac{nk}{N}\right) = \cos\left(-2\pi\frac{nk}{N}\right) = \cos\left(2\pi\frac{nk}{N}\right),$$

on a utilisé la périodicité et la parité du cosinus.
Par conséquent,

$$\boxed{\cos\left(2\pi\frac{nm}{N}\right),\ m \in \left\{\frac{N}{2}+1, \frac{N}{2}+2, \ldots, N-1\right\} \Leftrightarrow \cos\left(2\pi\frac{nk}{N}\right),\ k \in \left\{\frac{N}{2}-1, \ldots, 1\right\}},$$

Donc, **le nombre d'oscillations harmoniques est maximal pour $m = N/2$ et symétrique par rapport à cette valeur**. Par exemple, le graphe de H_m pour $m = 9, N = 16$ est exactement égal au graphe de H_m avec $m = 7, N = 16$. De la même manière, le graphe de $m = 15, N = 16$ est exactement égal au graphe de la figure 2.1, qui représente H_m avec $m = 1, N = 16$.

— Bien sûr, si N est *impaire*, alors ce qu'on vient de dire est valide pour $\left[\frac{N}{2}\right]$, la partie entière de $\frac{N}{2}$, i.e. l'entier le plus proche et non supérieur à $\frac{N}{2}$.

Les considérations précédentes expliquent pourquoi on appelle :

— **Hautes fréquences** : les valeurs de m proches de $\frac{N}{2}$;

— **Basses fréquences** : les valeurs de m proches de 0 ou de $N-1$.

Donc, si dans la formule de synthèse d'un signal discret $z \in \ell^2(\mathbb{Z}_N)$ on a des coefficients de Fourier $\hat{z}(m)$ avec un module élevé pour des valeurs de m proches de $N/2$, le signal sera caractérisé par des variations plutôt violentes (par exemple un son aigu, comme pour les *cymbales*). Par contre, si les coefficients de Fourier avec le module plus élevé sont en correspondance des valeurs de m proches de 0 et à $N-1$, alors le signal sera caractérisé par des variations plus douces (par exemple un bruit grave, comme pour les *tambours*).

La fréquence $m = N/2$ est dite **fréquence de Nyquist** [7]. Elle est la fréquence harmonique la plus élevée qui peut apparaitre dans la formule de synthèse de N échantillons d'un signal.

2.6.5 Filtrage de signaux dans la représentation fréquentielle

Grâce à la DFT on peut modifier très facilement le contenu fréquentiel d'un signal, par exemple pour augmenter l'importance des basses ou des hautes fréquences.

Le schéma standard consiste à passer à l'espace de Fourier avec la DFT et après à manipuler les coefficients de Fourier, selon nos besoins avec un *filtre* $f : \ell^2(\mathbb{Z}_N) \to \ell^2(\mathbb{Z}_N)$, qui peut être une transformation linéaire ou non-linéaire. On termine par l'application de la IDFT à la suite des coefficients de Fourier modifiés pour reconstruire le signal original avec les modifications fréquentielles désirées.

Le schéma du traitement des signaux dans la représentation fréquentielle est montré dans la figure 2.3.

On souligne que dans la composition de transformations IDFT $\circ$ f $\circ$ DFT seulement f peut changer l'énergie du signal, car la composition entre la transformée de Fourier et la transformation inverse donne l'identité et donc elle préserve l'énergie du signal original.

7. Harry Nyquist, ingénieur suédois (1889-1976).

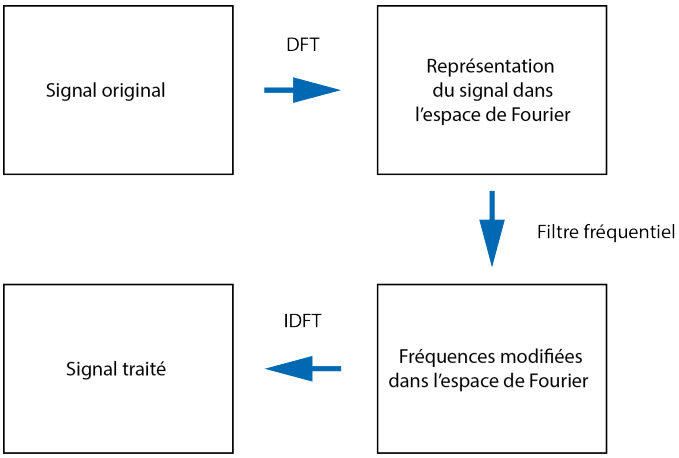


FIGURE 2.3 – Schéma d'un filtrage dans le domaine de Fourier.

Parmi toutes les filtres f, il existe un filtre très importante, qu'on va définir dans la sous-section 2.6.6 et qui nous permettra de définir le concept de multiplicateur de Fourier dans la sous-section 2.6.7.

2.6.6 L'opérateur de multiplication et sa représentation matricielle diagonale

Soit $w : \mathbb{Z}_N \to \mathbb{C}$ une suite fixé de $\ell^2(\mathbb{Z}_N)$.

Déf. 2.6.3 *On appelle* ***opérateur de multiplication*** *par la suite w, l'application linéaire ci-dessous :*

$$\begin{array}{rccc} M_w : & \ell^2(\mathbb{Z}_N) & \longrightarrow & \ell^2(\mathbb{Z}_N) \\ & z & \longmapsto & M_w(z) = w \cdot z, \end{array}$$

où $M_w(z) = w \cdot z : \mathbb{Z}_N \to \mathbb{C}$ est la suite définie par produit ponctuel (ou d'Hadamard) entre w et z :

$$\boxed{M_w z(n) = (w \cdot z)(n) = w(n) \cdot z(n)} \qquad \forall n \in \mathbb{Z}_N.$$

On observe que, si on représente z comme un vecteur colonne dans la base canonique de $\ell^2(\mathbb{Z}_N)$, alors **la matrice associée à l'opérateur M_w par rapport à la base canonique de $\ell^2(\mathbb{Z}_N)$ est une matrice diagonale** D_w qui a comme éléments diagonaux les composantes de la suite w, en fait :

$$D_w z = \begin{pmatrix} w(0) & & 0 \\ & \ddots & \\ 0 & & w(N-1) \end{pmatrix} \begin{pmatrix} z(0) \\ \vdots \\ z(N-1) \end{pmatrix} = \begin{pmatrix} w(0)z(0) \\ \vdots \\ w(N-1)z(N-1) \end{pmatrix}.$$

Exemple d'opérateur de multiplication : on considère la suite de $\ell^2(\mathbb{Z}_6)$ donnée par $z = (2, 3-i, 2i, 4+i, 0, 1)$ et la suite $w(n) = i^n$, $n \in \mathbb{Z}_6$, alors :

$$(w(0) = 1, w(1) = i, w(2) = -1, w(3) = -i, w(4) = 1, w(5) = i),$$

et donc

$$(M_w z)(n) = (1 \cdot 2, i \cdot (3-i), -1 \cdot 2i, -i \cdot (4+i), 1 \cdot 0, i \cdot 1) = (2, 3i+1, -2i, -4i+1, 0, i).$$

On peut maintenant introduire le multiplicateur de Fourier.

2.6.7 Le multiplicateur de Fourier

Ici on va donner un exemple remarquable de filtre fréquentiel, dit multiplicateur de Fourier.

Déf. 2.6.4 *Étant donnée une suite $w : \mathbb{Z}_N \to \mathbb{C}$, on appelle **multiplicateur de Fourier** par la suite w, l'opérateur ci-dessous :*

$$\begin{array}{rccc} T_{(w)} : & \ell^2(\mathbb{Z}_N) & \longrightarrow & \ell^2(\mathbb{Z}_N) \\ & z & \longmapsto & T_{(w)}(z) = \widetilde{w \cdot \hat{z}}, \end{array}$$

i.e. $T_{(w)}$ est l'opérateur donné par la composition suivante :

$$\boxed{T_{(w)} = \mathrm{IDFT} \circ M_w \circ \mathrm{DFT}},$$

i.e.

$$\begin{array}{ccccccccc} \ell^2(\mathbb{Z}_N) & \xrightarrow[\mathrm{DFT}]{} & \ell^2(\mathbb{Z}_N) & \xrightarrow[M_w]{} & \ell^2(\mathbb{Z}_N) & \xrightarrow[\mathrm{IDFT}]{} & \ell^2(\mathbb{Z}_N) \\ z & \mapsto & \mathrm{DFT}(z) = \hat{z} & \mapsto & M_w(\mathrm{DFT}(z)) = w \cdot \hat{z} & \mapsto & \mathrm{IDFT}(M_w(\mathrm{DFT}(z))) = \widetilde{w \cdot \hat{z}}. \end{array}$$

Si on applique la DFT aux deux côtés de la définition de $T_{(w)}$, on voit que l'action du multiplicateur de Fourier est diagonale dans la base de Fourier F :

$$\mathrm{DFT}\, T_{(w)} z = [T_{(w)} z]_F = M_w \circ \mathrm{DFT}\, z = M_w \hat{z}, \qquad \forall z \in \ell^2(\mathbb{Z}_N). \tag{2.26}$$

Alors, $T_{(w)}$ multiplie les coefficients de Fourier de z par les composantes de la suite w (ce qui explique le nom de l'opérateur). Donc, on peut :

- **Atténuer les basses fréquences** du signal z en choisissant une suite $w(m)$ avec $|w(m)|$ petit quand $m \simeq 0$ et $m \simeq N-1$;
- **Atténuer les hautes fréquences** du signal z en choisissant une suite $w(m)$ avec $|w(m)|$ petit quand $m \simeq N/2$;
- **Booster (amplifier) les basses fréquences** du signal z en choisissant une suite $w(m)$ avec $|w(m)|$ grand quand $m \simeq 0$ et $m \simeq N-1$;
- **Booster les hautes fréquences** du signal z en choisissant une suite $w(m)$ avec $|w(m)|$ grand quand $m \simeq N/2$.

Ceci est utilisé dans les égalisateurs graphiques utilisés par les musiciens pour régler le niveau des aigües et des basses dans un signal audio.

2.7 Propriétés de la DFT

Dans cette section on démontrera les plus importantes propriétés de la DFT. On commence par rappeler la propriété de translation de l'indice d'une sommation :

$$\sum_{i=n_0}^{n} a_i = \sum_{i=n_0-k}^{n-k} a_{i+k} = \sum_{i=n_0+k}^{n+k} a_{i-k}, \tag{2.27}$$

qu'on utilisera souvent, et on anticipe un lemme qui sera très utile dans le développement de plusieurs démonstrations.

Lemme 2.7.1 *Soit $f : \mathbb{Z} \to \mathbb{C}$ une fonction N-périodique, $N \in \mathbb{N}$:*

$$f(n + aN) = f(n) \qquad \forall a, n \in \mathbb{Z}.$$

Alors, pour tout $m \in \mathbb{Z}$:

$$\boxed{\sum_{n=m}^{m+N-1} f(n) = \sum_{n=0}^{N-1} f(n)},$$

i.e. ***la somme de f (N-périodique) sur tout intervalle de taille N est constante.***

Preuve. Si $m = 0$ il n'y a rien à prouver, soit alors $m \in \mathbb{Z}$, $m \neq 0$. On commence à considérer $m > 0$, alors :

$$\sum_{n=m}^{m+N-1} f(n) = \sum_{n=0}^{m+N-1} f(n) - \sum_{n=0}^{m-1} f(n) = \sum_{n=0}^{N-1} f(n) + \sum_{n=N}^{m+N-1} f(n) - \sum_{n=0}^{m-1} f(n),$$

mais, en utilisant (2.27),

$$\sum_{n=N}^{m+N-1} f(n) = \sum_{n=0}^{m-1} f(n+N) = \sum_{n=0}^{m-1} f(n)$$

grâce à la N-périodicité de f, donc :

$$\sum_{n=m}^{m+N-1} f(n) = \sum_{n=0}^{N-1} f(n) + \sum_{n=0}^{m-1} f(n) - \sum_{n=0}^{m-1} f(n) = \sum_{n=0}^{N-1} f(n).$$

Si $m < 0$, la démonstration est analogue. □

2.7.1 La périodicité de $\hat{z}$ et $\check{z}$

Par la suite, on examinera les plus importantes propriétés des objets de la théorie de Fourier discrète et on les appliquera à des problèmes pratiques. Néanmoins, il y a une propriété qu'on doit démontrer tout de suite pour pouvoir interpréter correctement l'analyse de Fourier discrète : la périodicité de $\hat{z}$ et $\check{z}$.

Par calcul direct, si $a \in \mathbb{Z}$, on obtient que :

$$\hat{z}(m + aN) = \sum_{n=0}^{N-1} z(n) e^{-2\pi i \frac{(m+aN)n}{N}} = \sum_{n=0}^{N-1} z(n) e^{-2\pi i \frac{mn}{N}} e^{-2\pi i \frac{aNn}{N}} = \sum_{n=0}^{N-1} z(n) e^{-2\pi i \frac{mn}{N}} e^{-2\pi a n i} = \hat{z}(m),$$

car $e^{-2\pi ani} = \cos(2\pi an) - i\sin(2\pi an) = 1$. Avec le même calcul on arrive à démontrer que $\check{z}(n + aN) = \check{z}(n)\ \forall a \in \mathbb{Z}$.

Donc, exactement comme pour les suites $z \in \ell^2(\mathbb{Z}_N)$, on peut étendre la définitions de $\hat{z}$ et $\check{z}$ à $\mathbb{Z}$ en définissant les deux *suites N-périodiques* :

$$\begin{array}{rccc} \hat{z}: & \mathbb{Z} & \longrightarrow & \mathbb{C} \\ & m & \longmapsto & \boxed{\hat{z}(m) = \hat{z}(m + aN)}, \end{array}$$

et

$$\begin{array}{rccc} \check{z}: & \mathbb{Z} & \longrightarrow & \mathbb{C} \\ & n & \longmapsto & \boxed{\check{z}(n) = \check{z}(n + aN)}, \end{array}$$

où $a \in \mathbb{Z}$ est tel que $m + aN \in \mathbb{Z}_N$, ou $n + aN \in \mathbb{Z}_N$, respectivement.

2.7.2 DFT et translation (invariance du spectre par translations)

On veut examiner maintenant comment varie la DFT d'un signal $z \in \ell^2(\mathbb{Z}_N)$ si on fait une translation de z avec une quantité différente de N. Pour formaliser ceci, on va introduire un autre opérateur de $\ell^2(\mathbb{Z}_N)$.

Déf. 2.7.1 *Soit $z \in \ell^2(\mathbb{Z}_N)$. On appelle **opérateur de translation à droite** de la quantité k, l'application linéaire ci-dessous :*

$$\begin{array}{rccc} R_k: & \ell^2(\mathbb{Z}_N) & \longrightarrow & \ell^2(\mathbb{Z}_N) \\ & z & \longmapsto & R_k(z), \end{array}$$

où $R_k(z) : \mathbb{Z}_N \to \mathbb{C}$ est la suite définie par la formule :

$$\boxed{R_k z(n) = z(n-k)} \qquad \forall n \in \mathbb{Z}_N.$$

Exemple d'opérateur de translation : $N = 6$, $k = 2$, $z = (2, 3-i, 2i, 4+i, 0, 1)$. Alors :

$$\begin{cases} R_2 z(0) = z(0-2) = z(-2) = z(-2+6) = z(4) = 0 \\ R_2 z(1) = z(1-2) = z(-1) = z(-1+6) = z(5) = 1 \\ \quad \vdots \end{cases}$$

on obtient : $R_2 z = (0, 1, 2, 3-i, 2i, 4+i)$, on voit que l'effet de R_2 sur z est juste de déplacer à droite de 2 positions les éléments de la suite. Droite = Right en anglais, c'est pour cela qu'on trouve souvent l'opérateur de translation à droite écrit avec la lettre R.

On observe que les 2 derniers éléments sont tournés dans les deux premières positions comme dans un cercle. Pour cette raison R_k est aussi connu comme **opérateur de translation circulaire** ou *opérateur de rotation.*

On veut comprendre comment caractériser la composition de l'opérateur de translation avec la DFT et, inversement, de la DFT avec l'opérateur de translation. On commence avec la dernière composition : $\mathrm{DFT}(z(n-k))$, i.e.

$$\boxed{\begin{array}{ccccc} \ell^2(\mathbb{Z}_N) & \xrightarrow[R_k]{} & \ell^2(\mathbb{Z}_N) & \xrightarrow[\text{DFT}]{} & \ell^2(\mathbb{Z}_N) \\ z & \longmapsto & R_k z & \longmapsto & (\text{DFT} \circ R_k)z = \text{DFT}(R_k z) = \widehat{R_k z}. \end{array}}$$

Le théorème suivant montre que l'action de l'opérateur R_k est transformée par la DFT en la multiplication par un exponentiel complexe.

Théorème 2.7.1 *Soit $z \in \ell^2(\mathbb{Z}_N)$ et soit $k \in \mathbb{Z}$. Alors :*

$$\boxed{\widehat{R_k z}(m) = e^{-2\pi i \frac{mk}{N}} \hat{z}(m)} \qquad \forall m \in \mathbb{Z}, \tag{2.28}$$

i.e. si on définit la suite $\omega_N^k \in \ell^2(\mathbb{Z}_N)$, $\omega_N^k(m) = \omega_N^{mk} = e^{-2\pi i \frac{mk}{N}}$ $\forall m \in \mathbb{Z}$, alors :

$$\boxed{\text{DFT} \circ R_k = M_{\omega_N^k} \circ \text{DFT}}. \tag{2.29}$$

Preuve.

$$\begin{aligned} \widehat{R_k z}(m) &= \sum_{n=0}^{N-1} (R_k z)(n) e^{-2\pi i \frac{mn}{N}} \\ &= \sum_{n=0}^{N-1} z(n-k) e^{-2\pi i \frac{mn}{N}} \\ &= \sum_{n=-k}^{N-k-1} z(n-k+k) e^{-2\pi i \frac{m(n+k)}{N}} \\ &= \sum_{n=-k}^{N-k-1} z(n) e^{-2\pi i \frac{mn}{N}} e^{-2\pi i \frac{mk}{N}}. \end{aligned}$$

Le facteur $e^{-2\pi i \frac{mk}{N}}$ ne dépend pas de l'indice n, donc on peut le sortir de la sommation :

$$\begin{aligned} \widehat{R_k z}(m) &= e^{-2\pi i \frac{mk}{N}} \sum_{n=-k}^{N-k-1} z(n) e^{-2\pi i \frac{mn}{N}} \\ &\underset{\text{(Lemme 2.7.1)}}{=} e^{-2\pi i \frac{mk}{N}} \sum_{n=0}^{N-1} z(n) e^{-2\pi i \frac{mn}{N}} \\ &= e^{-2\pi i \frac{mk}{N}} \hat{z}(m). \end{aligned}$$

On a pu appliquer le lemme 2.7.1 car, par hypothèse, z est N-périodique et l'exponentielle $e^{-2\pi i \frac{mn}{N}}$ est elle-même une fonction N-périodique. □

On observe que, si on écrit $\hat{z}(m) = |\hat{z}(m)| e^{i\text{Arg}(\hat{z}(m))}$, alors, comme $\left| e^{-2\pi i \frac{mk}{N}} \right| = 1$, le produit $e^{-2\pi i \frac{mk}{N}} \hat{z}(m)$ change seulement la phase de $\hat{z}(m)$, c'est pour cette raison qu'on dit que **la DFT transforme la translation dans un changement de phase**. Le fait que la phase des coefficients de Fourier soit modifiée par les translations implique que **le spectre de phase contient des informations sur la géométrie du signal original**.

Invariance du spectre par translation. Le théorème ci-dessus montre une limitation importante de la transformée de Fourier. En fait :

$$\left|e^{-2\pi i\frac{mk}{N}}\right| = 1 \quad \Longrightarrow \quad |\widehat{R_k z}(m)| = |\hat{z}(m)| \qquad \forall m, k \in \mathbb{Z},$$

donc, les modules des coefficients de Fourier de z et de toutes ses translations sont égaux. Par conséquent, **dans le module des coefficients de Fourier $|\hat{z}(m)|$, on a l'information sur l'importance de la fréquence m dans le signal z, mais pas de sa position dans le signal même**.

Pour souligner encore plus clairement ce comportement, reconsidérons l'impulsion unitaire et faisons une translation quelconque : $R_k\delta$, le spectre de ce signal est $|\widehat{R_k\delta_0}(m)| = |e^{-2\pi i\frac{mk}{N}}\hat{\delta}_0(m)|$, mais on a vue que $\hat{\delta}_0(m) = 1$ pour tout $m \in \mathbb{Z}_N$, donc $|\widehat{R_k\delta_0}(m)| = |e^{-2\pi i\frac{mk}{N}}| = 1$. La différence avec le cas de l'impulsion unitaire non-translatée c'est que, pour elle, le spectre est réel et donc $\left|\widehat{\delta_0}(m)\right| = \widehat{\delta_0}(m) = 1\ \forall m \in \mathbb{Z}_N$.

Ceci montre que le spectre de l'impulsion unitaire est le exactement le même que celui de toutes ses translations ! Donc, la seule information du spectre ne peut pas permettre de reconstruire la localisation spatiale du signal, pour le faire il faut utiliser aussi l'information donnée par la phase, qui, néanmoins, n'est pas une quantité simple à interpréter et manipuler.

Une alternative est donnée par deux transformations qui *localisent* la transformée de Fourier : la transformée de Gabor et la transformée en ondelettes. L'étude de ces opérateurs va au delà du but de ce cours.

Maintenant on analyse la composition entre l'opérateur de translation et la DFT : $\hat{z}(m-k)$, i.e.

$$\boxed{\begin{array}{ccccc} \ell^2(\mathbb{Z}_N) & \xrightarrow[\text{DFT}]{} & \ell^2(\mathbb{Z}_N) & \xrightarrow[R_k]{} & \ell^2(\mathbb{Z}_N) \\ z & \longmapsto & \text{DFT}(z) & \longmapsto & (R_k \circ \text{DFT})z = \hat{z}(m-k). \end{array}}$$

Théorème 2.7.2 *Avec les hypothèses du théorème 2.7.1 ça vaut la formule :*

$$\boxed{(R_k\hat{z})(m) = \hat{z}(m-k) = \left(\widehat{e^{2\pi i\frac{nk}{N}}z}\right)(m)}, \qquad \forall m \in \mathbb{Z}, \tag{2.30}$$

i.e.

$$\boxed{R_k \circ \text{DFT} = \text{DFT} \circ M_{\overline{\omega_N^k}}}. \tag{2.31}$$

Preuve.

$$(R_k\hat{z})(m) = \hat{z}(m-k) = \sum_{n=0}^{N-1} z(n)e^{-2\pi i\frac{(m-k)n}{N}} = \sum_{n=0}^{N-1}\left(e^{2\pi i\frac{kn}{N}}z(n)\right)e^{-2\pi i\frac{mn}{N}} = \left(\widehat{e^{2\pi i\frac{kn}{N}}z}\right)(m).$$

□

On peut résumer les propriétés qu'on vient d'analyser avec les couples de Fourier de la table suivante, qui montre que l'opération de translation dans la représentation originale de z devient un changement de phase dans l'espace de Fourier et que, vice-versa, l'opérations de

Représentation originale	Espace de Fourier
$z(n-k)$	$e^{-2\pi i \frac{km}{N}} \hat{z}(m)$
$e^{2\pi i \frac{kn}{N}} z(n)$	$\hat{z}(m-k)$

translation dans l'espace de Fourier correspond à un changement de phase (avec une phase conjuguée) dans la représentation originale de z.

Un **cas remarquable** est donné par la situation suivante : N pair et $k = N/2$, alors :

$$e^{-\frac{2\pi i m \frac{N}{2}}{N}} = e^{-\pi i m} = (e^{-\pi i})^m = (-1)^m$$

et :

$$e^{\frac{2\pi i n \frac{N}{2}}{N}} = e^{\pi i n} = (e^{\pi i})^n = (-1)^n.$$

Alors :

$$\mathrm{DFT}\left(z\left(n-\frac{N}{2}\right)\right) = (-1)^m \hat{z}(m), \quad \hat{z}\left(m-\frac{N}{2}\right) = (\widehat{(-1)^n z})(m), \tag{2.32}$$

on voit que multiplier par $(-1)^n$ la suite z correspond à faire une translation de $N/2$ de son spectre. Cette opération est souvent utilisée pour centrer le spectre en $m = 0$.

Pour terminer, une observation sur le rapport entre la formule (2.29) et la représentation diagonale de l'opérateur R_k. Si on compose les membres de gauche et de droite de la formule (2.29) avec la IDFT on obtient :

$$\mathrm{DFT} \circ R_k \circ \mathrm{IDFT} = M_{\omega_N^k}.$$

Si on écrit avec A_k et $D_{\omega_N^k}$ (diagonale, cfr. section 2.6.6) les matrice associées à l'opérateur R_k et $M_{\omega_N^k}$ par rapport à la base canonique, alors on peut réécrire l'équation antérieure comme ceci :

$$W_N \, A_k W_N^{-1} = D_{\omega_N^k},$$

qui montre que *la matrice A_k associée à l'opérateur de translation R_k est semblable à la matrice diagonale $D_{\omega_N^k}$.*

La matrice inversible qui réalise la conjugaison matricielle entre A_k et $D_{\omega_N^k}$ est la matrice de Sylvester W_N, donc on peut dire que l'action de *l'opérateur de translation R_k est diagonale dans l'espace de Fourier.*

Ce comportement est un cas particulier d'une classe plus ample, et très importante, d'opérateurs (dits stationnaires) qui sont diagonaux dans la base de Fourier.

2.7.3 DFT et conjugaison

Étant donnée une suite $z \in \ell^2(Z_N)$, on définit la suite complexe conjuguée $\bar{z}$ comme ceci : $\bar{z} = (\bar{z}(0), \bar{z}(1), \ldots, \bar{z}(N-1))$, i.e. $\bar{z}(n) = \overline{z(n)}$ $\forall n \in \mathbb{Z}_N$.

Le théorème suivant montre la relation entre DFT et conjugaison.

Théorème 2.7.3 *Pour toute $z \in \ell^2(Z_N)$ ça vaut que :*

$$\hat{\bar{z}}(m) = \overline{\hat{z}(-m)} = \overline{\hat{z}(N-m)} \qquad \forall m \in \mathbb{Z}_N.$$

Preuve.

$$\widehat{\widehat{z}}(m) = \sum_{n=0}^{N-1} \overline{z(n)} e^{-2\pi i \frac{mn}{N}} = \overline{\sum_{n=0}^{N-1} z(n) e^{2\pi i \frac{mn}{N}}} = \overline{\sum_{n=0}^{N-1} z(n) e^{-2\pi i \frac{(-m)n}{N}}} = \overline{\hat{z}(-m)},$$

$\overline{\hat{z}(N-m)} = \overline{\hat{z}(-m)}$, par périodicité. □

Corollaire 2.7.1 *$z \in \ell^2(\mathbb{Z}_N)$ est réelle, i.e. $z(n) \in \mathbb{R}$ $\forall n \in \mathbb{Z}_N$, si et seulement si*

$$\hat{z}(m) = \overline{\hat{z}(-m)} = \overline{\hat{z}(N-m)}, \qquad \forall m \in \mathbb{Z}_N.$$

$\hat{z} \in \ell^2(\mathbb{Z}_N)$ est réelle, i.e. $\hat{z}(m) \in \mathbb{R}$ $\forall m \in \mathbb{Z}_N$, si et seulement si

$$z(n) = \overline{z(-n)} = \overline{z(N-n)}, \qquad \forall n \in \mathbb{Z}_N.$$

Preuve. Comme la DFT est un isomorphisme de $\ell^2(\mathbb{Z}_N)$, z est réelle, i.e. $z = \bar{z}$, si et seulement si $\hat{z} = \widehat{\bar{z}}$, mais, grâce au théorème ci-dessus, ça vaut quand $\hat{z}(m) = \overline{\hat{z}(-m))} = \overline{\hat{z}(N-m)}$.

$\hat{z}$ est réelle si et seulement si $\hat{z} = \overline{\hat{z}}$, mais le théorème ci-dessus dit que $\overline{\hat{z}(-m)} = \widehat{\bar{z}}(m)$ $\forall m \in \mathbb{Z}_N$, ce qui implique $\overline{\hat{z}(m)} = \widehat{\bar{z}}(-m)$, par simple échange de variable $m \leftrightarrow -m$. Donc :

$$\mathrm{IDFT}(\overline{\hat{z}(m)}) = \mathrm{IDFT}(\widehat{\bar{z}}(-m)) = \mathrm{IDFT}(\mathrm{DFT}(\bar{z}(-m))) = \bar{z}(-n) = \overline{z(-n)}.$$

Par conséquent, $\hat{z}$ est réelle $\iff$ $\hat{z}(m) = \overline{\hat{z}(m)}$ $\iff$ $\mathrm{IDFT}(\hat{z}(m)) = \mathrm{IDFT}(\overline{\hat{z}(m)})$ $\iff$ $z(n) = \overline{z(-n)} = \overline{z(N-n)}$ $\forall n \in \mathbb{Z}_N$. □

Une conséquence immédiate du résultat ci-dessus est le corollaire suivant.

Corollaire 2.7.2 *Si $z, \hat{z} \in \ell^2(\mathbb{Z}_N)$ sont simultanément réelles si et seulement si elles sont symétriques par rapport à 0, i.e. $z(n) = z(-n)$ et $\hat{z}(m) = \hat{z}(-m)$, $\forall m, n \in \mathbb{Z}_N$.*

2.7.4 DFT et convolution

Une des plus importantes propriétés de la transformée de Fourier est liée à l'opération de convolution. Pour introduire cette opération on va rappeler la formule relative au produit de polynômes.

Si $P(x) = a_0 + a_1 x + \ldots + a_n x^n = \sum_{i=0}^{n} a_i x^i$ et $Q(x) = b_0 + b_1 x + \ldots + b_m x^m = \sum_{j=0}^{m} b_j x^j$, alors :

$$P(x)Q(x) = \sum_{\ell=0}^{n+m} c_\ell x^\ell, \quad \text{où} \quad c_\ell = \sum_{k=0}^{\ell} a_{\ell-k} b_k = \sum_{k=0}^{\ell} a_k b_{\ell-k}. \tag{2.33}$$

Exemple : $P(x) = a_0 + a_1 x + a_2 x^2$, $Q(x) = b_0 + b_1 x + b_2 x^2$, alors :

$$P(x)Q(x) = a_0 b_0 + (a_0 b_1 + a_1 b_0)x + (a_0 b_2 + a_1 b_1 + a_2 b_0)x^2 + (a_1 b_2 + a_2 b_1)x^3 + (a_2 b_2)x^4,$$

les coefficients des puissances de la variable x vérifient la formule (2.33). On observe que, dans les coefficients c_ℓ, on a une somme de produits des coefficients a_i et b_j, avec la particularité que la somme des indices $i + j$ est toujours égale à ℓ et que l'indice d'une variable croît et l'indice de l'autre décroît.

Cela c'est la caractéristique qui définit l'opération de convolution (discrète), qu'on va introduire dans l'espace $\ell^2(\mathbb{Z}_N)$.

Déf. 2.7.2 *Soient $z, w \in \ell^2(\mathbb{Z}_N)$. La convolution entre z et w, écrite $z * w$, est la suite de $\ell^2(\mathbb{Z}_N)$ avec composantes définies par :*

$$\boxed{(z * w)(n) = \sum_{k=0}^{N-1} z(n-k)w(k) = \sum_{k=0}^{N-1} w(n-k)z(k)}, \qquad \forall n \in \mathbb{Z}_N.$$

La convolution est *symétrique*, i.e. $z * w = w * z$, grâce à la commutativité du produit en $\mathbb{C}$.

Exemple : $z, w \in \ell^2(\mathbb{Z}_4)$, $z = (1, 1, 0, 2)$, $w = (i, 0, 1, 2)$, avec la périodicité canonique : $z(n + kN) = z(n)$ et $w(n + kN) = w(n)$ $\forall n \in \mathbb{Z}_N$ et $k \in \mathbb{Z}$. Alors :

$$\begin{aligned}
(z * w)(0) &= \sum_{k=0}^{4-1} z(0-k)w(k) = \sum_{k=0}^{3} z(-k)w(k) \\
&= z(0)w(0) + z(-1)w(1) + z(-2)w(2) + z(-3)w(3) \\
&= z(0)w(0) + z(4-1)w(1) + z(4-2)w(2) + z(4-3)w(3) \\
&= z(0)w(0) + z(3)w(1) + z(2)w(2) + z(1)w(3) \\
&= 1 \cdot i + 2 \cdot 0 + 0 \cdot 1 + 1 \cdot 2 \\
&= i + 2
\end{aligned}$$

On obtient également : $(z * w)(1) = 2 + i$, $(z * w)(2) = 1 + 2i$, $(z * w)(3) = 1 + 3i$, et donc : $(z * w) = (i + 2, 2 + i, 1 + 2i, 1 + 3i)$.

L'interaction entre DFT et convolution a la propriété magnifique décrite dans le théorème suivant.

Théorème 2.7.4 *Soient $z, w \in \ell^2(\mathbb{Z}_N)$. Alors :*

$$\boxed{\mathrm{DFT}\,(z * w)(m) = \hat{z}(m) \cdot \hat{w}(m)} \Longleftrightarrow \boxed{(z * w)(n) = \mathrm{IDFT}\,(\hat{z} \cdot \hat{w})(n)} \qquad \forall n, m \in \mathbb{Z} \quad (2.34)$$

$$\boxed{\mathrm{IDFT}\,(\hat{z} * \hat{w})(n) = Nz(n) \cdot w(n)} \Longleftrightarrow \boxed{(\hat{z} * \hat{w})(m) = N\,\mathrm{DFT}(z \cdot w)(m)} \qquad \forall n, m \in \mathbb{Z} \ (2.35)$$

i.e. la transformée de Fourier de la convolution entre z et w est le produit ponctuel des transformées de Fourier et, vice-versa, la transformée de Fourier inverse de la convolution entre $\hat{z}$ et $\hat{w}$ est N fois le produit ponctuel de z et w. Autrement dit, on a les couples de Fourier suivantes :

Représentation originale	Espace de Fourier
$z * w$	$\hat{z} \cdot \hat{w}$
$Nz \cdot w$	$\hat{z} * \hat{w}$

Preuve. Par définition :

$$\widehat{(z * w)}(m) = \sum_{n=0}^{N-1} (z * w)(n) e^{-2\pi i \frac{mn}{N}} = \sum_{n=0}^{N-1} \left(\sum_{k=0}^{N-1} z(n-k) w(k) \right) e^{-2\pi i \frac{mn}{N}}.$$

On réécrit l'exponentiel comme ceci :

$$e^{-2\pi i \frac{mn}{N}} = e^{-2\pi i \frac{m(n-k+k)}{N}} = e^{-2\pi i \frac{m(n-k)+mk}{N}} = e^{-2\pi i \frac{m(n-k)}{N}} e^{-2\pi i \frac{mk}{N}}.$$

Alors :

$$\begin{aligned}
\widehat{(z * w)}(m) &= \sum_{n=0}^{N-1} \sum_{k=0}^{N-1} z(n-k) w(k) e^{-2\pi i \frac{m(n-k)}{N}} e^{-2\pi i \frac{mk}{N}} \\
&= \sum_{k=0}^{N-1} w(k) e^{-2\pi i \frac{mk}{N}} \sum_{n=0}^{N-1} z(n-k) e^{-2\pi i \frac{m(n-k)}{N}} \\
&= \sum_{k=0}^{N-1} w(k) e^{-2\pi i \frac{mk}{N}} \sum_{n=-k}^{N-k-1} z(n-k+k) e^{-2\pi i \frac{m(n-k+k)}{N}} \\
&= \sum_{k=0}^{N-1} w(k) e^{-2\pi i \frac{mk}{N}} \sum_{n=-k}^{N-k-1} z(n) e^{-2\pi i \frac{mn}{N}} \\
&\underset{\text{(Lemme 2.7.1)}}{=} \sum_{k=0}^{N-1} w(k) e^{-2\pi i \frac{mk}{N}} \sum_{n=0}^{N-1} z(n) e^{-2\pi i \frac{mn}{N}} \\
&= \hat{w}(m)\hat{z}(m) = \hat{z}(m)\hat{w}(m),
\end{aligned}$$

on observe qu'on a pu appliquer le Lemme 2.7.1 car il est valide pour n'importe quel $k \in \mathbb{Z}$. Alors :

$$\widehat{(z * w)}(m) = \hat{z}(m)\hat{w}(m), \qquad \forall m \in \mathbb{Z}.$$

La preuve du fait que $\mathrm{IDFT}(\hat{z} * \hat{w})(n) = z(n) \cdot w(n)$ est très similaire. Par définition :

$$\mathrm{IDFT}(\hat{z} * \hat{w})(n) = \frac{1}{N} \sum_{m=0}^{N-1} (\hat{z} * \hat{w})(m) e^{2\pi i \frac{mn}{N}} = \frac{1}{N} \sum_{m=0}^{N-1} \left(\sum_{k=0}^{N-1} \hat{z}(m-k)\hat{w}(k) \right) e^{2\pi i \frac{mn}{N}}.$$

On réécrit l'exponentiel comme ceci :

$$e^{2\pi i \frac{mn}{N}} = e^{2\pi i \frac{n(m-k+k)}{N}} = e^{2\pi i \frac{n(m-k)+nk}{N}} = e^{2\pi i \frac{n(m-k)}{N}} e^{2\pi i \frac{nk}{N}}.$$

Alors :

$$\begin{aligned}
\mathrm{IDFT}(\hat{z} * \hat{w})(n) &= \frac{1}{N} \sum_{m=0}^{N-1} \sum_{k=0}^{N-1} \hat{z}(m-k)\hat{w}(k) e^{2\pi i \frac{n(m-k)}{N}} e^{2\pi i \frac{nk}{N}} \\
&= \frac{1}{N} \sum_{k=0}^{N-1} \hat{w}(k) e^{2\pi i \frac{nk}{N}} \sum_{m=0}^{N-1} \hat{z}(m-k) e^{2\pi i \frac{n(m-k)}{N}} \\
&= \frac{1}{N} \sum_{k=0}^{N-1} \hat{w}(k) e^{2\pi i \frac{nk}{N}} \sum_{m=-k}^{N-k-1} \hat{z}(m-k+k) e^{2\pi i \frac{n(m-k+k)}{N}} \\
&= \frac{1}{N} \sum_{k=0}^{N-1} \hat{w}(k) e^{2\pi i \frac{nk}{N}} \sum_{m=-k}^{N-k-1} \hat{z}(m) e^{2\pi i \frac{mn}{N}} \\
&\underset{\text{(Lemme 2.7.1)}}{=} \frac{1}{N} \sum_{k=0}^{N-1} \hat{w}(k) e^{2\pi i \frac{nk}{N}} \sum_{m=0}^{N-1} \hat{z}(m) e^{2\pi i \frac{mn}{N}} \\
&= N \left(\frac{1}{N} \sum_{k=0}^{N-1} \hat{w}(k) e^{2\pi i \frac{nk}{N}} \right) \cdot \left(\frac{1}{N} \sum_{m=0}^{N-1} \hat{z}(m) e^{2\pi i \frac{mn}{N}} \right) \\
&= N \, \mathrm{IDFT}\, \hat{w}(n) \cdot \mathrm{IDFT}\, \hat{z}(n) = N \, w(n) \cdot z(n) = N \, z(n) \cdot w(n).
\end{aligned}$$

□

Observations :

— Pendant les étapes de la preuve, on ne peut pas remplacer $\sum_{k=0}^{N-1} w(k) e^{-2\pi i \frac{mk}{N}}$ avec $\hat{w}(m)$ avant de la dernière étape, car dans la deuxième sommation il y a encore l'indice k. Seulement quand on élimine la présence de k on peut remplacer avec $\hat{w}(m)$.

— Les formules (2.34) montrent une espèce de propriété distributive par rapport à la convolution et au produit ponctuel : **quand la DFT est appliquée à un produit de convolution, elle est répartie sur les facteurs et la convolution devient un produit ponctuel**, vice-versa, **quand la IDFT est appliquée à un produit produit ponctuel, elle est repartie sur les facteur et le produit ponctuel devient une convolution**. Donc :

$$\mathrm{DFT}(\check{z} * \check{w}) = z \cdot w, \;\; \mathrm{IDFT}(z \cdot w) = \check{z} * \check{w} \qquad \forall z, w \in \ell^2(\mathbb{Z}_N). \tag{2.36}$$

— Grâce à la transformée de Fourier, **on peut transformer une opération complexe comme la convolution dans le simple produit de transformées de Fourier** (qu'on peut calculer rapidement avec la FFT).

Ce résultat est très utilisé dans les applications de traitement des signaux. Si on utilise la normalisation relative à la base orthonormale pour définir la DFT, alors on voit apparaitre des coefficients dans la formule de la DFT de la convolution. Ces coefficients peuvent être très grands, surtout quand on considère la DFT en dimensions supérieures à 1 et/ou signaux massifs, et cela peut créer des erreurs numériques pendant les calculs. C'est surtout à cause de la simplicité de cette formule que beaucoup d'auteurs et de logiciels privilégient la définition de coefficients de Fourier qu'on a donné dans ce cours plutôt qu'autres définitions.

— Souvent, la convolution est faite entre un signal z et un autre signal w qui est différent de 0 seulement dans un support de taille T. La valeur de T est importante pour choisir d'effectuer la convolution directement ou avec la FFT. La complexité de l'opération de convolution directe est $\mathcal{O}(NT)$, par contre, si on utilise la FFT, la complexité est $\mathcal{O}(N \log N)$. Par conséquent, il est avantageux de transformer la convolution en un produit ponctuel avec la FFT dès que $T > \log(N)$. Pour donner un exemple concret, si $z \in \ell^2(\mathbb{Z}_N)$, avec $N = 1000$, alors $\log(N) \simeq 7$ et donc il est conseillé de faire la convolution $z * w$ dans le domaine de Fourier dès que le support de w est supérieur à 7.

Si on fixe un vecteur dans la convolution, on peut définir l'endomorphisme de $\ell^2(\mathbb{Z}_N)$ suivant :

Déf. 2.7.3 *Fixée une suite $w \in \ell^2(\mathbb{Z}_N)$, on appelle **opérateur de convolution** avec w la transformation linéaire suivante :*

$$\begin{array}{rccc} T_w : & \ell^2(\mathbb{Z}_N) & \longrightarrow & \ell^2(\mathbb{Z}_N) \\ & z & \longmapsto & T_w(z) = z * w. \end{array}$$

Comme pour le cas de l'opérateur de translation, on peut donner une représentation diagonale de l'opérateur de convolution. Pour cela, on réécrit la formule (2.34) sans spécifier l'indice m (car elle vaut pour n'importe quel indice), i.e. $\mathrm{DFT}(z * w) = \hat{z} \cdot \hat{w}$, mais $\mathrm{DFT}(z * w) = (\mathrm{DFT} \circ T_w)z$ et $\hat{z} \cdot \hat{w} = \hat{w} \cdot \hat{z} = M_{\hat{w}}\, \hat{z} = (M_{\hat{w}} \circ \mathrm{DFT})z$, i.e. $(\mathrm{DFT} \circ T_w)z = (M_{\hat{w}} \circ \mathrm{DFT})z \ \forall z \in \ell^2(\mathbb{Z}_N)$, donc on peut écrire la relation opératorielle $\mathrm{DFT} \circ T_w = M_{\hat{w}} \circ \mathrm{DFT}$.

Si on compose avec la IDFT à gauche et à droite de l'expression antécédente on obtient :

$$\mathrm{DFT} \circ T_w \circ \mathrm{IDFT} = M_{\hat{w}}.$$

On va expliciter cette relation dans la base canonique B, comme dans le cas de l'opérateur de translation. DFT et IDFT deviennent W_N et W_N^{-1} et l'opérateur de multiplication $M_{\hat{w}}$ devient la matrice diagonale $D_{\hat{w}} = \mathrm{diag}(\hat{w}(0), \ldots, \hat{w}(N-1))$. Si la matrice A_w est la représentation de T_w dans la base B, i.e. $A_w = [T_w]_B$, alors :

$$W_N A_w W_N^{-1} = D_{\hat{w}},$$

ce qui montre que l'action de l'opérateur de convolution est diagonalisée dans la base de Fourier.

Les cas des opérateurs de translation et de convolution ne sont pas spéciaux, en fait il existe une catégorie spécifique d'opérateurs dont l'action est diagonale dans la base de Fourier. Ces opérateurs s'appellent *stationnaires* et ils sont examinés en détail dans la section suivante.

Le fait d'avoir déjà examiné les opérateurs de convolution et les multiplicateurs de Fourier nous permettra de montrer un résultat très importante : l'opérateur de convolution et le multiplicateur de Fourier sont les prototypes des opérateurs stationnaires, dans le sens que l'action de tout opérateur stationnaire peut être transformée en une convolution ou un multiplicateur de Fourier.

2.8 La DFT et les opérateurs stationnaires

La transformée de Fourier a une relation particulière avec une classe d'opérateurs qu'on dit « stationnaires ». En fait, elle permet de diagonaliser ces opérateurs et de montrer qu'ils sont équivalents à des opérateurs de convolution et, aussi, à des multiplicateurs de Fourier.

Après avoir défini précisément les opérateurs stationnaires, pour démontrer les résultats mentionnés ci-dessus, on devra aussi introduire une catégorie de matrices, dites circulantes, qui représentent les opérateurs stationnaires dans la base canonique de $\ell^2(\mathbb{Z}_N)$.

Pour motiver la définition d'opérateur stationnaire, imaginons de devoir transformer un signal audio z avec un appareil T. Si on transmet le signal z avec un retard Δt et le seul effet sur T est de retarder également la sortie du signal transformé de la quantité Δt, alors on dit que l'appareil T est stationnaire.

Autrement dit, si l'action de T sur le signal z est indépendante de l'instant de temps dans lequel on l'applique à A, alors T est stationnaire.

Si R_k est l'opérateur de translation de la quantité $k \in \mathbb{Z}$, alors la stationnarité de T est traduite par la relation suivante :

$$T(R_k z) = R_k(Tz), \qquad \forall z \in \ell^2(\mathbb{Z}_N).$$

En fait, le côté de gauche représente l'action de l'opérateur T sur le signal z décalé de la quantité k, tandis que le côté de droite représente le décalage de l'action de l'opérateur T sur le signal original z. Le diagramme commutatif suivant résume ce qu'on vient de discuter.

$$\begin{array}{ccc} \ell^2(\mathbb{Z}_N) & \xrightarrow{R_k} & \ell^2(\mathbb{Z}_N) \\ {\scriptstyle T}\downarrow & & \downarrow{\scriptstyle T} \\ \ell^2(\mathbb{Z}_N) & \xrightarrow[R_k]{} & \ell^2(\mathbb{Z}_N) \end{array}$$

Cette considération donne la motivation pour la définition suivante.

Déf. 2.8.1 *Un opérateur $T : \ell^2(\mathbb{Z}_N) \to \ell^2(\mathbb{Z}_N)$ est dit* ***stationnaire*** *(ou invariant par translation) si*

$$T(R_k z) = R_k(Tz), \qquad \forall z \in \ell^2(\mathbb{Z}_N),\ \forall k \in \mathbb{Z}, \tag{2.37}$$

i.e. T est stationnaire s'il commute avec tout opérateur de translation R_k :

$$\boxed{T \circ R_k = R_k \circ T}, \qquad \forall k \in \mathbb{Z}. \tag{2.38}$$

On démontrera dans 2.8.5 qu'un opérateur linéaire $T \in \mathrm{End}(\ell^2(\mathbb{Z}_N)$ est stationnaire si et seulement $(Tz)(n) = \sum_{k=0}^{N-1} a_k z(n-k) = \sum_{k=0}^{N-1} a_k R_k z(n)$, $n \in \{0, \ldots, N-1\}$, $a_k \in \mathbb{C}$.

Un exemple extrêmement important d'opérateur non stationnaire sur $\ell^2(\mathbb{Z}_N)$ est la transformée de Fourier discrète ! Pour prouver que **la DFT n'est pas un opérateur stationnaire**, il suffit de se rappeler comment elle interagit avec les opérateurs de translation R_k : à partir des formule (2.29) et (2.31) on a, respectivement,

$$\mathrm{DFT} \circ R_k = M_{\omega_N^k} \circ \mathrm{DFT} \quad \text{et} \quad R_k \circ \mathrm{DFT} = \mathrm{DFT} \circ M_{\overline{\omega_N^k}},$$

ce qui montre que la DFT ne commute pas avec les opérateurs de translations à cause du changement de phase engendré par l'action de ce type d'opérateurs.

2.8.1 La DFT et la diagonalisation des opérateurs stationnaires

On pourrait résumer dans un seul théorème les plus importantes propriétés de la DFT en relation aux opérateurs stationnaires, mais on préfère souligner le fait que la transformée de Fourier diagonalise les opérateurs stationnaires dans un énonce séparé.

Théorème 2.8.1 *Soit $T \in End(\ell^2(\mathbb{Z}_N))$ un opérateur linéaire stationnaire. Alors, T est diagonalisable, en fait chaque élément de la base orthogonale de Fourier F_m de $\ell^2(\mathbb{Z}_N)$ est un vecteur propre de T.*

Preuve. Pour chaque $m \in \{0, \ldots, N-1\}$ *fixé*, on considère l'élément m de la base orthogonale de Fourier : $F_m(n) = \frac{1}{N} e^{2\pi i \frac{mn}{N}}$.

Comme T est un endomorphisme, $TF_m \in \ell^2(\mathbb{Z}_N)$, et donc *on peut décomposer TF_m sur la base $(F_0, \ldots, F_{N-1})$ elle-même* :

$$(TF_m)(n) = \sum_{k=0}^{N-1} a_k F_k(n) = \frac{1}{N} \sum_{k=0}^{N-1} a_k e^{2\pi i \frac{kn}{N}}, \qquad \forall n \in \mathbb{Z}_N. \tag{2.39}$$

On considère maintenant l'action de l'opérateur de translation R_1 sur F_m :

$$\begin{aligned} R_1 F_m(n) = F_m(n-1) &= \frac{1}{N} e^{2\pi i \frac{m(n-1)}{N}} = e^{-2\pi i \frac{m}{N}} \cdot \frac{1}{N} e^{2\pi i \frac{mn}{N}} \\ &= e^{-2\pi i \frac{m}{N}} \cdot F_m(n). \end{aligned}$$

Si on applique T à $R_1 F_m$ on obtient :

$$\begin{aligned} TR_1 F_m(n) &= T\left(e^{-2\pi i \frac{m}{N}} \cdot F_m(n)\right) \\ &\underset{\text{Linéarité de } T}{=} e^{-2\pi i \frac{m}{N}} \left(TF_m\right)(n) \\ &\underset{\text{éq. (2.39)}}{=} e^{-2\pi i \frac{m}{N}} \sum_{k=0}^{N-1} a_k F_k(n) \\ &= \sum_{k=0}^{N-1} a_k e^{-2\pi i \frac{m}{N}} F_k(n). \end{aligned}$$

Maintenant on permute l'ordre de composition entre R_1 et T :

$$\begin{aligned} R_1 TF_m(n) &= TF_m(n-1) \\ &\underset{\text{éq. (2.39)}}{=} \frac{1}{N} \sum_{k=0}^{N-1} a_k e^{2\pi i \frac{k(n-1)}{N}} \\ &= \frac{1}{N} \sum_{k=0}^{N-1} a_k e^{-2\pi i \frac{k}{N}} \cdot e^{2\pi i \frac{kn}{N}} \\ &= \sum_{k=0}^{N-1} a_k e^{-2\pi i \frac{k}{N}} F_k(n). \end{aligned}$$

Mais T est stationnaire, donc $TR_1F_m = R_1TF_m$, i.e.

$$\sum_{k=0}^{N-1} a_k e^{-2\pi i \frac{m}{N}} F_k(n) = \sum_{k=0}^{N-1} a_k e^{-2\pi i \frac{k}{N}} F_k(n),$$

mais grâce au théorème d'unicité de la décomposition sur une base :

$$a_k e^{-2\pi i \frac{m}{N}} = a_k e^{-2\pi i \frac{k}{N}}, \qquad \forall k \in \mathbb{Z}_N, \ (m \text{ est fixé}). \tag{2.40}$$

On analyse cette égalité : si $k = m$, alors l'éq. (2.40) devient une identité et on a rien à discuter. Pour examiner le cas $k \neq m$ il faut rappeler que $m, k \in \{0, \ldots, N-1\}$, donc le cos et le sin des exponentiels complexes prennent leurs valeurs dans une période seulement, car la période suivante commence quand $m, k = N$! Donc :

$$k \neq m \quad \Longrightarrow \quad e^{-2\pi i \frac{m}{N}} \neq e^{-2\pi i \frac{k}{N}},$$

et l'éq. (2.40) peut être vérifiée si et seulement si $a_k = 0 \ \forall k \neq m$.

En conséquence, l'éq. (2.39) devient

$$TF_m(n) = a_m F_m(n), \qquad \forall n \in \mathbb{Z}_N,$$

c'est à dire, F_m est un vecteur propre de T avec une valeur propre a_m qui est donnée par l'm-ième coefficient de la décomposition de TF_m sur la base orthogonale de Fourier. Bien sûr, le coefficient a_m dépend de T.

On rappelle qu'on a fixé un indice m arbitraire, donc tout élément de la base orthogonale de Fourier est un vecteur propre de T, et, par conséquent, $\ell^2(\mathbb{Z}_N)$ a une base de vecteurs propres de T. Par définition, T est diagonalisable. □

Dans le théorème 2.8.2 on verra comment expliciter les valeurs propres a_m grâce à la DFT.

On peut donner une interprétation matricielle du théorème qu'on vient de voir : en fait, on sait que l'action de la DFT est représentée par la matrice W_N définie dans l'éq. (2.5) et que W_N est la matrice de passage de la base canonique B de $\ell^2(\mathbb{Z}_N)$ à la base de Fourier F de $\ell^2(\mathbb{Z}_N)$, avec inverse $W_N^{-1} = \frac{1}{N} W_N^*$, qui représente la matrice de passage de base de F à B.

Si A est la matrice associée à T par rapport à la base canonique de $\ell^2(\mathbb{Z}_N)$ et D est la matrice diagonale des valeurs propres de A, alors :

$$\boxed{D = W_N A W_N^{-1}}, \qquad \boxed{A = W_N^{-1} D W_N}, \tag{2.41}$$

en fait, si $[w]_F$ représente un vecteur quelconque $w \in \ell^2(\mathbb{Z}_N)$ par rapport à la base de Fourier F, alors :

$$W_N A z = [Az]_F \underset{(F \text{ diagonalise } A)}{=} D[z]_F = DW_N z, \quad \forall z \in \ell^2(\mathbb{Z}_N),$$

donc $W_N A = DW_N$, si et seulement si : $W_N A W_N^{-1} = DW_N W_N^{-1} = D$.

2.8.2 Matrices circulantes

Pour pouvoir énoncer et démontrer le théorème fondamental relatif au lien entre transformée de Fourier et opérateurs stationnaires, on doit introduire un dernier objet : les matrices circulantes.

On commence par généraliser la périodicité des suites de $\ell^2(\mathbb{Z}_N)$ aux matrices : donnée une matrice $A = (a_{mn})_{m,n=0}^{N-1}$, on dit que A est une **matrice N-périodique** si :

$$a_{m+kN,n} = a_{m,n} \quad \text{et} \quad a_{m,n+kN} = a_{m,n}, \qquad \forall m, n, k \in \mathbb{Z}.$$

Exemple : $a_{0,2} = a_{N,2} = a_{N,N+2}$.

Déf. 2.8.2 *Soit $A = (a_{mn})_{m,n=0}^{N-1}$ une matrice $N \times N$ périodique. A est dite **circulante** si :*

$$\boxed{a_{m+1,n+1} = a_{m,n}}, \qquad \forall m, n \in \mathbb{Z},$$

si on répète k fois la translation, on peut réécrire la définition comme ceci :

$$a_{m+k,n+k} = a_{m,n}, \qquad \forall m, n, k \in \mathbb{Z},$$

on observe que, comme $k \in \mathbb{Z}$, on peut définir une matrice périodique circulante aussi avec la propriété : $a_{m-k,n-k} = a_{m,n}$, $k \in \mathbb{Z}$.

L'interprétation de la définition est la suivante : la ligne (colonne) $m+1$ $(n+1)$ est obtenue de la ligne (colonne) m (n) par translation à droite (en bas) d'une position, comme on peut le voir dans la matrice suivante.

$$\begin{pmatrix} a_0 & a_1 & a_2 & \dots & a_{N-1} \\ a_{N-1} & a_0 & a_1 & \dots & a_{N-2} \\ a_{N-2} & a_{N-1} & a_0 & \dots & a_{N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & a_3 & \dots & a_0 \end{pmatrix}.$$

Exemple de matrice circulante :

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

Exemple de matrice non-circulante :

$$B = \begin{pmatrix} 2 & i & 3 \\ 3 & 2 & i \\ i & 2 & 3 \end{pmatrix}.$$

Pour être circulante, la 3-ième ligne devrait être $(i, 3, 2)$.

2.8.3 La caractérisation exhaustive des opérateurs stationnaires

Le théorème suivante, le plus important du chapitre, permettra d'expliciter les valeurs propres d'un opérateur stationnaire T d'une façon très simple et aussi de caractériser T comme opérateur de convolution, dans la représentation originale de z, et comme un multiplicateur, dans la représentation fréquentielle.

Théorème 2.8.2 *Soit $T : \ell^2(\mathbb{Z}_N) \to \ell^2(\mathbb{Z}_N)$ un endomorphisme. Les propriétés suivantes sont équivalentes.*

(1) *T est stationnaire ;*

(2) *La matrice A qui représente T dans la base canonique de $\ell^2(\mathbb{Z}_N)$ est circulante ;*

(3) *T est un opérateur de convolution ;*

(4) *T est un multiplicateur de Fourier ;*

(5) *La matrice D qui représente T dans la base orthogonale de Fourier F est diagonale.*

On note qu'on a déjà démontré l'implication (1) $\implies$ (5) On démontrera le théorème avec la stratégie suivante :

$$\boxed{(1) \implies (2) \implies (3) \implies (1)} \quad \text{et} \quad \boxed{(3) \iff (4)} \quad \text{et} \quad \boxed{(4) \iff (5)}$$

La *preuve du théorème a une importance fondamentale*, car elle donne une technique explicite pour trouver les valeurs propres de T et pour construire l'opérateur de convolution et le multiplicateur de Fourier qui représentent T.

Preuve.

$\boxed{(1) \implies (2)}$: soit A la matrice associée à T via la base canonique [8] $(e_n)_{n=0}^{N-1}$ de $\ell^2(\mathbb{Z}_N)$:

$$A = \begin{pmatrix} a_{0,0} & a_{0,1} & \cdots & a_{0,N-1} \\ a_{1,0} & a_{1,1} & \cdots & a_{1,N-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N-1,0} & a_{N-1,1} & \cdots & a_{N-1,N-1} \end{pmatrix}.$$

Par définition de matrice associée on a : $a_{m,n} = (Te_n)(m)$, i.e. la n-ième colonne de A est le vecteur Te_n.

En utilisant le fait que T est stationnaire, on doit démontrer que :

$$a_{m+1,n+1} = a_{m,n} \quad \iff \quad (Te_{n+1})(m+1) = (Te_n)(m), \qquad \forall m, n \in \mathbb{Z}_N.$$

On observe que :

$$(R_1 e_n)(m) = e_n(m-1) = \begin{cases} 1 & \text{si } n = m-1 \quad \iff \quad m = n+1 \\ 0 & \text{si } n \neq m-1 \quad \iff \quad m \neq n+1 \end{cases} = e_{n+1}(m) \quad \forall m \in \mathbb{Z}_N,$$

8. On rappelle que $e_n(m) = \delta_{n,m}$, $\forall n, m \in \mathbb{Z}_N$.

donc $e_{n+1} = R_1 e_n$ et alors :

$$a_{m+1,n+1} = (TR_1e_n)(m+1) \underset{(T \text{ stationnaire})}{=} R_1(Te_n)(m+1) = (Te_n)(m+1-1) = (Te_n)(m) = a_{m,n}.$$

Comme $a_{m+1,n+1} = a_{m,n}\ \forall m, n \in \mathbb{Z}_N$, alors A est circulante et l'implication $\boxed{(1) \implies (2)}$ est démontrée.

$\boxed{(2) \implies (3)}$: soit A une matrice périodique circulante, i.e. $a_{m,n} = a_{m-k,n-k}\ \forall n, m, k \in \mathbb{Z}$, on doit démontrer qu'il existe $h \in \ell^2(\mathbb{Z}_N)$ tel que $Az = z * h = T_h(z)\ \forall z \in \ell^2(\mathbb{Z}_N)$.

On va prouver que la suite h qu'on cherche est la première colonne de A, i.e. :

$$\boxed{h = Te_0} = \begin{pmatrix} a_{0,0} \\ \vdots \\ a_{N-1,0} \end{pmatrix}, \quad h(m) = a_{m,0}, \qquad \forall m \in \mathbb{Z}_N.$$

On observe que $h(m-n) = a_{m-n,0} = a_{m-n,n-n} \underset{(A \text{ circulante})}{=} a_{m,n}$, et alors, par définition de produit matrice par vecteur, on a :

$$(Az)(m) = \sum_{n=0}^{N-1} a_{m,n} z(n) = \sum_{n=0}^{N-1} h(m-n) z(n) = (h * z)(m)$$

et l'implication $\boxed{(2) \implies (3)}$ est démontrée.

$\boxed{(3) \implies (1)}$: il faut prouver qu'un opérateur de convolution T_w est stationnaire, i.e. :

$$(T_w \circ R_k)(z) = (R_k \circ T_w)(z), \quad \forall z \in \ell^2(\mathbb{Z}_n),\ \forall k \in \mathbb{Z}.$$

On calcule d'abord le côté gauche de l'équation :

$$(T_w R_k z)(m) = (w * R_k z)(m) = \sum_{n=0}^{N-1} w(m-n) R_k z(n) = \sum_{n=0}^{N-1} w(m-n) z(n-k).$$

On va faire le changement d'indice suivant : $\ell = n - k \Leftrightarrow n = k + \ell$, la variabilité de ℓ est :

$$\begin{cases} n = 0 \implies \ell = -k \\ \vdots \\ n = N-1 \implies \ell = N-1-k, \end{cases}$$

alors :

$$\begin{aligned}(T_w R_k z)(m) &= \sum_{\ell=-k}^{N-1-k} w(m-k-\ell) z(\ell) \underset{\text{Lemme 2.7.1}}{=} \sum_{\ell=0}^{N-1} w((m-k)-\ell) z(\ell) \\ &= (z * w)(m-k) = R_k(z * w)(m) = (R_k T_w z)(m),\end{aligned}$$

et donc l'implication $\boxed{(3) \implies (1)}$ est démontrée.

$\boxed{(3) \Longleftrightarrow (4)}$: on doit démonter qu'un opérateur linéaire $T : \ell^2(\mathbb{Z}_N) \to \ell^2(\mathbb{Z}_N)$ est un opérateur de convolution si et seulement si T est un multiplicateur de Fourier.

Si on fixe un élément $w \in \ell^2(\mathbb{Z}_N)$, quelconque, on a :

$$\begin{aligned} T_w(z) = z * w = \mathrm{IDFT}(\mathrm{DFT}(z * w)) &\underset{\text{(Th. 2.34)}}{=} \mathrm{IDFT}(\hat{z} \cdot \hat{w}) = \mathrm{IDFT}(\hat{w} \cdot \hat{z}) \\ &= (\mathrm{IDFT} \circ M_{\hat{w}} \circ \mathrm{DFT})(z) = T_{(\hat{w})}(z), \quad \forall w, z \in \ell^2(\mathbb{Z}_N), \end{aligned}$$

où $M_{\hat{w}}$ est l'opérateur de multiplication par la suite $\hat{w}$. Vice-versa :

$$\begin{aligned} T_{(w)}(z) = (\mathrm{IDFT} \circ M_w \circ \mathrm{DFT})(z) = \mathrm{IDFT}(w \cdot \hat{z}) &\underset{\text{éq. (2.36)}}{=} \check{w} * \check{\hat{z}} = \check{w} * z \\ &= T_{\check{w}}(z) \quad \forall w, z \in \ell^2(\mathbb{Z}_N), \end{aligned}$$

Cela montre que **l'opérateur de convolution avec w peut être interprété comme le multiplicateur de Fourier par $\hat{w}$** et vice-versa, **le multiplicateur de Fourier par w peut être interprété comme l'opérateur de convolution avec $\check{w}$** :

$$\boxed{T_w = T_{(\hat{w})}}, \quad \boxed{T_{(w)} = T_{\check{w}}} \qquad \forall w \in \ell^2(\mathbb{Z}_N).$$

La double implication $\boxed{(3) \Longleftrightarrow (4)}$ est donc démontrée.

Avant de passer à la dernière étape de la preuve, on résume ce qu'on a vu jusqu'à présent : un opérateur stationnaire $T : \ell^2(\mathbb{Z}_N) \to \ell^2(\mathbb{Z}_N)$ est représenté par une matrice circulante A relativement à la base canonique $(e_0, \ldots, e_{N-1})$ de $\ell^2(\mathbb{Z}_N)$.

La matrice A, à son tour, peut être représentée par l'opérateur de convolution T_h avec $h = Te_0$, la première colonne de A ou, comme l'on vient de le voir, par le multiplicateur de Fourier $T_{(\hat{h})}$, où $\hat{h}$ est la suite des coefficients de Fourier de h.

$\boxed{(4) \Longleftrightarrow (5)}$: On doit prouver que T est un multiplicateur de Fourier $T_{(w)}$, si et seulement si la matrice associée à T par rapport à la base orthogonale de Fourier F est diagonale.

L'implication directe a déjà été démontrée dans la formule (2.26), nous allons donc à démontrer tout simplement l'implication (5) $\Longrightarrow$ (4) : dire que $D = \mathrm{diag}(d_{n,n})$, $n = 0, \ldots, N-1$, est la matrice diagonale qui représente l'opérateur T dans la base de Fourier F veut dire que

$$[T(z)]_F = D[z]_F \iff \mathrm{DFT} \circ T(z) = M_w \circ \mathrm{DFT}(z),$$

étant M_w l'opérateur de multiplication par la suite $w(n) = d_{n,n}$, $n = 0, \ldots, N-1$. Si on applique la IDFT aux deux côtés :

$$T(z) = \mathrm{IDFT} \circ M_w \circ \mathrm{DFT}(z) \qquad \forall z \in \ell^2(\mathbb{Z}_N),$$

mais alors $T = T_{(w)}$ et l'implication (5) $\Longrightarrow$ (4) est démontrée.

Le théorème est donc complètement prouvé. □

Le théorème qu'on vient de démontrer donne une technique standard pour l'étude des opérateurs stationnaires T sur $\ell^2(\mathbb{Z}_N)$. On rappelle que la suite :

$$\delta \in \ell^2(\mathbb{Z}_N), \quad \delta(n) = e_0(n) = \delta_{0,n} = \begin{cases} 1 & \text{si } n = 0 \\ 0 & \text{si } n \neq 0 \end{cases} \qquad \forall n \in \mathbb{Z}_N,$$

est dite **impulsion unitaire**, alors **l'opérateur T est complètement déterminé par son action sur δ**, $h = T\delta$, qu'on appelle **réponse impulsionnelle**. $\hat{h}$, la DFT de la réponse impulsionnelle, est dite **fonction de transfert**.

En fait, grâce aux propriétés démontrées dans les théorèmes 2.8.1 et 2.8.2, on peut résumer l'analyse des opérateurs stationnaires comme ci-dessous.

Analyse des opérateurs stationnaires sur $\ell^2(\mathbb{Z}_N)$:

— T, opérateur stationnaire de $\ell^2(\mathbb{Z}_N)$;

— A, matrice **circulante** associée à T par rapport à la *base canonique* de $\ell^2(\mathbb{Z}_N)$;

— h, **réponse impulsionnelle de** T :

$$h = T\delta = \textbf{ première colonne de } A;$$

— T_h, **opérateur de convolution avec** h :

$$Tz = T_h z = h * z = z * h;$$

— $T_{(\hat{h})}$, **multiplicateur de Fourier par** $\hat{h}$, la **fonction de transfert** :

$$Tz = T_{(\hat{h})} z = \mathrm{IDFT}(\hat{h} \cdot \hat{z});$$

— Étant donné $h = T\delta$, on a le couple de Fourier suivante :

Représentation originale	Espace de Fourier
$h * z$	$\hat{h} \cdot \hat{z}$

— D, matrice diagonale qui représente T dans la base orthogonale F de Fourier de $\ell^2(\mathbb{Z}_N)$:

$$D = \frac{1}{N} W_N A W_N^* = \mathrm{diag}(\hat{h}(0), \ldots, \hat{h}(N-1)).$$

— Les valeurs propres de T (le spectre dans le sens de l'algèbre linéaire !) sont les composantes de la fonction de transfert, i.e. les coefficients de Fourier de la réponse impulsionnelle, i.e.

$$\textbf{Valeurs propres de } T \ : \ \{\hat{h}(0), \ldots, \hat{h}(N-1)\}.$$

2.8.4 Filtres passe-haut, passe-bas, passe-bande

La formule de synthèse pour un signal $z \in \ell^2(\mathbb{Z}_N)$ quelconque transformé via l'action d'un opérateur *stationnaire* $T \in \mathrm{End}(\ell^2(\mathbb{Z}_N))$ est :

$$Tz(n) = \sum_{m=0}^{N-1} \widehat{Tz}(m) F_m(n) \qquad n \in \mathbb{Z}_N, \tag{2.42}$$

F_m étant le vecteur d'indice m de la base orthogonale de Fourier de $\ell^2(\mathbb{Z}_N)$. Donc, $|\widehat{Tz}(m)|$ représente l'importance de l'harmonique de fréquence m dans la reconstruction de Tz et $\{\widehat{Tz}(m),\ m \in \mathbb{Z}_N\}$ représente le spectre du signal transformé Tz.

Pour comprendre comment lier le spectre de Tz et celui du signal original z, allons appliquer la DFT aux deux côtés de la formule $Tz = T_{(\hat{h})}z = \mathrm{IDFT}(\hat{h} \cdot \hat{z})$, où $h = T\delta_0$:

$$\mathrm{DFT}(Tz) = \mathrm{DFT} \circ \mathrm{IDFT}(\hat{h} \cdot \hat{z}) = \hat{h} \cdot \hat{z},$$

i.e.

$$\boxed{\widehat{Tz}(m) = \hat{h}(m) \cdot \hat{z}(m)}, \qquad \forall m \in \mathbb{Z}_N,$$

donc les coefficients de Fourier de Tz, la suite transformée par l'opérateur T, sont donnés par le produit des coefficients de Fourier de la suite originale z fois les coefficients de Fourier de la réponse impulsionnelle h.

Par conséquent, le **spectre de la suite transformée** Tz est :

$$\left\{ \left| \widehat{Tz}(m) \right| = |\hat{h}(m)| \cdot |\hat{z}(m)|,\ m \in \mathbb{Z}_N \right\}. \tag{2.43}$$

Cela permet de *comprendre l'action des filtres stationnaires T sur le contenu fréquentiel d'un signal z* :

- Si $\hat{h}(0) = 0$, Tz a **moyenne nulle**, car $\left|\widehat{Tz}(0)\right| = 0 \cdot |\hat{z}(0)| = 0$ et on sait que $\left|\widehat{Tz}(0)\right|$ est proportionnelle à la moyenne de Tz ;
- Si $|\hat{h}(0)| = 1$, alors T **conserve la moyenne** de z, i.e. $\langle Tz \rangle = \langle z \rangle$;
- Si $|\hat{h}(m)| > 1$ pour $m \simeq 0$ et $m \simeq N-1$ et $|\hat{h}(m)| \in [0, 1[$ pour $m \simeq N/2$, alors T augmente les basses fréquences et réduit les hautes fréquences (filtre **passe-bas**) ;
- Si $|\hat{h}(m)| > 1$ pour $m \simeq N/2$ et $|\hat{h}(m)| \in [0, 1[$ pour $m \simeq 0$ et $m \simeq N-1$, alors T augmente les hautes fréquences et réduit les basses fréquences (filtre **passe-haut**) ;
- Si $|\hat{h}(m)| > 1$ pour des valeurs intermédiaires de m, alors T augmente les fréquences moyennes (filtre **passe-bande**).

2.8.5 La caractérisation des opérateurs stationnaires via les opérateurs de translation

On a tous les résultats pour démontrer comment *caractériser un opérateur stationnaire comme une combinaison linéaire d'opérateurs de translation*, ou, d'une manière équivalente, comme un *polynôme de l'opérateur de translation* R_1, vu que $R_k = R_1 \circ \cdots \circ R_1$ k fois, i.e. R_1^k, $\forall k \in \mathbb{Z}$.

Théorème 2.8.3 *$T \in End(\ell^2(\mathbb{Z}_N))$ est stationnaire si et seulement si T a cette expression :*

$$(Tz)(n) = \sum_{k=0}^{N-1} a_k z(n-k) = \sum_{k=0}^{N-1} a_k R_k z(n) = \sum_{k=0}^{N-1} a_k (R_1)^k z(n), \qquad \forall n \in \{0, .., N-1\}, \quad (2.44)$$

où $a_k \in \mathbb{C}$.

Preuve.
$\boxed{\Longrightarrow}$: soit T stationnaire, alors on sait que $T = T_h$, où T_h est l'opérateur de convolution par rapport à la réponse impulsionnelle $h = T\delta$, i.e.

$$(Tz)(n) = \sum_{k=0}^{N-1} h(k) z(n-k),$$

mais alors ça suffit d'identifier les coefficients a_k de la formule $(Tz)(n) = \sum_{k=0}^{N-1} a_k z(n-k)$ avec $h(k)$ pour avoir la thèse.

$\boxed{\Longleftarrow}$: on peut vérifier que T écrit comme dans la formule (2.44) est stationnaire grâce à la linéarité de T et de R_k. En fait, $\forall n \in \{0, \ldots, N-1\}$:

$$(TR_m z)(n) = T(R_m z(n)) = T(z(n-m)) = \sum_{k=0}^{N-1} a_k z(n-k-m) = \sum_{k=0}^{N-1} a_k R_m z(n-k)$$
$$\underset{\text{(linéarité de } R_m)}{=} R_m \left(\sum_{k=0}^{N-1} a_k z(n-k) \right) = (R_m T z)(n),$$

donc : $T \circ R_m = R_m \circ T$ $\forall m \in \mathbb{Z}$. □

Comme $h(k) = T\delta(k)$, la preuve du théorème nous a montré la validité de la formule :

$$\boxed{(Tz)(n) = \sum_{k=0}^{N-1} T\delta(k) z(n-k)} \qquad \forall\, T \text{ stationnaire.}$$

2.8.6 Analyse fréquentielle des opérateurs de dérivation (discrète) première et seconde

Dans cette section on va analyser deux opérateurs stationnaires qui représentent la version discrète de la dérivée première et seconde. La comparaison de leurs valeurs propres montrera pourquoi l'opérateur de dérivation seconde est plus efficace pour amplifier les hautes fréquences dans les signaux numériques.

Déf. 2.8.3 *Étant donnée une suite $z \in \ell^2(\mathbb{Z}_N)$, on définit :*

$$T_1 z(n) = z(n+1) - z(n) \qquad \textit{Dérivée première discrète}$$

$$T_2 z(n) = z(n+1) - 2z(n) + z(n-1) \qquad \textit{Dérivée seconde discrète}$$

La *dérivée première discrète* est simplement *la différence en avant* de z, divisé par la différence des valeurs de n, mais comme $(n+1) - n = 1$ on peut éviter d'écrire le dénominateur.

La *dérivée seconde discrète* est la *différence en arrière de la dérivée première* de z, divisé par la différence des valeurs de n, qui est encore une fois 1, donc on ne l'écrit pas. Explicitement : $T_2 z(n) = T_1 z(n) - T_1 z(n-1) = z(n+1) - z(n) - [z(n) - z(n-1)] = z(n+1) - 2z(n) + z(n-1)$.

On commence par l'analyse de T_1. Pour calculer sa réponse impulsionnelle on doit appliquer T_1 à l'impulsion unitaire $\delta = e_0$:

$$h = T_1\delta = \begin{pmatrix} e_0(1) - \boxed{e_0(0)} & \longleftarrow n = 0 \\ e_0(2) - e_0(1) & \longleftarrow n = 1 \\ \vdots & \\ \boxed{e_0(N-1+1)} - e_0(N-1) & \longleftarrow n = N-1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

où on a utilisé le fait que $e_0(0) = e_0(N) = 1$. La matrice qui représente T_1 dans la base canonique de $\ell^2(\mathbb{Z}_N)$ est :

$$A_{T_1} = \begin{pmatrix} -1 & 1 & 0 & \dots & 0 \\ 0 & -1 & 1 & \dots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & -1 & 1 \\ 1 & 0 & \dots & 0 & -1 \end{pmatrix}.$$

On calcule maintenant la DFT de h, pour tout $m \in \mathbb{Z}_N$ ça vaut :

$$\begin{aligned} \hat{h}(m) &= \sum_{n=0}^{N-1} h(n) e^{-2\pi i \frac{mn}{N}} = -1 \cdot e^{-2\pi i \frac{m0}{N}} + 0 + \ldots + 1 \cdot e^{-2\pi i \frac{m(N-1)}{N}} \\ &= -1 + e^{2\pi i \frac{m}{N}} e^{-2\pi i \frac{mN}{N}} = e^{2\pi i \frac{m}{N}} - 1. \end{aligned}$$

donc les valeurs propres de T_1 sont $\{\hat{h}(m) = e^{2\pi i \frac{m}{N}} - 1, m = 0, 1, \ldots, N-1\}$ et sa représentation diagonale est :

$$D = \operatorname{diag}\left(0, e^{2\pi i \frac{1}{N}} - 1, e^{2\pi i \frac{2}{N}} - 1 \ldots, e^{2\pi i \frac{(N-1)}{N}} - 1\right).$$

Allons maintenant à interpréter l'action fréquentielle de T_1 grâce à la formule (2.43). Il faut calculer les modules des valeurs propres $(\hat{h}(m))_{m\in\mathbb{Z}_N}$. On observe que

$$e^{2\pi i\frac{m}{N}} - 1 = e^{\pi i\frac{m}{N}}(e^{\pi i\frac{m}{N}} - e^{-\pi i\frac{m}{N}}) = e^{\pi i\frac{m}{N}} 2i \sin\left(\pi\frac{m}{N}\right)$$

donc, $|\hat{h}(m)| = \left|e^{\pi i\frac{m}{N}}\right| \cdot \left|2i\sin\left(\pi\frac{m}{N}\right)\right| = 2\left|\sin\left(\pi\frac{m}{N}\right)\right|$, mais $m \in \mathbb{Z}_N$, $\frac{m}{N} < 1$, donc le sinus est toujours non négatif et on peut éliminer la valeur absolue. En résumé :

$$\boxed{\left\{|\hat{h}(m)| = 2\sin\left(\pi\frac{m}{N}\right),\ m \in \mathbb{Z}_N\right\}}.$$

En particulier :

- $|\hat{h}(0)| = 0$: par conséquent, le signal filtré $T_1 z$ a moyenne nulle ;
- $|\hat{h}(\frac{N}{2})| = 2$;
- $|\hat{h}(m)| < 2\ \forall m \neq \frac{N}{2}$;
- $|\hat{h}(m)| \to 0$ si $m \to 0$ ou $m \to N-1$;
- L'action de l'opérateur est symétrique par rapport à $\frac{N}{2}$.

Comme $m = N/2$ représente la plus haute fréquence du signal et $m = 0$ et $m = N-1$ représentent les plus basses fréquences, on déduit que T_1 réduit les basses fréquences de z et il augmente jusqu'à 2 fois les hautes fréquences. Donc, **l'opérateur dérivée première discrète est un filtre passe-haut**.

Maintenant on passe à l'analyse de T_2. Sa réponse impulsionnelle est donné par le vecteur :

$$h = T\delta = \begin{pmatrix} e_0(1) - 2\boxed{e_0(0)} + e_0(-1) \\ e_0(2) - 2e_0(1) + \boxed{e_0(0)} \\ \vdots \\ \boxed{e_0(N)} - 2e_0(N-1) + e_0(N-2) \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

La matrice associée à T_2 dans la base canonique de $\ell^2(\mathbb{Z}_N)$ est :

$$A_{T_2} = \begin{pmatrix} -2 & 1 & 0 & \dots & 1 \\ 1 & -2 & 1 & \dots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & -2 & 1 \\ 1 & 0 & \dots & 1 & -2 \end{pmatrix}.$$

On calcule maintenant la DFT de h :

$$\begin{aligned} \hat{h}(m) &= \sum_{n=0}^{N-1} h(n) e^{-2\pi i\frac{mn}{N}} = -2\cdot e^{-2\pi i\frac{m0}{N}} + 1\cdot e^{-2\pi i\frac{m}{N}} + 0 + \ldots + 1\cdot e^{-2\pi i\frac{m(N-1)}{N}} \\ &= -2 + e^{-2\pi i\frac{m}{N}} + e^{-2\pi im} e^{2\pi i\frac{m}{N}} = -2 + e^{2\pi i\frac{m}{N}} + e^{-2\pi i\frac{m}{N}} \\ &= -2 + 2\cdot\frac{e^{2\pi i\frac{m}{N}} + e^{-2\pi i\frac{m}{N}}}{2} = -2 + 2\cos\left(2\pi\frac{m}{N}\right). \end{aligned}$$

On veut comparer ces valeurs de $\hat{h}(m)$ avec celles de l'opérateur dérivée première, pour cela on réécrit $\hat{h}(m) = -4\left[\frac{1}{2} - \frac{1}{2}\cos\left(2\pi\frac{m}{N}\right)\right]$ et on utilise l'identité trigonométrique $\sin^2(\alpha) = \frac{1}{2} - \frac{1}{2}\cos(2\alpha)$ avec $\alpha = \pi\frac{m}{N}$, pour obtenir $\hat{h}(m) = -4\sin^2\left(\pi\frac{m}{N}\right)$. Donc les valeurs propres de T_2 sont $\{\hat{h}(m) = -4\sin^2\left(\pi\frac{m}{N}\right), m = 0, 1, \ldots, N-1\}$ et sa représentation diagonale est :

$$D = \text{diag}\left(0, -4\sin^2\left(\frac{\pi}{N}\right), -4\sin^2\left(\frac{2\pi}{N}\right), \ldots, -4\sin^2\left(\frac{(N-1)\pi}{N}\right)\right).$$

L'action fréquentielle de T_2 est définie par les modules de ses valeurs propres :

$$\boxed{\left\{|\hat{h}(m)| = 4\sin^2\left(\pi\frac{m}{N}\right),\ m \in \mathbb{Z}_N\right\}},$$

on voit que les modules des valeurs propres de l'opérateur de dérivée seconde sont les carrés de celles de l'opérateur de dérivée première. Par conséquent :

- $|\hat{h}(0)| = 0$: donc, comme pour la première dérivée, le signal filtré $T_2 z$ a moyenne nulle ;
- $|\hat{h}(\frac{N}{2})| = 4$;
- $|\hat{h}(m)| < 4\ \forall m \neq \frac{N}{2}$;
- $|\hat{h}(m)| \to 0$ si $m \to 0$ ou $m \to N-1$ et la convergence à zéro est plus rapide que pour l'opérateur de dérivée première, car dans ce cas le sinus est élevé au carré, comme on le voit dans la figure 2.4.
- L'action de l'opérateur est symétrique par rapport à $\frac{N}{2}$.

Donc, aussi **l'opérateur dérivée seconde discrète est un filtre passe-haut**, avec une action d'amplification des hautes fréquences et de réduction des basses fréquences quadratique par rapport à celle de l'opérateur de dérivée première discrète.

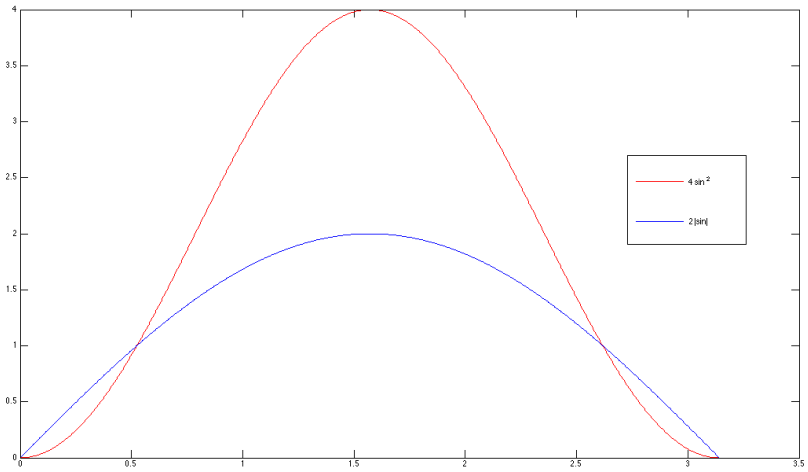


FIGURE 2.4 – Différence entre les fonctions sinusoïdales qui représentent les valeurs du spectre de l'opérateur dérivée première et seconde entre 0 et π.

2.9 La transformée de Fourier bidimensionnelle (DFT 2D)

La transformée de Fourier qu'on a considérée jusqu'à maintenant est appliquée à des signaux $z(n)$ qui dépendent seulement d'un paramètre n.

Dans les applications, on a des signaux très importants qui dépendent de plusieurs paramètres. Un exemple canonique est donné par les **images numériques** qui dépendent de deux paramètres : les *deux coordonnées spatiales d'un pixel*, comme dans la figure 2.5.

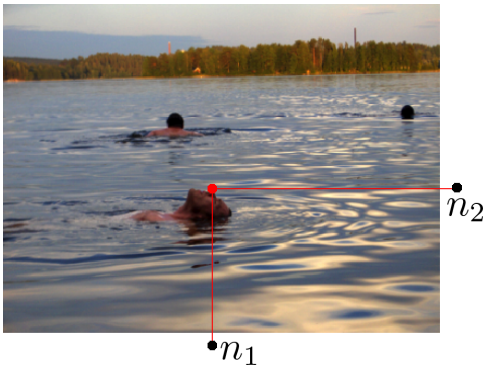


FIGURE 2.5 – Les deux coordonnées n_1, n_2 d'un pixel dans une image numérique.

On peut généraliser la théorie de la DFT pour examiner des signaux qui dépendent de n'importe quel nombre (fini) de paramètres. Néanmoins, pour simplifier l'explication, on va considérer seulement le cas bidimensionnel de deux paramètres n_1, n_2.

On va définir d'abord l'espace de travail : si $N_1, N_2 \in \mathbb{N}$, on définit

$$\ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2}) = \{z : \mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2} \implies \mathbb{C}\},$$

$z \in \ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$ est une suite complexe qui dépend de deux paramètres :

$$\begin{cases} n_1 \in \{0, 1, \ldots, N_1 - 1\} \\ n_2 \in \{0, 1, \ldots, N_2 - 1\}. \end{cases}$$

$\ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$ est un espace vectoriel de dimension $N_1 \cdot N_2$ avec les définitions de somme et de produit par un scalaire complexe définies comme dans le cas 1D et produit scalaire :

$$\langle z, w \rangle = \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} z(n_1, n_2)\overline{w(n_1, n_2)}, \quad \forall z, w \in \ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2}).$$

L'extension de la théorie de la DFT de 1D à 2D est basée sur la procédure de génération de bases de $\ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$ à partir de bases de $\ell^2(\mathbb{Z}_{N_1})$ et de $\ell^2(\mathbb{Z}_{N_2})$.

Théorème 2.9.1 *Soient $\{B_0, B_1, \ldots, B_{N_1-1}\}$, $\{C_0, C_1, \ldots, C_{N_2-1}\}$, une base orthonormale de $\ell^2(\mathbb{Z}_{N_1})$ et de $\ell^2(\mathbb{Z}_{N_2})$, respectivement.*

On définit, $\forall m_1 \in \{0, \ldots, N_1 - 1\}$ et $m_2 \in \{0, \ldots, N_2 - 1\}$ les suites de $\ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$ données par :

$$\boxed{D_{m_1,m_2}(n_1, n_2) = B_{m_1}(n_1) \cdot C_{m_2}(n_2)}.$$

*Alors, D_{m_1,m_2} est une base orthonormale de $\ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$ dite la **base produit tensoriel** de deux bases initiales.*

Preuve. Les suites D_{m_1,m_2}, $m_1 \in \{0,\dots,N_1-1\}$, $m_2 \in \{0,\dots,N_2-1\}$ sont $N_1 \cdot N_2$ éléments de $\ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$, qui a dimension $N_1 \cdot N_2$. Donc, pour démontrer qu'ils sont une base orthonormale, on doit tout simplement montrer que :

$$\langle D_{m_1,m_2}, D_{k_1,k_2} \rangle = \delta_{(m_1,m_2),(k_1,k_2)} = \delta_{m_1,k_1}\delta_{m_2,k_2} = \begin{cases} 1 & \text{si } (m_1,m_2) = (k_1,k_2) \\ 0 & \text{si } (m_1,m_2) \neq (k_1,k_2). \end{cases}$$

$$\begin{aligned} \langle D_{m_1,m_2}, D_{k_1,k_2} \rangle &\underset{\text{déf. de } \langle\,,\,\rangle}{=} \sum_{n_1=0}^{N_1-1}\sum_{n_2=0}^{N_2-1} D_{m_1,m_2}(n_1,n_2)\overline{D_{k_1,k_2}(n_1,n_2)} \\ &\underset{\text{déf. de } D}{=} \sum_{n_1=0}^{N_1-1}\sum_{n_2=0}^{N_2-1} B_{m_1}(n_1)C_{m_2}(n_2)\overline{B_{k_1}(n_1)C_{k_2}(n_2)} \\ &= \sum_{n_1=0}^{N_1-1} B_{m_1}(n_1)\overline{B_{k_1}(n_1)} \sum_{n_2=0}^{N_2-1} C_{m_2}(n_2)\overline{C_{k_2}(n_2)} \\ &= \underbrace{\langle B_{m_1}, B_{k_1}\rangle}_{\shortparallel}\underbrace{\langle C_{m_2}, C_{k_2}\rangle}_{\shortparallel} = \delta_{(m_1,m_2),(k_1,k_2)}. \\ &\qquad\quad \delta_{m_1,k_1} \qquad \delta_{m_2,k_2} \end{aligned}$$

□

Ce théorème a les corollaires suivantes, pour $m_1 \in \{0,1,\dots,N_1-1\}$ et $m_2 \in \{0,1,\dots,N_2-1\}$:

— La base orthonormale canonique de $\ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$ est :

$$\boxed{B = e_{m_1,m_2}(n_1,n_2) = \begin{cases} 1 & \text{si } (n_1,n_2) = (m_1,m_2) \\ 0 & \text{si } (n_1,n_2) \neq (m_1,m_2) \end{cases}}$$

— La base orthogonale de Fourier de $\ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$ est

$$\boxed{F_{m_1,m_2}(n_1,n_2) = \frac{1}{N_1N_2} e^{2\pi i \frac{m_1n_1}{N_1}} \cdot e^{2\pi i \frac{m_2n_2}{N_2}} = \frac{1}{N_1N_2} e^{2\pi i\left(\frac{m_1n_1}{N_1} + \frac{m_2n_2}{N_2}\right)}}$$

— La base orthonormale de Fourier de $\ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$ est

$$E_{m_1,m_2}(n_1,n_2) = \sqrt{N_1N_2} F_{m_1,m_2}(n_1,n_2).$$

— La base orthogonale des exponentiels complexes de $\ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$ est

$$\mathcal{E}_{m_1,m_2}(n_1,n_2) = N_1N_2 F_{m_1,m_2}(n_1,n_2).$$

Grâce à la théorie des espaces vectoriels avec produit scalaire complexe, on peut généraliser la définition des coefficients de Fourier, DFT et IDFT à $\ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$. Soit $z \in \ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$, alors :

$$\begin{aligned}\langle z, \mathcal{E}_{m_1,m_2}\rangle &= \sum_{n_1=0}^{N_1-1}\sum_{n_2=0}^{N_2-1} z(n_1,n_2)\overline{e^{2\pi i\frac{n_1m_1}{N_1}}e^{2\pi i\frac{n_2m_2}{N_2}}}\\ &= \sum_{n_1=0}^{N_1-1}\sum_{n_2=0}^{N_2-1} z(n_1,n_2)e^{-2\pi i\frac{m_1n_1}{N_1}}e^{-2\pi i\frac{m_2n_2}{N_2}}\\ &= \sum_{n_1=0}^{N_1-1}\sum_{n_2=0}^{N_2-1} z(n_1,n_2)e^{-2\pi i(\frac{m_1n_1}{N_1}+\frac{m_2n_2}{N_2})},\end{aligned}$$

donc on définit les **coefficients de Fourier** de $z \in \ell^2(\mathbb{Z}_{N_1}\times\mathbb{Z}_{N_2})$ comme ceci :

$$\boxed{\hat{z}(m_1,m_2) = \sum_{n_1=0}^{N_1-1}\sum_{n_2=0}^{N_2-1} z(n_1,n_2)e^{-2\pi i\left(\frac{m_1n_1}{N_1}+\frac{m_2n_2}{N_2}\right)}}.$$

Comme pour le cas 1D :

$$\boxed{\hat{z}(0,0) = N_1N_2\langle z\rangle},$$

où $\langle z\rangle$ est la moyenne de z. On observe que la quantité N_1N_2 peut être extrêmement élevée.

On peut également généraliser la formule de synthèse au cas 2D comme ceci :

$$\boxed{z(n_1,n_2) = \frac{1}{N_1N_2}\sum_{m_1=0}^{N_1-1}\sum_{m_2=0}^{N_2-1}\hat{z}(m_1,m_2)e^{2\pi i\left(\frac{m_1n_1}{N_1}+\frac{m_2n_2}{N_2}\right)}}.$$

En conséquence, les opérateurs DFT2D et IDFT2D peuvent être écrits via les formules suivantes :

$$\begin{array}{rccl}\hat{}: & \ell^2(\mathbb{Z}_{N_1}\times\mathbb{Z}_{N_2}) & \longrightarrow & \ell^2(\mathbb{Z}_{N_1}\times\mathbb{Z}_{N_2})\\ & z & \longmapsto & \hat{z},\ \hat{z}(m_1,m_2) = \sum_{n_1=0}^{N_1-1}\sum_{n_2=0}^{N_2-1} z(n_1,n_2)e^{-2\pi i\left(\frac{m_1n_1}{N_1}+\frac{m_2n_2}{N_2}\right)},\end{array}$$

et

$$\begin{array}{rccl}\check{}: & \ell^2(\mathbb{Z}_{N_1}\times\mathbb{Z}_{N_2}) & \longrightarrow & \ell^2(\mathbb{Z}_{N_1}\times\mathbb{Z}_{N_2})\\ & z & \longmapsto & \check{z},\ \check{z}(n_1,n_2) = \frac{1}{N_1N_2}\sum_{m_1=0}^{N_1-1}\sum_{m_2=0}^{N_2-1} z(m_1,m_2)e^{2\pi i\left(\frac{m_1n_1}{N_1}+\frac{m_2n_2}{N_2}\right)}.\end{array}$$

Il est clair que, si on augmente la dimension de 2 à $2 < d < +\infty$, les formules se généralisent comme ça :

$$\hat{z}(m_1,\ldots,m_d) = \sum_{n_1=0}^{N_1-1}\cdots\sum_{n_d=0}^{N_d-1} z(n_1,\ldots,n_d)e^{-2\pi i\sum_{k=1}^{d}\frac{m_kn_k}{N_k}},$$

$$\check{z}(n_1,\cdots,n_d) = \left(\prod_{k=1}^{d}N_k\right)^{-1}\sum_{m_1=0}^{N_1-1}\cdots\sum_{m_d=0}^{N_d-1} z(m_1,\ldots,m_d)e^{2\pi i\sum_{k=1}^{d}\frac{m_kn_k}{N_k}}.$$

2.9.1 Représentation matricielle de la DFT 2D : produit de Kronecker vs. itération de deux DFT 1D

La représentation matricielle de la DFT 2D dans la base canonique de $\ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$ peut être construite grâce aux matrices de Sylvester W_{N_1} et W_{N_2} associées respectivement à la DFT 1D en $\ell^2(\mathbb{Z}_{N_1})$ et en $\ell^2(\mathbb{Z}_{N_2})$.

L'opération dont on a besoin pour obtenir la représentation matricielle de la DFT 2D est dite « produit de Kronecker », qu'on va définir de suite.

Déf. 2.9.1 *Données deux matrices A de dimension $m \times n$ et B de dimension $p \times q$:*

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & \cdots & b_{1q} \\ \vdots & \ddots & \vdots \\ b_{p1} & \cdots & b_{pq} \end{pmatrix},$$

la matrice ***produit de Kronecker*** $A \otimes B$ *est la matrice de dimension $mp \times nq$ définie par :*

$$A \otimes B = \begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{pmatrix}.$$

On peut démontrer par calcul direct que la matrice associée à la DFT 2D dans la base canonique de $\ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$ est la matrice de dimension $N_1N_2 \times N_1N_2$ donnée par :

$$\boxed{W_{N_1,N_2} = W_{N_1} \otimes W_{N_2}} \quad \Longrightarrow \quad \boxed{\hat{z}(m_1, m_2) = W_{N_1} \otimes W_{N_2} z(n_1, n_2)}.$$

Malheureusement, le calcul de la matrice produit de Kronecker est trop lourd quand N_1 et N_2 sont grands. Dans la pratique, on préfère réécrire la DFT 2D comme l'itération de deux DFT 1D.

Pour comprendre comment l'on peut faire, *on doit interpréter $z \in \ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$ comme une matrice de N_2 vecteurs colonne de N_1 éléments* :

$$z(n_1, n_2) = \underbrace{\left.\begin{pmatrix} \vdots & \vdots & \cdots & \vdots \\ z(\cdot, 0) & z(\cdot, 1) & \cdots & z(\cdot, N_2 - 1) \\ \vdots & \vdots & \cdots & \vdots \end{pmatrix}\right\}}_{N_2 \text{ vecteurs colonne}} \quad N_1 \text{ éléments pour chaque vecteur colonne.}$$

Par définition de DFT 2D on peut écrire :

$$\hat{z}(m_1, m_2) = \sum_{n_2=0}^{N_2-1} \underbrace{\left(\sum_{n_1=0}^{N_1-1} z(n_1, n_2) e^{-2\pi i \frac{n_1 m_1}{N_1}} \right)}_{\hat{z}(m_1, n_2) = W_{N_1} z(n_1, n_2)} e^{-2\pi i \frac{n_2 m_2}{N_2}} = \sum_{n_2=0}^{N_2-1} W_{N_1} z(n_1, n_2) e^{-2\pi i \frac{n_2 m_2}{N_2}}. \tag{2.45}$$

L'explication de la formule ci-dessus est la suivante : la sommation par rapport à l'indice n_2 est la plus extérieure, donc *on fixe* n_2 à chaque fois. Avec n_2 fixé, $z(n_1, n_2)$ est un vecteur colonne, et donc la parenthèse qu'on a mis en évidence représente la DFT 1D de ce vecteur colonne,

qui peut être obtenue en appliquant la matrice W_{N_1} à $z(n_1, n_2)$, n_2 fixé, comme correctement écrit ci-dessus.

Maintenant le problème est que n_1 est fixe e que l'indice qui change est n_2, donc $W_{N_1}(n_1, n_2)$ est un vecteur ligne et on ne peut pas lui appliquer W_{N_2} pour obtenir la DFT, car, comme on l'a vu dans dans la section 2.5, la DFT 1D est réalisée par le produit de la matrice W_N avec une suite représentée via un *vecteur colonne* !

La solution à ce problème consiste à transposer les deux côtés de l'équation (2.45), ceci transforme le vecteur ligne $\hat{z}(m_1, n_2)$ en un vecteur colonne :

$$\hat{z}(m_1, m_2)^t = \sum_{n_2=0}^{N_2-1} (W_{N_1} z(n_1, n_2))^t e^{-2\pi i \frac{n_2 m_2}{N_2}},$$

maintenant $(W_{N_1} z(n_1, n_2))^t$ est un vecteur colonne, donc on peut calculer sa DFT en appliquant W_{N_2} :

$$\hat{z}(m_1, m_2)^t = W_{N_2}(W_{N_1} z(n_1, n_2))^t \underset{(AB)^t = B^t A^t}{=} W_{N_2}\, z(n_1, n_2)^t\, (W_{N_1})^t = W_{N_2} z(n_2, n_1) W_{N_1},$$

car $W_{N_1}^t = W_{N_1}$ (observer l'échange entre n_1 et n_2). Donc $\hat{z}(m_1, m_2)^t = W_{N_2}\, z(n_2, n_1)\, W_{N_1}$, et alors, pour trouver $\hat{z}(m_1, m_2)$, il suffit de transposer encore une fois les deux côtés :

$$\hat{z}(m_1, m_2) = (\hat{z}(m_1, m_2)^t)^t = (W_{N_2}\, z(n_2, n_1)\, W_{N_1})^t = W_{N_1}\, z(n_1, n_2)\, W_{N_2}.$$

La formule qui permet de calculer la DFT 2D d'une suite $z \in \ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$ est donc :

$$\boxed{\hat{z}(m_1, m_2) = W_{N_1} z(n_1, n_2) W_{N_2}}. \tag{2.46}$$

Il est important de souligner que l'équation (2.46) a du sens seulement si $\hat{z}(m_1, m_2)$ et $z(n_1, n_2)$ sont interprétées dans leur globalité comme des matrices $N_1 \times N_2$.

La formule (2.46) est différente de $W_{N_1} W_{N_2} z(n_1, n_2)$ ou $W_{N_2} W_{N_1} z(n_1, n_2)$, qui sont les formules naïves qu'on aurait pu imaginer pour implémenter la DFT 1D sur les colonnes et sur les lignes de z. Comme on l'a vu, la raison de la différence est due au fait que la DFT 1D matricielle nécessite toujours un vecteur colonne, d'où la nécessité de la transposition, qui amène vers la formule (2.46).

2.9.2 Les propriétés de la DFT 2D

Les propriétés de la DFT 1D, qu'on a vu dans la section 2.7 peuvent être généralisées sans difficulté à la DFT 2D.

Les démonstrations sont pratiquement identiques, mais avec une notation plus lourde, pour cela, on présente l'extension bidimensionnelle des résultats relatifs à la DFT 1D sans preuve.

Comme dans le cas 1D, pour pouvoir discuter les propriétés de la DFT 2D, on doit d'abord étendre la définition d'une suite $z \in \ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$ par périodicité à tout intervalle de longueur N_1 par rapport à la variable n_1 et de longueur N_2 par rapport à la variable n_2.

Cette extension est possible si on définit z hors de $\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2}$ comme ceci :

$$\boxed{z(n_1 + j_1 N_1, n_2 + j_2 N_2) = z(n_1, n_2)}, \qquad \forall n_1, n_2, j_1, j_2 \in \mathbb{Z}. \tag{2.47}$$

C'est aussi utile d'introduire l'opérateur de translation dans le cas bidimensionnel.

Déf. 2.9.2 *Soit $z \in \ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$ étendue par périodicité comme dans la formule (2.47) et $k_1, k_2 \in \mathbb{Z}$. L'opérateur de translation sur $\ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$ est défini par :*

$$\begin{array}{rccc} R_{k_1,k_2}: & \ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2}) & \longrightarrow & \ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2}) \\ & z & \longmapsto & R_{k_1,k_2} z, \end{array}$$

$$(R_{k_1,k_2} z)(n_1, n_2) = z(n_1 - k_1, n_2 - k_2).$$

Propriétés de la DFT 2D

Soit $z \in \ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$ étendue par périodicité comme dans la formule (2.47), alors, pour tout $n_1, n_2, m_1, m_2 \in \mathbb{Z}$:

— **Périodicité** de $\hat{z}$ et $\check{z}$:

$$\boxed{\hat{z}(m_1, m_2) = \hat{z}(m_1 + N_1, m_2) = \hat{z}(m_1, m_2 + N_2) = \hat{z}(m_1 + N_1, m_2 + N_2)}$$

et

$$\boxed{\check{z}(n_1, n_2) = \check{z}(n_1 + N_1, n_2) = \check{z}(n_1, n_2 + N_2) = \check{z}(n_1 + N_1, n_2 + N_2)}.$$

— DFT 2D et **translation** :

$$\boxed{\widehat{R_{k_1,k_2} z}(m_1, m_2) = e^{-2\pi i \left(\frac{m_1 k_1}{N_1} + \frac{m_2 k_2}{N_2}\right)} \hat{z}(m_1, m_2)} \qquad \forall k_1, k_2 \in \mathbb{Z},$$

i.e. si on définit la suite $\omega^{k_1,k_2}_{N_1,N_2} \in \ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$, $\omega^{k_1,k_2}_{N_1,N_2}(m_1, m_2) = e^{-2\pi i \left(\frac{m_1 k_1}{N_1} + \frac{m_2 k_2}{N_2}\right)}$ $\forall m_1, m_2 \in \mathbb{Z}$, alors :

$$\boxed{\text{DFT 2D} \circ R_{k_1,k_2} = M_{\omega^{k_1,k_2}_{N_1,N_2}} \circ \text{DFT 2D}},$$

où $M_{\omega^{k_1,k_2}_{N_1,N_2}}$ est l'opérateur de multiplication par $\omega^{k_1,k_2}_{N_1,N_2}$ dans $\ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$. Si on permute le sens de la composition on obtient :

$$\boxed{(R_{k_1,k_2} \hat{z})(m_1, m_2) = \hat{z}(m_1 - k_1, m_2 - k_2) = \text{DFT 2D}\left(e^{2\pi i \left(\frac{m_1 k_1}{N_1} + \frac{m_2 k_2}{N_2}\right)} z\right)(m_1, m_2)},$$

i.e.

$$\boxed{R_{k_1,k_2} \circ \text{DFT 2D} = \text{DFT 2D} \circ M_{\left(\omega^{k_1,k_2}_{N_1,N_2}\right)^*}}, \qquad \forall k_1, k_2 \in \mathbb{Z}.$$

On peut résumer les propriétés qu'on vient d'analyser avec les couples de Fourier suivants :

Représentation originale	Espace de Fourier
$z(n_1 - k_1, n_2 - k_2)$	$e^{-2\pi i \left(\frac{m_1 k_1}{N_1} + \frac{m_2 k_2}{N_2}\right)} \hat{z}(m_1, m_2)$
$e^{2\pi i \left(\frac{n_1 k_1}{N_1} + \frac{n_2 k_2}{N_2}\right)} z(n_1, n_2)$	$\hat{z}(m_1 - k_1, m_2 - k_2)$

Comme dans le cas de la DFT 1D, si on considère $k_1 = \frac{N_1}{2}$ et $k_2 = \frac{N_2}{2}$, alors $(-1)^{n_1+n_2} z(n_1, n_2)$ et $\hat{z}(m_1 - \frac{N_1}{2}, m_2 - \frac{N_2}{2})$. Cette transformation est utilisée pour la visualisation centrée du spectre de z.
De plus, encore comme dans le cas mono-dimensionnel, le spectre d'amplitude d'un signal bidimensionnel $z(n_1, n_2)$ et d'une translation quelconque $z(n_1 - k_1, n_2 - k_2)$ est exactement le même, car $\left| e^{-2\pi i \left(\frac{m_1 k_1}{N_1} + \frac{m_2 k_2}{N_2} \right)} \right| = 1$.
Donc, *le spectre d'amplitude* permet de connaître le contenu fréquentiel du signal, mais il *ne permet pas de savoir où les fréquences sont localisées.*

— DFT 2D et **conjugaison** :

$$\boxed{\widehat{\overline{z}}(m_1, m_2) = \overline{\hat{z}(-m_1, -m_2)} = \overline{\hat{z}(N_1 - m_1, N_2 - m_2)}}.$$

— DFT 2D et **convolution** :

$$\boxed{\widehat{(z * w)}(m_1, m_2) = \hat{z}(m_1, m_2)\hat{w}(m_1, m_2)},$$

où la convolution bidimensionnelle est définie comme ceci :

$$\begin{array}{rccc} T_z : & \ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2}) & \longrightarrow & \ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2}) \\ & w & \longmapsto & T_z w = z * w, \end{array}$$

$$\boxed{(z * w)(n_1, n_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} z(n_1 - k_1, n_2 - k_2) w(k_1, k_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} z(n_1, n_2) w(n_1 - k_1, n_2 - k_2)}.$$

2.9.3 La DFT 2D et les opérateurs stationnaires

La DFT 2D a les mêmes propriétés que la DFT 1D relativement aux opérateurs stationnaires.

Rigoureusement, un opérateur $T : \ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2}) \to \ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$ est stationnaire si :

$$T \circ R_{k_1,k_2} = R_{k_1,k_2} \circ T, \quad \forall k_1, k_2 \in \mathbb{Z}.$$

Pratiquement, un opérateur $T : \ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2}) \to \ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$ est stationnaire si son action sur une suite $z(n_1, n_2)$ est indépendante de la position des paramètres (n_1, n_2).

Si z est une image numérique, alors **un opérateur stationnaire est une transformation définie indépendamment de la position d'un pixel dans le support spatial de l'image**. Cela ne veut pas dire que le *résultat* de l'application de l'opérateur stationnaire ne change pas par rapport à la position du pixel, mais que la définition de l'opérateur ne change pas.

Un exemple d'opérateur stationnaire est donné par la transformation qui remplace l'intensité d'un pixel par sa moyenne locale, i.e. la moyenne calculé dans un voisinage d'une taille fixée du pixel.

Comme dans le cas mono-dimensionnel, les opérateurs stationnaires sur $\ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$ peuvent être caractérisés comme des opérateurs de convolution ou comme des multiplicateurs de Fourier.

Avant d'énoncer le théorème qui formalise cette relation, on définit le multiplicateur de Fourier, l'impulsion unitaire et la réponse impulsionnelle dans le cas bidimensionnel.

Déf. 2.9.3 *Soit $w \in \ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$, fixé, le multiplicateur de Fourier associé à w est défini comme ceci :*

$$\begin{array}{rccc} T_{(w)} : & \ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2}) & \longrightarrow & \ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2}) \\ & z & \longmapsto & T_{(w)}z = \widetilde{w \cdot \hat{z}} \ . \end{array}$$

Déf. 2.9.4 *L'impulsion unitaire δ dans $\ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$ est le premier vecteur de sa base canonique : $\delta = e_{0,0}$.*

Étant donné un opérateur linéaire T sur $\ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$, on appelle réponse impulsionnelle la suite $h = T\delta \in \ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$.

Théorème 2.9.2 *Soit $T : \ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2}) \longrightarrow \ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$ un opérateur linéaire, alors les conditions suivantes sont équivalentes :*

(1) *T est stationnaire ;*

(2) *T est l'opérateur de convolution avec la réponse impulsionnelle $h = T\delta$:*

$$Tz = T_h z = h * z = z * h \qquad \forall z \in \ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2}) \ ;$$

(3) *T est le multiplicateur de Fourier associé à $\hat{h}$:*

$$Tz = T_{(\hat{h})} z = \widetilde{\hat{h} \cdot \hat{z}} \qquad \in \ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2}) \ ;$$

(4) *T est diagonalisable, ses vecteurs propres sont la base orthogonale de Fourier F_{m_1,m_2} de $\ell^2(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2})$, et ses valeurs propres sont les composantes de $\hat{h}$.*

Observation : on peut étendre aussi le résultat sur les matrices circulantes, mais leur définition dans le cas bidimensionnel est plus difficile.

2.9.4 Les opérateurs gradient et Laplacien et leurs action sur les images numériques

Si on répète l'analyse des opérateurs de dérivation discrète qu'on a fait dans la section 2.8.6 dans le cas 2D, la dérivée première devient le **gradient**, i.e. $\vec{\nabla} = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y})$, et la dérivée seconde devient le **Laplacien**, i.e. $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$.

Le gradient est utilisé quand on veut détecteur les bords d'une image selon une direction particulière. Si on veut une détection de bords isotrope, i.e. uniforme par rapport aux directions, alors on utilise le Laplacien, qui est même plus efficace que le gradient dans la tâche d'intensification des détails fins, comme on l'a vu dans le cas 1D.

Même dans le cas 2D, les opérateurs différentiels ci-dessus annulent la moyenne d'une images, ce qui explique pourquoi leur sortie est une image noire partout, sauf à côté des bords, comme montré en Fig. 2.6.

FIGURE 2.6 – Première ligne de gauche à droite : image originale et image filtrée avec le Laplacien. Deuxième ligne de gauche à droite : image filtrée avec le gradient dans la direction verticale et horizontale, respectivement.

2.9.5 Visualisation du spectre d'amplitude en 2D

La visualisation du spectre d'un signal 2D nécessite une procédure de centralisation pour les mêmes raisons qu'on a discuté dans le cas 1D. Ceci peut être effectué via l'équivalent 2D de la formule (2.32), i.e. en considérant $(-1)^{n_1+n_2} z(n_1, n_2)$ au lieu de $z(n_1, n_2)$, comme vu dans la section 2.9.2.

Il faut observer, en particulier, que la symétrie mono-dimensionnelle de la DFT 1D par rapport aux fréquences $m \in \{0, 1, \ldots, N/2\}$ et $m \in \{N/2+1, N/2+2, \ldots, N-1\}$ devient une *symétrie miroir* 2D pour la DFT 2D.

Dans la figure 2.7 on montre trois images numériques en niveau de gris et leur spectre d'amplitude. Les points les plus lumineux correspondent à des valeurs des modules de coefficients de Fourier grandes et, vice-versa, les points les plus sombres à des valeurs petites.

Il y a plusieurs caractéristiques à observer :

- La **symétrie** du spectre : le contenu fréquentiel est répété dans les quadrants par symétrie miroir ;
- Les points les plus lumineux sont localisés vers le centre du spectre : ceci est dû au fait que les spectres qu'on montre sont centrés, donc la fréquence centrale a pour coordonnes $(m_1, m_2) = (0, 0)$ et $|\hat{z}(0, 0)| = N_1 N_2 \langle z \rangle$, i.e. $N_1 N_2$ fois la valeur moyenne de l'image. *Ceci explique pourquoi, pour visualiser le spectre, on doit appliquer une fonction compressive, par exemple le logarithme* : les valeurs de $|\hat{z}(0, 0)|$ sont tellement plus élevées que les autres, qu'il faut comprimer le rang de variabilité avec une fonction compressive comme le logarithme !
- Dès qu'on s'éloigne du centre, le spectre montre l'amplitude des coefficients qui correspondent aux fréquences les plus élevées, jusqu'aux fréquences maximales $(N_1/2, N_2/2)$, si N_1, N_2 sont paires, ou leurs parties entières $([N_1/2], [N_2/2])$ si N_1, N_2 sont impaires. L'image avec le contenu fréquentiel le plus élevé est l'image du mandrill, en fait, on voit que son spectre est le plus étendu parmi les trois images.
  En particulier, on observe des valeurs très intenses vers les bords, qui représentent des fréquences très élevées : elles correspondent aux fréquences des détails très fins du poil à côté des yeux du mandrill.
- Comme m_1 et m_2 représentent les fréquences verticales et horizontales, les bords (« edges » en anglais) verticaux et horizontaux des images produisent des coefficients de Fourier localisés sur les axes correspondants. C'est pour cela que la première image, qui a des forts gradients d'intensité verticaux entre les rochers et le ciel, a un spectre avec une forte prédominance de coefficients de Fourier intenses sur l'axe vertical.
  Dans l'image de la deuxième ligne (« Lena », une des images les plus utilisées en traitement d'images), on a des détails fins dans le chapeau, à 45° et à -45°. De manière cohérente, on trouve dans son spectre des structures diagonales.
- L'analyse du spectre qu'on vient de faire montre qu'on peut comprendre l'existence de structures géométriques dans une image en regardant son spectre de Fourier, mais on ne peut pas dire où ces structures sont situées dans l'image.

FIGURE 2.7 – Colonne de gauche : images originales. Colonne de droite : spectres d'amplitude centrés des images de gauches, visualisés avec une échelle logarithmique.

2.9.6 Un exemple remarquable de filtrage d'une image numérique dans l'espace de Fourier : le floutage

Le théorème 2.9.2 dit que tous les opérateurs stationnaires T qui agissent sur des images (interprétées comme des suites 2D finies) sont des convolutions « cachées » (entre l'image et la réponse impulsionnelle $h = T\delta$). De plus, ces convolutions peuvent être représentées comme des multiplicateurs de Fourier (multiplication dans l'espace de Fourier entre $\hat{h}$ et la DFT 2D de l'image).

Selon la suite h avec laquelle on fait la convolution, on obtient des résultats différents. C'est souvent beaucoup plus simple d'interpréter l'effet d'une convolution en considérant le multiplicateur de Fourier associé.

En particulier, on veut comprendre ce que veut dire d'opérer une *convolution avec une Gaussienne discrète*, qu'on écrit $h(n_1, n_2)$.

Comme on le verra dans le chapitre 6, la transformée de Fourier d'une Gaussienne d'écart-type σ est encore une Gaussienne, mais avec un écart-type inversement proportionnel à σ. Donc, on peut essayer de comprendre la signification de la convolution d' une image $z(n_1, n_2)$ avec une Gaussienne $h(n_1, n_2)$ en analysant la multiplication suivante dans l'espace de Fourier : $\hat{z}(m_1, m_2) \cdot \hat{h}(m_1, m_2)$.

Dans la figure 2.8 on peut voir trois images qui correspondent à des Gaussiennes 2D de taille 512×512, l'intensité du pixel de position (n_1, n_2) est $h(n_1, n_2) = \exp\left(-\frac{n_1^2+n_2^2}{2\sigma^2}\right)$ et l'écart type est σ =1,5,10, respectivement.

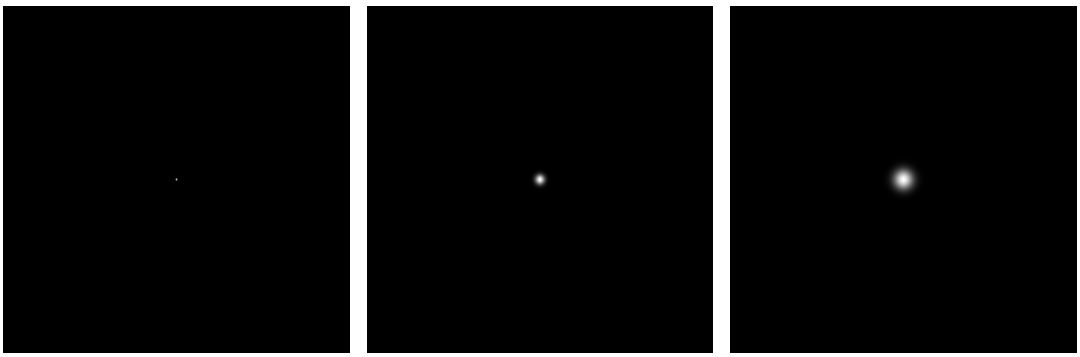


FIGURE 2.8 – De gauche à droite : images Gaussiennes bidimensionnelles avec un écart type 1,5,10.

Comme on l'a dit avant, les DFT 2D de h sont encore des Gaussiennes, mais avec un écart type proportionnel à $1, \frac{1}{5}, \frac{1}{10}$. Bien sûr, $h(0,0) = 1$ et les valeurs de $\hat{h}(m_1, m_2)$ décroissent dès qu'on s'éloigne du centre, donc la multiplication dans l'espace de Fourier $\hat{z}(m_1, m_2) \cdot \hat{h}(m_1, m_2)$ fait décroître l'importance des harmoniques avec $(m_1, m_2) \neq (0, 0)$, qui sont associées aux détailles de l'image. Alors, si on applique la IDFT 2D à $\hat{z}(m_1, m_2) \cdot \hat{h}(m_1, m_2)$ on peut reconstruire une image qui sera plus flou que l'image originelle.

Par conséquent, opérer la convolution avec une Gaussienne correspond, dans le traitement d'images, à une opération de floutage (« blurring » en anglais), comme on peut le voir dans la figure 2.9.

Le floutage peut être utile, par exemple, quand l'image originale est bruitée : le floutage rend le bruit moins percevable (mais aussi les bords moins nets).
La figure 2.10 montre une version continue du filtres fréquentiel de floutage.

FIGURE 2.9 – De gauche à droite : floutage de l'image de Lena par multiplication dans l'espace de Fourier avec les DFT des Gaussiennes d'écart type 1,5,10. Il faut observer que, comme la DFT d'une Gaussienne a un écart-type inversement proportionnel à l'écart-type originale, la DFT de la Gaussienne d'écart-type 10 a un écart-type petit et donc elle tend vers 0 rapidement. Donc, quand on fait le produit de la DFT de la Gaussienne d'écart-type 10 avec la DFT de l'image, on réduit énormément les détails de l'image.

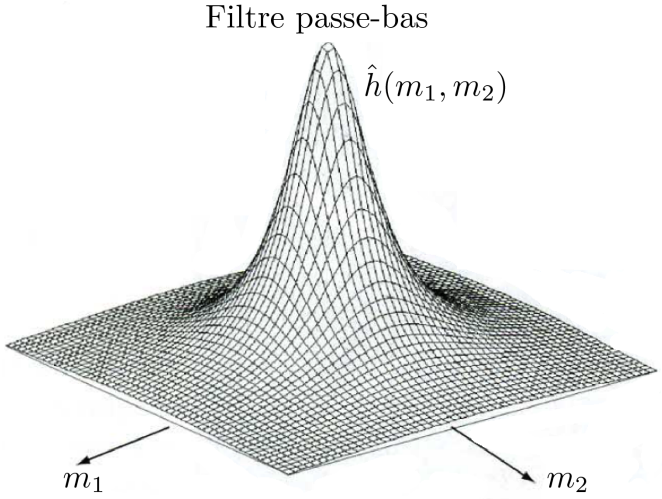


FIGURE 2.10 – Filtre de floutage, passe-bas dans le domaine fréquentiel.

Observation importante : même si opérer la convolution avec une Gaussienne amène à l'effet de floutage, *il ne faut pas penser que la convolution soit toujours liée à une opération de floutage.* En fait, on a vu que la convolution, comme le multiplicateur de Fourier, est le prototype des opérateurs stationnaires, qui peuvent flouter ou rehausser les fréquences d'un signal.

2.10 Résumé du chapitre 2

- Nous avons considéré l'espace $\ell^2(\mathbb{Z}_N)$ des suites $\mathbb{N}$-périodiques à valeurs complexes, qui est isomorphe à $\mathbb{C}^N$.

- On a introduit une base particulière de cet espace : celle des exponentiels complexes donnés par les puissances consécutives des racines complexes N-ièmes de l'unité. Cette

base permet de construire (via une simple normalisation) la base de Fourier de $\ell^2(\mathbb{Z}_N)$. Nous avons interprété ses éléments comme des ondes dites « harmoniques » qui oscillent avec des fréquences multiples d'une fréquence fondamentale.

- Les coefficients de Fourier d'un élément de $\ell^2(\mathbb{Z}_N)$ sont ses composantes par rapport à la base de Fourier. Comme ces coefficients sont complexes, il faut considérer leur module pour connaître l'importance d'une harmonique relative à una certaine fréquence dans la reconstruction (ou synthèse) de l'élément lui-même. L'ensemble des modules des coefficients de Fourier est dite spectre d'un élément de $\ell^2(\mathbb{Z}_N)$.

- La transformée de Fourier discrète (DFT) est l'endomorphisme de $\ell^2(\mathbb{Z}_N)$ qui envoie un élément de $\ell^2(\mathbb{Z}_N)$ dans la suite de ses coefficients de Fourier. La DFT est un isomorphisme, et son inverse est écrite IDFT.

- Il est possible d'associer à la DFT une matrice, dite matrice de Sylvester, elle est une matrice de Vandermonde, i.e. toutes les lignes et les colonnes sont données par des progressions géométriques.

- Nous avons donné une interprétation des concepts ci-dessus dans la théorie des signaux, en particulier, on a souligné le fait que la fréquence d'oscillation harmonique la plus élevée dans un signal discret donné par N échantillons est $N/2$ (ou de sa partie entière si N n'est pas pair), dit fréquence de Nyquist.

- La DFT transforme la translation en une multiplication par un facteur de phase, i.e. un exponentiel complexe de module unitaire, ceci implique que le spectre d'un signal est invariant par translation.

- La convolution est transformée dans le produit ponctuel par la DFT, ce qui permet de donner une expression diagonale de l'opérateur de convolution dans l'espace de Fourier.

- Pour terminer, nous avons vu que la DFT permet de diagonaliser les opérateurs stationnaires, i.e. ceux qui commutent avec les opérateurs de translation. En fait, via le théorème 2.8.2, il est possible de caractériser complètement un opérateur stationnaire comme une convolution ou comme un multiplicateur de Fourier et de déterminer ses valeurs propres.

Chapitre 3

Rappel sur la théorie de la mesure et de l'intégration à la Lebesgue

Dans ce chapitre on va rappeler les informations strictement essentielles de la théorie de la mesure et de l'intégration (qui sont le sujet du cours d'intégration) avec le but d'avoir une notation et un langage compréhensible et commun.

La rédaction est volontairement très synthétique, elle doit être accompagnée par un véritable support pédagogique dédié à la théorie de la mesure et de l'intégration pour les élèves qui n'ont pas encore eu la possibilité d'étudier ce champ extrêmement important des mathématiques.

Deux ouvrages conseillés sont ceux de M.Briane et G.Pages [5] et de Bartle [2].

3.1 Riemann vs Lebesgue

La différence principale de l'approche à la Riemann et à la Lebesgue est montrée dans la Figure 3.1.

Le principe fondateur de l'intégration à la Riemann est d'approcher l'aire de la surface comprise entre l'axe des abscisses et la courbe du graphe de la fonction f à l'aide de petits rectangles $[a_{i-1}, a_i] \times [0, \Phi_i]$ basés sur l'axe des abscisses et dont la hauteur Φ_i est proche de la hauteur moyenne de la fonction f sur $[a_{i-1}, a_i]$.

L'idée novatrice apportée par la théorie de l'intégration de Lebesgue est, à l'inverse, de commencer par découper en petits intervalles $[b_{j-1}, b_j]$ l'axe des ordonnées, puis d'approcher

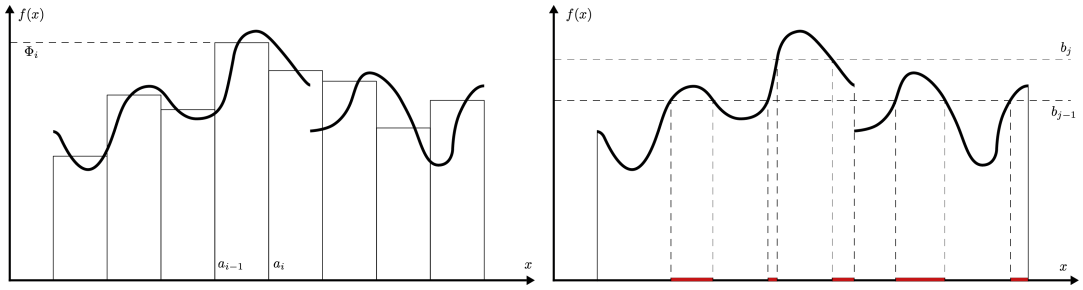


FIGURE 3.1 – Intégration à la Riemann (gauche) et à la Lebesgue (droite).

la surface située sous le graphe de f par

$$\int_a^b f \approx \sum_j \frac{b_{j-1}+b_j}{2} \cdot \text{longueur}\left(\{x \ : \ b_{j-1} \leqslant f(x) \leqslant b_j\}\right).$$

La principale difficulté réside dans le fait que les ensembles

$$E_j = \{x \in [a,b] \ : \ b_{j-1} \leqslant f(x) \leqslant b_j\}$$

en rouge dans l'image de droite de la figure 3.1, ne sont généralement pas des intervalles et que leur associer une longueur, ou, plus généralement, une mesure, est délicat voir, dans certains cas, impossible.

Il est cette considération qui montre l'exigence de construire une théorie de la mesure des ensembles pour pouvoir développer une théorie de l'intégration avec la stratégie décrite ci-dessus. Cette route est beaucoup plus longue et compliquée par rapport à celle tracée par Riemann, mais la contrepartie est que la définition d'intégrale à la Lebesgue est bien plus générale et les propriétés que l'on peut prouver sont incomparablement plus puissantes.

3.2 Tribu (σ-algèbre), espace mesurable, mesure et espace mesuré

Pour définir l'intégrale à la Lebesgue il faut donc, tout d'abord, définir les ensembles et les fonction qu'on peut mesurer. D'après l'étude faite au début du XXème siècle, on est arrivé à la formalisation contenue dans les définitions et résultats suivants.

Soit X un ensemble. On appelle **tribu** (ou **σ-algèbre**) sur X, un ensemble $\mathcal{A}$ de parties de X, i.e. $\mathcal{A} \subseteq \mathcal{P}(X)$, qui vérifie :

— $\varnothing, X \in \mathcal{A}$;

— $\mathcal{A}$ est stable par complémentaire : $E \in \mathcal{A} \Rightarrow E^c \in \mathcal{A}$;

— $\mathcal{A}$ est stable par union dénombrable : $(E_n)_{n\in\mathbb{N}} \in \mathcal{A} \Rightarrow \bigcup_{n\in\mathbb{N}} E \in \mathcal{A}$.

La définition implique que $\mathcal{A}$ est stable par intersection dénombrable.

Exemples triviales :

— $\mathcal{A} = \mathcal{P}(X)$: tribu des parties de X ;

— $\mathcal{A} = \{\varnothing, X\}$: tribu grossière.

Dans les livres de théorie de la mesure on démontre que les propriétés qui définissent une tribu sont nécessaires et suffisantes pour pouvoir « mesurer » les ensembles contenus dans la tribu elle-même, dans un sens qu'on va définir bientôt. C'est pour ça qu'on appelle le couple $(X, \mathcal{A})$ un **espace mesurable** et les éléments de $\mathcal{A}$ **ensembles mesurables**.

Pour donner un premier exemple significatif d'espace mesurable on doit introduire d'abord le concept de *relation d'ordre entre tribus* : si tout élément de la tribu $\mathcal{A}_1$ est contenu dans la tribu $\mathcal{A}_2$, alors on dit que $\mathcal{A}_1$ est plus petite que $\mathcal{A}_2$ et on écrit $\mathcal{A}_1 \subset \mathcal{A}_2$. Ce concept est utilisé pour définir la plus petite tribu engendrée par une collection de parties d'un ensemble : soit $S \subset \mathcal{P}(X)$, alors on appelle **tribu engendrée** par S l'intersection de toutes les tribus qui contiennent S.

Le cas très important d'un ensemble X doté d'une topologie mérite une discussion privilégiée.

L'existence d'une topologie permet de définir le concept de partie ouverte de X. Soit $\tau \subseteq \mathcal{P}(X)$ l'ensemble des parties ouvertes de X, bien évidemment τ n'est pas une tribu car le complémentaire d'un ouvert est fermé. Néanmoins, on peut considérer la tribu engendrée par τ, qu'on appelle **tribu de Borel**[1] et qu'on dénote avec $\mathcal{B}(X)$. Chaque élément de cette tribu (qui, on le rappelle encore une fois, est un sous-ensemble de X !) est appelé un **ensemble Borélien**.
Une fois qu'on a un espace mesurable $(X, \mathcal{A})$, on peut définir le concept de mesure positive, ou tout simplement **mesure** μ comme une fonction $\mu : \mathcal{A} \to [0, +\infty]$ telle que :

— $\mu(\varnothing) = 0$;
— μ est σ-additive (ou vérifie l'additivité dénombrable) : si $(E_n)_{n\in\mathbb{N}}$ est une famille dénombrable d'éléments de $\mathcal{A}$ deux à deux disjoints, alors :

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n).$$

La triplette $(X, \mathcal{A}, \mu)$ est dite un **espace mesuré**. Quand la tribu et la mesure sont clairement spécifiées, elles sont souvent admises.

Un exemple très simple mais déjà significatif de mesure est celle de Dirac dans l'espace mesurable $(\mathbb{R}, \mathcal{B}(\tau))$, i.e. $\mathbb{R}$ avec la tribu des Boréliens. La **mesure de Dirac** basée en $x_0 \in \mathbb{R}$ est définie par : $\delta_{x_0} : \mathcal{B}(\tau) \to \{0, 1\}$,

$$\delta_{x_0}(E) = \begin{cases} 1 \text{ si } x_0 \in E \\ 0 \text{ si } x_0 \notin E, \end{cases} \qquad \forall E \in \mathcal{B}(\tau).$$

Comme $\mathbb{R}$ lui-même est un élément de $\mathcal{B}(\tau)$, $\delta_{x_0}(\mathbb{R}) = 1$, donc la mesure de Dirac de $\mathbb{R}$ est 1 indépendamment du point de base. On a donc un exemple de **mesure finie**, i.e. telle que la mesure de l'espace mesurable est $< +\infty$.

En général, les mesures ne sont pas finies, mais plutôt σ-finies. Étant donné un espace mesuré $(X, \mathcal{A}, \mu)$, on dit que μ est une **mesure σ-finie** si X peut être écrit comme l'union dénombrable de parties $(E_n)_{n\in\mathbb{N}} \subset X$ qui ont toutes mesure finie via μ, i.e.

$$X = \bigcup_{n\in\mathbb{N}} E_n, \quad \mu(E_n) < +\infty \ \forall n \in \mathbb{N}.$$

Il y a plusieurs techniques pour construire une mesure, néanmoins elles ne sont pas ni simples ni rapides à décrire, on invite le lecteur à consulter, par exemple, le livre mentionné dans l'avant propos ou n'importe quel autre livre de théorie de la mesure.

1. Une description *explicite* de cette tribu est impossible.

3.3 Fonctions mesurables et propriétés possédées presque partout (p.p)

C'est maintenant le moment d'introduire les *morphismes des espaces mesurables*, i.e. les applications entre espaces mesurables qui en préservent la propriété essentielle : la mesurabilité des ensembles.

Soient $(X_1, \mathcal{A}_1)$, $(X_2, \mathcal{A}_2)$ deux espaces mesurables et soit $f : X_1 \to X_2$ une fonction quelconque. On dit que f est une **fonction mesurable** (par rapport aux tribus choisies) si la tribu image réciproque par f de la tribu $\mathcal{A}_2$ est incluse dans $\mathcal{A}_1$, i.e. [2]

$$\forall E \in \mathcal{A}_2 \ \Rightarrow \ f^{-1}(E) \in \mathcal{A}_1,$$

ce qui est équivalent, par définition, à dire que l'image réciproque d'une partie mesurable de X_2 (par rapport à $\mathcal{A}_2$) est une partie mesurable de X_1 (par rapport à $\mathcal{A}_1$).

Exemples :

- Les fonctions continues entre deux espaces topologiques sont évidemment mesurables par rapport aux tribus de Boréliens ;
- Quand on considère les fonctions à valeur réelles, i.e. $f : X \to \mathbb{R}$, où $(X, \mathcal{A})$ est un espace mesurable, alors on fixe la tribu de Borel sur $\mathbb{R}$ et on définit la mesurabilité de f par rapport à ce choix ;
- Une fonction à valeur complexes $f : X \to \mathbb{C}$ est mesurable si le sont sa partie réelle et sa partie imaginaire.

On vient maintenant à un concept cruciale, celui des propriétés définies presque partout. On dit qu'une fonction f définie sur un espace mesuré $(X, \mathcal{A}, \mu)$ possède une propriété **presque partout** (et on écrit **p.p**) si f a cette propriété sur $X \backslash E$, où $E \in \mathcal{A}$ a mesure nulle : $\mu(E) = 0$.

Exemples :

- f, g : fonctions mesurables définies sur $(X, \mathcal{A}, \mu)$, alors $f = g$ p.p si $f(x) = g(x)$ $\forall x \in U \in \mathcal{A}$ et $\mu(X \backslash U) = 0$;
- f est la limite ponctuelle p.p de la suite $(f_n)_{n\in\mathbb{N}}$ si $\lim\limits_{n\to+\infty} f_n(x) = f(x)$ $\forall x \in U \in \mathcal{A}$ et $\mu(X \backslash U) = 0$.

3.4 Fonctions intégrables, intégrales à la Lebesgue

Une fois donné un espace mesuré $(X, \mathcal{A}, \mu)$, la construction de l'intégrale d'une fonction mesurable à valeurs réelles ou complexes est plutôt rapide. On commence par considérer une fonction particulière, la **fonction indicatrice (ou caractéristique)** d'un ensemble $E \in \mathcal{A}$: $\chi_E : X \to \{0, 1\}$,

$$\chi_E(x) = \begin{cases} 1 \text{ si } x \in E \\ 0 \text{ si } x \notin E, \end{cases}$$

une notation équivalente est $\mathbb{1}_E$.

2. Il faut noter la similarité entre cette définition et celle d'une fonction continue dans le sens topologique du terme.

Les fonctions caractéristiques sont utilisées pour définir les **fonctions simples** ou **fonctions étagées** via combinaison linéaires. Plus précisément, soit $(E_k)_{k=1}^n$ une partition finie et disjointe de X, i.e. $E_k \cap E_{k'} = \varnothing \ \forall k \neq k'$ et

$$\bigcup_{k=1}^{n} E_k = X,$$

alors une fonction simple $s : X \to \mathbb{R}$ est définie comme ça :

$$s = \sum_{k=1}^{n} c_k \chi_{E_k},$$

$s(x) = c_k \ \forall x \in E_k$, donc s prend seulement un nombre fini de valeurs, si X est un sous-ensemble de $\mathbb{R}$, alors s est une fonction constantes par morceaux.

L'intégrale à la Lebesgue d'une fonction simple a la définition naturelle suivante :

$$\int_X s d\mu = \sum_{k=1}^{n} c_k \mu(E_k),$$

à noter que sans la définition de la mesure des ensembles E_k, l'intégrale de s ne serait pas bien définie.

L'importance des fonctions simples est exprimée dans le théorème suivant.

Théorème 3.4.1 *Soit $(X, \mathcal{A}, \mu)$ un espace mesuré et $f : X \to [0, +\infty]$ une fonction mesurable et non-négative, alors f on peut approcher f du bas avec une suite de fonctions simples, i.e. $\exists$ $(s_n)_{n\in\mathbb{N}}$, s_n fonction simple, telle que $(s_n)_{n\in\mathbb{N}} \nearrow f$, i.e.*

1. $0 \leqslant s_0(x) \leqslant s_1(x) \leqslant \ldots \leqslant s_n(x) \leqslant \ldots \leqslant f(x) \ \forall x \in X$;
2. $\lim_{n\to+\infty} s_n(x) = f(x) \ \forall x \in X$ *(limite ponctuel). Si f est borné, alors la convergence de la suite $(s_n)_{n\in\mathbb{N}}$ vers f est uniforme.*

La preuve (très élégante) du théorème est constructive, on y prouve que la suite de fonctions simples est donnée par :

$$s_n(x) = \begin{cases} \frac{k}{2^n} \text{ si } \frac{k}{2^n} \leqslant f(x) \leqslant \frac{k+1}{2^n}, \text{ pour } k = 0, 1, \ldots, 2^n - 1 \\ 2^n \text{ si } 2^n \leqslant f(x). \end{cases}$$

Ce théorème fondamentale nous permet de définir l'**intégrale d'une fonction mesurable non négative** $f : X \to [0, +\infty]$ comme cela :

$$\boxed{\int_X f d\mu = \sup_{0\leqslant s\leqslant f} \int_X s \, d\mu}, \qquad s \text{ simple},$$

f est dite **intégrable** (à la Lebesgue) si $\int_X f d\mu < +\infty$.

Si $f : X \to \overline{\mathbb{R}}$ est mesurable, alors on peut définir son intégrale en considérant sa partie positive :

$$f_+(x) = \begin{cases} f(x) \text{ si } f(x) \geqslant 0 \\ 0 \text{ si } f(x) < 0 \end{cases}$$

et sa partie négative :

$$f_-(x) = \begin{cases} -f(x) \text{ si } f(x) \leqslant 0 \\ 0 \text{ si } f(x) > 0 \end{cases}$$

en fait les deux sont des fonctions à valeurs non négatives pour lesquelles on sait définir l'intégrale. Comme $f = f_+ - f_-$, si f_+ et f_- sont intégrables, on peut définir l'**intégrale d'une fonction mesurable à valeurs réels étendues** comme ça :

$$\boxed{\int_X f d\mu = \int_X f_+ d\mu - \int_X f_- d\mu}.$$

Pour des fonctions $f : X \to \mathbb{C}$ mesurables on utilise la même stratégie, mais avec la partie positive et la partie négative de sa partie réelle $\mathfrak{Re}(f)$ et de sa partie imaginaire $\mathfrak{Im}(f)$, l'intégrale est alors définie comme cela :

$$\int_X f d\mu = \int_X \mathfrak{Re}(f) d\mu + i \int_X \mathfrak{Im}(f) d\mu.$$

Une condition nécessaire et suffisante pour l'intégrabilité d'une fonction à valeurs réelles ou complexes est son intégrabilité absolue :

$$\boxed{\int_X f d\mu < +\infty \iff \int_X |f| d\mu < +\infty}.$$

3.5 La caractérisation de la mesure de Lebesgue sur $\mathbb{R}$ et les ensembles de mesure de Lebesgue nulle

Comme on l'a déjà dit, la construction d'une mesure n'est pas, en général, un processus simple. Néanmoins, vu l'importance de la mesure de Lebesgue sur $\mathbb{R}$, on veut au moins résumer les caractéristiques de cette mesure. Heureusement, il existe un théorème qui donne la caractérisation de la mesure de Lebesgue via certaines propriétés qu'on doit décrire avant de citer ce résultat.

- **Mesure de Borel** : soit X un espace topologique, considérons l'espace mesurable $(X, \mathcal{B}(X))$, où $\mathcal{B}(X)$ est la tribu des Boréliens. Alors on dit qu'une mesure μ définie sur cet espace est une **mesure de Borel** si elle associe un nombre fini à toute partie compacte K de X ;
- **Mesure de Borel régulière** : (celle-ci est une propriété plus technique qu'il faut, néanmoins, citer pour pouvoir énoncer des résultats précisément), une mesure de Borel est dite régulière si, pour tout ensemble Borélien $E \in \mathcal{B}(X)$ il vaut que :
  1. $\mu(E) = \sup\{\mu(K),\ K \subset E,\ K \text{ compact}\}$;
  2. $\mu(E) = \inf\{\mu(O),\ E \subset O,\ O \text{ ouvert}\}$.

Au lieu de considérer un espace mesuré quelconque, on va mettre notre attention sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$, alors on dit que μ est une **mesure invariante par translations** si

$$\mu(E + a) = \mu(E),$$

pour tout ensemble Borélien $E \in \mathcal{B}(\mathbb{R})$ et tout $a \in \mathbb{R}$, où $E + a = \{x \in \mathbb{R} \ :\ x = e + a,\ \text{où } e \in E\}$.

On peut maintenant détourner la construction de la mesure de Lebesgue sur $\mathbb{R}$, qu'on écrit avec le symbole m, en citant le théorème qui la caractérise.

Théorème 3.5.1 *Si une mesure sur* $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$ *a les propriétés suivantes :*

1. μ *est une mesure de Borel régulière ;*
2. μ *est une invariante par translation ;*
3. μ *est normalisée, i.e.* $\mu[0,1] = 1$ *;*

alors μ *est la mesure de Lebesgue* m*.*

On peut donc dire que **la mesure de Lebesgue** sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ **est une mesure de Borel régulière invariante par translation et normalisée**, ce qui implique aussi que $m[a,b] = b - a$.

Une autre conséquence de ce théorème est que *la mesure de Lebesgue est* σ*-finie* : en fait, $\mathbb{R}$ peut être recouvert avec une partition d'intervalles compacts $[-n, n]$ avec $n \in \mathbb{N}$, tous de mesure finie ($\mu[-n, n] = 2n$).

La mesure de Lebesgue sur $\mathbb{R}$ se généralise sans problème à $\mathbb{R}^n$ et on peut démontrer que, si une fonction est Riemann-intégrable sur $\mathbb{R}^n$, alors elle est aussi Lebesgue-intégrable et les deux intégrales coïncident.

Des exemples importants d'ensembles de mesure de Lebesgue nulle en $\mathbb{R}^n$ sont les hypersurfaces de dimension $n-1$ en $\mathbb{R}^n$, par exemple les courbes en $\mathbb{R}^2$, les surfaces 2D en $\mathbb{R}^3 \ldots$et en $\mathbb{R}$ lui-même ? Comme $\mathbb{R}$ a la dimension (infinie) du continu, les parties de $\mathbb{R}$ qui ont dimension inférieure sont celles fines ou dénombrables, donc, en particulier :

$$m(\mathbb{N}) = m(\mathbb{Z}) = m(\mathbb{Q}) = 0,$$

cela veut dire que si on élimine d'un ensemble mesurable de $\mathbb{R}$, comme par exemple un intervalle $[a,b]$, un nombre infini dénombrable de point, la mesure de l'ensemble ne change pas.

Cette propriété permet de considérer dans la classe des fonctions Lebesgue-intégrables énormément plus de fonctions que dans le cas de l'intégrale à la Riemann. Imaginons le cas d'une fonction continue par morceaux sur un ensemble avec un nombre fini ou dénombrable de discontinuités sous forme de saut, comme c'est par exemple le cas pour une fonction étagée.

L'intégrale à la Riemann de cette fonction n'existe pas, néanmoins l'intégrale à la Lebesgue existe et il est la somme algébrique des intégrales de Riemann sur chaque morceau où la fonction est continue : comme on a seulement un nombre fini ou dénombrable de discontinuités, on peut tout simplement les ignorer, car elles constituent un ensemble de mesure nulle pour la mesure de Lebesgue, et donc n'influent pas sur le résultat final de l'intégration !

Il faut bien retenir que la théorie de l'intégration à la Lebesgue ne donne pas des outils plus avancées pour *le calcul explicite* des intégrales, sauf dans des cas particuliers, mais elle permet de définir les intégrales de fonction beaucoup moins régulières que dans le cas de Riemann.

Ce résultat et les célèbres théorèmes qu'on va rappeler dans la section suivante ont établi la supériorité de la théorie de l'intégration à la Lebesgue sur celle de Riemann.

3.6 Les trois théorèmes d'opération de limite dans la théorie de l'intégration

On résume dans cette section les trois théorèmes les plus importants qui concernent l'opération de limite dans la théorie de l'intégration, qui seront utilisés dans le chapitre 4.

Dans les théorèmes on supposera que $(X, \mathcal{A}, \mu)$ est un espace mesuré quelconque fixé.

Théorème 3.6.1 (Théorème de convergence monotone - Beppo Levi) *Soit $(f_n)_{n\in\mathbb{N}}$, avec $f_n : X \to \mathbb{R}$, une suite* ***monotone croissante*** *de fonctions intégrables. Si la suite des intégrales est* ***bornée****, i.e.*

$$\forall n \in \mathbb{N},\ \exists K \in \mathbb{R} \text{ tel que } \int_X f_n\ d\mu < K,$$

alors $\exists \lim_{n\to+\infty} f_n(x) < +\infty$ p.p. En plus, si on définit la fonction limite $f : X \to \mathbb{R}$ comme ceci

$$f(x) = \begin{cases} \lim_{n\to+\infty} f_n(x) & \text{si la limite est finie} \\ 0 & \text{autrement,} \end{cases}$$

alors f est intégrable et ça vaut l'échange entre la limite et l'intégrale, i.e.

$$\int_X f(x)\ d\mu = \lim_{n\to+\infty} \int_X f_n(x)\ d\mu.$$

Pour introduire le lemme de Fatou il faut rappeler que, donnée une suite quelconque $(x_n)_{n\in\mathbb{N}}$ de nombres réels, on appelle $\liminf$ la **limite inférieure** de la suite, i.e.

$$\liminf_{n\in\mathbb{N}}(x_n) = \inf\{x \in \mathbb{R}\ :\ x \text{ point d'adhérence pour } (x_n)_{n\in\mathbb{N}}\}.$$

Théorème 3.6.2 (Lemme de Fatou) *Soit $(f_n)_{n\in\mathbb{N}}$, avec $f_n : X \to \mathbb{R}$, une suite de fonctions intégrables positives et soit $\liminf_{n\to+\infty} \int_X f_n\ d\mu < +\infty$. Alors, la fonction f définie par*

$$f(x) = \begin{cases} \liminf_{n\to+\infty} f_n(x) & \text{si la limite inférieure est finie} \\ 0 & \text{autrement,} \end{cases}$$

est intégrable et ça vaut l'inégalité suivante

$$\int_X f\ d\mu \leqslant \liminf_{n\to+\infty} \int_X f_n\ d\mu.$$

Théorème 3.6.3 (Théorème de convergence dominée - Henri Lebesgue) *Soit $(f_n)_{n\in\mathbb{N}}$, où $f_n : X \to \mathbb{R}$, une suite de fonctions mesurables et soit $\Phi : X \to \mathbb{R}$ une fonction positive et intégrable telle que :*

$$|f_n| \leqslant \Phi \quad \text{p.p } \forall n \in \mathbb{N}.$$

Si la suite réelle $(f_n(x))_{n\in\mathbb{N}}$ est convergente $\forall x \in X$ et si $f(x) = \lim_{n\to+\infty} f_n(x)$, alors f_n et f sont intégrables et ça vaut l'échange entre la limite et l'intégrale, i.e.

$$\int_X f(x)\ d\mu = \lim_{n\to+\infty} \int_X f_n(x)\ d\mu.$$

3.7 Résumé du chapitre 3

- Nous avons proposé un très court rappel sur la théorie de la mesure et de l'intégration à la Lebesgue en discutant les concepts de tribu, ensemble mesurable, mesure, espace mesuré, fonction mesurable.

- On a mis l'accent sur une tribu particulière : celle des boréliens dans un espace topologique, i.e. la tribu engendrée par les ouverts de cet espace.

- Les propriétés possédées presque partout (p.p) jouent un rôle importante en théorie de la mesure : une propriété est vérifiée p.p si elle est valide sur un sous-ensemble mesurable tel que la mesure de son complémentaire est nulle.

- On a vu qu'il est possible de caractériser la mesure de Lebesgue m sur $\mathbb{R}$ par rapport à tribu des Boreliens comme une mesure de régulière, normalisée et invariante par translation. Des exemples remarquables d'ensemble de mesure de Lebesgue nulle en $\mathbb{R}$ sont les ensembles dénombrables, en particulier $m(\mathbb{N}) = m(\mathbb{Z}) = m(\mathbb{Q}) = 0$.

- Une fois qu'on a à disposition une mesure, la définition de l'intégrale d'une fonction mesurable est très facile et standard. Tout d'abord, on considère les fonctions simples, ou étagées, qui sont des combinaisons linéaires de fonctions indicatrices d'ensembles mesurables. Les fonctions simples approchent par le bas toute fonction mesurable non-négative. Ce résultat-clé permet de définir l'intégrale des fonctions mesurables non-négatives comme le *sup* de l'intégrale des fonctions simples non supérieures à la fonction elle-même. L'extension de cette définition aux fonctions à valeurs réelles quelconque est faite via la partie positive et négative et l'extension aux fonctions à valeurs complexes est faite avec la partie réelle et imaginaire.

- Pour terminer, on a rappelé les trois théorèmes fondamentaux liés à la relation entre limite et intégrale dans la théorie de Lebesgue : le théorème de convergence monotone et dominée (de Levi et Lebesgue, respectivement) et le lemme de Fatou.

Chapitre 4

Espaces de Banach et de Hilbert

Dans ce chapitre on va considérer des espaces normés et avec produit scalaire de dimension infinie. Nous nous intéresserons en particulier aux espaces appelés « complets », pour lesquels on peut formuler de théorèmes très importants, qui ne sont pas disponibles pour des espaces de dimension infinie qui n'ont pas la propriété de complétude.

Avant de commencer cette analyse, il faut souligner que toutes les propriétés qu'on a vu dans les espaces vectoriels de dimension finie avec produit scalaire qui font appel seulement à la nature algébrique du produit scalaire sont valides aussi quand la dimension de l'espace vectoriel est infinie, par exemple :

- Une famille de vecteurs orthogonaux est libre ;
- Si $\langle x, z\rangle = \langle y, z\rangle \ \forall z$, nécessairement les vecteurs x et y coïncident ;
- Le seul vecteur orthogonal à tous les autres vecteurs est le vecteur nul ;
- On peut itérer la procédure d'orthonormalisation de Gram-Schmidt et garantir l'existence d'un système infini de vecteurs orthogonaux entre eux et de norme unitaire à partir d'un ensemble infini quelconque de vecteurs.

La preuve des trois premières propriétés est identique à celle qu'on a vu pour les espaces vectoriels de dimension finie. Pour la dernière propriété il faut faire appel au *lemme de Zorn*.

Par contre, la généralisation de résultats valides pour sommes finies à des séries n'est pas évidente, car, dans ce cas, des arguments topologiques et pas seulement algébriques doivent être pris en considération.

Le chemin pour arriver à la définition et à l'analyse des espace de Hilbert passe, avant tout, par l'analyse de la compatibilité entre la structure linéaire et celle topologique d'un espace vectoriel avec produit scalaire.

4.1 La topologie métrique des espaces vectoriels avec produit scalaire

Comme on l'a vu, chaque espace vectoriel V avec produit scalaire peut être muni d'une norme, canoniquement induite par le produit scalaire. À partir d'une norme, on peut toujours définir, canoniquement, une **distance** ou **métrique** sur V :

$$d(x, y) = \|x - y\| = \sqrt{\langle x - y, x - y\rangle}, \qquad \forall x, y \in V.$$

La fonction d a les propriétés suivantes, $\forall u, x, y \in V$:

1. $d(x,y) \geqslant 0,\ d(x,y) = 0 \Leftrightarrow x = y$
2. $d(x,y) = d(y,x)$ (symétrie)
3. $d(x,y) \leqslant d(x,z) + d(z,y)$ (inégalité triangulaire).

Déf. 4.1.1 (Espace vectoriel métrique) *On appelle espace vectoriel métrique le couple (V,d) donnée par un espace vectoriel V et une fonction, la distance $d : V \times V \to \mathbb{R}_0^+ = [0, +\infty[$, qui satisfait les trois propriétés ci-dessus.*

Donc, un espace vectoriel avec produit scalaire (réel ou complexe) est automatiquement aussi un espace vectoriel normé et un espace vectoriel avec distance. On peut montrer de contre-exemples qui montrent que le contraire, en général, n'est pas vrai.

La présence d'une métrique donne la possibilité d'établir des relations entre points et sous-ensembles d'un espace plus fines que la simple relation d'appartenance. Comme on le sait, cela est la base pour la construction d'une topologie. Nous rappelons ci-dessous des définitions déjà connues avec, surtout, le but de fixer la notation et la nomenclature pour la suite.

Déf. 4.1.2 *Soit $(V, \langle , \rangle)$ un espace vectoriel avec produit scalaire et soit $\| \ \|$ la norme induite canoniquement. Alors :*

— *un **voisinage (ouvert) de rayon** ε **de** $x \in V$ est le sous ensemble de V défini par :*

$$U_\varepsilon(x) = \{y \in V \ : \ \|x - y\| < \varepsilon\},$$

*si $\| \ \|$ est la norme Euclidienne, alors $U_\varepsilon(x)$ est une sphère (ouverte) centrée en x et de rayon ε. Par extension, on appelle souvent $U_\varepsilon(x)$ une **boule ou sphère (ouverte)**, et on écrit $B(x,\varepsilon)$, pour toute norme $\| \ \|$;*

— *un sous-ensemble $O \subseteq V$ est dit **ouvert** si*

$$\forall x \in O \ \exists \varepsilon > 0 \text{ tel que } y \in U_\varepsilon(x) \ \Rightarrow \ y \in O;$$

— *un sous-ensemble $F \subseteq V$ est dit **fermé** si son complémentaire $F^c = V \backslash F$ est ouvert. Nous rappelons que ceci est équivalent à dire que toute suite convergente d'éléments de F a limite en F même ;*

— *l'**adhérence** (ou fermeture) de $E \subset V$ est $\overline{E} = \bigcap_{\alpha \in I} E_\alpha$, où E_α est un sous-ensemble fermé de V contenant E. $\overline{E}$ est le plus petit sous-ensemble fermé de V qui contient E ;*

— *le **bord** (ou **surface sphérique**) de $U_\varepsilon(x)$ est le sous ensemble de V défini par :*

$$\partial U_\varepsilon(x) = \{y \in V \ : \ \|x - y\| = \varepsilon\},$$

si on utilise le symbole $B(x,\varepsilon)$, alors son bord est noté $\partial B(x,\varepsilon)$;

— *le **voisinage (ou boule, ou sphère) fermé de rayon** ε **de** $x \in V$ est le sous ensemble de V défini par :*

$$\overline{U_\varepsilon(x)} = \{y \in V \ : \ \|x - y\| \leqslant \varepsilon\},$$

si on utilise le symbole $B(x,\varepsilon)$, alors on écrit $\overline{B(x,\varepsilon)}$.

On rappelle qu'une topologie sur V est un sous-ensemble de parties de V qui contient V même et $\varnothing$ et qui est stable par rapport aux unions quelconques et aux intersections finies. La topologie engendrée par les ouverts de V est la plus petite topologie qui contient les ouverts de V. Avec cette dernière topologie, par rapport aux ouvert définis ci-dessus, V est un espace topologique.

La topologie de V est *métrique*, i.e. les ouverts sont définis grâce à une fonction distance. On rappelle que cela assure que la topologie soit **séparée**, i.e. le fait que, pour toute couple $x, y \in V$, $x \neq y$, existent deux voisinages $U(x)$ et $V(y)$, de rayon égale ou différent, tels que $U(x) \cap V(y) = \varnothing$, on dit que les points sont séparés par les voisinages.

Un résultat standard de topologie garantit l'*unicité de la limite* des suites dans une topologie séparée, donc les suites de vecteurs de V, si convergent, ont une seule limite.

Dans une topologie métrique, l'unicité de la limite descend trivialement de l'inégalité triangulaire, en fait, si $x_n \underset{n\to+\infty}{\to} x$ et $x_n \underset{n\to+\infty}{\to} y$, alors :

$$0 \leqslant d(x,y) \leqslant d(x,x_n) + d(x_n,y) \underset{n\to+\infty}{\to} 0 + 0 = 0.$$

C'est utile de réécrire ici la définition de convergence d'une suite dans la topologie de V.

Déf. 4.1.3 *Soit $(V, \| \|)$ un espace vectoriel normé. Alors, une suite de vecteurs $(x_n)_{n\in\mathbb{N}} \subset V$ est **convergente**, ou **convergente en norme** $\| \|$, vers la limite x si :*

$$\forall \varepsilon > 0 \; \exists N_\varepsilon > 0 \; : \; n \geqslant N_\varepsilon \; \Rightarrow \; \|x_n - x\| < \varepsilon,$$

i.e. si, à partir de $n = N_\varepsilon$, $x_n \in U_\varepsilon(x)$. On symbolise ça avec les écritures plus compactes :

$$\lim_{n\to+\infty} x_n = x, \quad \|x_n - x\| \underset{n\to+\infty}{\to} 0.$$

Remarque 4.1.1 Le fait de demander que l'inégalité $\|x_n - x\| < \varepsilon$ soit valide $\forall \varepsilon > 0$ permet de reformuler la définition de convergence en rajoutant une *constante multiplicative strictement positive et finie* à ε, i.e. $x_n \underset{n\to+\infty}{\to} x$ si :

$$\forall \varepsilon > 0 \; \exists N_\varepsilon > 0 \text{ et } \exists m \in]0, +\infty[\; : \; n \geqslant N_\varepsilon \; \Rightarrow \; \|x_n - x\| < m\varepsilon,$$

en fait, si la propriété est valide pour toute ε positive et arbitrairement petite, alors on peut considérer $\tilde{\varepsilon} = \frac{\varepsilon}{m}$ et redéfinir la convergence par rapport à $\tilde{\varepsilon}$:

$$\forall \tilde{\varepsilon} > 0 \; \exists N_{\tilde{\varepsilon}} > 0 \; : \; n \geqslant N_{\tilde{\varepsilon}} \; \Rightarrow \; \|x_n - x\| < \tilde{\varepsilon},$$

vu que le symbole ε ou $\tilde{\varepsilon}$ n'a aucune importance car les deux sont des quantités petites comme on le veut, les deux définitions sont équivalentes.

Dans la suite on fera référence à cette considération en utilisant tout simplement la formule linguistique standard « **grâce à l'arbitrarité de ε...** ».

Rappelons aussi le concept de densité d'un sous-ensemble dans un espace vectoriel normé. Dans tous les livres de topologie on démontre que les propriétés contenues dans la définition suivante sont équivalentes.

Déf. 4.1.4 (Densité d'une partie) *Soit $(V, \|\ \|)$ un espace vectoriel normé. On dit qu'un sous-ensemble $E \subset V$ est dense dans V si une des propositions suivantes est vérifiée :*

1. $\forall x \in V, \exists (x_n)_{n\in\mathbb{N}} \subset E \ : \ x_n \underset{n\to+\infty}{\longrightarrow} x$ *(i.e.* $\|x_n - x\| \underset{n\to+\infty}{\longrightarrow} 0$*), c'est-à-dire :* ***tout élément de V peut être approché indéfiniment par une suite d'éléments de E****, en étant sa limite.*

2. $\forall x \in V$ *et* $\forall \varepsilon > 0,\ \exists y \in E : \|x - y\| < \varepsilon$*, i.e.* ***pour tout élément x de V il existe un élément y de E qui a une distance de x petite comme l'on veut****.*

3. *V est l'adhérence de E :* $\boxed{\overline{E} = V}$.

Rappelons aussi la définition de série et de série convergente.

Déf. 4.1.5 *Étant donné la suite de vecteurs $(x_n)_{n\in\mathbb{N}} \subset V$, on appelle* ***série*** *de terme général x_n la suite des sommes partielles $(S_n)_{n\in\mathbb{N}}$, où $S_n = \sum\limits_{k=0}^{n} x_k$, et on écrit :*

$$\sum_{n\in\mathbb{N}} x_n = \sum_{n=0}^{\infty} x_n = (S_n)_{n\in\mathbb{N}}.$$

La ***série*** $\sum\limits_{n\in\mathbb{N}} x_n$ *est dite* ***convergente****, ou convergente en norme $\|\ \|$, vers la somme x si la suite des sommes partielles $(S_n)_{n\in\mathbb{N}}$ est convergente vers x, i.e. :*

$$x = \lim_{n\to+\infty} \sum_{k=0}^{n} x_k = \sum_{n\in\mathbb{N}} x_n \quad \Leftrightarrow \quad \lim_{n\to+\infty} \|S_n - x\| = 0 \quad \Leftrightarrow \quad \|S_n - x\| \underset{n\to+\infty}{\to} 0.$$

La ***série*** $\sum\limits_{n\in\mathbb{N}} x_n$ *est dite* ***absolument convergente***[1] *si la suite de la norme des sommes partielles des normes, i.e.* $\left(\sum\limits_{k=0}^{n} \|x_k\|\right)_{n\in\mathbb{N}}$ *est convergente, dans ce cas on écrit :* $\sum\limits_{n\in\mathbb{N}} \|x_n\| < +\infty$.

On observe que, comme $S_n - x = \sum\limits_{k=0}^{n} x_k - \sum\limits_{k=0}^{\infty} x_n = -\sum\limits_{k=n+1}^{\infty} x_k$, alors :

$$\|S_n - x\| = \left\| \sum_{k=n+1}^{\infty} x_k \right\|, \tag{4.1}$$

donc, la définition explicite de série convergente dans un espace vectoriel normé est la suivante :

$$\forall \varepsilon > 0\ \exists N_\varepsilon > 0 \ : \ n \geqslant N_\varepsilon \ \Rightarrow \ \left\| \sum_{k=n+1}^{\infty} x_k \right\| < \varepsilon.$$

C'est important de rappeler aussi la définition de continuité d'une fonction entre espaces métriques et un résultat classique de la théorie des espaces métriques, qui permet de caractériser la continuité avec l'action de la fonction sur les suites.

1. La convergence absolue qu'on vient de définir devient l'habituelle convergence des modules de S_n quand $V = \mathbb{R}$ ou $V = \mathbb{C}$.

Déf. 4.1.6 (Limites et continuité pour fonctions entre espaces métriques) *Soient X et Y deux espaces métriques quelconques, $\bar{x} \in X$ et $\ell \in Y$, alors*

$$\boxed{\lim_{x\to\bar{x}} f(x) = \ell} \iff \forall \varepsilon > 0\ \exists \delta_\varepsilon > 0\ :\ x \in U_{\delta_\varepsilon}(\bar{x}) \cap X \quad \Longrightarrow \quad f(x) \in U_\varepsilon(\ell) \cap Y,$$

i.e. la limite de f en $\bar{x}$ est ℓ si f transforme les points de X arbitrairement proches de $\bar{x}$ en points de Y arbitrairement proches de ℓ.

Si $\ell = f(\bar{x})$, alors la fonction $f : X \to Y$ est dite continue en $\bar{x} \in V$, explicitement :

$$\forall \varepsilon > 0\ \exists \delta_\varepsilon > 0\ :\ x \in U_{\delta_\varepsilon}(\bar{x}) \cap X \quad \Rightarrow \quad f(x) \in U_\varepsilon(f(\bar{x})) \cap Y.$$

f est continue sur X si elle est continue en tout point de X.

Théorème 4.1.1 (Continuité par suites) *La fonction $f : X \to Y$, (X, d_X) et (Y, d_Y) espaces métriques quelconques, est continue en $\bar{x} \in X$ si et seulement si :*

$$\forall (x_n)_{n\in\mathbb{N}} \subseteq X \text{ telle que } \lim_{n\to+\infty} x_n = \bar{x} \quad \Rightarrow \quad \lim_{n\to+\infty} f(x_n) = f\left(\lim_{n\to+\infty} x_n\right) = f(\bar{x}),$$

i.e. :

$$\forall (x_n)_{n\in\mathbb{N}} \subseteq X\ :\ d_X(x_n, \bar{x}) \underset{n\to+\infty}{\to} 0 \quad \Rightarrow \quad d_Y(f(x_n), f(\bar{x})) \underset{n\to+\infty}{\to} 0.$$

Nous remarquons que l'opération de limite sur la suite $(x_n)_{n\in\mathbb{N}}$ est effectuée dans l'espace métrique (X, d_X), tandis que celle sur la suite $(f(x_n))_{n\in\mathbb{N}}$ est effectuée dans l'espace métrique (Y, d_Y).

La possibilité d'échanger l'ordre entre la limite et la fonction (continue) dans l'expression

$$\boxed{\lim_{n\to+\infty} f(x_n) = f\left(\lim_{n\to+\infty} x_n\right)}$$

sera essentielle pour prouver beaucoup de résultats dans la suite.

Preuve.

$\boxed{\Longrightarrow}$: soit f continue en $\bar{x}$ et soit $(x_n)_{n\in\mathbb{N}} \subseteq X$ une suite arbitraire d'éléments de X telle que $\lim_{n\to+\infty} x_n = \bar{x}$, alors, par définition de limite de suite, pour n suffisamment grand, x_n appartient à un voisinage de rayon petit comme l'on veut, disons $\delta > 0$, de $\bar{x}$, i.e. il existe $N \in \mathbb{N}$ tels que $n \geqslant N \implies x_n \in U_\delta(\bar{x})$. De l'autre côté, grâce à la continuité de f, les éléments x_n qui appartiennent au voisinage $U_\delta(\bar{x})$ sont transformés par f dans des points qui appartiennent à un voisinage de rayon arbitrairement petit, disons $\varepsilon > 0$, de $f(\bar{x})$, i.e. $n \geqslant N \implies f(x_n) \in U_\varepsilon(f(\bar{x})) \cap Y$, i.e. $\lim_{n\to+\infty} f(x_n) = f(\bar{x})$.

$\boxed{\Longleftarrow}$: soit vrai que, pour toute suite $(x_n)_{n\in\mathbb{N}} \subseteq X$ telle que $\lim_{n\to+\infty} x_n = \bar{x} \in X$, ça vaut $\lim_{n\to+\infty} f(x_n) = f(\bar{x})$, il faut démontrer que ceci implique que f est continue en $\bar{x} \in X$.

Supposons, par l'absurde, que f ne soit pas continue en $\bar{x}$ et démontrons que ça amène à une contradiction. La négation [2] de la continuité de f en $\bar{x}$ correspond à dire que $\forall \delta > 0$, $\exists \varepsilon_\delta > 0$ tel que $x \in U_\delta(\bar{x}) \subset X$ implique $f(x) \notin U_{\varepsilon_\delta}(f(\bar{x}))$.

2. Rappeler que la négation d'une proposition mathématique est réalisée en échangeant les quantificateurs universels et existentiels entre eux et en considérant l'affirmation complémentaire de celle de la proposition initiale, donc la négation de $(\forall A\ \exists B \implies C)$ est $(\forall B\ \exists A \implies \bar{C})$, où $\bar{C}$ est la négation de l'affirmation C.

Vu que les valeurs de δ sont arbitraires, nous pouvons considérer la suite $(\delta_n)_{n\geqslant 1}$ définie par $\delta_n = \frac{1}{n}\ \forall n \geqslant 1$, qui entraîne l'existence d'une suite $(x_n)_{n\geqslant 1} \subset X$ telle que $x_n \in U_{\delta_n}(\bar{x})$ et $f(x_n) \notin U_{\varepsilon_{\delta_n}}(f(\bar{x}))$.

Nous sommes arrivés à la contradiction : d'un côté, quand $n \to +\infty$, $\delta_n \to 0$ et donc $x_n \to \bar{x}$, mais de l'autre côté $f(x_n) \not\to f(\bar{x})$, i.e. l'hypothèse que f ne soit pas continue nous a permis de construire une suite d'éléments de X qui converge vers $\bar{x}$ sans que $f(x)$ soit convergente vers $f(\bar{x})$, mais ceci est contraire à notre hypothèse de départ et donc la possibilité que f ne soit pas continue doit être rejetée. □

Si V, W sont deux espaces vectoriels normés, alors ils sont automatiquement deux espaces métriques par rapport à la distance canoniquement induite par la norme et les définitions et donc résultats rappelés ci-dessus restent valides.

4.2 La continuité des opérations fondamentales des espaces vectoriels avec produit scalaire

La présence simultanée d'une structure linéaire et d'une topologique métrique sur un espace vectoriel avec produit scalaire pose, d'une manière naturelle, la question de la *compatibilité* entre les deux structures, i.e. les opérations linéaires de V en tant qu'espace vectoriel sont elles continues dans la topologie de V canoniquement engendrée par son produit scalaire ? Le théorème suivant dit que la réponse à cette question est affirmative.

Théorème 4.2.1 *Soit $(V, \langle\ ,\ \rangle)$ un espace vectoriel sur $\mathbb{K}$ avec produit scalaire. Considérons la topologie induite par le produit scalaire sur V, la topologie Euclidienne usuelle sur $\mathbb{K}$ et la topologie produit sur $V \times V$ et sur $\mathbb{K} \times V$. Alors,*

le produit scalaire

$$\begin{array}{rccc} \langle\ ,\ \rangle : & V \times V & \longrightarrow & \mathbb{K} \\ & (x, y) & \longmapsto & \langle x, y\rangle, \end{array}$$

la norme

$$\begin{array}{rccc} \|\ \| : & V & \longrightarrow & \mathbb{R}_0^+ \\ & x & \longmapsto & \|x\|, \end{array}$$

la somme

$$\begin{array}{rccc} + : & V \times V & \longrightarrow & V \\ & (x, y) & \longmapsto & x + y, \end{array}$$

et le produit par un scalaire

$$\begin{array}{rccc} \cdot_{\mathbb{K}} : & \mathbb{K} \times V & \longrightarrow & V \\ & (k, x) & \longmapsto & kx, \end{array}$$

sont des fonctions continues.

Preuve. Dans toutes les preuves qu'on verra, la technique consiste, à chaque fois, à majorer une norme opportune avec une expression qui contient la norme de la différence entre une suite et sa limite, qui tends bien sûr vers 0.

— *Continuité du produit scalaire* : il faut démontrer que si $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ sont deux suites quelconques d'éléments de V qui convergent vers x et y, respectivement, alors la suite de scalaires $(\langle x_n, y_n\rangle)_{n\in\mathbb{N}}$ converge vers $\langle x, y\rangle$. Pour faire ça, on doit d'abord écrire une simple manipulation algébrique qui vaut pour tout $n \in \mathbb{N}$:

$$\begin{aligned}\langle x_n, y_n\rangle - \langle x, y\rangle &= \langle x_n - x + x, y_n - y + y\rangle - \langle x, y\rangle \\ &= \langle x_n - x, y_n - y\rangle + \langle x_n - x, y\rangle + \langle x, y_n - y\rangle + \cancel{\langle x, y\rangle} - \cancel{\langle x, y\rangle} \\ &= \langle x_n - x, y_n - y\rangle + \langle x_n - x, y\rangle + \langle x, y_n - y\rangle.\end{aligned}$$

On peut écrire la majoration suivante :

$$|\langle x_n, y_n\rangle - \langle x, y\rangle| \leqslant |\langle x_n - x, y_n - y\rangle| + |\langle x_n - x, y\rangle| + |\langle x, y_n - y\rangle| \quad \forall n \in \mathbb{N},$$

mais, grâce à l'inégalité de Cauchy-Schwarz :

$$|\langle x_n, y_n\rangle - \langle x, y\rangle| \leqslant \|x_n - x\|\|y_n - y\| + \|x_n - x\|\|y\| + \|x\|\|y_n - y\| \quad \forall n \in \mathbb{N}.$$

Comme l'égalité vaut pour tout $n \in \mathbb{N}$, on peut considérer la limite $n \to +\infty$ aux deux côtés : par hypothèse, $\|x_n - x\| \underset{n\to+\infty}{\to} 0$ et $\|y_n - y\| \underset{n\to+\infty}{\to} 0$, donc le côté de droite tend vers 0, et alors :

$$|\langle x_n, y_n\rangle - \langle x, y\rangle| \underset{n\to+\infty}{\to} 0,$$

ce qui prouve la continuité du produit scalaire.

— *Continuité de la norme* : il faut démontrer que si $(x_n)_{n\in\mathbb{N}}$ est une suites quelconque d'éléments de V qui converge vers x, alors la suite de réels positifs $(\|x_n\|)_{n\in\mathbb{N}}$ converge vers $\|x\|$. Pour faire ça, on peut utiliser la majoration de la norme vu dans la formule (1.3) :

$$|\|x_n\| - \|x\|| \leqslant \|x_n - x\|,$$

mais $\|x_n - x\| \underset{n\to+\infty}{\to} 0$, donc $\|x_n\| \underset{n\to+\infty}{\to} \|x\|$.

— *Continuité de la somme* : il faut démontrer que si $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ sont deux suites quelconques d'éléments de V qui convergent vers x et y, respectivement, alors la suite somme $(x_n + y_n)_{n\in\mathbb{N}}$ converge vers $x + y$. Pour cela, on écrit :

$$\|(x_n + y_n) - (x + y)\| = \|(x_n - x) + (y_n - y)\| \leqslant \|x_n - x\| + \|y_n - y\| \underset{n\to+\infty}{\to} 0.$$

— *Continuité du produit par un scalaire* : il faut démontrer que si $(x_n)_{n\in\mathbb{N}}$ et $(k_n)_{n\in\mathbb{N}}$ sont deux suites quelconques d'éléments de V et de $\mathbb{K}$, respectivement, qui convergent vers x et k, respectivement, alors la suite produit $(k_n x_n)_{n\in\mathbb{N}}$ converge vers kx. Encore une fois, on a besoin d'une petite manipulation algébrique :

$$\begin{aligned}\|k_n x_n - kx\| &= \|k_n(x_n - x + x) - kx\| = \|k_n(x_n - x) + k_n x - kx\| \\ &= \|k_n(x_n - x) + x(k_n - k)\| \leqslant |k_n|\|x_n - x\| + \|x\||k_n - k| \underset{n\to+\infty}{\to} 0.\end{aligned}$$

□

Allons voir des conséquences immédiates du théorème qu'on vient de voir : tout d'abord, la

continuité de la somme et du produit par un scalaire implique la continuité de la différence, car $x - y = x + (-1)y$.

Si $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ sont deux suites de $(V, \langle\ ,\ \rangle)$ qui convergent vers les éléments de x et y, respectivement, alors la continuité du produit scalaire et de la norme et le théorème 4.1.1 implique ces deux formules qu'on utilisera souvent dans la suite :

$$\boxed{\lim_{n\to+\infty}\langle x_n, y_n\rangle = \big\langle \lim_{n\to+\infty} x_n, \lim_{n\to+\infty} y_n\big\rangle = \langle x, y\rangle}, \tag{4.2}$$

$$\boxed{\lim_{n\to+\infty} \|x_n\| = \left\|\lim_{n\to+\infty} x_n\right\| = \|x\|}. \tag{4.3}$$

Le cas des séries mérite une observation séparée : étant données les séries convergentes $\sum\limits_{n\in\mathbb{N}} x_n$, $\sum\limits_{m\in\mathbb{N}} y_m$, on peut utiliser le fait qu'une série est la suite de ses sommes partielles pour écrire :

$$\boxed{\Big\langle \sum_{n\in\mathbb{N}} x_n, \sum_{m\in\mathbb{N}} y_m\Big\rangle = \Big\langle \lim_{N\to+\infty}\sum_{n=0}^{N} x_n, \lim_{K\to+\infty}\sum_{m=0}^{K} y_m\Big\rangle = \lim_{N,K\to+\infty}\sum_{n=0}^{N}\sum_{m=0}^{K}\langle x_n, y_m\rangle}, \tag{4.4}$$

et

$$\boxed{\left\|\sum_{n\in\mathbb{N}} x_n\right\| = \left\|\lim_{N\to+\infty}\sum_{n=0}^{N} x_n\right\| = \lim_{N\to+\infty}\left\|\sum_{n=0}^{N} x_n\right\|}. \tag{4.5}$$

Si on élève au carré les membres de l'équation (4.5) on obtient :

$$\left\|\sum_{n\in\mathbb{N}} x_n\right\|^2 = \left\|\lim_{N\to+\infty}\sum_{n=0}^{N} x_n\right\|^2 = \left(\lim_{N\to+\infty}\left\|\sum_{n=0}^{N} x_n\right\|\right)^2 = \lim_{N\to+\infty}\left\|\sum_{n=0}^{N} x_n\right\|^2,$$

où la dernière égalité est possible grâce au fait que les quantités $\left\|\sum\limits_{n=0}^{N} x_n\right\|$ sont réelles et que la puissance carrée est continue en $\mathbb{R}$.

Si, au lieu d'une suite $(x_n)_{n\in\mathbb{N}}$ quelconque, nous considérons une *famille orthogonale de vecteurs* $(u_n)_{n\in\mathbb{N}}$, alors on peut utiliser le théorème de Pythagore généralisé 1.4.1 pour écrire $\lim\limits_{N\to+\infty}\left\|\sum\limits_{n=0}^{N} u_n\right\|^2 = \lim\limits_{N\to+\infty}\sum\limits_{n=0}^{N}\|u_n\|^2 = \sum\limits_{n\in\mathbb{N}}\|u_n\|^2$, ce qui amène à la formule très utile suivante :

$$\boxed{\left\|\sum_{n\in\mathbb{N}} u_n\right\|^2 = \sum_{n\in\mathbb{N}}\|u_n\|^2} \qquad (u_n)_{n\in\mathbb{N}} \text{ : famille orthogonale de vecteurs.} \tag{4.6}$$

Dans le chapitre 5 on utilisera souvent la formule (4.6). Il est important de souligner que cette formule ne vaut pas, en général, si $(u_n)_{n\in\mathbb{N}}$ n'est pas un système orthogonal de vecteurs et si on considère la norme au lieu de son carré.

La **possibilité d'échanger l'opération de limite avec celle de produit scalaire et de norme** sera indispensable pour prouver beaucoup de théorèmes dans la suite. Ceci souligne *l'importance de la compatibilité entre la structure linéaire et celle topologique* dans un espace vectoriel avec produit scalaire.

On va montrer tout de suite un résultat dans lequel on utilise la continuité de la norme. Dans le chapitre 1 on a annoncé que la formule du parallélogramme permet de caractériser les normes issues d'un produit scalaire, i.e. les normes hilbertiennes. Nous avons tous les instruments pour formaliser cette affirmation, qui est appelée théorème de Fréchet-von Neumann-Jordan dans [15].

Théorème 4.2.2 (M. Fréchet, J. von Neumann, P. Jordan) *Soit V un espace vectoriel sur $\mathbb{K}$ (de dimension finie ou infinie) et soit $\| \ \|$ une norme sur V. $\| \ \|$ est une norme hilbertienne si et seulement elle satisfait la formule du parallélogramme.*

Si la norme satisfait la formule du parallélogramme, alors le produit scalaire qui l'induit est univoquement déterminé par les formules de polarisation dans le cas réel et complexe, respectivement :

$$\langle v, w\rangle = \frac{1}{4}(\|v+w\|^2 - \|v-w\|^2),$$

$$\langle v, w\rangle = \frac{1}{4}\left[\|v+w\|^2 - \|v-w\|^2 + i\left(\|v+iw\|^2 - \|v-iw\|^2\right)\right].$$

Preuve. L'implication directe est évidente, donc on doit seulement démontrer l'implication inverse, i.e. que si une norme $\| \ \|$ satisfait la formule du parallélogramme, alors elle est induite par un produit scalaire de la manière canonique : $\| \cdot \| = \sqrt{\langle \cdot, \cdot \rangle}$.

On commence avec le cas réel. S'il existe un produit scalaire qui induit la norme, alors il doit avoir la forme suivante

$$p(v, w) = \frac{1}{4}(\|v+w\|^2 - \|v-w\|^2), \qquad \forall v, w \in V,$$

on observe que p est composition de fonctions algébriques (somme et différence), de la norme et de la puissance carrée, qui sont toutes fonctions continues, donc p est elle-même une fonction continue de ses arguments.

Il faut vérifier que cette définition satisfait les propriétés définitoires d'un produit scalaire réel.

Tout d'abord on observe que la symétrie $p(v, w) = p(w, v)$ est évidente et pareil pour la définie positivité, vu que $p(v, v) = \frac{1}{4}\|2v\|^2 = \|v\|^2 \geqslant 0$ et $p(v, v) = 0$ si et seulement si $v = 0_V$.

Il reste à vérifier la bilinéarité. Grâce à la symétrie, toute propriété de p qu'on démontre par rapport au premier argument, vaut aussi pour le deuxième, donc on va se fixer seulement sur la première entrée de p.

On commence par observer que, grâce à la formule du parallélogramme, on peut écrire $\forall v, w, z \in V$:

$$\|(v+z)+w\|^2 + \|(v-z)+w\|^2 = \|(v+w)+z\|^2 + \|(v+w)-z\|^2 = 2\|v+w\|^2 + 2\|z\|^2$$

et

$$\|(v+z)-w\|^2 + \|(v-z)-w\|^2 = \|(v-w)+z\|^2 + \|(v-w)-z\|^2 = 2\|v-w\|^2 + 2\|z\|^2$$

donc

$$\begin{aligned} p(v+z,w)+p(v-z,w) &= \frac{1}{4}(\|v+z+w\|^2-\|v+z-w\|^2+\|v-z+w\|^2-\|v-z-w\|^2) \\ &= \frac{1}{4}[2(\|v+w\|^2+\|z\|^2)-2(\|v-w\|^2+\|z\|^2)] \\ &= 2\frac{1}{4}(\|v+w\|^2-\|v-w\|^2) \\ &= 2p(v,w), \end{aligned} \tag{4.7}$$

si on choisit $v = z$, alors on obtient $p(v+v,w) + \cancelto{0}{p(v-v,w)} = p(2v,w) \underset{(4.7)}{=} 2p(v,w)$ $\forall v, w \in V$, i.e.

$$2p(v,w) = p(2v,w), \qquad \forall v, w \in V. \tag{4.8}$$

Soient maintenant $v_1, v_2 \in V$ tels que $v = \frac{1}{2}(v_1+v_2)$ et $z = \frac{1}{2}(v_1-v_2)$, alors $v+z = v_1$ et $v-z = v_2$, alors

$$\begin{aligned} p(v_1,w)+p(v_2,w) = p(v+z,w)+p(v-z,w) &\underset{(4.7)}{=} 2p(v,w) \underset{(4.8)}{=} p(2v,w) \\ &\underset{(\text{déf. } v)}{=} p(v_1+v_2,w), \end{aligned}$$

comme v, w, z sont des vecteurs arbitraires, aussi v_1, v_2 sont arbitraires, et alors ce qu'on vient de démontrer, i.e. $p(v_1+v_2,w) = p(v_1,w)+p(v_2,w)$, prouve l'additivité de p.

Pour démontrer l'homogénéité on commence à observer que, si on itère $n \in \mathbb{N}$ fois le raisonnement qui a conduit à $p(2v,w) = 2p(v,w)$, alors on obtient $p(nv,w) = np(v,w)$.

De plus, pour tout $m \in \mathbb{N}$, $m \neq 0$, ça vaut $p(v,w) = p(m\frac{v}{m},w)$, mais aussi $p(m(\frac{v}{m}),w) = mp(\frac{v}{m},w)$, que, combiné avec la formule $p(nv,w) = np(v,w)$ donne $p(\frac{n}{m}v,w) = \frac{n}{m}p(v,w)$ $\forall n, m \in \mathbb{N}$, $m \neq 0$, i.e. p est homogène par rapport à tout nombre $r \in \mathbb{Q}$, $r \geqslant 0$: $p(rv,w) = rp(v,w)$.

Pour étendre l'homogénéité à tout rationnel on utilise l'argument suivant : si $r < 0$, alors, en réécrivant $rv = -|r|v = |r|(-v)$, alors

$$\begin{aligned} rp(v,w)-p(rv,w) &= rp(v,w)-p(|r|(-v),w) = rp(v,w)-|r|p(-v,w) = rp(v,w)+rp(-v,w) \\ &= r(p(v,w)+p(-v,w)) \underset{(\text{additivité})}{=} rp(v-v,w) = rp(0_V,w) = 0, \end{aligned}$$

donc l'homogénéité vaut aussi pour les rationnels négatifs, et alors pour tous les rationnels. Maintenant on utilise le fait que $\mathbb{Q}$ est dense en $\mathbb{R}$, donc, pour tout $\alpha \in \mathbb{R}$ il existe une suite de rationnels $(r_n)_{n\in\mathbb{N}} \subset \mathbb{Q}$ tels que $r_n \underset{n\to+\infty}{\longrightarrow} \alpha$ et, grâce à la continuité de p on a

$$\alpha p(v,w) = \lim_{n\to+\infty} r_n p(v,w) = p(\lim_{n\to+\infty} r_n v, w) = p(\alpha v, w), \quad \forall \alpha \in \mathbb{R},\ v, w \in V.$$

En résumé, p est un produit scalaire sur V compatible avec sa norme si $\mathbb{K} = \mathbb{R}$.

Nous considérons maintenant le cas complexe $\mathbb{K} = \mathbb{C}$. Comme observé dans le cas réel, s'il existe un produit scalaire qui induit la norme, alors il doit avoir la forme suivante

$$\tilde{p}(v,w) = \frac{1}{4}\left[\|v+w\|^2-\|v-w\|^2+i\left(\|v+iw\|^2-\|v-iw\|^2\right)\right] = p(v,w)+ip(v,iw),$$

$\forall v, w \in V$.

Grâce aux observations de la section 1.1, pour vérifier que $\tilde{p}(v, w)$ soit un produit scalaire complexe, ça suffit de vérifier la propriété hermitienne, i.e. $\tilde{p}(v, w) = \overline{\tilde{p}(w, v)}$, vu que la linéarité dans la première variable et la définie positivité de p impliquent la validité de ces propriétés aussi pour $\tilde{p}$.

$\tilde{p}$ est une forme hermitienne si et seulement si $\tilde{p}(w, v) = \overline{\tilde{p}(v, w)} = p(v, w) - ip(v, iw)$, vu que $p(v, w)$ et $p(v, iw) \in \mathbb{R}$! Mais $\tilde{p}(w, v) = p(w, v) + ip(iw, v) = p(v, w) + ip(iw, v)$, car on sait que $p(v, w) = p(w, v)$, donc $\tilde{p}(w, v) = p(v, w) + ip(iw, v)$. Si on compare les formules :

$$\tilde{p}(w, v) = \overline{\tilde{p}(v, w)} = p(v, w) - ip(v, iw) \quad \text{et} \quad \tilde{p}(w, v) = p(v, w) + ip(iw, v),$$

on voit que $\tilde{p}$ est une forme hermitienne si et seulement si $p(v, iw) = -p(iw, v)$ $\forall v, w \in V$.

Calculons alors :

$$\begin{aligned}
\tilde{p}(v, iw) &= \frac{1}{4}(\|v + iw\|^2 - \|v - iw\|^2) = \frac{1}{4}(|i| \, \|v + iw\|^2 - |i| \, \|v - iw\|^2) \\
&= \frac{1}{4}(\|iv - w\|^2 - \|iv + w\|^2) = -\frac{1}{4}(\|w + iv\|^2 - \|w - iv\|^2) \\
&= -p(w, iv),
\end{aligned}$$

en ayant utilisé le fait que $\|w - iv\| = \|iv - w\|$. En résumé, $\tilde{p}$ est le produit scalaire associé à notre norme dans le cas complexe. □

Allons introduire maintenant un objet mathématique très importante.

Déf. 4.2.1 (Espace vectoriel topologique) *On appelle espace vectoriel topologique un espace vectoriel V muni d'une topologie compatible avec la structure linéaire de V, i.e. telle que les opérations linéaires de somme et de multiplication par un scalaire sont des fonctions continues.*

Grâce à la continuité des opérations fondamentales dans un espace vectoriel avec produit scalaire il suit que ce type d'espace est toujours un espace vectoriel topologique. On peut dire la même chose pour un espace vectoriel normé, la preuve de la continuité des opérations linéaires est exactement la même. En fait, par rapport aux arguments topologiques, il n'y a aucune différence entre un espace vectoriel avec produit scalaire et un espace vectoriel normé, car c'est toujours la norme qui intervient dans les preuves des propriétés de continuité.

La différence majeure entre un espace vectoriel avec produit scalaire et un espace vectoriel normé consiste en la structure géométrique subjacente à l'espace même, qui est beaucoup plus riche pour le premier type d'espace.

4.2.1 L'équivalence des topologies séparées des espaces vectoriels de dimension finie

Dans les preuves du théorème 4.2.1, la dimension de l'espace vectoriel ne jouait aucun rôle, donc les considérations qu'on a fait dans la sous-section antécédente sont valides pour tout espace vectoriel, de dimension finie ou infinie.

Néanmoins, en dimension finie, on peut assurer un résultat d'essentielle unicité de la topologie (séparable) d'un espace vectoriel topologique.

Théorème 4.2.3 (Tychonov) *Soit V un espace vectoriel topologique* ***séparée*** *de dimension finie n sur le corps $\mathbb{K}$. Alors, fixée une base $B = (b_1, \ldots, b_n)$ quelconque de V, l'isomorphisme linéaire définit par*

$$\begin{array}{rccc} I: & V & \longrightarrow & \mathbb{K}^n \\ & x = [x]_B = \sum_{i=1}^{n} x_i b_i & \longmapsto & \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \end{array}$$

est un homéomorphisme (ou isomorphisme topologique), i.e. une application bicontinue (continue, inversible, avec inverse continue) si on considère sur $\mathbb{K}^n$ la topologie Euclidienne usuelle.

Nous avons dit qu'un espace vectoriel avec produit scalaire, normé sont ou métrique est un espace vectoriel topologiques séparé, donc le théorème de Tychonov a comme conséquence immédiate que *tous les produits scalaires, toutes les normes et les distances que l'on peut définir sur un espace vectoriel de dimension finie sont topologiquement équivalents*, i.e. ils génèrent la même topologie, qui est, à un isomorphisme près, la topologie Euclidienne ! Des contre-exemples montrent que, en dimension infinie, ceci n'est plus vrai.

L'exemple le plus simple d'indépendance topologique par rapport au choix de la norme dans des espaces vectoriels de dimension finie est fourni par les espaces vectoriels de dimension 1, comme montré par le résultat suivant.

Théorème 4.2.4 *Si V est un espace vectoriels normé mono-dimensionnel sur le corps $\mathbb{K}$, alors deux normes quelconques définies sur V sont multiples l'une de l'autre par un scalaire réel strictement positif.*

Preuve. Soient $\| \ \|_1$, $\| \ \|_2$ deux normes sur V. Par définition, $\|0_V\|_1 = \|0_V\|_2 = 0$, soit alors $v \in V$ arbitraire différent du vecteur nul. Les deux normes mesurerons, en général, un module différent du vecteur v : $\|v\|_1 = k_1$, $\|v\|_2 = k_2$, écrivons $\frac{\|v\|_1}{\|v\|_2} = \frac{k_1}{k_2} = k \in \mathbb{R}^+$, alors $\|v\|_1 = k\|v\|_2$.

Comme V a dimension unitaire, pour tout autre vecteur $w \in V$ il existe $\lambda \in \mathbb{K}$ tel que $w = \lambda v$. Alors, l'homogénéité de la norme nous permet d'écrire :

$$\|w\|_1 = \|\lambda v\|_1 = |\lambda|\|v\|_1 = |\lambda|k\|v\|_2 = k\|\lambda v\|_2 = k\|w\|_2,$$

i.e. pour tout $w \in V$ et pour chaque couple de normes $\| \ \|_1$, $\| \ \|_2$ sur V, il existe une constante $k \in \mathbb{R}^+$ telle que $\|w\|_1 = k\|w\|_2$. □

4.3 Suites de Cauchy et complétude : espaces de Banach et de Hilbert

Les études conduits entre la fin du siècle XIX et le début du siècle XX ont montré que, parmi les espaces vectoriels métriques, normés et avec produit scalaire de dimension infinie, ceux qui sont plus 'similaires' aux espaces Euclidiens de dimension finie peuvent être caractérisés par une propriété plutôt simple : les suites convergentes (dans l'espace même) sont identifiables avec les suites de Cauchy.

Déf. 4.3.1 *Si (X, d) est un espace métrique (pas forcement vectoriel), alors une suite $(x_n)_{n\in\mathbb{N}}$ est dite* ***de Cauchy*** *si :*

$$\forall \varepsilon > 0 \; \exists N_\varepsilon > 0 \; : \; n, m \geqslant N_\varepsilon \; \Rightarrow \; d(x_n, x_m) < \varepsilon,$$

i.e. la distance entre deux éléments de la suite devient indéfiniment petite au fur et à mesure que les indices des deux éléments grandissent.

On dit que (X, d) est un espace métrique ***complet*** *si toutes les suites de Cauchy sont convergentes vers des limites contenues en X.*

Voyons de suite la relation entre espaces métriques complets et fermés.

Théorème 4.3.1 *Si (X, d) est un espace métrique complet et (E, d), $E \subseteq X$ un sous-espace métrique fermé de X, alors (E, d) est complet.*

Preuve. Soit $(x_n)_{n\in\mathbb{N}} \subseteq E$ une suite de Cauchy, en particulier $(x_n)_{n\in\mathbb{N}} \subseteq X$ et alors, comme X est complet, $(x_n)_{n\in\mathbb{N}}$ converge vers un point limite $x \in X$. Mais les limites des suites de E appartiennent à $\overline{E}$ et, étant E fermé, $\overline{E} = E$, donc $x \in E$, i.e. toutes les suites de Cauchy d'éléments de E convergent en E lui-même. □

Théorème 4.3.2 *Si (X, d) est un espace métrique quelconque et (E, d), $E \subseteq X$ un sous-espace métrique complet de X, alors (E, d) est fermé.*

Preuve. Soit $x \in \overline{E}$, alors il existe une suite $(x_n)_{n\in\mathbb{N}} \subseteq E$ qui converge vers x. Vu que la suite est convergente, elle est de Cauchy en E, mais E est complet et donc $(x_n)_{n\in\mathbb{N}}$ doit être convergente vers un élément $y \in E$. Par unicité de la limite, $x = y$ et alors $x \in E$, i.e. $\overline{E} = E$. □

Un espace vectoriel avec produit scalaire, ou normé, est aussi un espace vectoriel métrique, dans ce cas la définition de suite de Cauchy peut être réécrite comme ça :

$$\forall \varepsilon > 0 \; \exists N_\varepsilon > 0 \; : \; n, m \geqslant N_\varepsilon \; \Rightarrow \; \|x_n - x_m\| < \varepsilon,$$

que certains auteurs écrivent d'une manière synthétique comme ça :

$$\lim_{n,m\to+\infty} \|x_n - x_m\| = 0.$$

Dans les cours d'analyse on démontre que $(\mathbb{R}^n, |\ |)$ **et** $(\mathbb{C}^n, |\ |)$ **sont des espaces métriques complets pour tout** $n \in \mathbb{N}$ **fini**. Grâce au théorème de Tychonov 4.2.3, les espaces vectoriels topologiques réels ou complexes séparées de dimension finie n sont topologiquement équivalents aux espaces Euclidiens $\mathbb{R}^n$ ou $\mathbb{C}^n$, respectivement, par conséquent **pour les espaces vectoriels pre-Hilbertiens de dimension finie la complétude n'est jamais un problème** : **les suites convergentes dans ces espaces sont toutes et seules les suites de Cauchy**.

Pour montrer avec simplicité l'existence d'espaces métriques non complets nous avons besoin d'un résultat basique sur les suites de Cauchy.

Théorème 4.3.3 *Toute suite convergente d'un espace métrique est nécessairement de Cauchy.*

Preuve. Si $x_n \underset{n\longrightarrow+\infty}{\rightarrow} \bar{x}$, alors, grâce à l'arbitrarité de ε et à l'inégalité triangulaire :

$$\forall \varepsilon > 0 \; \exists N_\varepsilon > 0 \; : \; n, m \geqslant N_\varepsilon \implies d(x_n, x_m) \leqslant d(x_n, \bar{x}) + d(\bar{x}, x_m) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

□

Ce résultat nous permet de démontrer que les espaces métriques [3] $(\mathbb{Q}, |\ |)$ et $(]0,1[, |\ |)$ ne sont pas complets. Pour vérifier que $(\mathbb{Q}, |\ |)$ n'est pas complet prenons, par exemple, la suite $((1+\frac{1}{n})^n)_{n\in\mathbb{N}}$, elle est une suite rationnelle car $\mathbb{Q}$ est stable par rapport aux opérations de somme, division et puissance et à leur composition, de plus, il est bien connu qu'elle converge vers e, la base des logarithmes naturels, donc, grâce au théorème 4.3.3 elle est une suite de Cauchy en $\mathbb{Q}$, interprétée comme sous-ensemble de $\mathbb{R}$.

Néanmoins e est un nombre irrationnel, i.e. $e \in \mathbb{R}\backslash\mathbb{Q}$, ce qui montre l'existence d'au moins une suite de Cauchy de $\mathbb{Q}$ qui converge hors de $\mathbb{Q}$ lui-même.

De la même manière, dans $(]0,1[, |\ |)$ nous pouvons considérer la suite $(\frac{1}{n})_{n\geqslant 1}$, évidemment contenue en $]0,1[$ et convergente vers 0, donc de Cauchy sur $]0,1[\subset\mathbb{R}$, mais $0 \notin]0,1[$.

Donc, malgré le fait que toute suite convergente soit nécessairement de Cauchy, l'inverse n'est pas toujours vrai, c'est à cause de ça qu'il est utile d'introduire une définition pour caractériser les espaces pour lesquels la condition de Cauchy est nécessaire et suffisante pour la convergence [4].

Déf. 4.3.2 (Espaces de Hilbert et de Banach) *Soit V un espace vectoriel de dimension finie ou infinie.*

- *Si $(V, \|\ \|)$ est un espace vectoriel normé et complet, on dit qu'il est un espace de Banach ;*
- *Si $(V, \langle\ ,\ \rangle)$ est un espace vectoriel avec produit scalaire et complet, on dit qu'il est un espace de Hilbert.*

Comme conséquence du théorème de Tychonov, les **espaces vectoriels réels ou complexes normés de dimension finie** sont **tous des espaces de Banach**, les **espaces vectoriels réels ou complexes avec produit scalaire de dimension finie** sont **tous des espaces de Hilbert**. Les **espaces de Hilbert** (de dimension finie ou infinie) **sont**, en tant qu'espaces vectoriels normés complets, **aussi espaces de Banach**, le vice-versa n'est pas vrai en général car, dans un espace de Banach, nous ne pouvons pas assurer l'existence d'un produit scalaire.

Nous allons rappeler deux résultats relatifs aux suites de Cauchy qui seront très utiles dans la suite. On commence par rappeler qu'une suite dans un espace métrique est dite bornée si tous les éléments de la suite peuvent être inscrits dans un voisinage de rayon fini d'un seul élément de la suite, comme précisé dans la définition suivante.

Déf. 4.3.3 *Une suite $(x_n)_{n\in\mathbb{N}}$ dans une espace métrique (X,d) est dite **bornée** s'il existe $x^* \in X$ et $M \geqslant 0$ tel que $d(x_n, x^*) \leqslant M \ \forall n \in \mathbb{N}$.*

3. On rappelle que $\mathbb{Q}$ n'est pas un espace vectoriel réel ou complexe, car il n'est pas stable par rapport au produit par un scalaire réel ou complexe, donc, pour $\mathbb{Q}$ nous ne pouvons pas appliquer le théorème de Tychonov.

4. On définit aussi les espaces de **Fréchet**, qui sont des espaces vectoriels topologiques localement convexes complets par rapport à une topologie invariant par translation.

Théorème 4.3.4 *Une suite de Cauchy est toujours bornée.*

Preuve. Par définition, si $(x_n)_{n\in\mathbb{N}}$ est une de suite de Cauchy, il existe $N_\varepsilon > 0$ tel que la distance entre x_{N_ε} et tous les éléments x_n de la suite avec $n \geqslant N_\varepsilon$ est inférieure à ε, i.e. $d(x_n, x_{N_\varepsilon}) < \varepsilon$ $\forall n \geqslant N_\varepsilon$. x_{N_ε} est alors un bon candidat pour jouer le rôle de x^* dans la définition de suite bornée.

Pour démontrer que c'est le cas, observons que les éléments de la suite correspondants à une valeur de l'indice n inférieure à N_ε appartiennent à X, donc leur distance de x_{N_ε} est finie et nous pouvons définir la valeur :

$$r = \max\{d(x_{N_\varepsilon}, x_0), d(x_{N_\varepsilon}, x_1), \ldots, d(x_{N_\varepsilon}, x_{N_\varepsilon - 1})\}.$$

Si maintenant on définit $M = \max\{\varepsilon, r\}$, on obtient $d(x_n, x_{N_\varepsilon}) \leqslant M$ $\forall n \in \mathbb{N}$. □

Le deuxième résultat fait référence aux sous-suites, ou suites extraites.

Déf. 4.3.4 *Soit $(x_n)_{n\in\mathbb{N}}$ une suite dans une espace métrique (X, d) et soit $\varphi : \mathbb{N} \to \mathbb{N}$ une fonction strictement croissante, i.e. $\varphi(n) > n$ pour tout $n \in \mathbb{N}$. Alors, la suite définie par $(x_{\varphi(n)})_{n\in\mathbb{N}}$ est dite* ***sous-suite****, ou* ***suite extraite****, de la suite initiale $(x_n)_{n\in\mathbb{N}}$.*

Comme simple exercice, le lecteur est invité à démontrer que, si une suite $(x_n)_{n\in\mathbb{N}}$ dans une espace métrique (X, d) est convergente, alors aussi toutes ses sous-suites le sont et elles convergent à la même limite.

L'important résultat suivant montre que, pour les suites de Cauchy, l'ordre de cette implication peut être reversé.

Théorème 4.3.5 *Toute suite de Cauchy d'un espace métrique (X, d) qui possède au moins une sous-suite convergente, est convergente elle-même et à la même limite.*

Preuve. Supposons que $(x_n)_{n\in\mathbb{N}}$ soit une suite de Cauchy en (X, d) qui admette une sous-suite convergente $(x_{\varphi(n)})_{n\in\mathbb{N}}$, où $\varphi : \mathbb{N} \to \mathbb{N}$ est l'application strictement croissante définissant cette sous-suite. Soit a la limite de la sous-suite, i.e. $a = \lim\limits_{n\to+\infty} x_{\varphi(n)}$.

Pour tout $n \in \mathbb{N}$, par la propriété triangulaire on a : $d(x_n, a) \leqslant d(x_n, x_{\varphi(n)}) + d(x_{\varphi(n)}, a)$, nous voulons démontrer qu'il est possible de majorer la somme de deux termes de droite avec une quantité ε petite comme l'on veut.

Pour faire ça, allons utiliser la définition de suite de Cauchy pour $(x_n)_{n\in\mathbb{N}}$ pour écrire :

$$\forall \varepsilon > 0\ \exists N_\varepsilon \in \mathbb{N} \text{ tel que : } m, n \geqslant N_\varepsilon \implies d(x_m, x_n) < \frac{\varepsilon}{2},$$

mais, comme φ est strictement croissante, $\varphi(n) > n \geqslant N_\varepsilon$, donc $d(x_n, x_{\varphi(n)}) < \frac{\varepsilon}{2}$.

Vu que la sous-suite $(x_{\varphi(n)})_{n\in\mathbb{N}}$ est supposée être convergente vers a, ça vaut que :

$$\forall \varepsilon > 0\ \exists K_\varepsilon \in \mathbb{N} \text{ tel que : } n \geqslant K_\varepsilon \implies d(x_{\varphi(n)}, a) < \frac{\varepsilon}{2},$$

mais alors ça suffit de considérer $n \geqslant \max\{N_\varepsilon, K_\varepsilon\}$ pour avoir $d(x_n, a) \leqslant d(x_n, x_{\varphi(n)}) + d(x_{\varphi(n)}, a) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ $\forall \varepsilon > 0$, i.e. $x_n \underset{n\to+\infty}{\longrightarrow} a$. □

Ce théorème a des applications remarquables dans la théorie de l'*optimisation*, où on recherche la solution optimale à un problème, réel ou abstrait, via la minimisation d'une fonction opportune. Quand cette fonction est trop compliquée pour permettre une description analytique de ses points de minima, nous sommes obligés d'approcher la solution avec un *algorithme itératif*, qui arrive à un point de minimum de la fonction en passant par *une suite de points*. Souvent, pour démontrer que l'algorithme itératif converge, nous utilisons le théorème 4.3.5 en démontrant que la suite de points définis par l'algorithme est de Cauchy et après en sélectionnant une opportune sous-suite, plus simple à analyser que celle initiale, que l'on démontre être convergente.

4.3.1 La complétude des espaces vectoriel

Pour un espace vectoriel métrique le fait de ne pas être complet n'est pas un problème si grave, car il peut toujours être complété d'une manière essentiellement unique, comme le dit le résultat suivant.

Théorème 4.3.6 (Complétion d'un espace vectoriel métrique non complet) *Si (V, d) est un espace vectoriel métrique non complet, alors il existe un espace vectoriel métrique complet $(\hat{V}, \hat{d})$ et une injection isométrique $\iota : V \longrightarrow \hat{V}$, i.e.*

$$\begin{cases} x_1, x_2 \in V,\ x_1 \neq x_2 \implies \iota(x_1) \neq \iota(x_2) \\ \forall x_1, x_2 \in V,\ d(x_1, x_2) = \hat{d}(\iota(x_1), \iota(x_2)) \end{cases}$$

tels que $\overline{\iota(V)} = \hat{V}$, i.e. l'image de V via ι est dense dans $\hat{V}$.

Corollaire 4.3.1 *Tout espace pré-hilbertien V peut être complété à un espace de Hilbert $\mathcal{H}$.*

Preuve. Nous allons montrer la preuve dans le cas d'intérêt pour nous, celui d'un espace pré-hilbertien, la preuve générale est conduite d'une manière analogue, en remplaçant la norme de la différence entre deux vecteurs par leur distance.

La complétion d'un espace pré-hilbertien V est, par définition, l'espace $\mathcal{H}_1$ de toutes les suites de Cauchy $(x_n)_{n\in\mathbb{N}}$ modulo la relation d'équivalence $\sim$ définie comme ça : deux suites de Cauchy $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ d'éléments de V sont équivalentes si $\lim_{n\to+\infty} \|x_n - y_n\| = 0$.

On écrit la complétion de V comme ça : $\mathcal{H} = \mathcal{H}_1/\sim$ et ses éléments comme ça : $[x]$. On va définir une norme sur $\mathcal{H}$ de la manière suivante :

$$\forall [x] \in \mathcal{H}, \quad \|[x]\| = \lim_{n\to+\infty} \|x_n\|,$$

où $(x_n)_{n\in\mathbb{N}}$ est une suite de Cauchy quelconque dans la classe d'équivalence $[x]$. Cette définition est indépendante du choix de la suite de Cauchy utilisée comme représentant pour la classe d'équivalence, vu que, si $(y_n)_{n\in\mathbb{N}} \in [x]$, alors, vu que $|\,\|x_n\| - \|y_n\|\,| \leqslant \|x_n - y_n\|$, si on passe à la limite on obtient :

$$\lim_{n\to+\infty} |\,\|x_n\| - \|y_n\|\,| \leqslant \lim_{n\to+\infty} \|x_n - y_n\| = 0,$$

i.e. $\lim_{n\to+\infty} \|x_n\| = \lim_{n\to+\infty} \|y_n\|$.

Maintenant on va définir un produit scalaire sur $\mathcal{H}$ compatible avec la norme ci-dessus :

$$\langle [x], [y] \rangle = \lim_{n \to +\infty} \langle x_n, y_n \rangle,$$

où $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ sont deux suites de Cauchy quelconques dans les classes d'équivalences $[x]$ et $[y]$, respectivement.

Pour vérifier que ce produit scalaire est bien défini, il faut vérifier que la limite qui le définit existe et qu'elle ne dépend pas des représentants choisis.

On commence par l'existence de la limite : pour faire ça, il suffit de démontrer que la suite $\langle x_n, y_n \rangle$ (suite en $\mathbb{K}$!) est de Cauchy, en fait, vu que $\mathbb{K}$ est complet, la limite existe. Observons que, $\forall n, m \in \mathbb{N}$, grâce à l'inégalité triangulaire et à celle de Cauchy-Schwarz on peut écrire :

$$\begin{aligned} |\langle x_n, y_n \rangle - \langle x_m, y_m \rangle| &= |\langle x_n, y_n \rangle - \langle x_n, y_m \rangle + \langle x_n, y_m \rangle - \langle x_m, y_m \rangle| \\ &= |\langle x_n, y_n - y_m \rangle + \langle x_n - x_m, y_m \rangle| \leqslant |\langle x_n, y_n - y_m \rangle| + |\langle x_n - x_m, y_m \rangle| \\ &\leqslant \|x_n\| \, \|y_n - y_m\| + \|x_n - x_m\| \, \|y_m\| \underset{n,m \to +\infty}{\longrightarrow} 0, \end{aligned}$$

car $\|x_n\|$ et $\|y_n\|$ sont bornées vu que $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ sont de Cauchy.

Maintenant on vérifie que la limite ne dépend pas des représentants choisis : soient $(\xi_n)_{n\in\mathbb{N}}$ et $(\eta_n)_{n\in\mathbb{N}}$ deux autres représentants des classes d'équivalences $[x]$ et $[y]$, respectivement.

Observons que, avec des manipulations algébriques directes, on peut écrire :

$$\langle x_n, y_n \rangle = \langle x_n - \xi_n + \xi_n, y_n - \eta_n + \eta_n \rangle = \langle x_n - \xi_n, y_n \rangle + \langle \xi_n, y_n - \eta_n \rangle + \langle \xi_n, \eta_n \rangle,$$

mais

$$|\langle x_n - \xi_n, y_n \rangle + \langle \xi_n, y_n - \eta_n \rangle| \leqslant \|x_n - \xi_n\| \, \|y_n\| + \|\xi_n\| \, \|y_n - \eta_n\| \underset{n,m \to +\infty}{\longrightarrow} 0$$

vu que $(x_n)_{n\in\mathbb{N}}, (\xi_n)_{n\in\mathbb{N}} \in [x]$ et que $(y_n)_{n\in\mathbb{N}}, (\eta_n)_{n\in\mathbb{N}} \in [y]$, et donc :

$$\langle [x], [y] \rangle = \lim_{n \to +\infty} \langle x_n, y_n \rangle = \lim_{n \to +\infty} \langle \xi_n, \eta_n \rangle.$$

Grâce à la continuité du produit scalaire sur V, toutes ses propriétés sont transportées par l'opération de limite sur $\mathcal{H}$.

Il reste tout simplement l'isométrie à vérifier :

$$\|[x]\| = \lim_{n \to +\infty} \|x_n\| = \lim_{n \to +\infty} \sqrt{\langle x_n, x_n \rangle} = \sqrt{\langle [x], [x] \rangle}, \qquad \forall [x] \in \mathcal{H}.$$

□

Une preuve alternative peut être consulté dans le livre [8].

4.3.2 Une caractérisation de la complétude d'espaces vectoriels normés via séries

Dans cette section nous allons examiner un critère de complétude pour les espaces vectoriels normés, très utile dans la pratique, qui fait appel aux séries.

Explicitement, la définition de condition de Cauchy pour la suite des sommes partielles d'une série $\sum_{n\in\mathbb{N}}$ est :

$$\forall \varepsilon > 0\ \exists N_\varepsilon > 0\ :\ n, m \geqslant N_\varepsilon\ \Rightarrow\ \left\| \sum_{k=0}^{n} x_k - \sum_{k=0}^{m} x_k \right\| < \varepsilon,$$

les deux indices n et m varient *indépendamment* l'un de l'autre et nous pouvons supposer, sans perdre de généralité, que l'un soit toujours supérieur à l'autre. Pour fixer les idées, supposons que $n > m$, alors, $\left\| \sum_{k=0}^{n} x_k - \sum_{k=0}^{m} x_k \right\| = \left\| \sum_{k=m+1}^{n} x_k \right\|$, ceci implique que la **condition de Cauchy pour les séries** peut être réécrite comme ça :

$$\forall \varepsilon > 0\ \exists N_\varepsilon > 0\ :\ n > m \geqslant N_\varepsilon\ \Rightarrow\ \left\| \sum_{k=m+1}^{n} x_k \right\| < \varepsilon. \tag{4.9}$$

Nous allons utiliser cette observation dans la preuve du résultat suivant.

Théorème 4.3.7 (Caractérisation complétude espaces normés via séries) *Un espace vectoriel normé $(V, \| \|)$ est complet si et seulement si toute série d'éléments de V absolument convergente est aussi (simplement) convergente en V.*

Preuve. La preuve de l'implication directe est, presque, un simple exercice d'écriture de la condition de Cauchy (4.9). La preuve de l'implication inverse, par contre, est bien plus compliquée.

$\boxed{\Rightarrow}$: Supposons que $(V, \| \|)$ soit complet et démontrons que si $\sum_{n\in\mathbb{N}} \|x_n\|$ est convergente, alors aussi $\sum_{n\in\mathbb{N}} x_n$ est convergente en V.

Par complétude, la convergence de $\sum_{n\in\mathbb{N}} \|x_n\|$ équivaut à la condition de Cauchy (4.9), i.e.

$$\forall \varepsilon > 0\ \exists N_\varepsilon > 0\ :\ n > m \geqslant N_\varepsilon\ \Longrightarrow\ \sum_{k=m+1}^{n} \|x_k\| < \varepsilon, \tag{4.10}$$

mais $\left\| \sum_{k=m+1}^{n} x_k \right\| \leqslant \sum_{k=m+1}^{n} \|x_k\|$, donc aussi la suite des sommes partielles $\left(S_n = \sum_{k=0}^{n} x_k \right)_{n\in\mathbb{N}}$ est de Cauchy, i.e. $\sum_{n\in\mathbb{N}} x_n$ est convergente.

$\boxed{\Leftarrow}$: supposons maintenant que toute série d'éléments de V absolument convergente soit aussi simplement convergente en V, il faut démontrer que ceci implique que V est complet, i.e. que toute suite de Cauchy $(x_n)_{n\in\mathbb{N}} \subset V$, i.e.

$$\forall \varepsilon > 0\ \exists N_\varepsilon > 0\ :\ n, m \geqslant N_\varepsilon\ \Longrightarrow\ \|x_n - x_m\| < \varepsilon$$

converge en V, i.e. qu'il existe $\bar{x} \in V$ tel que $(x_n)_{n\in\mathbb{N}} \underset{n\to+\infty}{\to} \bar{x}$.

La condition de Cauchy doit être valide pour toute valeur de $\varepsilon > 0$, et donc, aussi pour $\varepsilon_k = \frac{1}{2^k}$, $k \in \mathbb{N}$, par conséquent, toute suite de Cauchy de V vérifie

$$\forall k \geqslant 0\ \exists \tilde{N}_k > 0\ :\ n, m \geqslant \tilde{N}_k\ \Longrightarrow\ \|x_n - x_m\| < \frac{1}{2^k}. \tag{4.11}$$

Vu que $\frac{1}{2^{k+1}} < \frac{1}{2^k}$, sûrement $\tilde{N}_{k+1} \geqslant \tilde{N}_k$, mais alors il suffit de définir $N_k = \inf_{\ell \in \mathbb{N}} \{\tilde{N}_{k+\ell} > \tilde{N}_k\}$ pour tout $k \in \mathbb{N}$ pour avoir une suite de nombres naturels strictement croissante $(N_k)_{k \geqslant 0}$. Grâce à ça, nous allons définir la suite extraite $(x_{N_k})_{k \in \mathbb{N}} \subset V$ de $(x_n)_{n \in \mathbb{N}}$ qui vérifie :

$$\forall k \geqslant 0, \ \left\| x_{N_k} - x_{N_{k+1}} \right\| < \frac{1}{2^k}. \tag{4.12}$$

L'intérêt de considérer cette sous-suite est que, si elle converge en V, i.e. s'il existe $\bar{x} \in V$ tel que $\lim_{k \to +\infty} x_{N_k} = \bar{x}$, alors, grâce au théorème 4.3.5, aussi la suite initiale $(x_n)_{n \in \mathbb{N}}$ converge vers $\bar{x} \in V$.

Pour terminer la preuve, donc, il faut démontrer que la suite extraite $(x_{N_k})_{k \in \mathbb{N}}$ est convergente en V. Comme on n'a aucune information sur la convergence de la suite originale $(x_n)_{n \in \mathbb{N}}$, on ne peut pas penser de démontrer directement la convergence de $(x_{N_k})_{k \in \mathbb{N}}$, il faut nécessairement passer par une astuce qui fait entrer en jeu l'hypothèse que les séries absolument convergentes en V impliquent la convergence simple des séries en V.

Le **lien avec les séries** est obtenu via une technique géniale dans sa simplicité : **la réécriture de la suite extraite $(x_{N_k})_{k \in \mathbb{N}}$ comme une suite de sommes partielles télescopiques**. Pour faire ça on utilise $(x_{N_k})_{k \in \mathbb{N}}$ pour définir une nouvelle suite $(y_k)_{k \in \mathbb{N}} \subset V$ de la manière suivante :

$$\begin{cases} y_0 = x_{N_0} \\ y_k = x_{N_k} - x_{N_{k-1}}, \ \forall k \geqslant 1, \end{cases} \qquad \text{donc : } (y_k)_{k \in \mathbb{N}} = (x_{N_0}, x_{N_1} - x_{N_0}, x_{N_2} - x_{N_1}, \ldots),$$

alors

$$\sum_{j=0}^{k} y_j = \underbrace{\cancel{x_{N_0}}}_{y_0} + \underbrace{\cancel{x_{N_1}} - \cancel{x_{N_0}}}_{y_1} + \underbrace{\cancel{x_{N_2}} - \cancel{x_{N_1}}}_{y_2} + \ldots + \underbrace{x_{N_k} - \cancel{x_{N_{k-1}}}}_{y_k} = x_{N_k},$$

et ça vaut $\forall k \in \mathbb{N}$, donc $\left(\sum_{j=0}^{k} y_j \right)_{k \in \mathbb{N}} = (x_{N_k})_{k \in \mathbb{N}}$.

En résumé, la complétude de V, i.e. la convergence en V d'une suite de Cauchy arbitraire $(x_n)_{n \in \mathbb{N}}$, est impliqué par la convergence en V de la suite $(x_{N_k})_{k \in \mathbb{N}}$; celle-ci est équivalente à la convergence en V de $\left(\sum_{j=0}^{k} y_j \right)_{k \in \mathbb{N}}$, i.e. à la convergence simple de la série $\sum_{k=0}^{\infty} y_k$. Grâce à l'hypothèse de départ, si on prouve que $\sum_{k \in \mathbb{N}} \|y_k\|$ est convergente, le théorème sera complètement démontré. Pour faire ça, allons expliciter les termes de la série $\sum_{k \in \mathbb{N}} \|y_k\|$:

$$\sum_{k \in \mathbb{N}} \|y_k\| = \|y_0\| + \sum_{k=1}^{\infty} \|y_k\| = \|y_0\| + \sum_{k=1}^{\infty} \left\| x_{N_k} - x_{N_{k-1}} \right\| = \|y_0\| + \sum_{k=0}^{\infty} \left\| x_{N_{k+1}} - x_{N_k} \right\|,$$

mais, grâce à l'inégalité (4.12), ça vaut $\left\| x_{N_{k+1}} - x_{N_k} \right\| < \frac{1}{2^k} = \left(\frac{1}{2} \right)^k \ \forall k \geqslant 0$, donc :

$$\sum_{k \in \mathbb{N}} \|y_k\| = \|y_0\| + \sum_{k=0}^{\infty} \left\| x_{N_{k+1}} - x_{N_k} \right\| < \|y_0\| + \sum_{k=0}^{\infty} \left(\frac{1}{2} \right)^k = \|y_0\| + \frac{1}{1 - \frac{1}{2}} = \|y_0\| + 2 < +\infty,$$

où on a utilisé la formule da la somme de la série géométrique.

Donc, $\sum_{k\in\mathbb{N}} \|y_k\|$ est une série à termes réels positifs et bornée, et alors elle converge grâce à un résultat classique de la théorie des séries. □

Si, dans le théorème précédent, $(V, \| \ \|)$ s'avère être complet, alors la suite de Cauchy $(x_n)_{n\in\mathbb{N}} \subset V$ considéré au début de la preuve est convergente et ainsi le sera la sous-suite $(x_{N_k})_{k\in\mathbb{N}}$ et, bien sûr, vers la même limite. Nous formalisons ce résultat important dans le corollaire suivant.

Corollaire 4.3.2 *Soit $(V, \| \ \|)$ un espace vectoriel normé complet et soit $(x_n)_{n\in\mathbb{N}} \subset V$ une suite convergente vers $x_0 \in V$, alors il existe une suite extraite $(x_{n_k})_{k\in\mathbb{N}}$ qui converge vers x_0.*

L'exponentiel matriciel

Nous allons examiner ici un exemple important d'application du théorème de caractérisation des espaces vectoriels normés complets : la définition de l'exponentiel matriciel.

Déf. 4.3.5 (Exponentiel d'une matrice) *Soit $A \in M(n, \mathbb{K})$ une matrice carrée*[5] *avec des coefficients dans le corps $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Alors l'exponentiel de A est la matrice définie comme ça :*

$$\boxed{e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}}.$$

La bonne définition de e^A est très simple à démontrer si nous utilisons le théorème que nous venons de prouver. Considérons, par exemple, la **norme de Frobenius** de A :

$$\|A\| = \left(\sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2},$$

elle est la norme Euclidienne du vecteur de $\mathbb{K}^{n^2}$ obtenu par A via ordre lexicographique, i.e. en mettant en séquence l'une après l'autre, les lignes de A. On va prouver que la série $I_n + A + \frac{A}{2} + \frac{A^2}{3!} + \frac{A^k}{k!} + \dots$ est convergente dans la topologie de $\mathrm{M}(n, \mathbb{K})$ engendrée par cette norme et donc elle est convergente par rapport à toute autre norme, grâce au théorème de Tychonov.

$\mathrm{M}(n, \mathbb{K})$ est homéomorphe à $\mathbb{K}^{n^2}$ Euclidien, que l'on sait être un espace normé complet, donc, pour montrer que e^A est bien défini, ça suffit de montrer que la série qui définit e^A converge absolument, vu que le théorème 4.3.7 implique la convergence simple.

La preuve de la convergence absolue est très simple : il suffit de considérer l'inégalité suivante $\|A^k\| \leqslant \|A\|^k$, vérifiée par la norme de Frobenius pour tout $k \in \mathbb{N}$ et pour toute matrice $A \in \mathrm{M}(n, \mathbb{K})$, pour écrire

$$\sum_{k=0}^{\infty} \frac{\|A^k\|}{k!} \leqslant \sum_{k=0}^{\infty} \frac{\|A\|^k}{k!} = e^{\|A\|},$$

où on a utilisé le fait que $\|A\|$ est un nombre réel $\geqslant 0$ et que le rayon de convergence de la série exponentielle en $\mathbb{R}$ est infini.

5. L'exigence que A soit carrée vient du fait qu'on doit considérer des puissances de A, qui ne sont pas définies, pour des raisons dimensionnelles, si A n'est pas carrée.

4.3.3 Le théorème de point fixe de Banach

Le résultat que l'on va discuter dans cette section est de grande importance pour beaucoup de champs différents des mathématiques (analyse, topologie, équations différentielles, etc.).

Tout d'abord il faut rappeler la définition suivante.

Déf. 4.3.6 (Contraction) *Soient (X_1, d_1) et (X_2, d_2) deux espaces métriques quelconque et soit $k \in]0, 1[$ une constante réelle. L'application $f : X_1 \to X_2$ est dite être une contraction de rapport k si, pour tout $x, y \in X_1$, ça vaut que :*

$$d_2(f(x), f(y)) \leqslant k d_1(x, y). \tag{4.13}$$

*La plus petite valeur de k pour laquelle ça vaut (4.13) est dite être la **constante de Lipschitz** de f.*

Il est immédiat de vérifier qu'**une contraction est toujours une fonction continue** : pour toute valeur $\varepsilon > 0$, fixons un élément arbitraire $\bar{x} \in X_1$ et considérons les éléments $y \in X_1$ tels que $d_1(\bar{x}, y) < \varepsilon$, alors, par définition de contraction, $d_2(f(\bar{x}), f(y)) \leqslant k d_1(\bar{x}, y) < k\varepsilon < \varepsilon$, car $k \in]0, 1[$, donc la fonction f est continue en $\bar{x}$. Comme $\bar{x}$ est un élément arbitraire de X_1, f est continue sur tout X_1.

Remarque 4.3.1 À partir de la définition, il suit que la distance (dans le codomaine) des images d'un couple d'éléments via une contraction est inférieure à leur distance initiale. Néanmoins, ceci ne veut pas dire que nous pouvons redéfinir une contraction en demandant seulement cette propriété, i.e. la définition ci-dessus n'est pas équivalente à demander que, pour tout $x, y \in X_1$, $x \neq y$, $d_2(f(x), f(y)) < d_1(x, y)$, en fait, si f satisfait cette requête, nous disons qu'elle est une *contraction au sens faible*, ou une *application qui réduit les distances entre points*.

Pour comprendre la différence, subtile, entre le deux définitions, observons que si f est une contraction au sens faible, alors, pour toute couple $x, y \in X_1$, $x \neq y$, il existe $k_{x,y} \in]0, 1[$ tel que $d_2(f(x), f(y)) \leqslant k_{x,y}\, d_1(x, y)$, i.e. $k_{x,y}$ n'est pas une constante, comme requis par la définition de contraction. Les deux définitions coïncident si et seulement si $\sup\limits_{x,y \in X_1} k_{x,y} \equiv \bar{k} \in]0, 1[$, néanmoins, cette condition n'est pas forcement vérifiée : le sup existe sûrement car $\{k_{x,y},\ x, y \in X_1,\ x \neq y\}$ est un sous-ensemble borné de $\mathbb{R}$, mais il peut être 1 et non pas strictement inférieur à 1, comme demandé dans la définition de contraction.

Parmi toutes les contractions, celles qui ont domaine et image dans un espace métrique complet ont la propriété splendide décrite dans le théorème classique suivant.

Théorème 4.3.8 (Théorème de point fixe de Banach) *Soit (X, d) un espace métrique complet et $f : X \to X$ une contraction en X de rapport $k \in]0, 1[$, alors f admet un seul point fixe, i.e. il existe un seul $\bar{x} \in X$ tel que $f(\bar{x}) = \bar{x}$.*

Preuve. Soit $a \in X$ un élément *quelconque*. Définissons la suite $(x_n)_{n \in \mathbb{N}} \subset X$ définie par récurrence comme ça :

$$\begin{cases} x_0 = a \\ x_n = f(x_{n-1}),\ n \geqslant 1. \end{cases}$$

Nous allons d'abord montrer très simplement que si cette suite admet une limite en X, alors cette limite est un point fixe pour f. L'unicité du point fixe sera une conséquence banale de la définition de contraction. Nous allons vérifier à la fin que, en effet, la suite $(x_n)_{n\in\mathbb{N}}$ est convergente, ce qui est plus compliqué à prouver.

— *S'il existe* $X \ni \bar{x} = \lim\limits_{n\to+\infty} x_n$*, alors* $\bar{x}$ *est un point fixe pour* f : la preuve de cette affirmation consiste en utiliser un simple argument de continuité. En fait, en ayant déjà remarqué qu'une contraction est continue, si nous faisons tendre n vers $+\infty$ dans la définition de la suite, i.e. $x_n = f(x_{n-1})$, nous avons :

$$\lim_{n\to+\infty} x_n = \lim_{n\to+\infty} f(x_{n-1}) \iff \bar{x} = f(\lim_{n\to+\infty} x_{n-1}) \iff \bar{x} = f(\bar{x}),$$

i.e. $\bar{x}$ est un point fixe pour f, ce qui explique l'intérêt de considérer la suite définie par récurrence $(x_n)_{n\in\mathbb{N}}$ qu'on a considéré ci-dessus.

— *Unicité du point fixe* : soient $\bar{x}, \bar{y} \in X$ deux point fixes pour f, i.e. $f(\bar{x}) = \bar{x}$, $f(\bar{y}) = \bar{y}$, allons démontrer que leur distance est nulle, i.e. que $\bar{x} = \bar{y}$, grâce à la définie positivité de la distance, en utilisant la définition de contraction :

$$d(\bar{x}, \bar{y}) = d(f(\bar{x}), f(\bar{y})) \leqslant kd(\bar{x}, \bar{y}),$$

mais $k \in]0, 1[$, donc cette inégalité à du sens si et seulement si $d(\bar{x}, \bar{y}) = 0$, i.e. $\bar{x} = \bar{y}$.

— *Convergence de la suite* : ici c'est où nous utilisons l'hypothèse que (X, d) soit complet, en fait, grâce à ça, si nous démontrons que $(x_n)_{n\in\mathbb{N}}$ est une suite de Cauchy, alors elle sera convergente. Pour cela, nous observons tout d'abord que, pour tout $n \geqslant 1$ la définition de la suite et l'hypothèse que f soit une contraction permettent d'écrire

$$d(x_{n+1}, x_n) = d(f(x_n), f(x_{n-1})) \leqslant kd(x_n, x_{n-1}),$$

d'où, par itération,

$$\begin{aligned} d(x_{n+1}, x_n) &\leqslant kd(x_n, x_{n-1}) \\ &\leqslant k^2 d(x_{n-1}, x_{n-2}) \\ &\leqslant \ldots \leqslant k^n d(x_1, x_0), \end{aligned} \tag{4.14}$$

i.e. la distance entre deux éléments consécutifs, x_{n+1} et x_n, de la suite $(x_n)_{n\in\mathbb{N}}$ est majoré par $k^n d(x_1, x_0)$, à noter que *la puissance de k est égale à l'indice le plus petit.* Soient maintenant $n, m \in \mathbb{N}$ deux indices naturels arbitraires mais différents. Sans perdre en généralité, nous pouvons imaginer que $n < m$, alors $m - n = p \in \mathbb{N}$ et on peut écrire $m = n + p$ de manière que $d(x_m, x_n) = d(x_{n+p}, x_n)$. En itérant la propriété triangulaire de la distance, nous avons :

$$d(x_{n+p}, x_n) \leqslant d(x_{n+p}, x_{n+p-1}) + d(x_{n+p-1}, x_{n+p-2}) + \ldots + d(x_{n+1}, x_n).$$

Nous observons que tous les termes du côté droit de l'inégalité sont des distances entre deux éléments consécutifs de la suite $(x_n)_{n\in\mathbb{N}}$, grâce à ça, nous pouvons appliquer la majoration donnée par (4.14) et écrire :

$$\begin{cases} d(x_{n+p}, x_{n+p-1}) \leqslant k^{n+p-1} d(x_1, x_0) \\ d(x_{n+p-1}, x_{n+p-2}) \leqslant k^{n+p-2} d(x_1, x_0) \\ \qquad\vdots \\ d(x_{n+1}, x_n) \leqslant k^n d(x_1, x_0) \end{cases}$$

i.e.

$$\begin{aligned} d(x_{n+p}, x_n) &\leqslant (k^{n+p-1} + k^{n+p-2} + \ldots + k^n) d(x_1, x_0) \\ &= (k^{p-1} + k^{p-2} + \ldots + 1) k^n d(x_1, x_0) \\ &= \left(\sum_{j=0}^{p-1} k^j \right) k^n d(x_1, x_0) \\ &\underset{k^j > 0}{\leqslant} \left(\sum_{j=0}^{+\infty} k^j \right) k^n d(x_1, x_0), \end{aligned}$$

mais $\sum_{j=0}^{+\infty} k^j$ est une série géométrique de raison $k \in]0, 1[$ et donc elle converge vers $\frac{1}{1-k}$, par conséquent

$$d(x_{n+p}, x_n) \leqslant \frac{k^n}{1-k} d(x_1, x_0).$$

En rappelant que $d(x_{n+p}, x_n) = d(x_m, x_n)$, $m > n \in \mathbb{N}$ arbitraires, nous avons que

$$d(x_m, x_n) \leqslant \frac{k^n}{1-k} d(x_1, x_0) \underset{n \to +\infty}{\longrightarrow} 0,$$

ce qui implique que $(x_n)_{n \in \mathbb{N}}$ est une suite de Cauchy et alors elle converge à un élément $\bar{x} \in X$ par l'hypothèse de complétude de X. □

Nous insistons sur le fait que *le premier élément de la suite* $(x_n)_{n \in \mathbb{N}}$ *est totalement arbitraire* : même s'il est choisi très loin du point fixe $\bar{x}$, la suite arrivera au point fixe par la limite. Évidemment, quand il est possible, un choix judicieux du point de départ x_0 permet le plus souvent d'accélérer la vitesse de convergence de la suite.

Voyons un exercice très intéressant proposé dans [13].

Exercice 4.1

1. Donner un exemple d'espace métrique (X, d) et de contraction $f : X \to X$ n'ayant aucun point fixe.
2. Donner un exemple d'espace métrique (X, d) complet et d'application $f : X \to X$ qui réduit strictement les distances, i.e. telle que $d(f(x), f(y)) < d(x, y)$ $\forall x, y \in X$, $x \neq y$, et qui n'admet aucun point fixe.
3. Montrer que le problème de Cauchy

$$\begin{cases} x'(t) = \frac{1}{2} \sin x(t) \\ x(0) = 1, \end{cases} \tag{4.15}$$

possède une unique solution $\varphi : [-1, 1] \to \mathbb{R}$.

Solution de l'exercice 4.1

Pour le point 1. et 2. il s'agit, évidemment, de mettre en défaut le théorème de point fixe en enlevant une hypothèse : en 1. nous considérons un espace métrique non complet, en 2. nous considérons une application qui réduit strictement les distances, que, comme nous l'avons vu, est une hypothèse moins stricte par rapport à demander que l'application soit une contraction.

1. Il faut évidemment considérer un espace métrique non-complet. Nous avons déjà vu dans le cours que $(]0,1[, |\ |)$ n'est pas complet. Dans cet espace métrique considérons, par exemple, la fonction $f :]0,1[\to]0,1[$, $f(x) = \frac{1}{2}x$, alors :

$$\forall x, y \in]0,1[, \quad |f(x) - f(y)| = \left|\frac{1}{2}x - \frac{1}{2}y\right| = \frac{1}{2}|x-y| \leqslant \frac{1}{2}|x-y|,$$

donc f est une contraction de rapport $k = 1/2$. Néanmoins, l'équation de point fixe pour f, i.e. $f(x) = x$, n'a évidemment pas de solutions en $]0,1[$ car $\frac{1}{2}x = x$ si et seulement si $x = 0 \notin]0,1[$.

2. Considérons l'espace métrique $(X,d) = ([0,+\infty[, |\ |)$ et l'application $f : [0,+\infty[\to [0,+\infty[$ définie par : $f(x) = \sqrt{x^2+1}$. Si on fixe deux éléments arbitraires $x, y \in [0,+\infty[$, alors, grâce au théorème des accroissements finis, il existe un élément $\xi \in [0,+\infty[$ strictement inclut entre x et y, tel que :

$$f(x) - f(y) = (x-y)f'(\xi) = (x-y)\frac{\xi}{\sqrt{\xi^2+1}}.$$

Comme $\sqrt{\xi^2+1} > \sqrt{\xi^2} = \xi \in [0,+\infty[$, alors $\frac{\xi}{\sqrt{\xi^2+1}} < 1$, i.e. $|f(x) - f(y)| < |x-y|$, c'est-à-dire, f décroît strictement les distances.
Néanmoins, en $[0,+\infty[$ l'équation de point fixe pour f, i.e. $x = \sqrt{x^2+1}$, peut être écrite comme $x^2 = x^2 + 1$, i.e. $1 = 0$, qui n'a évidemment pas de solutions, donc f n'admet pas de point fixe.

3. De la théorie des équations différentielles il est connu que résoudre le problème de Cauchy (4.15) est équivalent à déterminer une fonction $\varphi \in \mathcal{C}([-1,1])$ qui satisfait l'équation intégrale de Volterra suivant :

$$\varphi(t) = \frac{1}{2}\int_0^t \sin\varphi(s)ds + 1. \tag{4.16}$$

Il est instructif de vérifier cela. D'un côté, si φ est une solution de (4.15) alors, par définition, φ est différentiable, et donc continue. En intégrant de 0 à t les deux côtés de l'équation différentielle on obtient $\int_0^t \varphi'(s)ds = \frac{1}{2}\int_0^t \sin\varphi(s)ds$, i.e. $\varphi(t) - \varphi(0) = \frac{1}{2}\int_0^t \sin\varphi(s)ds$. La condition initiale (4.15) donne $\varphi(0) = 1$ et alors φ satisfait $\varphi(t) = \frac{1}{2}\int_0^t \sin\varphi(s)ds + 1$.
De l'autre côté, supposons que φ satisfait (4.16), alors la fonction intégrale $\frac{1}{2}\int_0^t \sin\varphi(s)ds + 1$ est dérivable $\forall t \in [-1,1]$, car $\sin\circ\varphi$ est continue et l'opération d'intégration rend dérivable une fonction continue. En dérivant (4.16) nous obtenons $\varphi'(t) = \frac{1}{2}\sin\varphi(t)$ avec $\varphi(0) = 1$, i.e. φ satisfait (4.15).

Ce qu'on vient de vérifier, motive l'intérêt vers l'espace $\mathcal{C}([-1,1])$, qui est un espace de Banach quand on le muni de la norme $\|f\| = \sup_{t\in[-1,1]} |f(t)|$. Considérons l'application suivante :

$$\begin{array}{rccc} F: & \mathcal{C}([-1,1]) & \longrightarrow & \mathcal{C}([-1,1]) \\ & f & \longmapsto & F(f), \end{array}$$

où $F(f)$ est la fonction continue sur $[-1,1]$ à valeurs réelles définie par l'expression analytique $F(f)(t) = \frac{1}{2}\int_0^t \sin\varphi(s)ds + 1$, pour tout $t \in [-1,1]$. Il est clair que, si nous arrivons à démontrer que F est une contraction, alors la preuve du fait qu'il existe une seule solution du problème de Cauchy (4.15) sera terminée.
Pour cela, considérons deux fonctions $f, g \in \mathcal{C}([-1,1])$ quelconques et $t \in [-1,1]$ arbitraire, alors :

$$\begin{aligned}
|F(f)(t) - F(g)(t)| &= \frac{1}{2}\left|\int_0^t [\sin f(s) - \sin g(s)]ds\right| \\
&\leqslant \frac{1}{2}\left|\int_0^t |\sin f(s) - \sin g(s)|ds\right| \\
&\text{(en utilisant la formule } \sin p - \sin q = 2\sin\left(\frac{p-q}{2}\right)\cos\left(\frac{p+q}{2}\right)) \\
&= \left|\int_0^t \left|\sin\frac{f(s)-g(s)}{2}\ \cos\frac{f(s)+g(s)}{2}\right| ds\right| \\
&(|\cos(\alpha)| \leqslant 1) \\
&\leqslant \left|\int_0^t \left|\sin\frac{f(s)-g(s)}{2}\right| ds\right| \\
&(|\sin(\alpha)| \leqslant |\alpha|) \\
&\leqslant \left|\int_0^t \left|\frac{f(s)-g(s)}{2}\right| ds\right| \\
&\leqslant \frac{1}{2}\left|\int_0^t \|f-g\|ds\right| \\
&\leqslant \frac{\|f-g\|}{2}\left|\int_0^t ds\right| = \frac{\|f-g\|}{2}|t| \\
&(t \in [-1,1] \implies |t| \leqslant 1) \\
&\leqslant \frac{\|f-g\|}{2}.
\end{aligned}$$

En résumé : $|F(f)(t) - F(g)(t)| \leqslant \frac{\|f-g\|}{2}\ \forall t \in [-1,1]$, donc :

$$\|F(f) - F(g)\| = \sup_{t\in[-1,1]} |F(f)(t) - F(g)(t)| \leqslant \frac{1}{2}\|f-g\|,$$

i.e., F est effectivement une contraction et la thèse de l'exercice est prouvée. □

4.4 Exemples remarquables d'espaces de Banach et de Hilbert

Dans cette section on va introduire des espaces fonctionnels d'importante fondamentale en mathématique. En particulier, on démontrera que certains de ces espaces sont de Banach et certains autres sont de Hilbert. Une fois fait, on énoncera des théorèmes de densité parmi ces espaces.

4.4.1 Les espaces L^p et ℓ^p et leur complétude

Dans les définitions suivantes $\mathbb{K}$ sera soit $\mathbb{R}$, soit $\mathbb{C}$. Soit $(X, \mathcal{A}, \mu)$ un espace mesuré. Pour tout $1 \leqslant p < +\infty$, on définit :

$$\boxed{\mathcal{L}^p(X, \mathcal{A}, \mu) = \left\{ f : X \to \mathbb{K},\ f \text{ mesurable} \ : \ \int_X |f|^p d\mu < +\infty \right\}}.$$

L'ensemble $\mathcal{L}^p(X, \mathcal{A}, \mu)$ devient un espace vectoriel si on définit la **structure vectorielle ponctuelle**, i.e. $\forall \alpha, \beta \in \mathbb{K}, \forall f, g \in \mathcal{L}^p(X, \mathcal{A}, \mu)$:

$$\begin{array}{rccl} \alpha f + \beta g : & X & \to & \mathbb{K} \\ & x & \mapsto & (\alpha f + \beta g)(x) = \alpha f(x) + \beta g(x). \end{array}$$

Cette opération de combinaison linéaire est bien définie grâce à la célèbre **inégalité de Minkowski**[6] pour les intégrales (dont on assumera la preuve) :

$$\boxed{\left(\int_X |f+g|^p d\mu \right)^{1/p} \leqslant \left(\int_X |f|^p d\mu \right)^{1/p} + \left(\int_X |g|^p d\mu \right)^{1/p}}, \tag{4.17}$$

de plus, la multiplication par les scalaires α, β n'affecte bien sûr pas l'intégrabilité de f, g et donc la définition est cohérente.

Si on écrit :

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

alors les propriétés de l'intégrale de Lebesgue donnent :

— Positivité (non définie) et homogénéité :

$$\|f\|_p \geqslant 0,\ \|\lambda f\|_p = |\lambda| \|f\|_p \quad \forall f \in \mathcal{L}^p(X, A, \mu),\ \lambda \in \mathbb{K}.$$

— **L'inégalité de Minkowski (4.17) devient l'inégalité triangulaire**[7] **pour** $\| \ \|_p$:

$$\|f+g\|_p \leqslant \|f\|_p + \|g\|_p, \quad \forall f, g \in \mathcal{L}^p(X, A, \mu).$$

— Cependant,

$$\|f\|_p = 0 \not\Longrightarrow f = 0 \text{ (la fonction nulle)},$$

en fait toute fonction $g \in \mathcal{L}^p(X, A, \mu)$ nulle p.p. est telle que $\|g\|_p = 0$ et donc le fait d'avoir une valeur nulle de $\| \ \|_p$ n'implique pas que d'être la fonction nulle partout $f(x) = 0\ \forall x \in X$.

6. Par itération, on arrive à pouvoir écrire l'**inégalité de Minkowski généralisée**, qu'on utilisera dans la suite :

$$\left(\int_X \left| \sum_{k=1}^n f_k \right|^p d\mu \right)^{1/p} \leqslant \sum_{k=1}^n \left(\int_X |f_k|^p d\mu \right)^{1/p}.$$

7. Par itération :

$$\left\| \sum_{k=1}^n f_k \right\|_p \leqslant \sum_{k=1}^n \|f_k\|_p. \tag{4.18}$$

Donc, $\| \ \|_p$ **est une semi-norme sur** $\mathcal{L}^p(X, \mathcal{A}, \mu)$. Le fait d'avoir une semi-norme ne permet pas d'utiliser la propriété suivante (très utile dans la pratique) : si la norme de la différence entre deux éléments d'une espace normé est nulle, alors les deux éléments coïncident, vu qu'ils peuvent différer sur un ensemble de mesure nulle. La stratégie de passer par la norme pour montrer l'unicité d'un objet mathématique est utilisée quand c'est difficile de prouver l'unicité en travaillant directement avec les éléments de l'espace normé.

La solution au problème consiste à opérer le quotient sur un sous-espace opportunément choisi, i.e.

$$\mathcal{N} = \{f : X \to \mathbb{K},\ f \text{ mesurable } \ : \ f = 0 \ \text{ p.p}\},$$

l'espace quotient :

$$\boxed{L^p(X, \mathcal{A}, \mu) = \mathcal{L}^p(X, \mathcal{A}, \mu)/\mathcal{N}},$$

formé par les classes d'équivalence de fonctions mesurables sur X, absolument intégrables en puissance p et égales p.p, est donc un **espace vectoriel normé avec norme** $\| \ \|_p$.

Grâce aux considérations de l'annexe A, il est immédiat de voir que, fixé un représentant f d'une classe d'équivalence de $L^p(X, \mathcal{A}, \mu)$, toutes les autres fonctions g qui appartiennent à la même classe peuvent être écrites comme ça : $g = f + h$, où $h : X \to \mathbb{K}$ est nulle p.p.

Par simplicité de notation il est usuel d'identifier une fonction représentante et la classe d'équivalence à laquelle elle appartient en les écrivant avec le même symbole. De plus, quand ce n'est pas important de spécifier X, $\mathcal{A}$ et μ on écrira tout simplement L^p.

Remarque 4.4.1 Soit $X \subseteq \mathbb{K}^n$ avec la mesure de Lebesgue. Considérons deux fonctions $f, g \in \mathcal{L}^p(X, \mathcal{A}, \mu)$ *continues* sur X et différentes au moins dans le point $x_0 \in X : f(x_0) \neq g(x_0)$. Par définition de continuité :

$$\forall \varepsilon > 0 \ \exists \delta_\varepsilon > 0 \ : \ x \in U_{\delta_\varepsilon}(x_0) \implies f(x) \in U_\varepsilon(f(x_0)) \text{ et } g(x) \in U_\varepsilon(g(x_0))$$

mais, par la propriété de séparabilité de $\mathbb{K}^n$, $\exists \varepsilon > 0$ tel que $U_\varepsilon(f(x_0)) \cap U_\varepsilon(g(x_0)) = \varnothing$, i.e. $\exists \delta_\varepsilon > 0$ tel que $x \in U_{\delta_\varepsilon}(x_0)$ implique $f(x) \neq g(x)$, i.e. si deux fonctions f et g, continues, sont différentes en un point x_0, alors elles sont différentes aussi sur un voisinage $U_{\delta_\varepsilon}(x_0)$ de rayon $\delta_\varepsilon > 0$. Ce voisinage a *une mesure de Lebesgue non nulle*, et donc les deux fonctions ne sont pas égales p.p.

Autrement dit : **deux fonctions continues sur** $X \subset \mathbb{K}^n$ **ne peuvent pas être égales p.p.** : soit elles sont la même fonction, soit elles sont différentes sur un ensemble de mesure de Lebesgue non nulle. Donc, **deux fonctions continues de** $\mathcal{L}^p(X, \mathcal{A}, \mu)$ **qui sont différentes au moins dans un point sont deux éléments différents de** $L^p(X, \mathcal{A}, \mu)$, car elles sont des représentants de deux classes d'équivalence différentes.

Si $p = 2$, alors on peut définir un produit scalaire sur $L^2(X, \mathcal{A}, \mu)$:

$$\boxed{\langle f, g \rangle = \int_X fg \ d\mu} \quad \text{si } \ \mathbb{K} = \mathbb{R} \quad \text{ et } \quad \boxed{\langle f, g \rangle = \int_X f\overline{g} \ d\mu} \quad \text{si } \mathbb{K} = \mathbb{C}.$$

Ces produits scalaires sont bien définis grâce à la célèbre **inégalité de Hölder pour les intégrales** (dont on assumera la preuve) : si $p, q > 0$ sont des *exposantes conjugués*, i.e. $\frac{1}{p} + \frac{1}{q} = 1$, alors ça vaut que

$$\boxed{\int_X |fg| \ d\mu \leqslant \left(\int_X |f|^p \ d\mu \right)^{1/p} \left(\int_X |g|^q \ d\mu \right)^{1/q}}, \tag{4.19}$$

évidemment $p = q = 2$ sont deux nombres conjugués et alors le produit scalaire introduit ci-dessus est bien défini. La preuve qu'il vérifie les axiomes de produit scalaire est laissé comme simple exercice, on remarque seulement que l'inégalité de Hölder pour $p = q = 2$ implique la validité de *l'inégalité de Cauchy-Schwarz pour l'espace* $L^2(X, \mathcal{A}, \mu)$, en fait, pour toute $f, g \in L^2(X, \mathcal{A}, \mu)$:

$$|\langle f, g\rangle_{L^2}| = \left|\int_X f\bar{g}\, d\mu\right| \leqslant \int_X |fg|\, d\mu \underset{(4.19)}{\leqslant} \left(\int_X |f|^2\, d\mu\right)^{1/2} \left(\int_X |g|^2\, d\mu\right)^{1/2} = \|f\|_2 \|g\|_2.$$

Un cas particulier des espaces L^p est celui des espaces ℓ^p, qui sont définis avec les choix suivants :

— X : un ensemble dénombrable, typiquement $X = \mathbb{N}$ ou $X = \mathbb{Z}$;

— $\mathcal{A} = \mathcal{P}(X)$, l'ensemble des parties de X ;

— μ : **mesure de comptage**, i.e. $\mu : \mathcal{P}(X) \to [0, +\infty]$, $\mu(A) = \text{card}(A)$ $\forall A \in \mathcal{P}(X)$ qui a un cardinal fini et $\mu(A) = +\infty$ si $\text{card}(A)$ n'est pas fini.

Avec ces choix, toute fonction $f : X \to \mathbb{K}$, est mesurable et identifiable avec une suite d'éléments de $\mathbb{K}$, qu'on écrira $(x_n)_{n\in\mathbb{N}}$. Donc, explicitement [8] :

$$\boxed{\ell^p(\mathbb{N}, \mathbb{K}) = \left\{(x_n)_{n\in\mathbb{N}},\ \sum_{n\in\mathbb{N}} |x_n|^p < +\infty\right\}}$$

et pareil si on échange $\mathbb{N}$ avec $\mathbb{Z}$.

Quand ce n'est pas important de spécifier $\mathbb{N}$ ou $\mathbb{Z}$ ou un autre ensemble dénombrable, on écrira tout simplement ℓ^p. La structure linéaire de ces espaces est la même des espaces L^p, i.e. elle est définie ponctuellement, et la norme de $(x_n)_{n\in\mathbb{N}} \in \ell^p(\mathbb{N}, \mathbb{K})$ est :

$$\boxed{\left\|(x_n)_{n\in\mathbb{N}}\right\|_p = \left(\sum_{n\in\mathbb{N}} |x_n|^p\right)^{1/p}},$$

et pareil si on échange $\mathbb{N}$ avec $\mathbb{Z}$. L'inégalité triangulaire pour cette norme est garantie par **l'inégalité de Minkowski pour les séries** :

$$\boxed{\left(\sum_{n\in\mathbb{N}} |x_n + y_n|^p\right)^{1/p} \leqslant \left(\sum_{n\in\mathbb{N}} |x_n|^p\right)^{1/p} + \left(\sum_{n\in\mathbb{N}} |y_n|^p\right)^{1/p}}, \tag{4.20}$$

Comme pour les espaces L^p, si $p = 2$, on a la possibilité de définir un produit scalaire sur ℓ^2 :

$$\boxed{\langle x_n, y_n\rangle = \sum_{n\in\mathbb{N}} x_n y_n} \quad \text{si } \mathbb{K} = \mathbb{R} \quad \text{et} \quad \boxed{\langle x_n, y_n\rangle = \sum_{n\in\mathbb{N}} x_n \overline{y_n}} \quad \text{si } \mathbb{K} = \mathbb{C},$$

et pareil si on échange $\mathbb{N}$ avec $\mathbb{Z}$, ou n'importe quel autre ensemble dénombrable.

8. Les espaces $\ell^p(\mathbb{N}, \mathbb{K})$ sont des sous-espaces vectoriels de l'espace vectoriel $\mathbb{K}^{\mathbb{N}} := \{(x_n)_{n\in\mathbb{N}},\ x_n \in \mathbb{K}\ \forall n \in \mathbb{N}\}$ des suite à valeurs en $\mathbb{K}$ muni de la structure linéaire définie ponctuellement. Pareil si on échange $\mathbb{N}$ avec $\mathbb{Z}$, dans ce cas on parle de suites bilatérales.

La bonne définition du produit scalaire est garantie par **l'inégalité de Hölder pour les séries** : si $p, q > 0$, $\frac{1}{p} + \frac{1}{q} = 1$, alors ça vaut que

$$\boxed{\sum_{n\in\mathbb{N}} |x_n y_n| \leqslant \left(\sum_{n\in\mathbb{N}} |x_n|^p\right)^{1/p} \left(\sum_{n\in\mathbb{N}} |y_n|^q\right)^{1/q}}.$$

Remarque 4.4.2

- Le produit scalaire de $\ell^2(\mathbb{N}, \mathbb{K})$ est la généralisation en dimension infinie du produit scalaire de $\ell^2(\mathbb{Z}_N)$;
- Le rôle des inégalités de Minkowski et de Hölder dans la définition des espaces L^p et ℓ^p devrait être clair : l'inégalité de Minkowski sert pour garantir l'existence d'une structure linéaire et celle de Hölder pour assurer la bonne définition du produit scalaire dans le cas $p = 2$;
- Les normes $\| \ \|_p$ avec $p \neq 2$ ne sont pas hilbertiennes : il est possible de trouver des exemples d'éléments des espaces L^p, $p \neq 2$, pour lesquels l'identité du parallélogramme n'est pas vérifiée.

On va maintenant démontrer que **les espaces L^p et ℓ^p avec $1 \leqslant p < +\infty$, $p \neq 2$, sont des espaces de Banach, et pour $p = 2$ ils sont des espaces de Hilbert**.

La complétude de l'espaces $L^2([0, 1])$ a été démontré indépendamment par Ernst Sigismund Fisher (1875-1954, mathématicien autrichien) et Frigyes Riesz [9] en 1907. Riesz a démontré en 1910 que tous les espaces $L^p[0, 1]$ sont complets.

Théorème 4.4.1 (Riesz-Fisher) *Pour tout $1 \leqslant p < +\infty$, les espaces $(L^p(X, \mathcal{A}, \mu), \| \ \|_p)$ et $(\ell^p, \| \ \|_p)$ sont complets.*

Preuve. Riesz a utilisé l'artillerie lourde pour prouver ce résultat : le théorème de caractérisation des espaces vectoriels normés complets, le lemme de Fatou, l'inégalité de Minkowski généralisée, le théorème de convergence monotone et celui de convergence dominée sont mélangés pour compléter la démonstration.

Considérons une série quelconque $\sum_{k=0}^{\infty} f_k$ en $L^p(X, \mathcal{A}, \mu)$, $1 \leqslant p \leqslant +\infty$, qui soit absolument convergente, i.e.

$$\sum_{k=0}^{\infty} \|f_k\|_p = M < +\infty,$$

alors, grâce au théorème 4.3.7, pour montrer que $L^p(X, \mathcal{A}, \mu)$ est complet, il suffit de démontrer que $\sum_{k\in\mathbb{N}} f_k$ converge en norme, i.e que $\exists S \in L^p(X, \mathcal{A}, \mu)$ telle que :

$$\left\| \sum_{k=0}^{n} f_k - S \right\|_p \underset{n\to+\infty}{\longrightarrow} 0. \tag{4.21}$$

9. **Frigyes Riesz** (1880-1956), mathématicien hongrois qui a donné des contributions d'importance énorme dans le développement de l'analyse fonctionnelle...et pas que...

Le premier pas pour déterminer la fonction S consiste à définir la suite

$$(g_n)_{n\in\mathbb{N}},\ g_n = \sum_{k=0}^{n} |f_k|, \quad \forall n \in \mathbb{N}.$$

Grâce à l'inégalité de Minkowski généralisée (4.18) on a que :

$$\left(\int_X (g_n)^p d\mu\right)^{1/p} = \|g_n\|_p = \left\|\sum_{k=0}^{n} |f_k|\right\|_p \leqslant \sum_{k=0}^{n} \|f_k\|_p \leqslant \sum_{k=0}^{\infty} \|f_k\|_p \underset{\text{(par hypothèse)}}{=} M < +\infty,$$

donc

$$\int_X (g_n)^p d\mu \leqslant M^p, \qquad \forall n \in \mathbb{N} \tag{4.22}$$

i.e. $(g_n)^p$ est une suite de fonctions monotones croissantes de fonctions intégrables et la suite des intégrales est bornée.

Le théorème de convergence monotone 3.6.1 nous assure que la fonction limite ponctuelle $\lim_{n\to+\infty}(g_n)^p(x)$ est finie p.p. sur X, i.e. $\forall x \in E \subseteq X$ et $\mu(X\backslash E) = 0$. Nécessairement, $\forall x \in E$, aussi la limite ponctuelle

$$g(x) \equiv \lim_{n\to+\infty} g_n(x) \left(= \lim_{n\to+\infty} [(g_n)^p(x)]^{1/p}\right)$$

existe finie.

Vu que, $\forall x \in E$, $\sum_{k=0}^{\infty} f_k(x) \leqslant \sum_{k=0}^{\infty} |f_k(x)| = g(x)$, la série $\sum_{k=0}^{\infty} f_k(x)$ converge p.p. sur X.

Nous allons maintenant construire la fonction $S : X \to \mathbb{K}$ que l'on cherche :

$$S(x) = \begin{cases} \sum_{k=0}^{\infty} f_k(x) & x \in E \\ 0 & x \in X\backslash E. \end{cases}$$

Cette définition assure que S est mesurable, le fait que $S \in L^p(X, \mathcal{A}, \mu)$, i.e. que $\int_X S^p d\mu$ existe fini p.p., est une conséquence du théorème de convergence dominé 3.6.3 et du lemme de Fatou 3.6.2. Pour démontrer ça allons considérer la suites de sommes partielles pour S^p, i.e. $(S_n)^p = \left(\sum_{k=0}^{n} f_k\right)^p \leqslant \left(\sum_{k=0}^{n} |f_k|\right)^p = (g_n)^p$.

$(g_n)^p$ est une suite positive croissante, donc

$$(S_n)^p(x) \leqslant (g_n)^p(x) \leqslant \lim_{n\to+\infty}(g_n)^p(x) = g^p(x), \qquad \forall x \in E. \tag{4.23}$$

Par monotonie $\lim_{n\to+\infty}(g_n)^p(x) = \liminf_{n\to+\infty}(g_n)^p(x)$ et donc, grâce au lemme de Fatou on obtient

$$\int_X g^p d\mu \leqslant \lim_{n\to+\infty} \int_X (g_n)^p d\mu \underset{\text{éq.(4.22)}}{\leqslant} \lim_{n\to+\infty} M^p = M^p < +\infty.$$

La fonction positive et mesurable g^p est donc intégrable p.p. sur X, grâce à cette information et à l'équation (4.23), i.e. $(S_n)^p \leqslant g^p\ \forall n \in \mathbb{N}$, on peut invoquer le théorème de convergence dominée pour garantir que S^p, la limite p.p. de $(S_n)^p$, converge sur X, i.e. $S \in L^p(X, \mathcal{A}, \mu)$.

Pour terminer la preuve du théorème il faut démontrer que la fonction S est celle qui vérifie l'équation (4.21), i.e.

$$\left\| \sum_{k=0}^{n} f_k - S \right\|_p \underset{n\to+\infty}{\longrightarrow} 0 \iff \lim_{n\to+\infty} \int_E \left| \sum_{k=0}^{n} f_k - \sum_{k=0}^{\infty} f_k \right|^p d\mu = 0,$$

où on a évité d'écrire l'intégration sur $X\backslash E$ car $\mu(X\backslash E) = 0$. Avec les notations qu'on a utilisé, on peut réécrire d'une manière plus simple la condition de convergence en norme $\| \ \|_p$ de la série $\sum_{k=0}^{\infty} f_k$ à S comme ça :

$$\lim_{n\to+\infty} \|S_n - S\|_p = 0 \iff \lim_{n\to+\infty} \int_E |S_n - S|^p \, d\mu = 0,$$

il est évident que, si on démontre qu'on a le droit d'échanger l'intégrale avec la limite, alors le résultat est vrai, vu que, dans ce cas :

$$\lim_{n\to+\infty} \int_E |S_n - S|^p \, d\mu = \int_E \lim_{n\to+\infty} |S_n - S|^p \, d\mu \underset{(S \text{ ne dépend pas de } n)}{=} \int_E |\lim_{n\to+\infty} S_n - S|^p \, d\mu = 0.$$

Pour obtenir la possibilité d'effectuer l'échange, on observe qu'on peut écrire la majoration suivante :

$$|S_n(x) - S(x)|^p \leqslant (|S_n(x)| + |S(x)|)^p \leqslant (g(x) + g(x))^p = (2g(x))^p = 2^p (g(x))^p \quad \forall x \in E,$$

comme $\int_X g^p d\mu \leqslant M^p < +\infty$, cette majoration assure que la suite $(|S_n(x) - S(x)|^p)_{n\in\mathbb{N}}$ vérifie les conditions du théorème de convergence dominée, qui garantit la possibilité d'échanger limite et intégrale.

Ceci, comme dit ci-dessus, assure que la série $\sum_{k=0}^{\infty} f_k$ en $L^p(X, \mathcal{A}, \mu)$, supposée être absolument convergente, est aussi simplement convergente. Donc, tous les espaces $L^p(X, \mathcal{A}, \mu)$ avec $1 \leqslant p < \infty$ sont complets.

Comme les espaces ℓ^p sont des cas particuliers des espaces L^p, le résultat vaut aussi pour ces espaces $\forall 1 \leqslant p < \infty$. □

Exercice 4.2

Soit $a = (a_n)_{n\in\mathbb{N}}$ une suite de nombres réels strictement positifs, on note $\ell^2_a(\mathbb{N}, \mathbb{C})$ l'espace vectoriel formé des suites de nombres complexes $(u_n)_{n\in\mathbb{N}}$ vérifiant $\sum_{n\in\mathbb{N}} a_n |u_n|^2 < +\infty$. Montrer que l'application définie par

$$\langle u, v \rangle_{\ell^2_a} = \sum_{n\in\mathbb{N}} a_n u_n \overline{v_n}$$

est bien définie sur $\ell^2_a(\mathbb{N}, \mathbb{C}) \times \ell^2_a(\mathbb{N}, \mathbb{C})$ (c'est-à-dire que $\langle u, v \rangle$ existe pour tout $u, v \in \ell^2_a(\mathbb{N}, \mathbb{C})$), en déduire qu'il s'agit d'un produit scalaire.

Solution de l'exercice 4.2

Vu que $u, v \in \ell^2_a(\mathbb{N}, \mathbb{C})$, $\sqrt{a}u$ et $\sqrt{a}\overline{v}$ appartiennent à $\ell^2(\mathbb{N}, \mathbb{C})$, alors :

$$\langle u, v\rangle_{\ell^2_a} = \sum_{n\in\mathbb{N}} \sqrt{a_n}u_n\overline{\sqrt{a_n}v_n} = \langle\sqrt{a_n}u_n, \sqrt{a_n}v_n\rangle_{\ell^2} < +\infty.$$

La sesquilinéarité et la symétrie conjuguée de $\langle u, v\rangle_{\ell^2_a}$ descendent immédiatement des propriétés analogues du produit scalaire de $\ell^2(\mathbb{N}, \mathbb{C})$. Allons vérifier explicitement seulement la définie positivité. Si $u \in \ell^2_a(\mathbb{N}, \mathbb{C})$, alors $\langle u, u\rangle_{\ell^2_a} = \sum_{n\in\mathbb{N}} a_n|u_n|^2 \geqslant 0$ car somme de termes positifs. Cette formule montre aussi que $\langle u, u\rangle_{\ell^2_a} = 0 \iff a_n|u_n|^2 = 0$ pour tout $n \in \mathbb{N}$, mais $a_n > 0$ pour tout $n \in \mathbb{N}$ par hypothèse, donc $|u_n|^2 = 0 \iff u_n = 0\ \forall n \in \mathbb{N}$, i.e. $u = 0_{\ell^2_a}$. □

Exercice 4.3

Soit $s \in \mathbb{R}$, $s > 0$ et

$$H^s = \left\{ u = (u_n)_{n\in\mathbb{N}} \subset \mathbb{C}\ \forall n \in \mathbb{N}\ :\ \sum_{n\in\mathbb{N}} (1 + n^2)^s|u_n|^2 < +\infty \right\},$$

H^s est un espace de Hilbert qui apparait souvent dans la résolution des équations différentielles à l'aide de la transformée de Fourier.

1. Montrer que H^s est un sous-espace vectoriel de $\ell^2(\mathbb{N}, \mathbb{C})$.
2. On note $\phi : H^s \times H^s \to \mathbb{C}$ l'application définie par

$$\phi(u, v) := \sum_{n\in\mathbb{N}} (1 + n^2)^s u_n\overline{v_n} \qquad \forall u, v \in H^s,$$

sans se préoccuper pour le moment de sa bonne définition, i.e. de la convergence de la série. Pour toute suite $w = (w_n)_{n\in\mathbb{N}} \in H^s$, on définit la suite $\tilde{w}$ par :

$$\tilde{w}_n = (1 + n^2)^{s/2} w_n \qquad \forall n \in \mathbb{N}.$$

(a) Montrer que $\tilde{w} \in \ell^2(\mathbb{N}, \mathbb{C})$ et que ça vaut :

$$\phi(u, v) = \langle\tilde{u}, \tilde{v}\rangle_{\ell^2} \qquad \forall u, v \in H^s,$$

où on a noté $\langle\ ,\ \rangle_{\ell^2}$ le produit scalaire usuel de $\ell^2(\mathbb{N}, \mathbb{C})$.

(b) En déduire que ϕ est bien définie sur $H^s \times H^s$, puis qu'il s'agit d'un produit scalaire que l'on notera avec $\phi = \langle\ ,\ \rangle_{H^s}$.

3. On veux montrer que $(H^s, \langle\ ,\ \rangle_{H^s})$ est un espace de Hilbert. Pour cela, fixons une suite de Cauchy arbitraire $(u_m)_{m\in\mathbb{N}}$ dans H^s.

(a) Montrer que $(\tilde{u}_m)_{m\in\mathbb{N}}$ est de Cauchy dans $\ell^2(\mathbb{N}, \mathbb{C})$.

(b) En déduire que $(\tilde{u}_m)_{m\in\mathbb{N}}$ converge dans $\ell^2(\mathbb{N}, \mathbb{C})$ vers une limite qu'on notera $\tilde{l}$.

(c) On définit la suite $l = (l_n)_{n\in\mathbb{N}}$ par :

$$l_n = \frac{1}{(1 + n^2)^{s/2}}\tilde{l}_n \qquad \forall n \in \mathbb{N}.$$

Montrer que l appartient à H^s, que $(u_m)_{m\in\mathbb{N}}$ converge vers l dans H^s et conclure.

Solution de l'exercice 4.3

1. Pour montrer que $H^s \subset \ell^2(\mathbb{N},\mathbb{C})$ nous allons démontrer, dans l'ordre qui suit, que $u \in H^s \implies u \in \ell^2(\mathbb{N},\mathbb{C})$, que $H^s \neq \varnothing$ et que H^s est stable par rapport aux combinaisons linéaires de ses éléments.

   Pour toute suite $u = (u_n)_{n\in\mathbb{N}} \subset \mathbb{C}$ ça vaut $0 \leqslant |u_n|^2 \leqslant (1+n^2)|u_n|^2$ pour tout $n \in \mathbb{N}$, donc

   $$\sum_{n\in\mathbb{N}} |u_n|^2 \leqslant \sum_{n\in\mathbb{N}} (1+n^2)^s |u_n|^2 \underset{\text{déf. de } H^s}{<} +\infty \implies u \in \ell^2(N,\mathbb{C}).$$

   Évidemment, $0_{\ell^2} \in H^s$, et alors $H^s \neq \varnothing$. Pour terminer, soit $\lambda \in \mathbb{C}$ et $u, v \in H^s$, alors :

   $$0 \leqslant |u_n + \lambda v_n|^2 \leqslant (|u_n| + |\lambda||v_n|)^2 \leqslant 2(|u_n|^2 + |\lambda|^2|v_n|^2),$$

   où, dans la dernière inégalité, on a utilisé le fait que les modules sont des nombres réels et que, pour tout $a, b \in \mathbb{R}$, $0 \leqslant (a-b)^2 = a^2 + b^2 - 2ab = 2a^2 - a^2 + 2b^2 - b^2 - 2ab$, i.e. $a^2 + b^2 + 2ab \leqslant 2a^2 + 2b^2$, i.e. $(a+b)^2 \leqslant 2(a^2+b^2)$, si on écrit $a = |u_n|$ et $b = |\lambda||v_n|$ on obtient la dernière inégalité écrite dans la formule ci-dessus. En passant aux séries nous pouvons écrire :

   $$\sum_{n\in\mathbb{N}} (1+n^2)^s |u_n + \lambda v_n|^2 \leqslant 2\left(\sum_{n\in\mathbb{N}} (1+n^2)^s |u_n|^2 + |\lambda|^2 \sum_{n\in\mathbb{N}} (1+n^2)^s |v_n|^2\right) < +\infty$$

   car $u, v \in H^s$, donc $u + \lambda v \in H^s$ et alors H^s est bien un sous-espace vectoriel de $\ell^2(\mathbb{N},\mathbb{C})$.

2. (a) $\tilde{w} \in \ell^2(\mathbb{N},\mathbb{C})$ si $\sum\limits_{n\in\mathbb{N}} |\tilde{w}_n|^2 < +\infty$, mais

   $$\sum_{n\in\mathbb{N}} |\tilde{w}_n|^2 = \sum_{n\in\mathbb{N}} (1+n^2)^s |w_n|^2 < +\infty$$

   car $w \in H^s$, donc $\tilde{w} \in \ell^2(\mathbb{N},\mathbb{C})$. Soient maintenant $u, v \in H^s$, alors :

   $$\phi(u,v) = \sum_{n\in\mathbb{N}} (1+n^2)^s u_n \overline{v_n} = \sum_{n\in\mathbb{N}} (1+n^2)^{s/2} u_n \overline{(1+n^2)^{s/2} v_n} = \sum_{n\in\mathbb{N}} \tilde{u}_n \overline{\tilde{v}_n} = \langle \tilde{u}, \tilde{v} \rangle_{\ell^2}.$$

   (b) Nous avons :

   $$\begin{aligned}\phi(u,v) = \sum_{n\in\mathbb{N}} (1+n^2)^s u_n v_n &\leqslant \sum_{n\in\mathbb{N}} |(1+n^2)^s u_n \overline{v}_n| = \sum_{n\in\mathbb{N}} |(1+n^2)^{s/2} u_n (1+n^2)^{s/2} \overline{v}_n| \\ &= \sum_{n\in\mathbb{N}} |\tilde{u}_n \overline{\tilde{v}_n}| = \langle \tilde{u}, \tilde{v} \rangle_{\ell^2} \underset{\text{Cauchy-Schwarz}}{\leqslant} \|\tilde{u}\|_{\ell^2} \|\tilde{v}\|_{\ell^2} < +\infty,\end{aligned}$$

   et alors $\phi(u,v)$ est bien défini pour tous $u, v \in H^2$. Grâce au fait que $\phi(u,v) = \langle \tilde{u}, \tilde{v} \rangle_{\ell^2}$ nous avons que ϕ est un produit scalaire : elle est hermitienne et sesquilinéaire vu que $\langle\ ,\ \rangle_{\ell^2}$ a ces propriétés et, pour la définie positivité, il suffit d'observer que, pour tout $u \in H^s$, $\phi(u) = 0$ implique $\sum\limits_{n\in\mathbb{N}} (1+n^2)^{s/2} u_n (1+n^2)^{s/2} \overline{u_n} = \langle \tilde{u}, \tilde{u} \rangle_{\ell^2} = \|\tilde{u}\|_{\ell^2} = 0$, i.e. $\tilde{u} = 0$, c'est-à-dire $(1+n^2)^{s/2} u_n = 0 \iff u_n = 0\ \forall n \in \mathbb{N}$. Donc ϕ est un produit scalaire complexe sur H^2 que l'on note avec $\phi(u,v) = \langle u, v \rangle_{H^s}$.

3. (a) Pour montrer que si $u = (u_m)_{m\in\mathbb{N}}$ est une suite de Cauchy arbitraire en H^s, alors $(\tilde{u}_m)_{m\in\mathbb{N}}$ est une suite de Cauchy en $\ell^2(\mathbb{N},\mathbb{C})$ nous allons écrire la condition de Cauchy au carré pour u :

$$\forall\varepsilon>0\ \exists N_\varepsilon\in\mathbb{N}\ :\ m,k\leqslant N_\varepsilon\implies \|u_m-u_k\|^2_{H^s}<\varepsilon^2,$$

mais $\|u_m-u_k\|^2_{H^s}=\langle u_m-u_k,u_m-u_k\rangle_{H^s}\underset{(2.(a))}{=}\langle\widetilde{u_m-u_k},\widetilde{u_m-u_k}\rangle_{\ell^2}$, et

$$\widetilde{u_m-u_k}=(1+n^2)^{s/2}(u_m-u_k)=(1+n^2)^{s/2}u_m-(1+n^2)^{s/2}u_k=\tilde{u}_m-\tilde{u}_k,$$

alors $\|u_m-u_k\|^2_{H^s}=\langle\widetilde{u_m-u_k},\widetilde{u_m-u_k}\rangle_{\ell^2}=\langle\tilde{u}_m-\tilde{u}_k,\tilde{u}_m-\tilde{u}_k\rangle_{\ell^2}=\|\tilde{u}_m-\tilde{u}_k\|^2_{\ell^2}$, ce qui implique que $(\tilde{u}_m)_{m\in\mathbb{N}}$ est une suite de Cauchy en $\ell^2(\mathbb{N},\mathbb{C})$.

(b) $\ell^2(\mathbb{N},\mathbb{C})$ étant complet, la suite de Cauchy $(\tilde{u}_m)_{m\in\mathbb{N}}$ converge vers un élément de $\ell^2(\mathbb{N},\mathbb{C})$ qu'on écrit $\tilde{l}$.

(c) Considérons la suite $l=\tilde{l}/(1+n^2)^{s/2}$ et démontrons qu'elle appartient à H^s en calculant sa norme au carré en H^s :

$$\|l\|^2_{H^s}=\sum_{n\in\mathbb{N}}(1+n^2)^s|l_n|^2=\sum_{n\in\mathbb{N}}\cancel{(1+n^2)^s}\frac{|\tilde{l}_n|^2}{\cancel{((1+n^2)^s}}=\sum_{n\in\mathbb{N}}|\tilde{l}_n|^2<+\infty,$$

donc on a bien $l\in H^2$.

Montrons maintenant que $(u_m)_{m\in\mathbb{N}}$ converge vers l : grâce encore au résultat du point 2.(a) nous avons $\langle u_m-l,u_m-l\rangle_{H^s}=\langle\widetilde{u_m-l},\widetilde{u_m-l}\rangle_{\ell^2}$, mais on a vu aussi que $\widetilde{u_m-l}=\tilde{u}_m-\tilde{l}$ donc ça vaut $\|u_m-l\|^2_{H^s}=\|\tilde{u}_m-\tilde{l}\|^2_{\ell^2}\underset{m\to+\infty}{\to}0$, grâce à 3(b), i.e. $(u_m)_{m\in\mathbb{N}}$ converge vers l in H^s. Étant $(u_m)_{m\in\mathbb{N}}$ une suite de Cauchy arbitraire en H^s, nous avons démontré que cet espace est de Hilbert. □

4.4.2 Les espaces L^∞ et ℓ^∞ et leur complétude

On a volontairement laissé de côté le cas $p=\infty$, qu'on va examiner séparément. Comme d'habitude, soit $(X,\mathcal{A},\mu)$ un espace mesuré et $\mathbb{K}=\mathbb{R}$ ou $\mathbb{C}$. On commence par définir l'espace :

$$\boxed{\mathcal{L}^\infty(X,\mathcal{A},\mu)=\{f:X\to\mathbb{K}\ :\ \exists M\in\mathbb{R},M\geqslant 0,\ \text{tel que}\ |f(x)|\leqslant M\ \text{p.p}\}}.$$

On appelle les fonctions de $\mathcal{L}^\infty(X,\mathcal{A},\mu)$ des **fonctions essentiellement bornées**, i.e. bornées sur le complémentaire d'un ensemble de mesure μ nulle.

Comme pour les espaces L^p, on introduit la relation d'équivalence :

$$f,g\in\mathcal{L}^\infty(X,\mathcal{A},\mu),\ f\sim g\ \text{si}\ f=g\ \text{p.p},$$

alors, l'espace quotient

$$L^\infty(X,\mathcal{A},\mu)=\mathcal{L}^\infty(X,\mathcal{A},\mu)/_\sim$$

est un espace normé ayant pour norme :

$$\boxed{\|f\|_\infty=\inf\{M\geqslant 0\ :\ |f(x)|\leqslant M\ \text{p.p}\}}$$

qu'on appelle supess(f) : **borne supérieure essentielle** de f, qui satisfait par définition :

$$|f(x)| \leqslant ||f||_\infty \text{ p.p}$$

pour toute $f \in L^\infty(X, \mathcal{A}, \mu)$.

Le symbole $\| \|_\infty$ dérive du fait que si $1 \leqslant p < +\infty$ et $f \in L^p \cap L^\infty$ alors :

$$\|f\|_\infty = \lim_{p \to +\infty} \|f\|_p.$$

Comme pour les espaces L^p, on va faire un commentaire à part pour les fonctions continues : si une fonction continue est telle que $|f(x)| > M$ alors, par continuité, il existe un voisinage de taille positif dans lequel f n'est pas bornée par M, donc **une fonction continue et essentiellement bornée est bornée tout court**.

On définit aussi :

$$\boxed{\ell^\infty(\mathbb{N}, \mathbb{K}) = L^\infty(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu_{\text{comptage}}) = \{(x_n)_{n\in\mathbb{N}} \ : \ x_n \in \mathbb{K} \ \ \forall n \in \mathbb{N}, \ \exists M \geqslant 0 \ : \ |x_n| \leqslant M\}},$$

i.e. ℓ^∞ est **l'espace des suites bornées** (une définition similaire est donnée si on échange $\mathbb{N}$ avec $\mathbb{Z}$). $\ell^\infty(\mathbb{N}, \mathbb{K})$ est un espace normé avec :

$$\boxed{\|(x_n)_{n\in\mathbb{N}}\|_\infty = \sup_{n\in\mathbb{N}} |x_n|}.$$

Théorème 4.4.2 *$(L^\infty(X, \mathcal{A}, \mu), \| \|_\infty)$ et $(\ell^\infty(\mathbb{N}, \mathbb{K}), \| \|_\infty)$ sont des espaces de Banach.*

Preuve. On va développer la preuve pour $L^\infty(X, \mathcal{A}, \mu)$, le fait que $\ell^\infty(\mathbb{N}, \mathbb{K})$ soit de Banach sera une implication automatique, en tant que cas particulier de $L^\infty(X, \mathcal{A}, \mu)$.

Il faut montrer que si $(f_n)_{n\in\mathbb{N}}$ une suite de Cauchy d'éléments de $L^\infty(X, \mathcal{A}, \mu)$, alors elle converge vers un élément de $L^\infty(X, \mathcal{A}, \mu)$.

Par définition d'une suite de Cauchy :

$$\forall \varepsilon > 0 \ \exists N_\varepsilon > 0 \ : \ n, m \geqslant N_\varepsilon \implies \|f_n - f_m\|_\infty < \varepsilon, \tag{4.24}$$

on va l'utiliser plus tard.

Considérons les ensembles des points où les fonctions de la suite « se comportent mal » :

$$A_k = \{x \in X : |f_k(x)| > ||f_k||_\infty\}, \ B_{n,m} = \{x \in X : |f_n(x) - f_m(x)| > ||f_n - f_m||_\infty\}$$

par définition de $L^\infty(X, \mathcal{A}, \mu)$, $\mu(A_k) = \mu(B_{n,m}) = 0$ et

$$\forall x \in A_k^c : |f_k(x)| \leqslant ||f_k||_\infty, \ \forall x \in B_{n,m}^c : |f_n(x) - f_m(x)| \leqslant ||f_n - f_m||_\infty.$$

Pour éliminer la dépendance des indices k, n, m, on va construire l'ensemble :

$$E = \bigcup_{k\in\mathbb{N}} A_k \cup \bigcup_{n,m\in\mathbb{N}} B_{n,m},$$

qui a mesure nulle, $\mu(E) = 0$, en tant qu'union dénombrable d'ensembles de mesure nulle.

On observe maintenant que

$$\forall x \in E^c,\ \forall n, m \geqslant N_\varepsilon\ : |\ f_n(x) - f_m(x)\ | \leqslant \|f_n - f_m\|_\infty \underset{\text{(éq. (4.24)}}{<} \varepsilon, \tag{4.25}$$

donc $(f_n(x))_{n\in\mathbb{N}}$ est une suite de Cauchy d'éléments de $\mathbb{K}$, qui est complet, par conséquent il existe la limite ponctuelle $f(x) = \lim_{n\to+\infty} f_n(x)$.

L'éq. (4.25) vaut, en particuliers, si on fait tendre $n \to +\infty$, donc $\forall \varepsilon > 0$ ça vaut que :

$$\forall x \in E^c,\ \forall m \geqslant N_\varepsilon\ : |\ \lim_{n\to+\infty}\ f_n(x) - f_m(x)\ | = |\ f(x) - f_m(x)\ | < \varepsilon,$$

qui est la définition de convergence uniforme de la suite $(f_n)_{n\in\mathbb{N}} \subset L^\infty$ vers f sur E^c. Un résultat standard de l'analyse garantie que si une suite de fonctions bornées converge uniformément à une fonction, alors même la fonction limite est bornée, ceci implique, dans notre cas, que f est essentiellement bornée sur E^c.

Il reste tout simplement d'étendre la définition de f à une fonction $\tilde{f}$ définie sur tout X (car les éléments de $L^\infty(X, \mathcal{A}, \mu)$ sont définis sur tout X) en préservant la propriété d'être essentiellement bornée, mais cela est très simple car ça suffit de poser :

$$\tilde{f}(x) = \begin{cases} f(x) & \text{si } x \in E^c \\ 0 & \text{si } x \in E, \end{cases}$$

comme $\mu(E) = 0$, $\tilde{f} : X \to \mathbb{K}$ est le représentant d'une classe d'équivalence de $L^\infty(X, \mathcal{A}, \mu)$ à laquelle converge la suite de Cauchy $(f_n)_{n\in\mathbb{N}}$ et ainsi le théorème est prouvé. □

Exercice 4.4

On considère une suite $a = (a_k)_{k\in\mathbb{Z}}$ et, pour tout $u \in \ell^\infty(\mathbb{Z}, \mathbb{C})$, on appelle $a * u$ la suite bilatérale définie pour $k \in \mathbb{Z}$ par :

$$(a * u)_k = \sum_{m\in\mathbb{Z}} a_m u_{k-m}.$$

On pose, pour tout $f \in \ell^\infty(\mathbb{Z}, \mathbb{C})$, $T(u) := a * u + f$.

1. On suppose dans cette question que $a = \delta_1$, la suite définie par $a_1 = 1$ et $a_j = 0$ si $j \neq 1$. Calculer $a * u$ en fonction de u.
2. On suppose d'abord que $a = (a_k)_{k\in\mathbb{Z}} \in \ell^1$ vérifiant $\|a\|_1 = \sum_{k\in\mathbb{Z}} |a_k| < 1$.
   (a) Montrer que $(a * u)_k$ est bien définie pour tout $k \in \mathbb{Z}$ et que $a * u \in \ell^\infty$.
   (b) Montrer que $T : \ell^\infty(\mathbb{Z}, \mathbb{C}) \to \ell^\infty(\mathbb{Z}, \mathbb{C})$ est une contraction.
   (c) En déduire qu'il existe une unique solution $u \in \ell^\infty(\mathbb{Z}, \mathbb{C})$ de l'équation $T(u) = u$.
3. On suppose maintenant que $a = (a_k)_{k\in\mathbb{Z}} \in \ell^2$ vérifiant $\|a\|_2 := \sum_{k\in\mathbb{Z}} |a_k|^2 < 1$.
   (a) Montrez à l'aide d'un exemple qu'on peut avoir $a \notin \ell^1(\mathbb{Z}, \mathbb{C})$.
   (b) Montrer que $(a * u)_k$ est bien définie pour tout $k \in \mathbb{Z}$ et que $a * u \in \ell^\infty(\mathbb{Z}, \mathbb{C})$.

(c) En déduire que pour tout $u \in \ell^2(\mathbb{Z},\mathbb{C})$, $T(u) \in \ell^\infty(\mathbb{Z},\mathbb{C})$ et que si $u, v \in \ell^2(\mathbb{Z},\mathbb{C})$, alors $\|T(u) - T(v)\|_\infty < \|u - v\|_2$.

(d) On prend maintenant $a = \frac{1}{2}\delta_1$ et $f = \mathbf{1}$ la suite constante $f_j = 1$ pour tout $j \in \mathbb{Z}$. Calculer $T(u)$ en fonction de u et déterminer $\lim_{k\to+\infty}(T(u))_k$.

(e) En déduire qu'il n'existe pas $u \in \ell^2(\mathbb{Z},\mathbb{C})$ tel que $T(u) = u$. Est-ce que ça est en contradiction avec le théorème de point fixe ?
Indication : il n'est pas nécessaire de déterminer u pour répondre à cette question.

(f) Déterminer $u \in \ell^\infty(\mathbb{Z},\mathbb{C})$ tel que $T(u) = u$.

Solution de l'exercice 4.4

1. Par définition,
$$(\delta_1 * u)_k = \sum_{m\in\mathbb{Z}} \delta_{m,1} u_{k-m} = u_{k-1}.$$

2. (a) Par calcul direct :
$$(a * u)_k = \sum_{m\in\mathbb{Z}} a_m u_{k-m} \leqslant \sum_{m\in\mathbb{Z}} |a_m u_{k-m}| \leqslant \|u\|_\infty \sum_{m\in\mathbb{Z}} |a_m| = \|u\|_\infty \|a\|_1 < +\infty,$$
car $a \in \ell^1(\mathbb{Z},\mathbb{C})$ et $u \in \ell^\infty(\mathbb{Z},\mathbb{C})$. De plus, comme la majoration ne dépend pas de k, $\|a * u\|_\infty = \sup_{k\in\mathbb{Z}}\{(a * u)_k\} \leqslant \|u\|_\infty \|a\|_1 < +\infty$.

(b) Encore une fois, par calcul direct, on a $\|T(u) - T(v)\|_\infty = \|a * u + f - a * v - f\|_\infty = \|a * u - a * v\|_\infty$, mais $u \mapsto a * u$ est linéaire, donc, grâce à ce que l'on a vu dans la question précédente : $\|T(u) - T(v)\|_\infty = \|a * (u - v)\|_\infty \leqslant \|a\|_1 \|u - v\|_\infty$. Comme $\|a\|_1 < 1$ par hypothèse, T est bien une contraction.

(c) Comme $(\ell^\infty(\mathbb{Z},\mathbb{C}), \|\ \|_\infty)$ est un espace normée (et donc métrique) complet, le théorème de point fixe nous assure l'existence d'un seul élément $\bar{u} \in \ell^\infty(\mathbb{Z},\mathbb{C})$ tel que $T(\bar{u}) = \bar{u}$, i.e. $\bar{u} = a * \bar{u} + f$.

3. (a) Probablement, l'exemple le plus simple de suite $a \in \ell^2(\mathbb{Z},\mathbb{C})$ telle que $a \notin \ell^1(\mathbb{Z},\mathbb{C})$ est donné par la suite $a_k = \begin{cases} 0 & k \leqslant 0 \\ \frac{1}{k} & \text{sinon.} \end{cases}$ En fait, dans ce cas, $\sum_{k\in\mathbb{Z}} |a_k| = \sum_{k=1}^{\infty} \frac{1}{k}$ la série harmonique, que l'on sait être divergente, donc $a \notin \ell^1(\mathbb{Z},\mathbb{C})$. Or, $\sum_{k\in\mathbb{Z}} |a_k|^2 = \sum_{k=1}^{\infty} \frac{1}{k^2}$ qui est convergente, i.e. $a \in \ell^2(\mathbb{Z},\mathbb{C})$.

(b) Grâce aux hypothèses que nous avons, pour tout $k \in \mathbb{Z}$ fixé, il est possible d'appliquer l'inégalité de Cauchy-Schwarz comme ça :
$$\sum_{m\in\mathbb{Z}} |a_m u_{k-m}| \leqslant \left(\sum_{m\in\mathbb{Z}} |a_m|\right)^{1/2} \left(\sum_{m\in\mathbb{Z}} |u_{k-m}|\right)^{1/2} \leqslant \|a\|_2^2 \left(\sum_{n\in\mathbb{Z}} |u_n|\right)^{1/2} = \|a\|_2 \|u\|_2,$$
où on a utilisé un changement de variable $n = k - m$, avec $k \in \mathbb{Z}$ fixé et $m \in \mathbb{Z}$, donc $n \in \mathbb{Z}$. Comme $(a * u)_k = \sum_{m\in\mathbb{Z}} a_m u_{k-m} \leqslant \sum_{m\in\mathbb{Z}} |a_m u_{k-m}| < +\infty$, pour tout

$k \in \mathbb{Z}$ fixé, la suite $a * u$ est bien définie. De plus, comme dans la question 2 (a), comme la majoration ne dépend pas de k, $\|a*u\|_\infty = \sup_{k\in\mathbb{Z}}\{(a*u)_k\} \leqslant \|a\|_2\|u\|_2 < +\infty$.

(c) $T(u) = a * u + f$ est somme de deux éléments de $\ell^\infty(\mathbb{Z},\mathbb{C})$ (f par hypothèse et $a * u$ on vient de le démontrer), donc $T(u) \in \ell^\infty(\mathbb{Z},\mathbb{C})$. À nouveau, $u \mapsto a * u$ est clairement linéaire, donc, grâce au résultat de la question précédente :

$$\|T(u) - T(v)\|_\infty = \|a * u - a * v\|_\infty = \|a * (u - v)\|_\infty \leqslant \|a\|_2\|u - v\|_2 < \|u - v\|_2,$$

car, par hypothèse, $\|a\|_2 < 1$.

(d) En utilisant le résultat de la question 1, on a $(T(u))_k = \frac{u_{k-1}}{2} + \mathbf{1}$. De plus, comme $u \in \ell^2(\mathbb{Z},\mathbb{C})$, $\sum_{k\in\mathbb{Z}} |u_k|^2$ converge, et alors nécessairement $u_k \underset{k\to+\infty}{\longrightarrow} 0$, ce qui implique $(T(u))_k \underset{k\to+\infty}{\longrightarrow} \mathbf{1}$.

(e) Si on avait $T(u) = u$, alors on aurait $u_k = \frac{u_{k-1}}{2} + 1$ et, en prenant la limite pour k qui tend vers l'infini aux deux côtés, on obtiendrait l'absurde d'avoir $0 = 1$. Il n'y aucune contradiction avec le théorème de point fixe, car dans l'inégalité $\|T(u) - T(v)\|_\infty < \|u - v\|_2$ on n'a pas $\|T(u) - T(v)\|_2 \ldots$ Évidemment, comme il n'y a pas de point fixe, T ne peut pas être une contraction sur $\ell^2(\mathbb{Z},\mathbb{C})$.

(f) Une suite $u \in \ell^\infty(\mathbb{Z},\mathbb{C})$ telle que $T(u) = u$ est une suite bornée telle que $u_k = \frac{u_{k-1}}{2} + 1$ (il s'agit d'une suite « arithmético-géométrique »). On pose $u_k = v_k + \alpha$, avec v_k et α inconnue pour le moment, alors $v_k + \alpha = \frac{u_{k-1}}{2} + 1 = \frac{v_{k-1}+\alpha}{2} + 1$, i.e. $v_k = \frac{v_{k-1}}{2} + 1 - \frac{\alpha}{2}$ donc, en choisissant $\alpha = 2$, on obtient une suite géométrique $v_k = \frac{v_{k-1}}{2}$, soit, grâce à un résultat standard sur les suites géométriques, $v_k = 2^{-k}v_0$, mais $v_0 = u_0 - \alpha$ et $\alpha = 2$, donc $v_0 = u_0 - 2$, ce qui implique $u_k = 2^{-k}(u_0 - 2) + 2$. Pour tout $k \geqslant 0$, $2^{-k} < 1$, mais, pour $k < 0$, 2^{-k} n'est pas borné, donc, pour avoir u_k bornée, il faut annuler son facteur, i.e. imposer $u_0 - 2 = 0$. Finalement, et on trouve que la seule suite $u \in \ell^\infty(\mathbb{Z},\mathbb{C})$ telle que $T(u) = u$, i.e. le seul point fixe pour la contraction $T : \ell^\infty(\mathbb{Z},\mathbb{C}) \to \ell^\infty(\mathbb{Z},\mathbb{C})$, est la suite constante à 2, $u_k = 2$ pour tout $k \in \mathbb{Z}$. □

4.4.3 Relations d'inclusion parmi les espaces ℓ^p

Allons introduire l'espace fonctionnel suivant :

$$\boxed{\ell^0(\mathbb{N},\mathbb{K}) \equiv \ell_{\text{fin}}(\mathbb{N},\mathbb{K}) = \{(x_n)_{n\in\mathbb{N}} \subset \mathbb{K},\ \exists N \in \mathbb{N}\ :\ x_n = 0\ \forall n \geqslant N\}}, \tag{4.26}$$

i.e. suites avec un nombre fini d'éléments $\neq 0$. Évidemment, $\ell^0(\mathbb{N},\mathbb{K}) \subset \ell^p(\mathbb{N},\mathbb{K})\ \forall p \geqslant 1$.

Théorème 4.4.3 *Soient $p, q \in \mathbb{R}$, $1 \leqslant p \leqslant q < \infty$, alors :*

$$\boxed{\ell^0(\mathbb{N},\mathbb{K}) \subset \ell^1(\mathbb{N},\mathbb{K}) \subset \ldots \subset \ell^p(\mathbb{N},\mathbb{K}) \subset \ldots \subset \ell^q(\mathbb{N},\mathbb{K}) \subset \ldots \subset \ell^\infty(\mathbb{N},\mathbb{K})}.$$

Preuve. Vu que $\ell^0(\mathbb{N},\mathbb{K}) \subset \ell^1(\mathbb{N},\mathbb{K})$, passons tout de suite à démontrer que $\ell^p(\mathbb{N},\mathbb{K}) \subset \ell^\infty(\mathbb{N},\mathbb{K})$ $\forall 1 \leqslant p < \infty$: presque évident car

$$(x_n)_{n\in\mathbb{N}} \in \ell^p(\mathbb{N},\mathbb{K}) \iff \sum_{n\in\mathbb{N}} |x_n|^p < +\infty$$

donc, nécessairement, $|x_n| \underset{n\to+\infty}{\longrightarrow} 0$, i.e. $|x_n|$ est bornée et alors $(x_n)_{n\in\mathbb{N}} \in \ell^\infty(\mathbb{N},\mathbb{K})$.

Il reste seulement à démontrer que : $\ell^p(\mathbb{N},\mathbb{K}) \subset \ell^q(\mathbb{N},\mathbb{K})$ si $1 \leqslant p \leqslant q$: comme $|x_n| \underset{n\to+\infty}{\longrightarrow} 0$, alors, en particulier, $\exists N \in \mathbb{N}$ tel que $|x_n| \leqslant 1$, $\forall\, n \geqslant N$ donc $|x_n|^q \leqslant |x_n|^p$ $\forall\, n \geqslant N$ et ceci implique que :

$$\sum_{n\in\mathbb{N}} |x_n|^q \leqslant \sum_{n\in\mathbb{N}} |x_n|^p$$

donc la convergence de $\sum\limits_{n\in\mathbb{N}} |x_n|^p$ implique celle de $\sum\limits_{n\in\mathbb{N}} |x_n|^q$, i.e. $\ell^p \subset \ell^q$. □

Remarque 4.4.3 La complétude d'un espace métrique de dimension infinie dépend de la métrique que nous choisissons pour l'espace. Pour vérifier cette affirmation, nous allons examiner la complétude de $(\ell^1, \|\ \|_\infty)$, i.e. de ℓ^1 interprété comme sous-espace de ℓ^∞ et muni de la norme de ce dernier espace.

Exercice 4.5

Démontrer que $(\ell^1, \|\ \|_\infty)$ n'est pas complet.

Solution de l'exercice 4.5

Vu que $\ell^1 \subset \ell^\infty$, pour résoudre l'exercice, ça suffit de démontrer que ℓ^1 n'est pas un sous-ensemble fermé de ℓ^∞ par rapport à la norme $\|\ \|_\infty$, i.e. qu'il existe au moins une suite convergente (et donc de Cauchy) hors de $(\ell^1, \|\ \|_\infty)$.

Les éléments de ℓ^1 sont des suites $x \equiv (x_n)_{n\in\mathbb{N}}$, donc une suite d'éléments de ℓ^1 est une suite de suites. Pour tout $m \in \mathbb{N}$ fixé, nous allons écrire cette suite comme ça : $x^m \equiv (x_n^m)_{n\in\mathbb{N}}$.

Allons vérifier que la suite d'éléments de ℓ^1 définie par :

$$x_n^m = \begin{cases} 0 & \text{si } n = 0 \\ \frac{1}{n} & \text{si } 1 \leqslant n \leqslant m \\ 0 & \text{si } n > m, \end{cases}$$

converge en $\ell^\infty \setminus \ell^1$. Pour tout $m \in \mathbb{N}$ fixé, la suite x^m est définie explicitement comme ça :

$$\left(0, 1, \frac{1}{2}, \ldots, \frac{1}{m}, 0, 0, \ldots\right),$$

ce qui montre que $x^m \in \ell^1$ pour tout $m \in \mathbb{N}$ fixé.

Considérons maintenant la suite $x^* \equiv (x_n^*)_{n\in\mathbb{N}}$ définie par

$$x_n^* = \begin{cases} 0 & \text{si } n = 0 \\ \frac{1}{n} & \text{si } n \geqslant 1. \end{cases}$$

Évidemment $(x_n^*)_{n\in\mathbb{N}}$ est bornée et donc elle appartient à ℓ^∞, mais $\|(x_n^*)_{n\in\mathbb{N}}\|_1 = \sum_{n=1}^{\infty} \frac{1}{n} = +\infty$, donc $(x_n^*)_{n\in\mathbb{N}} \notin \ell^1$. Si nous montrons que $(x^m)_{m\in\mathbb{N}}$ converge vers x^* en norme $\| \ \|_\infty$, l'exercice sera résolu.

Pour cela, calculons :

$$\|x^m - x^*\|_\infty = \sup_{n\in\mathbb{N}} |x_n^m - x_n^*| = \sup_{n>m} \frac{1}{n},$$

en fait, jusqu'à $n = m$, la différence $x_n^m - x_n^*$ est nulle, mais quand $n > m$, la différence devient $|0 - \frac{1}{n}| = \frac{1}{n}$. Par définition de sup, $\sup_{n>m} \frac{1}{n} = \sup\{\frac{1}{m+1}, \frac{1}{m+2}, \ldots\} = \frac{1}{m+1}$ et donc :

$$\|x^m - x^*\|_\infty = \frac{1}{m+1} \underset{m\to+\infty}{\longrightarrow} 0.$$

□

4.4.4 Relations d'inclusion dans les espaces L^p

En général, il n'existe aucune relation d'inclusion parmi les espaces $L^p(X, \mathcal{A}, \mu)$. En fait, par exemple, considérons $L^1(\mathbb{R})$, $L^2(\mathbb{R})$ et les fonctions :

$$f(x) = \begin{cases} x^{-2/3} & \text{si } 0 < x < 1 \\ 0 & \text{autrement} \end{cases}, \quad g(x) = \begin{cases} x^{-2/3} & \text{si } x > 1 \\ 0 & \text{autrement.} \end{cases}$$

Il est clair que $f \in L^1(\mathbb{R})$, mais $f \notin L^2(\mathbb{R})$, car [10]

$$\int_{\mathbb{R}} |f(x)|dx = \int_0^1 \frac{1}{x^{2/3}}dx < +\infty, \quad \int_{\mathbb{R}} |f(x)|^2 dx = \int_0^1 \frac{1}{x^{4/3}}dx = +\infty$$

et que $g \in L^2(\mathbb{R})$, mais $g \notin L^1(\mathbb{R})$, car

$$\int_{\mathbb{R}} |g(x)|dx = \int_1^{+\infty} \frac{1}{x^{2/3}}dx = +\infty, \quad \int_{\mathbb{R}} |g(x)|^2 dx = \int_1^{+\infty} \frac{1}{x^{4/3}}dx < +\infty.$$

Nous obtenons des inclusions parmi les espaces L^p en rajoutant des conditions supplémentaires. Vu que les espaces $L^1(\mathbb{R})$ et $L^2(\mathbb{R})$ sont particulièrement importants, nous examinons des conditions, qui sont souvent vérifiées dans les applications, dans le théorème suivant.

Théorème 4.4.4 *Les affirmations suivantes sont vraies :*

1. *$f \in L^1(\mathbb{R})$, f bornée, alors $f \in L^2(\mathbb{R})$;*

2. *$f \in L^2(\mathbb{R})$, f nulle au-dehors d'un intervalle fini, alors $f \in L^1(\mathbb{R})$.*

10. Nous rappelons que, si $a > 0$ et $b \in \mathbb{R}$, $\int_0^a \frac{1}{x^\alpha}dx < +\infty$ et $\int_b^{+\infty} \frac{1}{x^\beta}dx < +\infty$ si et seulement si $\alpha < 1$ et $\beta > 1$.

Preuve.

1. Si f est en $L^1(\mathbb{R})$ et elle est bornée, disons $|f(x)| \leqslant M\ \forall x \in \mathbb{R}$, $M \geqslant 0$, alors :

$$\int_{\mathbb{R}} |f(x)|^2 dx = \int_{\mathbb{R}} |f(x)| \cdot |f(x)| dx \leqslant \int_{\mathbb{R}} M|f(x)| dx = M\|f\|_1 < +\infty,$$

donc $f \in L^2(\mathbb{R})$.

2. Si f est en $L^2(\mathbb{R})$ et elle est nulle au-dehors d'un intervalle fini, disons $f(x) = 0$ $\forall x \notin [a,b]$, alors :

$$\begin{aligned}\int_{\mathbb{R}} |f(x)| dx = \int_{[a,b]} |f(x)| dx &\underset{\mathbf{1}(x)=1\ \forall x \in [a,b]}{=} \int_{[a,b]} \mathbf{1}(x) \cdot |f(x)| dx = \langle \mathbf{1}, |f| \rangle_{L^2[a,b]} \\ &\underset{\text{(Cauchy-Schwarz)}}{\leqslant} \left(\int_{[a,b]} dx \right)^{1/2} \left(\int_{[a,b]} |f(x)|^2 dx \right)^{1/2} = \sqrt{b-a}\, \|f\|_2 < +\infty,\end{aligned}$$

donc $f \in L^1(\mathbb{R})$. □

L'affirmation 1. reste valide pour toute $f \in L^1(\mathbb{R}^n)$, $n \geqslant 1$, l'affirmation 2. reste valide si l'on remplace un intervalle avec une partie de mesure finie de $\mathbb{R}^n$.

Plus en général, dans le cas particulier où $\mu(X) < +\infty$, on peut arriver à une chaîne d'inclusions très utile.

Théorème 4.4.5 *Si $(X, \mathcal{A}, \mu)$ est un espace mesuré avec **mesure finie**, $\mu(X) < +\infty$, et si $q > p > 1$, alors :*

$$\boxed{L^\infty(X, \mathcal{A}, \mu) \subset \ldots \subset L^q(X, \mathcal{A}, \mu) \subset \ldots \subset L^p(X, \mathcal{A}, \mu) \subset \ldots \subset L^1(X, \mathcal{A}, \mu)}.$$

Preuve. On va vérifier la thèse d'abord dans le cas de L^∞, ensuite L^1 et L^2 (qui permet de comprendre très clairement l'astuce qu'on va utiliser), et enfin L^p et L^q.
Si $f \in L^\infty(X, \mathcal{A}, \mu)$, alors $\int_X |f|^p d\mu \leqslant \int_X \|f\|_\infty^p d\mu = \|f\|_\infty^p \mu(X) < +\infty$, donc $f \in L^p(X, \mathcal{A}, \mu)$.
Si $f \in L^2(X, \mathcal{A}, \mu)$, alors

$$\int_X |f| d\mu = \int_X |\mathbf{1} \cdot f| d\mu \underset{\text{inég. Hölder (4.19)}}{\leqslant} \left(\int_X 1^2 d\mu \right)^{\frac{1}{2}} \left(\int_X |f|^2 d\mu \right)^{\frac{1}{2}} = \sqrt{\mu(X)} \|f\|_2 < +\infty.$$

donc $f \in L^1(X, \mathcal{A}, \mu)$.

Soient $E = \{x \in X\ :\ |f(x)| \geqslant 1\}$ et $F = \{x \in X\ :\ |f(x)| \leqslant 1\}$, alors $X = E \cup F$, et soit $p < q$. Alors $|f(x)|^p \leqslant |f(x)|^q\ \forall x \in E$ et $|f(x)|^p \leqslant 1\ \forall x \in F$. Donc, si $f \in L^q$:

$$\int_X |f(x)|^p d\mu \leqslant \int_E |f(x)|^q d\mu + \int_F 1\, d\mu \leqslant \int_X |f(x)|^q d\mu + \int_X 1\, d\mu = \|f\|_q^q + \mu(X) < +\infty,$$

i.e $f \in L^p$. □

4.4.5 Théorèmes de densité dans L^p(X,A,μ)

Nous commençons la description des variétés denses en L^p avec les fonctions étagées.

Fonctions étagées Soit $(X, \mathcal{A}, \mu)$ un espace mesuré quelconque et $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Une fonction constante par morceaux sur X à valeurs en $\mathbb{K}$ est dite étagée ou **simple**. Pour tout $N \in \mathbb{N}$, la définition rigoureuse de l'espace de ces fonctions est la suivante :

$$\Sigma = \left\{ s : X \to \mathbb{K} \ : \ \exists (\alpha_i)_{i=1}^N \in \mathbb{K} \ : \ s = \sum_{i=1}^N \alpha_i \chi_{E_i}, \ E_i \text{ mesurable et } \mu(E_i) < +\infty \text{ si } \alpha_i \neq 0 \right\},$$

la fonction $\chi_{E_i} = \begin{cases} 1 \text{ si } x \in E_i \\ 0 \text{ si } x \notin E_i \end{cases}$ est la fonction indicatrice de E_i.

Théorème 4.4.6 $\boxed{\overline{\Sigma} = L^p(X, A, \mu)}$ $\forall 1 \leqslant p < \infty$, *où l'adhérence est à interpréter par rapport à la topologie de* $L^p(X, A, \mu)$ *en considérant* $\Sigma \subset L^p(X, A, \mu)$.

Intersections $L^p \cap L^q$ et $\ell^p \cap \ell^q$

Théorème 4.4.7 *Soit* $(X, \mathcal{A}, \mu)$ *un espace mesuré quelconque et* $\mathbb{K} = \mathbb{R}$ *ou* $\mathbb{C}$, *alors :*

$$\boxed{\overline{L^p(X, \mathcal{A}, \mu) \cap L^q(X, \mathcal{A}, \mu)} = \begin{cases} L^p(X, \mathcal{A}, \mu) \\ L^q(X, \mathcal{A}, \mu) \end{cases}} \qquad \forall 1 \leqslant p, q \leqslant \infty,$$

où, dans le premier cas, l'intersection est à interpréter comme un sous-ensemble de $L^p(X, A, \mu)$ *et l'adhérence par rapport à la topologie métrique engendrée par la norme* $\| \ \|_p$ *; dans le deuxième cas l'intersection est à interpréter comme un sous-ensemble de* $L^q(X, A, \mu)$ *et l'adhérence par rapport à la topologie relative à la norme* $\| \ \|_q$.

En particulier, comme les espaces ℓ^p sont emboîtés, ça vaut que :

$$\boxed{\overline{\ell^p(\mathbb{N}, \mathbb{K})} = \ell^q(\mathbb{N}, \mathbb{K})} \qquad \forall 1 \leqslant p < q < \infty.$$

Comme avant, pour tout q fixé, il faut interpréter ℓ^p comme un sous-espace de ℓ^q et interpréter l'adhérence par rapport à la norme $\| \ \|_q$.

Théorème 4.4.8 *Pour tout* $p \in \mathbb{R}$, $1 \leqslant p < +\infty$ *:*

$$\boxed{\overline{\ell^0(\mathbb{N}, \mathbb{K})} = \ell^p(\mathbb{N}, \mathbb{K})},$$

i.e. $\ell^0(\mathbb{N}, \mathbb{K})$ *est dense en* $\ell^p(\mathbb{N}, \mathbb{K})$ *par rapport à la topologie engendrée par la norme* $\| \ \|_p$.

Preuve. Soit $(x_n)_{n\in\mathbb{N}}$ une suite arbitraire de $\ell^p(\mathbb{N},\mathbb{K})$. Considérons la suite

$$x_n^N := \begin{cases} x_n & \text{si } n < N \\ 0 & \text{sinon,} \end{cases}$$

alors

$$\|x_n - x_n^N\|_p = \sum_{n\in\mathbb{N}} |x_n - x_n^N|^p = \sum_{n=N}^{+\infty} |x_n|^p \underset{N\to+\infty}{\to} 0,$$

car c'est le reste d'une série convergente (vu que $(x_n)_{n\in\mathbb{N}}$ appartient à $\ell^p(\mathbb{N},\mathbb{K})$), i.e. la densité de $\ell^0(\mathbb{N},\mathbb{K})$ en $\ell^p(\mathbb{N},\mathbb{K})$. □

Fonctions test Soit $\Omega \subseteq \mathbb{R}^n$ un ouvert.

Déf. 4.4.1

$$\mathcal{C}_c^\infty(\Omega) = \{f : \Omega \to \mathbb{K},\ f \text{ indéfiniment dérivable sur } \Omega \text{ et } \mathrm{supp}(f) \text{ compact en } \mathbb{R}^n\}$$

où $supp(f) = \overline{\{x \in \Omega : f(x) \neq 0\}}$ *est dit le support de* f*.*

Les fonctions de $\mathcal{C}_c^\infty(\Omega)$ sont dites **fonctions test**, car elles sont tellement régulières, qu'elles sont utilisées pour tester l'action et les propriétés de certains opérateurs « sauvages ». Les fonctions test jouent un rôle fondamental dans la théorie de distributions et dans l'analyse des équations différentielles. La fonction identiquement nulle est évidemment une fonction test, sinon, il n'est pas simple de donner des exemples explicites de fonctions test. Pour tout valeur de $\varepsilon > 0$, l'exemple canonique de fonction test sur $\mathbb{R}$ est donné par :

$$f(x) = \begin{cases} \exp\left(-\frac{1}{1-\left(\frac{x}{\varepsilon}\right)^2}\right) & \text{si } |x| < \varepsilon \\ 0 & \text{si } |x| \geqslant \varepsilon. \end{cases}$$

Pour la suite nous allons avoir besoin d'un symbole simple pour dénoter la dérivée partielle d'une fonction de n variables par rapport à un *multi-indice* $\mathbf{l} = (l_1, l_2, \ldots, l_d) \in \mathbb{N}^d$ de longueur $|\mathbf{l}| = l_1 + l_2 + \ldots + l_d$. La notation canonique est la suivante :

$$D^{\mathbf{l}} f(x) = \frac{\partial^{|\mathbf{l}|} f}{\partial x_1^{l_1} \partial x_2^{l_2} \ldots \partial x_d^{l_d}}(x) \qquad \forall x \in \mathbb{R}^n,$$

donc $D^{\mathbf{l}} f(x)$ est la dérivée partielle de f en x l_1 fois par rapport à x_1, l_2 fois par rapport à x_2, etc. Ce symbole intervient dans la définition d'une topologie (non banale) sur $\mathcal{C}_c^\infty(\Omega)$ par rapport à laquelle la convergence d'une suite $(f_n)_{n\in\mathbb{N}}$ de fonctions tests à la fonction test f est équivalente à la réalisation de ces deux conditions :

— il existe un ensemble compact $K \subset \Omega$ tel que $\mathrm{supp}(f_n) \subseteq K$ pour tout $n \in \mathbb{N}$;
— $\forall x \in \mathbb{R}^n$, $\forall \mathbf{l} \in \mathbb{N}^d : D^{\mathbf{l}} f_n(x) \underset{n\to+\infty}{\to} D^{\mathbf{l}} f(x)$, uniformément.

L'espace $\mathcal{C}_c^\infty(\Omega)$ muni de cette topologie s'écrit habituellement $\mathcal{D}(\Omega)$ et il est **complet**. Il vaut le résultat de densité suivant.

Théorème 4.4.9 *Si on considère la tribu Borélienne et la mesure de Lebesgue, alors :*

$$\boxed{\overline{\mathcal{C}_c^\infty(\Omega)} = L^p(\Omega)} \qquad \forall 1 \leqslant p < \infty,$$

où, comme d'habitude, il faut interpréter $\mathcal{C}_c^\infty(\Omega)$ comme un sous-espace de $L^p(\Omega)$ et considérer l'adhérence par rapport à la topologie engendrée par la norme $\| \ \|_p$. Par définition d'adhérence, $\mathcal{C}_c^\infty(\Omega)$ *n'est pas complet par rapport à la topologie engendrée par la norme* $\| \ \|_p$, vu qu'il y a des suites d'éléments de $\mathcal{C}_c^\infty(\Omega)$ qui convergent vers des éléments de $L^p(\Omega) \backslash \mathcal{C}_c^\infty(\Omega)$.

L'espace de Schwartz Nous allons commencer avec le cas d'une fonction d'une seule variable réelle pour simplicité, surtout par rapport à la notation.

Soient $k, l \in \mathbb{N}$ et $f \in \mathcal{C}^\infty(\mathbb{R})$, alors, pour tout $x \in \mathbb{R}$, on écrit

$$f^{k,l}(x) = x^k \frac{d^l f}{dx^l}(x).$$

Déf. 4.4.2 (Espace de Schwartz, $n = 1$**)** *L'espace des fonctions* $f \in \mathcal{C}^\infty(\mathbb{R})$ *telles que*

$$\lim_{|x| \to +\infty} |f^{k,l}(x)| = 0 \qquad \forall k, l \in \mathbb{N}.$$

est dit espace de Schwartz, ou espace des fonctions à décroissance rapide. La notation canonique pour cet espace est $\mathcal{S}(\mathbb{R})$.

Un élément quelconque de $\mathcal{S}(\mathbb{R})$ est, donc, une fonction infiniment dérivable partout telle que, si on considère sa dérivées de n'importe quel ordre et on la multiplie par n'importe quelle puissances de sa variable, elle convergent vers 0 quand la variable tend vers $\pm\infty$. Pour vérifier cette caractéristique, une fonction doit décroître très rapidement à l'infini, d'où l'appellation alternative des fonctions de $\mathcal{S}(\mathbb{R})$.

Il est évident que $\mathcal{D}(\mathbb{R}) \subset \mathcal{S}(\mathbb{R})$, car les fonctions test sont nulles à l'infini, mais l'inclusion est stricte, comme le montre l'exemple le plus important de fonction à décroissance rapide, i.e. la Gaussienne $f(x) = e^{-x^2}$, qui n'appartient pas à $\mathcal{D}(\mathbb{R})$, car son support n'est pas compact.

Considérons maintenant une fonction de n variables réelles $f \in \mathcal{C}^\infty(\mathbb{R}^n)$. Dans ce cas, étant donnés deux multi-indices $\mathbf{l}, \mathbf{k} \in \mathbb{N}^n$, nous écrivons

$$f^{\mathbf{k},\mathbf{l}}(x) = x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n} D^{\mathbf{l}} f(x) \qquad \forall x \in \mathbb{R}^n.$$

Déf. 4.4.3 (Espace de Schwartz, n **arbitraire)** *L'espace des fonctions* $f \in \mathcal{C}^\infty(\mathbb{R}^n)$ *telles que*

$$\lim_{\|x\| \to +\infty} |f^{\boldsymbol{k},\boldsymbol{l}}(x)| = 0 \qquad \forall \boldsymbol{k}, \boldsymbol{l} \in \mathbb{N}^n,$$

est dit espace de Schwartz, ou espace des fonctions à décroissance rapide. La notation canonique pour cet espace est $\mathcal{S}(\mathbb{R}^n)$.

Par construction, $\mathcal{S}(\mathbb{R}^n)$ est stable par rapport à la dérivation partielle et par rapport à la multiplication par un polynôme. Le fait que les fonctions de $\mathcal{S}(\mathbb{R}^n)$ (et leurs dérivées) décroissent plus rapidement que l'inverse d'un polynôme, explique pourquoi souvent on les appelle des *fonctions à décroissance plus que polynomiale.*

Comme dans le cas $n = 1$, $\mathcal{D}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$ et l'inclusion est stricte, car la Gaussienne $f(x) = e^{-\|x\|^2}$ appartient à $\mathcal{S}(\mathbb{R}^n)$, mais elle n'appartient pas à $\mathcal{D}(\mathbb{R}^n)$.

Sur $\mathcal{S}(\mathbb{R}^n)$ il est possible de définir une topologie dans laquelle une suite $(f_n)_{n\in\mathbb{N}}$ de fonctions de $\mathcal{S}(\mathbb{R}^n)$ converge vers $f \in \mathcal{S}(\mathbb{R}^n)$ si $f_n^{\mathbf{k},\mathbf{l}} \underset{n\to+\infty}{\to} f$ uniformément $\forall \mathbf{k}, \mathbf{l} \in \mathbb{N}^n$. *Par rapport à cette topologie*, **l'espace de Schwartz est complet**.

Comme pour l'espace des fonctions test, l'espace de Schwartz apparaît dans la théorie des distributions (qui a été formalisée par Laurent Schwartz lui-même) et dans celle des équations différentielles en dérivées partielles.

Le fait que $\mathcal{D}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$ et que $\mathcal{D}(\mathbb{R}^n)$ est $\| \ \|_p$-dense en $L^p(\mathbb{R}^n)$ implique le résultat suivant.

Théorème 4.4.10 *Si on considère la tribu Borelienne et la mesure de Lebesgue, alors :*

$$\boxed{\overline{\mathcal{S}(\mathbb{R}^n)} = L^p(\mathbb{R}^n)} \qquad \forall 1 \leqslant p < \infty,$$

en interprétant $\mathcal{S}(\mathbb{R}^n)$ comme un sous-espace de $L^p(\mathbb{R}^n)$ et en considérant l'adhérence par rapport à la topologie engendrée par la norme de $\| \ \|_p$.

Par définition d'adhérence, $\mathcal{S}(\mathbb{R}^n)$ *n'est pas complet par rapport à la topologie engendrée par la norme* $\| \ \|_p$: il y a des suites de $\mathcal{S}(\mathbb{R}^n)$ qui convergent vers des éléments de $L^p(\mathbb{R}^n)\backslash\mathcal{S}(\mathbb{R}^n)$.

4.5 Résumé du chapitre 4

- Nous avons examiné la compatibilité entre la structure topologique d'un espace vectoriel avec produit scalaire et la structure linéaire : les opérations de somme et multiplication par un scalaire sont continues dans la topologie engendrée par le produit scalaire, ainsi que le produit scalaire lui-même et la norme induite canoniquement. Cette compatibilité est essentielle car elle permet d'échanger la limite avec les opérations ci-dessus, ce qui est fondamentale pour la théorie et les applications.

- On a vu que les espaces vectoriels de dimension finie ont tous la même structure topologique : celle Euclidienne.

- Nous avons introduit les espaces de Hilbert et de Banach comme les espaces vectoriels avec produit scalaire, ou norme, respectivement, tels que toutes les suites de Cauchy convergent dans l'espace lui-même (complétude). Tout espace vectoriel avec produit scalaire (resp. normé) de dimension finie est un espace de Hilbert (resp. de Banach). Tout espace de Hilbert est de Banach, mais le vice-versa n'est pas vrai en général.

- Il existe une caractérisation très utile des espace vectoriel normés complets : ils sont touts et seulement les espaces dans lesquels les suites absolument convergentes sont aussi simplement convergentes.

- Pour toute contraction définie sur un espace métrique complet il existe toujours un point fixe et celui-ci est unique.

- Exemples d'espaces de Hilbert sont L^2 et ℓ^2, exemples d'espaces de Banach qui ne sont pas de Hilbert sont L^p et ℓ^p, avec $1 \leqslant p \leqslant \infty$, $p \neq 2$. L'inégalité de Minkowski permet de définir une structure linéaire sur tous ces espaces, celle de Hölder permet de définir un produit scalaire quand $p = 2$.

- Les espaces ℓ^p sont emboîtés l'un dans l'autre avec p croissant, par contre, il n'y a pas de relation d'inclusion en général dans les espaces L^p, avec l'exception remarquable des espaces de mesure finie, pour lesquels les espaces L^p sont emboîtés, mais en sens contraire des espaces ℓ^p, i.e. avec p décroissant.

- Pour terminer, on a montré que les espaces L^p coïncident avec la fermeture de nombreux espaces fonctionnels très utilisés, comme l'espace des fonctions test et l'espace de Schwartz.

Chapitre 5

La structure géométrique des espaces de Hilbert

Dans ce chapitre on montrera que, parmi les espaces vectoriels de dimension infinie, les espaces de Hilbert ont la structure géométrique la plus proche de celles des espaces Euclidiens $\mathbb{K}^n$, qu'on a examiné dans le premier chapitre.

Cela n'est plus vrai pour les espaces de Banach de dimension infinie, qui peuvent avoir des propriétés structurales beaucoup plus compliquées que celles des espaces de Hilbert.

La structure géométrique riche des espaces de Hilbert nous permettra d'étendre la transformée de Fourier discrète à des espaces de dimension infinie avec le concept de série et de transformée de Fourier continue.

Des ouvrages conseillés pour approfondir les sujets traités dans ce chapitre et dans les chapitres suivants sont [1], [3], [6], [11] et [12].

Le premier pas dans l'analyse de la structure géométrique des espaces de Hilbert consiste à étudier le concept de complément orthogonal.

5.1 Le complément orthogonal dans un espace de Hilbert et ses propriétés

L'ensemble des vecteurs orthogonaux aux vecteurs d'un sous-ensemble d'un espace de Hilbert joue un rôle fondamental pour comprendre les propriétés géométriques de ces espaces.

Allons introduire la définition et les propriétés de cet ensemble.

Déf. 5.1.1 *Soit $\mathcal{H}$ un espace de Hilbert et $M \subseteq \mathcal{H}$ un sous ensemble quelconque. Le complément orthogonal de M en $\mathcal{H}$ est :*

$$M^{\perp} = \{x \in \mathcal{H} : \langle x, y\rangle = 0 \ \ \forall\, y \in M\}\,,$$

i.e. $M^{\perp}$ contient tous les éléments de $\mathcal{H}$ qui sont orthogonaux à tout élément de M.

Examinons ses propriétés, après avoir rappelé que $\mathrm{span}(M)$ est le sous espace vectoriel de $\mathcal{H}$ engendré par M, i.e. l'ensemble des combinaisons de linéaires (finies) de vecteurs de M. Dans le théorème qui suit on écrira $(M^{\perp})^{\perp} = M^{\perp\perp}$ et $(M^{\perp\perp})^{\perp} = M^{\perp\perp\perp}$.

Théorème 5.1.1 (Propriétés du complément orthogonal) *Soit $\mathcal{H}$ un espace de Hilbert et $M \subseteq \mathcal{H}$ un* ***sous ensemble*** *quelconque. Alors :*

1. $\{0_\mathcal{H}\}^\perp = \mathcal{H}$ *et* $\mathcal{H}^\perp = \{0_\mathcal{H}\}$

2. $M \cap M^\perp = \begin{cases} \{0_\mathcal{H}\} & \text{si } 0_\mathcal{H} \in M \\ \varnothing & \text{si } 0_\mathcal{H} \notin M \end{cases}$

3. $M^\perp$ *est un sous espace vectoriel fermé de* $\mathcal{H}$

4. *Si* $N \subseteq \mathcal{H}$*, alors* $M \subseteq N \Rightarrow N^\perp \subseteq M^\perp$ ($\perp$ reverse les relations d'inclusion entre ensembles)

5. $M \subseteq M^{\perp\perp}$ (différence par rapport à la dimension finie)

6. $(\overline{M})^\perp = M^\perp$

7. $M^{\perp\perp\perp} = M^\perp$

8. $M^\perp = (\operatorname{span} M)^\perp = (\overline{\operatorname{span} M})^\perp$

9. *Si* $\overline{M} = \mathcal{H} \Rightarrow M^\perp = \{0_\mathcal{H}\}$ (le complément orthogonal d'un sous-ensemble dense est le vecteur nul).

Avant de développer la preuve on observe que le fait que $M^\perp$ soit toujours fermé est une propriété extrêmement utile pour démontrer qu'une variété linéaire de $\mathcal{H}$ est fermée : il suffit de démontrer qu'elle coïncide avec le complément orthogonal d'un sous-ensemble de $\mathcal{H}$.

Preuve.

1. La propriété descend du fait que $0_\mathcal{H}$ est le seul vecteur de $\mathcal{H}$ orthogonale à tous les autres.

2. $0_\mathcal{H}$ est le seul vecteur orthogonal à lui-même.

3. $M^\perp$ est un *sous-espace vectoriel* : en fait, si $x, x' \in M^\perp$, alors $\langle x, y\rangle = \langle x', y\rangle = 0$ $\forall y \in M$, donc $\langle \alpha x + \beta x', y\rangle = \alpha\langle x, y\rangle + \beta\langle x', y\rangle = 0$ $\forall y \in M$, i.e. $M^\perp$ est un sous-espace vectoriel car il est stable par rapport aux combinaison linéaires de ses éléments.

   $M^\perp$ est *fermé* : il faut montrer que $M^\perp$ contient tous les points limites de suites en $M^\perp$. Soit $(x_n)_{n\in\mathbb{N}} \subset M^\perp$ une suite convergente (et donc de Cauchy) vers la limite x, alors, comme $M^\perp \subseteq \mathcal{H}$, et $\mathcal{H}$ est complet, $x \in \mathcal{H}$. Pour tout $y \in M$, $\langle x_n, y\rangle = 0$ $\forall n \in \mathbb{N}$, donc, grâce à la continuité du produit scalaire on peut écrire :

   $$0 = \lim_{n\to+\infty} \langle x_n, y\rangle = \Big\langle \lim_{n\to+\infty} x_n, y\Big\rangle = \langle x, y\rangle$$

   donc $x \perp y$ $\forall y \in M$, i.e. $x \in M^\perp$.

4. Comme $M \subseteq N$, les vecteurs de $\mathcal{H}$ orthogonaux aux vecteurs de N sont aussi orthogonaux aux vecteurs de M (mais le contraire n'est pas nécessairement vrai) donc $y \in \mathcal{H}$, $y \in N^\perp$ implique $y \in M^\perp$, i.e. $N^\perp \subseteq M^\perp$.

5. Tout vecteur de M est orthogonale à tout vecteur de $M^\perp$ par définition, mais on peut avoir aussi d'autres vecteurs de $\mathcal{H}$ orthogonaux à $M^\perp$, donc $M \subseteq M^{\perp\perp}$.

6. On démontrera l'égalité des ensembles en démontrant les deux inclusions en sens inverse :

• $(\overline{M})^\perp \subseteq M^\perp$: cela découle de $M \subseteq \overline{M}$ et de la propriété **4.**

• $M^\perp \subseteq (\overline{M})^\perp$: il faut démontrer que $y \in M^\perp \implies y \in (\overline{M})^\perp$. Allons traduire cela : les éléments de $\overline{M}$ sont l'union de tous les éléments de M avec les limites des suites en M, donc il faut démontrer que si $y \in M$ est orthogonale à tous les éléments $(x_n)_{n\in\mathbb{N}} \subset M^\perp$ d'une suite convergente arbitraire en $M^\perp$, alors y à orthogonale aussi à la limite de cette suite. Pour prouver ceci on utilise encore la continuité du produit scalaire : par hypothèse $\langle x_n, y\rangle = 0 \ \forall n \in \mathbb{N}$, donc :

$$0 = \lim_{n\to+\infty} \langle x_n, y\rangle = \langle \lim_{n\to+\infty} x_n, y\rangle,$$

i.e. $y \perp \lim\limits_{n\to+\infty} x_n$.

7. Par **5.** $M \subseteq M^{\perp\perp}$ et, par **4.**, $M^{\perp\perp\perp} \subseteq M^\perp$ pour n'importe quel sous-ensemble M de $\mathcal{H}$. Si maintenant on écrit $N = M^\perp$, on peut réécrire la dernière inclusion comme $N^{\perp\perp} \subseteq N$, ce qui implique, par **4.**, $N^\perp \subseteq N^{\perp\perp\perp}$. M, et donc N, sont des sous-ensembles arbitraires de $\mathcal{H}$, par conséquent les inclusions $M^{\perp\perp\perp} \subseteq M^\perp$, $N^\perp \subseteq N^{\perp\perp\perp}$ impliquent l'égalité entre un sous-ensemble de $\mathcal{H}$ et son complément tri-orthogonal.

8. Considérons un élément arbitraire de $\text{span}(M)$: $y_0 = \sum\limits_{i=1}^{n} \alpha_i y_i$, $y_i \in M$ et $\alpha_i \in \mathbb{K}$ $\forall i = 1, \ldots, n$. Soit $x \in M^\perp$ quelconque fixé, alors, grâce à la sesquilinéairité (ou bilinéarité) de $\langle\ ,\ \rangle$ (selon $\mathbb{K} = \mathbb{C}$ ou $\mathbb{K} = \mathbb{R}$) on peut écrire :

$$\langle x, y_0\rangle = \langle x, \sum_{i=1}^{n} \alpha_i y_i\rangle = \sum_{i=1}^{n} \overline{\alpha_i}\langle x, y_i\rangle \underset{x\perp y_i}{=} 0,$$

donc $x \in (\text{span}(M))^\perp$, mais comme x est arbitraire, ceci implique : $M^\perp \subseteq (\text{span}(M))^\perp$. Néanmoins, $M \subseteq \text{span}(M)$, donc, grâce à **4.**, $(\text{span}(M))^\perp \subseteq M^\perp$, i.e. $(\text{span}(M))^\perp = M^\perp$, en plus, grâce à **6.**, $(\text{span}(M))^\perp = (\overline{\text{span}(M)})^\perp = M^\perp$.

8. $M^\perp = (\overline{M})^\perp = \mathcal{H}^\perp = \{0_\mathcal{H}\}$, car le seul vecteur orthogonal à tous les vecteurs de $\mathcal{H}$ est le vecteur nul. □

5.2 Le théorème de projection sur un convexe fermé et ses conséquences

On va maintenant examiner une propriété géométrique très importante qu'on connaît déjà dans le cas d'un espace Euclidien : l'existence de la projection orthogonale d'un vecteur, interprétée comme le vecteur d'un sous-espace le plus proche à un vecteur donné.

On va énoncer et prouver le résultat sur un sous-ensemble plus général qu'un sous-espace vectoriel : un sous-ensemble convexe fermé.

Déf. 5.2.1 *Un sous-ensemble S d'un espace vectoriel est dit* ***convexe*** *si*

$$\forall x, y \in S,\ \forall \lambda \in [0,1]\ :\ \lambda x + (1-\lambda) y \in S,$$

i.e. si S est stable par rapport aux combinaisons convexes, i.e combinaisons linéaires pour lesquelles la somme des coefficients est 1.

Géométriquement, un sous-espace convexe peut être caractérisé par le fait que toute couple de points peut être connecté via une segment de droite, tout en restant dans le sous-espace même. On observe que, évidemment, *un sous-espace vectoriel est toujours convexe*, car il est stable par rapport à toute combinaison linéaire, et donc, en particulier, à celles convexes.

On remarque que la *demi somme* entre x et y, i.e. $\frac{x+y}{2}$ est une combinaison convexe avec $\lambda = 1/2$.

Théorème 5.2.1 *Soit $\mathcal{H}$ un espace de Hilbert et S un sous-ensemble propre*[1] *convexe et fermé de $\mathcal{H}$. Alors, $\forall x \in \mathcal{H}$ (fixé) il existe un seul point $y_0 \in S$ tel que :*

$$\|x - y_0\| = \inf_{y \in S} \|x - y\|,$$

i.e y_0 minimise la distance entre x et les points de S.

Avant de donner la preuve du théorème nous remarquons que *ce résultat est vrai aussi pour tout sous-espace vectoriel fermé de $\mathcal{H}$* et, donc, que le théorème de projection sur un convexe fermé généralise la propriété 3) du théorème 1.6.1 aux espaces de Hilbert de dimension infinie.

Déf. 5.2.2 *Le vecteur y_0 du théorème précédent est dit la* ***projection orthogonale*** *de x sur S et indiqué avec : $y_0 = P_S(x)$. La quantité réelle non-négative $d(x,S) = \|x - P_S(x)\|$ est dite la* ***distance entre*** *x* ***et le sous-ensemble propre convexe et fermé*** *S.*

Évidemment, si $x \in S$ alors $P_S(x) = x$ et $d(x,S) = 0$, donc le théorème a un intérêt quand $x \notin S$.

Preuve.
$\boxed{\exists}$: notons, par simplicité de lecture[2], $\delta \equiv \inf_{y \in S} \|x - y\|$. Nous allons démontrer l'existence de y_0 avec une technique non-constructive typique de l'école de Hilbert. Considérons une suite $(y_n)_{n \in \mathbb{N}} \subset S$ qui satisfait l'équation[3] :

$$\lim_{n \to +\infty} \|x - y_n\| = \delta. \tag{5.1}$$

1. Si $S = \mathcal{H}$, alors le théorème est banalement vérifié avec $y_0 = x$.
2. $\delta(x,S)$ serait une notation plus correcte, car δ, en général, change, en général, avec x et S.
3. Par exemple, $\delta^2 \leqslant \|x - y_n\|^2 \leqslant \delta^2 + \frac{1}{n}$, $\forall n \in \mathbb{N}$.

L'intérêt d'examiner une telle suite est que, grâce à la continuité de la norme, nous pouvons réécrire (5.1) comme ça :

$$\delta = \left\| \lim_{n\to+\infty}(x-y_n) \right\| = \left\| \lim_{n\to+\infty} x - \lim_{n\to+\infty} y_n \right\| = \left\| x - \lim_{n\to+\infty} y_n \right\|,$$

et alors, pour démontrer l'existence de y_0, il suffit de poser : $y_0 := \lim_{n\to+\infty} y_n$.

Pour faire ça, on commence par observer que S est fermé et donc *est lui même complet*, alors, pour démontrer l'existence de la limite de y_n, ça suffit de démontrer que $(y_n)_{n\in\mathbb{N}}$ est une suite de Cauchy en S.

Nous démontrerons que $(y_n)_{n\in\mathbb{N}}$ est une suite de Cauchy à l'aide de la formule du parallélogramme (1.2.4) (qui vaut car la norme est hilbertienne, cfr. théorème 4.2.2) pour les éléments $x-y_n$ et $x-y_m$:

$$\|(x-y_n)+(x-y_m)\|^2 + \|(x-y_n)-(x-y_m)\|^2 = 2(\|(x-y_n)\|^2 + \|(x-y_m)\|^2),$$

qui peut être réécrite comme ça :

$$\|2x-y_n-y_m\|^2 + \|y_n-y_m\|^2 = 2(\|(x-y_n)\|^2 + \|(x-y_m)\|^2),$$

i.e. :

$$\|y_n-y_m\|^2 = 2(\|(x-y_n\|^2 + \|x-y_m\|^2) - \|2x-y_n-y_m\|^2 \tag{5.2}$$

mais $2x-y_n-y_m = 2\left(x-\frac{1}{2}(y_n+y_m)\right)$, donc (5.2) devient :

$$\|y_n-y_m\|^2 = 2(\|(x-y_n\|^2 + \|x-y_m\|^2) - 4\left\|x-\frac{1}{2}(y_n+y_m)\right\|^2, \tag{5.3}$$

mais $\frac{1}{2}(y_n+y_m) \in S$ par convexité et alors $\left\|x-\frac{1}{2}(y_n+y_m)\right\| \geqslant \delta$ par définition de δ, i.e. $-4\left\|x-\frac{1}{2}(y_n+y_m)\right\|^2 \leqslant -4\delta^2$ et alors l'eq. (5.3) entraîne :

$$\|y_n-y_m\|^2 \leqslant 2(\|x-y_n\|^2 + \|x-y_m\|^2) - 4\delta^2.$$

Comme $\lim_{n\to+\infty}\|x-y_n\|^2 = \lim_{m\to+\infty}\|x-y_m\|^2 = \delta^2$, le côté de droit de l'inégalité précédente tend vers 0 pour n et m suffisamment grands, i.e. $(y_n)_{n\in\mathbb{N}}$ est une suite de Cauchy.

$\boxed{!}$: il nous reste à démontrer qu'il existe seulement un y_0 qui satisfait l'équation (5.1). Soit y_1 un autre élément de S vérifiant $\|x-y_1\| = \delta$, alors, si on écrit encore la formule du parallélogramme, mais cette fois-ci relative à $x-y_0$ et $x-y_1$, on obtient :

$$\|(x-y_0)+(x-y_1)\|^2 + \|(x-y_0)-(x-y_1)\|^2 = 2(\|x-y_0\|^2 + \|x-y_1\|^2)$$

i.e.

$$\|2x-y_0-y_1\|^2 + \|y_1-y_0\|^2 = 4\delta^2,$$

alors :

$$\begin{aligned} 0 \leqslant \|y_1-y_0\|^2 &= 4\delta^2 - \|2x-y_0-y_1\|^2 = 4\delta^2 - \left\|2\left(x-\frac{y_0+y_1}{2}\right)\right\|^2 \\ &= 4\delta^2 - 4\left\|x-\frac{y_0+y_1}{2}\right\|^2 = 4\left(\delta^2 - \left\|x-\frac{y_0+y_1}{2}\right\|^2\right). \end{aligned}$$

On observe que $\frac{y_0+y_1}{2} \in S$ par convexité, et, comme $\delta^2 = \inf_{y\in S} \|x-y\|^2$, sûrement c'est vrai que $\delta^2 \leqslant \left\|x - \frac{y_0+y_1}{2}\right\|^2$ et donc $\delta^2 - \left\|x - \frac{y_0+y_1}{2}\right\|^2 \leqslant 0$, i.e.

$$0 \leqslant \|y_1 - y_0\|^2 = 4\left(\delta^2 - \left\|x - \frac{y_0+y_1}{2}\right\|^2\right) \leqslant 0,$$

mais alors $y_1 = y_0$. □

Comme on l'a vu, l'utilisation de la formule du parallélogramme est indispensable pour effectuer la preuve du théorème. Vu que, par le théorème 4.2.2, seulement les normes hilbertiennes vérifient cette formule, la preuve que l'on vient de voir ne peut pas être reproduite pour les espaces de Banach. En fait, *pour un espace de Banach quelconque, de dimension infinie, le théorème de la projection sur un convexe fermé n'est pas valide.*

Le théorème de projection sur un convexe fermé a des conséquences très importantes que nous détaillerons dans la suite.

Pour le moment, nous remarquons que **ce théorème** nous garantit l'existence et l'unicité de la projection orthogonale y_0, mais il **ne donne aucune information sur comment expliciter les éléments de la suite** $(y_n)_{n\in\mathbb{N}}$ **de** S **convergeant vers** y_0.

Il est donc intéressant d'avoir une caractérisation géométrique de y_0, qui est fournie par le théorème suivant. Dans la section 5.3 on verra une autre application remarquable de ce résultat.

Théorème 5.2.2 *Soit $\mathcal{H}$ un espace de Hilbert réel, S un sous-ensemble propre convexe et fermé de $\mathcal{H}$ et x un élément fixé de $\mathcal{H}$. Alors y_0 est la projection orthogonale de x sur S, i.e. $\|x-y_0\| = \inf_{y\in S}\|x-y\|$, si et seulement si*

$$\forall y \in S, \quad \langle x-y_0, y-y_0\rangle \leqslant 0,$$

i.e. [4], si et seulement si l'angle ϑ entre les vecteurs $x-y_0$ et $y-y_0$ est obtus, comme on le montre dans la figure 5.1.

Si $\mathcal{H}$ est *complexe*, alors il faut remplacer $\langle x-y_0, y-y_0\rangle \leqslant 0$ avec $\Re(\langle x-y_0, y-y_0\rangle) \leqslant 0$.

Preuve. On fera la preuve dans le cas réel, en laissant comme simple exercice de la compléter pour le cas complexe.

$\boxed{\Rightarrow}$: il faut démontrer que, si $\|x-y_0\| = \inf_{y\in S}\|x-y\|$, alors $\langle x-y_0, y-y_0\rangle \leqslant 0 \ \forall y \in \mathcal{H}$. Pour cela, considérons un $y \in S$ quelconque et utilisons la convexité de S pour garantir que $\lambda y + (1-\lambda)y_0 \in S \ \forall \lambda \in [0,1]$. Ainsi, par hypothèse, et grâce aux propriétés de bilinéarité et de symétrie du produit scalaire réel on obtient :

$$\begin{aligned}
\|x-y_0\|^2 &\leqslant \|x-[\lambda y + (1-\lambda)y_0]\|^2 = \|x-y_0-\lambda(y-y_0)\|^2 \\
&= \langle x-y_0-\lambda(y-y_0), x-y_0-\lambda(y-y_0)\rangle \\
&= \langle x-y_0, x-y_0\rangle - \lambda\langle x-y_0, y-y_0\rangle \\
&\quad - \underbrace{\lambda\langle y-y_0, x-y_0\rangle}_{=\lambda\langle x-y_0, y-y_0\rangle} + \lambda^2\langle y-y_0, y-y_0\rangle \\
&= \|x-y_0\|^2 + \lambda^2\|y-y_0\|^2 - 2\lambda\langle x-y_0, y-y_0\rangle.
\end{aligned}$$

4. Vu que $\langle x-y_0, y-y_0\rangle = \|x-y_0\|\|y-y_0\|\cos(\vartheta)$, où ϑ est l'angle entre les vecteurs $x-y_0$ et $y-y_0$.

Ainsi :

$$\|x - y_0\|^2 \leqslant \|x - y_0\|^2 + \lambda^2 \|y - y_0\|^2 - 2\lambda\langle x - y_0, y - y_0\rangle,$$

si on simplifie et on divise par $\lambda \in]0,1[$, on obtient

$$0 \leqslant \lambda \|y - y_0\|^2 - 2\langle x - y_0, y - y_0\rangle,$$

i.e.

$$\langle x - y_0, y - y_0\rangle \leqslant \frac{\lambda}{2} \|y - y_0\|^2,$$

pour tout λ de $]0,1]$. Si maintenant on prend la limite par $\lambda \to 0$ aux deux membres de l'inégalité on obtient : $\lim\limits_{\lambda\to 0}\langle x - y_0, y - y_0\rangle = \langle x - y_0, y - y_0\rangle \leqslant \lim\limits_{\lambda\to 0} \frac{\lambda}{2} \|y - y_0\|^2 = 0$, ce qui conclut la preuve de l'implication directe.

$\boxed{\Leftarrow}$: il faut démontrer que, si $\langle x - y_0, y - y_0\rangle \leqslant 0 \ \forall y \in \mathcal{H}$, alors $\|x - y_0\| = \inf_{y\in S} \|x - y\|$. Pour cela, soit $x \in \mathcal{H}$ fixé et $y_0 \in S$ tel que $\forall y \in S, \langle x - x_0, y - y_0\rangle \leqslant 0$, alors :

$$\begin{aligned}
\|x - y\|^2 &= \|x - y_0 + y_0 - y\|^2 = \|x - y_0 - (y - y_0)\|^2 \\
&= \langle x - y_0 - (y - y_0), x - y_0 - (y - y_0)\rangle \\
&= \langle x - y_0, x - y_0\rangle - \langle x - y_0, y - y_0\rangle - \langle y - y_0, x - y_0\rangle + \langle y - y_0, y - y_0\rangle \\
&= \|x - y_0\|^2 + \|y - y_0\|^2 - 2\langle x - y_0, y - y_0\rangle,
\end{aligned}$$

où, dans la dernière étape, on a utilisé encore la symétrie du produit scalaire réel. On a donc :

$$\|x - y_0\|^2 - \|x - y\|^2 \quad = 2\underbrace{\langle x - y_0, y - y_0\rangle}_{\leqslant 0 \text{ par hypothèse}} - \underbrace{\|y - y_0\|^2}_{\geqslant 0} \leqslant 0,$$

i.e. $\|x - y_0\|^2 \leqslant \|x - y\|^2 \ \forall y \in S$, c'est-à-dire : $\|x - y_0\| = \inf\limits_{y\in S} \|x - y\|$. □

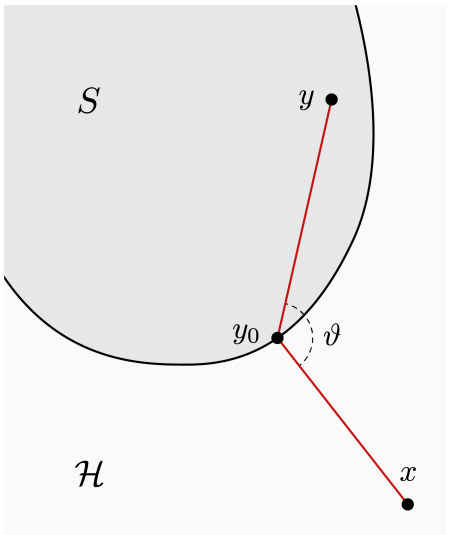


FIGURE 5.1 – Visualisation géométrique bidimensionnelle de la propriété vérifiée par la projection sur un sous-ensemble convexe fermé de $\mathcal{H}$.

Le corollaire suivant montre que tout complémentaire d'un sous-espace vectoriel propre et fermé d'un espace de Hilbert n'est pas trivial et il généralise la propriété 2) du théorème 1.6.1 aux espaces de Hilbert de dimension infinie.

Théorème 5.2.3 *Soit $\mathcal{H}$ un espace de Hilbert et S un sous-espace vectoriel* ***fermé*** *et* ***propre*** *de $\mathcal{H}$, i.e. $S \neq \varnothing$ et $S \neq \mathcal{H}$. Alors $S^\perp$ n'est pas réduit à $\{0_\mathcal{H}\}$, en fait : $\forall x \in \mathcal{H}$, $x \notin S$ fixé, le vecteur $u = x - P_S(x)$ est non nul et il appartient à $S^\perp$:*

$$\forall x \in \mathcal{H}\backslash S, \quad u = x - P_S(x) \neq 0_\mathcal{H} \text{ et } u \in S^\perp.$$

Le vecteur $\boxed{u = x - P_S(x)}$ est appelé « **vecteur résidu** » et le fait que $u \perp S$ justifie pleinement le nom de « projection orthogonale » pour $P_S(x)$.

Preuve. Un sous-espace vectoriel est convexe, donc il vaut le théorème de projection sur un convexe fermé et alors $\exists\ P_S(x) \in S$, tel que $\|u\| = \|x - P_S(x)\| = \inf_{y \in S} \|x - y\| \equiv \delta$.

Comme $x \notin S$ et $P_S(x) \in S$, $u \neq 0_\mathcal{H}$. Si on démontre que $u \in S^\perp$, i.e. que $\langle u, s\rangle = 0\ \forall s \in S$, le théorème sera démontré.

Pour arriver à ce résultat on commence par observer que : $\forall k \in \mathbb{K}, \forall s \in S$,

$$\|u + ks\|^2 = \|x - P_S(x) + ks\|^2 = \|x - \underbrace{(P_S(x) - ks)}_{\in S}\|^2 \geqslant \delta^2 = \|u\|^2$$

où on a utilisé la définition de δ et le fait que S est un sous-espace vectoriel.

Ainsi :

$$\|u + ks\|^2 - \|u\|^2 \geqslant 0, \qquad \forall k, \in \mathbb{K}, \forall s \in S,$$

mais :

$$\begin{aligned} \|u + ks\|^2 &= \langle u + ks, u + ks\rangle \\ &= \langle u, u\rangle + \langle u, ks\rangle + \langle ks, u\rangle + \langle ks, ks\rangle \\ &= \|u\|^2 + \bar{k}\langle u, s\rangle + k\langle s, u\rangle + |k|^2 \|s\|^2, \end{aligned}$$

on a donc que $\|u + ks\|^2 - \|u\|^2 \geqslant 0$ si et seulement si :

$$\bar{k}\langle u, s\rangle + k\langle s, u\rangle + |k|^2 \|s\|^2 \geqslant 0.$$

k étant arbitraire, on peut prendre $k = \langle u, s\rangle t$ avec $t \in \mathbb{R}$ *quelconque*. Ainsi l'inéquation ci-dessus devient :

$$\begin{aligned} & \overline{\langle u, s\rangle} t\langle u, s\rangle + \langle u, s\rangle t \underbrace{\langle s, u\rangle}_{=\overline{\langle u, s\rangle}} + |\langle u, s\rangle|^2 t^2 \|s\|^2 \geqslant 0 \\ \iff\ & |\langle u, s\rangle|^2 t + |\langle u, s\rangle|^2 t + |\langle u, s\rangle|^2 t^2 \|s\|^2 \geqslant 0 \\ \iff\ & t^2 \left(|\langle u, s\rangle|^2 \|s\|^2\right) + 2t|\langle u, s\rangle|^2 \geqslant 0 \quad \forall t \in \mathbb{R}. \end{aligned}$$

Il ne serait pas intelligent de simplifier $|\langle u, s\rangle|^2$, car c'est bien $\langle u, s\rangle$ qu'on veut calculer... Par contre, la bonne stratégie pour terminer la preuve consiste à reconnaître en $t^2 \left(|\langle u, s\rangle|^2 \|s\|^2\right) + 2t|\langle u, s\rangle|^2$ une fonction polynomiale de degré 2 de la forme $P(t) = at^2 + bt + c$ avec :

$$\begin{cases} a = |\langle u, s\rangle|^2 \|s\|^2 \\ b = 2|\langle u, s\rangle|^2 \\ c = 0, \end{cases}$$

son discriminant vaut : $\Delta = b^2 - 4ac = 4|\langle u, s\rangle|^4 \geqslant 0$. Donc, pour que $P(t)$ soit positive $\underline{\forall t \in \mathbb{R}}$, il faut que $\Delta \leqslant 0$, ce qui n'est possible que si $\Delta = 0$, mais

$$\Delta = 0 \iff 4|\langle u, s\rangle|^4 = 0 \ \forall s \in S \ \ (u \text{ étant fixé car } x \text{ l'est}),$$

i.e. $\langle u, s\rangle = 0 \ \forall s \in S$, d'où $u \in S^\perp$. □

5.2.1 Une caractérisation des sous-espaces vectoriels fermés dans les espaces de Hilbert

Nous allons utiliser le théorème 5.2.3 pour déduire une caractérisation très utile des sous-espaces vectoriels fermés dans les espaces de Hilbert.

On commence avec un résultat intermédiaire.

Lemme 5.2.1 *Soit $\mathcal{H}$ un espace de Hilbert, et M un sous-ensemble quelconque de $\mathcal{H}$, alors*

$$\overline{\text{span}(M)} = \overline{\text{span}(M)}^{\perp\perp}$$

Preuve. Si l'on pose

$$\begin{cases} S = \overline{\text{span}(M)} \\ T = \overline{\text{span}(M)}^{\perp\perp}, \end{cases}$$

alors le théorème 5.1.1 nous garantit que $S \subseteq T$, si on prouve que $S \subset T$ est une condition impossible, ça restera seulement $S = T$.

Pour cela, nous observons d'abord que S est un sous-espace vectoriel fermé de T et que T, en tant que complément orthogonal d'un sous-ensemble de $\mathcal{H}$, il est un sous-espace vectoriel fermé de $\mathcal{H}$ et donc il est lui-même un espace de Hilbert.

Si, par l'absurde, $S \subset T$, alors on peut appliquer le théorème 5.2.3 au couple S et T, pour assurer qu'il existe $u \in T$, $u \neq 0_\mathcal{H}$ et $u \in S^\perp$. Mais ceci implique que $u \in S^\perp \cap T = \overline{\text{span}(M)}^\perp \cap \overline{\text{span}(M)}^{\perp\perp} = \{0_\mathcal{H}\}$, ce qui est en contradiction avec le fait que $u \neq 0_\mathcal{H}$. Finalement, on a bien : $\overline{\text{span}(M)} = \overline{\text{span}(M)}^{\perp\perp}$. □

On a maintenant toutes les informations pour prouver une caractérisation très utile des sous-espaces vectoriels fermés dans les espaces de Hilbert : **les sous-espaces vectoriels fermés sont précisément ceux qui coïncident avec leur complément bi-orthogonal**.

On remarque la puissance de cette caractérisation : elle permet d'analyser une propriété topologique (la fermeture, qui fait référence à la convergence des suites), qui peut être très compliquée à examiner directement, via l'analyse de l'orthogonalité, et donc du produit scalaire.

Théorème 5.2.4 *Soit $\mathcal{H}$ un espace de Hilbert, et M un **sous-espace vectoriel** de $\mathcal{H}$.*

1. $M^{\perp\perp} = \overline{\text{span}(M)}$
2. $\boxed{\overline{M} = M^{\perp\perp}}$
3. *M est un sous-espace vectoriel fermé de $\mathcal{H}$ si et seulement si $M = M^{\perp\perp}$.*

Preuve.

1. Comme M est un sous-espace vectoriel, $M \equiv \text{span}(M)$. De plus, grâce à la propriété **8.** du théorème 5.1.1 : $M^\perp = \text{span}(M)^\perp = \overline{\text{span}(M)}^\perp$. Alors, le lemme précédent implique : $M^{\perp\perp} = \overline{\text{span}(M)}^{\perp\perp} = \overline{\text{span}(M)}$.

2. On vient de démontrer que $M^{\perp\perp} = \overline{\text{span}(M)}$ et on sait que $M \equiv \text{span}(M)$, donc $M = M^{\perp\perp}$.

3. On va montrer la double implication :

$\boxed{\Rightarrow}$: on sait du point **1.** que $\overline{M} = M^{\perp\perp}$, mais si M est fermé alors $M = \overline{M}$, et donc $M = M^{\perp\perp}$.

$\boxed{\Leftarrow}$: si $M = M^{\perp\perp}$, alors M est automatiquement un sous-espace vectoriel fermé en tant que complément orthogonal de $M^\perp$. □

Corollaire 5.2.1 *Soient $\mathcal{H}$ un espace de Hilbert et M, N deux parties quelconques de $\mathcal{H}$.*

1. *Ça vaut que :*
$$\boxed{(M \cup N)^\perp = M^\perp \cap N^\perp}\,.$$
2. *Si, de plus, M et N sont deux* ***sous-espaces vectoriels fermés*** *de $\mathcal{H}$, alors :*
$$\boxed{(M \cap N)^\perp = \overline{\text{span}\,(M^\perp \cup N^\perp)}}\,. \tag{5.4}$$

Preuve.

1. Allons prouver les deux relations d'inclusion :

   $(M \cup N)^\perp \subseteq M^\perp \cap N^\perp$: soit $x \in (M \cup N)^\perp$ et $y \in M$, alors y appartient aussi à $M \cup N$, donc $\langle x, y\rangle = 0$, i.e. $x \in M^\perp$. Si maintenant on prend $y \in N$, le même argument montre que $x \in N^\perp$. Mais alors $x \in M^\perp$ et $x \in N^\perp$, i.e. $x \in M^\perp \cap N^\perp$.

   $M^\perp \cap N^\perp \subseteq (M \cup N)^\perp$: soit $x \in M^\perp \cap N^\perp$, alors $x \in M^\perp$ et $x \in N^\perp$. Si $y \in M \cup N$, alors $y \in M$ ou $y \in N$, mais dans les deux cas $\langle x, y\rangle = 0$, i.e. $x \in (M \cup N)^\perp$.

2. La relation que l'on vient de déterminer vaut pour toutes parties de $\mathcal{H}$ et donc aussi pour $M^\perp$ et $N^\perp$, dans ce cas la 1. devient : $\left(M^\perp \cup N^\perp\right)^\perp \subseteq M^{\perp\perp} \cap N^{\perp\perp} \underset{\text{th. (5.2.4)}}{=} \overline{\text{span}(M)} \cap \overline{\text{span}(N)} = M \cap N$ car M et N sont supposés être sous-espaces vectoriels fermés ! Si maintenant on passe à l'orthogonal on obtient :
$$(M \cap N)^\perp = \left(M^\perp \cup N^\perp\right)^{\perp\perp} \underset{\text{th. (5.2.4)}}{=} \overline{\text{span}(M^\perp \cup N^\perp)}$$

□

5.3 Sous-ensemble polaire et bipolaire d'un espace de Hilbert

Dans cette section nous allons retrouver le résultat sur la caractérisation d'une partie fermée d'une espace de Hilbert d'une manière alternative qui passe par l'analyse d'un concept, celui de l'ensemble polaire, qui a une grande importance dans la théorie de l'optimisation convexe.

Déf. 5.3.1 (Polaire et bipolaire) *Soit $\mathcal{H}$ un espace de Hilbert et M une partie non vide quelconque de $\mathcal{H}$. L'ensemble polaire de M, noté avec M° est le sous-ensemble de $\mathcal{H}$ défini par*[5] *:*

$$M^\circ := \{x \in \mathcal{H} \ : \ \forall y \in M, \ \Re(\langle x, y\rangle) \leqslant 1\} \equiv \{x \in \mathcal{H} \ : \ \sup_{y \in M} \ \Re(\langle x, y\rangle) \leqslant 1\}.$$

Le bipolaire de M est le polaire du polaire, i.e.

$$M^{\circ\circ} := (M^\circ)^\circ = \{h \in \mathcal{H} \ : \ \forall x \in M^\circ, \ \Re(\langle h, x\rangle) \leqslant 1\} \equiv \{h \in \mathcal{H} \ : \ \sup_{x \in M^\circ} \ \Re(\langle h, x\rangle) \leqslant 1\}.$$

Le résultat contient des propriétés remarquables du polaire et du bipolaire. Avant de l'énoncer, nous rappelons la définition suivante.

Déf. 5.3.2 (Enveloppe convexe fermé) *L'enveloppe convexe fermé d'une partie M de $\mathcal{H}$ est l'adhérence de l'intersection de toutes les parties convexes de $\mathcal{H}$ qui contiennent M. Il est le plus petit sous-ensemble convexe fermé de $\mathcal{H}$ qui contient M.*

Théorème 5.3.1 *Soit $\mathcal{H}$ un espace de Hilbert et M une partie non vide quelconque de $\mathcal{H}$.*

1. *M° est un sous-ensemble convexe fermé de $\mathcal{H}$ qui contient le vecteur nul $0_\mathcal{H}$.*
2. *$M^{\circ\circ}$ coïncide avec l'enveloppe convexe fermé $\mathcal{C}$ de $M \cup \{0_\mathcal{H}\}$.*
3. *Si M est une partie convexe de $\mathcal{H}$ qui contient $0_\mathcal{H}$, alors $\overline{M} = M^{\circ\circ}$.*
4. *Si M est un sous-espace vectoriel de $\mathcal{H}$, alors $M^\circ = M^\perp$.*

Preuve.

1. Le fait que M° contient $0_\mathcal{H}$ est évident car $\langle 0_\mathcal{H}, y\rangle = 0 < 1$ pour tout $y \in M$. Pour vérifier la convexité, considérons $\lambda \in [0, 1]$ et $x_1, x_2 \in M^\circ$, alors, grâce à la linéarité à gauche du produit scalaire :

$$\Re(\langle x_1 + (1-\lambda)x_2, y\rangle) = \lambda\Re(\langle x_1, y\rangle) + (1-\lambda)\Re(\langle x_2, y\rangle) \leqslant \lambda + (1-\lambda) = 1,$$

ce qui montre que M° est convexe. Ça reste seulement à démontrer la fermeture, pour cela nous observons d'abord que, pour tout $y \in \mathcal{H}$ fixé, l'application $\phi : \mathcal{H} \to \mathbb{R}$, $x \mapsto \phi_y(x) := \Re(\langle x, y\rangle)$ est continue. Si l'on écrit :

$$\mathcal{H}_y := \phi_y^{-1}\{]-\infty, +1]\} = \{x \in \mathcal{H} \ : \ \Re(\langle x, y\rangle) \leqslant 1\},$$

5. Évidemment, la partie réelle $\Re$ du produit scalaire peut être éliminée si $\mathcal{H}$ est un espace de Hilbert réel.

alors $\mathcal{H}_y$ est un sous-ensemble fermé de $\mathcal{H}$ en tant qu'image réciproque d'un sous-ensemble fermé de $\mathbb{R}$ (rappeler que $]-\infty,+1]$ est fermé vu que son ensemble complémentaire en $\mathbb{R}$ est $]1,+\infty[$, qui est ouvert). Par définition, les éléments de M° doivent appartenir à $\mathcal{H}_y$ pour tout $y \in M$, i.e. $M^\circ = \bigcap_{y\in M} \mathcal{H}_y$ et donc il est fermé en $\mathcal{H}$ car il est intersection de parties fermées de $\mathcal{H}$.

2. Montrons les inclusions opposées.

$\boxed{\mathcal{C} \subseteq M^{\circ\circ}}$: nous allons d'abord montrer que $M \subseteq M^{\circ\circ}$. Pour faire ça, remarquons tout d'abord que $M^{\circ\circ} = \bigcap_{x\in M^\circ} \mathcal{H}_x$. Puis, supposons $y \in M$ et $x \in M^\circ$ arbitraires mais fixés, alors, en particulier, $x \in \mathcal{H}_y$, i.e. $\Re(\langle x, y\rangle) \leqslant 1$ mais, comme $\Re(\langle x, y\rangle) = \Re(\langle y, x\rangle)$, ça vaut aussi $\Re(\langle y, x\rangle) \leqslant 1$, i.e. $y \in \mathcal{H}_x$. Grâce à l'arbitrarité de $x \in M^\circ$, $y \in \bigcap_{x\in M^\circ} \mathcal{H}_x = M^{\circ\circ}$. Mais alors : $y \in M \implies y \in M^{\circ\circ}$, i.e. $M \subseteq M^{\circ\circ}$.

Grâce à 1., $M^{\circ\circ}$, en tant qu'ensemble polaire, est convexe, fermé et il contient $\{0_\mathcal{H}\}$, mais, comme on vient de le voir $M \subseteq M^{\circ\circ}$, et alors $M^{\circ\circ}$ est un ensemble convexe, fermé qui il contient $M \cup \{0_\mathcal{H}\}$. Vu que $\mathcal{C}$, l'enveloppe convexe fermée de $M \cup \{0_\mathcal{H}\}$, est le plus petit sous-ensemble convexe de $\mathcal{H}$ qui contient $M \cup \{0_\mathcal{H}\}$, surement il est contenu en $M^{\circ\circ}$.

$\boxed{M^{\circ\circ} \subseteq \mathcal{C}}$: le fait d'avoir souligné que $0_\mathcal{H} \in \mathcal{C}$ est utile dans cette étape de la preuve. En fait, grâce au théorème 5.2.2, pour tout $x \in \mathcal{H}$ ça vaut

$$\Re(\langle x - P_\mathcal{C}x, 0_\mathcal{H} - P_\mathcal{C}x\rangle) \leqslant 0 \iff \Re(\langle x - P_\mathcal{C}x, -P_\mathcal{C}x\rangle) \leqslant 0 \iff \Re(\langle x - P_\mathcal{C}x, P_\mathcal{C}x\rangle) \geqslant 0,$$

et, pour tout $y \in M$, ça vaut aussi $\Re(\langle x - P_\mathcal{C}x, y - P_\mathcal{C}x\rangle) \leqslant 0$, i.e. $\Re(\langle x - P_\mathcal{C}x, y - P_\mathcal{C}x\rangle) \leqslant \varepsilon$ pour tout $\varepsilon > 0$, i.e., par linéarité du produit scalaire,

$$\Re(\langle x - P_\mathcal{C}x, y - P_\mathcal{C}x\rangle) = \Re(\langle x - P_\mathcal{C}x, y\rangle) - \Re(\langle x - P_\mathcal{C}x, P_\mathcal{C}x\rangle) \leqslant \varepsilon,$$

c'est-à-dire, vu que $\varepsilon + \Re(\langle x - P_\mathcal{C}x, P_\mathcal{C}x\rangle)$ est un nombre réel > 0,

$$\Re(\langle x - P_\mathcal{C}x, y\rangle) \leqslant \varepsilon + \Re(\langle x - P_\mathcal{C}x, P_\mathcal{C}x\rangle) \iff \frac{\Re(\langle x - P_\mathcal{C}x, y\rangle)}{\varepsilon + \Re(\langle x - P_\mathcal{C}x, P_\mathcal{C}x\rangle)} \leqslant 1,$$

qui peut être réécrit comme ça :

$$\Re\left(\left\langle \frac{x - P_\mathcal{C}x}{\varepsilon + \Re(\langle x - P_\mathcal{C}x, P_\mathcal{C}x\rangle)}, y \right\rangle\right) \leqslant 1 \qquad \forall y \in M,\ \forall \varepsilon > 0,$$

i.e. l'élément $z(x) := \frac{x - P_\mathcal{C}x}{\varepsilon + \Re(\langle x - P_\mathcal{C}x, P_\mathcal{C}x\rangle)} \in M^\circ$ pour tout $x \in \mathcal{H}$.

Vu que ce résultat vaut pour n'importe quel $x \in \mathcal{H}$, nous pouvons l'appliquer quand $\underline{x \in M^{\circ\circ}}$, dans ce cas, par définition, nous avons $\Re(\langle x, z(x)\rangle) \leqslant 1$, i.e.

$$\Re(\langle x, z(x)\rangle) = \Re\left(\left\langle x, \frac{x - P_\mathcal{C}x}{\varepsilon + \Re(\langle x - P_\mathcal{C}x, P_\mathcal{C}x\rangle)} \right\rangle\right) \leqslant 1$$

d'où

$$\Re(\langle x, x - P_\mathcal{C}x\rangle) \leqslant \varepsilon + \Re(\langle x - P_\mathcal{C}x, P_\mathcal{C}x\rangle) = \varepsilon + \Re(\langle P_\mathcal{C}x, x - P_\mathcal{C}x\rangle)$$

ce qui donne

$$\Re(\langle x - P_\mathcal{C}x, x - P_\mathcal{C}x\rangle) \leqslant \varepsilon \iff \|x - P_\mathcal{C}x\|^2 \leqslant \varepsilon \quad \forall \varepsilon > 0,$$

mais ça est possible si et seulement si $x - P_{\mathcal{C}}x = 0_{\mathcal{H}}$, i.e. $x = P_{\mathcal{C}}x$, mais comme $P_{\mathcal{C}}x \in \mathcal{C}$, $\underline{x \in \mathcal{C}}$. On a alors prouvé que $x \in M^{\circ\circ} \implies x \in \mathcal{C}$, i.e. $M^{\circ\circ} \subseteq \mathcal{C}$.

3. D'après 2., si M est une partie convexe de $\mathcal{H}$ qui contient $0_{\mathcal{H}}$, alors $M^{\circ\circ}$ est le plus petit convexe fermé qui contient M. Or, si M est convexe, aussi $\overline{M}$ est convexe et, de plus, il est le plus petit fermé qui contient M. Par conséquent, $\overline{M} = M^{\circ\circ}$.

4. Montrons les inclusions opposées.

$\boxed{M^{\perp} \subseteq M^{\circ}}$: évident car si $\underline{x \in M^{\perp}}$, alors, pour tout $y \in M$, $\langle x, y\rangle = 0 < 1$, donc $\underline{x \in M^{\circ}}$.

$\boxed{M^{\circ} \subseteq M^{\perp}}$: soit $\underline{x \in M^{\circ}}$, on doit démontrer que $x \in M^{\perp}$, i.e. que $\langle x, y\rangle = 0 \ \forall y \in M$. Pour faire ça, on utilise le fait que M est supposé être un sous-espace vectoriel de $\mathcal{H}$: si $y \in M$, alors aussi $ty \in M \ \forall t \in \mathbb{R}\backslash\{0\}$. Vu que $x \in M^{\circ}$ et $ty \in M$, ça doit valoir que

$$\Re(\langle x, ty\rangle) \leqslant 1 \iff t\Re(\langle x, y\rangle) \leqslant 1 \ \forall t \in \mathbb{R}\backslash\{0\} \iff \Re(\langle x, y\rangle) = 0 \quad \forall y \in M,$$

si $\mathcal{H}$ est un espace de Hilbert réel on a terminé la preuve. Si $\mathcal{H}$ est complexe, nous devons démontrer que aussi la partie imaginaire du produit scalaire est nulle. Pour cela, utilisons le théorème 1.1.2 qui garantit que $\Im(\langle x, y\rangle) = \Re(\langle x, iy\rangle)$, donc $\Im(\langle x, y\rangle) = \Re(\langle x, iy\rangle) = 0$ car nous avons démontré avant que $\Re(\langle x, z\rangle) = 0 \ \forall z \in M$ et $z = iy \in M$ quand $y \in M$ si M est un sous-espace vectoriel complexe. En conclusion, $\langle x, y\rangle = 0 \ \forall y \in M$ et alors $\underline{x \in M^{\perp}}$. □

Les propriétés 3. et 4. du théorème précédent impliquent la propriété 2. du théorème 5.2.4, i.e. $\overline{M} = M^{\perp\perp}$, en fait, d'un côté, $M^{\circ} = M^{\perp}$, donc, si on répète deux fois l'opération du polaire on obtient $M^{\circ\circ} = M^{\perp\perp}$, mais $M^{\circ\circ} = \overline{M}$, donc $\overline{M} = M^{\perp\perp}$.

5.4 Le théorème de la projection (orthogonale) dans un espace de Hilbert

Allons maintenant à énoncer et démontrer le plus important corollaire du théorème de la projection orthogonale sur un convexe fermé.

Théorème 5.4.1 (Théorème de la projection orthogonale) *Soit $\mathcal{H}$ un espace de Hilbert sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Soit S un sous-espace vectoriel propre et **fermé** de $\mathcal{H}$. Alors :*

$$\mathcal{H} = S \oplus S^{\perp},$$

i.e. $\forall x \in \mathcal{H}$, $\exists s \in S$, $\exists t \in S^{\perp}$: $x = s + t$, et cette décomposition est unique, i.e. si :

$$\begin{cases} x = s + t \\ x = s' + t' \end{cases} \quad \text{avec } s, s' \in S,\ t, t' \in S^{\perp}, \quad \text{alors : } \begin{cases} s = s' \\ t = t'. \end{cases}$$

Si S n'est pas un sous-espace propre, alors on a la décomposition triviale $\mathcal{H} = \mathcal{H} \oplus \{0_{\mathcal{H}}\}$.

Preuve. Soient $x \in \mathcal{H}$, fixé, $P_S(x) \in S$ le projeté orthogonal de x sur S et u le vecteur résidu : $u = x - P_S(x) \in S^\perp$. Grâce au théorème précédent, on peut décomposer x ainsi :

$$x = \underbrace{P_S(x)}_{\in S} + \underbrace{x - P_S(x)}_{\in S^\perp}.$$

Il reste maintenant à prouver qu'une décomposition comme celle ci-dessus est unique. Considérons les décompositions : $x = s + t$ et $x = s' + t'$, avec $s, s' \in S, t, t' \in S^\perp$, alors $s + t = s' + t'$, i.e.

$$\underbrace{s - s'}_{\in S} = \underbrace{t' - t}_{\in S^\perp},$$

S et $S^\perp$ étant des espaces vectoriels, ils sont stables par soustraction, d'où le résultat entre accolades. Or, on a : $S \ni s - s' = t' - t \in S^\perp$, mais alors $s - s' \in S \cap S^\perp$ et $t' - t \in S \cap S^\perp$, mais $S \cap S^\perp = \{0_\mathcal{H}\}$ et ainsi, nécessairement, $s' = s$ et $t = t$. □

Observation importante : à la lumière de ce théorème, l'écriture $x = x - y + y$, qui peut sembler triviale, assume une signification bien plus profonde quand on peut prouver que $y = P_S x$, S sous-espace fermé de $\mathcal{H}$. En fait, dans ce cas, comme conséquence immédiate du théorème de la projection et de Pythagore, on obtient la formule $\|x\|^2 = \|x - y\|^2 + \|y\|^2$, qui est extrêmement utile pour la théorie et pour les applications des espaces de Hilbert.

Voyons un exercice dans lequel on applique les résultats qu'on vient d'apprendre.

Exercice 5.1

Soit Ω un sous-ensemble borné de $\mathbb{R}^n$ et considérons l'ensemble M des fonctions $f : \Omega \to \mathbb{R}$, $f \in L^2(\Omega)$, constantes p.p. Démontrer que :

1. M est un sous-espace vectoriel fermé de $L^2(\Omega)$.
2. $\forall f \in L^2(\Omega)$, la projection de f sur M est la fonction constante p.p. et égale à la moyenne de f sur Ω, i.e. ça vaut que : $P_M f = \frac{1}{|\Omega|} \int_\Omega f(x)dx$, $|\Omega| = m(\Omega)$, m : mesure de Lebesgue sur $\mathbb{R}^n$.
3. Le complément orthogonale de M en $L^2(\Omega)$ est donné par les fonctions h de $L^2(\Omega)$ à moyenne nulle, i.e. $\int_\Omega h(x)dx = 0$.

Solution de l'exercice 5.1

1. On peut caractériser M comme le sous-espace vectoriel de $L^2(\Omega)$ engendré par la fonction constante $\mathbf{1}(x) = 1$ p.p. sur Ω. Comme il y a un seul générateur, M est isomorphe à $\mathbb{R}$, qui est fermé.
2. Soit $f \in L^2(\Omega)$, alors, grâce au théorème de la projection, si on écrit $f = f - P_M f + P_M f$, nous avons que $f - P_M f \in M^\perp$. Soit alors g un élément de M tel que : $g(x) = c \neq 0$

p.p. sur Ω, calculons le produit scalaire entre $f - P_M f$ et g

$$\begin{aligned}
0 &= \langle f - P_M f, g\rangle_{L^2(\Omega)} \\
&= \int_\Omega (f(x) - P_M f) g(x) dx \quad (P_M f \in M \implies \text{const. p.p., donc on interprète } P_M f \in \mathbb{R}) \\
&= \int_\Omega f(x)g(x)dx - \int_\Omega (P_M f) g(x) dx = c \int_\Omega f(x)dx - c(P_M f) \int_\Omega dx \\
&= c\left(\int_\Omega f(x)dx - (P_M f)|\Omega| \right),
\end{aligned}$$

i.e., vu que $c \neq 0$,

$$P_M f = \frac{1}{|\Omega|} \int_\Omega f(x)dx.$$

3. Soit $h \in M^\perp$ quelconque, alors, par définition : $\langle h, g\rangle_{L^2(\Omega)} = 0 \ \forall g \in M$. Prenons encore $g(x) = c \neq 0$ p.p., alors

$$0 = \langle h, g\rangle_{L^2(\Omega)} = \int_\Omega h(x)g(x)dx = c\int_\Omega h(x)dx, \qquad \forall k \in \mathbb{R},$$

donc $\int_\Omega h(x)dx = 0$.

Grâce à ce que l'on vient de prouver et au théorème de la projection orthogonale, nous pouvons représenter *d'une manière unique* toute fonction $f \in L^2(\Omega)$, $\Omega \subset \mathbb{R}^n$ tel que $m(\Omega) < +\infty$, comme ça :

$$f = \langle f\rangle_\Omega + h,$$

où $\langle f\rangle_\Omega$ est la fonction constante sur Ω et égale à la moyenne de f sur Ω et $h \in L^2(\Omega)$ telle que $\langle h\rangle_\Omega = 0$. □

5.5 Systèmes orthonormés et bases hilbertiennes

Nous avons vu, dans les chapitres 1 et 2, que la présence d'une base orthonormale dans un espace Euclidien permet de calculer facilement les composantes des vecteurs et de caractériser les opérateurs de projection orthogonale. De plus, avec le choix de la base de Fourier, on a pu définir les coefficients de Fourier et la transformée de Fourier discrète.

Nous allons voir ici quelles sont les conditions qu'il faut rajouter pour étendre tout ça aux espaces de Hilbert de dimension infinie. On commence par une définition.

Déf. 5.5.1 (s.o.n) *Une famille orthonormale d'éléments d'un espace de Hilbert est dite un système orthonormale (ou orthonormée) et écrite s.o.n.*

On va analyser les s.o.n. dans les espaces de Hilbert dits séparables, qu'on va définir ci-dessous.

Déf. 5.5.2 *Un espace de Hilbert $\mathcal{H}$ est dit* ***séparable*** *s'il existe un sous-ensemble $E \subseteq \mathcal{H}$* ***dénombrable*** *et* ***dense*** *en $\mathcal{H}$:* $\mathrm{card}(E) = \aleph_0$, $\overline{E} = \mathcal{H}$.

Les espaces de Hilbert séparables sont ceux qui apparaissent dans la grande majorité de cas, nous donnerons un exemple d'espace de Hilbert non-séparable dans la section 5.5.3. L'avantage principal de travailler avec les espaces de Hilbert séparables est contenu dans le théorème suivant.

Théorème 5.5.1 *Tous les s.o.n. d'un espace de Hilbert $\mathcal{H}$ séparable et de dimension infinie sont dénombrables.*

Preuve. Soit M un s.o.n infini de $\mathcal{H}$. Comme $\mathcal{H}$ est séparable, il existe un sous-ensemble $E \subseteq \mathcal{H}$ dénombrable et dense en $\mathcal{H}$: $\overline{E} = \mathcal{H}$.

Grâce à la caractérisation *2.* de la densité donnée dans la définition 4.1.4, nous pouvons garantir que, fixé un élément $x \in M$ quelconque et fixé arbitrairement $\varepsilon > 0$, $\exists u_x \in E$ tel que $\|x - u_x\| < \varepsilon$. La preuve sera complète si on démontre que la correspondance définie par la fonction

$$\begin{array}{rccc} \imath : & M & \longrightarrow & E \\ & x & \longmapsto & \imath(x) = u_x \end{array}$$

est injective, en fait, si c'est le cas, M est en correspondance bijective avec $\imath(M) \subseteq E$ qui est une partie infinie d'un ensemble dénombrable, et donc dénombrable elle-même.

Soit alors $y \in M$ quelconque mais tel que $y \neq x$, et $u_y \in E$ tel que $\|x - u_y\| < \varepsilon$ pour tout $\varepsilon > 0$ arbitraire mais fixé. Vu que $x \neq y$ sont deux points distincts arbitrairement choisis dans M, l'injectivité de $\imath$ correspond au fait que $\imath(x) \neq \imath(y)$, i.e. $u_x \neq u_y$. Pour démontrer ça, observons que, comme x et y appartiennent à un s.o.n, leur distance vaut $\sqrt{2}$ et alors nous pouvons écrire

$$\begin{aligned} \sqrt{2} = \|x - y\| &= \|x - u_x + u_y - y + u_x - u_y\| \\ &\underset{\text{inég. triang.}}{\leqslant} \|x - u_x\| + \|y - u_y\| + \|u_x - u_y\| \\ &< 2\varepsilon + \|u_x - u_y\|, \end{aligned}$$

i.e. $\|u_x - u_y\| > \sqrt{2} - 2\varepsilon$. $\sqrt{2} - 2\varepsilon > 0 \iff \varepsilon < \sqrt{2}/2$, donc, il suffit de fixer $\varepsilon \in]0, \sqrt{2}/2[$, $\|u_x - u_y\| > 0$ pour avoir $u_x \neq u_y$. □

Ce théorème explique pourquoi, dans la suite, on utilisera une valeur discrète $n \in \mathbb{N}$ ou $\mathbb{Z}$ pour indexer les éléments d'un s.o.n. dans un espace de Hilbert séparable.

Convention : dorénavant, un espace de Hilbert $\mathcal{H}$ sera *toujours implicitement considéré séparable*, sauf si autrement spécifié.

Les deux plus importantes propositions relatives aux s.o.n sont l'inégalité de Bessel et le théorème de Fischer-Riesz.

5.5.1 L'inégalité de Bessel et les coefficients de Fourier

L'expansion d'un vecteur $v \in \mathbb{K}^n$, $n < +\infty$, par rapport à une base orthonormale $(u_i)_{i=1}^n$ s'écrit $v = \sum_{i=1}^{n} \langle v, u_n \rangle u_n$, où $\langle v, u_n \rangle$ sont les composantes de v dans cette base. De plus, ça vaut l'identité de Plancherel : $\|v\|^2 = \sum_{i=1}^{n} |\langle v, u_n \rangle|^2$. Si on veut qu'un s.o.n d'un espace de Hilbert

$\mathcal{H}$ de dimension infinie soit l'extension d'une base orthonormale, alors il faut que, pour tout élément $x \in \mathcal{H}$, la suite $(\langle x, u_n\rangle)_{n\in\mathbb{N}}$ soit décroissante vers 0 quand $n \to +\infty$, sinon la série $\sum_{n\in\mathbb{N}} \langle x, u_n\rangle u_n$ ne pourra pas être convergente ! Le résultat suivant nous assure que ceci est toujours vrai, ce qui n'est pas toujours garanti est la validité de l'identité de Plancherel...

Théorème 5.5.2 (Inégalité de Bessel) *Soit $(u_n)_{n\in\mathbb{N}} \subset \mathcal{H}$ un s.o.n dans un espace de Hilbert $\mathcal{H}$. Alors, $\forall x \in \mathcal{H}$ ça vaut :*

$$\sum_{n\in\mathbb{N}} |\langle x, u_n\rangle|^2 \leqslant \|x\|^2, \tag{5.5}$$

plus précisément, l'écart entre les deux membres de l'inégalité peut être quantifié comme ça :

$$\|x\|^2 - \sum_{n\in\mathbb{N}} |\langle x, u_n\rangle|^2 = \left\| x - \sum_{n\in\mathbb{N}} \langle x, u_n\rangle u_n \right\|^2. \tag{5.6}$$

Preuve. Si on prouve que l'écart $\|x\|^2 - \sum_{n\in\mathbb{N}} |\langle x, u_n\rangle|^2$ est égal à une norme au carré, qui est $\geqslant 0$, l'inégalité de Bessel sera automatiquement démontrée.

Pour simplifier la notation, écrivons $\lambda_n = \langle x, u_n\rangle \iff \overline{\lambda_n} = \langle u_n, x\rangle \ \forall n \in \mathbb{N}$ et considérons $N \in \mathbb{N}$ quelconque. Par calcul direct, nous avons :

$$\left\| x - \sum_{n=0}^{N} \lambda_n u_n \right\|^2 = \|x\|^2 - \langle x, \sum_{n=0}^{N} \lambda_n u_n\rangle - \langle \sum_{n=0}^{N} \lambda_n u_n, x\rangle + \left\| \sum_{n=0}^{N} \lambda_n u_n \right\|^2.$$

En appliquant la sesquilinéarité aux deux termes intermédiaires, et le théorème de Pythagore généralisé au dernier terme, nous pouvons réécrire l'égalité ci-avant comme ça :

$$\left\| x - \sum_{n=0}^{N} \lambda_n u_n \right\|^2 = \|x\|^2 - \sum_{n=0}^{N} \overline{\lambda_n}\langle x, u_n\rangle - \sum_{n=0}^{N} \lambda_n \langle u_n, x\rangle + \sum_{n=0}^{N} |\lambda_n|^2 \|u_n\|^2,$$

mais, grâce aux définitions de λ_n et $\overline{\lambda_n}$ et au fait que $\|u_n\|^2 = 1$ pour tout n, la dernière égalité devient :

$$\left\| x - \sum_{n=0}^{N} \lambda_n u_n \right\|^2 = \|x\|^2 - \sum_{n=0}^{N} \overline{\lambda_n}\lambda_n - \sum_{n=0}^{N} \lambda_n\overline{\lambda_n} + \sum_{n=0}^{N} |\lambda_n|^2 = \|x\|^2 - \sum_{n=0}^{N} |\lambda_n|^2,$$

i.e.

$$\|x\|^2 - \sum_{n=0}^{N} |\langle x, u_n\rangle|^2 = \left\| x - \sum_{n=0}^{N} \langle x, u_n\rangle u_n \right\|^2.$$

Comme on n'a pas posé des restrictions sur $N \in \mathbb{N}$, cette égalité vaut pour N arbitrairement grand, i.e.

$$\|x\|^2 - \sum_{n\in\mathbb{N}} |\langle x, u_n\rangle|^2 = \left\| x - \sum_{n\in\mathbb{N}} \langle x, u_n\rangle u_n \right\|^2.$$

□

Nous allons maintenant généraliser la définition de coefficients de Fourier donnée dans le chapitre 2.

Déf. 5.5.3 (Coefficients de Fourier généralisés) *Les scalaires $\langle x, u_n\rangle \in \mathbb{K}$ sont dits les coefficients de Fourier généralisés de x par rapport au s.o.n $(u_n)_{n\in\mathbb{N}}$ et ils sont écrits comme ça :*

$$\boxed{\hat{x}(n) = \langle x, u_n\rangle} \quad \forall n \in \mathbb{N}.$$

On peut donc reformuler l'inégalité de Bessel en disant que, pour tout $x \in \mathcal{H}$, la suite

$$\hat{x} \equiv (\hat{x}_n)_{n\in\mathbb{N}}$$

appartient à $\ell^2(\mathbb{N}, \mathbb{K})$ et que

$$\|\hat{x}\|_2 \leqslant \|x\| \quad \forall x \in \mathcal{H}.$$

Nous observons aussi que, nécessairement, *la suite des coefficients de Fourier est une suite décroissante vers 0.* Pour des espaces de Hilbert où x est identifiable avec une fonction, l'analyse de la vitesse de décroissance des coefficients de Fourier donne des informations très intéressantes sur la régularité de x.

L'équation (5.6) donne l'estimation de l'écart entre $\|x\|^2$ et ${\|\hat{x}\|_2}^2$ et, réécrite avec la notation qu'on vient d'introduire, implique immédiatement le corollaire suivant.

Corollaire 5.5.1 *Soit $\mathcal{H}$ un espace de Hilbert et $(u_n)_{n\in\mathbb{N}}$ un s.o.n. quelconque de $\mathcal{H}$. Alors :*

$$\left\| x - \sum_{n\in\mathbb{N}} \hat{x}(n) u_n \right\|^2 = \|x\|^2 - {\|\hat{x}\|_2}^2,$$

en particulier, $\sum\limits_{n\in\mathbb{N}} \hat{x}(n)u_n$ converge vers x si et seulement si $\|\hat{x}\|_2 = \|x\|$.

5.5.2 Le théorème de Fischer - Riesz

Le théorème (fondamentale) suivant est souvent appelé théorème de Fischer-Riesz dans la littérature, par exemple dans le livre classique [7].

Théorème 5.5.3 (Fischer - Riesz) *Soit $\mathcal{H}$ un espace de Hilbert, $(u_n)_{n\in\mathbb{N}}$ un s.o.n de $\mathcal{H}$ et $(k_n)_{n\in\mathbb{N}}$ une suite de scalaires de $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.*

1. *Alors,*

$$\sum_{n\in\mathbb{N}} k_n u_n \text{ converge (en norme } \|\ \| \text{ de } \mathcal{H}) \iff \sum_{n\in\mathbb{N}} |k_n|^2 \text{ converge (en } \mathbb{K}),$$

   i.e. $\sum\limits_{n\in\mathbb{N}} k_n u_n$ *converge* $\iff (k_n)_{n\in\mathbb{N}} \in \ell^2(\mathbb{N}, \mathbb{K})$.

2. *Si* $\sum\limits_{n\in\mathbb{N}} k_n u_n$ *converge vers la somme* x, *i.e.* $x = \sum\limits_{n\in\mathbb{N}} k_n u_n$, *alors :*

$$k_n = \langle x, u_n\rangle = \hat{x}(n)$$

   et

$$\|x\|^2 = \sum_{n\in\mathbb{N}} |k_n|^2,$$

   i.e. l'inégalité de Bessel devient l'égalité de Plancherel $\|x\|^2 = \sum\limits_{n\in\mathbb{N}} |\langle x, u_n\rangle|^2 = {\|\hat{x}\|_2}^2$.

Preuve.

1. On doit vérifier que l'étude de la convergence de $\sum_{n\in\mathbb{N}} k_n u_n$ et de $\sum_{n\in\mathbb{N}} |k_n|^2$ sont équivalents. Pour cela on utilise le fait que $\mathcal{H}$ et $\mathbb{K}$ sont complets, donc la condition de Cauchy est nécessaire et suffisante pour la convergence des suites, en rappelant que la série $\sum_{n\in\mathbb{N}} k_n u_n$ est la suite $(S_N)_{N\in\mathbb{N}} = \left(\sum_{n=1}^{N} k_n u_n\right)_{N\in\mathbb{N}}$ des sommes partielles.

La condition Cauchy pour $S_N = \sum_{n=1}^{N} k_n u_n$ est :

$$\forall \varepsilon > 0\ \exists K_\varepsilon > 0\ :\ r > s \geqslant K_\varepsilon \implies \|S_r - S_s\| = \left\| \sum_{n=s+1}^{r} k_n u_n \right\| < \varepsilon.$$

Vu que $\left\| \sum_{n=s+1}^{r} k_n u_n \right\| < \varepsilon \quad \Leftrightarrow \quad \left\| \sum_{n=s+1}^{r} k_n u_n \right\|^2 < \varepsilon^2 \equiv \delta$, car l'inégalité concerne deux nombres réels positifs, on peut re-définir la condition de Cauchy pour $(S_N)_{N\in\mathbb{N}}$ comme ça :

$$\forall \delta > 0\ \exists K_\delta > 0\ :\ r > s \geqslant K_\delta \implies \left\| \sum_{n=s+1}^{r} k_n u_n \right\|^2 < \delta.$$

Maintenant on utilise l'orthogonalité des u_n pour développer la norme au carré en utilisant le théorème de Pythagore généralisé et le fait que $\|u_n\| = 1\ \forall n \in \mathbb{N}$:

$$\left\| \sum_{n=s+1}^{r} k_n u_n \right\|^2 = \sum_{n=s+1}^{r} \|k_n u_n\|^2 = \sum_{n=s+1}^{r} |k_n|^2 \|u_n\|^2 = \sum_{n=s+1}^{r} |k_n|^2,$$

mais alors la condition de Cauchy pour la suite des sommes partielles de la séries $\sum_{n\in\mathbb{N}} k_n u_n$ peut être réécrite comme :

$$\forall \delta > 0\ \exists K_\delta > 0\ :\ r > s \geqslant K_\delta \implies \sum_{n=s+1}^{r} |k_n|^2 < \delta,$$

mais celle-ci est la condition de Cauchy pour la suite des sommes partielles de la séries $\sum_{n\in\mathbb{N}} |k_n|^2$, donc l'étude de la convergence des deux séries est équivalent.

2. Supposons la convergence de la séries $\sum_{m\in\mathbb{N}} k_m u_m$ vers la somme x, alors, par la continuité du produit scalaire :

$$\langle x, u_n \rangle = \langle \sum_{m\in\mathbb{N}} k_m u_m, u_n \rangle = \sum_{m\in\mathbb{N}} k_m \langle u_m, u_n \rangle = \sum_{m\in\mathbb{N}} k_m \delta_{m,n} = k_n,$$

$\forall n \in \mathbb{N}$, d'où : $x = \sum_{n\in\mathbb{N}} \langle x, u_n \rangle u_n = \sum_{n\in\mathbb{N}} \hat{x}(n) u_n$. Le fait que **2.** implique que l'inégalité de Bessel devient l'égalité de Plancherel est une conséquence immédiate du corollaire (5.5.1). Alternativement, nous pouvons le démontrer avec la continuité de la norme :

$$\|x\|^2 = \| \sum_{n\in\mathbb{N}} \langle x, u_n \rangle u_n \|^2 = \sum_{n\in\mathbb{N}} |\langle x, u_n \rangle|^2 \|u_n\|^2 = \sum_{n\in\mathbb{N}} |\langle x, u_n \rangle|^2 = \|\hat{x}\|_2{}^2.$$

□

Corollaire 5.5.2 *Soit $\mathcal{H}$ un espace de Hilbert et $(u_n)_{n\in\mathbb{N}}$ un s.o.n de $\mathcal{H}$. Alors, la série $\sum_{n\in\mathbb{N}} \hat{x}(n)u_n$ est toujours convergente (par rapport à la norme $\|\ \|$ de $\mathcal{H}$).*

Preuve. Grâce à l'inégalité de Bessel, $(\hat{x}(n))_{n\in\mathbb{N}} \in \ell^2(\mathbb{N},\mathbb{K})$, i.e. $\sum_{n\in\mathbb{N}} |\hat{x}(n)|^2$ est convergente en $\mathbb{K}$, mais alors, grâce au théorème de Fischer-Riesz, la série $\sum_{n\in\mathbb{N}} \hat{x}(n)u_n$ est convergente en $\mathcal{H}$. □

Exemple remarquable : montrons un cas explicite de série $\sum_{n\in\mathbb{N}} \hat{x}(n)u_n$, étant $(u_n)_{n\in\mathbb{N}}$ un s.o.n dans un espace de Hilbert $\mathcal{H}$, qui converge vers un élément $y \in \mathcal{H}$, $y \neq x$.

Prenons : $\mathcal{H} = L^2[-\pi,\pi]$, $u_n(t) = \frac{1}{\sqrt{\pi}}\sin(nt)$, $n \in \mathbb{N}$ et $t \in [-\pi,\pi]$. On peut vérifier que $(u_n)_n \in \mathbb{N}$ est un système orthonormal pour $\mathcal{H}$. Soit $x(t) = \cos(t)$, alors, par calcul direct :

$$\sum_{n\in\mathbb{N}} \hat{x}(n)u_n = \sum_{n=1}^{\infty} \frac{1}{\pi}\left(\int_{-\pi}^{\pi} \cos(t)\sin(nt)dt\right)\sin(nt),$$

mais $\int_{-\pi}^{\pi} \cos(t)\sin(nt)dt = 0$ car c'est l'intégrale sur un domaine symétrique d'une fonction impaire, donc

$$\sum_{n\in\mathbb{N}} \hat{x}(n)u_n = \sum_{n=1}^{\infty} 0 \cdot \sin(nt) = \mathbf{0},$$

où $\mathbf{0}$ est la fonction identiquement nulle sur $[-\pi,\pi]$, qui, évidemment, est différente de la fonction $\cos(t)$, donc $\sum_{n\in\mathbb{N}} \hat{x}(n)u_n \neq x$.

5.5.3 Les caractérisations d'une base hilbertienne

L'exemple que l'on vient de voir montre qu'un s.o.n. quelconque dans une espace de Hilbert $\mathcal{H}$ n'est pas suffisant pour garantir la convergence en norme de la série des coefficients de Fourier de $x \in \mathcal{H}$ multipliés par les éléments du s.o.n à x lui-même.

On peut se demander d'une manière naturelle s'il existe une condition qui assure la convergence.

Dans la section 1.5, on a vu que, en dimension finie, cette condition est que le s.o.n. soit une base orthonormale, i.e. un système maximal de vecteurs unitaires linéairement indépendants et orthogonaux, où « maximal » veut dire qu'il n'existe pas un autre vecteur unitaire orthogonale aux vecteurs de la base considérée.

Cette propriété caractérise les « bases » aussi dans le cas d'un espace de Hilbert de dimension infinie, mais dans ce cas on utilise une nomenclature différente.

Déf. 5.5.4 (s.o.n.c) *Soit $(u_n)_{n\in\mathbb{N}} \subset \mathcal{H}$ un s.o.n. d'un espace de Hilbert $\mathcal{H}$. Si $(u_n)_{n\in\mathbb{N}}$ n'est pas une partie propre d'un autre système orthonormal de $\mathcal{H}$, i.e. s'il n'existe pas un autre vecteur unitaire orthogonal aux vecteurs $(u_n)_{n\in\mathbb{N}}$, alors on l'appelle un système orthonormal* ***complet*** *(s.o.n.c. dorénavant), ou* ***total****, ou* ***base hilbertienne****.*

Remarquablement, la propriété d'être une base hilbertienne, dans le sens qu'on vient de définir, est équivalente à 5 autres propriétés, comme on le voit dans le résultas suivant, qui généralise le théorème 1.8.2 aux espaces de Hilbert de dimension infinie.

Théorème 5.5.4 *Soit $(u_n)_{n\in\mathbb{N}}$ un système orthonormé d'un espace de Hilbert $\mathcal{H}$. Les affirmations suivantes sont équivalentes :*

1. *$(u_n)_{n\in\mathbb{N}}$ est une base hilbertienne*

2. *$\langle x, u_n\rangle \equiv \hat{x}(n) = 0\ \forall n \in \mathbb{N} \iff x = 0_{\mathcal{H}}$, i.e. $0_{\mathcal{H}}$ est le seul vecteur orthogonal à tous les vecteurs d'un s.o.n.c. (ou, le seul vecteur $x \in \mathcal{H}$ qui a coefficients de Fourier généralisés tous nuls sur un s.o.n.c. de $\mathcal{H}$ est le vecteur nul)*

3. *$\overline{\text{span}((u_n)_{n\in\mathbb{N}})} = \mathcal{H}$, i.e. $(u_n)_{n\in\mathbb{N}}$ génère un sous-espace vectoriel dense en $\mathcal{H}$*

4. *$\forall x \in \mathcal{H}$:*

$$x = \sum_{n\in\mathbb{N}} \langle x, u_n\rangle u_n = \sum_{n\in\mathbb{N}} \hat{x}(n) u_n \qquad \textbf{\textit{Expansion en série de Fourier généralisée}}$$

5. *$\forall x, y \in \mathcal{H}$:*

$$\langle x, y\rangle = \sum_{n\in\mathbb{N}} \langle x, u_n\rangle\langle u_n, y\rangle = \langle \hat{x}, \hat{y}\rangle_{\ell^2(\mathbb{N},\mathbb{K})} \qquad \textbf{\textit{Identité de Parseval}}$$

6. *$\forall x \in \mathcal{H}$:*

$$\|x\|^2 = \sum_{n\in\mathbb{N}} |\langle x, u_n\rangle|^2 = \|x\|_2{}^2 \qquad \textbf{\textit{Identité de Plancherel.}}$$

Preuve. On procédera de la manière suivante : **1.** $\Rightarrow$ **2.** $\Rightarrow$ **3.** $\Rightarrow$ **4.** $\Rightarrow$ **5.** $\Rightarrow$ **6.** $\Rightarrow$ **1.**

(**1.** $\Rightarrow$ **2.**) : par l'absurde, si **1.** est vrai et **2.** non, alors $\exists x^* \in \mathcal{H}$, $x^* \neq 0_{\mathcal{H}}$ tel que : $\langle x^*, u_n\rangle = 0\ \forall n \in \mathbb{N}$, mais alors le vecteur $u^* = \dfrac{x^*}{\|x^*\|}$ est un vecteur unitaire et $\perp$ à tous les éléments de $(u_n)_{n\in\mathbb{N}}$ et donc $(u^*, (u_n)_{n\in\mathbb{N}})$ serait un s.o.n plus grand que $(u_n)_{n\in\mathbb{N}}$, ce qui est contraire à la totalité de $(u_n)_{n\in\mathbb{N}}$.

(**2.** $\Rightarrow$ **3.**) : **2.** $\Rightarrow ((u_n)_{n\in\mathbb{N}})^\perp = \{0_{\mathcal{H}}\} \iff \left(\overline{\text{span}((u_n)_{n\in\mathbb{N}})}\right)^\perp = \{0_{\mathcal{H}}\}$, grâce à la propriété **8.** du théorème 5.1.1, si on prend le complément orthogonal aux deux côtés on obtient $\left(\overline{\text{span}((u_n)_{n\in\mathbb{N}})}\right)^{\perp\perp} = \{0_{\mathcal{H}}\}^\perp = \mathcal{H}$, mais alors $\mathcal{H} = \left(\overline{\text{span}((u_n)_{n\in\mathbb{N}})}\right)^{\perp\perp} = \overline{\text{span}((u_n)_{n\in\mathbb{N}})}$, grâce au théorème 5.2.4.

(**3.** $\Rightarrow$ **4.**) : considérons x, calculons les produits scalaires avec $(u_n)_{n\in\mathbb{N}}$ et écrivons la série $\sum_{n\in\mathbb{N}} \langle x, u_n\rangle u_n$, qu'on sait être convergente vers un certain $y \in \mathcal{H}$, on doit utiliser la **3.** pour démontrer que $y = x$. La deuxième partie du théorème de Fischer-Riesz assure que $\langle x, u_n\rangle = \langle y, u_n\rangle\ \forall n \in \mathbb{N}$, i.e. $\langle x - y, u_n\rangle = 0$, $\forall n \in \mathbb{N}$, i.e. $x - y \in ((u_n)_{n\in\mathbb{N}})^\perp = \left(\overline{\text{span}((u_n)_{n\in\mathbb{N}})}\right)^\perp \underset{(\mathbf{3.})}{=} \mathcal{H}^\perp = \{0_{\mathcal{H}}\}$, i.e. $y = x$.

(**4.** $\Rightarrow$ **5.**) : on considère $x, y \in \mathcal{H}$ quelconques et on écrit leurs séries de Fourier généralisées : grâce à **4.** on a :

$$x = \sum_{n\in\mathbb{N}} \langle x, u_n\rangle u_n \quad y = \sum_{m\in\mathbb{N}} \langle y, u_m\rangle u_m$$

donc

$$\langle x, y\rangle = \langle \sum_{n\in\mathbb{N}} \langle x, u_n\rangle u_n, \sum_{m\in\mathbb{N}} \langle y, u_m\rangle u_m\rangle,$$

Par continuité du produit scalaire et sa linéarité :

$$\langle x, y\rangle = \sum_{n\in\mathbb{N}} \langle x, u_n\rangle \langle u_n, \sum_{m\in\mathbb{N}} \langle y, u_m\rangle u_m\rangle,$$

puis, par continuité du produit scalaire et sa sesquilinéairité :

$$\langle x, y\rangle = \sum_{n\in\mathbb{N}} \sum_{m\in\mathbb{N}} \langle x, u_n\rangle \overline{\langle y, u_m\rangle} \langle u_n, u_m\rangle = \sum_{n\in\mathbb{N}} \sum_{m\in\mathbb{N}} \langle x, u_n\rangle \langle u_m, y\rangle \delta_{n,m},$$

i.e. :

$$\langle x, y\rangle = \sum_{n\in\mathbb{N}} \langle x, u_n\rangle \langle u_n, y\rangle.$$

(**5.** $\Rightarrow$ **6.**) : ça suffit de considérer $y = x$ en **5.** : $\|x\|^2 = \langle x, x\rangle = \sum_{n\in\mathbb{N}} \langle x, u_n\rangle\langle u_n, x\rangle = \sum_{n\in\mathbb{N}} \langle x, u_n\rangle\overline{\langle x, u_n\rangle} = \sum_{n\in\mathbb{N}} |\langle x, u_n\rangle|^2$.

(**6.** $\Rightarrow$ **1.**) : par l'absurde, si **6.** est vraie et **1.** non, alors $\exists u^* \in \mathcal{H}$, $\|u^*\| = 1$ et $\langle u^*, u_n\rangle = 0$ $\forall n \in \mathbb{N}$, on aurait alors $\sum_{n\in\mathbb{N}} |\langle u^*, u_n\rangle|^2 = 0$, ce qui est contraire à **6.** car elle dit que $\sum_{n\in\mathbb{N}} |\langle u^*, u_n\rangle|^2 = \|u^*\|^2 = 1$. □

Observation importante sur la propriété **4.** : l'expansion en série de Fourier généralisée sur une base hilbertienne est l'extension du théorème de décomposition d'un vecteur sur une base orthonormale dans un espace Euclidien de dimension finie d, comme le montre le tableau suivant.

$\mathbb{K}^d$	$\mathcal{H}$: Espace de Hilbert
Base orthonormale : $(u_i)_{i=1,\ldots,d}$	Base hilbertienne : $(u_n)_{n\in\mathbb{N}}$
Expansion : $\forall x \in \mathbb{K}^d \quad x = \sum_{i=1}^{d} \langle x, u_i\rangle u_i$	Série de Fourier : $\forall x \in \mathcal{H} \quad x = \sum_{n\in\mathbb{N}} \langle x, u_n\rangle u_n$
Composantes : $\langle x, u_i\rangle$	Coefficients de Fourier : $\langle x, u_n\rangle$

La généralisation de la base canonique de l'espace $\ell^2(\mathbb{Z}_N)$ qu'on a introduit dans la section 2.1 est **la base hilbertiennes canonique de** $\mathcal{H} = \ell^2(\mathbb{Z}, \mathbb{K})$ donnée par les vecteurs $(e_k)_{k\in\mathbb{Z}}$, $e_k(n) = \delta_{k,n}$ $\forall n \in \mathbb{Z}$:

$$(e_1 = (1, 0, 0, \ldots),\ e_2 = (0, 1, 0, \ldots),\ \ldots).$$

L'orthonormalité est évidente, la complétude descend, par exemple, du fait qu'il est immédiat de voir que le seul vecteur orthogonal à $e_1, e_2, \ldots$ est le vecteur nul.

Théorème 5.5.5 *Tout espace de Hilbert $\mathcal{H}$ admet une base hilbertienne.*

Preuve. Notons avec O la collection de toutes les familles orthonormées de $\mathcal{H}$. O est un ensemble ordonné par inclusion. Si $\Phi \subset O$ est linéairement ordonné, alors la réunion de tous les éléments de Φ est un extrême supérieur. Le lemme de Zorn [11] garantit qu'il existe un élément maximal en O. □

Exemple d'un espace de Hilbert non-séparable. Nous rappelons que nous avons implicitement assumé que l'espace de Hilbert du théorème 5.5.4 soit séparable, donc, tout espace de Hilbert qui ne vérifie pas une des propriétés qui caractérisent une base hilbertienne est non-séparable. Utilisons, en particulier, la propriété 2. du théorème 5.5.4 pour montrer un exemple d'espace de Hilbert non-séparable. Définissons l'espace suivant :

$$\mathcal{H} = \{f : \mathbb{R} \to \mathbb{K} \ : \ \exists\, E_f \subset \mathbb{R},\ \mathrm{card}(E_f) \leqslant \aleph_0 \ : \ f|_{E_f} \in \ell^2(\mathbb{N}, \mathbb{K}) \text{ et } \ f|_{\mathbb{R}\setminus E_f} = \mathbf{0}_{\mathbb{R}\setminus E_f}\},$$

i.e. l'espace constitué par toute fonction f définies sur $\mathbb{R}$ à valeur en $\mathbb{K}$, qui s'annule partout sauf sur un sous-ensemble fini ou dénombrable E_f de $\mathbb{R}$ et telle que la suite (finie ou infinie) $f : E_f \to \mathbb{K}$ est carré sommable.

$\mathcal{H}$ est un espace vectoriel, par rapport aux opérations linéaires définies ponctuellement, qui peut être muni du produit scalaire suivant :

$$\langle f, g\rangle = \sum_{x\in E_f\cap E_g} f(x)\overline{g(x)} \qquad f, g \in \mathcal{H},$$

bien défini vu que la somme est, par définition de $\mathcal{H}$, soit une somme finie, soit une série convergente (évidemment, si $\mathbb{K} = \mathbb{R}$ l'opération de conjugaison devient l'identité). Il est simple de vérifier que $\mathcal{H}$ est un espace de Hilbert par rapport à la topologie induite par ce produit scalaire.

Supposons, par l'absurde, que $\mathcal{H}$ est séparable, alors toute base hilbertienne est dénombrable. Soit alors $u \equiv (u_n)_{n\in\mathbb{N}}$ une base hilbertienne de $\mathcal{H}$ sous l'hypothèse de séparabilité et soit $U := \bigcup_{n\in\mathbb{N}} U_n$, où les ensembles $U_n \subset \mathbb{R}\ \forall n \in \mathbb{N}$ sont tels que $u_n|_{U_n} \in \ell^2(\mathbb{N}, \mathbb{K})$ et $u_n|_{\mathbb{R}\setminus U_n} = \mathbf{0}_{\mathbb{R}\setminus U_n}$
Si on montre qu'il existe un élément f_u de $\mathcal{H}$ qui est orthogonal à tout u_n et qui n'est pas la fonction identiquement nulle sur $\mathbb{R}$, alors on montre que la propriété 2. du théorème 5.5.4 ne vaut pas, cette contradiction implique que $\mathcal{H}$ ne peut pas être séparable.

Pour construire un tel élément, observons que U est la réunion d'ensembles dénombrables ou finis, donc lui aussi est soit dénombrable soit fini. Alors, il suffit de considérer un point quelconque $\bar{x} \in \mathbb{R}\setminus U$ et définir $f_u : \mathbb{R} \to \mathbb{K}$ comme ça

$$f_u(x) = \begin{cases} 1 & \text{si } x = \bar{x} \\ 0 & \text{sinon,} \end{cases}$$

pour avoir un élément de $\mathcal{H}$ tel que $\langle u_n, f_u\rangle = 0\ \forall n \in \mathbb{N}$, mais $f \neq \mathbf{0}_{\mathbb{R}}$.

Le fait que toutes les bases hilbertiennes d'un espace de Hilbert séparable $\mathcal{H}$ de dimension non-finie soient dénombrables ne doit pas faire penser que $\mathcal{H}$ lui-même a une dimension dénombrable *en tant qu'espace vectoriel*.

C'est-à-dire, si on oublie que $\mathcal{H}$ est un espace de Hilbert et qu'on le traite simplement comme un espace vectoriel, alors sa dimension est, par définition, le cardinal d'une base au sens algébrique, i.e. un sous-ensemble $B \subset \mathcal{H}$ d'éléments linéairement indépendants de $\mathcal{H}$ tel que tout élément de $\mathcal{H}$ peut être obtenu comme combinaison linéaire *finie* d'éléments de la base B. En fait, ça vaut le théorème suivant, dont on assume la preuve.

Théorème 5.5.6 *Si le cardinal commun des bases hilbertiennes d'un espace de Hilbert $\mathcal{H}$ (séparable ou non) est $\aleph_0$, i.e. dénombrable, alors la dimension de $\mathcal{H}$ en tant qu'espace vectoriel ne peut pas être inférieure à $\aleph_1$, i.e continue.*

Comme conséquence du théorème, les espaces de Hilbert séparables ont, au moins, le cardinal du continu en tant qu'espaces vectoriels, i.e. un système maximal de vecteurs linéairement indépendants a, au moins, le cardinal du continu ! Demander l'orthonormalité rajoute la contrainte que la distance entre les éléments de la base doit être $\sqrt{2}$, qui fait descendre le cardinal au dénombrable.

Néanmoins, il faut souligner encore une fois qu'avec une base hilbertienne on peut reconstruire tout élément d'un espace de Hilbert via série de Fourier généralisée dans le sens de la norme hilbertienne, ce qui est bien différent de la possibilité de reconstruire tout élément à travers d'une combinaison linéaire finie.

Ceci montre que c'est une base hilbertienne l'objet le plus adapté à donner une « paramétrisation » d'un espace de Hilbert via les coefficients de Fourier généralisés, et non pas une base dans le sens des espaces vectoriels.

La raison est qu'une base hilbertienne possède des propriétés qui la rendent compatible avec la riche structure géométrique de l'espace de Hilbert générée par le produit scalaire, par contre une base dans le sens vectoriel classique tient en considération seulement sa structure linéaire.

À cause de cette différence, certains auteurs introduisent la nomenclature suivante.

Déf. 5.5.5 (Dimension orthogonale) *Soit $\mathcal{H}$ un espace de Hilbert. On appelle dimension orthogonale de $\mathcal{H}$ le cardinal commun de toutes les bases hilbertiennes de $\mathcal{H}$.*

Il est clair que la dimension orthogonale coïncide avec celle ordinaire pour un espace de Hilbert de dimension finie, mais ceci n'est plus vrai en dimension infinie.

5.5.4 Isomorphismes entre espaces de Hilbert

ll nous reste à discuter une dernière propriété qui montre l'analogie entre espaces de Hilbert et espaces Euclidiens de dimension finie : *l'existence d'un prototype* pour ces espaces.

Nous avons vu que la dimension d'un espace vectoriel V de dimension finie d est suffisante pour le caractériser à un isomorphisme près. En fait, nous savons que, une fois qu'on fixe une base quelconque de V, la correspondance $I : V \to \mathbb{K}^d$ qui associe à chaque vecteur v de V ses composantes (en $\mathbb{K}^d$) par rapport à la base choisie est un isomorphisme. C'est dans ce sens que $\mathbb{K}^d$ est le prototype des espaces vectoriels sur $\mathbb{K}$ de dimension $d < +\infty$. Pour les espaces de Hilbert (séparables) de dimension infinie, le prototype est $\ell^2(\mathbb{N}, \mathbb{K})$.

Pour rendre rigoureuse cette affirmation, on doit d'abord définir le concept d'isomorphisme entre espaces de Hilbert. La présence du produit scalaire implique une modification dans la définition canonique d'isomorphisme entre espaces vectoriels.

Déf. 5.5.6 (Isomorphisme entre espaces de Hilbert) *Soient $\mathcal{H}$ et $\mathcal{H}'$ deux espaces de Hilbert sur le même corps $\mathbb{K}$. La transformation $U : \mathcal{H} \to \mathcal{H}'$ est un isomorphisme d'espaces de Hilbert si :*

1. *U est linéaire*
2. *U est bijective*
3. *U préserve le produit scalaire, i.e.*

$$\langle U(x), U(y)\rangle_{\mathcal{H}'} = \langle x, y\rangle_{\mathcal{H}} \qquad \forall x, y \in \mathcal{H}.$$

La requête **3.** implique (dans le cas particulier $x = y$) que U préserve les normes, i.e.

$$\|U(x)\|_{\mathcal{H}'} = \|x\|_{\mathcal{H}} \qquad \forall x \in \mathcal{H}.$$

Ceci implique aussi :

$$\|U(x) - U(y)\|_{\mathcal{H}'} = \|U(x-y)\|_{\mathcal{H}'} = \|x - y\|_{\mathcal{H}} \qquad \forall x, y \in \mathcal{H},$$

i.e. U préserve les distances, on dit dans ce cas que U est une **isométrie**.

Nous observons que la propriété de préservation de la norme implique que $\|U(x)\|_{\mathcal{H}'} = 0 \iff \|x\|_{\mathcal{H}} = 0$, mais alors, grâce à la définie positivité de la norme, ceci est équivalent à $U(x) = 0_{\mathcal{H}'} \iff x = 0_{\mathcal{H}}$, i.e. $\ker(U) = \{0_{\mathcal{H}}\}$ et donc U est injective.

Nous pourrions alors redéfinir un isomorphisme U entre espaces de Hilbert comme une transformation linéaire, qui préserve le produit scalaire (donc injective) et surjective. En réalité, demander la linéarité serait redondant, comme montré par le résultat suivant.

Théorème 5.5.7 *Soient V, V' deux espaces vectoriels, de dimension finie ou infinie, sur le même corps $\mathbb{K}$ avec produit scalaire. Si la transformation $U : V \to V'$ est surjective et préserve le produit scalaire, alors elle est linéaire.*

Preuve. $\forall x, y, z \in V$ et $\forall \alpha, \beta \in \mathbb{K}$:

$$\begin{aligned} 0 = \langle 0, z\rangle &= \langle \alpha x + \beta y - \alpha x - \beta y, z\rangle = \langle \alpha x + \beta y, z\rangle - \alpha\langle x, z\rangle - \beta\langle y, z\rangle \\ &\underset{(U \text{ preserve } \langle\,\rangle)}{=} \langle U(\alpha x + \beta y), U(z)\rangle - \alpha\langle U(x), U(z)\rangle - \beta\langle U(y), U(z)\rangle \\ &\underset{(\text{linéarité } \langle\,\rangle)}{=} \langle U(\alpha x + \beta y) - \alpha U(x) - \beta U(y), U(z)\rangle. \end{aligned}$$

Comme U est surjective par hypothèse, si on laisse varier $z \in V$, $U(z)$ représente tout élément de V', donc, $U(\alpha x + \beta y) - \alpha U(x) - \beta U(y)$ est orthogonal à tous les éléments de V', i.e. $U(\alpha x + \beta y) - \alpha U(x) - \beta U(y) = 0_{\mathcal{H}'}$, d'où :

$$U(\alpha x + \beta y) = \alpha U(x) + \beta U(y) \qquad \forall x, y \in V,\ \forall \alpha, \beta \in \mathbb{K},$$

U est linéaire. □

La définition d'isomorphisme entre espaces de Hilbert peut être reformulée comme ci-dessous.

Déf. 5.5.7 (Définition alternative d'isomorphisme entre espaces de Hilbert) *Soient $\mathcal{H}$ et $\mathcal{H}'$ deux espaces de Hilbert sur le même corps $\mathbb{K}$. La transformation $U : \mathcal{H} \to \mathcal{H}'$ est un isomorphisme d'espaces de Hilbert si :*

1. *U est surjective*
2. *U préserve le produit scalaire.*

Être isomorphe est une relation d'équivalence dans l'ensemble des espaces de Hilbert sur le même corps $\mathbb{K}$, en particulier ça vaut le théorème suivant (dont on assumera la preuve).

Théorème 5.5.8 *$\mathcal{H}, \mathcal{H}'$: espaces de Hilbert sur le même corps $\mathbb{K}$. $\mathcal{H}$ est isomorphe à $\mathcal{H}'$ si et seulement si la dimension orthogonale de $\mathcal{H}$ est la même de $\mathcal{H}'$.*

5.5.5 $\ell^2(\mathbb{N}, \mathbb{K})$ comme prototype des espaces de Hilbert séparables de dimension infinie

Lemme 5.5.1 *Soit $(u_n)_{n\in\mathbb{N}}$ une base hilbertienne de $\mathcal{H}$, alors, pour toute suite $(k_n)_{n\in\mathbb{N}}$ de $\ell^2(\mathbb{N}, \mathbb{K})$ il existe $x \in \mathcal{H}$ tel que : $(k_n)_{n\in\mathbb{N}} = (\langle x, u_n\rangle)_{n\in\mathbb{N}}$.*

Preuve. Pour tout $x \in \mathcal{H}$, $(k_n)_{n\in\mathbb{N}} = (\langle x, u_n\rangle)_{n\in\mathbb{N}}$ appartient à $\ell^2(\mathbb{N}, \mathbb{K})$ grâce à l'inégalité de Bessel. Inversement, si $(k_n)_{n\in\mathbb{N}} \in \ell^2(\mathbb{N}, \mathbb{K})$, alors, grâce à 1. du théorème de Fischer-Riesz, $\sum_{n\in\mathbb{N}} k_n u_n$ converge, mais comme $(u_n)_{n\in\mathbb{N}}$ est une base hilbertienne de $\mathcal{H}$, la convergence est vers x, et alors la 2. du théorème de Fischer-Riesz garantit que $(k_n)_{n\in\mathbb{N}} = (\langle x, u_n\rangle)_{n\in\mathbb{N}}$. □

Théorème 5.5.9 *Si l'espace de Hilbert $\mathcal{H}$ a dimension orthogonale dénombrable $\aleph_0$, alors $\mathcal{H}$ est isomorphe à $\ell^2(\mathbb{N}, \mathbb{K})$.*

Preuve. Soit $(u_n)_{n\in\mathbb{N}}$ une base hilbertienne dénombrable de $\mathcal{H}$. Considérons l'application :

$$\begin{array}{rccl} U : & \mathcal{H} & \longrightarrow & \ell^2(\mathbb{N}, \mathbb{K}) \\ & x & \longmapsto & U(x) = (\langle x, u_n\rangle)_{n\in\mathbb{N}}. \end{array}$$

U est surjective grâce au Lemme 5.5.1. U préserve le produit scalaire grâce à l'identité de Parseval, en fait :

$$\forall x, y \in \mathcal{H} \ : \ \langle x, y\rangle_{\mathcal{H}} = \sum_{n\in\mathbb{N}} \langle x, u_n\rangle\langle u_n, y\rangle = \sum_{n\in\mathbb{N}} \langle x, u_n\rangle\overline{\langle y, u_n\rangle} \equiv \langle U(x), U(y)\rangle_{\ell^2(\mathbb{N},\mathbb{K})}.$$

□

5.6 La base hilbertienne de Fourier en L^2

L'exemple le plus connu, et le plus important dans les applications, de base hilbertienne est la base de Fourier, qu'on va définir ci-dessous dans le cadre de l'espace de Hilbert L^2 avec des choix différents du domaine des fonctions carrés intégrables.

5.6.1 $L^2[-\pi,\pi]$ ou $L^2[0,2\pi]$

Commençons avec $\mathcal{H} = L^2[-\pi,\pi]$ ou $L^2[0,2\pi]$ et $\mathbb{K} = \mathbb{C}$, alors :

$$u_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx}, \qquad n \in \mathbb{Z},$$

est un s.o.n.c dit la base hilbertienne de Fourier de $L^2[-\pi,\pi]$ ou $L^2[0,2\pi]$. À noter que ce s.o.n.c. complète le s.o.n. $\frac{1}{\sqrt{\pi}}\sin(nx)$, $n \in \mathbb{N}$ qu'on avait utilisé pour démontrer que, pour les s.o.n., la convergence de la série de Fourier généralisée (en norme hilbertienne) vers l'élément qui définit les coefficients de Fourier généralisés n'est pas assuré.

L'orthonormalité est simple à prouver, en fait (considérons $L^2[-\pi,\pi]$, pour $L^2[0,2\pi]$ la preuve est la même) :

$$\langle u_n, u_m \rangle = \int_{-\pi}^{\pi} u_n(x)\overline{u_m(x)}dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{inx}e^{-imx}dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(n-m)x}dx.$$

— Si $n = m$, alors $e^{i(n-m)x} = e^0 = 1$ et donc $\langle u_n, u_n \rangle = \|u_n\|^2 = 1$;
— Si $n \neq m$, alors, en écrivant $y = i(n-m)$, on peut réécrire le produit scalaire comme ceci :

$$\langle u_n, u_m \rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{yx}dx = \frac{1}{2\pi y}\,[e^{yx}]_{x=-\pi}^{x=\pi} = \frac{1}{2\pi i(n-m)}[e^{i(n-m)\pi} - e^{i(m-n)\pi}] = 0.$$

En résumé, $\langle u_n, u_m \rangle = \delta_{n,m}$, ce qui prouve l'orthonormalité. La preuve que le système soit complet est beaucoup plus compliquée et on l'assumera.

L'expansion en série de Fourier ici s'écrit :

$$\forall f \in L^2[-\pi,\pi] \;:\; f = \sum_{n\in\mathbb{Z}} \langle f, u_n\rangle u_n,$$

où :

$$\langle f, u_n \rangle \equiv \hat{f}(n) = \frac{1}{\sqrt{2\pi}}\int_{-\pi}^{\pi} f(x)e^{-inx}dx,$$

est l'n-ième **coefficient de Fourier** de f. On souligne que la convergence de la série doit être interprétée comme ça :

$$\int_{-\pi}^{\pi} \left| f(x) - \sum_{n=-N}^{N} \hat{f}(n)u_n(x) \right|^2 dx \underset{N\to+\infty}{\to} 0.$$

Déf. 5.6.1 *Soit $\mathcal{H} = L^2[-\pi,\pi]$ ou $L^2[0,2\pi]$. L'application :*

$$F \equiv \hat{}: \begin{array}{ccc} \mathcal{H} & \longrightarrow & \ell^2(\mathbb{Z},\mathbb{C}) \\ f & \longmapsto & (\hat{f}(n))_{n\in\mathbb{Z}}. \end{array}$$

est dite transformée de Fourier sur $\mathcal{H} = L^2[-\pi,\pi]$ ou $L^2[0,2\pi]$.

On observe que F coïncide avec la transformation qui implémente l'isomorphisme entre $L^2[-\pi,\pi]$ et son prototype $\ell^2(\mathbb{Z},\mathbb{C})$!

La base hilbertienne de Fourier de $L^2([-\pi,\pi])$ et $L^2([0,2\pi])$ peut être écrite en termes de fonctions réelles :

$$\begin{cases} u_0 \equiv \frac{1}{\sqrt{2\pi}} \\ \cos_n(x) \equiv \frac{1}{\sqrt{\pi}}\cos(nx), \quad n \in \mathbb{N} \\ \sin_n(x) \equiv \frac{1}{\sqrt{\pi}}\sin(nx), \quad n \in \mathbb{N}, \end{cases}$$

il faut observer que l'exponentiel complexe avec paramètre $n \in \mathbb{Z}$ est remplacé par deux suites réelles avec paramètre $n \in \mathbb{N}$ ce qui n'est pas surprenant grâce à la formule de Euler $e^{i\vartheta} = \cos\vartheta + i\sin\vartheta$.

Cette base a l'avantage de ne pas présenter l'unité imaginaire et, en plus, le développement en série de Fourier d'une fonction :

— paire est fait avec u_0 et $\cos_n$

— impaire est fait avec $\sin_n$,

en fait, supposons f paire, alors $\hat{f}(n) = \frac{1}{\sqrt{\pi}}\int_{-\pi}^{\pi} f(x)\sin(nx)dx = 0 \ \forall n$, car $f(x)\sin(nx)$ est impaire et $[-\pi,\pi]$ est un domaine symétrique. Une considération analogue vaut dans le cas de f impaire.

5.6.2 $L^2(\mathbb{T})$

Le choix de considérer l'intervalle $[-\pi,\pi]$ ou $[0,2\pi]$ a été motivé seulement par le fait qu'il est très simple de démontrer l'orthonormalité du système $(\frac{1}{\sqrt{2\pi}}e^{inx})_{n\in\mathbb{Z}}$. Néanmoins, toutes les propriétés qu'on a énoncées pour ce système de fonctions restent valides si on change $[-\pi,\pi]$ ou $[0,2\pi]$ avec n'importe quel intervalle de taille 2π.

De plus, les propriétés restent valides si on considère des fonctions définies sur tout l'intervalle réel, i.e. $f : \mathbb{R} \to \mathbb{C}$, mais 2π-périodiques. Pour formaliser cela, allons à définir un espace de Hilbert très utile :

$$L^2(\mathbb{T}) = \left\{ f : \mathbb{R} \to \mathbb{C} \ , \ f \text{ mesurable} \ , \ f(x+2\pi) = f(x) \ , \int_0^{2\pi} |f(x)|^2 \, \mathrm{d}x < +\infty \right\} / \sim$$

où, comme d'habitude, $f \sim g$ si $f = g$ p.p. Par périodicité, l'intégration peut être faite sur n'importe quel intervalle de taille 2π.

Le symbole $\mathbb{T}$ représente le *tore* 1D, qui peut être identifié avec la circonférence unitaire. Toute fonction $f : \mathbb{R} \to \mathbb{C}$ 2π-périodique peut être identifiée avec une fonction définie sur $\mathbb{T}$ grâce au diagramme suivant :

$$\begin{array}{ccc} \mathbb{R} & \xrightarrow{f} & \mathbb{C} \\ {\scriptstyle p}\downarrow & \nearrow_{\tilde{f}} & \\ \mathbb{T} & & \end{array}$$

$$\begin{array}{rccc} p: & \mathbb{R} & \longrightarrow & \mathbb{T} \\ & x & \longmapsto & (\cos x, \sin x), \end{array}$$

$f : \mathbb{R} \to \mathbb{C}$ 2π-périodique, $\tilde{f} : \mathbb{T} \to \mathbb{C}$, $\tilde{f}(p(x)) = f(x)$.

$L^2(\mathbb{T})$ est isomorphe à $L^2[0,2\pi]$ ou $L^2[-\pi,\pi]$ via l'application qui réalise la restriction de $f \in L^2(\mathbb{T})$, $f : \mathbb{R} \to \mathbb{C}$, à l'intervalle $[0,2\pi]$ ou $[-\pi,\pi]$ (ou n'importe quel intervalle de taille 2π) :

$$\begin{array}{rccc} I: & L^2(\mathbb{T}) & \longrightarrow & L^2[0,2\pi] \\ & f & \longmapsto & If = f|_{[0,2\pi]} \text{ ou } f|_{[-\pi,\pi]}. \end{array}$$

grâce à I on peut transporter le s.o.n.c. de Fourier de $L^2[0,2\pi]$ ou de $L^2[-\pi,\pi]$ sur $L^2(\mathbb{T})$:

$$\left(\frac{1}{\sqrt{2\pi}}e^{inx}\right)_{n\in\mathbb{Z}} : \text{ base hilbertienne pour } L^2(\mathbb{T}),$$

et étendre la définition de la transformée de Fourier sur $L^2(\mathbb{T})$.

Déf. 5.6.2 *La transformation :*

$$\begin{array}{rccc} F \equiv \hat{}: & L^2(\mathbb{T}) & \longrightarrow & \ell^2(\mathbb{Z},\mathbb{C}) \\ & f & \longmapsto & Ff = \widetilde{f}, \end{array}$$

$Ff(n) = \hat{f}(n) = (\langle f, u_n\rangle)_{n\in\mathbb{Z}} = (\frac{1}{\sqrt{2\pi}}\int_0^{2\pi} f(x)e^{-inx}\,dx)_{n\in\mathbb{Z}}$. *est dite transformée de Fourier sur* $L^2(\mathbb{T})$.

On sait que cette transformation est un isomorphisme entre espaces de Hilbert et que : $||\widehat{f}||_{\ell^2(\mathbb{Z},\mathbb{C})} = \sum_{n\in\mathbb{Z}} |\langle f, u_n\rangle|^2 = ||f||_{L^2(\mathbb{T})}$.

5.6.3 $L^2[a,b]$

Si au lieu d'une fonction 2π-périodique on doit gérer une fonction $f \in L^2[a,b]$, $a,b \in \mathbb{R}$, $a < b$, alors il faut changer le s.o.n.c de Fourier pour avoir des fonctions $(b-a)$-périodiques. L'astuce consiste tout simplement en multiplier la variable de f par une quantité opportune, la pulsation, qui rend f une fonction $(b-a)$-périodique. On définit :

- $T = b - a$: période
- $\nu = \frac{1}{T}$: fréquence
- $\omega = 2\pi\nu = \frac{2\pi}{T}$: pulsation

et on observe que

$$\begin{aligned} e^{i\omega n(x+T)} &= \cos[\omega n(x+T)] + i\sin[\omega n(x+T)] = \cos[\omega nx + \omega nT] + i\sin[\omega nx + \omega nT] \\ &= \cos\left[\omega nx + \frac{2\pi}{T}nT\right] + i\sin[\omega nx + \frac{2\pi}{T}nT] = \cos[\omega nx + 2\pi n] + i\sin[\omega nx + 2\pi n] \\ &= \cos(\omega nx) + i\sin(\omega nx) = e^{i\omega nx}, \end{aligned}$$

donc $x \mapsto e^{i\omega nx}$ est une fonction T-périodique. Grâce à ces considérations, on peut démontrer qu'un s.o.n.c de $L^2[a,b]$ est donné par l'ensemble de fonctions :

$$\begin{array}{rccc} u_n: & [a,b] & \longrightarrow & \mathbb{C} \\ & x & \longmapsto & u_n(x) = \frac{1}{\sqrt{b-a}}e^{2\pi ni\frac{x-a}{b-a}}, \quad n \in \mathbb{Z}. \end{array}$$

dans le cas complexe, et :

$$\begin{cases} u_0 = \frac{1}{\sqrt{b-a}} \\ \cos_n(x) \equiv \sqrt{\frac{2}{b-a}} \cos\left(2\pi n \frac{x-a}{b-a}\right), \quad n \in \mathbb{N} \\ \sin_n(x) \equiv \sqrt{\frac{2}{b-a}} \sin\left(2\pi n \frac{x-a}{b-a}\right), \quad n \in \mathbb{N}, \end{cases}$$

dans le cas réel.

Dans le cas particulier de l'espace de Hilbert $L^2[\ell, \ell]$, on peut écrire les bases hilbertiennes de Fourier comme ceci :

$$u_n(x) = \frac{1}{\sqrt{2\ell}} e^{\pi i n \frac{x}{\ell}}, \qquad n \in \mathbb{Z}$$

dans le cas complexe, et :

$$\begin{cases} u_0 = \frac{1}{\sqrt{2\ell}} \\ \cos_n(x) \equiv \sqrt{\frac{1}{\ell}} \cos\left(\pi n \frac{x}{\ell}\right), \quad n \in \mathbb{N} \\ \sin_n(x) \equiv \sqrt{\frac{1}{\ell}} \sin\left(\pi n \frac{x}{\ell}\right), \quad n \in \mathbb{N}, \end{cases}$$

dans le cas réel.

5.6.4 Série de Fourier réelle

Si on utilise la base hilbertienne réelle de $L^2(\mathbb{T})$, i.e.

$$\begin{cases} u_0 \equiv \frac{1}{\sqrt{2\pi}} \\ \cos_n(x) \equiv \frac{1}{\sqrt{\pi}} \cos(nx), \quad n \in \mathbb{N} \\ \sin_n(x) \equiv \frac{1}{\sqrt{\pi}} \sin(nx), \quad n \in \mathbb{N}, \end{cases}$$

alors l'expansion en série de Fourier réelle pour tout élément $f \in L^2(\mathbb{T})$ est la suivante :

$$\boxed{f(t) \underset{L^2(\mathbb{T})}{=} \frac{a_0}{2} + \sum_{n=1}^{+\infty} a_n \cos(nt) + \sum_{n=1}^{+\infty} b_n \sin(nt)},$$

avec

$$a_0 = \frac{1}{\pi} \int_{\mathbb{T}} f(t)dt \quad \Longrightarrow \quad \frac{a_0}{2} = \frac{1}{2\pi} \int_{\mathbb{T}} f(t)dt = \langle f \rangle_{\mathbb{T}} \quad \text{(moyenne de } f\text{)}$$

$$a_n = \frac{1}{\pi} \int_{\mathbb{T}} f(t) \cos(nt)dt \qquad \forall n = 1, 2, \ldots$$

$$b_n = \frac{1}{\pi} \int_{\mathbb{T}} f(t) \sin(nt)dt \qquad \forall n = 1, 2, \ldots$$

Les coefficients a_0, a_n, b_n, $n = 1, 2, \ldots$ sont dits **coefficients de Fourier réels** de f.

L'égalité doit bien sûr être interprétée dans le sens de $L^2(\mathbb{T})$, i.e.

$$\int_{\mathbb{T}} \left[f(t) - \left(\frac{a_0}{2} + \sum_{n=1}^{N} a_n \cos(nt) + \sum_{n=1}^{N} b_n \sin(nt) \right) \right]^2 dt \underset{N \to +\infty}{\longrightarrow} 0,$$

l'expression

$$\boxed{S_N(t) = \frac{a_0}{2} + \sum_{n=1}^{N} a_n \cos(nt) + \sum_{n=1}^{N} b_n \sin(nt)}$$

est dite **polynôme trigonométrique d'ordre** N. S_N est une fonction 2π-périodique, comme les éléments de $L^2(\mathbb{T})$.

Pour comprendre la présence de la constante $\frac{1}{\pi}$ dans les coefficients de Fourier réels on va considérer, par exemple, l'expansion de f par rapport au système de cosinus :

$$\sum_{n=1}^{+\infty} \langle f, \frac{1}{\sqrt{\pi}} \cos(nt) \rangle \frac{1}{\sqrt{\pi}} \cos(nt) = \sum_{n=1}^{+\infty} \left(\frac{1}{\pi} \int_{\mathbb{T}} f(t) \cos(nt) dt \right) \cos(nt),$$

la même considération vaut pour le système de sinus et pour la constante.

La constante $\frac{1}{\pi}$ est incorporée dans la définition des coefficients de Fourier pour permettre d'identifier $\frac{a_0}{2}$ avec la moyenne de f et donc pour pouvoir *interpréter la série de Fourier réelle comme la superposition entre la moyenne de f et des combinaisons d'ondes harmoniques avec fréquence croissante*, en particulier :

— $t \mapsto a_1 \cos(t) + b_1 \sin(t)$ est dite **harmonique fondamentale** ;

— $t \mapsto a_n \cos(nt) + b_n \sin(nt)$ est dite **harmonique d'ordre** n.

Un *diapason* peut émettre un son « pur », i.e. constitué par la seule harmonique fondamentale, par contre la grande majorité des instruments musicaux produit des sons qui peuvent être décrits par une série de Fourier, i.e. une superposition d'harmoniques avec fréquences multiples d'une fréquente fondamentale.
Grâce au théorème de la projection orthogonale et à l'identité de Plancherel on peut dire que l'erreur quadratique moyenne (i.e. la norme L^2) entre f et le polynôme trigonométrique d'ordre N est :

$$E_N = \int_{\mathbb{T}} [f(t) - S_N(t)]^2 \, dt = \int_{\mathbb{T}} f(t)^2 dt - \pi \left[\frac{a_0}{2} + \sum_{n=1}^{N} \left(a_n^2 + b_n^2 \right) \right],$$

comme $E_N \underset{N \to +\infty}{\longrightarrow} 0$, ça vaut que :

$$\int_{\mathbb{T}} f(t)^2 dt = \pi \left[\frac{a_0}{2} + \sum_{n=1}^{+\infty} \left(a_n^2 + b_n^2 \right) \right],$$

qui est une identité entre une intégrale et une série numérique très utile pour déterminer l'une grâce au calcul de l'autre.

Si on considère $L^2[a,b]$ et on écrit $T = b - a$, et $\omega = \frac{2\pi}{T}$, alors on sait que sa base hilbertienne de Fourier réelle est :

$$\left\{ \frac{1}{\sqrt{T}}, \sqrt{\frac{2}{T}} \cos(\omega nt), \sqrt{\frac{2}{T}} \sin(\omega nt), \ n = 1, 2, 3, \dots \right\},$$

par rapport à cette base hilbertienne, l'expansion en série de Fourier de $f \in L^2([a,b])$, f (T-périodique) est :

$$f(t) \underset{L^2[a,b]}{=} \frac{a_0}{2} + \sum_{n=1}^{+\infty} a_n \cos(\omega nt) + \sum_{n=1}^{+\infty} b_n \sin(\omega nt),$$

avec :

$$\frac{a_0}{2} = \frac{1}{T}\int_a^b f(t)dt = \langle f\rangle_{[a,b]} \quad \text{(moyenne de } f\text{)}$$

$$a_n = \frac{2}{T}\int_a^b f(t)\cos(\omega nt)dt, \;\; b_n = \frac{2}{T}\int_a^b f(t)\sin(\omega nt)dt \qquad \forall n = 1, 2, \ldots$$

Les polynômes de Fourier dans ce cas sont des fonctions T-périodiques.

Exercice 5.2

La famille $(e_k : [-\pi, \pi] \to \mathbb{C})_{k\in\mathbb{Z}}$ des exponentiels non-normalisés $\left(e_k(t) := e^{ikt}\right)_{k\in\mathbb{Z}}$ est une base hilbertienne de $L^2[-\pi, \pi]$ si on muni cet espace du produit scalaire défini par $\langle f, g\rangle_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}dx$.

1. Écrire la série de Fourier associée à la fonction $\phi : [-\pi, \pi] \to \mathbb{C}$, $t \mapsto \cos(3t) - \sin(5t)$.
2. Soit $\mathbb{N}^* = \mathbb{N}\backslash\{0\}$. Soit $(\psi_k : \mathbb{R} \to \mathbb{R})_{k\in\mathbb{N}^*}$ la famille définie par : $\psi_k(t) = \sin(kt)$.
   (a) Soit $f \in L^2[0, \pi]$ telle que $\int_0^\pi f(t)\psi_k(t)dt = 0 \;\forall k \in \mathbb{N}^*$ et soit

   $$g(t) = \begin{cases} f(t) & \text{si } 0 \leqslant t < \pi \\ -f(-t) & \text{si } -\pi < t < 0. \end{cases}$$

   Montrer que $\int_{-\pi}^{\pi} g(t)e^{-ikt}dt = 0 \;\forall k \in \mathbb{N}^*$.

   (b) Montrer que $(\psi_k)_{k\in\mathbb{N}^*}$ est un système complet de $L^2[0, \pi]$ muni du produit scalaire $\langle f, g\rangle = \int_0^\pi f(x)\overline{g(x)}dx$, i.e. une famille, non-orthogonale, d'éléments de $L^2[0, \pi]$ telle que

   $$\overline{\text{span}((\psi_k)_{k\in\mathbb{N}^*})} = L^2[0, \pi] \Leftrightarrow (\text{span}((\psi_k)_{k\in\mathbb{N}^*}))^\perp = \{0_{L^2[0,\pi]}\}$$
   $$\Leftrightarrow \langle f, \psi_k\rangle = 0 \;\forall f \in L^2[0, 2\pi], \forall k \in \mathbb{N}^* \Rightarrow f \equiv 0_{L^2[0,\pi]}.$$

3. Construire une base hilbertienne de $L^2[0, \pi]$ à partir de la famille $(\psi_k)_{k\in\mathbb{N}^*}$.
4. Utiliser le résultat obtenu au point précédent pour déterminer une suite de coefficients réels $(a_k)_{k\in\mathbb{N}^*}$ telle que $\sum_{k=1}^{+\infty} a_k\psi_k = 1$ (égalité dans le sens de $L^2[0, \pi]$).
5. En utilisant l'identité de Plancherel, montrer la validité de la formule suivante :

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

Solution de l'exercice 5.2

1. Nous allons réécrire ϕ comme ça :

   $$\phi(x) = \frac{1}{2}(e^{3it} + e^{-3it}) - \frac{1}{2i}(e^{5it} + e^{-5it}),$$

   i.e. $\phi = \frac{1}{2}(e_3 + e_{-3}) - \frac{1}{2i}(e_5 - e_{-5})$, où l'égalité est dans le sens de $L^2[-\pi, \pi]$, est la série de Fourier de la fonction ϕ par l'unicité de la décomposition.

2. Étudions les deux points séparément.

   (a) Par calcul direct :

$$\int_{-\pi}^{\pi} g(t)e^{-ikt}dt = \int_{-\pi}^{0} -f(-t)e^{-ikt}dt + \int_{0}^{\pi} f(t)e^{-ikt}dt,$$

   si on change la variable dans la première intégrale $s = -t$, $ds = -dt$, on obtient $\int_{-\pi}^{0} -f(-t)e^{-ikt}dt = \int_{\pi}^{0} f(s)e^{iks}ds = \int_{0}^{\pi} -f(s)e^{iks}ds = \int_{0}^{\pi} -f(t)e^{ikt}dt$ et alors

$$\int_{-\pi}^{\pi} g(t)e^{-ikt}dt = \int_{0}^{\pi} -f(t)e^{ikt}dt + \int_{0}^{\pi} f(t)e^{-ikt}dt = \int_{0}^{\pi} f(t)(e^{-ikt} - e^{ikt})dt,$$

   grâce à la formule de Euler pour le sinus, nous avons :

$$\int_{-\pi}^{\pi} g(t)e^{-ikt}dt = -2i\int_{0}^{\pi} f(t)\sin(kt)dt = -2i\int_{0}^{\pi} f(t)\psi_k(t)dt = 0$$

   par définition des fonctions ψ_k.

   (b) D'après le point précédent, $\langle g, e_k\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t)e^{-ikt}dt = 0\ \forall k \in \mathbb{N}^*$, allons vérifier que ça vaut aussi pour $k = 0$ et pour $k \in -\mathbb{N}^*$, ou $-k \in \mathbb{N}^*$:

$$\begin{aligned}\langle g, e_0\rangle = \langle g, e_0\rangle &= \frac{1}{2\pi}\int_{-\pi}^{0}(\cos(-3t) - \sin(-5t))dt = \frac{1}{2\pi}\int_{-\pi}^{0}(\cos(3t) + \sin(5t))dt \\ &= \frac{1}{2\pi}\left(\frac{\sin(3t)}{3}\Big|_{-\pi}^{0} - \frac{\cos(5t)}{3}\Big|_{-\pi}^{0}\right) = 0,\end{aligned}$$

$$\begin{aligned}\langle g, e_{-k}\rangle &= \frac{1}{2\pi}\int_{-\pi}^{0}(\cos(3t) + \sin(5t))\overline{e^{-ikt}}dt + \frac{1}{2\pi}\int_{0}^{\pi}(\cos(3t) - \sin(5t))\overline{e^{-ikt}}dt \\ &= \frac{1}{2\pi}\int_{-\pi}^{0}(\cos(3t) + \sin(5t))e^{ikt}dt + \frac{1}{2\pi}\int_{0}^{\pi}(\cos(3t) - \sin(5t))e^{ikt}dt \\ &= \frac{1}{2\pi}\int_{\pi}^{0} -(\cos(3s) - \sin(5s))e^{-iks}ds + \frac{1}{2\pi}\int_{0}^{-\pi} -(\cos(3s) + \sin(5s))e^{-iks}ds \\ &= \frac{1}{2\pi}\int_{0}^{\pi}(\cos(3t) - \sin(5t))e^{-ikt}dt + \frac{1}{2\pi}\int_{-\pi}^{0}(\cos(3t) + \sin(5t))e^{-ikt}dt \\ &\equiv \langle g, e_k\rangle = 0\ \forall k \in \mathbb{N}^*,\end{aligned}$$

   donc, $\langle g, e_k\rangle = 0\ \forall k \in \mathbb{Z}$ et, comme $(e_k)_{k\in\mathbb{Z}}$ est une base hilbertienne de $L^2[-\pi, \pi]$, $g = 0_{L^2[-\pi,\pi]}$. Mais, par sa définition, $g = 0_{L^2[-\pi,\pi]}$ si et seulement si $f = 0_{L^2[0,\pi]}$, i.e. ce que l'on devait vérifier pour prouver que $(\psi_k)_{k\in\mathbb{N}*}$ est un système complet de $L^2[0, \pi]$.

3. Le fait que $(\psi_k)_{k\in\mathbb{N}*}$ est un système complet de $L^2[0, \pi]$ nous permet de trouver une base hilbertienne du même espace tout simplement en examinant les propriétés

d'orthonormalité de ce système. Pour tout $n, m \in \mathbb{N}^*$:

$$\begin{aligned}
\langle \psi_n, \psi_m \rangle &= \int_0^\pi \sin(nt)\sin(mt)dt \quad (t \mapsto \sin(nt)\sin(mt) \text{ est paire}) \\
&= \frac{1}{2}\int_{-\pi}^{\pi} \sin(nt)\sin(mt)dt \\
&= \frac{1}{2}\int_{-\pi}^{\pi} \frac{e^{int}-e^{-int}}{2i}\,\frac{e^{imt}-e^{-imt}}{2i}dt \\
&= -\frac{1}{8}\int_{-\pi}^{\pi} (e^{int}-e^{-int})(e^{imt}-e^{-imt})dt \\
&= -\frac{1}{8}\int_{-\pi}^{\pi} (e^{int}-e^{-int})(\overline{e^{-imt}-e^{imt}})dt \\
&= -\frac{2\pi}{8}\langle e_n - e_{-n}, e_{-m} - e_m \rangle_0 \\
&= -\frac{\pi}{4}\left(\langle e_n, e_{-m}\rangle_0 - \langle e_n, e_m\rangle_0 - \langle e_{-n}, e_{-m}\rangle_0 + \langle e_{-n}, e_m\rangle_0\right) \\
&= \begin{cases} 0 & \text{si } n \neq m \\ -\frac{\pi}{4}(-1-1) = \frac{\pi}{2} & \text{si } n = m, \end{cases}
\end{aligned}$$

donc $\|\psi_n\| = \sqrt{\frac{\pi}{2}}\ \forall n \in \mathbb{N}^*$ et alors $\left(\sqrt{\frac{2}{\pi}}\psi_n\right)_{n\in\mathbb{N}^*}$ est une base hilbertienne de $L^2[0,\pi]$.

4. Nous allons interpréter 1 comme la fonction constante $\mathbf{1} \in L^2[0,\pi]$, $\mathbf{1}(t) = 1\ \forall t \in [0,\pi]$ et nous allons la décomposer sur la base hilbertienne $\left(\sqrt{\frac{2}{\pi}}\psi_n\right)_{n\in\mathbb{N}^*}$ de $L^2[0,\pi]$ que l'on vient de déterminer :

$$\mathbf{1} = \sum_{k=1}^{+\infty} \langle \mathbf{1}, \sqrt{\frac{2}{\pi}}\psi_k\rangle \sqrt{\frac{2}{\pi}}\psi_k = \sum_{k=1}^{+\infty} \langle \mathbf{1}, \frac{2}{\pi}\psi_k\rangle\, \psi_k,$$

ce qui montre que $1 = \sum\limits_{k=1}^{+\infty} a_k\, \psi_k$, avec

$$a_k = \langle \mathbf{1}, \frac{2}{\pi}\psi_k\rangle = \frac{2}{\pi}\int_0^\pi \sin(kt)dt = \frac{2}{\pi k}\left[-\cos(kt)\right]_0^\pi = \frac{2}{\pi k}\left[1-(-1)^k\right],$$

i.e. la suite cherchée est : $a_k = \begin{cases} 0 & k \text{ pair} \\ \frac{4}{\pi k} & k \text{ impair} \end{cases}$.

5. L'identité de Plancherel pour $\mathbf{1}$ donne :

$$\|\mathbf{1}\|^2 = \sum_{k=1}^{+\infty} \left|\langle \mathbf{1}, \sqrt{\frac{2}{\pi}}\psi_k\rangle\right|^2,$$

mais $\|\mathbf{1}\|^2 = \int_0^\pi 1dt = \pi$ et $\langle \mathbf{1}, \sqrt{\frac{2}{\pi}}\psi_k\rangle = \sqrt{\frac{\pi}{2}}a_k$, d'où

$$\pi = \sum_{k=0}^{+\infty} \frac{\pi}{2}|a_{2k+1}|^2 = \sum_{k=0}^{+\infty} \frac{\pi}{2}\left(\frac{4}{\pi}\right)^2 \frac{1}{(2k+1)^2} \iff \sum_{k=0}^{+\infty} \frac{1}{(2k+1)^2} = \frac{\pi^2}{8}.$$

□

5.6.5 Convergence ponctuelle de la série de Fourier réelle : le théorème de Dirichlet

La communauté mathématique a initialement été très sceptique vis-à-vis de la série de Fourier : il était considéré absurde qu'une série dont les termes sont des fonctions trigonométriques, et donc infiniment dérivables, pouvait approcher des fonctions non dérivables ou, pire encore, non continues. De plus, Fourier n'a pas pu prouver des résultats de convergence pour sa série.

En fait, les résultats de convergence en norme qu'on a discuté dans les sections antécédentes sont venu après et, en tout cas, ils ne sont pas suffisantes pour garantir la convergence ponctuelle de la série. Les premières conditions de ce type ont été découvertes par le mathématicien allemand Dirichlet [6] (Düren 1805 − Göttingen 1859) en 1829. La démonstration constructive de Dirichlet a une grande importance dans l'analyse de Fourier, le lecteur intéressé peut consulter [14].

Ici nous nous limitons à citer d'une manière rigoureuse le théorème de Dirichlet. Pour cela, on va introduire une notation et deux définitions. La notation est la suivante : si t_0 est un point de discontinuité d'une fonction f réelle de variable réelle, alors on écrit les limites à droite et à gauche comme ceci :

$$f(t_0^+) = \lim_{t \to t_0^+} f(t), \quad f(t_0^-) = \lim_{t \to t_0^-} f(t).$$

Déf. 5.6.3 (Fonction de Dirichlet) *Soit $f : \mathbb{R} \to \mathbb{R}$. On dit que f est une fonction de Dirichlet si elle vérifie les trois conditions suivantes :*

1. *f est T-périodique, $T \in \mathbb{R}^+$;*
2. *f est continue par morceaux, i.e. il existe seulement un nombre fini de points dans lesquels f n'est pas continue ;*
3. *Pour tout $t_0 \in \mathbb{R}$,*

$$f(t_0) = \frac{f(t_0^+) + f(t_0^-)}{2}, \tag{5.7}$$

i.e. en tout point $t_0 \in \mathbb{R}$, la valeur de f en t_0 est la moyenne de la limite à droite et de la limite à gauche de f en t_0.

La condition (5.7) dit que la valeur en t_0 est la moyenne de la limite à droite et de la limite à gauche de f en t_0. Cette condition est trivialement vérifiée en tout point où f est continue. C'est seulement aux éventuels points de discontinuité de f que cette condition est non triviale.

Déf. 5.6.4 (Dérivée généralisée) *Soit f une fonction de Dirichlet et $t_0 \in \mathbb{R}$. On dit que f possède une dérivée généralisée à droite en t_0 s'il existe (finie) la limite :*

$$\lim_{h \to 0^+} \frac{f(t_0 + h) - f(t_0^+)}{h},$$

de la même manière, on dit que f possède une dérivée généralisée à gauche en t_0 si'il existe (finie) la limite :

$$\lim_{h \to 0^-} \frac{f(t_0 + h) - f(t_0^-)}{h}.$$

6. C'est dans l'effort de démontrer la convergence ponctuelle de la série de Fourier que Dirichlet a donné la définition « moderne » de fonction comme correspondance univoque entre deux ensembles.

On peut finalement énoncer le théorème de Dirichlet.

Théorème 5.6.1 (Théorème de Dirichlet (1829)) *Soit f une fonction de Dirichlet et $t_0 \in \mathbb{R}$. Si la fonction f possède des dérivées généralisées à droite et à gauche au point t_0, alors la série de Fourier réelle évaluée en t_0 converge vers $f(t_0)$.*

Les conditions du théorème sont dites « conditions de Dirichlet » et sont seulement suffisantes mais pas nécessaires pour la convergence ponctuelle de la série de Fourier réelle. Trouver des conditions nécessaires et suffisantes pour la convergence ponctuelle de la série de Fourier reste un problème toujours ouvert. Néanmoins, heureusement, les conditions de Dirichlet sont vérifiées par la grande majorité des fonctions qu'on rencontre dans les applications pratiques.

Il faut souligner que, si on élimine la requête (5.7), alors la série de Fourier converge vers $f(t_0) = \frac{f(t_0^+)+f(t_0^-)}{2}$.

Une observation finale relativement aux possibles conséquences du manque de continuité de f : le grand mathématicien russe Kolmogorov (Tambov 1903 − Moscow 1987) a construit en 1923 une fonction avec de discontinuités pathologiques qui font que sa série de Fourier est divergente en chaque point !

5.6.6 Le phénomène de Gibbs et les sommes de Cesàro

L'énoncé du théorème de Dirichlet ne doit pas faire penser que le comportement de la série de Fourier dans un voisinage d'une discontinuité d'une fonction soit « régulière ». En fait, à chaque fois qu'on approche la discontinuité d'une fonction, des oscillations, dite de **Gibbs**, apparaissent. Même en augmentant la quantité de coefficients de Fourier les oscillations restent. Si une fonction f est de Dirichlet, alors les oscillations à gauche et à droite de la discontinuité se compensent et leur moyenne coïncide avec la valeur de f dans la discontinuité.

On peut démontrer que l'écart entre la valeur de la fonction f et la valeur du polynôme trigonométrique S_N dans un voisinage proche comme l'on veut d'une discontinuité reste proche du 18% même quand $N \to +\infty$. L'étude du phénomène de Gibbs a des subtilités mathématiques non-triviales et il est hors du but de ce cours. Pour connaître les détails du phénomène de Gibbs un excellente référence est [14].

Nous nous limitons à montrer, dans la figure 5.2, l'effet de Gibbs pour la fonction créneau.

Il y a la possibilité d'éliminer les oscillations de Gibbs si on considère, au lieu de la somme usuelle, une sommation dite de Cesàro (Naples, 1859 − Torre Annunziata, 1906), dans laquelle on opère de moyennes des sommes partielles qui 'lissent' les oscillations de Gibbs.

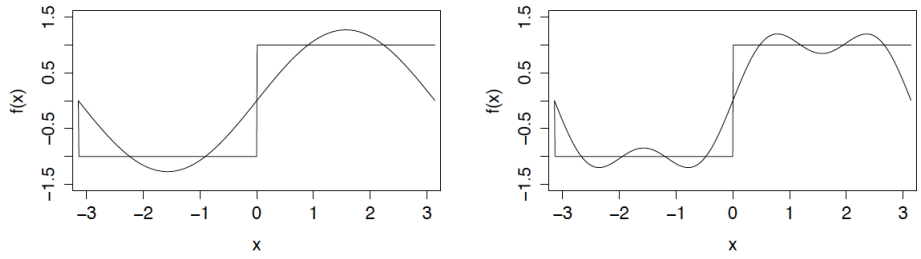
5.6.7 La vitesse de convergence à 0 des coefficients de Fourier

On commence avec un résultat général.

Lemme 5.6.1 (Riemann-Lebesgue) *Soit $f \in L^1[a,b]$, alors :*

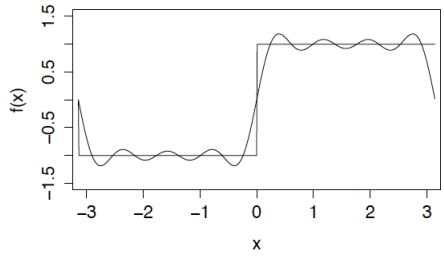
$$\lim_{n\to+\infty} \int_a^b f(t)\cos(nt)dt = \lim_{n\to+\infty} \int_a^b f(t)\sin(nt)dt = \lim_{n\to+\infty} \int_a^b f(t)e^{int}dt = 0.$$

L'interprétation géométrique du lemme de Riemann-Lebesgue est que la fonction $f(t)\cos(nt)$ ou $f(t)\sin(nt)$ oscille avec une fréquence tellement élevée, quand $n \to +\infty$, que les valeurs autour de la moyenne se compensent et donc l'intégrale converge vers 0.

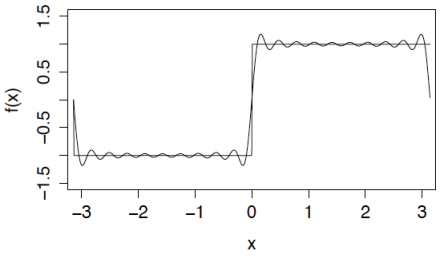


(a) $n = 1$.

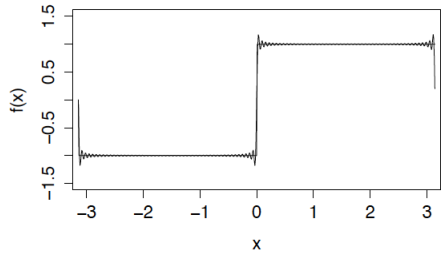
(b) $n = 2$.



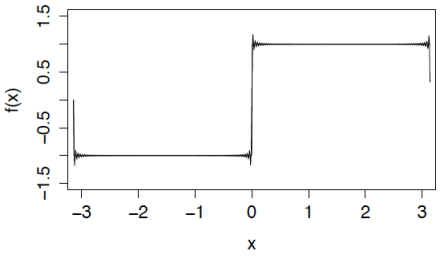
(c) $n = 4$.



(d) $n = 10$.



(e) $n = 50$.



(f) $n = 80$.

FIGURE 5.2 – Le phénomène de Gibbs pour la fonction créneau (courtoisie de Eric Luçon).

Un corollaire immédiat du lemme est que **les coefficients de Fourier de la série de Fourier d'une fonction** $f \in L^1[a,b]$ (et $(b-a)$-périodique, bien sûr), **décroissent vers 0** quand $n \to +\infty$.

Le théorème suivant montre que la régularité de f a un impact important sur la vitesse de décroissance des coefficients de Fourier.

Théorème 5.6.2 *Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction :*

— *de classe $\mathcal{C}^p([a,b])$, i.e. f dérivable p fois sur $[a,b]$ avec p dérivées continues ;*

— *$(b-a)$-périodique ;*

— *qui possède des dérivées généralisées égales dans les extrêmes de l'intervalle $[a,b]$.*

Alors, les coefficients de Fourier de f, a_n, b_n, $n = 1, 2, \ldots$ vérifient :

$$a_n, b_n = o\left(\frac{1}{n^p}\right),$$

i.e., ils décroissent vers 0 plus rapidement que $\frac{1}{n^p}$.

Ce résultat est très important car nous dit que si f est « bien lisse », alors on peut l'approximer avec précision même en considérant très peu de coefficients de sa série de Fourier.

Par contre, si f n'est pas trop lisse, alors la convergence vers 0 des coefficients de Fourier de f est lente et il faut considérer une grande quantité de coefficients de Fourier pour avoir une bonne approximation de f.

Le vice-versa est vrai aussi sous des hypothèses opportunes qu'on n'a pas le temps de discuter. Le concept le plus important à retenir c'est que plus la convergence vers 0 des coefficients de Fourier d'une fonction est rapide, plus la fonction est lisse.

Preuve.

On considère seulement les coefficients a_n, la preuve est identique pour les coefficients b_n. De plus, par simplicité, on développe la preuve dans le cas $b = \pi$, $a = -\pi$, car on peut toujours se ramener à ce cas avec la transformation linéaire de variable suivante :

$$\boxed{s(t) = \frac{b+a}{2} + \frac{b-a}{2\pi}t},$$

en fait $s(\pi) = b$ et $s(-\pi) = a$.

Avec cette convention, on intègre par parties l'expression des a_n, $n = 0, 1, 2, \ldots$, avec $u = f(t)$ et $dv = \cos(\omega nt)dt$, donc $du = f'(t)dt$ et $v = \frac{1}{\omega n}\sin(\omega nt)$.

On obtient :

$$a_n = \frac{1}{\pi n}\left[f(t)\sin(nt)\right]_{-\pi}^{\pi} - \frac{1}{\pi n}\int_{-\pi}^{\pi} f'(t)\sin(nt)dt = \frac{1}{\pi n}\int_{-\pi}^{\pi} f'(t)\cos\left(\frac{\pi}{2} + nt\right)dt,$$

car $\sin(n\pi) = \sin(-n\pi) = 0$ et $\cos\left(\frac{\pi}{2} + \alpha\right) = -\sin(\alpha)\ \forall \alpha \in \mathbb{R}$.

En procédant encore une fois à l'intégration par parties on obtient :

$$a_n = \frac{1}{\pi n}\left\{\frac{1}{n}\left[f'(t)\sin\left(\frac{\pi}{2} + nt\right)\right]_{-\pi}^{\pi} - \frac{1}{n}\int_{-\pi}^{\pi} f''(t)\sin\left(\frac{\pi}{2} + nt\right)dt\right\},$$

mais le premier terme dans les accolades est nul car, par hypothèse, $f'(-\pi) = f'(\pi)$, donc

$$\begin{aligned}\left[f'(t)\sin\left(\frac{\pi}{2} + nt\right)\right]_{-\pi}^{\pi} &= f'(\pi)\sin\left(\frac{\pi}{2} + n\pi\right) - f'(-\pi)\sin\left(\frac{\pi}{2} - n\pi\right)\\ &= f'(\pi)\left[\sin\left(\frac{\pi}{2} + n\pi\right) - \sin\left(\frac{\pi}{2} - n\pi\right)\right]\\ &= f'(\pi)\left[\sin\left(\frac{\pi}{2} + n\pi\right) - \sin\left(\frac{\pi}{2} - n + 2n\pi\right)\right]\\ &= f'(\pi)\left[\sin\left(\frac{\pi}{2} + n\pi\right) - \sin\left(\frac{\pi}{2} + n\pi\right)\right] = 0.\end{aligned}$$

Par ailleurs, on peut réécrire le deuxième terme dans les accolades comme ceci :

$$-\frac{1}{n}\int_{-\pi}^{\pi} f''(t)\sin\left(\frac{\pi}{2} + nt\right)dt = \frac{1}{n}\int_{-\pi}^{\pi} f''(t)\cos\left(\frac{\pi}{2} + \frac{\pi}{2} + nt\right)dt = \frac{1}{n}\int_{-\pi}^{\pi} f''(t)\cos\left(\frac{\pi}{2}\cdot 2 + nt\right)dt,$$

or,

$$a_n = \frac{1}{\pi n^2}\int_{-\pi}^{\pi} f''(t)\cos\left(\frac{\pi}{2}\cdot 2 + nt\right)dt.$$

En résumé, une intégration par parties de a_n amène à l'expression

$$a_n = \frac{1}{\pi n}\int_{-\pi}^{\pi} f'(t)\cos\left(\frac{\pi}{2}+nt\right)dt,$$

deux intégrations par parties de a_n amènent à l'expression

$$a_n = \frac{1}{\pi n^2}\int_{-\pi}^{\pi} f''(t)\cos\left(\frac{\pi}{2}\cdot 2+nt\right)dt,$$

si on intègre de a_n p fois par parties, on arrive à l'expression

$$a_n = \frac{1}{\pi n^p}\int_{-\pi}^{\pi} f^{(p)}(t)\cos\left(\frac{\pi}{2}\cdot p+nt\right)dt.$$

D'une façon analogue, on trouve

$$b_n = \frac{1}{\pi n^p}\int_{-\pi}^{\pi} f^{(p)}(t)\sin\left(\frac{\pi}{2}\cdot p+nt\right)dt.$$

On observe maintenant que, grâce aux identités trigonométriques $\cos\left(\frac{\pi}{2}+\alpha\right) = -\sin(\alpha)$ et $\sin\left(\frac{\pi}{2}+\alpha\right) = \cos(\alpha)$, $\forall \alpha \in \mathbb{R}$, les intégrales

$$\varepsilon_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f^{(p)}(t)\cos\left(\frac{\pi}{2}\cdot p+nt\right)dt,\ \ \tilde{\varepsilon}_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f^{(p)}(t)\sin\left(\frac{\pi}{2}\cdot p+nt\right)dt$$

sont, par définition, les coefficients de Fourier de la fonction $f^{(p)}$ au signe près. Par hypothèse $f^{(p)}$ est continue sur $[-\pi, pi]$ et donc, comme le domaine $[-\pi,\pi]$ est compact, $f^{(p)} \in L^2[-\pi,\pi]$ et donc, grâce au lemme de Riemann-Lebesgue, ses coefficients de Fourier convergent vers 0 quand $n \to +\infty$, or $\varepsilon_n \underset{n\to\infty}{\longrightarrow} 0$ et $\tilde{\varepsilon}_n \underset{n\to\infty}{\longrightarrow} 0$, mais alors

$$a_n = \frac{\varepsilon_n}{n^p} \underset{n\to\infty}{\longrightarrow} 0,\ \ b_n = \frac{\tilde{\varepsilon}_n}{n^p} \underset{n\to\infty}{\longrightarrow} 0,$$

i.e. $a_n, b_n = o\left(\frac{1}{n^p}\right)$. □

Ce résultat a été utilisé par A. Krylov (1863-1945) qui l'a mis à la base de sa méthode d'amélioration de la convergence des séries de Fourier pour des fonctions avec de discontinuités.

5.6.8 Transformée de Fourier en $L^2(\mathbb{T})$ et translation

Allons analyser maintenant la relation entre translation et transformée de Fourier pour une fonction $f \in L^2(\mathbb{T})$. Le résultat sera qualitativement identique à ce qu'on a déjà vu dans le cas de la transformée de Fourier discrète dans la section 2.7.2.

Théorème 5.6.3 (Transformée de Fourier et translation) *Soit $f \in L^2(\mathbb{T})$, alors :*

1. *Si $g_a(x) = f(x-a)$, $a \in \mathbb{R}$, alors : $\hat{g}_a(n) = e^{-ina}\hat{f}(n)$, $\forall n \in \mathbb{Z}$*

2. *Si $g_k(x) = e^{ikx}f(x)$, $k \in \mathbb{Z}$, alors : $\hat{g}_k(n) = \hat{f}(n-k)$, $\forall n \in \mathbb{Z}$.*

Preuve. On va faire seulement la preuve pour **1.**, car la preuve pour **2.** est analogue. Il s'agit d'un calcul direct qui utilise l'invariance de la mesure de Lebesgue par translations :

$$\begin{aligned}\hat{g_a}(n) = \langle g_a, u_n \rangle &= \int_0^{2\pi} \frac{f(x-a)}{\sqrt{2\pi}} e^{-inx} dx \\ &= \int_{-a}^{2\pi-a} \frac{f(x)}{\sqrt{2\pi}} e^{-in(x+a)} d(x+a) = e^{-ina} \int_0^{2\pi} \frac{f(x)}{\sqrt{2\pi}} e^{-inx} dx \\ &= e^{-ina} \widehat{f}(n).\end{aligned}$$

□

Déf. 5.6.5 *L'ensemble* $\left\{ |\hat{f}(n)|,\ n \in \mathbb{Z} \right\}$ *est dit le spectre (d'amplitude) de* $f \in L^2(\mathbb{T})$.

$|\hat{f}(n)|$ représente l'importance de l'harmonique de fréquence n, i.e. e^{inx}, dans la reconstruction de f, ce qui peut être compris grâce à la formule : $f = \sum_{n\in\mathbb{Z}} \hat{f}(n) \frac{e^{inx}}{\sqrt{2\pi}}$.

La propriété qu'on vient de prouver montre que le spectre de f donne des informations sur la présence de certaines fréquences en f, mais pas sur leur « position » ! En fait, le signal translaté $g_a(x) = f(x-a)$ a le même spectre que f, car $|\widehat{g_a}(n)| = |\widehat{f}(n)|$.

La théorie des « ondelettes » permet d'obtenir des informations localisées par rapport à leur fréquence et à leur position.

5.7 Résumé du chapitre 5

- Nous avons étendu les propriétés des espaces vectoriels avec produit scalaire de dimension finie aux espaces de Hilbert.

- Un rôle important dans cette extension est joué par le complément orthogonale à un sous-ensemble ou sous-espace vectoriel d'un espace de Hilbert.

- Le résultat clé pour l'extension de la structure géométrique des espaces Euclidiens de dimension finie à la dimension infinie est le théorème de projection sur un sous-ensemble convexe fermé d'un espace de Hilbert. Dans la preuve de ce théorème on utilise d'une manière essentielle la formule du parallélogramme, ce qui implique la nécessité d'avoir une norme hilbertienne, d'ici l'exigence d'un espace de Hilbert pour la validité du théorème.

- Quand le convexe fermé du théorème ci-dessus est aussi un sous-espace vectoriel, alors la différence entre le vecteur projeté et la projection appartient au complément orthogonal du sous-espace, comme en dimension finie, et ceci permet d'étendre le théorème de projection orthogonale aux espaces de Hilbert de dimension infinie.

- Le théorème de la projection orthogonale est utilisé pour donner une caractérisation très utile des sous-espaces vectoriels fermés des espaces de Hilbert : ils sont ceux qui coïncident avec leur complément bi-orthogonal.

- Nous avons examiné les systèmes orthonormés (s.o.n) dans les espaces de Hilbert séparables, i.e. ceux qui possèdent au moins un sous-ensemble dénombrable et dense. Un s.o.n d'un espace de Hilbert séparable est dénombrable. Tous les espaces de Hibert qu'on examinera seront supposés séparables sans devoir le spécifier.

- Si on veut qu'un s.o.n $(u_n)_{n\in\mathbb{N}}$ soit la généralisation d'une base orthonormée pour un espaces de Hilbert $\mathcal{H}$ de dimension infinie, alors on doit garantir que, pour tout $x \in \mathcal{H}$, la suite des coefficients de Fourier $(\hat{x}(n) = \langle x, u_n\rangle)_{n\in\mathbb{N}}$ soit décroissante vers 0, sinon l'expansion $\sum_{n\in\mathbb{N}} \langle x, u_n\rangle u_n$ ne pourra pas être convergente. L'inégalité de Bessel assure cela, comme conséquence du fait que la suite des coefficients de Fourier relatifs à un s.o.n. quelconque d'un espace de Hilbert appartient à ℓ^2. Néanmoins, l'inégalité de Bessel dit que l'identité de Plancherel n'est pas forcement vérifiée pour un s.o.n quelconque, car, en général, ça vaut que $\sum_{n\in\mathbb{N}} |\langle x, u_n\rangle|^2 u_n \leqslant \|x\|^2$.

- Le théorème de Fischer-Riesz dit que ça vaut l'identité de Plancherel quand la série $\sum_{n\in\mathbb{N}} \langle x, u_n\rangle u_n$ converge vers x et on a vu, avec un contre-exemple, que ceci n'est pas assuré pour un s.o.n quelconque. La condition qui assure la validité de l'expansion sur $\sum_{n\in\mathbb{N}} \langle x, u_n\rangle u_n$ est que le s.o.n $(u_n)_{n\in\mathbb{N}}$ soit complet, i.e. qu'il ne soit pas une partie propre d'un autre s.o.n de $\mathcal{H}$. Les s.o.n complets s'écrivent s.o.n.c ou bases hilbertiennes.

- Il est possible de caracteriser une base hilbertienne $(u_n)_{n\in\mathbb{N}}$ avec cinq conditions équivalentes : le fait que le seul vecteur orthogonale à tous les éléments d'une base hilbertienne est le vecteur nul ; le fait que le sous-espace engendré par la base hilbertienne soit dense en $\mathcal{H}$; la possibilité d'opérer l'expansion en série de Fourier généralisée ; l'identité de Parseval et celle de Plancherel.

- Un isomorphisme d'espaces de Hilbert est une transformation surjective et qui préserve le produit scalaire. On a vu que préserver les produits scalaires implique l'isométrie et donc l'injectivité et le mélange entre surjectivité et préservation du produit scalaire implique la linéarité. Tous les espaces de Hilbert séparables sur le même corps sont isomorphes entre eux, en particulier, le prototype d'un espace de Hilbert séparable de dimension infinie sur le corps $\mathbb{K}$ est $\ell^2(\mathbb{N},\mathbb{K})$. Ce résultat est l'extension à la dimension infinie du fait que $\mathbb{K}^n$ est le prototype de tous les espaces vectoriels de dimension finie n sur $\mathbb{K}$.

- La série et la transformée de Fourier usuelle sur les espaces $L^2([a,b])$ sont définies comme un cas particulier de la théorie qu'on a développé avant, la particularité étant le choix d'une base hilbertienne donné par des exponentiels complexes, ou par un système de cosinus et sinus (plus une fonction constante). Ceci vaut aussi pour les fonctions définies sur tout $\mathbb{R}$ mais périodiques.

- Comme pour les suites de $\ell^2(\mathbb{Z}_N)$, aussi pour les fonctions de L^2 ci-dessus, le spectre de Fourier (l'ensemble des modules des coefficients de Fourier) est invariant par translation, ce qui amène à l'exigence de développer une théorie fréquentielle localisée. À cette exigence répond la théorie des ondelettes.

Chapitre 6

Les opérateurs linéaires bornés dans les espaces de Hilbert

Une fonction $A : V \to W$, V et W espaces vectoriels normés sur le même corps $\mathbb{K}$, est dite **opérateur linéaire** entre V et W si

$$\forall \alpha, \beta \in \mathbb{K},\ A(\alpha x + \beta y) = \alpha A(x) + \beta A(y), \qquad \forall x, y \in V,$$

pour simplifier la notation, dans la suite parfois on pourra omettre la parenthèse et écrire simplement Ax au lieu de $A(x)$. V est le *domaine* de A et l'ensemble

$$\mathrm{Im}(A) = A(V) = \{y \in W\ :\ \exists x \in V\ :\ y = Ax\} \subseteq W$$

est le **codomaine** ou l'**image** de A, W est l'ensemble d'arrivé de A.

Allons voir des exemples basiques :

1. L'**opérateur identité** : $id : V \to V$, $id(x) = x\ \forall x \in V$ et l'**opérateur nul** : $\mathbf{0} : V \to V$, $\mathbf{0}(x) = 0_V\ \forall x \in V$.

2. L'**opérateur de dérivation** : il est défini sur un espace de fonctions différentiables qui peut changer selon l'application à laquelle nous sommes intéressés. Pour avoir un exemple concret considérons l'opérateur de dérivation première : $D_1 f(t) = \frac{df}{dt}(t) = f'(t)$, un domaine bien défini pour D_1 pourrait être $\mathrm{dom}(D_1) = \{f \in L^2[a,b] \cap \mathcal{C}^1[a,b]\ :\ f' \in L^2[a,b]\}$, $a < b$ constantes réelles, alors :

$$\begin{array}{rccc} D_1 : & \mathrm{dom}(D_1) \subset L^2[a,b] & \longrightarrow & L^2[a,b] \\ & f & \longmapsto & D_1 f. \end{array}$$

   De même, on peut définir l'opérateur

$$D_n f(t) = \frac{d^n f}{dt^n}(t) = f^{(n)}(t)$$

   sur le domaine $\mathrm{dom}(D_n) = \{f \in L^2[a,b] \cap \mathcal{C}^1[a,b]\ :\ f^{(n)} \in L^2[a,b]\}$, $a < b$ constantes réelles, i.e. :

$$\begin{array}{rccc} D_n : & \mathrm{dom}(D_n) \subset L^2[a,b] & \longrightarrow & L^2[a,b] \\ & f & \longmapsto & D_n f. \end{array}$$

   Les opérateurs de dérivation partielle sont définis d'une manière analogue.

3. L'**opérateur d'intégration** : il est typiquement défini on considérant une fonction, dite « kernel » (noyau, en français), $k(s,t)$, $k \in L^2([a,b] \times [a,b])$, $a < b$ constantes réelles, alors l'opérateur d'intégration avec kernel k est

$$\begin{array}{rccl} T_k: & L^2[a,b] & \longrightarrow & L^2[a,b] \\ & f & \longmapsto & T_k f, \ \text{où } T_k f(s) = \int_a^b k(s,t) f(t) dt. \end{array}$$

4. Les **opérateurs linéaires en dimension finie**. Soit $A : \mathbb{K}^n \to \mathbb{K}^n$ un opérateur linéaire et soit $(u_1, \ldots, u_n)$ une base orthonormée de $\mathbb{K}^n$. Tout $x \in \mathbb{K}^n$ peut être écrit comme $x = \sum_{j=1}^{n} \lambda_j u_j$, avec $\lambda_j \in \mathbb{K}$ $\forall j$ et, par linéarité, $Ax = \sum_{j=1}^{n} \lambda_j A u_j$. Donc :

$$\langle Ax, u_j \rangle = \sum_{j=1}^{n} \lambda_j \langle Au_j, u_i \rangle = \sum_{j=1}^{n} \alpha_{ij} \lambda_j, \qquad \forall i = 1, \ldots, n, \tag{6.1}$$

où : $\alpha_{ij} = \langle Au_j, u_i \rangle$. Ceci montre que l'action de A est complètement déterminée par la matrice d'élément $(\alpha_{ij})_{i,j=1,\ldots,n}$ et, vice-versa, pour toute matrice d'éléments $(\alpha_{ij})_{i,j=1,\ldots,n}$, la formule (6.1) définit un opérateur linéaire sur $\mathbb{K}^n$.

Ce que l'on vient de rappeler dans le dernier exemple est la célèbre relation entre opérateurs linéaires sur $\mathbb{K}^n$ et matrices $n \times n$ avec éléments en $\mathbb{K}$. Vu que $\mathbb{K}^n$ est le prototype de tous les espaces vectoriels V de dimension n sur $\mathbb{K}$, on peut dire que la théorie des opérateurs linéaires sur les espaces vectoriels de dimension finie est, essentiellement, une théorie de matrices.

On verra que pour les opérateurs linéaires dits « continus » on peut encore exprimer leur action via une matrice, mais cette fois-ci, la matrice a une nombre infini (dénombrable) de lignes et de colonnes.

L'examen de la continuité d'un opérateur linéaire définit entre deux espaces vectoriels normés V et W est motivé à cause de la présence d'une topologie engendrée par la norme. Si V et W ont une dimension finie, le problème a une solution très simple : tout opérateur linéaire défini sur un espace vectoriel normé de dimension finie est continu.

Néanmoins, on verra dans la section 6.2.1 que si V a une dimension infinie, alors un opérateur linéaire, même très simple, peut ne pas être continu.

Dans les sections suivantes on examinera les principales propriétés des opérateurs linéaires. En particulier, on montrera que, pour eux, la propriété d'être continu et celle d'être borné sont équivalentes.

6.1 Les propriétés fondamentales des opérateurs linéaires bornés entre espaces vectoriels normés

Allons commencer par introduire les définitions formelles d'opérateur continu et borné. $(V, \| \ \|_V)$ et $(W, \| \ \|_W)$ seront deux espaces vectoriels normés fixés.

Déf. 6.1.1 *Soit $A : V \to W$ un opérateur linéaire.*

— *A **est continu en** $x_0 \in V$ si*

$$\forall \varepsilon > 0 \ \exists \delta_\varepsilon > 0 \ : \ \|x - x_0\|_V < \delta_\varepsilon \implies \|Ax - Ax_0\|_W = \|A(x - x_0)\|_W < \varepsilon.$$

— *A **est continu sur** V si A est continu en tout élément de V.*

— *A **est borné** si $\exists c \in \mathbb{R}$, $c \geqslant 0$, tel que :*

$$\|Ax\|_W \leqslant c\|x\|_V \qquad \forall x \in V,$$

i.e., tout vecteur $x \in V$ est transformé par A dans un vecteur Ax dont la norme en W est majorée par un multiple positif de la norme de x en V.

Comme pour les fonctions définies sur les espaces métriques, on peut caractériser la continuité d'un opérateur linéaire à travers de son interaction avec les suites.

Théorème 6.1.1 *L'opérateur linéaire $A : V \to W$ est continu en $x_0 \in V$ si et seulement si*

$$\forall (x_n)_{n\in\mathbb{N}} \subset V, \ x_n \underset{n\to+\infty}{\longrightarrow} x_0 \implies Ax_n \underset{n\to+\infty}{\longrightarrow} Ax_0,$$

i.e. :

$$\forall (x_n)_{n\in\mathbb{N}} \subset V, \ \|x_n - x_0\|_V \underset{n\to+\infty}{\longrightarrow} 0 \implies \|Ax_n - Ax_0\|_W \underset{n\to+\infty}{\longrightarrow} 0.$$

Avant d'analyser en détail les propriétés des opérateurs linéaires continus, allons montrer que **tout opérateur linéaire continu sur un espace de Hilbert peut être représenté par une matrice infinie** : on va utiliser le même argument de l'exemple 4. vu précédemment, i.e. soit $\mathcal{H}$ un espace de Hilbert, $A : \mathcal{H} \to \mathcal{H}$ un opérateur linéaire continu et $(u_n)_{n\in\mathbb{N}}$ une base hilbertienne de $\mathcal{H}$. Alors, pour tout $x \in \mathcal{H}$, $x = \sum\limits_{n\in\mathbb{N}} \langle x, u_n\rangle u_n$ et, grâce à la continuité et à la linéarité de A :

$$Ax = A\left(\sum_{n\in\mathbb{N}} \langle x, u_n\rangle u_n\right) \underset{\text{(continuité)}}{=} \sum_{n\in\mathbb{N}} A(\langle x, u_n\rangle u_n) \underset{\text{(linéarité)}}{=} \sum_{n\in\mathbb{N}} \langle x, u_n\rangle Au_n.$$

De plus, grâce à la continuité du produit scalaire :

$$\langle Ax, u_m\rangle = \langle \sum_{n\in\mathbb{N}} \langle x, u_n\rangle Au_n, u_m\rangle = \sum_{n\in\mathbb{N}} \langle Au_n, u_m\rangle\langle x, u_m\rangle = \sum_{n\in\mathbb{N}} \alpha_{nm}\langle x, u_m\rangle, \qquad \forall m \in \mathbb{N},$$

où $\alpha_{nm} = \langle Au_n, u_m\rangle$, donc la matrice infinie d'éléments $(\alpha_{mn})_{n,m\in\mathbb{N}}$ est la représentation de l'opérateur linéaire continu A par rapport à la base hilbertienne $(u_n)_{n\in\mathbb{N}}$.

Différemment que dans la dimension finie, il n'est pas simple de savoir quand une matrice infinie correspond à un opérateur linéaire borné, c'est à cause de cette difficulté que les matrices infinies ne sont pratiquement jamais utilisées pour l'étude des opérateurs linéaires dans les espaces de Hilbert.

Le théorème suivant permet de simplifier remarquablement l'analyse de la continuité des opérateurs linéaires.

Théorème 6.1.2 *Soit $A : V \to W$ un opérateur linéaire et $x_0 \in V$ un élément fixé quelconque. Alors, A est continu en x_0 si et seulement si A est continu sur tout V.*

Ce théorème garantit qu'il suffit de prouver la continuité d'un opérateur linéaire dans un seul point (choisi arbitrairement) pour garantir la continuité sur tout l'espace vectoriel de définition.

Preuve.
$\boxed{\Leftarrow}$: trivial, car si A est continu sur V, alors, par définition, A est continu dans tout point de V !

$\boxed{\Rightarrow}$: soit A continu en x_0. Pour démontrer que A est continu en V on doit prouver que la continuité de A en x_0 implique sa continuité en tout élément arbitraire $x \in V$. Soit $(x_n)_{n\in\mathbb{N}} \subset V$, $x_n \underset{n\to+\infty}{\to} x$, on doit démontrer que ceci implique $\|A(x_n) - A(x)\| \underset{n\to+\infty}{\to} 0$.

Observons que la suite $(x_n - x + x_0)_{n\in\mathbb{N}}$ converge vers x_0 car $(x_n)_{n\in\ \mathbb{N}}$ converge vers x, donc grâce à la continuité de A en x_0, ça vaut $A(x_n - x + x_0) \underset{n\to+\infty}{\to} A(x_0)$, c'est-à-dire, $\|A(x_n - x + x_0) - A(x_0)\| \underset{n\to+\infty}{\to} 0$ mais, grâce à la linéarité de A, $\|A(x_n - x + x_0) - A(x_0)\| = \|A(x_n) - A(x) + \cancel{A(x_0)} - \cancel{A(x_0)}\| = \|A(x_n) - A(x)\| \underset{n\to+\infty}{\to} 0$. □

Donc, pour vérifier la continuité [1] (ou...le manque de continuité !) d'un opérateur linéaire $A : V \to W$ il suffit de la vérifier en un point quelconque de V. Ce point est souvent 0_V, le vecteur nul de V, car ce choix permet de simplifier les calculs dans beaucoup de cas.

Allons utiliser tout de suite cette propriété dans la preuve d'un théorème qui montre la relation entre opérateurs linéaires continus et bornés.

Théorème 6.1.3 *Un opérateur linéaire $A : V \to W$ est borné si et seulement s'il est continu.*

Preuve.
$\underline{A \text{ borné} \Rightarrow A \text{ continu} \Leftrightarrow A \text{ continu en } 0_V}$. Soit $(x_n)_{n\in\mathbb{N}} \subset V$, $x_n \underset{n\to+\infty}{\to} 0_V$, i.e. $\|x_n\|_V \underset{n\to+\infty}{\to} 0$, mais A est supposé être borné, donc $\exists c \in \mathbb{R}^+$ tel que :

$$\|Ax_n - A0_V\|_W \underset{A0_V=0_W}{=} \|Ax_n\|_W \leqslant c\|x_n\|_V \underset{n\to+\infty}{\to} 0,$$

donc on a démontré que pour toute suite $(x_n)_{n\in\mathbb{N}} \subset V$, $x_n \underset{n\to+\infty}{\to} 0_V$, $Ax_n \underset{n\to+\infty}{\to} A(0_V)$, ce qui correspond à la continuité de A en 0_V, et donc sur tout V grâce au théorème précédent.

$\underline{A \text{ continu} \Leftrightarrow A \text{ continu en } 0_V \Rightarrow A \text{ borné}}$. Ici c'est utile de considérer la définition originale de continuité et de l'expliciter en $x_0 = 0_V$:

$$\forall \varepsilon > 0\ \exists \delta_\varepsilon > 0\ :\ \|x - 0_V\|_V < \delta_\varepsilon \Rightarrow \|Ax - A0_V\|_W < \varepsilon,$$

i.e.

$$\forall \varepsilon > 0\ \exists \delta_\varepsilon > 0\ :\ \|x\|_V < \delta_\varepsilon \Rightarrow \|Ax\|_W < \varepsilon.$$

1. On rappelle que, pour un opérateur linéaire, la continuité et la **continuité uniforme** sont deux conditions équivalentes.

Comme l'expression ci-dessus est valide pour tout ε, on a la possibilité de considérer $\varepsilon = 1$. Allons écrire par simplicité $\delta_{\varepsilon=1} \equiv K > 0$. Avec ces choix, l'hypothèse de continuité de A en 0_V donne lieu à l'implication :

$$\|x\|_V < K \implies \|Ax\|_W < 1, \tag{6.2}$$

qui s'approche de la définition d'opérateur borné.

Tout ce qui reste à faire est de construire un x particulier qui satisfait (6.2) et qui nous permet de manipuler l'inégalité $\|Ax\|_W < 1$ afin d'arriver à montrer que A est borné.

Pour faire ça, considérons un nombre réel positif $0 < \sigma < K$, comme ça $K - \sigma > 0$, et un élément $y \in V$ *quelconque*.

Analysons la norme du vecteur $(K-\sigma)\frac{y}{\|y\|_V}$:

$$\left\|(K-\sigma)\frac{y}{\|y\|_V}\right\|_V = \frac{K-\sigma}{\|y\|_V}\|y\|_V = K - \sigma < K,$$

i.e. $(K-\sigma)\frac{y}{\|y\|_V}$ est un vecteur de V avec norme strictement inférieure à K, alors la relation (6.2) implique

$$\left\|A\left((K-\sigma)\frac{y}{\|y\|_V}\right)\right\|_W < 1 \iff \frac{K-\sigma}{\|y\|_V}\|Ay\|_W < 1 \iff \|Ay\|_W < \frac{1}{K-\sigma}\|y\|_V.$$

Comme $y \in V$ est arbitraire et $K - \sigma > 0$, on peut poser $c \equiv \frac{1}{K-\sigma}$ et obtenir la définition de A borné :

$$\|Ay\|_W < c\|y\|_W \quad \forall y \in V.$$

□

On peut donc **utiliser les adjectifs borné et continu d'une manière équivalente pour les opérateurs linéaires** entre espaces vectoriels normés.

Ci-dessus on a voulu spécifier l'espace vectoriel dans lequel on considère la norme, néanmoins, pour rendre la notation plus simple, dans la suite on évitera cette spécification et on écrira tout simplement le symbole $\|\ \|$.

6.1.1 La continuité des opérateurs linéaires définis sur un espace vectoriel normé de dimension finie

Le résultat suivant montre que tout opérateur linéaire défini sur un espace vectoriel de dimension finie est continue (et donc borné).

Théorème 6.1.4 *Si V est un espace vectoriel normé de dimension finie N et W est un espace vectoriel normé (de dimension quelconque), alors tout opérateur linéaire $A : V \to W$ est borné (et donc continu).*

Preuve. L'espace V étant de dimension finie, toutes les normes sur V sont équivalentes grâce au théorème de Tychonov (4.2.3), donc, il suffit de démontrer que $A : V \to W$ est continu par rapport à une norme particulière, pour le démontrer pour toutes les autres normes.

Pour cela, soit $(u_1, \ldots, u_N)$ est une base de V, on peut écrire tout $x \in V$ comme $x = \sum_{n=1}^{N} x_n u_n$, $x_n \in \mathbb{K}$, et considérer sur V la norme suivante : $\|x\| = \left\| \sum_{n=1}^{N} x_n u_n \right\| \equiv \sup_{n=1,\ldots,N} |x_n|$. La linéarité de A et l'inégalité triangulaire impliquent alors :

$$\begin{aligned}\|Ax\| = \left\| A\left(\sum_{n=1}^{N} x_n u_n \right) \right\| = \left\| \sum_{n=1}^{N} x_n A u_n \right\| &\leqslant \sum_{n=1}^{N} \|x_n A u_n\| = \sum_{n=1}^{N} |x_n| \, \|A u_n\| \\ &\leqslant \sum_{n=1}^{N} \sup_{n=1,\ldots,N} |x_n| \, \|A u_n\| = \left(\sup_{n=1,\ldots,N} |x_n| \right) \left(\sum_{n=1}^{N} \|A u_n\| \right) \underset{\text{déf de} \|x\|}{=} \left(\sum_{n=1}^{N} \|A u_n\| \right) \|x\|,\end{aligned}$$

ce qui montre que A est borné, i.e. continu. □

Il n'y a, donc, pas de problèmes de continuité quand on considère des opérateurs linéaires définis sur des espaces vectoriels normés de dimension finie, *indépendamment de la dimension de l'espace image*. Comme on va le voir dans la suite, la situation est bien différente si la dimension du domaine de l'opérateur linéaire est infini.

6.2 La norme opératorielle, la convergence des suites d'opérateurs et les algèbres de Banach

Déf. 6.2.1 *Soit $A : V \to W$, $V \neq \{0_V\}$, un opérateur linéaire borné. On définit la* ***norme opératorielle*** *de A de quatre manières différentes (mais équivalentes) :*

$$\|A\| = \inf \{c \geqslant 0 : \|Ax\| \leqslant c\|x\|, \ \forall x \in V\} = N_1 \tag{6.3}$$

$$\|A\| = \sup_{\|x\| \leqslant 1} \|Ax\| = N_2 \tag{6.4}$$

$$\|A\| = \sup_{\|x\| = 1} \|Ax\| = N_3 \tag{6.5}$$

$$\|A\| = \sup_{x \neq 0_V} \frac{\|Ax\|}{\|x\|} = N_4 \tag{6.6}$$

Pour un opérateur non borné A on écrira $\|A\| = +\infty$ et, évidemment, pour l'opérateur nul $\mathbf{0}$ ça vaut $\|\mathbf{0}\| = 0$. La définition ci-dessus est bien posée grâce au théorème suivant.

Théorème 6.2.1 *Les quatre définitions ci-dessus coïncident*[2].

Preuve. On va montrer que $N_1 \leqslant N_4 \leqslant N_3 \leqslant N_2 \leqslant N_1 \ldots$ de la droite vers la gauche. Dans toutes les preuves on utilisera le fait que, par définition, le sup d'un ensemble est le plus petit des majorants de l'ensemble lui-même.

$\boxed{N_2 \leqslant N_1}$: par définition de N_1, i.e. (6.3), on peut écrire $\|Ax\| \leqslant N_1\|x\|$ $\forall x \in V$, donc, en particulier, pour les vecteurs x tels que $\|x\| \leqslant 1$ c'est vrai que $\|Ax\| \leqslant N_1$, i.e. N_1 est un majorant pour l'ensemble $\{\|Ax\|,\ x \in V,\ \|x\| \leqslant 1\}$. Par définition, le sup est le plus petit des majorants d'un ensemble, donc $N_2 = \sup\limits_{\|x\|\leqslant 1} \|Ax\| \leqslant N_1$.

$\boxed{N_3 \leqslant N_2}$: considérons $x \in V$ tel que $\|x\| = 1$ et la suite $x_n = \left(1 - \frac{1}{n}\right) x$, $n \geqslant 1$. D'un côté : $\|x_n\| = \left(1 - \frac{1}{n}\right) \|x\| = \left(1 - \frac{1}{n}\right) \leqslant 1$ et donc $\|Ax_n\| \leqslant \sup\limits_{\|y\|\leqslant 1} \|Ay\| = N_2$. Si on passe à la limite on obtient : $\lim\limits_{n\to+\infty} \|Ax_n\| \leqslant \lim\limits_{n\to+\infty} N_2 = N_2$. De l'autre côté, il est évident que $x_n \underset{n\to+\infty}{\to} x$ et alors, par la continuité de A et de la norme : $\lim\limits_{n\to+\infty} \|Ax_n\| = \|A \lim\limits_{n\to+\infty} x_n\| = \|Ax\|$.

Si on mélange les informations qu'on a obtenu on peut écrire $\|Ax\| \leqslant N_2$, i.e. N_2 est un majorant pour l'ensemble $\{\|Ax\|,\ x \in V,\ \|x\| = 1\}$. La quantité N_3 est définie être le sup de cet ensemble, i.e. le plus petit majorant, par conséquent $N_3 \leqslant N_2$.

$\boxed{N_4 \leqslant N_3}$: considérons $x \in V$, $x \neq 0_V$, alors $\left\| \frac{x}{\|x\|} \right\| = 1$ et $\left\| A\frac{x}{\|x\|} \right\| \leqslant \sup\limits_{\|y\|=1} \|Ay\| = N_3$, mais $\left\| A\frac{x}{\|x\|} \right\| = \frac{\|Ax\|}{\|x\|}$, donc $\frac{\|Ax\|}{\|x\|} \leqslant N_3$, i.e. c est un majorant pour l'ensemble $\left\{ \frac{\|Ax\|}{\|x\|},\ x \in V,\ x \neq 0_V \right\}$. Vu que N_4 est le sup de cet ensemble, i.e. le plus petit des majorants, alors $N_4 \leqslant N_3$.

$\boxed{N_1 \leqslant N_4}$: pour tout $x \neq 0_V$, $\frac{\|Ax\|}{\|x\|} \leqslant N_4$, or $\|Ax\| \leqslant N_4\|x\|$ mais, par définition de N_1 ça vaut que $N_1 \leqslant N_4$. □

Remarque 6.2.1

1. Dans la définition $\|A\| = \inf\{c \geqslant 0 : \|Ax\| \leqslant c\|x\|,\ \forall x \in V\}$, il faut faire attention à la spécification $\forall x \in V$: sans cette condition la norme de A serait trivialement nulle pour tout opérateur linéaire, car $A0 = 0$. Par contre, le fait de considérer tous les vecteurs transformés Ax, $x \in V$, permet à la norme de A d'être $\neq 0$ (sauf, bien sûr, quand A est l'opérateur identiquement nul).
2. Nous soulignons aussi la différence entre l'expression $\|A\|$, qui représente la norme opératorielle de l'application linéaire $A : V \to W$ et l'expression $\|x\|$, qui représente la norme d'un vecteur $x \in V$. Certains auteurs utilisent un symbole différent pour la norme opératorielle, comme, par exemple, $|||A|||$, mais nous utiliserons le même symbole, i.e. $\|\ \|$.

2. Une cinquième caractérisation de la norme d'un opérateur borné est possible si $V = W = \mathcal{H}$, espace de Hilbert, cfr. formule (6.19).

Allons maintenant vérifier que la norme opératorielle est bien définie sur l'ensemble des opérateurs linéaires de V à W et que cet espace est stable par rapport aux opérations linéaires définies ponctuellement, i.e. $(A+B)x = Ax + Bx$ et $(\alpha A)x = \alpha Ax$, pour tout $\alpha \in \mathbb{K}$ et pour tout $x \in V$.

— *Définie positivité* : évidemment $\|A\| \geqslant 0$ pour tout opérateur borné A grâce à la déf. (6.3), de plus, grâce à la déf. (6.6), $\|A\| = \sup\limits_{x \neq 0_V} \frac{\|Ax\|}{\|x\|} = 0$ si et seulement si $\|Ax\| = 0$ $\forall x \in V$, $x \neq 0_V$ (si $x = 0_V$ alors $Ax = 0$ par linéarité). Donc, grâce à la définie positivité de la norme de W, $\|A\| = 0 \iff Ax = 0\ \forall x \in V$, i.e. si et seulement si A est l'opérateur nul $\mathbf{0}(x) = 0\ \forall x \in V$.

— *Homogénéité* : conséquence directe de l'homogénéité de la norme de W, en fait $\forall \alpha \in \mathbb{K}$, en utilisant, par exemple, la déf. (6.5) on obtient

$$\|\alpha A\| = \sup_{\|x\|=1} \|\alpha Ax\| = \sup_{\|x\|=1} |\alpha| \|Ax\| = |\alpha| \sup_{\|x\|=1} \|Ax\| = |\alpha| \|A\|,$$

i.e.

$$\boxed{\|\alpha A\| = |\alpha| \|A\|} \qquad \forall \alpha \in \mathbb{K}. \tag{6.7}$$

— *Inégalité triangulaire* : on observe que, comme conséquence de la définition (6.3), on peut écrire :

$$\boxed{\|Ax\| \leqslant \|A\| \|x\|} \quad \forall x \in V, \tag{6.8}$$

donc, en utilisant ceci et l'inégalité triangulaire de la norme de W, pour toute couple d'opérateurs $A, B : V \to W$ et pour tout $x \in V$ on peut écrire

$$\|(A+B)x\| := \|Ax + Bx\| \leqslant \|Ax\| + \|Bx\| \leqslant \|A\|\|x\| + \|B\|\|x\| = (\|A\| + \|B\|)\|x\|,$$

ce qui implique, grâce à la déf. (6.3), que

$$\boxed{\|A + B\| \leqslant \|A\| + \|B\|}. \tag{6.9}$$

L'inégalité (6.9) et l'homogénéité (6.7) montrent que l'ensemble des opérateurs linéaires bornés est stable par rapport aux combinaisons linéaires et donc il est un espace vectoriel lui-même, qui devient normé grâce à la norme opératorielle.

Déf. 6.2.2 *On écrit avec le symbole $\mathcal{B}(V, W)$ l'espace vectoriel normé des opérateurs linéaires continus de V à W avec norme opératorielle. Si $V = W$, on écrit simplement $\mathcal{B}(V)$.*

Dans la littérature mathématique on utilise la lettre $\mathcal{B}$ pour « bounded » (borné, en anglais). Une autre notation souvent utilisée est $\mathcal{L}(V, W)$.

La définition suivante est une conséquence immédiate du fait que $\mathcal{B}(V, W)$ est un espace vectoriel normé.

Déf. 6.2.3 (Convergence dans $\mathcal{B}(V, W)$) *Une suite d'opérateurs bornés $(A_n)_{n\in\mathbb{N}} \subset \mathcal{B}(V, W)$ converge vers l'opérateur borné $A \in \mathcal{B}(V, W)$ si :*

$$\|A_n - A\| \underset{n\to+\infty}{\longrightarrow} 0,$$

où $\|A_n - A\|$ est la norme opératorielle de la différence entre A_n et A.

Exercice 6.1

Utiliser la définition (6.4) pour prouver qu'une condition nécessaire de convergence pour une suite d'opérateurs de $(A_n)_{n\in\mathbb{N}} \subset \mathcal{B}(V,W)$ vers $A \in \mathcal{B}(V,W)$ est que

$$\lim_{n\to+\infty} \|(A_n - A)x\| = 0 \quad \forall x \in \overline{B(0,1)}. \tag{6.10}$$

Solution de l'exercice 6.1

Pour obtenir cette condition, on commence par observer que, vu que le sup est un majorant d'un ensemble, ça vaut que

$$\boxed{\|A\| \geqslant \|Ax\| \quad \forall x \in \overline{B(0,1)}}\,, \qquad \overline{B(0,1)} := \{x \in V \ : \ \|x\| \leqslant 1\}. \tag{6.11}$$

L'inégalité (6.11) implique $\|A_n - A\| \geqslant \|(A_n - A)x\|\ \forall x \in \overline{B(0,1)}$, donc, s'il existe au moins un $x \in \overline{B(0,1)}$ tel que $\lim_{n\to+\infty} \|(A_n - A)x\| > 0$, alors $\lim_{n\to+\infty} \|A_n - A\| > 0$ et on ne peut pas avoir la convergence de la suite $(A_n)_{n\in\mathbb{N}}$ vers A. La propriété (6.10) est donc nécessaire pour la convergence de $(A_n)_{n\in\mathbb{N}} \subset \mathcal{B}(V,W)$ vers $A \in \mathcal{B}(V,W)$. □

Dans le cas $V = W$ on peut rajouter une troisième opération sur $\mathcal{B}(V)$, le **produit** :

$$(AB)(x) := (A \circ B)(x) = A(B(x)) \quad \forall x \in V,$$

i.e., *le produit en $\mathcal{B}(V)$ correspond à l'opération de composition fonctionnelle entre opérateurs linéaires*. On observe que :

$$\|(AB)x\| = \|A(Bx)\| \underset{(Bx\ :\ \text{vecteur de V})}{\leqslant} \|A\|\|Bx\| \leqslant \|A\|\|B\|\|x\| \quad \forall x \in V,$$

et donc, grâce à la déf. (6.3),

$$\boxed{\|AB\| \leqslant \|A\|\|B\|} \tag{6.12}$$

et ainsi, en prenant $A = B$, $\|A^2\| \leqslant \|A\|^2$, en itérant ces considérations on arrive à la formule :

$$\boxed{\|A^n\| \leqslant \|A\|^n} \qquad \forall n \in \mathbb{N}.$$

Par conséquent, $\mathcal{B}(V)$ est stable par rapport à l'opération de produit défini ci-dessus, donc $\mathcal{B}(V)$ **est une algèbre associative normée avec unité**, où l'unité est l'opérateur identité.

Nous rappelons qu'une *algèbre* $\mathfrak{A}$ sur le corps $\mathbb{K}$ est un espace vectoriel sur $\mathbb{K}$ muni d'une opération binaire $\cdot : \mathfrak{A} \times \mathfrak{A} \to \mathfrak{A}$, communément appelée *produit*, compatible avec les opérations linéaires, ce qui est équivalent à dire que $\cdot$ est bilinéaire, i.e. pour tout $a, b, c \in \mathfrak{A}$ et $k \in \mathbb{K}$ ça vaut :

— $(a + b) \cdot c = a \cdot c + b \cdot c$ et $a \cdot (b + c) = a \cdot b + a \cdot c$;
— $(ka) \cdot b = k(a \cdot b)$, $a \cdot (kb) = k(a \cdot b)$.

Théorème 6.2.2 *Soit $(V, \|\ \|)$ un espace vectoriel normé quelconque, alors, les opérations de somme, produit par un scalaire et produit dans l'algèbre $\mathcal{B}(V)$ sont continues par rapport à la norme opératorielle.*

Preuve. Le théorème 4.2.1 s'applique aussi dans le cas de l'algèbre $\mathcal{B}(V)$, donc la somme et le produit par un scalaire sont continues, la seule preuve qui reste à faire est celle relative à la continuité du produit opératoriel. Si $(A_n)_{n\in\mathbb{N}}$ et $(B_n)_{n\in\mathbb{N}}$ sont deux suites d'opérateurs de $\mathcal{B}(V)$ convergentes vers $A \in \mathcal{B}(V)$ et $B \in \mathcal{B}(V)$, respectivement, i.e. $\|A_n - A\| \underset{n\to+\infty}{\to} 0$, $\|B_n - B\| \underset{n\to+\infty}{\to} 0$, alors il faut démontrer que $A_nB_n \underset{n\to+\infty}{\to} AB$, i.e. que $\|A_nB_n - AB\| \underset{n\to+\infty}{\to} 0$:

$$\|A_nB_n - AB\| = \|A_n(B_n - B) + (A_n - A)B\| \underset{(6.9),(6.12)}{\leqslant} \|A_n\|\|B_n - B\| + \|A_n - A\|\|B\| \underset{n\to+\infty}{\to} 0.$$

□

La présence d'une norme sur $\mathcal{B}(V, W)$ génère une topologie et ceci amène d'une manière naturelle à se demander sous quelles conditions cet espace est complet. Le résultat suivant donne une condition suffisante pour la complétude de $\mathcal{B}(V, W)$.

Théorème 6.2.3 *Soient V, W deux espaces vectoriels normés. Si W est complet alors $\mathcal{B}(V, W)$ est complet.*

Avant de montrer la preuve, on souligne que, en particulier, **le théorème vaut pour $\mathcal{B}(\mathcal{H})$ ou $\mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$**, si $\mathcal{H}, \mathcal{H}_1, \mathcal{H}_2$ sont des espaces de Hilbert.

Preuve. Soit $(A_n)_{n\in\mathbb{N}}$ une suite de Cauchy d'opérateurs de $\mathcal{B}(V, W)$, i.e. :

$$\forall \varepsilon > 0 \quad \exists N_\varepsilon > 0 \ : \ \forall m, n \geqslant N_\varepsilon \ : \ \|A_n - A_m\| < \varepsilon,$$

pour prouver le théorème il faut démontrer que $(A_n)_{n\in\mathbb{N}}$ converge en $\mathcal{B}(V, W)$ en utilisant l'hypothèse de complétude de W.

On commence par observer que $\forall x \in V$ fixé, ça vaut que

$$\forall m, n \geqslant N_\varepsilon \ : \ \|A_nx - A_mx\| \leqslant \|A_n - A_m\|\|x\| < \varepsilon\|x\|, \tag{6.13}$$

et donc, grâce à l'arbitrarité de ε, $(A_nx)_{n\in\mathbb{N}}$ est une suite de Cauchy en W.

Par hypothèse, W est complet, donc il existe $\lim_{n\to\infty} A_nx \in W$ et cela permet de définir l'opérateur limite A associé à la suite $(A_n)_{n\in\mathbb{N}}$:

$$\begin{array}{rccl} A: & V & \longrightarrow & W \\ & x & \longmapsto & A(x) = \lim\limits_{n\to+\infty} A_nx. \end{array}$$

On va démontrer que $(A_n)_{n\in\mathbb{N}}$ converge à A en norme opératorielle et que $A \in \mathcal{B}(V, W)$, ceci terminera la preuve.

On commence par observer que $\forall n \geqslant N_\varepsilon$, ça vaut :

$$\|A_n(x) - A(x)\| = \|A_n(x) - \lim_{m\to+\infty} A_m(x)\| \underset{(\text{continutité } \|\ \|)}{=} \lim_{m\to+\infty} \|A_n(x) - A_m(x)\| \underset{(6.13)}{<} \varepsilon\|x\|, \tag{6.14}$$

où dans la dernière inégalité on a utilisé le fait que m tend vers $+\infty$, donc sûrement $m \geqslant N_\varepsilon$.

Par conséquent,

$$\forall \varepsilon > 0 \ \exists N_\varepsilon > 0 \ : \ n \geqslant N_\varepsilon \implies \|A_n - A\| = \sup_{\|x\|=1} \|(A_n - A)x\| = \sup_{\|x\|=1} \|A_n(x) - A(x)\| < \varepsilon,$$

i.e. $(A_n)_{n\in\mathbb{N}}$ converge vers A en norme opératorielle.

Il reste simplement à vérifier que $A \in \mathcal{B}(V,W)$. Soit $x \in V$ quelconque, alors, comme l'inéquation (6.14) vaut pour tout $n \geqslant N_\varepsilon$ on peut écrire :

$$\|Ax\| = \|Ax - A_{N_\varepsilon}x + A_{N_\varepsilon}x\| \leqslant \|Ax - A_{N_\varepsilon}x\| + \|A_{N_\varepsilon}x\| \underset{(6.14)}{<} \varepsilon\|x\| + \|A_{N_\varepsilon}\|\|x\|,$$

i.e. $\|Ax\| < (\varepsilon + \|A_{N_\varepsilon}\|)\|x\|$ $\forall x \in V$, et donc A est borné. □

Déf. 6.2.4 (Algèbre de Banach) *Une algèbre $\mathfrak{A}$ sur le corps $\mathbb{K}$ est une algèbre de Banach si les propriétés suivantes sont vérifiées $\forall a, b, c \in \mathfrak{A}$:*

- *$\mathfrak{A}$ est une algèbre associative, i.e. $a \cdot (b \cdot c) = (a \cdot b) \cdot c$;*
- *$\mathfrak{A}$, en tant qu'espace vectoriel, admet une norme $\|\ \|$ par rapport à laquelle il est un espace de Banach ;*
- *$\|a \cdot b\| \leqslant \|a\|\,\|b\|$.*

Grâce à ce qu'on a vu avant, si V est un espace de Banach, $\mathcal{B}(V)$ est une algèbre associative avec unité et complète par rapport à la norme opératorielle, et donc $\mathcal{B}(V)$ est une algèbre de Banach avec unité. Évidemment, pour tout espace de Hilbert $\mathcal{H}$, $\mathcal{B}(\mathcal{H})$ **est une algèbre de Banach avec unité**.

Allons maintenant présenter une propriété très importante du noyau des opérateurs de $\mathcal{B}(V,W)$.

Théorème 6.2.4 *Soient V, W deux espaces vectoriels normés et soit $A \in \mathcal{B}(V,W)$, alors* $\ker(A)$ ***est un sous-espace vectoriel fermé*** *de V.*

Preuve. Soit $(v_n)_{n\in\mathbb{N}} \subset \ker(A)$ une arbitraire suite convergente, il faut démontrer que sa limite, $\bar{v} = \lim\limits_{n\to+\infty} v_n$, reste en $\ker(A)$. A est borné donc continu et alors $\lim\limits_{n\to+\infty} Av_n = A(\lim\limits_{n\to+\infty} v_n) = A\bar{v}$, mais $Av_n = 0$ $\forall n \in \mathbb{N}$ car $v_n \in \ker(A)$, donc $A\bar{v} = \lim\limits_{n\to+\infty} 0 = 0$, ce qui implique que $\bar{v} \in \ker(A)$. □

Pour montrer l'utilité remarquable de ce théorème, allons l'exploiter tout de suite dans le prochain exercice pour montrer que **le théorème de la projection sur un sous-espace vectoriel propre et fermé n'est plus valide sans l'hypothèse de complétude**.

Exercice 6.2

Soit T l'opérateur linéaire (en fait une fonctionnelle linéaire) définit par :

$$\begin{array}{rccc} T: & \ell^2(\mathbb{N},\mathbb{C}) & \longrightarrow & \mathbb{C} \\ & x = (x_n)_{n\in\mathbb{N}} & \longmapsto & T(x) = \sum\limits_{n\in\mathbb{N}} \frac{x_n}{n+1}. \end{array}$$

1. Montrer que T est continue.
2. Soit $F = \left\{(x_n)_{n\in\mathbb{N}} \in \ell^2(\mathbb{N},\mathbb{C}) \ : \ \sum\limits_{n\in\mathbb{N}} \frac{x_n}{n+1} = 0\right\}$. Montrer que F est un sous-espace vectoriel propre et fermé de $\ell^2(\mathbb{N},\mathbb{C})$.

3. Montrer qu'il existe $u \in \ell^2(\mathbb{N},\mathbb{C})$ tel que $F = \{u\}^\perp$ et en déduire l'expression explicite de $F^\perp$.
4. On rappelle (cfr. déf. (4.26)) que $\ell^0(\mathbb{N},\mathbb{C})$ est le sous-espace vectoriel de $\ell^2(\mathbb{N},\mathbb{C})$ formé des suites $(x_n)_{n\in\mathbb{N}}$ nulles à partir d'un certain rang, que l'on munit de la topologie induite par $\ell^2(\mathbb{N},\mathbb{C})$. Soit $G = F \cap \ell^0(\mathbb{N},\mathbb{C})$.
   (a) Montrer que G est un sous-espace vectoriel propre et fermé de $\ell^0(\mathbb{N},\mathbb{C})$.
   (b) En rappelant la formule (5.4), démontrer que le complément orthogonal de G dans $\ell^0(\mathbb{N},\mathbb{C})$, i.e. $G^{\perp_{\ell^0}} := G^\perp \cap \ell^0$, est réduit au vecteur nul : $G^{\perp_{\ell^0}} = \{0_{\ell^2(\mathbb{N},\mathbb{C})}\}$.
   (c) En déduire que $\ell^0(\mathbb{N},\mathbb{C})$ **n'est pas complet dans la topologie héritée de** $\ell^2(\mathbb{N},\mathbb{C})$.

Solution de l'exercice 6.2

1. Notons avec $u = (u_n)_{n\in\mathbb{N}}$ la suite définie par $u_n = \frac{1}{n+1}$ $\forall n \in \mathbb{N}$, évidemment appartenant à $\ell^2(\mathbb{N},\mathbb{C})$. Pour tout $x \in \ell^2(\mathbb{N},\mathbb{C})$ ça vaut :
$$|T(x)| = \left|\sum_{n\in\mathbb{N}} x_n \frac{1}{n+1}\right| = \left|\sum_{n\in\mathbb{N}} x_n u_n\right| = |\langle x,u\rangle| \underset{\text{Cauchy-Schwarz}}{\leqslant} \|x\|\|u\|,$$
donc T est borné, avec $\|T\| \leqslant \|u\|$, i.e. continu.
2. Par définition, $F = \ker(T)$ et donc il est un sous-espace vectoriel fermé de $\ell^2(\mathbb{N},\mathbb{C})$, vu qu'on vient de démontrer la continuité de T. Un exemple d'élément de $\ell^2(\mathbb{N},\mathbb{C})$ qui n'appartient pas à F est le premier vecteur de la base canonique de $\ell^2(\mathbb{N},\mathbb{C})$, i.e. $e_1 = (1,0,0,\dots)$, vu que $\sum\limits_{n\in\mathbb{N}} \frac{e_1(n)}{n+1} = \frac{1}{2} \neq 0$. Par conséquent, F est un sous-espace vectoriel propre et fermé de $\ell^2(\mathbb{N},\mathbb{C})$.
3. Si u est défini comme dans la question 1. alors il est évident que
$$F = \{x \in \ell^2(\mathbb{N},\mathbb{C}) \ : \ \langle x,u\rangle = 0\} \equiv \{u\}^\perp,$$
donc $F^\perp = \{u\}^{\perp\perp} \underset{(5.2.4)}{=} \overline{\text{span}\{u\}}$, mais $\text{span}\{u\} = \left\{\frac{\lambda}{n+1},\ \lambda \in \mathbb{C}\right\}$ est un sous-espace vectoriel mono-dimensionnel, et donc il est fermé, par conséquent $\overline{\text{span}\{u\}} = \text{span}\{u\}$ et $F^\perp = \left\{\frac{\lambda}{n+1},\ \lambda \in \mathbb{C}\right\}$.
4. (a) Avec un moment de réflexion on peut comprendre que G peut être écrit explicitement comme ça :
$$G = F \cap \ell^0(\mathbb{N},\mathbb{C}) = \{(x_n)_{n\in\mathbb{N}},\ \exists\, N \in \mathbb{N} \ : \ x_n = 0\ \forall n > N \text{ et } \sum_{n=0}^{N} \frac{x_n}{n+1} = 0\},$$
ce qui montre que $G = \ker T|_{\ell^0(\mathbb{N},\mathbb{C})}$, mais comme la restriction d'un opérateur linéaire continue est continue elle-même, G est un sous-espace vectoriel fermé de $\ell^0(\mathbb{N},\mathbb{C})$. Pour démontrer que G est propre, il suffit de considérer encore e_1 : $e_1 \in \ell^0(\mathbb{N},\mathbb{C})$, mais $\sum\limits_{n=0}^{N} \frac{e_1(n)}{n+1} = \frac{1}{2} \neq 0$.

(b) Nous avons :

$$G^{\perp_{\ell^0}} = G^{\perp} \cap \ell^0(\mathbb{N},\mathbb{C}) = (F \cap \ell^0(\mathbb{N},\mathbb{C}))^{\perp} \cap \ell^0(\mathbb{N},\mathbb{C})$$
$$\underset{(5.4)}{=} \overline{\mathrm{span}(F^{\perp} \cup \ell^0(\mathbb{N},\mathbb{C})^{\perp}} \cap \ell^0(\mathbb{N},\mathbb{C}),$$

mais nous avons vu (cfr. th. 4.4.8) que $\ell^0(\mathbb{N},\mathbb{C})$ est dense en $\ell^2(\mathbb{N},\mathbb{C})$, donc $\ell^0(\mathbb{N},\mathbb{C}))^{\perp} = \{0_{\ell^2(\mathbb{N},\mathbb{C})}\}$ qui est déjà inclus en $F^{\perp}$ en tant que sous-espace vectoriel de $\ell^2(\mathbb{N},\mathbb{C})$, mais alors

$$\overline{\mathrm{span}(F^{\perp} \cup \ell^0(\mathbb{N},\mathbb{C})^{\perp})} = \overline{\mathrm{span}(F^{\perp})} = \overline{F^{\perp}} = F^{\perp} = \left\{ \frac{\lambda}{n+1},\ \lambda \in \mathbb{C} \right\},$$

car $F^{\perp}$ est fermé et grâce au résultat de la question 3. Alors,

$$G^{\perp_{\ell^0}} = \left\{ \frac{\lambda}{n+1},\ \lambda \in \mathbb{C} \right\} \cap \ell^0(\mathbb{N},\mathbb{C}) = \{0_{\ell^2(\mathbb{N},\mathbb{C})}\},$$

car évidemment la suite $\frac{\lambda}{n+1} \notin \ell^0(\mathbb{N},\mathbb{C})$.

(c) G est un sous-espace vectoriel propre et fermé de $\ell^0(\mathbb{N},\mathbb{C})$ muni de la topologie héritée par $\ell^2(\mathbb{N},\mathbb{C})$, néanmoins on vient de démontrer que $G^{\perp_{\ell^0}}$, le complément orthogonal de G en $\ell^0(\mathbb{N},\mathbb{C})$, est réduit au seul vecteur nul, mais ceci est en contradiction avec le résultat du théorème 5.2.3 (un corollaire du théorème de la projection sur une partie propre convexe fermée d'un espace de Hilbert) qui dit que le complément orthogonal d'un sous-espace vectoriel propre et fermé d'un espace de Hilbert n'est pas réduit au seul vecteur nul. Il est clair que la seule hypothèse qui est mise en défaut est la complétude de $\ell^0(\mathbb{N},\mathbb{C})$ par rapport à la topologie héritée de $\ell^2(\mathbb{N},\mathbb{C})$.

□

Voyons maintenant comment un opérateur linéaire continue entre deux espaces vectoriels normés interagit avec les suites de Cauchy.

Théorème 6.2.5 *Soient V et W deux espaces vectoriels normés quelconques, $A \in \mathcal{B}(V,W)$ et $(x_n)_{n\in\mathbb{N}} \subset V$ une suite de Cauchy, alors $(Ax_n)_{n\in\mathbb{N}}$ est une suite de Cauchy en W.*

Preuve. Par hypothèse : $\forall \varepsilon > 0\ \exists N_\varepsilon > 0 : \forall n, m \geqslant N_\varepsilon : \|x_n - x_m\| < \varepsilon$. Considérons maintenant $(Ax_n)_{n\in\mathbb{N}}$, et analysons $\|Ax_n - Ax_m\| = \|A(x_n - x_m)\| \leqslant \|A\|\|x_n - x_m\| < \|A\|\varepsilon, \forall n, m \geqslant N_\varepsilon$. Grâce à l'arbitrarité de ε, $(Ax_n)_{n\in\mathbb{N}}$ est une suite de Cauchy d'éléments de W. □

Ce résultat peut *aider à prouver la complétude d'un espace vectoriel normé*, comme on le montre dans l'exercice suivant, qui peut être vu comme la suite de l'exercice 4.2.

Exercice 6.3

Nous rappelons que, fixé une suite $a = (a_n)_{n\in\mathbb{N}}$ de nombres réels strictement positifs, on note

$$\ell^2_a(\mathbb{N},\mathbb{C}) := \{u \in \mathbb{C}^{\mathbb{N}} \ :\ \sum_{n\in\mathbb{N}} a_n |u_n|^2 < +\infty \iff \sqrt{a}u \in \ell^2(\mathbb{N},\mathbb{C})\}.$$

Dans l'exercice 4.2 on avait vérifié que

$$\langle u, v\rangle_{\ell^2_a} = \sum_{n\in\mathbb{N}} a_n u_n \overline{v_n} \quad \text{et} \quad \|u\|^2_{\ell^2_a} = \sum_{n\in\mathbb{N}} a_n |u_n|^2$$

sont un produit scalaire et une norme sur $\ell^2_a(\mathbb{N},\mathbb{C})$.

1. Montrer que l'opérateur

$$\begin{array}{rccl} \imath_a : & \ell^2_a(\mathbb{N},\mathbb{C}) & \hookrightarrow & \ell^2(\mathbb{N},\mathbb{C}) \\ & u & \mapsto & \imath_a(u) := \sqrt{a}u \equiv (\sqrt{a_n}u_n)_{n\in\mathbb{N}}, \end{array}$$

est linéaire, continu, de norme égale à 1 et bijectif. Expliciter l'opérateur inverse de $\imath_a$ et vérifier qu'il soit continu et de norme égale à 1.

2. En déduire que $\ell^2_a(\mathbb{N},\mathbb{C})$ est un espace de Hilbert.
3. Soit a et b deux suites de nombres réels strictement positifs telles que $a_n \underset{n\to+\infty}{=} O(b_n)$. Montrer que $\ell^2_b(\mathbb{N},\mathbb{C}) \subset \ell^2_a(\mathbb{N},\mathbb{C})$ et que cette injection est continue.

Solution de l'exercice 6.3

1. La linéarité suit d'un simple exercice d'écriture : si $u, v \in \ell^2_a(\mathbb{N},\mathbb{C})$ et $\lambda \in C$, alors $\imath(u+\lambda v) = \sqrt{a}(u+\lambda v) = \sqrt{a}u + \lambda\sqrt{a}v = \imath(u) + \lambda\imath(v)$. Examinons la continuité : pour tout $u \in \ell^2_a(\mathbb{N},\mathbb{C})$, $\imath(u) \in \ell^2(\mathbb{N},\mathbb{C})$ et sa norme vaut

$$\|\imath_a(u)\|^2_{\ell^2} = \sum_{n\in\mathbb{N}} |\sqrt{a_n}u_n|^2 = \sum_{n\in\mathbb{N}} a_n|u_n|^2 = \|u\|^2_{\ell^2_a} \iff \|\imath_a(u)\|_{\ell^2} = \|u\|^2_{\ell^2_a}$$

ce qui montre la continuité de $\imath_a$ et le fait que sa norme soit unitaire. Nous terminons avec la bijectivité : observons que l'opérateur

$$\begin{array}{rccl} j_{1/a} : & \ell^2(\mathbb{N},\mathbb{C}) & \to & \ell^2_a(\mathbb{N},\mathbb{C}) \\ & v & \mapsto & \imath_{1/a}(v) := \frac{1}{\sqrt{a}}v \equiv (\frac{1}{\sqrt{a_n}}v_n)_{n\in\mathbb{N}}, \end{array}$$

est bien défini car $a > 0$ et $v/\sqrt{a} \in \ell^2_a(\mathbb{N},\mathbb{C}) \iff \sqrt{a}v/\sqrt{a} = v \in \ell^2(\mathbb{N},\mathbb{C})$. De plus, il est tel que $j_{1/a}\circ\imath_a : \ell^2_a(\mathbb{N},\mathbb{C}) \to \ell^2_a(\mathbb{N},\mathbb{C})$, $j_{1/a}\circ\imath_a(u) = \frac{\sqrt{a}}{\sqrt{a}}u = u \ \forall u \in \ell^2_a(\mathbb{N},\mathbb{C})$; vice-versa, pour tout $v \in \ell^2(\mathbb{N},\mathbb{C})$, $\imath_a \circ j_{1/a} : \ell^2(\mathbb{N},\mathbb{C}) \to \ell^2(\mathbb{N},\mathbb{C})$, $\imath_a \circ j_{1/a}(v) = \frac{\sqrt{a}}{\sqrt{a}}v = v$, i.e. $j_{1/a}\circ\imath_a = id_{\ell^2_a(\mathbb{N},\mathbb{C})}$ et $\imath_a \circ j_{1/a} = id_{\ell^2(\mathbb{N},\mathbb{C})}$. Donc, $\imath_a$ est bijectif avec inverse $j_{1/a}$, évidemment continu lui aussi et avec norme unitaire car

$$\|j_{1/a}(v)\|^2_{\ell^2_a} = \sum_{n\in\mathbb{N}} \frac{a_n}{a_n}|v_n|^2 = \|v\|^2_{\ell^2} \iff \|j_{1/a}(v)\|_{\ell^2_a} = \|v\|^2_{\ell^2} \qquad \forall v \in \ell^2(\mathbb{N},\mathbb{C}). \quad (6.15)$$

2. Grâce à la continuité de $\imath_a$ et au théorème 6.2.5, $\imath_a$ transforme des suites de Cauchy de $\ell^2_a(\mathbb{N},\mathbb{C})$ en suites de Cauchy de $\ell^2(\mathbb{N},\mathbb{C})$. Soit alors $(u_m)_{m\in\mathbb{N}}$ une suite de Cauchy quelconque d'éléments de $\ell^2_a(\mathbb{N},\mathbb{C})$, $\imath_a((u_m)_{m\in\mathbb{N}})$ est une suite de Cauchy de $\ell^2(\mathbb{N},\mathbb{C})$, que l'on sait être complet, donc $\exists\, L \in \ell^2(\mathbb{N},\mathbb{C})$ tel que $\imath_a((u_m)_{m\in\mathbb{N}}) \underset{m\to+\infty}{\to} L$, i.e.

$$\begin{aligned} 0 &= \lim_{m\to+\infty} \|\imath_a((u_m)_{m\in\mathbb{N}}) - L\|_{\ell^2} \underset{(6.15)}{=} \lim_{m\to+\infty} \|j_{1/a}(\imath_a((u_m)_{m\in\mathbb{N}}) - L)\|_{\ell^2_a} \\ &\underset{j_{1/a}\text{ linéaire}}{=} \lim_{m\to+\infty} \|j_{1/a}\circ\imath_a((u_m)_{m\in\mathbb{N}}) - j_{1/a}(L)\|_{\ell^2_a} = \lim_{m\to+\infty} \|(u_m)_{m\in\mathbb{N}} - j_{1/a}(L)\|_{\ell^2_a}, \end{aligned}$$

i.e. $(u_m)_{m\in\mathbb{N}}$ converge en $\ell^2_a(\mathbb{N},\mathbb{C})$ vers $j_{1/a}(L)$, et alors $\ell^2_a(\mathbb{N},\mathbb{C})$ est un espace de Hilbert.

3. Nous devons démontrer que, si $a_n \underset{n\to+\infty}{=} O(b_n)$, alors $u \in \ell^2_b(\mathbb{N},\mathbb{C}) \implies \ell^2_a(\mathbb{N},\mathbb{C})$, i.e.

$$\sum_{n\in\mathbb{N}} b_n|u_n|^2 < +\infty \implies \sum_{n\in\mathbb{N}} a_n|u_n|^2 < +\infty,$$

pour tout $u \in \ell^2_b(\mathbb{N},\mathbb{C})$. Par définition, $a_n \underset{n\to+\infty}{=} O(b_n)$ si et seulement s'il existe $C_1 > 0$ et $N \in \mathbb{N}$ tels que, pour tout $n \geqslant N$ ça vaut $a_n \leqslant C_1 b_n$.
Pour arriver à démontrer ce que l'on veut, il n'est pas compliqué de voir qu'il faut multiplier les deux côtés de l'inégalité précedente par $|u_n|^2$, en obtenant $a_n|u_n|^2 \leqslant C_1 b_n|u_n|^2$ pour tout $n \geqslant N$, i.e. $\sum\limits_{n=N}^{+\infty} a_n|u_n|^2 \leqslant C_1 \sum\limits_{n=N}^{+\infty} b_n|u_n|^2$. La somme des premiers N termes, de $n = 0$ à $n = N-1$ est finie, donc il existe sûrement une constante $C_2 > 0$ suffisamment grande telle que $\sum\limits_{n=0}^{N-1} a_n|u_n|^2 \leqslant C_2 \sum\limits_{n=0}^{N-1} b_n|u_n|^2$, mais alors il suffit de considérer $C := \max(C_1, C_2) > 0$ pour avoir $\sum\limits_{n\in\mathbb{N}} a_n|u_n|^2 \leqslant C \sum\limits_{n\in\mathbb{N}} b_n|u_n|^2$, ce qui montre que, si $u \in \ell^2_b(\mathbb{N},\mathbb{C})$, alors $u \in \ell^2_a(\mathbb{N},\mathbb{C})$. De plus, l'inégalité précédente peut être réécrite comme $\|u\|^2_{\ell^2_a} \leqslant C\|u\|^2_{\ell^2_b}$ et alors l'injection canonique $\iota : \ell^2_b(\mathbb{N},\mathbb{C}) \hookrightarrow \ell^2_a(\mathbb{N},\mathbb{C})$ vérifie $\|\iota(u)\|_{\ell^2_a} \leqslant \sqrt{C}\|u\|_{\ell^2_b}$ pour tout $u \in \ell^2_b(\mathbb{N},\mathbb{C})$, i.e. elle est bornée, et donc continue.

□

On termine cette section avec *un résultat extrêmement utile* qui permet de caractériser l'égalité entre opérateurs continus sur un espace vectoriel avec produit scalaire de dimension quelconque avec l'égalité de leur action sur les vecteurs à l'intérieur d'un produit scalaire.

Théorème 6.2.6 *Soient $A, B : V \to W$ deux opérateurs linéaires définis sur un espace vectoriel avec produit scalaire de dimension quelconque, alors :*

$$A = B \iff \langle x, Ay\rangle = \langle x, By\rangle \quad \forall x, y \in \mathcal{H} \iff \langle x, Ax\rangle = \langle x, Bx\rangle \quad \forall x \in \mathcal{H}.$$

Preuve. Prouvons d'abord la première équivalence. Par linéarité de A ça vaut que $\langle x, Ay\rangle = \langle x, By\rangle \ \forall x, y \in \mathcal{H} \iff \langle x, (A-B)y\rangle = 0 \ \forall x, y \in \mathcal{H}$. Fixons un élément $y \in \mathcal{H}$ quelconque et écrivons $u = (A-B)y \in \mathcal{H}$, alors $\langle x, u\rangle = 0 \ \forall x \in \mathcal{H}$ est vrai si et seulement si $u = 0$, i.e. $(A-B)y = 0 \ \forall y \in \mathcal{H}$, i.e. $A - B = \mathbf{0}$, mais alors $A = B$.

La preuve de la deuxième équivalence est analogue : il suffit de considérer $u = (A-B)x$ au lieu de $u = (A-B)y$. □

6.2.1 Un exemple (classique) d'opérateur linéaire non borné sur un espace vectoriel de dimension infinie

Malgré le fait qu'on va s'occuper seulement des opérateurs linéaires bornés sur un espace de Hilbert, il est impératif de montrer au moins un exemple d'opérateur linéaire non borné : on va vérifier qu'une des plus simples opérations, celle de dérivation, faite sur la plus simple des bases hilbertiennes, celle de Fourier, ne donne pas lieu à un opérateur borné !

En fait, soit $u_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx}$, $(n \in \mathbb{Z})$, la base hilbertienne de Fourier de $L^2[0, 2\pi]$. Considérons l'opérateur de dérivation première sur l'espace vectoriel (de *dimension infinie*) engendré par la base hilbertienne de Fourier de $L^2[0, 2\pi]$:

$$\begin{array}{rccl} D: & \mathrm{span}((u_n)_{n\in\mathbb{Z}}) & \longrightarrow & L^2[0,2\pi] \\ & u_n & \longmapsto & Du_n = \frac{d}{dx}u_n, \end{array}$$

où $\frac{d}{dx}u_n(x) = \frac{in}{\sqrt{2\pi}}e^{inx}$, qui est intégrable en module carré sur $[0, 2\pi]$.

On montera que la norme de D n'est pas finie. Pour calculer sa norme, on peut choisir la définition (6.5) de norme opératorielle, i.e.

$$\|D\| = \sup_{\|u_n\|=1} \|Du_n\|,$$

mais la condition $\|u_n\| = 1$ ne détermine aucune contrainte car tout élément u_n de la base hilbertienne de Fourier de $L^2[0, 2\pi]$ a norme unitaire, donc $\|D\|$ est simplement le sup de l'ensemble de valeurs $\|Du_n\|$ par rapport à l'indice entier n, i.e. :

$$\|D\| = \sup_{n\in\mathbb{Z}} \|Du_n\| = \sup_{n\in\mathbb{Z}} \left(\int_0^{2\pi} \left| \frac{in}{\sqrt{2\pi}} e^{inx} \right|^2 dx \right)^{1/2} = \sup_{n\in\mathbb{Z}} \left(|in|^2 \int_0^{2\pi} \left| \frac{1}{\sqrt{2\pi}} e^{inx} \right|^2 dx \right)^{1/2},$$

i.e.

$$\|D\| = \sup_{n\in\mathbb{Z}} |in| = \sup_{n\in\mathbb{Z}} n = +\infty,$$

ce qui implique que l'opérateur de dérivation définit ci-dessus n'est pas borné, et, donc, pas continu.

6.3 Inversibilité des opérateurs linéaires

Dans l'exercice 6.3 nous avons pu apprécier l'importance de l'analyse de l'inverse d'un opérateur linéaire. Dans cette section nous allons examiner en profondeur ce sujet.

Déf. 6.3.1 *Soient V, W deux espaces vectoriels normés sur le même corps $\mathbb{K}$ et soit $A : V \to \mathrm{Im}(A) \subseteq W$ linéaire. L'opérateur inverse de A est l'opérateur $A^{-1} : \mathrm{Im}(A) \subseteq W \to V$ tel que $\forall x \in V$,*

$$\begin{array}{rccl} A^{-1}: & \mathrm{Im}(A) \subseteq W & \longrightarrow & V \\ & Ax & \longmapsto & A^{-1}(Ax) = x. \end{array}$$

Si A^{-1} existe, on dit que A est inversible.

Pour tout $x \in V$, ça vaut que $A^{-1}(Ax) = x$ et que $A(A^{-1}(Ax)) = A(x)$, donc l'inversibilité de A peut être définie d'une manière équivalente avec les conditions

$$A^{-1} \circ A = id_V \text{ et } A \circ A^{-1} = id_{\mathrm{Im}(A)}.$$

Dans le cas particulier $W = V$ et $\mathrm{Im}(A) = V$, l'inversibilité de A est équivalente à l'existence d'un opérateur $A^{-1} : V \to V$ tel que :

$$A \circ A^{-1} = A^{-1} \circ A = id_V.$$

Si $A : V \to V$, on utilise le symbole $\mathrm{GL}(V)$ pour dénoter *l'ensemble des opérateurs linéaires continus et bijectifs avec inverse continue*, qu'on appelle **l'ensemble des éléments réguliers de** V.

Dans le théorème suivant on rappelle des propriétés élémentaires de l'inverse (les preuves sont identiques à celles déjà connues dans le cas de dimension finie).

Théorème 6.3.1 *Soient V, W deux espaces vectoriels normés et soit $A : V \to \mathrm{Im}(A) \subseteq W$ linéaire.*

1. *Si A^{-1} existe, alors il est unique.*
2. *Si A^{-1} existe, alors il est un opérateur linéaire.*
3. *A^{-1} existe si et seulement si $\ker(A) = \{0_V\}$, i.e. condition nécessaire et suffisante pour l'inversibilité de A sur son image est que son noyau soit réduit au vecteur nul de V.*

Preuve.

1. Supposons que $B_1, B_2 : \mathrm{Im}(A) \subseteq W \to V$ soient deux opérateurs inverses de A, alors : $B_1 = B_1 \circ id_{\mathrm{Im}(A)} = B_1 \circ (A \circ B_2) = (B_1 \circ A) \circ B_2 = id_{\mathrm{Im}(A)} \circ B_2 = B_2$.

2. Pour tout $w_1, w_2 \in \mathrm{Im}(A)$ et $k \in \mathbb{K}$ on a :

$$\begin{aligned} A^{-1}(w_1 + kw_2) &= A^{-1}(AA^{-1}(w_1) + kAA^{-1}(w_2)) \\ &\quad \text{(linéarité de } A\text{)} \\ &= A^{-1}A(A^{-1}(w_1) + kA^{-1}(w_2)) \\ &= A^{-1}(w_1) + kA^{-1}(w_2). \end{aligned}$$

3. Nous savons qu'il est possible définir l'inverse de A sur son image si et seulement si A est injectif. Vérifions que cela est équivalent à $\ker(A) = 0_V$. D'un côté, si $Ax = 0_W$, alors $x = A^{-1}Ax = A^{-1}0_W = 0_V$ par linéarité de A^{-1}, donc si A^{-1} existe, le noyau de A est réduit au vecteur nul de V. Vice-versa, supposons $\ker(A) = \{0_V\}$ et soient $x_1, x_2 \in V$ tels que $Ax_1 = Ax_2$, alors $Ax_1 - Ax_2 = 0_W$, i.e. par linéarité de A, $A(x_1 - x_2) = 0_W$, mais si $\ker(A) = \{0_V\}$ alors $x_1 - x_2 = 0_V$, i.e. $x_1 = x_2$, ce qui prouve l'injectivité de A. □

La condition $\ker(A) = \{0_V\}$ est nécessaire et suffisante pour l'inversibilité d'un opérateur linéaire sur son espace image $\mathrm{Im}(A)$ en dimension finie et infinie. De plus, en dimension finie, l'inverse d'un opérateur linéaire, s'il existe, il est toujours borné.

Par contre, en dimension infinie, la condition $\ker(A) = \{0_V\}$ n'implique aucune relation entre la continuité de A et celle de A^{-1} : A **peut être borné et avoir un inverse non borné** ou, vice-versa, A **peut être non borné avec un inverse borné** ! Un exemple classique de cette situation est donné par les opérateurs de dérivation et d'intégration. Voyons un exemple explicite : soit $A : \ell^2(\mathbb{N}, \mathbb{K}) \to \ell^2(\mathbb{N}, \mathbb{K})$ l'opérateur linéaire défini par $A(x_1, x_2, x_3, \dots, x_n, \dots) = (x_1, x_2/2, x_3/3, \dots, x_n/n, \dots)$, i.e. $A((x_n)_{n\in\mathbb{N}^*}) = (x_n/n)_{n\in\mathbb{N}^*}$. A est borné et $\|A\| \leqslant 1$, en fait, pour tout $x = (x_n)_{n\in\mathbb{N}} \in \ell^2(\mathbb{N}, \mathbb{K})$:

$$\|Ax\|_{\ell^2}^2 = \sum_{n\in\mathbb{N}} \frac{|x_n|^2}{n^2} \leqslant \sum_{n\in\mathbb{N}} |x_n|^2 = \|x\|_{\ell^2}^2.$$

L'opérateur $A^{-1} : \ell^2(\mathbb{N}, \mathbb{K}) \to \ell^2(\mathbb{N}, \mathbb{K})$, $A^{-1}((y_n)_{n\in\mathbb{N}^*}) = (ny_n)_{n\in\mathbb{N}^*}$ est évidemment l'inverse de A. Néanmoins, A^{-1} n'est pas borné : pour le vérifier, il suffit de considérer l'élément général de la base canonique de $\ell^2(\mathbb{N}, \mathbb{K})$, i.e. $e_n = (0, 0, \dots, 1, 0, 0, \dots)$ avec 1 dans la position n, et d'observer que, d'un côté $\|e_n\|_{\ell^2} = 1$ $\forall n \in \mathbb{N}$ et, de l'autre côté, $\|A^{-1}e_n\|_{\ell^2} = n$, donc $\|A^{-1}\| = \sup_{n\in\mathbb{N}} \|A^{-1}e_n\|_{\ell^2} = +\infty$.

Il existe une caractérisation très utile de l'inversibilité bornée d'un opérateur linéaire. Il faut souligner que cette caractérisation vaut indépendamment de la continuité de l'opérateur, ceci est très utile dans les applications.

Théorème 6.3.2 (Inversibilité bornée d'un opérateur linéaire) *Si V et W sont deux espaces vectoriels normés et si $A : V \to W$ est un opérateur linéaire (non nécessairement borné), alors $\exists A^{-1} \in \mathcal{B}(\mathrm{Im}(A), V)$ si et seulement si $\exists \mu > 0$ tel que $\|Ax\| \geqslant \mu\|x\|$ $\forall x \in V$.*

Preuve.

$\boxed{\Rightarrow}$: supposons que $\exists A^{-1} \in \mathcal{B}(\mathrm{Im}(A), W)$, alors, par définition $\exists m > 0$ tel que $\forall y \in \mathrm{Im}(A)$: $\|A^{-1}y\| \leqslant m\|y\|$. Comme A est inversible et $y \in \mathrm{Im}(A)$, $\exists x \in V$ tel qu'on peut écrire $y = Ax$, donc $\underbrace{\|A^{-1}Ax\|}_{\|x\|} \leqslant m\|Ax\|$, i.e. $\|Ax\| \geqslant \underbrace{\frac{1}{m}}_{=\mu>0}\|x\|$ et, comme y est un élément arbitraire de $\mathrm{Im}(A)$, l'inégalité vaut pour tout $x \in V$.

$\boxed{\Leftarrow}$: vice-versa, soit $\|Ax\| \geqslant \underset{(>0)}{\mu}\|x\|$ $\forall x \in V$, alors, en particulier, si $x \in \ker(A)$,

$$\|Ax\| = \|0\| = 0 \geqslant \mu\|x\| \underset{(\mu>0)}{\iff} \|x\| = 0 \iff x = 0_V \implies \ker(A) = \{0_V\},$$

i.e. $\exists A^{-1} : \mathrm{Im}(A) \to V$. Ça reste à démontrer que A^{-1} est borné. Pour tout $y \in \mathrm{Im}(A)$ tel que $x = A^{-1}y$, alors : $\|Ax\| \geqslant \mu\|x\| \iff \|AA^{-1}y\| \geqslant \mu\|A^{-1}y\| \iff \|y\| \geqslant \mu\|A^{-1}y\| \iff \|A^{-1}y\| \leqslant \frac{1}{\mu}\|y\|$, $\forall y \in \mathrm{Im}(A)$, i.e. A^{-1} est borné. □

L'interprétation de la condition du théorème est la suivante : tout d'abord, le fait que $\|Ax\| \geqslant \mu\|x\|$ permet de garantir que le noyau de A est réduit au seul vecteur nul. De plus, l'inégalité $\|Ax\| \geqslant \mu\|x\|$ est inversée par rapport à celle qui définit un opérateur borné, elle est donc bien adaptée pour permettre à l'opérateur inverse de A d'être borné !

Une conséquence immédiate du théorème qu'on vient de voir est qu'un opérateur linéaire $A : V \to W$ satisfait les deux conditions : $A \in \mathcal{B}(V, W)$ et $\exists\ A^{-1} \in \mathcal{B}(\mathrm{Im}(A), V)$ si et seulement s'il satisfait la condition :

$$\exists a, b > 0,\ a \leqslant b\ :\ \boxed{a\|x\| \leqslant \|Ax\| \leqslant b\|x\|} \quad \forall x \in V,$$

i.e. la norme de tous les vecteurs de V, transformés par l'action de A, est bornée par la norme du vecteur lui-même multipliée par deux constantes positives. Cette considération a une conséquence importante sur les images des opérateurs linéaires bornés définis sur les espaces de Banach.

Théorème 6.3.3 *Soit V un espace de Banach et W un espace vectoriel normé quelconque. Soit $A \in \mathcal{B}(V, W)$. Si A est inversible avec inverse borné, alors* $\mathrm{Im}(A)$ *est un sous-espace vectoriel fermé de W.*

Preuve. Grâce au théorème 6.3.2, la condition $\exists A^{-1} \in \mathcal{B}(\mathrm{Im}(A), V)$ est équivalente à

$$\exists a > 0\ :\ \|x\| \leqslant a\|Ax\| \quad \forall x \in V,$$

il faut démontrer que cette condition implique que $\mathrm{Im}(A)$ est fermé, i.e. que si $(y_n)_{n\in\mathbb{N}} \subset \mathrm{Im}(A)$ est telle que $y_n \underset{n\to+\infty}{\to} y$, alors $y \in \mathrm{Im}(A)$. Comme $y_n \in \mathrm{Im}(A)$, alors il existe $(x_n)_{n\in\mathbb{N}} \subset V$ tel que $y_n = Ax_n \ \forall n \in \mathbb{N}$, donc :

$$\|x_n - x_m\| \leqslant a \, \|A(x_n - x_m)\| = \|Ax_n - Ax_m\| = \|y_n - y_m\| \underset{n,m\to+\infty}{\to} 0,$$

car $(y_n)_{n\in\mathbb{N}}$ est une suite convergente, et donc elle est de Cauchy, mais alors aussi la suite $(x_n)_{n\in\mathbb{N}}$ est de Cauchy et, comme V est de Banach, il existe $x \in V$ tel que $x_n \underset{n\to+\infty}{\to} x$. Grâce à la continuité de A on obtient :

$$Ax = A \lim_{n\to+\infty} x_n = \lim_{n\to+\infty} Ax_n = \lim_{n\to+\infty} y_n = y,$$

i.e. $y \in \mathrm{Im}(A)$. □

On termine cette section avec une deuxième condition suffisante pour la continuité de l'inverse d'un opérateur linéaire. Pour cela on a besoin d'un résultat intermédiaire qui est l'un plus importantes théorèmes de l'analyse fonctionnelle (on assumera la démonstration, qui est une conséquence d'un autre théorème très important : le théorème de catégorie de Baire).

Théorème 6.3.4 (Théorème de Banach-Schauder - application ouverte) *Soient V et W deux espaces de Banach. Si $A \in \mathcal{B}(V, W)$ est surjectif, alors A est une application ouverte, i.e. A transforme sous-ensembles ouverts de V en sous-ensembles ouverts de W.*

Théorème 6.3.5 (Théorème de l'inverse continu en espaces de Banach) *Soient V et W deux espaces de Banach. Si $A \in \mathcal{B}(V, W)$ est bijectif, i.e. $\ker(A) = \{0_V\}$ et A est surjectif, alors $A^{-1} \in \mathcal{B}(W, V)$, i.e. A^{-1} est continu.*

Preuve. On rappelle que la caractérisation topologique de la continuité est la suivante : une fonction entre deux espaces topologiques est continue si et seulement si la contre-image de tout sous-ensemble ouvert est un ouvert. Par définition, les contre-images de A^{-1} sont les images de A, donc A^{-1} est continu si et seulement si toute image via A d'un ouvert est ouvert, mais ça est garanti par le théorème de Banach-Schauder. □

Le théorème de l'inverse continu permet de caractériser les opérateurs qui appartiennent à l'ensemble $\mathrm{GL}(V)$ pour un espace de Banach V quelconque.

Théorème 6.3.6 (Caracterisation de GL(V)) *Soit V un espace de Banach et $GL(V)$ l'ensemble des éléments réguliers de l'algèbre de Banach $\mathcal{B}(V)$ (bijections linéaires avec inverse continu). Pour un opérateur $A \in \mathcal{B}(V)$ les conditions suivantes sont équivalentes :*

1. $A \in \mathrm{GL}(V)$.

2. $\exists$ *un opérateur linéaire B définit sur tout V tel que : $BA = id_V$ et $AB = id_V$.*

Le cas échéant, B est unique et $B = A^{-1}$.

Preuve.

1. $\implies$ 2. Si $A \in \mathrm{GL}(V)$, alors ça suffit de considérer $B = A^{-1}$ pour prouver l'implication.

2. $\Longrightarrow$ **1.** L'hypothèse $BA = id_V$ implique que $\ker(A) = \{0\}$, i.e. que A est injectif. En fait, si, par l'absurde $x \neq 0$ et $Ax = 0$, alors $BAx = 0$, ce qui est en contradiction avec le fait que $BAx = id_V(x) = x \neq 0$. De plus, l'hypothèse $AB = id_V$ implique que $\mathrm{Im}(A) = V$, en fait, pour tout $x \in V$ ça vaut $A(Bx) = AB(x) = id_V(x) = x$, donc tout $x \in V$ peut être vu comme le transformé d'un élément de V, i.e. Bx, via A, i.e. A est surjectif. Donc l'existence de B tel que les deux hypothèses $BA = id_V$ et $AB = id_V$ sont valides implique que A est une bijection linéaire et que $\forall x \in V$, $B(Ax) = x$, i.e. $B = A^{-1}$. Donc A est borné par hypothèse, inversible et surjectif, grâce au théorème de l'inverse continu $B = A^{-1}$ est continu, et alors $A \in \mathrm{GL}(V)$.

Il reste à démontrer l'unicité : soient B et B' deux opérateurs qui vérifient la **2.**, alors $A^{-1} = A^{-1}AB$ et $A^{-1} = A^{-1}AB'$ et donc $A^{-1} = B = B'$. □

Évidemment, si $A \in \mathrm{GL}(V)$, alors aussi $A^{-1} \in \mathrm{GL}(V)$ et si $A, B \in \mathrm{GL}(V)$, alors $AB \in \mathrm{GL}(V)$ car $(AB)^{-1} = B^{-1}A^{-1}$ vu que $ABB^{-1}A^{-1} = id_V$ et $B^{-1}A^{-1}AB = id_V$. Donc, $\mathrm{GL}(V)$ est stable par rapport au produit et à l'inversion et son élément unité est id_V, i.e. $\mathrm{GL}(V)$ est un groupe.

Déf. 6.3.2 *Le groupe $GL(V)$ est dit le groupe général linéaire de V.*

6.4 L'espace dual d'un espace de Hilbert et le théorème de représentation de Riesz

Reconsidérons $\mathcal{B}(V, W)$, où V, W sont deux espaces vectoriels normés. On a vu que $\mathcal{B}(V, W)$ est un espace de Banach par rapport à la norme opératorielle si W est un espace de Banach. Allons considérer le cas particulier dans lequel W est le corps $\mathbb{K}$ sur lequel V est défini en tant qu'espace vectoriel.

Comme $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ est complet, $\mathcal{B}(V, \mathbb{K})$ est un espace de Banach, qu'on appelle le **dual** de V et qu'on écrit V^* (parfois dans la littérature mathématique on trouve la notation V' pour l'espace dual). On appelle les éléments de V^* les « **fonctionnelles linéaires bornés** sur V ».

On peut se demander comment le processus de « dualisation » de V peut s'itérer. En fait, il est possible de démontrer que, *pour tout espace de Hilbert $\mathcal{H}$, la dualisation est une involution*, i.e. $\mathcal{H}^{**} \simeq \mathcal{H}$, où $\simeq$ est un isomorphisme entre espaces de Hilbert. $\mathcal{H}^{**}$ est appelé le **bidual** de $\mathcal{H}$.

Ceci n'est plus vrai pour les espaces de Banach, ceux qui ont la propriété d'être isomorphes à leur bidual sont dits **espaces de Banach réflexifs**. Les espaces de Banach $L^p(X, \mathcal{A}, \mu)$ sont réflexifs pour $1 < p < \infty$, mais $L^1(X, \mathcal{A}, \mu)$ et $L^\infty(X, \mathcal{A}, \mu)$ ne le sont pas.

Chaque fonctionnelle $\varphi \in V^*$ transforme un élément de V en un scalaire de $\mathbb{K}$, la notation pour représenter cette transformation est la suivante :

$$\begin{array}{rccc} \varphi : & V & \longrightarrow & \mathbb{K} \\ & x & \longmapsto & \varphi(x) = \langle \varphi, x \rangle, \end{array}$$

la notation $\langle \varphi, x \rangle$ vient du fait que si V est un espace de Hilbert, alors toute fonctionnelle linéaire continue $\varphi \in V^*$ agit sur les vecteurs de V comme un produit scalaire ! Cette affirmation est la thèse d'un célèbre résultat de Riesz qu'on énonce et démontre ci-dessous.

Théorème 6.4.1 (Représentation de Riesz) *Soient $\mathcal{H}$ un espace de Hilbert sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, et $\mathcal{H}^*$ le dual de $\mathcal{H}$. Alors :*

$$\begin{array}{rccc} T: & \mathcal{H} & \longrightarrow & \mathcal{H}^* \\ & x & \longmapsto & T_x, \end{array}$$

où :

$$\begin{array}{rccc} T_x: & \mathcal{H} & \longrightarrow & \mathbb{K} \\ & y & \longmapsto & T_x(y) = \langle y, x \rangle \end{array}$$

est un isomorphisme entre $\mathcal{H}$ et $\mathcal{H}^*$ interprétés comme espaces de Banach, i.e. T est bijective, préserve les normes et :

— *si $\mathbb{K} = \mathbb{R}$, alors T est linéaire ;*

— *si $\mathbb{K} = \mathbb{C}$, alors T est anti-linéaire.*

La fonctionnelle T_x est dite le « **représentant de Riesz** » de x en $\mathcal{H}^*$.

Avant de donner la preuve, allons comprendre la raison de l'anti-linéarité dans le cas $\mathbb{K} = \mathbb{C}$. On commence à analyser l'opération de somme :

$$\begin{array}{rccc} T: & \mathcal{H} & \longrightarrow & \mathcal{H}^* \\ & x_1 + x_2 & \longmapsto & T_{x_1+x_2}, \end{array}$$

$$\begin{array}{rccc} T_{x_1+x_2}: & \mathcal{H} & \longrightarrow & \mathbb{C} \\ & y & \longmapsto & T_{x_1+x_2}(y) = \langle y, x_1 + x_2 \rangle = \langle y, x_1 \rangle + \langle y, x_2 \rangle = T_{x_1}(y) + T_{x_2}(y), \end{array}$$

donc $\boxed{T_{x_1+x_2} = T_{x_1} + T_{x_2}}$.

On considère maintenant l'opération de produit par un scalaire $k \in \mathbb{C}$:

$$\begin{array}{rccc} T: & \mathcal{H} & \longrightarrow & \mathcal{H}^* \\ & kx & \longmapsto & T_{kx}, \end{array}$$

$$\begin{array}{rccc} T_{kx}: & \mathcal{H} & \longrightarrow & \mathbb{C} \\ & y & \longmapsto & T_{kx}(y) = \langle y, kx \rangle = \bar{k}\langle y, x \rangle = \bar{k} T_x(y), \end{array}$$

donc $\boxed{T_{kx} = \bar{k} T_x}$.

Par conséquent,

$$\begin{array}{rccc} T: & \mathcal{H} & \longrightarrow & \mathcal{H}^* \\ & x_1 + x_2 & \longmapsto & T_{x_1} + T_{x_2}, \end{array} \qquad \begin{array}{rccc} T: & \mathcal{H} & \longrightarrow & \mathcal{H}^* \\ & kx & \longmapsto & \bar{k} T_x \end{array}$$

ceci explique pourquoi T est anti-linéaire si $\mathbb{K} = \mathbb{C}$. Bien évidemment, si $\mathbb{K} = \mathbb{R}$ cette distinction n'existe plus et T est linéaire.

Le théorème de représentation de Riesz est appelé comme ça car il permet de représenter toutes les fonctionnelles linéaires continues sur un espace de Hilbert via des produits scalaires, en particulier, pour toute fonctionnelle linéaire continue φ sur $\mathcal{H} = L^2(X, \mathcal{A}, \mu)$ existe un seul élément $f \in L^2(X, \mathcal{A}, \mu)$ tel que $\varphi = T_f$ avec :

$$\begin{array}{rccc} T_f: & L^2(X, \mathcal{A}, \mu) & \longrightarrow & \mathbb{K} \\ & g & \longmapsto & T_f(g) = \langle g, f \rangle = \int_X g \bar{f} d\mu. \end{array}$$

Plus en général, on sait que tout espace de Hilbert séparable de dimension infinie est isomorphe à $\ell^2(\mathbb{N}, \mathbb{K})$, qui a un produit scalaire défini par une série.

Ces observations expliquent pourquoi, **dans les applications de l'analyse fonctionnelle à des problèmes concrets, les fonctionnelles linéaires continues sont très souvent représentées par des sommes finies, des séries ou des intégrales !**

Une dernière observation avant de passer à la preuve du théorème : si on considère le produit scalaire « *à la physicien* », i.e. anti-linéaire par rapport à la première entrée et linéaire par rapport à la deuxième, alors la définition de T_x devient : $T_x(y) = \langle x, y\rangle$.

Preuve. Ayant déjà examiné le caractère linéaire ou anti-linéaire de T, on commence la preuve par vérifier la *bonne définition* de T, i.e. que T_x est une fonctionnelle linéaire bornée sur $\mathcal{H}$. Soient $\alpha, \beta \in \mathbb{K}$, $y, y_1, y_2 \in \mathcal{H}$:

— T_x est linéaire [3] :

$$T_x(\alpha y_1 + \beta y_2) = \langle \alpha y_1 + \beta y_2, x\rangle = \alpha\langle y_1, x\rangle + \beta\langle y_2, x\rangle = \alpha T_x(y_1) + \beta T_x(y_2).$$

— T_x est borné. Pour prouver ça, on observe que $\|T_x(y)\| = |T_x(y)|$ car $T_x(y) \in \mathbb{K}$! Donc :

$$\|T_x(y)\| = |T_x(y)| = |\langle y, x\rangle| \underset{\text{(Cauchy-Schwarz)}}{\leqslant} \|x\|\|y\|. \tag{6.16}$$

Le fait que T_x soit un opérateur linéaire borné entre les espaces de Hilbert $\mathcal{H}$ et $\mathbb{K}$ nous autorise à calculer la norme opératorielle de T_x. C'est par rapport à cette norme que T *est une isométrie*, i.e. $\|T_x\|_{\mathcal{B}(\mathcal{H})} = \|x\|_{\mathcal{H}}\ \forall x \in \mathcal{H}$. On va éliminer tout de suite le cas banal du vecteur nul : si $x = 0_{\mathcal{H}}$ alors $T_{0_{\mathcal{H}}}$ est la fonctionnelle nulle car $T_{0_{\mathcal{H}}}(y) = \langle y, 0_{\mathcal{H}}\rangle = 0\ \forall y \in \mathcal{H}$, donc : $\|0_{\mathcal{H}}\| = 0 = \|T_{0_{\mathcal{H}}}\|$.

Soit donc $x \in \mathcal{H}$, $x \neq 0_{\mathcal{H}}$, on va prouver que $\|T_x\| \leqslant \|x\|$ et que $\|x\| \leqslant \|T_x\|$, dans cet ordre.

— $\|T_x\| \leqslant \|x\|$: grâce à (6.16) on peut écrire $\|T_x(y)\| \leqslant \|x\|\|y\|\ \forall y \in \mathcal{H}$, donc

$$\|T_x\| = \sup_{\|y\|=1} |T_x(y)| \leqslant \sup_{\|y\|=1} \|x\|\|y\| = \|x\|.$$

— $\|x\| \leqslant \|T_x\|$: l'astuce ici consiste à écrire

$$\|x\|^2 = \langle x, x\rangle \underset{\text{(déf de } T_x)}{=} T_x(x) \underset{T_x(x)=\|x\|^2 \geqslant 0\,!}{=} |T_x(x)| = \|T_x(x)\| \underset{(T_x \text{ borné})}{\leqslant} \|T_x\|\|x\|,$$

mais comme $\|x\| \neq 0$, on peut diviser par $\|x\|$ le premier et le dernier membre de l'expression ci-dessus, en obtenant $\|x\| \leqslant \|T_x\|$.

En résumé, $\|T_x\| = \|x\|\ \forall x \in \mathcal{H}$, donc T est une isométrie et, par conséquent, T est injective.

La dernière étape de la preuve consiste en démontrer que T *est surjective*, i.e. que pour tout $\varphi \in \mathcal{H}^*$ il existe un $x \in \mathcal{H}$ tel que $\varphi = T_x$. L'argument utilisé par Riesz pour démontrer la surjectivité de T est particulièrement élégant.

Tout d'abord, si φ est la fonctionnelle identiquement nulle $\mathbf{0}$, alors $\varphi = T_0$.

Soit alors φ une fonctionnelle non-identiquement nulle et *analysons son noyau* :

3. Si on avait définit $T_x(y) = \langle x, y\rangle$, alors on aurait $T_x(y)$, i.e. $T_x(\alpha y_1 + \beta y_2) = \bar{\alpha}\langle x, y_1\rangle + \bar{\beta}\langle x, y_2\rangle = \bar{\alpha}T_x(y_1) + \bar{\beta}T_x(y_2)$, i.e. T_x serait une fonctionnelle anti-linéaire. Par conséquent, on ne peut pas échapper à l'antilinéarité, soit de T, soit de T_x.

— $0_{\mathcal{H}} \in \ker(\varphi)$ par linéarité de φ, donc $\ker(\varphi) \neq \varnothing$;
— vu que $\varphi \neq \mathbf{0}$, il existe au moins un vecteur de $\mathcal{H}$ qui n'est pas annulé par φ, i.e. $\ker(\varphi) \neq \mathcal{H}$;
— comme on l'a vu dans le théorème 6.2.4, $\ker(\varphi)$ est toujours fermé.

Par conséquent, $\ker(\varphi)$ *est un sous-espace propre et fermé de* $\mathcal{H}$, grâce à cette observation on peut utiliser le théorème 5.2.3 pour assurer que $\ker(\varphi)^{\perp} \neq \{0_{\mathcal{H}}\}$, i.e. il existe au moins un $u \neq 0_{\mathcal{H}}$, $u \in \ker(\varphi)^{\perp}$.

Nous observons maintenant que, comme $\ker(\varphi) \cap \ker(\varphi)^{\perp} = \{0_{\mathcal{H}}\}$ et comme $u \neq 0_{\mathcal{H}}$, $u \notin \ker(\varphi)$, pour tout $y \in \mathcal{H}$ il est bien défini le vecteur $z = y - \frac{\varphi(y)}{\varphi(u)}u$.

$z \in \ker(\varphi)$, en fait, par linéarité, $\varphi(z) = \varphi(y - \frac{\varphi(y)}{\varphi(u)}u) = \varphi(y) - \frac{\varphi(y)}{\cancel{\varphi(u)}}\cancel{\varphi(u)} = 0$, en résumé :

$$\begin{cases} u \in \ker(\varphi)^{\perp} \\ z = y - \frac{\varphi(y)}{\varphi(u)}u \in \ker(\varphi), \end{cases}$$

d'où :

$$0 = \langle z, u\rangle = \langle y - \frac{\varphi(y)}{\varphi(u)}u, u\rangle = \langle y, u\rangle - \langle \frac{\varphi(y)}{\varphi(u)}u, u\rangle = \langle y, u\rangle - \frac{\varphi(y)}{\varphi(u)}\|u\|^2,$$

i.e. :

$$\varphi(y) = \frac{\varphi(u)}{\|u\|^2}\langle y, u\rangle = \langle y, \frac{\overline{\varphi(u)}}{\|u\|^2}u\rangle \qquad \forall y \in \mathcal{H}.$$

Par conséquent, pour tout vecteur $u \in \ker(\varphi)^{\perp}$, $u \neq 0_{\mathcal{H}}$, le vecteur $x = \frac{\overline{\varphi(u)}}{\|u\|^2}u$, est tel que

$$\varphi(y) = \langle y, x\rangle = T_x(y), \qquad \forall y \in \mathcal{H},$$

i.e. $\varphi = T_x$. Ceci montre que T est surjective et ça termine la preuve. □

Pendant la partie finale de la preuve, en réalité, on a prouvé intrinsèquement un résultat encore plus fin : **le complément orthogonal du noyau d'une fonctionnelle linéaire bornée sur un espace de Hilbert $\mathcal{H}$ est une droite en $\mathcal{H}$.**

Corollaire 6.4.1 *Soit $\mathcal{H}$ un espace de Hilbert et $\varphi \in \mathcal{H}^*$, $\varphi \not\equiv \mathbf{0}$. Alors* $\ker(\varphi)^{\perp}$ *est un sous-espace vectoriel mono-dimensionnel de $\mathcal{H}$, i.e.* $\dim(\ker(\varphi)^{\perp}) = 1$. *Un générateur de cet espace est le vecteur résidu $x - P_{\ker\varphi}x$, où $x \in \mathcal{H}$ est tel que $\varphi = T_x$ via isomorphisme de Riesz.*

Preuve. Dans la partie finale de la preuve du théorème de représentation de Riesz on a démontré que si φ n'est pas la fonctionnelle identiquement nulle, alors, pour n'importe quel $u \in \ker(\varphi)^{\perp}$, $u \neq 0_{\mathcal{H}}$, $x = \frac{\overline{\varphi(u)}}{\|u\|^2}u$ est le vecteur de $\mathcal{H}$ qu'on identifie avec φ via la formule $\varphi = T_x$.

Si, par l'absurde, $\ker(\varphi)^{\perp}$ a dimension supérieure à 1, alors il existe au moins un autre générateur, disons $u' \neq u$, $u' \neq 0_{\mathcal{H}}$, $u' \in \ker(\varphi)^{\perp}$, u et u' linéairement indépendants. Vu que $\ker(\varphi)^{\perp}$ est un espace vectoriel, on peut appliquer l'algorithme de Gram-Schmidt pour orthonormaliser le couple (u, u') et obtenir le couple $(\tilde{u}, \tilde{u}') \in \ker(\varphi)^{\perp} \times \ker(\varphi)^{\perp}$, $\|\tilde{u}\| = \|\tilde{u}'\| = 1$ et $\tilde{u} \perp \tilde{u}'$. Allons définir les vecteurs

$$x = \frac{\overline{\varphi(\tilde{u})}}{\|\tilde{u}\|^2}\tilde{u} = \overline{\varphi(\tilde{u})}\tilde{u}, \quad x' = \frac{\overline{\varphi(\tilde{u}')}}{\|\tilde{u}'\|^2}\tilde{u}' = \overline{\varphi(\tilde{u}')}\tilde{u}',$$

qui sont eux-mêmes orthogonaux, donc on peut utiliser le théorème de Pythagore pour estimer la norme au carré de leur différence :

$$\|x - x'\|^2 = \|x + (-x')\|^2 = \|x\|^2 + \|x'\|^2 = |\varphi(\tilde{u})|^2\|\tilde{u}\|^2 + |\varphi(\tilde{u}')|^2\|\tilde{u}'\|^2 > 0,$$

car $\varphi(\tilde{u})$, $\varphi(\tilde{u}')$ et les normes de $\tilde{u}$ et de $\tilde{u}'$ sont $\neq 0$. Par conséquent, $x \neq x'$ et alors on aurait deux vecteurs différents de $\mathcal{H}$, x et x', associés à la même fonctionnelle $\varphi \in \mathcal{H}^*$, ce qui est en contradiction avec l'injectivité de la transformation de Riesz.

De plus, vu que $x = \frac{\overline{\varphi(u)}}{\|u\|^2}u$ et $u \in \ker(\varphi)^\perp$, $x \notin \ker\varphi$ et alors le théorème 5.2.3 nous assure que le vecteur résidu de la projection orthogonale de x sur $\ker\varphi$, i.e. $x - P_{\ker\varphi}x$, appartient à $\ker(\varphi)^\perp$. □

Remarque 6.4.1

À la lumière de ce qu'on vient de discuter, on peut expliciter l'inverse de la transformation de Riesz comme ça :

$$\begin{array}{rccl} T^{-1}: & \mathcal{H}^* & \longrightarrow & \mathcal{H} \\ & \varphi & \longmapsto & T^{-1}(\varphi) = x = \frac{\overline{\varphi(u)}}{\|u\|^2}u, \end{array}$$

où $u \neq 0_{\mathcal{H}}$ est un vecteur quelconque de $\ker(\varphi)^\perp$. Vu que $\dim(\ker(\varphi)^\perp) = 1$, pour vérifier que la définition soit bien posée, la seule chose qu'il faut faire est de vérifier que, si $k \in \mathbb{K}$, $k \neq 0$, alors le vecteur x associé à φ via $u' = ku$ (en tant qu'élément arbitraire du sous-espace mono-dimensionnel $\ker(\varphi)^\perp$) est le même :

$$x' = \frac{\overline{\varphi(u')}}{\|u'\|^2}u' = \frac{\overline{k}\,\overline{\varphi(u)}}{|k|^2\|u\|^2}ku = \frac{\cancel{\overline{k}k}\,\overline{\varphi(u)}}{\cancel{|k|^2}\|u\|^2}u = \frac{\overline{\varphi(u)}}{\|u\|^2}u = x,$$

donc la définition de T^{-1} ne dépend pas du choix du vecteur $u \neq 0_{\mathcal{H}} \in \ker(\varphi)^\perp$.

6.4.1 Le produit scalaire induit sur le dual d'un espace de Hilbert

Dans le théorème de représentation de Riesz nous avons vu qu'un espace de Hilbert $\mathcal{H}$ et son dual $\mathcal{H}^*$ peuvent être identifiés comme espaces de Banach, vu que l'isométrie de la transformation T fait appel seulement à la norme de $\mathcal{H}$ et de $\mathcal{H}^*$. Néanmoins, nous pouvons aller encore plus loin et les identifier en tant qu'espaces de Hilbert.

Pour faire ça, on doit tout d'abord introduire un produit scalaire sur $\mathcal{H}^*$. Ceci peut être fait grâce à l'isomorphisme de Riesz $T : \mathcal{H} \to \mathcal{H}^*$: toute fonctionnelle linéaire bornée de $\mathcal{H}^*$ est l'image d'un vecteur de $\mathcal{H}$ et, vu qu'on sait comment faire le produit scalaire en $\mathcal{H}$, il n'y aura aucune ambiguïté si on définit le produit scalaire sur $\mathcal{H}^*$ comme ça :

$$\langle\varphi, \psi\rangle_{\mathcal{H}^*} := \langle T^{-1}\varphi, T^{-1}\psi\rangle_{\mathcal{H}}, \qquad \forall\varphi, \psi \in \mathcal{H}^*.$$

Le fait que T préserve la norme, garantit que cette définition de produit scalaire est compatible avec la structure d'espace de Banach déjà existant sur $\mathcal{H}^*$. En fait, si $\varphi = T_x$, i.e. φ est la fonctionnelle qui peut être identifiée avec l'image du vecteur $x \in \mathcal{H}$ via T, alors :

$$\|\varphi\|^2 = \langle T^{-1}(T_x), T^{-1}(T_x)\rangle = \langle x, x\rangle = \|x\|^2 = \|T_x\|^2,$$

où la dernière égalité est une conséquence du théorème de représentation de Riesz.

La compatibilité entre la structure d'espace vectoriel avec produit scalaire et espace vectoriel normé complet implique que $\mathcal{H}^*$, muni du produit scalaire induit par l'isomorphisme de Riesz T, est lui-même un espace de Hilbert et, comme ça, T devient un isomorphisme (anti-linéaire) entre les espaces de Hilbert $\mathcal{H}$ et $\mathcal{H}^*$.

Le théorème de représentation de Riesz est l'un des résultat les plus importantes de l'analyse fonctionnelle. Dans les deux sections suivantes, nous allons examiner une extension de ce résultat (le théorème de Lax-Milgram) et une conséquence extrêmement significative : la possibilité d'associer sans ambiguïté à chaque opérateur de $\mathcal{B}(\mathcal{H})$ un opérateur, dit adjoint, qui va jouer un rôle fondamental dans l'analyse des opérateurs de projection et unitaires.

6.5 Formes bilinéaires, sesquilinéaires et formes quadratiques associées

Le concept de forme quadratique associée à une forme bilinéaire, ou sesquilinéaire, aurait pu être introduit dans le chapitre 1, nous avons décidé de traiter seulement maintenant ce sujet parce que le lien entre opérateurs linéaires bornés dans les espaces de Hilbert et formes quadratiques amène directement à la définition d'opérateur adjoint, que l'ont verra dans la section 6.6.

Déf. 6.5.1 (Forme quadratique) *Soit $\phi : V \times V \to \mathbb{R}$ (resp. $\phi : V \times V \to \mathbb{C}$) une forme bilinéaire (resp. sésquilinéaire) sur l'espace vectoriel réel (resp. complexe) V. La fonction $\Phi : V \to \mathbb{R}$, resp. $\Phi : V \to \mathbb{C}$, définie par restriction de ϕ sur la diagonal de $V \times V$, i.e.*

$$\Phi(x) := \phi(x, x)$$

est dite la forme quadratique associée à ϕ.

Si nous rajoutons la requête de définie positivité et symétrie (resp. symétrie conjuguée), ϕ devient un produit scalaire $\langle\ ,\ \rangle$ et, dans ce cas, $\Phi(x) = \langle x, x\rangle = \|x\|^2$ pour tout $v \in V$, i.e. Φ est le carré de la norme canoniquement associée à ϕ. Cette observation explique pourquoi Φ est dite forme quadratique.

Venons-en maintenant au concept de forme bornée.

Déf. 6.5.2 *Si $(V, \|\ \|)$ est un espace vectoriel normé, alors la forme $\phi : V \times V \to \mathbb{K}$ que l'on suppose être bilinéaire si $\mathbb{K} = \mathbb{R}$ et sesquilinéaire si $\mathbb{K} = \mathbb{C}$ est dite bornée s'il existe une constante $k > 0$ telle que :*

$$|\phi(x, y)| \leqslant m\|x\|\|y\|, \qquad \forall x, y \in V.$$

Le cas échéant, la norme de ϕ est définie par la formule :

$$\|\phi\| := \inf\{m > 0\ :\ |\phi(x, y)| \leqslant m\|x\|\|y\|,\ \forall x, y \in V\}$$

Comme pour le cas des opérateurs de $\mathcal{B}(\mathcal{H})$, la norme de ϕ peut être réécrite avec une formule équivalente et très utile :

$$\|\phi\| = \sup_{\|x\|=\|y\|=1} \phi(x, y),$$

qui entraîne

$$|\phi(x, y)| \leqslant \|\phi\|\,\|x\|\,\|y\|, \qquad \forall x, y \in \mathcal{H}.$$

Déf. 6.5.3 (Formes quadratiques bornées et leur norme) *Si $(V, \| \|)$ est un espace vectoriel normé, alors la forme quadratique Φ est dite bornée s'il existe une constante $k > 0$ telle que :*

$$|\Phi(x)| \leqslant k\|x\|^2, \qquad \forall x \in V.$$

La norme d'une forme quadratique bornée est définie par

$$\|\Phi\| := \inf\{k > 0 \ : \ |\Phi(x)| \leqslant k\|x\|^2, \ \forall x \in V\}.$$

Comme pour la norme de ϕ, nous avons la possibilité de réécrire la norme de Φ comme ça :

$$\|\Phi\| := \sup_{\|x\|=1} |\Phi(x)|,$$

ce qui entraîne

$$\|\Phi(x)\| \leqslant \|\Phi\|\|x\|^2, \qquad \forall x \in V. \tag{6.17}$$

2 Comme pour le produit scalaire et sa norme, la formule de polarisation qui permet de décrire complètement une forme bilinéaire (sesquilinéaire) via sa forme quadratique associée.

Théorème 6.5.1 *Soit ϕ une forme bilinéaire (resp. sesquilinéaire) sur V et soit Φ sa forme quadratique associée. Alors, pour tout $x, y \in V$:*

$$4\phi(x, y) = \Phi(x + y) - \Phi(x - y),$$

respectivement :

$$4\phi(x, y) = \Phi(x + y) - \Phi(x - y) + i\Phi(x + iy) - i\Phi(x - iy).$$

La preuve est identique à celle déjà vue dans la section 1.2.1 où nous avons souligné que seulement la bilinéarité ou la sesquilinéarité de la forme ϕ interviennent dans la démonstration de la formule de polarisation.

Le résultat suivant, corollaire immédiat de la formule de polarisation, donne une condition équivalente à celle du théorème 6.2.6 pour les formes bilinéaires ou sesquilinéaires.

Corollaire 6.5.1 *Soient ϕ_1 et ϕ_2 deux formes bilinéaires ou sesquilinéaires sur V. Alors :*

$$\phi_1 = \phi_2 \iff \Phi_1 = \Phi_2, \quad \textit{i.e. } \phi_1(x, y) = \phi_2(x, y) \ \forall x, y \in V \iff \phi_1(x, x) = \phi_2(x, x) \ \forall x \in V,$$

i.e. l'égalité des formes quadratiques est nécessaire et suffisante pour caractériser l'égalité des formes auxquelles elles sont associées.

Allons examiner une conséquence importante de ce corollaire.

Théorème 6.5.2 *Une forme sesquilinéaire $\phi : V \times V \to \mathbb{C}$ est hermitienne si et seulement si sa forme quadratique associée Φ est réelle, i.e. $\Phi(x) \in \mathbb{R}$ $\forall x \in V$.*

Preuve. Allons prouver les deux implications :

$\boxed{\Longrightarrow}$: soit ϕ hermitienne, i.e. $\phi(x,y) = \overline{\phi(y,x)}$ $\forall x, y \in V$, alors

$$\Phi(x) = \phi(x,x) = \overline{\phi(x,x)} = \overline{\Phi(x)}, \qquad \forall x \in V,$$

i.e. Φ est réelle.

$\boxed{\Longleftarrow}$: supposons maintenant que $\Phi(x) = \overline{\Phi(x)}$ et allons définir une forme sesquilinéaire $\psi : V \times V \to \mathbb{C}$ de la manière suivante : $\psi(x,y) = \overline{\phi(y,x)}$. Si on démontre que $\psi = \phi$, on prouve que ϕ est sesquilinéaire. Pour cela, examinons la forme quadratique Ψ associée à ψ :

$$\Psi(x) = \overline{\phi(x,x)} = \overline{\Phi(x)} = \Phi(x), \qquad \forall x \in V,$$

mais, grâce au corollaire 6.5.1, $\Psi = \Phi$ implique $\psi = \phi$. □

Comme cas particulier du théorème qu'on vient de prouver, nous avons que si une forme sesquilinéaire ϕ est positive, et donc réelle, elle ne peut que être hermitienne. Cette considération donne une ultérieure justification à la définition de produit scalaire complexe qu'on a donné dans le chapitre 1.

Dans le théorème suivant on examine la relation entre la bornitude d'une forme bilinéaire ou sesquilinéaire ϕ et celle da sa forme quadratique associée.

Théorème 6.5.3 *Une forme bilinéaire ou sesquilinéaire ϕ sur un espace vectoriel normé $(V, \| \|)$ est bornée si et seulement si la forme quadratique associée Φ est bornée. De plus,*

- *si ϕ est réelle, alors : $\|\phi\| = \|\Phi\|$;*
- *si ϕ est complexe, alors sa norme est contenue dans l'intervalle de valeurs entre la norme de Φ et son double : $\|\Phi\| \leqslant \|\phi\| \leqslant 2\|\Phi\|$.*

Preuve. Nous allons prouver la première inégalité en considérant une forme bilinéaire réelle ou sesquilinéaire complexe.

$\boxed{\|\Phi\| \leqslant \|\phi\|,\ \phi \text{ réelle ou complexe}}$: par définition on a

$$\|\Phi\| = \sup_{\|x\|=1} |\Phi(x)| = \sup_{\|x\|=1} |\phi(x,x)| \underset{(*)}{\leqslant} \sup_{\|x\|=\|y\|=1} |\phi(x,y)| = \|\phi\|,$$

où $(*)$ est dû au fait que la borne supérieure est calculée sur un ensemble de valeurs plus grand. Si ϕ est bornée, alors aussi Φ est bornée et la première inégalité vaut.

$\boxed{\|\phi\| \leqslant \|\Phi\|,\ \phi \text{ bilinéaire réelle}}$: supposons maintenant que Φ soit bornée, alors, par la formule de polarisation on a

$$\begin{aligned} |\phi(x,y)| &= \frac{1}{4}|\Phi(x+y) - \Phi(x-y)| \underset{(6.17)}{\leqslant} \frac{1}{4}\|\phi\|(\|x+y\|^2 + \|x-y\|^2) \\ &\leqslant \frac{1}{4}\|\phi\| 2(\|x\|^2 + \|y\|^2) = \frac{1}{2}\|\phi\|(\|x\|^2 + \|y\|^2) \end{aligned}$$

en ayant appliqué la formule du parallélogramme $\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$. D'où :

$$\|\phi\| = \sup_{\|x\|=\|y\|=1} \frac{1}{2}|\phi(x,y)| \leqslant \sup_{\|x\|=\|y\|=1} \|\Phi\|(\|x\|^2 + \|y\|^2) = \|\Phi\|.$$

Donc Φ bornée implique ϕ bornée et ça vaut que $\|\varphi\| \leqslant \|\Phi\|$.

$\boxed{\|\phi\| \leqslant 2\|\Phi\|,\ \phi \text{ sesquilinéaire complexe}}$: supposons que Φ soit bornée, alors, encore par la formule de polarisation on a

$$\begin{aligned}|\phi(x,y)| &= \frac{1}{4}|\Phi(x+y) - \Phi(x-y) + i\Phi(x+iy) - i\Phi(x-iy)| \\ &\underset{(6.17)}{\leqslant} \frac{1}{4}\|\phi\|(\|x+y\|^2 + \|x-y\|^2 + \|x+iy\|^2 + \|x-iy\|^2).\end{aligned}$$

Par la formule du parallélogramme dans ce cas : $\|x+iy\|^2 + \|x-iy\|^2 = 2(\|x\|^2 + \|iy\|^2) = 2(\|x\|^2 + |i|^2\,\|y\|^2) = 2(\|x\|^2 + \|y\|^2)$, donc $\|x+y\|^2 + \|x-y\|^2 + \|x+iy\|^2 + \|x-iy\|^2 = 4(\|x\|^2 + \|y\|^2)$ et alors

$$|\phi(x,y)| \leqslant \|\Phi\|(\|x\|^2 + \|y\|^2),$$

ce qui implique :

$$\|\phi\| = \sup_{\|x\|=\|y\|=1} |\phi(x,y)| \leqslant \sup_{\|x\|=\|y\|=1} \|\Phi\|(\|x\|^2 + \|y\|^2) = 2\|\Phi\|.$$

Donc Φ bornée implique ϕ bornée et ça vaut que $\|\varphi\| \leqslant 2\|\Phi\|$. □

Si une forme sesquilinéaire (complexe) ϕ est aussi hermitienne, alors nous savons que sa forme quadratique associée Φ est réelle. Le théorème que l'on vient de démontrer nous assure l'égalité des normes de ϕ et de Φ quand ϕ est une forme bilinéaire réelle (et donc aussi Φ est réelle). Ces considérations nous amènent naturellement vers l'idée qu'une forme sesquilinéaire (complexe) hermitienne pourrait avoir une norme coïncidente avec celle de sa forme quadratique (réelle). Le résultat suivant nous confirme que c'est effectivement le cas.

Théorème 6.5.4 *Si une forme sesquilinéaire $\phi : V \times V \to \mathbb{C}$, où $(V, \|\ \|)$ est un espace vectoriel normé, est bornée et hermitienne, alors $\|\phi\| = \|\Phi\|$.*

Preuve. Nous avons vu que l'inégalité $\|\Phi\| \leqslant \|\phi\|$ est toujours valide. Il faut tout simplement montrer la validité de l'inégalité opposée quand ϕ est hermitienne.

Considérons encore la formule de polarisation

$$\phi(x,y) = \frac{1}{4}(\Phi(x+y) - \Phi(x-y) + i\Phi(x+iy) - i\Phi(x-iy)),$$

vu que Φ est réelle, la partie réelle des deux côtés vaut :

$$\Re(\phi(x,y)) = \frac{1}{4}(\Phi(x+y) - \Phi(x-y)),$$

grâce à l'éq. (6.17) et à la formule du parallélogramme nous pouvons écrire la majoration suivante :

$$|\Re(\phi(x,y))| \leqslant \frac{1}{4}\|\Phi\|(\|x+y\|^2 + \|x-y\|^2) = \frac{1}{2}\|\Phi\|(\|x\|^2 + \|y\|^2). \tag{6.18}$$

Si $\theta \in [0, 2\pi[$ est tel que $\phi(x,y) = |\phi(x,y)|e^{i\theta}$, alors, par linéarité à gauche de ϕ :

$$0 \leqslant |\phi(x,y)| = e^{-i\theta}\phi(x,y) = \phi(e^{-i\theta}x, y),$$

i.e. $\phi(e^{-i\theta}x, y)$ est une quantité réelle positive, donc elle coïncide avec sa partie réelle et aussi avec son module, d'où : $|\phi(x,y)| = |\Re(\phi(e^{-i\theta}x, y))|$. En utilisant l'éq. (6.18) on obtient :

$$\|\phi\| = \sup_{\|x\|=\|y\|=1} |\phi(x,y)| = \sup_{\|x\|=\|y\|=1} |\Re(\phi(e^{-i\theta}x, y))| \leqslant \sup_{\|x\|=\|y\|=1} \frac{1}{2}\|\Phi\|(\|x\|^2 + \|y\|^2) = \|\Phi\|.$$

□

Venons maintenant à l'importante relation entre formes bilinéaires ou sesquilinéaires bornées définies sur un espace de Hilbert $\mathcal{H}$ et opérateurs de $\mathcal{B}(\mathcal{H})$ avec deux résultats qui seront essentiels pour définir l'adjoint d'un opérateur borné.

Théorème 6.5.5 *Pour tout $A \in \mathcal{B}(\mathcal{H})$ fixé, la forme bilinéaire sur (si $\mathcal{H}$ est réel) ou sesquilinéaire (si $\mathcal{H}$ est complexe) ϕ_A sur $\mathcal{H}$ définie par*

$$\phi_A(x,y) = \langle Ax, y\rangle \quad \textit{ou} \quad \phi_A(x,y) = \langle x, Ay\rangle$$

est bornée et ça vaut que $\|\phi_A\| = \|A\|$.

Preuve. Considérons la définition $\phi_A(x,y) = \langle Ax, y\rangle$, la preuve relative à l'autre définition est analogue. Observons que :

$$|\phi_A(x,y)| = |\langle Ax, y\rangle| \underset{\text{(Cauchy-Schwarz)}}{\leqslant} \|Ax\|\|y\| \underset{(6.11)}{\leqslant} \|A\|\|x\|\|y\|, \quad \forall x, y \in \mathcal{H},$$

donc ϕ_A est bornée et

$$\|\phi_A\| = \sup_{\|x\|=\|y\|=1} |\phi_A(x,y)| \leqslant \sup_{\|x\|=\|y\|=1} \|A\|\|x\|\|y\| = \|A\|,$$

donc $\|\phi_A\| \leqslant \|A\|$. Allons maintenant prouver l'égalité des normes en démontrant que $\|A\| \leqslant \|\phi_A\|$. Pour cela, observons que $\phi_A(x, Ax) = \langle Ax, Ax\rangle = \|Ax\|^2 \geqslant 0$ et alors ça vaut aussi que $\|Ax\|^2 = |\phi_A(x, Ax)|$, donc, vu que ϕ_A est bornée :

$$\|Ax\|^2 = |\phi_A(x, Ax)| \leqslant \|\phi_A\|\|x\|\|Ax\|.$$

Si $Ax \neq 0$, alors on peut diviser les deux côtés de l'inégalité précédente par $\|Ax\|$ en obtenant $\|Ax\| \leqslant \|\phi_A\|\|x\|$. Si $Ax = 0$, alors l'inégalité $\|Ax\| \leqslant \|\phi_A\|\|x\|$ s'écrit $0 \leqslant \|\phi_A\|\|x\|$ ce qui est trivialement vrai. Par conséquent, l'inégalité $\|Ax\| \leqslant \|\phi_A\|\|x\|$ vaut sans aucune contrainte et alors nous pouvons écrire :

$$\|A\| = \sup_{\|x\|=1} \|Ax\| \leqslant \sup_{\|x\|=1} \|\phi_A\|\|x\| = \|\phi_A\|,$$

i.e. $\|A\| \leqslant \|\phi_A\|$. □

Si nous écrivons $\mathrm{Bil}_b(\mathcal{H})$, resp. $\mathrm{Sesq}_b(\mathcal{H})$, pour indiquer l'espace vectoriel (par rapport aux opérations linéaires définies ponctuellement) des formes bilinéaires, resp. sesquilinéaires, bornes sur $\mathcal{H}$, alors la correspondance

$$\begin{array}{ccc} \mathcal{B}(\mathcal{H}) & \longrightarrow & \mathrm{Bil}_b(h) \\ A & \longmapsto & \phi_A \end{array}, \quad \text{ou} \quad \begin{array}{ccc} \mathcal{B}(\mathcal{H}) & \longrightarrow & \mathrm{Sesq}_b(h) \\ A & \longmapsto & \phi_A \end{array}$$

est une inclusion isométrique.

La correspondence définie par $\mathcal{B}(\mathcal{H}) \ni A \mapsto \phi_A \in \mathrm{Bil}_b(\mathcal{H})$ est linéaire, par contre, celle donnée par $\mathcal{B}(\mathcal{H}) \ni A \mapsto \phi_A \in \mathrm{Sesq}_b(\mathcal{H})$ est linéaire si l'on définit $\phi_A(x,y) = \langle Ax, y\rangle$ et antilinéaire si l'on définit $\phi_A(x,y) = \langle x, Ay\rangle$.

Grâce à l'isométrie, nous pouvons rajouter une autre caractérisation de la norme d'un opérateur $A \in \mathcal{B}(\mathcal{H})$.

Corollaire 6.5.2 (Cinquième caractérisation de la norme d'un opérateur de $\mathcal{B}(\mathcal{H})$) *Pour tout $A \in \mathcal{B}(\mathcal{H})$ ça vaut :*

$$\boxed{\|A\| = \sup_{\|x\|=\|y\|=1} |\langle Ax, y\rangle|}. \tag{6.19}$$

Le résultat suivant dit que l'application qui associe un opérateur borné à une forme bilinéaire ou sesquilinéaire bornée n'est pas seulement une inclusion isométrique, mais elle est aussi surjective, i.e. toute forme bilinéaire ou sesquilinéaire bornée sur un espace de Hilbert $\mathcal{H}$ est définie par un et un seul opérateur de $\mathcal{B}(\mathcal{H})$. En résumé : **la correspondance opérateur borné $\leftrightarrow$ forme bilinéaire ou sesquilinéaire bornée est un isomorphisme isométrique**.

Théorème 6.5.6 *Soit $\mathcal{H}$ un espace de Hilbert sur $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Pour toute forme $\phi : \mathcal{H} \times \mathcal{H} \to \mathbb{K}$ bilinéaire, si $\mathbb{K} = \mathbb{R}$, ou sesquilinéaire si $\mathbb{K} = \mathbb{C}$, bornée il existe un unique opérateur $B \in \mathcal{B}(\mathcal{H})$ tel que $\phi = \phi_B$, i.e. ça vaut :*

$$\phi(x,y) = \langle Bx, y\rangle \qquad ou \qquad \phi(x,y) = \langle x, By\rangle, \qquad \forall x, y \in \mathcal{H}.$$

Preuve. Pour fixer les idées, nous allons considérer la définition $\phi_B(x,y) = \langle x, By\rangle$, la preuve pour l'autre définition est complètement analogue.

Injectivité : le théorème précédent garantit que, pour tout $B \in \mathcal{B}(\mathcal{H})$, $\phi_B(x,y) = \langle x, By\rangle$ est une forme bilinéaire ou sesquilinéaire bornée. Soient maintenant $B_1, B_2 \in \mathcal{B}(\mathcal{H})$ tels que $\phi = \phi_{B_1} = \phi_{B_2}$, i.e. $\phi(x,y) = \langle x, B_1 y\rangle = \langle x, B_2 y\rangle$ $\forall x, y \in \mathcal{H}$, alors, grâce au théorème 6.2.6, $B_1 = B_2$.

Surjectivité : fixées arbitrairement une forme bilinéaire ou sesquilinéaire $\phi : \mathcal{H} \times \mathcal{H} \to \mathbb{K}$ et un élément $y \in \mathcal{H}$, l'application

$$\begin{array}{rccc} \phi_y : & \mathcal{H} & \longrightarrow & \mathbb{K} \\ & x & \longmapsto & \phi_y(x) := \phi(x,y) \end{array}$$

est évidemment une fonctionnelle linéaire bornée sur $\mathcal{H}$, i.e. $\phi_y \in \mathcal{H}^*$. Grâce au théorème de représentation de Riesz, il existe un seul élément $\xi_y \in \mathcal{H}$ tel que $\phi_y = T_{\xi_y} = T(\xi_y)$, où T est l'isomorphisme de Riesz et $T_{\xi_y} \in \mathcal{H}^*$ est le représentant de Riesz de $\xi_y \in \mathcal{H}$, dont l'action sur n'importe quel $x \in \mathcal{H}$ l'on sait est définie par $T_{\xi_y}(x) = \langle x, \xi_y\rangle$.

En résumé, $\forall x, y \in \mathcal{H}$, ça vaut que $\phi(x,y) = \phi_y(x) = \langle x, \xi_y\rangle$, et alors la surjectivité que l'ont cherche sera prouvé si nous montrons que l'application

$$\begin{array}{rccc} B : & \mathcal{H} & \longrightarrow & \mathcal{H} \\ & y & \longmapsto & By := \xi_y \end{array}$$

est un opérateur linéaire borné sur $\mathcal{H}$, vu que, dans ce cas ça vaut $\phi(x,y) = \langle x, By\rangle \ \forall x, y \in \mathcal{H}$.

Soient $x, y_1, y_2 \in \mathcal{H}$ et $\alpha_1, \alpha_2 \in \mathbb{K}$ quelconques, alors :

$$\begin{aligned}\langle x, \xi_{\alpha_1 y_1 + \alpha_2 y_2}\rangle &= \phi(x, \alpha_1 y_1 + \alpha_2 y_2) = \overline{\alpha_1}\phi(x, y_1) + \overline{\alpha_2}\phi(x, y_2) \\ &= \overline{\alpha_1}\langle x, \xi_{y_1}\rangle + \overline{\alpha_2}\langle x, \xi_{y_2}\rangle\langle x, \alpha_1\xi_{y_1}\rangle + \langle x, \alpha_2\xi_{y_2}\rangle \\ &= \langle x, \alpha_1\xi_{y_1} + \alpha_2\xi_{y_2}\rangle,\end{aligned}$$

ce qui montre la linéarité de la correspondance $\mathcal{H} \ni y \mapsto \xi_y = By \in \mathcal{H}$.

Pour montrer que B est borné, observons que, comme ϕ est bornée, il existe $k > 0$ tel que

$$|\langle x, By\rangle| = |\phi(x, y)| < k\|x\|\|y\| \qquad \forall x, y \in \mathcal{H},$$

grâce à l'arbitrarité de x, l'inégalité vaut aussi quand $x = Ay$, i.e.

$$\|By\|^2 = |\langle By, By\rangle| < k\|By\|\|y\| \qquad \forall y \in \mathcal{H},$$

d'où $\|By\| < k\|y\| \ \forall y \in \mathcal{H}$ tel que $By \neq 0$, et quand $By = 0$ l'inégalité $\|By\| < k\|y\|$ est trivialement vraie, donc ça vaut $\|By\| < k\|y\| \ \forall y \in \mathcal{H}$, i.e. B est borné. □

6.5.1 Le théorème de Lax-Milgram et ses conséquences

En 1954 P. Lax et A.N. Milgram en 1954 ont donné une preuve simple et élégante d'une conséquence remarquable du théorème précédent, qui généralise le théorème de représentation de Riesz à formes bilinéaires ou sesquilinéaires.

Une des hypothèses nécessaires pour ce résultat est définie ci-dessous.

Déf. 6.5.4 (Forme coercive ou V-elliptique) *Soit $(V, \|\ \|)$ un espace vectoriel normé, une forme bilinéaire ou sesquilinéaire $\phi : V \times V \to \mathbb{K}$, $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ est dite coercive ou V-elliptique s'il existe une constante $K > 0$ telle que :*

$$\Phi(x) \geqslant K\|x\|^2, \qquad \forall x \in V.$$

Évidemment, **un produit scalaire sur V est une forme coercive**, car, dans ce cas $\Phi(x) = \langle x, x\rangle = \|x\|^2 \geqslant K\|x\|^2 \ \forall x \in V$ avec $0 < K \leqslant 1$.

Un exemple moins évident : si $z \in \mathcal{C}([0,1], \mathbb{R})$ est telles que $\min\limits_{t\in[0,1]} z(t) > 0$, alors la forme bilinéaire

$$\begin{array}{rccl}\phi : & L^2[0,1] \times L^2[0,1] & \longrightarrow & \mathbb{R} \\ & (x, y) & \longmapsto & \phi(x, y) := \int_0^1 x(t)\overline{y}(t)z(t)dt\end{array}$$

est coercive car

$$\Phi(x) = \int_0^1 |x(t)|^2 z(t)dt \geqslant \int_0^1 |x(t)|^2 \min_{t\in[0,1]} z(t)dt = \min_{t\in[0,1]} z(t) \int_0^1 |x(t)|^2 dt = K\|x\|^2,$$

où $K = \min\limits_{t\in[0,1]} z(t)$.

Théorème 6.5.7 (Lax-Milgram) *Soit $\mathcal{H}$ un espace de Hilbert sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ et soit $\phi : \mathcal{H} \times \mathcal{H} \to \mathbb{K}$ une forme bilinéaire, si $\mathbb{K} = \mathbb{R}$, ou sesquilinéaire, si $\mathbb{K} = \mathbb{C}$, bornée et coercive. Alors, pour toute fonctionnelle bornée $\varphi \in \mathcal{H}^*$ existe un seul élément $u_\varphi \in \mathcal{H}$ tel que :*

$$\varphi(x) = \phi(x, u_\varphi), \qquad \forall x \in \mathcal{H}.$$

Preuve. Grâce au théorème 6.5.6 nous savons qu'il existe un opérateur $A \in \mathcal{B}(\mathcal{H})$ tel que :

$$\phi(x, y) = \langle x, Ay \rangle, \qquad \forall x, y \in \mathcal{H}. \tag{6.20}$$

De l'autre côté, le théorème de représentation de Riesz garantit que, pour toute fonctionnelle linéaire bornée $\varphi \in \mathcal{H}^*$, il existe un seul élément $T^{-1}(\varphi) \in \mathcal{H}$, où T est l'isomorphisme de Riesz, tel que :

$$\varphi(x) = \langle x, T^{-1}(\varphi) \rangle, \qquad \forall x \in \mathcal{H}. \tag{6.21}$$

L'idée pour démontrer le théorème passe par une simple comparaison des équations (6.20) et (6.21) : en fait, si l'opérateur $A : \mathcal{H} \to \mathcal{H}$ est un isomorphisme, alors il existe un unique élément de $\mathcal{H}$, que l'on écrit $u_\varphi \in \mathcal{H}$ car il dépend de φ, qui satisfait $Au_\varphi = T^{-1}(\varphi)$, mais alors :

$$\varphi(x) \underset{(6.21)}{=} \langle x, T^{-1}(\varphi) \rangle \underset{(T^{-1}(\varphi) = Au_\varphi)}{=} \langle x, Au_\varphi \rangle \underset{(6.20)}{=} \phi(x, u_\varphi), \qquad \forall x \in \mathcal{H},$$

i.e. la thèse du théorème de Lax-Milgram.

Allons alors démontrer que A est un isomorphisme. L'injectivité est une simple conséquence de la coercivité, en fait :

$$0 \leqslant K\|x\|^2 \leqslant \Phi(x) = \phi(x, x) = \langle x, Ax \rangle \underset{\langle x, Ax \rangle \geqslant 0}{=} |\langle x, Ax \rangle| \underset{\text{Cauchy-Schwarz}}{\leqslant} \|x\| \, \|Ax\|,$$

d'où : $\|x\| \leqslant \frac{1}{K} \|Ax\|$ pour tout $x \neq 0$, et pour $x = 0$ l'inégalité vaut trivialement, donc elle vaut pour tout $x \in \mathcal{H}$. Ceci entraîne que A est injectif : en fait, étant donnés $x_1, x_2 \in \mathcal{H}$ arbitraires, la condition $Ax_1 = Ax_2$ implique, par linéarité, que $A(x_1 - x_2) = 0$, mais alors $\|x_1 - x_2\| \leqslant \frac{1}{K} \|A(x_1 - x_2)\| = 0$, i.e. $x_1 = x_2$.

La surjectivité de A, i.e. le fait que $\mathrm{Im}(A) = \mathcal{H}$, est légèrement plus compliqué à démontrer. Le premier argument à utiliser passe encore par l'inégalité que l'on vient de démontrer. Précisément, soit $(x_n)_{n \in \mathbb{N}} \subset \mathcal{H}$ une suite quelconque d'éléments de $\mathcal{H}$, alors $(Ax_n)_{n \in \mathbb{N}} \subset \mathrm{Im}(A)$ est une suite arbitraire d'éléments de l'image de A. Supposons que cette suite soit convergente en $\mathcal{H}$, i.e. qu'il existe $y \in \mathcal{H}$ tel que : $\|Ax_n - y\| \underset{n \to +\infty}{\to} 0$, en particulier, en tant que suite convergente, $(Ax_n)_{n \in \mathbb{N}}$ est de Cauchy, i.e. pour tout $\varepsilon > 0$ $\exists N_\varepsilon \in \mathbb{N}$ tel que $n, m \geqslant N_\varepsilon$ implique $\|Ax_n - Ax_m\| < \varepsilon$, mais alors ça vaut aussi $\|x_n - x_m\| \leqslant \frac{1}{K} \|Ax_n - Ax_m\| < \varepsilon$ pour tout $n, m \geqslant N_\varepsilon$, i.e. si $(Ax_n)_{n \in \mathbb{N}}$ converge en $\mathcal{H}$, alors $(x_n)_{n \in \mathbb{N}}$ est une suite de Cauchy en $\mathcal{H}$. Vu que $\mathcal{H}$ est complet, $(x_n)_{n \in \mathbb{N}}$ converge elle-même en $\mathcal{H}$, i.e. il existe $x \in \mathcal{H}$ tel que $\lim_{n \to +\infty} \|x_n - x\| = 0$. A est borné et donc continu, par conséquent,

$$A\left(\lim_{n \to +\infty} \|x_n - x\| \right) = \lim_{n \to +\infty} \|Ax_n - Ax\| = 0,$$

mais alors, par l'unicité de la limite dans un espace métrique, on obtient $y = Ax \in \mathrm{Im}(A)$, i.e. $\mathrm{Im}(A)$ est un sous-espace vectoriel fermé de $\mathcal{H}$ car il contient les limites de toutes ses suites.

La fermeture de $\mathrm{Im}(A)$ nous permet d'utiliser le théorème 5.2.3 qui établit que, si, par l'absurde, $\mathrm{Im}(A)$ est un sous-espace vectoriel propre de $\mathcal{H}$, alors il existe un vecteur non nul $\xi \in \mathcal{H} \backslash \mathrm{Im}(A)$ orthogonal à $\mathrm{Im}(A)$, i.e. $\langle \xi, Ay \rangle = 0 \ \forall y \in \mathcal{H}$. En prenant $y = \xi$ on obtient

$$0 = \langle \xi, A\xi \rangle = \Phi(\xi) \underset{\text{coercivité}}{\geqslant} K\|\xi\|^2 > 0,$$

car $\xi \neq 0$ et $K > 0$, d'où l'absurde. □

Le théorème de Lax-Milgram est très utilisé dans la résolution des équations différentielle en dérivées partielles (EDP) exprimés sous forme variationnelle. Grossièrement, l'idée consiste à réécrire une EDP sous la forme de minimisation d'une fonctionnelle exprimée par une intégrale et chercher ce que l'on appelle la *solution au sens faible* de l'EDP sous la forme d'un minimiseur de la fonctionnelle.

Dans ce type de démarche, un corollaire presque immédiat du théorème de Lax-Milgram (souvent cité comme une partie intégrante du théorème lui-même) est extrêmement utile.

Corollaire 6.5.3 (Lax-Milgram cas symétrique) *Soient :*

- $\mathcal{H}$ *un espace de Hilbert réel ;*
- $\varphi \in \mathcal{H}^*$ *;*
- $\phi : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ *une forme bilinéaire bornée, coercive et symétrique ;*
- $\Phi : \mathcal{H} \to \mathbb{R}$ *la forme quadratique associée à* ϕ*.*

Alors, le vecteur $u_\varphi \in \mathcal{H}$ *tel que* $\varphi(x) = \phi(x, u_\varphi) \ \forall x \in \mathcal{H}$ *est le seul élément de* $\mathcal{H}$ *qui minimise la fonctionnelle linéaire*

$$\begin{array}{rccl} J_\varphi : & \mathcal{H} & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & J_\varphi(x) := \frac{1}{2}\Phi(x) - \varphi(x), \end{array}$$

i.e. $\exists!\ u_\varphi \in \mathcal{H}$ *tel que :*

$$J_\varphi(u_\varphi) = \min_{x \in \mathcal{H}} J_\varphi(x) \quad \Longleftrightarrow \quad u_\varphi = \arg\min_{x \in \mathcal{H}} J_\varphi(x).$$

Preuve. Allons perturber la fonctionnelle J_φ autour de u_φ avec un vecteur $w \in \mathcal{H}$ quelconque :

$$\begin{aligned} J_\varphi(u_\varphi + w) &= \frac{1}{2}\phi(u_\varphi + w, u_\varphi + w) - \varphi(u_\varphi + w) \\ &= \frac{1}{2}[\phi(u_\varphi, u_\varphi) + \phi(u_\varphi, w) + \phi(w, u_\varphi) + \phi(w, w)] - \varphi(u_\varphi) - \varphi(w) \\ &\underset{(\phi \text{ symétrique})}{=} \frac{1}{2}[\phi(u_\varphi, u_\varphi) + 2\phi(u_\varphi, w) + \phi(w, w)] - \varphi(u_\varphi) - \varphi(w) \\ &= \frac{1}{2}\phi(u_\varphi, u_\varphi) - \varphi(u_\varphi) + \frac{1}{2}\phi(w, w) + \phi(w, u_\varphi) - \varphi(w) \\ &\quad \left(\frac{1}{2}\phi(u_\varphi, u_\varphi) - \varphi(u_\varphi) = J(u_\varphi) \ \text{ and } u_\varphi \text{ satisfait } \varphi(w) = \phi(w, u_\varphi)\right) \\ &= J(u_\varphi) + \frac{1}{2}\Phi(w) \\ &\underset{(\phi \text{ coercive})}{\geqslant} J(u_\varphi) + \frac{K}{2}\|w\|^2 \underset{(\frac{K}{2}\|w\|^2 \geqslant 0)}{\geqslant} J(u_\varphi), \end{aligned}$$

i.e. $J(u_\varphi) \leqslant J_\varphi(u_\varphi + w)\ \forall w \in \mathcal{H}$, donc u_φ est le seul minimiseur de J. □

Vu qu'un produit scalaire réel est une forme bilinéaire bornée, coercive et symétrique, et que sa forme quadratique associée est la norme au carré (exprimé typiquement sous une forme intégrale), ce résultat garantit que, pour toute fonctionnelle réelle de la forme :

$$J_\varphi(x) = \frac{1}{2}\|x\|^2 - \varphi(x),$$

où $\varphi \in \mathcal{H}^*$, il existe un seul minimiseur $u_\varphi \in \mathcal{H}$.

Le théorème de Lax-Milgram et sa version symétrique sont à la base des **méthodes aux éléments finis**, l'idée est la suivante : si φ n'a pas une expression simple, alors chercher directement le minimiseur (solution faible d'une EDP) u_φ dans l'entier espace de Hilbert $\mathcal{H}$ peut être très compliqué, dans ce cas on approche le problème en cherchant une suite $u_\varphi^{(n)}$ dans $\mathcal{H}_n$, un sous-espace de dimension finie de $\mathcal{H}$ (d'où le nom 'éléments finis').

Dans le cas où ϕ est symétrique et définie positive, $u_\varphi^{(n)}$ est la projection orthogonale de u sur $\mathcal{H}_n$ au sens du produit scalaire défini par ϕ.

Une fois définie une base (typiquement orthonormée) $(h_i)_{i=1}^n$ de $\mathcal{H}_n$, le problème se ramène à la résolution du système linéaire $Au_\varphi^{(n)} = b$, où $A_{ij} = \phi(h_j, h_i)$ et $b_i = \varphi(h_i)$.

Terminons cette section avec l'observation que le théorème de Lax-Milgram que l'on a présenté peut être obtenu comme corollaire d'un théorème démontré par J.L. Lions et G. Stampacchia en 1967 dans le cadre des inégalités variationnelles.

6.6 L'opérateur adjoint et ses propriétés

Dans cette section nous allons examiner une conséquence très importante du théorème de représentation de Riesz et des résultats que l'on vient d'examiner dans la section précédente : la possibilité d'associer à A un autre opérateur, dit « adjoint », qui a une importance fondamentale dans la théorie et dans les applications de l'analyse fonctionnelle.

Considérons un opérateur $A \in \mathcal{B}(\mathcal{H})$, alors, grâce au théorème 6.5.5, la forme bilinéaire ou sesquilinéaire définie par $\phi(x, y) = \langle x, Ay\rangle$ est bornée. Grâce au théorème 6.5.6, il existe un unique opérateur borné B tel que, pour tout $x, y \in \mathcal{H}$, ça vaut $\phi(x, y) = \langle Bx, y\rangle$, d'où : $\langle x, Ay\rangle = \phi(x, y) = \langle Bx, y\rangle$. Avec les mêmes arguments, en sélectionnant les options alternatives dans les théorèmes 6.5.5 et 6.5.6, on arrive à l'équation : $\langle Ax, y\rangle = \phi(x, y) = \langle x, By\rangle$ $\forall x, y \in \mathcal{H}$.

L'opérateur B a un symbole et un nom spécifiques.

Déf. 6.6.1 *Soit $A \in \mathcal{B}(\mathcal{H})$. On appelle **opérateur adjoint** de A, et on l'écrit avec*[4] *$A^\dagger$, l'opérateur $A^\dagger \in \mathcal{B}(\mathcal{H})$ tel que :*

$$\boxed{\langle A^\dagger x, y\rangle = \langle x, Ay\rangle} \quad \textit{et} \quad \boxed{\langle Ax, y\rangle = \langle x, A^\dagger y\rangle} \qquad \forall x, y \in \mathcal{H}.$$

*L'application $^\dagger : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$, $A \mapsto A^\dagger$ est dite **adjonction**.*

4. L'origine du symbole †, *dagger* en anglais, poignard en français, vient de la relation étroite entre l'opérateur adjoint $A^\dagger$ et l'opérateur transposé ou dual A^t, pour plus d'information le lecteur est invité à consulter l'appendice B. Un autre symbole très utilisé est A^*.

Théorème 6.6.1 *L'adjonction est un automorphisme antilinéaire de $\mathcal{B}(\mathcal{H})$, en fait elle vérifie les propriétés suivantes : pour tout $A, B \in \mathcal{B}(\mathcal{H})$ et $k \in \mathbb{K}$,*

1. $(A+B)^\dagger = A^\dagger + B^\dagger$
2. $(kA)^\dagger = \bar{k}A^\dagger$
3. $(AB)^\dagger = B^\dagger A^\dagger$
4. $(A)^{\dagger\dagger} = A$
5. $\|A^\dagger A\| = \|A\|^2$, $\|AA^\dagger\| = \|A^\dagger\|^2$
6. $\|A^\dagger\| = \|A\|$.

Preuve.

1. et 2. : conséquences immédiates de la sesquilinéarité du produit scalaire complexe (si l'espace de Hilbert est réel, bien sur $\bar{k} = k$, comme conséquence de la bilinéarité).
3. : $\langle (AB)^\dagger x, y\rangle = \langle x, ABy\rangle = \langle A^\dagger x, By\rangle = \langle B^\dagger A^\dagger x, y\rangle \ \forall x, y \in \mathcal{H}$, d'où la 3.
4. : vu que $A^{\dagger\dagger} = (A^\dagger)^\dagger$, $\langle A^{\dagger\dagger} x, y\rangle = \langle (A^\dagger)^\dagger x, y\rangle = \langle x, A^\dagger y\rangle = \overline{\langle A^\dagger y, x\rangle} = \overline{\langle y, Ax\rangle} = \langle Ax, y\rangle \ \forall x, y \in \mathcal{H}$, d'où la 4.
5. : montrons d'abord que $\|A\|^2 \leqslant \|A^\dagger A\|$: soit $x \in \mathcal{H}$, $\|x\| = 1$, alors

$$\begin{aligned}\|Ax\|^2 = \langle Ax, Ax\rangle = |\langle Ax, Ax\rangle| = |\langle A^\dagger Ax, x\rangle| &\underset{\text{Cauchy-Schwarz}}{\leqslant} \|A^\dagger Ax\|\|x\| = \|A^\dagger Ax\| \\ &\leqslant \|A^\dagger A\|\|x\| = \|A^\dagger A\|,\end{aligned}$$

donc, vu que $\|A\|^2 = \sup_{\|x\|=1} \|Ax\|^2$, $\|A\|^2 \leqslant \|A^\dagger A\|$.

Montrons maintenant que $\|A^\dagger A\| \leqslant \|A\|^2$. Commençons par observer que, pour tout $x, y \in \mathcal{H}$, $\|x\| = \|y\| = 1$ ça vaut que

$$\begin{aligned}\Re\langle Ax, Ay\rangle \leqslant \sqrt{(\Re\langle Ax, Ay\rangle)^2 + (\mathfrak{Im}(\langle Ax, Ay\rangle)^2} = |\langle Ax, Ay\rangle| \\ \underset{\text{Cauchy-Schwarz}}{\leqslant} \|Ax\|\|Ay\| \leqslant \|A\|\|x\|\|A\|\|y\| = \|A\|^2.\end{aligned} \tag{6.22}$$

Si $\langle A^\dagger Ax, y\rangle = |\langle A^\dagger Ax, y\rangle| e^{i\vartheta}$, ϑ étant la phase de $\langle A^\dagger Ax, y\rangle$, alors

$$\mathbb{R} \ni |\langle A^\dagger Ax, y\rangle| = e^{-i\vartheta}\langle A^\dagger Ax, y\rangle = \langle A^\dagger Ax, e^{i\vartheta} y\rangle,$$

i.e. $\langle A^\dagger Ax, e^{i\vartheta} y\rangle \in \mathbb{R}$ et donc $\langle A^\dagger Ax, e^{i\vartheta} y\rangle = \Re\langle A^\dagger Ax, e^{i\vartheta} y\rangle \underset{(6.22)}{\leqslant} \|A\|^2$, vu que $\|e^{i\vartheta} y\| =$ 1. Grâce au fait que $\langle A^\dagger Ax, y\rangle = \langle A^\dagger Ax, e^{i\vartheta} y\rangle$ on peut écrire :

$$|\langle A^\dagger Ax, y\rangle| \leqslant \|A\|^2, \quad \forall x, y \in \mathcal{H}, \ \|x\| = \|y\| = 1. \tag{6.23}$$

Prenons maintenant un $\xi \in \mathcal{H}$ quelconque et utilisons cette dernière inégalité pour estimer la norme de $A^\dagger A\xi$:

$$\begin{aligned}\|A^\dagger A\xi\| &= \frac{1}{\|A^\dagger A\xi\|}\frac{1}{\|\xi\|}\|A^\dagger A\xi\|^2\|\xi\| = \frac{1}{\|A^\dagger A\xi\|}\frac{1}{\|\xi\|}\langle A^\dagger A\xi, A^\dagger A\xi\rangle\|\xi\| \\ &= \frac{1}{\|A^\dagger A\xi\|}\frac{1}{\|\xi\|}|\langle A^\dagger A\xi, A^\dagger A\xi\rangle|\|\xi\| = \left|\langle A^\dagger A\frac{\xi}{\|\xi\|}, \frac{A^\dagger A\xi}{\|A^\dagger A\xi\|}\rangle\right|\|\xi\|,\end{aligned}$$

si on écrit $x = \frac{\xi}{\|\xi\|}$ et $y = \frac{A^\dagger A\xi}{\|A^\dagger A\xi\|}$ et on observe que ces deux vecteurs sont unitaires, alors on peut utiliser l'inégalité (6.23) pour écrire $\|A^\dagger A\xi\| \leqslant \|A\|^2\|\xi\|$, pour tout $\xi \in \mathcal{H}$, ce qui implique que $\|A^\dagger A\| = \sup_{\|\xi\|=1} \|A^\dagger A\xi\| \leqslant \|A\|^2$. D'où $\|A^\dagger A\| = \|A\|^2\ \forall A \in \mathcal{B}(\mathcal{H})$. Si on écrit $B = A^\dagger$, alors $B \in \mathcal{B}(\mathcal{H})$ et $\|B^\dagger B\| = \|B\|^2$, i.e. $\|A^{\dagger\dagger}A^\dagger\| = \|A^\dagger\|^2$, mais $A^{\dagger\dagger} = A$, donc $\|AA^\dagger\| = \|A^\dagger\|^2$ pour tout $A \in \mathcal{B}(\mathcal{H})$.

6. : d'un côté on a

$$\|A\|^2 = \|A^\dagger A\| \underset{(6.12)}{\leqslant} \|A^\dagger\|\|A\| \implies \frac{\|A\|^2}{\|A\|} \leqslant \frac{\|A^\dagger\|\|A\|}{\|A\|} \iff \|A\| \leqslant \|A^\dagger\|,$$

de l'autre côté

$$\|A^\dagger\|^2 = \|AA^\dagger\| \underset{(6.12)}{\leqslant} \|A\|\|A^\dagger\| \implies \frac{\|A^\dagger\|^2}{\|A^\dagger\|} \leqslant \frac{\|A\|\|A^\dagger\|}{\|A^\dagger\|} \iff \|A^\dagger\| \leqslant \|A\|.$$

□

Un corollaire immédiat des propriétés 1. et 6. est que **l'adjonction** $^\dagger : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ **est une fonction continue**, en fait, supposons que $(A_n)_{n\in\mathbb{N}} \subset \mathcal{B}(\mathcal{H})$ soit une suite de $\mathcal{B}(\mathcal{H})$ convergente vers $A \in \mathcal{B}(\mathcal{H})$, i.e. $\|A_n - A\| \underset{n\to+\infty}{\to} 0$, alors :

$$\|A_n{}^\dagger - A^\dagger\| \underset{(1.)}{=} \|(A_n - A)^\dagger\| \underset{(6.)}{=} \|A_n - A\| \underset{n\to+\infty}{\to} 0.$$

L'algèbre de Banach $\mathcal{B}(\mathcal{H})$ munie de l'opération d'adjonction devient une $\mathbb{C}^*$-algèbre, comme formalisé ci-dessous.

Déf. 6.6.2 ($\mathbb{C}^*$-algèbre) *Une algèbre de Banach $\mathfrak{A}$ est dite $\mathbb{C}^*$-algèbre s'il est possible définir sur $\mathfrak{A}$ une involution, i.e. une transformation $j : \mathfrak{A} \to \mathfrak{A}$ telle que, $\forall a, b \in \mathfrak{A}$ et $\forall k \in \mathbb{C}$:*

1. $j(a + b) = j(a) + j(b)$
2. $j(ka) = \bar{k}j(a)$
3. $j(ab) = j(b)j(a)$
4. $j(j(a)) = a$.

La théorie des $\mathbb{C}^*$-algèbres est très importante dans l'analyse fonctionnelle et ses applications, néanmoins, une discussion des $\mathbb{C}^*$-algèbres est hors du but de cet ouvrage.

Allons examiner les opérateurs qui sont « insensibles » par rapport à l'adjonction, dans le sens que ils sont l'adjoint d'eux mêmes, ces opérateurs sont très utilisés et c'est pour ça qu'on leur donne un nom spécifique.

Déf. 6.6.3 (Opérateurs auto-adjoints ou hermitiens) *$A \in \mathcal{B}(\mathcal{H})$ est dit un opérateur **auto-adjoint** (a.a.) ou **hermitien** si $A^\dagger = A$, i.e. si :*

$$\langle Ax, y\rangle = \langle x, Ay\rangle, \qquad \forall x, y \in \mathcal{H}.$$

Pour avoir une idée de l'importance des opérateurs auto-adjoints, ça suffit de dire que **les observables physiques de la mécanique quantique sont représentées par des opérateurs auto-adjoints définis sur un espace de Hilbert**.

Deux opérateurs auto-adjoints remarquables sont : $A^\dagger A$ et $AA^\dagger$.

Théorème 6.6.2 *Soit $A \in \mathcal{B}(\mathcal{H})$, alors $A^\dagger A$ et $AA^\dagger$ sont auto-adjoints.*

Preuve. Il suffit d'appliquer les propriétés $(AB)^\dagger = B^\dagger A^\dagger$ et $A^{\dagger\dagger} = A$, en fait :

$$(A^\dagger A)^\dagger = A^\dagger A^{\dagger\dagger} = A^\dagger A$$

et

$$(AA^\dagger)^\dagger = A^{\dagger\dagger} A^\dagger = AA^\dagger.$$

□

Dans le théorème suivant on montre sous quelles conditions la propriété d'être auto-adjoint est stable par rapport aux opérations de l'algèbre $\mathcal{B}(\mathcal{H})$. Avant d'énoncer le résultat, on introduit la notation suivante : $\forall A, B \in \mathcal{B}(\mathcal{H})$ on définit l'opérateur $[A, B] = AB - BA$, dit **le commutateur entre A et B**. On dit que A et B commutent si $[A, B] = \mathbf{0}$, l'opérateur nul, dans ce cas : $AB = BA$.

Théorème 6.6.3 *Soient $A, B \in \mathcal{B}(\mathcal{H})$, A, B auto-adjoints, alors :*

- *$\alpha A + \beta B$ est auto-adjoint si et seulement si $\alpha, \beta \in \mathbb{R}$;*
- *AB est auto-adjoint si et seulement si $[A, B] = \mathbf{0}$.*

Preuve. La première propriété est une conséquence très simple de la propriété 2. de l'adjonction et de la sesquilinéairité du produit scalaire. Allons démontrer la deuxième.

$\boxed{\Rightarrow}$: AB a.a., i.e. $AB = (AB)^\dagger$, alors $(AB)^\dagger = B^\dagger A^\dagger \underset{A,B \text{ a.a.}}{=} BA$, donc $AB = BA$.

$\boxed{\Leftarrow}$: $\forall x, y \in \mathcal{H}$ ça vaut : $\langle ABx, y\rangle = \langle Bx, A^\dagger y\rangle = \langle x, B^\dagger A^\dagger y\rangle \underset{A,B \text{ a.a.}}{=} \langle x, BAy\rangle \underset{[A,B]=\mathbf{0}}{=} \langle x, ABy\rangle$, donc $AB = (AB)^\dagger$.

□

Allons voir un exercice dans lequel on va appliquer plusieurs résultats qu'on a vu jusqu'ici.

Exercice 6.4

Soient $(u_n)_{n\in\mathbb{N}}$ un s.o.n. dans l'espace de Hilbert $\mathcal{H}$, $(\lambda_n)_{n\in\mathbb{N}} \subset \mathbb{C}$ et $A : \mathcal{H} \to \mathcal{H}$,

$$Ax = \sum_{n\in\mathbb{N}} \lambda_n \langle x, u_n\rangle u_n, \qquad \forall x \in \mathcal{H}\,.$$

1. Démontrer que si la suite $(\lambda_n)_{n\in\mathbb{N}}$ est bornée, alors $A \in \mathcal{B}(\mathcal{H})$;
2. Calculer l'adjoint $A^\dagger$ de A. En déduire une condition nécessaire est suffisante pour que l'opérateur A soit anti-autoadjoint, i.e. $A + A^\dagger = 0$;

3. On considère, pour tout $n \in \mathbb{N}$, l'opérateur A_n défini par

$$A_n x = \sum_{k=0}^{n} \lambda_k \langle x, u_k \rangle u_k \, .$$

(a) Pour $n \in \mathbb{N}$, calculer $A_n u_{n+1} - A u_{n+1}$. En déduire une condition nécessaire pour qu'on ait $A_n \underset{n\to+\infty}{\longrightarrow} A$ dans $\mathcal{B}(\mathcal{H})$.

(b) En supposant que $\lim\limits_{n\to+\infty} \lambda_n = 0$, prouver que $A_n \underset{n\to+\infty}{\longrightarrow} A$ dans $\mathcal{B}(\mathcal{H})$.

Solution de l'exercice 6.4

1. Vu que $(u_n)_{n\in\mathbb{N}}$ un s.o.n. dans l'espace de Hilbert, le théorème de Fischer-Riesz 5.5.3 garantit que la bonne définition de Ax, i.e. la convergence de la série (en $\mathcal{H}$) $\sum\limits_{n\in\mathbb{N}} \lambda_n \langle x, u_n \rangle u_n$ est équivalente à la convergence de la série (en $\mathbb{C}$) $\sum\limits_{n\in\mathbb{N}} |\lambda_n \langle x, u_n \rangle|^2 = \sum\limits_{n\in\mathbb{N}} |\lambda_n|^2 |\langle x, u_n \rangle|^2$. Si $(\lambda_n)_{n\in\mathbb{N}}$ est une suite bornée, i.e. $\sup\limits_{n\in\mathbb{N}} |\lambda_n| = M < +\infty$, alors, $\sum\limits_{n\in\mathbb{N}} |\lambda_n \langle x, u_n \rangle|^2 \leqslant M^2 \|x\|^2 < +\infty$, grâce à l'inégalité de Bessel (5.5).
Analysons maintenant la bornitude de A :

$$\|Ax\|^2 = \left\| \sum_{n\in\mathbb{N}} \lambda_n \langle x, u_n \rangle u_n \right\|^2 \underset{\text{(th. Pythagore)}}{=} \sum_{n\in\mathbb{N}} \|\lambda_n \langle x, u_n \rangle u_n\|^2 = \sum_{n\in\mathbb{N}} |\lambda_n|^2 |\langle x, u_n \rangle|^2,$$

et on peut utiliser encore la bornitude de $(\lambda_n)_{n\in\mathbb{N}}$ et l'inégalité de Bessel pour écrire $\|Ax\| \leqslant M\|x\|$, $\forall x \in \mathcal{H}$, mais alors $\|A\| = \sup\limits_{\|x\|=1} \|Ax\| \leqslant M$, ce qui montre que $A \in \mathcal{B}(\mathcal{H})$.

2. Soient $x, y \in \mathcal{H}$, alors

$$\langle Ax, y \rangle = \sum_{n\in\mathbb{N}} \lambda_n \langle x, u_n \rangle \langle u_n, y \rangle = \left\langle x, \sum_{n\in\mathbb{N}} \overline{\lambda_n}\, \overline{\langle u_n, y \rangle} u_n \right\rangle = \left\langle x, \sum_{n\in\mathbb{N}} \overline{\lambda_n} \langle y, u_n \rangle u_n \right\rangle,$$

par conséquent $A^\dagger x = \sum\limits_{n\in\mathbb{N}} \overline{\lambda_n} \langle x, u_n \rangle u_n$, $\forall x \in \mathcal{H}$.

Par continuité on peut écrire $(A + A^\dagger)x = \sum\limits_{n\in\mathbb{N}} (\lambda_n + \overline{\lambda_n}) \langle x, u_n \rangle u_n$, donc $A + A^\dagger$ est l'opérateur nul si et seulement si $\lambda_n + \overline{\lambda_n} = 0$ $\forall n \in \mathbb{N}$. En écrivant $\lambda_n = a_n + i b_n$, $a_n, b_n \in \mathbb{R}$ $\forall n \in \mathbb{N}$, on voit que la condition $\lambda_n + \overline{\lambda_n} = 0$ est équivalente à $a_n = -a_n$ $\forall n \in \mathbb{N}$, i.e. $a_n = 0$ $\forall n \in \mathbb{N}$, tandis qu'on n'a aucune contrainte sur b_n. Par conséquent, A est anti-autoadjoint si et seulement si $\lambda_n \in i\mathbb{R}$ pour tout $n \in \mathbb{N}$, i.e. λ_n est une suite imaginaire pure.

3. (a) En utilisant les faits suivants :

$$\|u_n\| = 1, \quad \langle u_k, u_{n+1} \rangle = 0 \qquad \forall n \in \mathbb{N},\ k \neq n+1,$$

on déduit que $A_n u_{n+1} = 0$, $A u_{n+1} = \lambda_{n+1} u_{n+1}$ $\forall n \in \mathbb{N}$. Par conséquent, grâce au résultat 6.10, nous pouvons écrire :

$$\|A_n - A\|_{\mathcal{B}(\mathcal{H})} \geqslant \|(A_n - A)u_{n+1}\|_{\mathcal{H}} = |\lambda_{n+1}| \qquad \forall n \in \mathbb{N},$$

i.e., une condition nécessaire pour que $\lim\limits_{n\to+\infty} \|A_n - A\|_{\mathcal{B}(\mathcal{H})} = 0$ est $\lim\limits_{n\to+\infty} \lambda_n = 0$.

(b) Calculons, pour tout $n \in \mathbb{N}$ et $x \in \mathcal{H}$,

$$\|A_n x - Ax\|^2 = \left\| \sum_{k=n+1}^{\infty} \lambda_n \langle x, u_k \rangle u_k \right\|^2 = \sum_{k=n+1}^{\infty} |\lambda_k|^2 \, |\langle x, u_k \rangle|^2 \leqslant \left(\sup_{k \geqslant n+1} |\lambda_k| \right)^2 \|x\|^2,$$

grâce à l'inégalité de Bessel. Par conséquent $\|A_n - A\|_{\mathcal{B}(\mathcal{H})} \leqslant \sup_{k \geqslant n+1} |\lambda_k|$, $\forall n \in \mathbb{N}$. En utilisant le fait que $\lim_{n \to +\infty} \lambda_n = 0$, on obtient donc bien que

$$\lim_{n \to +\infty} \|A_n - A\|_{\mathcal{B}(\mathcal{H})} = 0.$$

□

Venons maintenant à la norme d'un opérateur auto-adjoint.

Théorème 6.6.4 *Soit $A \in \mathcal{B}(\mathcal{H})$ un opérateur auto-adjoint. Alors :*

$$\|A\| = \sup_{\|x\|=1} |\langle Ax, x \rangle|.$$

Preuve. Par simplicité, écrivons :

$$s_A = \sup_{\|x\|=1} |\langle Ax, x \rangle|.$$

$\boxed{s_A \leqslant \|A\|}$: par l'inégalité de Cauchy-Schwarz on a que

$$|\langle Ax, x \rangle| \leqslant \|x\| \|Ax\| \leqslant \|A\| \|x\|^2,$$

donc

$$s_A = \sup_{\|x\|=1} |\langle Ax, x \rangle| \leqslant \sup_{\|x\|=1} \|A\| \|x\|^2 = \|A\|,$$

et alors $s_A \leqslant \|A\|$.

$\boxed{\|A\| \leqslant s_A}$: nous allons utiliser le fait que, $\forall z \in \mathbb{C}$, $z + \bar{z} = 2\Re e(z)$ pour écrire $\forall x, y \in \mathcal{H}$

$$4\Re(\langle Ax, y \rangle) = 4\frac{1}{2}[\langle Ax, y \rangle + \overline{\langle Ax, y \rangle}] = 2[\langle Ax, y \rangle + \langle y, Ax \rangle].$$

Par calcul direct, on peut vérifier que ça vaut l'égalité suivante :

$$2[\langle Ax, y \rangle + \langle y, Ax \rangle] = \langle A(x+y), x+y \rangle - \langle A(x-y), x-y \rangle,$$

donc

$$\begin{aligned} 4\Re(\langle Ax, y \rangle) &= \langle A(x+y), x+y \rangle - \langle A(x-y), x-y \rangle \\ &= \|x+y\|^2 \langle A\frac{x+y}{\|x+y\|}, \frac{x+y}{\|x+y\|} \rangle - \|x-y\|^2 \langle A\frac{x-y}{\|x-y\|}, \frac{x-y}{\|x-y\|} \rangle \\ &\leqslant \|x+y\|^2 s_A + \|x-y\|^2 s_A = s_A(\|x+y\|^2 + \|x-y\|^2) \underset{(1.2.4)}{=} s_A 2(\|x\|^2 + \|y\|^2), \end{aligned}$$

i.e. $\Re(\langle Ax, y\rangle) \leqslant \frac{1}{2} s_A(\|x\|^2 + \|y\|^2)\ \forall x, y \in \mathcal{H}$. Vu la validité de l'inégalité pour chaque couple de vecteurs de $\mathcal{H}$, nous pouvons considérer, en particulier, le couple x, z où $z = e^{i\vartheta} y$, $\vartheta \in \mathbb{R}$ quelconque. Comme $\|z\| = \|y\|$, l'inégalité précédente devient

$$\Re(\langle Ax, e^{i\vartheta} y\rangle) \leqslant \frac{1}{2} s_A(\|x\|^2 + \|y\|^2). \tag{6.24}$$

Utilisons maintenant un argument similaire à celui qu'on a déjà utilisé dans la preuve de la propriété 5. de l'adjonction, i.e. écrivons $\langle Ax, y\rangle = |\langle Ax, y\rangle| e^{i\vartheta}$, ϑ étant la phase de $\langle Ax, y\rangle$, alors

$$\mathbb{R} \ni |\langle Ax, y\rangle| = e^{-i\vartheta} \langle Ax, y\rangle = \langle Ax, e^{i\vartheta} y\rangle \underset{\text{(car réel)}}{=} \Re(\langle Ax, e^{i\vartheta} y\rangle),$$

donc $|\langle Ax, y\rangle| = \Re(\langle Ax, e^{i\vartheta} y\rangle)$ et alors l'inégalité (6.24) peut être réécrite comme ça :

$$|\langle Ax, y\rangle| \leqslant \frac{1}{2} s_A(\|x\|^2 + \|y\|^2).$$

Allons introduire dans cette inégalité le vecteur $y = \frac{\|x\|}{\|Ax\|} Ax$. Dans le côté de gauche on obtient :

$$|\langle Ax, y\rangle| = |\langle Ax, \frac{\|x\|}{\|Ax\|} Ax\rangle| = \frac{\|x\|}{\|Ax\|} |\langle Ax, Ax\rangle| = \frac{\|x\|}{\|Ax\|} \|Ax\|^2 = \|x\| \|Ax\|,$$

tandis qu'à droite on a :

$$\frac{1}{2} s_A(\|x\|^2 + \|y\|^2) = \frac{1}{2} s_A(\|x\|^2 + \frac{\|x\|^2}{\|Ax\|^2} \|Ax\|^2) = \frac{1}{2} s_A(\|x\|^2 + \|x\|^2) = s_A \|x\|^2,$$

donc, $\forall x \in \mathcal{H}$ ça vaut $\|x\| \|Ax\| \leqslant s_A \|x\|^2$, et, si $x \neq 0_{\mathcal{H}}$, alors $\|Ax\| \leqslant s_A \|x\|$, d'où :

$$\|A\| = \sup_{x \neq 0_{\mathcal{H}}} \frac{\|Ax\|}{\|x\|} \leqslant s_A \sup_{x \neq 0_{\mathcal{H}}} \frac{\|x\|}{\|x\|} = s_A,$$

en définitive $\|A\| \leqslant s_A$. □

Le théorème suivant montre une propriété des opérateurs adjoints fondamentale dans la théorie de l'optimisation.

Théorème 6.6.5 *Soit $A \in \mathcal{B}(\mathcal{H})$, alors :*

$$\ker(A) = (\mathrm{Im}(A^\dagger))^\perp \quad \text{et} \quad \overline{\mathrm{Im}(A^\dagger)} = (\ker(A))^\perp,$$

donc

$$\mathcal{H} = \ker(A) \oplus \overline{\mathrm{Im}(A^\dagger)} \quad \text{et} \quad \mathcal{H} = \ker(A^\dagger) \oplus \overline{\mathrm{Im}(A)}.$$

Preuve.

$\boxed{\ker(A) \subseteq (\mathrm{Im}(A^\dagger))^\perp}$: soit $x \in \mathcal{H}$ quelconque et $\underline{y \in \ker(A)}$, alors $Ay = 0_{\mathcal{H}}$ et donc on peut écrire

$$0 = \langle x, 0_{\mathcal{H}}\rangle = \langle x, Ay\rangle = \langle A^\dagger x, y\rangle,$$

i.e. $y \perp A^\dagger x\ \forall x \in \mathcal{H}$, comme $\mathrm{Im}(A^\dagger) = \{A^\dagger x,\ x \in \mathcal{H}\}$, cela implique que $\underline{y \in (\mathrm{Im}(A^\dagger))^\perp}$.

$\boxed{(\mathrm{Im}(A^\dagger))^\perp \subseteq \ker(A)}$: soit $\underline{y \in (\mathrm{Im}(A^\dagger))^\perp}$, alors $\langle A^\dagger x, y\rangle = 0\ \forall x \in \mathcal{H}$, mais $\langle A^\dagger x, y\rangle = \langle x, Ay\rangle$, donc $\langle x, Ay\rangle = 0\ \forall x \in \mathcal{H}$, i.e. $Ay = 0_{\mathcal{H}}$, mais alors $\underline{y \in \ker(A)}$.

Par conséquent : $\ker(A) = (\mathrm{Im}(A^\dagger))^\perp = (\overline{\mathrm{Im}(A^\dagger)})^\perp$, mais alors, si on opère encore une fois le complément orthogonale : $\ker(A)^\perp = (\mathrm{Im}(A^\dagger))^{\perp\perp} = \overline{\mathrm{Im}(A^\dagger)}$. On observe que considérer l'adhérence de $\mathrm{Im}(A^\dagger)$ est indispensable, car $\ker(A)^\perp$ est un sous-espace fermé de $\mathcal{H}$ et $\mathrm{Im}(A^\dagger)$ ne l'est, en général, pas.

Les décompositions orthogonales de $\mathcal{H}$ en somme directe de sous-espaces sont une conséquence immédiate du théorème de la projection orthogonale. □

Terminons avec l'analyse de la relation entre l'inversion et l'adjonction. Nous rappelons de la section 6.3 que, si V est un espace de Banach, alors $\mathrm{GL}(V)$ est son groupe général linéaire, i.e. le groupe des opérateur linéaires continus et bijectifs avec inverse continue.

Théorème 6.6.6 *Soit $\mathcal{H}$ un espace de Hilbert et soit $A \in GL(\mathcal{H})$. Alors $A^\dagger$ est inversible et :*

1. *ça vaut :*
$$\boxed{(A^\dagger)^{-1} = (A^{-1})^\dagger}$$
i.e. pour les opérateurs de $GL(\mathcal{H})$ les opérations d'inversion et d'adjonctions commutent entre elles : l'inverse de l'adjoint est l'adjoint de l'inverse ;
2. *si, en plus, $A \in GL(\mathcal{H})$ est auto-adjoint, alors A^{-1} est autoadjoint.*

Preuve.

1. Il faut prouver que, pour tout $x \in \mathcal{H}$, $(A^{-1})^\dagger A^\dagger x = A^\dagger (A^{-1})^\dagger x = x$. Pour cela, considérons, $\forall x, y \in \mathcal{H}$:
$$\langle y, (A^{-1})^\dagger A^\dagger x\rangle = \langle A^{-1}y, A^\dagger x\rangle = \langle AA^{-1}y, x\rangle = \langle y, x\rangle$$
$$\langle y, A^\dagger (A^{-1})^\dagger x\rangle = \langle Ay, (A^{-1})^\dagger x\rangle = \langle A^{-1}Ay, x\rangle = \langle y, x\rangle,$$
d'où, grâce à (6.2.6) $(A^{-1})^\dagger A^\dagger x = A^\dagger (A^{-1})^\dagger x = x$.

2. Conséquence immédiate de la propriété 1. : si $A = A^\dagger$, alors $A^{-1} = (A^{-1})^\dagger$.

□

6.7 Les opérateurs de projection orthogonale dans un espace de Hilbert

On a déjà rencontré les opérateurs de projection orthogonale dans un espace de Hilbert $\mathcal{H}$. On veut, d'un côté, généraliser le concept de projection et, de l'autre côté, caractériser les projections orthogonales. L'opérateur adjoint sera fondamentale pour arriver à ce résultat.

Pour avoir une idée claire et simple du concept d'opérateur de projection, non nécessairement orthogonale, imaginons d'être dans un espace Euclidien de dimension finie, par exemple $\mathbb{R}^2$ et de projeter un vecteur dans la direction d'un autre vecteur.

Imaginons d'itérer le processus, i.e. de vouloir « projeter la projection », il est clair que cette opération n'a aucun effet sur le vecteur projeté. Cette propriété est prise comme définition du concept de projection [5].

5. Beaucoup d'auteurs utilisent la définition **projection oblique** pour la distinguer du concept, plus restrictif, de projection orthogonale.

Déf. 6.7.1 *Soit $A \in \mathcal{B}(\mathcal{H})$, A est un projecteur, ou un opérateur de projection, s'il est idempotent, i.e. $A^2 = A$.*

La présence d'un produit scalaire en $\mathcal{H}$ permet de cibler une projection particulière, celle orthogonale. Les résultats du chapitre 5 ont montré que la complétude de $\mathcal{H}$ par rapport à la topologie engendrée par le produit scalaire permet de donner deux définitions équivalentes de projection orthogonales.

Déf. 6.7.2 *Soit $\mathcal{H}$ un espace de Hilbert et S un sous-espace vectoriel propre et fermé de $\mathcal{H}$. La fonction :*

$$\begin{array}{rccc} P_S: & \mathcal{H} & \longrightarrow & S \\ & x & \longmapsto & P_S(x), \end{array}$$

où : $\|x - P_S(x)\| = \inf_{y \in S} \|x - y\|$, i.e. $P_S(x)$ est l'élément de S qui minimise la distance à $x \in \mathcal{H}$ par rapport à la norme induite par le produit scalaire de $\mathcal{H}$, est dite opérateur de projection orthogonale sur le sous-espace S.

D'une manière équivalente, si on considère la décomposition de $\mathcal{H}$ suivante : $\mathcal{H} = S \oplus S^\perp$ et l'écriture $x = x' + x''$, avec $x \in \mathcal{H}$, $x' \in S$, $x'' \in S^\perp$, alors l'opérateur de projection orthogonale P_S est défini via $P_S(x) = x'$.

Allons voir un exemple de projecteur. Soit $\mathcal{H} = L^2[-a, a]$, avec $a \in \mathbb{R}$ munit de la tribu de Borel et de la mesure de Lebesgue. On peut démontrer que les fonctions paires et impaires sont orthogonales pour le produit scalaire de $L^2[-a, a]$. On a ensuite la décomposition suivante :

$$f(x) = \underbrace{\frac{f(x) + f(-x)}{2}}_{\text{partie paire}} + \underbrace{\frac{f(x) - f(-x)}{2}}_{\text{partie impaire}},$$

donc, le projecteur de $f \in L^2[-a, a]$ sur le sous-espace $\mathcal{P} \subset L^2[-a, a]$ des fonctions paires est défini par $P_{\mathcal{P}} f(x) = \frac{f(x)+f(-x)}{2}$ et le projecteur sur le sous-espace $\mathcal{I} \subset L^2[-a, a]$ des fonctions impaires est définit par $P_{\mathcal{I}} f(x) = \frac{f(x)-f(-x)}{2}$.

Examinons les propriétés de l'opérateur P_S :

1. $\boxed{P_S^2 = P_S}$ (**idempotence**). $\forall x \in \mathcal{H}$, on a : $P_S^2(x) = P_S(\underbrace{P_S(x)}_{\in S, \text{ par définition}}) = P_S(x)$. Donc P_S est bien un projecteur.

2. $\boxed{P_S|_S = id_S}$. C'est évident : l'élément $P_S(x) \in S$ qui minimise la distance à $x \in S$ est lui même. Autrement dit, si $x = x' \in S$ alors, $P_S(x) = P_S(x') = x'$.

3. P_S **est un opérateur linéaire continu**. Soit $x_1 = x_1' + x_1'' \in \mathcal{H}$, avec $x_1' \in S$ et $x_1'' \in S^\perp$ et $x_2 = x_2' + x_2'' \in \mathcal{H}$, avec $x_2' \in S$ et $x_2'' \in S^\perp$. Pour tout $\alpha, \beta \in \mathbb{K}$ on a :

$$\alpha x_1 + \beta x_2 = \underbrace{\alpha x_1' + \beta x_2'}_{\in S} + \underbrace{\alpha x_1'' + \beta x_2''}_{\in S^\perp}$$

et donc,

$$\begin{aligned} P_S(\alpha x_1 + \beta x_2) &= \alpha x_1' + \beta x_2' \\ &= \alpha P_S(x_1) + \beta P_S(x_2), \end{aligned}$$

P_S est donc un opérateur linéaire. Pour montrer qu'il est continu, on montre qu'il est borné : si $x = x' + x'' \in \mathcal{H}$ quelconque, avec $x' \in S$, $x'' \in S^\perp$, alors, par le théorème de Pythagore, $\|x\|^2 = \|x'\|^2 + \|x''\|^2$ et $\|P_S x\|^2 = \|x'\|^2 \leqslant \|x'\|^2 + \|x''\|^2 = \|x\|^2$, i.e. $\|P_S x\| \leqslant \|x\|$ $\forall x \in \mathcal{H}$.

4. P_S **est une fonction 1-lipschitzienne**, i.e :

$$\|P_S(x) - P_S(y)\| \leqslant \|x - y\|, \quad \forall x, y \in \mathcal{H}.$$

Il suffit de remarquer que $\forall x, y \in \mathcal{H}$, la projection de $x - y$, i.e. $P_S(x - y)$ et le vecteur résidu $(x - y) - P_S(x - y)$ sont orthogonaux, car l'une appartient à S et l'autre à $S^\perp$, donc on peut appliquer le théorème de Pythagore pour écrire :

$$\begin{aligned} \|x - y\|^2 &= \|(x - y) - P_S(x - y) + P_S(x - y)\|^2 \\ &= \|(x - y) - P_S(x - y)\|^2 + \|P_S(x - y)\|^2 \\ &\geqslant \|P_S(x - y)\|^2 \\ &= \|P_S(x) - P_S(y)\|^2. \end{aligned}$$

On retrouve ainsi : $\|P_S(x) - P_S(y)\| \leqslant \|x - y\|, \forall x, y \in \mathcal{H}$.

5. Les projecteurs orthogonaux non banals ont norme unitaire :

$$\boxed{\|P_S\| = \begin{cases} 1 \text{ si } S \neq \{0_\mathcal{H}\} \\ 0 \text{ si } S = \{0_\mathcal{H}\}. \end{cases}}$$

Si $S = \{0_\mathcal{H}\}$ alors $P_S \equiv \mathbf{0}$ et donc sa norme est 0. Sinon, vu que P_S est borné, $\|P_S x\| \leqslant \|x\|$ $\forall x \in \mathcal{H}$, i.e. $\frac{\|P_S x\|}{\|x\|} \leqslant 1$ $\forall x \in \mathcal{H}$. Mais, si $S \neq \{0_\mathcal{H}\}$, alors il existe $\bar{x} \in S$, $\bar{x} \neq 0_\mathcal{H}$ et $P_S \bar{x} = \bar{x}$, i.e. $\|P_S \bar{x}\| = \|\bar{x}\|$, i.e. $\frac{\|P_S \bar{x}\|}{\|\bar{x}\|} = 1$. Alors, $\|P_S\| = \sup\limits_{x \in \mathcal{H},\, x \neq 0} \frac{\|P_S x\|}{\|x\|} = 1$.

6. $\boxed{\mathrm{Im}(P_S) = S}$. Évident, par définition de l'opérateur de projection.

7. $\boxed{\ker P_S = S^\perp}$. On montre la double inclusion :
$\underline{x \in \ker P_S} \implies P_S(x) = 0$ et $0 \perp y$, $\forall y \in S$, donc $\underline{x \in S^\perp}$.
$\underline{x \in S^\perp} \implies x = 0 + x = P_S(x) + P_{S^\perp}(x)$, par unicité de la décomposition orthogonale, donc $P_S(x) = 0$, et alors $\underline{x \in \ker P_S}$.

8. Une conséquence immédiate des deux propriétés précédentes et du théorème de la projection est :

$$\boxed{\mathcal{H} = \mathrm{Im}(P_S) \oplus \ker(P_S)} \qquad \forall S \text{ sous-espace fermé de } \mathcal{H}.$$

9. $\boxed{P_S + P_{S^\perp} = id_{\mathcal{H}}}$ (**décomposition de l'identité**). C'est évident. Pour $x \in \mathcal{H}$ quelconque, on a toujours la décomposition $x = x' + x''$ avec $x' = P_S(x)$ et $x'' = P_{S^\perp}(x)$. On a donc $P_S(x) + P_{S^\perp}(x) = (P_S + P_{S^\perp})(x) = x' + x'' = x$, $\forall x \in \mathcal{H}$. Comme conséquence immédiate on a que :

$$\boxed{P_{S^\perp} = id_{\mathcal{H}} - P_S} \qquad \forall S \text{ sous-espace fermé de } \mathcal{H}.$$

Pour tout $x \in \mathcal{H}$, le vecteur résidu de la projection de x sur S est obtenu via $P_{S^\perp} = (id_{\mathcal{H}} - P_S)(x) = x - P_S(x)$.

10. **Caractérisation du sous-espace de projection** :

$$\boxed{S = \{x \in \mathcal{H} : P_S(x) = x\} = \{x \in \mathcal{H} : \|P_S(x)\| = \|x\|\}},$$

c'est-à-dire, les éléments de S sont les points fixes de P_S en $\mathcal{H}$, qui peuvent se caractériser eux-mêmes pour être les éléments de $\mathcal{H}$ qui ont une norme égale à celle de leur projection sur S. La caractérisation est une conséquence immédiate de l'idempotence, il faut simplement vérifier l'équivalence : $P_S(x) = x \iff \|P_S(x)\| = \|x\|$:
$\boxed{\Rightarrow}$: évidemment, $P_S(x) = x \Rightarrow \|P_S(x)\| = \|x\|$.

$\boxed{\Leftarrow}$: on pose à nouveau $\mathcal{H} \ni x = x' + x''$ avec $P_S(x) = x'$, alors

$$\begin{aligned}
\|P_S(x)\| = \|x\| \quad &\Longrightarrow \quad \|P_S(x)\|^2 = \|x\|^2 \implies \left\|x'\right\|^2 = \left\|x' + x''\right\|^2 \\
&\Longrightarrow \quad \left\|x'\right\|^2 = \left\|x'\right\|^2 + \left\|x''\right\|^2 \text{ (car } x' \perp x'') \\
&\Longleftrightarrow \quad \left\|x''\right\|^2 = 0 \\
&\Longleftrightarrow \quad x'' = 0 \text{ (propriété de la norme)} \\
&\Longleftrightarrow \quad x = x' \text{ (par l'écriture de } x),
\end{aligned}$$

on a donc $\|P_S(x)\| = \|x\| \Rightarrow x = x' = P_S(x)$.

11. $\boxed{P_S = P_S^\dagger}$ (**auto-adjoint**). Pour le prouver on utilise deux fois le théorème de la projection, sur $x \in \mathcal{H}$ et sur $y \in \mathcal{H}$: $x = x' + x''$, $y = y' + y''$, $x', y' \in S$, $x'', y'' \in S^\perp$:

$\langle P_S x, y\rangle = \langle x', y' + y''\rangle = \langle x', y'\rangle + \cancel{\langle x', y''\rangle} = \langle x - x'', y'\rangle = \langle x, y'\rangle - \cancel{\langle x'', y'\rangle} = \langle x, y'\rangle$,

mais $y' = P_S y$ et donc : $\langle P_S x, y\rangle = \langle x, P_S y\rangle$ $\forall x, y \in \mathcal{H}$.

Parmi les propriétés qu'on vient d'examiner, il y en a deux qui caractérisent les opérateurs linéaires bornés comme projecteurs orthogonales.

Théorème 6.7.1 (Caractérisation « algébrique » des projecteurs orthogonales) *Soit $A \in \mathcal{B}(\mathcal{H})$. Les affirmations suivantes sont équivalentes.*

1. *A est un projecteur orthogonale*
2. $A^\dagger A = A$
3. $A^\dagger A = A^\dagger$
4. *A est auto-adjoint et idempotent, i.e. $A^\dagger = A$ et $A^2 = A$.*

Preuve. La boucle logique suivante : **1.** $\Longrightarrow$ **2.** $\Longrightarrow$ **3.** $\Longrightarrow$ **4.** $\Longrightarrow$ **1.** démontrera le théorème.

$\boxed{\textbf{1.} \Longrightarrow \textbf{2.}}$: A est un projecteur orthogonale, donc $A^\dagger = A$ et $A^2 = A$, alors $A^\dagger A = AA = A^2 = A$.

$\boxed{\textbf{2.} \Longrightarrow \textbf{3.}}$: si $A^\dagger A = A$, alors $(A^\dagger A)^\dagger = A^\dagger$, mais on sait que $A^\dagger A$ est auto-adjoint, donc $(A^\dagger A)^\dagger = A^\dagger A = A^\dagger$.

$\boxed{\textbf{3.} \Longrightarrow \textbf{4.}}$: si $A^\dagger A = A^\dagger$, alors $(A^\dagger A)^\dagger = A^{\dagger\dagger} = A$, mais $A^\dagger A$ est auto-adjoint, donc $(A^\dagger A)^\dagger = A^\dagger A = A$. Néanmoins, par hypothèse, $A^\dagger A = A^\dagger$, donc $A^\dagger = A$, i.e. A est auto-adjoint. En utilisant encore une fois de l'hypothèse de départ $A^\dagger A = A^\dagger$, le fait que A soit auto-adjoint implique que $A^2 = A$, i.e. A est idempotent.

$\boxed{\textbf{4.} \Longrightarrow \textbf{1.}}$: soit $A \in \mathcal{B}(\mathcal{H})$ auto-adjoint et idempotent, on doit démontrer que A est un projecteur orthogonale. Par définition, un projecteur orthogonale projette sur un sous-espace vectoriel fermé, donc il faut tout d'abord démontrer que $\mathrm{Im}(A)$, le sous-espace candidat à être « le lieu » de la projection, est fermé, étant données les hypothèses du théorème.

On montrera que *la continuité et l'idempotence d'un opérateur linéaire A impliquent la fermeture de son image*, ce qui est remarquable, car l'idempotence semble ne pas avoir une relation évidente avec le concept de fermeture d'un sous-espace vectoriel.

Soit $(x_n)_{n\in\mathbb{N}} \subset \mathrm{Im}(A)$ une suite convergente à $x_0 \in \mathcal{H}$, il faut démontrer que $x_0 \in \mathrm{Im}(A)$. Comme chaque $x_n \in \mathrm{Im}(A)$, alors, $\forall n \in \mathbb{N}$, il existe $\xi_n \in \mathcal{H}$ tel que $x_n = A\xi_n$ et donc on peut réécrire $x_n \xrightarrow[n\to+\infty]{} x_0$ comme $A\xi_n \xrightarrow[n\to+\infty]{} x_0$. La continuité de A implique que $A^2\xi_n \xrightarrow[n\to+\infty]{} Ax_0$, mais $A^2 = A$, donc $A\xi_n \xrightarrow[n\to+\infty]{} Ax_0$. Mais alors on a $A\xi_n \xrightarrow[n\to+\infty]{} Ax_0$ et $A\xi_n \xrightarrow[n\to+\infty]{} x_0$, l'unicité de la limite implique que $Ax_0 = x_0$, i.e. $x_0 \in \mathrm{Im}(A)$, donc $\mathrm{Im}(A)$ est fermé.

La propriété $A^\dagger = A$ ici intervient pour aider l'idempotence afin de montrer que A projette et il le fait d'une manière orthogonale. Pour prouver cela, allons essayer d'écrire une décomposition *orthogonale* de $\mathcal{H}$ relative à $\mathrm{Im}(A)$: pour tout $x \in \mathcal{H}$ considérons $x = Ax + (x - Ax)$, alors $Ax \in \mathrm{Im}(A)$ par définition, il faut démontrer que $x - Ax$ est orthogonale à tout vecteur de la forme $A\xi$, $\xi \in \mathcal{H}$:

$$\langle A\xi, x - Ax\rangle = \langle \xi, A^\dagger(x - Ax)\rangle \underset{A \text{ a.a.}}{=} \langle \xi, A(x - Ax)\rangle = \langle \xi, Ax - A^2x\rangle \underset{A^2=A}{=} \langle \xi, Ax - Ax\rangle = 0,$$

donc $x - Ax \in \mathrm{Im}(A)^\perp$ et alors $A = P_{\mathrm{Im}(A)}$ grâce au théorème de la projection orthogonale. $\square$

Exercice 6.5

Soit $E \subset \ell^2(\mathbb{N}, \mathbb{C})$ l'ensemble :

$$E = \{x = (x_n)_{n\in\mathbb{N}} \in \ell^2(\mathbb{N}, \mathbb{C}) \ : \ x_0 + x_1 = 0\}.$$

1. Montrer que E est un sous-espace vectoriel fermé de $\ell^2(\mathbb{N}, \mathbb{C})$.

2. Décrire explicitement $E^\perp$ et déterminer l'opérateur $P_E : \ell^2(\mathbb{N}, \mathbb{C}) \to E$ de projection orthogonale sur E. Déterminer $\|P_E\|$ et $P_E^\dagger$.

3. Soit $x = (x_n)_{n\in\mathbb{N}} \in \ell^2(\mathbb{N},\mathbb{C})$ tel que :

$$x_n = \begin{cases} 1 & \text{si } n = 0 \\ 0 & \text{sinon,} \end{cases}$$

calculer la distance entre x et le sous-espace E, i.e. $\delta = \inf_{y\in E} \{\|x - y\|\}$.

4. Soit $A : \ell^2(\mathbb{N},\mathbb{C}) \to \ell^2(\mathbb{N},\mathbb{C})$ l'opérateur défini par

$$(A(x_0, x_1, x_2, \dots))_n = \begin{cases} -x_1 & \text{si } n = 0 \\ x_n & \text{sinon.} \end{cases}$$

Déterminer $\|A\|$ et $A^\dagger$. (Indication : calculer $A(0, 1, 0, 0, ...)$).

5. Montrer que $A^2 = A$ et déterminer $\mathrm{Im}(A)$. L'opérateur A est-il une projecteur orthogonal ?

Solution de l'exercice 6.5

1. Montrons d'abord que E est un sous-espace vectoriel de $\ell^2(\mathbb{N},\mathbb{C})$: considérons $\lambda \in \mathbb{C}$ quelconque et $x, y \in E$ arbitraires, alors, vu que la structure linéaire de $\ell^2(\mathbb{N},\mathbb{C})$ est définie ponctuellement, ça vaut que

$$z := \lambda x + y = (\lambda x_n)_{n\in\mathbb{N}} + (y_n)_{n\in\mathbb{N}} = (\lambda x_n + y_n)_{n\in\mathbb{N}},$$

donc $z_0 = \lambda x_0 + y_0$ et $z_1 = \lambda x_1 + y_1$, et alors $z_0 + z_1 = \lambda(x_0 + x_1) + (y_0 + y_1) = \lambda\cdot 0 + 0 = 0$ car $x, y \in E$, ce qui montre que E est fermé par rapport aux combinaisons linéaires de ses éléments.

Pour montrer que E est fermé nous allons utiliser une stratégie très pratique quand il y a des *contraintes d'annulation*, comme dans le cas de l'équation $x_0 + x_1 = 0$. L'idée consiste à **identifier cette contrainte avec la condition qui définit le noyau d'un opérateur linéaire continu entre espaces vectoriels normés**, que l'on sait être un sous espace vectoriel fermé du domaine de l'opérateur grâce au théorème (6.2.4). Dans notre cas, il est simple d'identifier la somme des opérateurs de projection sur la première et deuxième composante, i.e. $A := P_0 + P_1 : \ell^2(\mathbb{N},\mathbb{C}) \to \mathbb{C}$, $A(x) := P_0(x) + P_1(x) = x_0 + x_1$, comme l'opérateur linéaire continu (en tant que somme d'opérateurs linéaires continus) entre deux espaces de Hilbert tel que $\ker(A) = \{x \in \ell^2(\mathbb{N},\mathbb{C}) : Ax = 0 \iff x_0 + x_1 = 0\} = E$, ce qui montre la fermeture de E.

2. Dû à la contrainte $x_0 + x_1 = 0$, nous pouvons écrire une suite $x \in E$ comme ça : $(x_0, -x_0, x_2, x_3, \dots)$, bien sûr tout en respectant le fait que $x \in \ell^2(\mathbb{N},\mathbb{C})$, ceci implique que la base hilbertienne canonique de $\ell^2(\mathbb{N},\mathbb{C})$, i.e. $e = ((1, 0, 0, \dots), (0, 1, 0, \dots), \dots)$ peut être utilisée pour construire une bases hilbertienne de E comme ça :

$$\tilde{e} := ((1, -1, 0, \dots), (-1, 1, 0, \dots), (0, 0, 1, 0, \dots), \dots)$$

et donc : $E^\perp = \{y \in \ell^2(\mathbb{N},\mathbb{C}) : \langle y, \tilde{e}_n\rangle = 0 \ \forall n \in \mathbb{N}\}$.
Soit $y = (y_0, y_1, y_2, \dots) \in \ell^2(\mathbb{N},\mathbb{C})$, alors

— $n = 0 : \langle y, \tilde{e}_0 \rangle = y_0 - y_1 + 0 + \cdots = y_0 - y_1$ nul si et seulement si $y_0 = y_1$
— $n = 1 : \langle y, \tilde{e}_1 \rangle = y_0 - y_1 + 0 + \cdots = -y_0 + y_1$ nul si et seulement si $y_0 = y_1$, comme pour $n = 0$
— $n = 2 : \langle y, \tilde{e}_2 \rangle = 0 + 0 + y_2 + 0 + \cdots = y_2$ nul si et seulement si $y_2 = 0$,
il est clair que, pour tout $n \geqslant 2$, $\langle y, \tilde{e}_n \rangle = y_n$, qui est nul si et seulement si $y_n = 0$. Par conséquent, le seul vecteur $y \in \ell^2(\mathbb{N}, \mathbb{C})$ qui est orthogonal à tous les éléments de la base hilbertienne $\tilde{e}$ de E est $y = (y_0, y_0, 0, \dots)$, i.e.

$$E^\perp = \{(y, y, 0, 0, \dots),\ y \in \mathbb{C}\}.$$

Pour déterminer l'opérateur de projection orthogonale sur E nous allons utiliser le théorème de la projection : $\ell^2(\mathbb{N}, \mathbb{C}) = E \oplus E^\perp$. Il suffit de décomposer le vecteur arbitraire $z = (z_n)_{n\in\mathbb{N}} \in \ell^2(\mathbb{N}, \mathbb{C})$ dans la somme de deux vecteurs, l'un qui appartient à E est l'autre à $E^\perp$. Pour faire ça nous observons que, donné $z = (z_0, z_1, z_2, \dots)$, $z \in E$ si les deux premières composantes sont l'une inverse de l'autre, et $z \in E^\perp$ si les deux premières composantes sont égales et elles sont nulles à partir de la troisième, alors

$$z = (z_0, z_1, z_2, z_3, \dots) = (a, -a, z_2, z_3, \dots) + (b, b, 0, 0, \dots),$$

avec les conditions :

$$\begin{cases} a + b = z_0 \\ b - a = z_1, \end{cases}$$

qui a comme solution $a = (z_0 - z_1)/2$ et $b = (z_0 + z_1)/2$, i.e.

$$(z_0, z_1, z_2, \dots) = \left(\frac{z_0 - z_1}{2}, -\frac{z_0 - z_1}{2}, z_2, \dots\right) + \left(\frac{z_0 + z_1}{2}, \frac{z_0 + z_1}{2}, 0, 0, \dots\right),$$

avec le premier vecteur dans E et le deuxième en $E^\perp$, donc

$$P_E(z_0, z_1, z_2, \dots) = \left(\frac{z_0 - z_1}{2}, -\frac{z_0 - z_1}{2}, z_2, \dots\right)$$

est l'expression explicite du projecteur orthogonal sur E. Finalement, sans faire un seul calcul, nous pouvons assurer que P_E a norme unitaire, $\|P^E\| = 1$, en tant que projecteur orthogonale, et aussi que $P_E^\dagger = P_E$, car les projecteurs orthogonaux sont auto-adjoints.

3. Soit $x = (x_n)_{n\in\mathbb{N}}$ l'élément de $\ell^2(\mathbb{N}, \mathbb{C})$ tel que :

$$x_n = \begin{cases} 1 & \text{si } n = 0 \\ 0 & \text{sinon} \end{cases}$$

Comme E est un sous-espace vectoriel fermé de $\ell^2(\mathbb{N}, \mathbb{C})$, la distance entre x et E est bien définie grâce au théorème de la projection. $P_E(x)$ représente le vecteur de E le plus proche de x, donc cette distance est égale à $\delta = \|x - P_E(x)\|_{\ell^2}$.

$$P_E(x) = P_E((1, 0, \dots)) = \left(\frac{1 - 0}{2}, -\frac{1 - 0}{2}, 0, \dots\right) = \left(\frac{1}{2}, -\frac{1}{2}, 0, \dots\right),$$

et alors

$$\delta = \left\|(1, 0, \dots) - \left(\frac{1}{2}, -\frac{1}{2}, 0, \dots\right)\right\|_{\ell^2} = \left\|\left(\frac{1}{2}, \frac{1}{2}, 0, \dots\right)\right\|_{\ell^2} = \sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2} = \frac{1}{\sqrt{2}}.$$

4. Tout d'abord nous observons que x_0 n'intervient pas dans l'action de A, donc

$$A(x_0, x_1, x_2, \dots) = A(y_0, x_1, x_2, \dots) = \begin{cases} -x_1 & \text{si } n = 0 \\ x_n & \text{sinon} \end{cases} \qquad \forall y_0 \in \mathbb{C}.$$

En particulier cela vaut pour $y_0 = 0$, par conséquent nous pouvons restreindre l'action de A sur les éléments de $\ell^2(\mathbb{N}, \mathbb{C})$ de la forme suivante $x = (0, x_1, x_2, \dots)$. Avec cette spécification, par calcul direct nous obtenons :

$$\begin{aligned} \|Ax\|_{\ell^2} &= \sqrt{(-x_1)^2 + x_1^2 + x_2^2 + \dots} = \sqrt{2x_1^2 + x_2^2 + \dots} \leqslant \sqrt{2x_1^2 + 2x_2^2 + \dots} \\ &\leqslant \sqrt{2}\sqrt{0^2 + x_1^2 + x_2^2 + \dots} = \sqrt{2}\|x\|_{\ell^2}. \end{aligned}$$

Cette majoration implique que la définition (6.3) de norme opératorielle dans notre cas devient :

$$\|A\| = \inf\{0 < c \leqslant \sqrt{2} \ : \ \|Ax\|_{\ell^2} \leqslant c\|x\|_{\ell^2} \ \ \forall x = (0, x_1, x_2, \dots) \in \ell^2(\mathbb{N}, \mathbb{C})\}.$$

L'inf est le sup de l'ensemble de minimiseurs, donc, si on trouve un vecteur $x \in \ell^2(\mathbb{N}, \mathbb{C})$ pour lequel $\|Ax\|_{\ell^2} = \sqrt{2}$, la norme de A sera obligatoirement $\sqrt{2}$. En suivant la suggestion, nous allons calculer

$$\|A(0, 1, 0, \dots)\|_{\ell^2} = \|(-1, 1, 0, \dots)\|_{\ell^2} = \sqrt{(-1)^2 + 1^2 + 0^2 + \dots} = \sqrt{2},$$

et alors $\|A\| = \sqrt{2}$.
Passons au calcul de $A^\dagger$. Pour tout $x, y \in \ell^2(\mathbb{N}, \mathbb{C})$ (ici on ne considère pas nécessairement x de la forme $(0, x_1, x_2, \dots)$) ça vaut :

$$\begin{aligned} \langle Ax, y\rangle_{\ell^2} &= \langle(-x_1, x_1, x_2, \dots), (y_0, y_1, y_2, \dots)\rangle_{\ell^2} = -x_1\overline{y_0} + x_1\overline{y_1} + x_2\overline{y_2} + \cdots \\ &= x_0 \cdot 0 + x_1(\overline{y_1 - y_0}) + x_2\overline{y_2} + \cdots = \langle x, (0, y_1 - y_0, y_2, ...)\rangle = \langle x, A^\dagger y\rangle, \end{aligned}$$

et alors l'opérateur adjoint de A est :

$$A^\dagger(y) = (0, y_1 - y_0, y_2, \dots) \qquad \forall y \in \ell^2(\mathbb{N}, \mathbb{C}).$$

5. On a $A^2x = AAx = A(-x_1, x_1, x_2, \dots) = (-x_1, x_1, x_2, \dots) = Ax$ pour tout $x \in \ell^2(\mathbb{N}, \mathbb{C})$, donc A est idempotent. De plus, il est clair que $\mathrm{Im}(A) = E$, où E est le sous-espace définit au début de l'exercice. Donc, A est un opérateur de projection sur E, néanmoins, $\|A\| = \sqrt{2} \neq 1$ et pour ça il ne peut pas être un projecteur orthogonal ! Par conséquent, A est un opérateur de projection oblique sur E. Ci-dessous, nous montrons explicitement la différence de l'action de A et P_E :

$$Ax = (-x_1, x_1, x_2, \dots) \qquad \text{opérateur de projection oblique sur } E$$

$$P_E\, x = \left(\frac{x_0 - x_1}{2}, -\frac{x_0 - x_1}{2}, x_2, \dots\right) \qquad \text{opérateur de projection orthogonale sur } E.$$

□

6.7.1 Les opérateurs de multiplication bornés et leur relation avec les projecteurs orthogonaux

Nous allons voir une application concrète du dernier théorème et, en même temps, nous en profitons pour introduire une nouvelle catégorie d'opérateurs linéaires très utilisés.

Déf. 6.7.3 *Soit $\mathcal{H} = L^2(X, \mathcal{A}, \mu)$ et $g \in L^\infty(X, \mathcal{A}, \mu)$. L'opérateur de multiplication par g est défini par :*

$$\begin{array}{rccc} M_g : & L^2(X,\mathcal{A},\mu) & \longrightarrow & L^2(X,\mathcal{A},\mu) \\ & f & \longmapsto & M_g f = f \cdot g, \end{array}$$

où $f \cdot g(x) = f(x)g(x)\ \forall x \in X$ (multiplication ponctuelle).

Allons examiner les propriétés de M_g :

— M_g est **borné** $\forall g \in L^\infty(X, \mathcal{A}, \mu)$:

$$\|M_g f\|_2^2 = \int_X |f(x)g(x)|^2 d\mu(x) \leqslant \int_X \left(\sup_{x\in X} |g(x)|^2\right) |f(x)|^2 d\mu(x) = \|g\|_\infty^2 \, \|f\|_2^2 < +\infty,$$

donc[6] $\|M_g\|_2 \leqslant \|g\|_\infty$ et alors $M_g \in \mathcal{B}(L^2(X, \mathcal{A}, \mu))\ \forall g \in L^\infty(X, \mathcal{A}, \mu)$.

— $\forall g, h \in L^\infty(X, \mathcal{A}, \mu)$, grâce à la commutativité du produit ponctuel, **les opérateurs de multiplication commutent**, i.e. $M_g M_h = M_h M_g$, i.e. $[M_g, M_h] = \mathbf{0}$.

— $\ker(M_g) = \{f \in L^2(X, \mathcal{A}, \mu)\ :\ M_g(f) = f \cdot g = 0_{L^2(X,\mathcal{A},\mu)}\}$. Si on définit l'ensemble

$$N_g = \{x \in X\ :\ g(x) = 0\},$$

il est clair que $(f \cdot g)(x) = 0\ \forall x \in N_g$, par conséquent, comme $g(x) \neq 0\ \forall x \in {N_g}^c = X \backslash N_g$, pour avoir obtenir la fonction nulle sur X via le produit $f \cdot g$ il faut tout simplement demander que f soit nulle sur ${N_g}^c$ (se rappeler que f est une classe d'équivalence de fonctions égales presque partout). En résumé :

$$\boxed{\ker(M_g) = \{f \in L^2(X, \mathcal{A}, \mu)\ :\ f(x) = 0\ \forall x \in {N_g}^c\}}.$$

— Examinons l'inversibilité de M_g : pour rendre le noyau de M_g soit banale, il faut que le seul élément de $\ker(M_g)$ soit la classe d'équivalence dans laquelle appartient la fonction identiquement nulle, mais cela correspond à demander que $\mu(N_g) = 0$, car, dans ce cas $\mu({N_g}^c) = \mu(X) - \mu(N_g) = \mu(X)$ donc $\ker(M_g) = \{f \in L^2(X, \mathcal{A}, \mu)\ :\ f(x) = 0 \text{ p.p}\}$.

— Si $\mu(N_g) = 0$, alors il existe l'opérateur inverse de M_g : ${M_g}^{-1} : \mathrm{Im}(M_g) \to L^2(X, \mathcal{A}, \mu)$ qui peut être caractérisé à l'aide de la fonction $\frac{1}{g} : X \to \mathbb{K}$, $\frac{1}{g}(x) = \begin{cases} \frac{1}{g(x)} & \text{si } g(x) \neq 0 \\ 0 & \text{sinon.} \end{cases}$

Par définition, $\mathrm{Im}(M_g) = \{h \in L^2(X, \mathcal{A}, \mu)\ :\ \exists f \in L^2(X, \mathcal{A}, \mu)\ :\ h = M_g(f) = f \cdot g\}$, il est clair que $\frac{1}{g} \cdot h = \frac{1}{g(x)} \cdot f \cdot g = f$, cette simple observation nous permet de caractériser, à la fois, $\mathrm{Im}(M_g)$ et l'action de ${M_g}^{-1}$:

$$\boxed{\mathrm{Im}(M_g) = \{h \in L^2(X, \mathcal{A}, \mu)\ :\ \frac{1}{g} \cdot h \in L^2(X, \mathcal{A}, \mu)\}} \quad \text{et} \quad \boxed{{M_g}^{-1} = M_{\frac{1}{g}}}.$$

6. Il est possible de démontrer que, si $(X, \mathcal{A}, \mu)$ est un espace mesuré avec mesure σ-finie, i.e. $X = \bigcup_{k\in\mathbb{N}} A_k$, où $\mu(A_k) < +\infty$, alors $\|M_g\|_2 = \|g\|_\infty$.

— Calculons l'adjoint de M_g : $\forall f, h \in L^2(X, \mathcal{A}, \mu)$, $g \in L^\infty(X, \mathcal{A}, \mu)$,

$$\begin{aligned}\langle M_g^\dagger f, h\rangle = \langle f, M_g h\rangle = \langle f, gh\rangle &= \int_X f(x)\overline{g(x)}\,\overline{h(x)}d\mu(x) = \int_X \Big(\overline{g(x)}f(x)\Big)\,\overline{h(x)}d\mu(x)\\ &= \langle \bar{g}\cdot f, h\rangle = \langle M_{\bar{g}} f, h\rangle,\end{aligned}$$

i.e. $\boxed{M_g^\dagger = M_{\bar{g}}}$, grâce au théorème 6.2.6.

— Calculons M_g^2 :

$$M_g^2 f = M_g(M_g f) = M_g(fg) = fg^2 = M_{g^2} f, \quad \forall f \in L^2(X, \mathcal{A}, \mu),\ g \in L^\infty(X, \mathcal{A}, \mu).$$

Donc l'opérateur linéaire borné M_g est auto-adjoint et idempotent si et seulement si : $\bar{g} = g$ et $g^2 = g$. La première condition impose que g soit une fonction à valeurs réelles, la seule fonction à valeurs réelles qui est égale à son carré est une fonction qui prend seulement les valeurs 0 et 1, i.e. la fonction indicatrice d'un sous-ensemble mesurable de $\mathbb{R}^n$, qui est certainement un élément de $L^\infty(X, \mathcal{A}, \mu)$.
En résumé : l'opérateur de multiplication M_g **est un opérateur de projection orthogonale si et seulement si** $g = \chi_E$**,** $E \subseteq X$ **mesurable**. M_{χ_E} est inversible si et seulement si $\mu(E^c) \neq 0$, i.e. $\mu(E) \neq \mu(X)$.
Indépendamment de l'inversibilité, calculons l'image de M_{χ_E} : la condition qui détermine ce sous-espace est $\frac{1}{\chi_E}\cdot h \in L^2(X, \mathcal{A}, \mu)$, mais, par définition, $\frac{1}{\chi_E}(x) = 0\ \forall x \in X$ tel que $\chi_E(x) = 0$, i.e. $\forall x \in E^c$ et, dans ce cas, $\frac{1}{\chi_E}\cdot h \in L^2(X, \mathcal{A}, \mu)$. Quand $x \in E$, $\frac{1}{\chi_E}(x) = 1$, et donc la condition définitoire de $\mathrm{Im}(M_g)$ devient $h \in L^2(E, \mathcal{A}, \mu)$.

En conclusion, pour tout ensemble mesurable $E \subseteq X$, l'opérateur de multiplication et de projection orthogonale M_{χ_E} est, explicitement,

$$\begin{array}{rccl} M_{\chi_E}: & L^2(X, \mathcal{A}, \mu) & \longrightarrow & L^2(E, \mathcal{A}, \mu) \\ & f & \longmapsto & M_{\chi_E} f = \begin{cases} f(x) & x \in E \\ 0 & \text{sinon.} \end{cases} \end{array}$$

6.7.2 La réalisation géométrique des opérateurs de projection orthogonale via systèmes orthonormés

On a maintenant la possibilité de démontrer une autre analogie entre les espaces de Hilbert et les espaces Euclidiens en dimension finie : la réalisation géométrique des projecteurs orthogonaux sur un sous-espace vectoriel engendré par une famille orthogonale, déjà examiné dans le chapitre 1.

Rappelons que le projecteur orthogonal d'un espace vectoriel pré-hilbertien V de dimension finie n dans un sous-espace vectoriel S de dimension s est :

$$P_S(x) = \sum_{i=1}^{s} \langle x, u_i\rangle u_i,$$

où $(u_i)_{i=1}^s$ est une base orthonormale quelconque de S.

Dans un espace de Hilbert on a le résultat suivant.

Théorème 6.7.2 *Soit $A \in \mathcal{B}(\mathcal{H})$, $A \neq \mathbf{0}$. Les affirmations suivantes sont équivalentes :*

1. *A est un projecteur orthogonale*
2. *Il existe un s.o.n[7] $(u_n)_{n\in\mathbb{N}}$ en $\mathcal{H}$ tel que*

$$Ax = \sum_{n\in\mathbb{N}} \langle x, u_n\rangle u_n \qquad \forall x \in \mathcal{H}.$$

Le cas échéant, A projette sur le sous-espace fermé $\overline{\text{span}(u_n,\, n \in \mathbb{N})}$.

Preuve.

1. $\implies$ 2. : on commence par observer qu'un projecteur orthogonale A est surjectif si et seulement s'il est l'opérateur identité. En fait, la condition $\text{Im}(A) = \mathcal{H}$ implique, par les propriétés 6. et 7. des projecteurs orthogonaux, que $\text{Im}(A)^\perp = \ker(A) = \{0_\mathcal{H}\}$, donc, grâce à la propriété 8., ça vaut : $Ax = x\ \forall x \in \mathcal{H}$, i.e. $A = id_\mathcal{H}$. Dans ce cas, tout s.o.n.c. $(u_n)_{n\in\mathbb{N}}$ de $\mathcal{H}$ réalise A car, d'un côté, $Ax = x$ et de l'autre côté, grâce au théorème 5.5.4 de caractérisation des s.o.n.c, on peut écrire $x = \sum_{n\in\mathbb{N}} \langle x, u_n\rangle u_n$. Vu que un s.o.n.c est, en particulier, un s.o.n. l'implication **1.** $\implies$ **2.** quand $\text{Im}(A) = \mathcal{H}$ est vraie.

Soit alors A un projecteur orthogonale tel que $\text{Im}(A) \subset \mathcal{H}$, i.e. $\text{Im}(A)$ est un sous-espace vectoriel fermé (par définition de projecteur orthogonale) et propre de $\mathcal{H}$, et donc il est lui-même un espace de Hilbert inclus proprement en $\mathcal{H}$.

Soit $(u_n)_{n\in\mathbb{N}}$ un s.o.n.c quelconque de $\mathbf{Im}(A)$. Vu nos hypothèses, $(u_n)_{n\in\mathbb{N}}$ est seulement un s.o.n. (et non pas, en général, un s.o.n.c) de $\mathcal{H}$. Pour tout $y \in \text{Im}(A)$ ça vaut la décomposition :

$$y = \sum_{n\in\mathbb{N}} \langle y, u_n\rangle u_n,$$

mais $\text{Im}(A) = \{Ax,\ x \in \mathcal{H}\}$, donc, en utilisant le fait que A, en tant que projecteur orthogonale, est auto-adjoint :

$$Ax = \sum_{n\in\mathbb{N}} \langle Ax, u_n\rangle u_n \underset{(A \text{ a.a.})}{=} \sum_{n\in\mathbb{N}} \langle x, Au_n\rangle u_n, \quad \forall x \in \mathcal{H},$$

mais comme A est l'identité $\text{Im}(A)$ et $u_n \in \text{Im}(A)\ \forall n \in \mathbb{N}$, alors $Au_n = u_n$, donc :

$$Ax = \sum_{n\in\mathbb{N}} \langle x, u_n\rangle u_n, \quad \forall x \in \mathcal{H},$$

i.e. le projecteur orthogonale A est réalisé sur le s.o.n. $(u_n)_{n\in\mathbb{N}}$ de $\mathcal{H}$ comme décrit dans le point **2.**

2. $\implies$ 1. : pour toute couple $x, y \in \mathcal{H}$, soit $(u_n)_{n\in\mathbb{N}}$ un s.o.n. de $\mathcal{H}$ tel que :

$$Ax = \sum_{m\in\mathbb{N}} \langle x, u_m\rangle u_m,\ Ay = \sum_{n\in\mathbb{N}} \langle y, u_n\rangle u_n,$$

alors, par la continuité du produit scalaire,

$$\langle x, Ay\rangle = \langle x, \sum_{n\in\mathbb{N}} \langle y, u_n\rangle u_n\rangle = \sum_{n\in\mathbb{N}} \overline{\langle y, u_n\rangle}\langle x, u_n\rangle = \sum_{n\in\mathbb{N}} \langle x, u_n\rangle\langle u_n, y\rangle.$$

7. On rappelle que, malgré que l'on écrive $(u_n)_{n\in\mathbb{N}}$, le s.o.n. peut être fini, i.e. avoir seulement un nombre fini de $u_n \neq 0$.

Mais, en utilisant encore la continuité du produit scalaire, comme déjà vu dans la preuve de l'identité de Parseval (théorème 5.5.4),

$$\begin{aligned}\langle x, A^\dagger A y\rangle = \langle Ax, Ay\rangle &= \langle \sum_{m\in\mathbb{N}} \langle x, u_m\rangle u_m, \sum_{n\in\mathbb{N}} \langle y, u_n\rangle u_n\rangle = \sum_{m\in\mathbb{N}}\sum_{n\in\mathbb{N}} \langle x, u_m\rangle\overline{\langle y, u_n\rangle}\langle u_m, u_n\rangle \\ &= \sum_{m\in\mathbb{N}}\sum_{n\in\mathbb{N}} \langle x, u_m\rangle\langle u_n, y\rangle\, \delta_{n,m} = \sum_{n\in\mathbb{N}} \langle x, u_n\rangle\langle u_n, y\rangle,\end{aligned}$$

donc $\langle x, Ay\rangle = \langle x, A^\dagger A y\rangle\ \forall x, y \in \mathcal{H}$, i.e. $A = A^\dagger A$ grâce au théorème 6.2.6. La caractérisation algébrique des projecteurs orthogonaux, théorème 6.7.1, permet de dire que A est un projecteur orthogonale.

Supposons que **1**. et **2**. soient vérifiés, alors :

1. $\implies$ $\mathrm{Im}(A) = \ker(A)^\perp$ et $\ker(A) = \mathrm{Im}(A)^\perp$;

2. $\implies$ $\mathrm{Im}(A) \subseteq \overline{\mathrm{span}(u_n,\ n\in\mathbb{N})}$, évident, et que $\ker(A) \subseteq (\,\overline{\mathrm{span}(u_n,\ n\in\mathbb{N})}\,)^\perp$, ce qui n'est pas totalement évident : pour $x = 0_\mathcal{H}$ c'est vrai, soit $x \in \ker(A)$, $x \neq 0_\mathcal{H}$, alors $Ax = 0 = \sum\limits_{n\in\mathbb{N}} \langle x, u_n\rangle u_n = \lim\limits_{N\to+\infty} \sum\limits_{n=1}^{N} \langle x, u_n\rangle u_n$, mais les vecteurs u_n sont linéairement indépendants car orthogonaux, donc, pour tout $N \in \mathbb{N}$, la combinaison linéaire $\sum\limits_{n=1}^{N} \langle x, u_n\rangle u_n$ est nulle si et seulement si les coefficients $\langle x, u_n\rangle$ sont nuls, i.e. $x \in (u_n,\ n\in\mathbb{N})^\perp = (\mathrm{span}(u_n,\ n\in\mathbb{N}))^\perp = (\,\overline{\mathrm{span}(u_n,\ n\in\mathbb{N})}\,)^\perp$, grâce aux propriétés du complément orthogonale.

En résumé : d'un côté $\mathrm{Im}(A) \subseteq \overline{\mathrm{span}(u_n,\ n\in\mathbb{N})}$, de l'autre $\ker(A) \subseteq (\,\overline{\mathrm{span}(u_n,\ n\in\mathbb{N})}\,)^\perp$, donc $\overline{\mathrm{span}(u_n,\ n\in\mathbb{N})} \subseteq \ker(A)^\perp = \mathrm{Im}(A)$, i.e. $\mathrm{Im}(A) = \overline{\mathrm{span}(u_n,\ n\in\mathbb{N})}$. □

Remarque : dans la preuve du théorème on a pu voir que tout s.o.n $(u_n)_{n\in\mathbb{N}}$ de $\mathrm{Im}(A)$ permet de réaliser un projecteur dans le sens défini par le théorème. Cela veut dire que, malgré que chaque terme de la somme peut être différent, l'action globale de l'opérateur sera la même pour tout s.o.n $(u_n)_{n\in\mathbb{N}}$ de $\mathrm{Im}(A)$.

Voyons une application remarquable de ce résultat avec l'exercice suivant, qui montre comment on peut trouver la meilleure approximation linéaire d'une parabole sur un intervalle réel grâce à la théorie de la projection orthogonale dans les espaces de Hilbert.

Exercice 6.6

Soient $f(x) = 1$ (fonction constante et égale à 1) e $g(x) = x$ deux éléments de $L^2[0,1]$. Calculer :

1. L'angle ϑ entre f et g.
2. Leur distance en $L^2[0,1]$.
3. La projection P_W de la fonction $h \in L^2[0,1]$, $h(x) = x^2$, sur l'espace $W = \mathrm{span}(f, g)$. Interpréter ce résultat.

Solution de l'exercice 6.6

1. L'angle entre f et g est calculé à partir de la définition de produit scalaire : $\langle f, g\rangle = \|f\|\,\|g\|\cos(\vartheta)$, donc il faut calculer $\langle f, g\rangle$, $\|f\|$, $\|g\|$:

$$\langle f, g\rangle = \int_0^1 x dx = \left[\frac{x^2}{2}\right]_0^1 = \frac{1}{2},$$

$$\|f\| = \left(\int_0^1 dx)\right)^{1/2} = 1, \quad \|g\| = \left(\int_0^1 x^2 dx)\right)^{1/2} = \left(\left[\frac{x^3}{3}\right]_0^1\right)^{1/2} = \frac{1}{\sqrt{3}},$$

en conclusion, $\cos(\vartheta) = \frac{\langle f,g\rangle}{\|f\|\|g\|} = \frac{\sqrt{3}}{2}$, donc $\vartheta = \frac{\pi}{6}$.

2. La distance :

$$d(f,g) = \|f-g\| = \left(\int_0^1 (f(x)-g(x))^2 dx\right)^{1/2} = \left(\int_0^1 (1-x)^2 dx\right)^{1/2}$$
$$= \left(-\left[\frac{(1-x)^3}{3}\right]_0^1\right)^{1/2} = \frac{1}{\sqrt{3}} = \frac{\sqrt{3}}{3}.$$

3. La projection sur W : utilisons la caractérisation de la projection qu'on vient de voir. Il faut construire une base orthonormée de W, pour cela on utilise la procédure de Gram-Schmidt en commençant par la fonction f, qui est un générateur de W et qui, en plus, a norme unitaire. Le deuxième (et dernier) élément de la base orthonormée de W sera alors :

$$\tilde{g}(x) = \frac{g(x) - \langle f, g\rangle f(x)}{\|g(x) - \langle f, g\rangle f(x)\|} = \frac{x - \frac{1}{2}}{\left\|x - \frac{1}{2}\right\|},$$

mais

$$\left\|x - \frac{1}{2}\right\| = \left(\int_0^1 \left(x - \frac{1}{2}\right)^2 dx\right)^{1/2} = \left(\frac{1}{3}\left[\left(x - \frac{1}{2}\right)^3\right]_0^1\right)^{1/2} = \frac{1}{2\sqrt{3}},$$

donc $\tilde{g}(x) = 2\sqrt{3}\left(x - \frac{1}{2}\right)$ et la base orthonormée cherchée est $B = (1, \sqrt{3}\,(2x-1))$. La projection orthogonale de $h(x) = x^2$ sur W est alors :

$$P_W h = \langle h, f\rangle f + \langle h, \tilde{g}\rangle \tilde{g},$$

par calcul direct on obtient $\langle h, f\rangle = \frac{1}{3}$ et $\langle h, \tilde{g}\rangle = \frac{\sqrt{3}}{6}$, donc :

$$P_W h(x) = \frac{1}{3} + \frac{\sqrt{3}}{6}\sqrt{3}\,(2x-1) = x - \frac{1}{6}.$$

L'*interprétation du résultat* est la suivante : les fonctions $[0,1] \underset{f}{\mapsto} 1$ et $[0,1] \underset{g}{\mapsto} x$ sont les générateurs de l'espace W des fonctions linéaires (les droites !) définies sur l'intervalle $[0,1]$. En fait, toute fonction linéaire $\ell : [0,1] \to \mathbb{K}$ peut s'écrire comme $\ell(x) = \alpha + \beta x$, $x \in [0,1]$ et $\alpha, \beta \in \mathbb{K}$, mais $\alpha + \beta x = \alpha f(x) + \beta g(x)\ \forall x \in [0,1]$, donc $\ell = \alpha f + \beta g$!

La fonction $[0,1] \underset{g}{\mapsto} x^2$ est une parabole définie sur le même intervalle. Par définition de projection orthogonale,

$$P_W h = \arg\min_{w \in W} \|h - w\|,$$

i.e. $P_W h$ est l'élément de W qui a la distance L^2 minimale entre h et les droites, on peut affirmer que la droite d'équation $y = x - \frac{1}{6}$ est la meilleure approximation de la parabole d'équation $y = x^2$, dans le sens de la norme L^2, sur l'intervalle $[0,1]$.

□

Dans la figure 6.1 on peut voir le graphe de cette approximation.

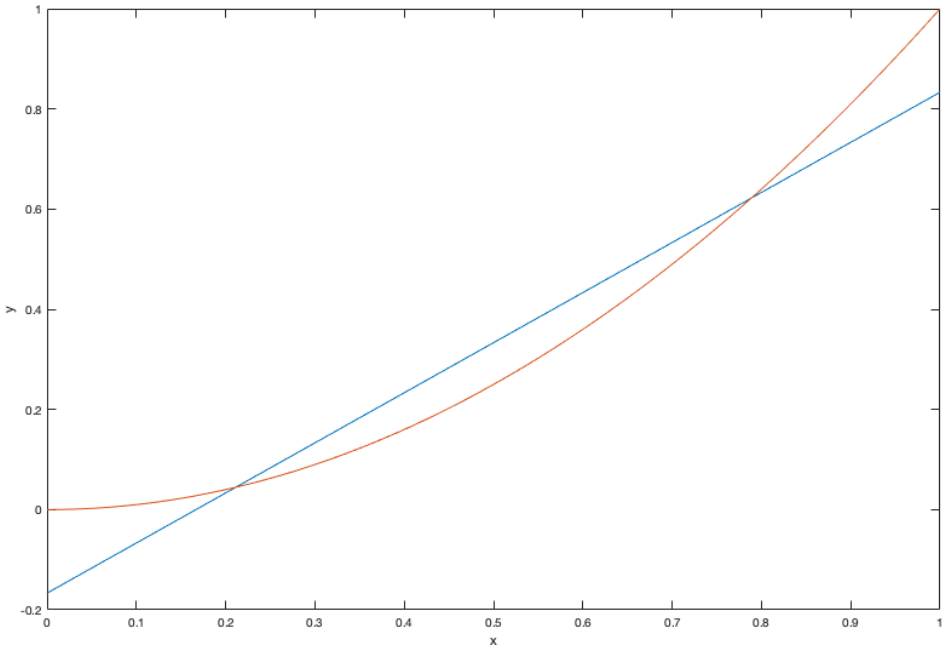


FIGURE 6.1 – La droite d'équation $y = x - \frac{1}{6}$ (en bleu) est la meilleure approximation de la parabole d'équation $y = x^2$ (en rouge) par rapport à la norme hilbertienne de $L^2[0,1]$.

Terminons avec une liste de propriétés des opérateurs de projection orthogonale (pour les preuves, consulter par exemple [1]). Pour tout $A, B \in \mathcal{B}(\mathcal{H})$, on écrit

$$[A,B] = AB - BA,$$

$[A,B]$ est dit le *commutateur* de A et B. Si $[A,B] = \mathbf{0}$, l'opérateur nul, i.e. $AB = BA$, alors on dit que A et B *commutent*.

Soient R et S deux sous-espaces vectoriels fermés de l'espace de Hilbert $\mathcal{H}$ et P_R, P_S les opérateurs de projection orthogonale sur R et S, respectivement.

Théorème 6.7.3 (Somme de projecteurs orthogonaux) *Les affirmations suivantes sont équivalentes :*

1. $P_R + P_S$ *est un projecteur orthogonal ;*
2. $P_R P_S = P_S P_R = \mathbf{0}$ *;*
3. $P_R(x) = 0 \ \forall x \in S$ *et* $P_S(x) = 0 \ \forall x \in R$ *;*
4. $R \perp S$.

De plus, si $P_R + P_S$ est un projecteur orthogonal, alors il projette sur $R + S$, le sous-espace somme de R et S.

Théorème 6.7.4 (Produit de projecteurs orthogonaux) *Les affirmations suivantes sont équivalentes :*

1. $P_R P_S$ *est un projecteur orthogonal ;*
2. $P_S P_R$ *est un projecteur orthogonal ;*
3. $[P_R, P_S] = \mathbf{0}$ *;*
4. $R = (R \cap S) \oplus (R \cap S^\perp)$ *;*
5. $S = (R \cap S) \oplus (R^\perp \cap S)$.

Si $P_R P_S$ et $P_S P_R$ sont des projecteurs orthogonaux, alors ils projettent sur $R \cap S$.

Théorème 6.7.5 (Différence de projecteurs orthogonaux) *Les affirmations suivantes sont équivalentes :*

1. $P_R - P_S$ *est un projecteur ;*
2. $P_R P_S = P_S P_R = P_S$ *;*
3. $R \cap S = S$, *i.e.* $S \subset R$.

Si $P_R - P_S$ est un projecteur orthogonale, alors il projette sur $R \cap S^\perp$.

Théorème 6.7.6 (Mélange de somme, différence et produit de projecteurs) *Si*

$$[P_S, P_R] = \mathbf{0},$$

alors $P_R + P_S - P_R P_S$ est un projecteur orthogonal qui projette sur $\mathrm{span}(R \cup S)$.

6.8 Les opérateurs isométriques et unitaires

Dans cette section on va déterminer les propriétés des opérateurs isométriques et unitaires dans un espace de Hilbert de dimension infinie, ainsi qu'une caractérisation algébrique et géométrique de ces opérateurs. L'opérateur adjoint va jouer encore une fois un rôle fondamentale dans la caractérisation algébrique, tandis que pour celle géométrique les protagonistes seront les s.o.n et les bases hilbertiennes.

Dans les espaces vectoriels V de dimension finie, un opérateur linéaire qui conserve le produit scalaire, i.e. $A : V \to V$, $\langle Ax, Ay\rangle = \langle x, y\rangle$, $\forall x, y \in V$ conserve aussi la norme des vecteurs (ça suffit de considérer $x = y$), i.e. $\|Ax\| = \|x\|$ $\forall x \in V$. Pour démontrer que ça vaut aussi l'inverse, considérons la formule de polarisation [8] (1.2.5), $\forall x, y \in V$:

$$\langle x, y\rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + i\,\|x+iy\|^2 - i\,\|x-iy\|^2\right)$$

et allons remplacer x, y par Ax, Ay, en utilisant, en plus, la linéarité de A :

$$\langle Ax, Ay\rangle = \frac{1}{4}\left(\|A(x+y)\|^2 - \|A(x-y)\|^2 + i\,\|A(x+iy)\|^2 - i\,\|A(x-iy)\|^2\right)$$

mais on a supposé que A conserve la norme, donc :

$$\langle Ax, Ay\rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + i\,\|x+iy\|^2 - i\,\|x-iy\|^2\right) = \langle x, y\rangle.$$

La norme génère canoniquement une distance, ou métrique : $d(x, y) = \|x-y\|$, c'est pour ça qu'un opérateur qui préserve le produit scalaire ou la norme est dit *isométrique*. Le seul vecteur qui a norme nulle est le vecteur 0_V, donc un opérateur A isométrique ne transforme jamais un vecteur non nul, qui a norme > 0, en 0, i.e. $\ker(A) = \{0_V\}$, donc $\dim(\ker(A)) = 0$ et alors, grâce au théorème de nullité + rang, $\dim(\mathrm{Im}(A)) = \dim(V)$. C'est-à-dire, *un endomorphisme isométrique en dimension finie est automatiquement surjectif.*

En un espace de Hilbert de dimension infinie c'est encore vrai que l'opérateur $A \in \mathcal{B}(\mathcal{H})$ conserve le produit scalaire si et seulement s'il est isométrique, mais il n'est plus vrai qu'un opérateur $A : \mathcal{H} \to \mathcal{H}$ isométrique est toujours surjectif. Un contre-exemple est offert par l'opérateur $A \in \mathcal{B}(\mathcal{H})$ défini par : $Au_n = u_{2n}$, où $(u_n)_{n\in\mathbb{N}}$ est une base hilbertienne quelconque de $\mathcal{H}$. Évidemment, A est isométrique, mais $\mathrm{Im}(A) = \overline{\mathrm{span}(u_k,\ k \in \mathbb{N},\ \ \text{pair})} \subset \mathcal{H}$, où l'inclusion est stricte car $(u_k,\ k \in \mathbb{N},\ k \text{ pair})$ n'est pas un s.o.n. complet, car sous-ensemble propre de $(u_n)_{n\in\mathbb{N}}$.

Ces considérations justifient les définitions suivantes :

Déf. 6.8.1 *L'opérateur $A : \mathcal{H} \to \mathcal{H}$ est dit :*

- ***isométrique**, si $\langle Ax, Ay\rangle = \langle x, y\rangle$, $\forall x, y \in \mathcal{H}$, ou, d'une manière équivalente, si $\|Ax\| = \|x\|$, $\forall x \in \mathcal{H}$;*
- ***unitaire**, si A est isométrique et surjectif.*

Calculons la norme d'un opérateur isométrique :

$$\|A\| = \sup_{\|x\|=1} \|Ax\| = \sup_{\|x\|=1} \|x\| = 1,$$

comme un opérateur unitaire est aussi isométrique, on peut dire que : **la norme des opérateurs isométriques et unitaires est 1**.

8. On considère le cas complexe ici, le cas réel est encore plus simple.

Exemples basiques d'opérateurs unitaires. Considérons $\mathbb{R}^n$ avec la tribu borélienne et la mesure de Lebesgue. Fixé un élément $a \in \mathbb{R}^n$ **tout opérateur de translation**

$$\begin{array}{rccl} T_a: & L^2(\mathbb{R}^n) & \longrightarrow & L^2(\mathbb{R}^n) \\ & f & \longmapsto & T_a f, \text{ où } T_a f(x) = f(x-a),\ \forall x \in \mathbb{R}^n, \end{array}$$

est unitaire, en fait il est évidemment bien défini, linéaire, et isométrique grâce à l'invariance de la mesure de Lebesgue par rapport aux translations. En plus, il est surjectif, car pour tout élément $g \in L^2(\mathbb{R}^n)$ il suffit de considérer $f \in L^2(\mathbb{R}^n)$, $f(x) = g(x+a)\ \forall x \in \mathbb{R}^n$, pour avoir $T_a f = g$.

Soit maintenant $R \in O(n)$ une matrice de rotation de $\mathbb{R}^n$, où $O(n)$ est le groupe orthogonale de dimension n, i.e. le groupe des matrices carrés R de dimension n orthogonales, i.e. telles que $R^t = R^{-1}$. **Tout opérateur de rotation**

$$\begin{array}{rccl} T_R: & L^2(\mathbb{R}^n) & \longrightarrow & L^2(\mathbb{R}^n) \\ & f & \longmapsto & T_R f, \text{ où } T_R f(x) = f(Rx),\ \forall x \in \mathbb{R}^n, \end{array}$$

est unitaire, grâce au fait que le Jacobien de la transformation, i.e. le déterminant de R, a valeur absolu 1 et donc les intégrales qui permettent le calcul de la norme de $T_R f$ et de f sont égales. De plus, il est surjectif car pour tout élément $g \in L^2(\mathbb{R}^n)$ il suffit de considérer $f \in L^2(\mathbb{R}^n)$, $f(x) = g(R^t x)\ \forall x \in \mathbb{R}^n$, pour avoir $T_R f = g$.

On souligne, comme cas particulier d'opérateur de rotation, le cas de la matrice inverse de l'identité : $P = -I$, dans ce cas $T_P f = f_P$, avec $f_P(x) = f(-x)$. T_P est dit **opérateur de parité**.

6.8.1 Les caractérisations des opérateurs isométriques et unitaires

On présente ici des résultats qui donnent une caractérisation très utile des opérateurs isométriques et unitaires.

Théorème 6.8.1 (Caractérisation algébrique des opérateurs isométriques) *$A \in \mathcal{B}(\mathcal{H})$ est un opérateur isométrique si et seulement si $A^\dagger A = id_{\mathcal{H}}$.*

Preuve. Soit A isométrique, alors $\forall x \in \mathcal{H}$:

$$\langle A^\dagger A x, x\rangle = \langle Ax, Ax\rangle = \|Ax\|^2 \underset{A \text{ isométrique}}{=} \|x\|^2 = \langle x, x\rangle,$$

i.e. $A^\dagger A = id_{\mathcal{H}}$.

Vice-versa, si $A^\dagger A = id_{\mathcal{H}}$, alors, $\forall x, y \in \mathcal{H}$:

$$\langle Ax, Ay\rangle = \langle A^\dagger A x, y\rangle = \langle x, y\rangle,$$

i.e. A préserve le produit scalaire et donc il est isométrique. □

Cette caractérisation a une conséquence très utilisée, par exemple, dans la théorie de l'optimisation.

Théorème 6.8.2 *Soit $A \in \mathcal{B}(\mathcal{H})$ isométrique, alors $AA^\dagger$ est un projecteur orthogonale.*

Preuve. On utilise le théorème de caractérisation algébrique des projecteurs orthogonaux : vu qu'on sait déjà que $AA^\dagger$ est auto-adjoint pour tout $A \in \mathcal{B}(\mathcal{H})$, il faut tout simplement vérifier que $AA^\dagger$ est borné et idempotent.

1. $AA^\dagger$ borné : $\forall x \in \mathcal{H}$ ça vaut

$$\|AA^\dagger x\| = \|A(A^\dagger x)\| \underset{A \text{ isométrique}}{=} \|A^\dagger x\| \leqslant \left\|A^\dagger\right\| \|x\| = \|A\| \|x\| \underset{\|A\|=1}{=} \|x\|.$$

2. $AA^\dagger$ idempotent : $(AA^\dagger)^2 = AA^\dagger(AA^\dagger) = A(A^\dagger A)A^\dagger \underset{A^\dagger A = id_{\mathcal{H}}}{=} AA^\dagger$.

□

$AA^\dagger$ projette sur son image, qui peut être caractérisée d'une manière très simple.

Théorème 6.8.3 *Soit $A \in \mathcal{B}(\mathcal{H})$ isométrique. L'image du projecteur orthogonale $AA^\dagger$ est $\mathrm{Im}(A)$, donc l'image d'un $A \in \mathcal{B}(\mathcal{H})$ isométrique est un sous-espace vectoriel fermé de $\mathcal{H}$.*

Preuve. Il faut prouver que $\mathrm{Im}(AA^\dagger) = \mathrm{Im}(A)$. On commence par observer que, en général, $\forall A, B \in \mathcal{B}(\mathcal{H})$, ça vaut que $\mathrm{Im}(AB) \subseteq \mathrm{Im}(A)$ et $\mathrm{Im}(AB) = \mathrm{Im}(A)$ si et seulement si B est surjectif, en fait :

$$\mathrm{Im}(A) = \{y \in \mathcal{H} \ : \ \exists x \in \mathcal{H} \ : \ Ax = y\}, \ \mathrm{Im}(AB) = \{z \in \mathcal{H} \ : \ \exists x \in \mathrm{Im}(B) \ : \ Ax = y\},$$

ce qui montre que $\mathrm{Im}(AB) = \mathrm{Im}(A)$ si et seulement si B est surjectif, sinon, les images de A sont non inférieures à celles de AB.

Si on prend $B = A^\dagger$, alors $\underline{\mathrm{Im}(AA^\dagger) \subseteq \mathrm{Im}(A)}$ $\forall A \in \mathcal{B}(\mathcal{H})$. Soit maintenant A isométrique, i.e. $A^\dagger A = id_{\mathcal{H}}$ et $y \in \mathrm{Im}(A)$, alors $\exists x \in \mathcal{H}$ tel que :

$$y = Ax = A(A^\dagger A)x = AA^\dagger(Ax),$$

i.e. $y \in \mathrm{Im}(AA^\dagger)$, d'où $\underline{\mathrm{Im}A \subseteq \mathrm{Im}(AA^\dagger)}$. Par conséquent, $\mathrm{Im}(AA^\dagger)$=$\mathrm{Im}(A)$.

Comme $AA^\dagger$ est un projecteur orthogonale, on sait que son image est fermée, donc $\mathrm{Im}(A)$ est fermée pour tout opérateur isométrique. □

Le fait que un opérateur isométrique ait une image fermée peut être démontré même directement avec une preuve très similaire à celle du théorème 6.3.3.

Si A est unitaire, i.e. isométrique et surjectif, alors $\mathrm{Im}(AA^\dagger) = \mathrm{Im}(A) = \mathcal{H}$ et alors $AA^\dagger = id_{\mathcal{H}}$.

Grâce au fait que $\ker(A^\dagger) = \mathrm{Im}(A)^\perp$ on a immédiatement des conditions suffisantes pour garantir l'inversibilité ou pas de l'adjoint d'un opérateur de $\mathcal{B}(\mathcal{H})$.

Théorème 6.8.4 *Soit $A \in \mathcal{B}(\mathcal{H})$.*

- *Si A est isométrique non surjectif, alors $\ker(A^\dagger) = \mathrm{Im}(A)^\perp \neq \{0_{\mathcal{H}}\}$, i.e. $A^\dagger$ n'est pas inversible.*
- *Si A est unitaire, alors $\ker(A^\dagger) = \mathcal{H}^\perp = \{0\}$, i.e. $A^\dagger$ est inversible.*

A la lumière de ces résultats, reprenons l'exemple de l'opérateur $Au_n = u_{2n}$ où $(u_n)_{n\in\mathbb{N}}$ est un s.o.n.c de $\mathcal{H}$. Comme $\|Au_n\| = \|u_{2n}\| = 1 = \|u_n\|$, A est isométrique mais non unitaire, car $\mathrm{Im}(A) = \overline{\mathrm{span}(u_k, \ k \text{ pair})} \subset \mathcal{H}$.

Calculons $A^\dagger$: $\langle Au_n, u_m\rangle = \langle u_n, A^\dagger u_m\rangle$, mais $\langle Au_n, u_m\rangle = \langle u_{2n}, u_m\rangle = \delta_{2n,m}$, et alors on peut écrire $\langle u_n, A^\dagger u_m\rangle = \delta_{2n,m}$, i.e.

$$A^\dagger u_m = \begin{cases} u_{\frac{m}{2}} & \text{si } m = 2n \\ 0 & \text{si } m \neq 2n. \end{cases}$$

On voit que $\ker(A^\dagger) = \overline{\text{span}(u_m,\ m \text{ impair})} = \text{Im}(A)^\perp$, ce qui confirme les résultats ci-dessus.

Allons voir maintenant une caractérisation algébrique encore plus fine des opérateurs unitaires.

Théorème 6.8.5 (Caractérisation algébrique des opérateurs unitaires) *$A \in \mathcal{B}(\mathcal{H})$, les affirmations suivantes sont équivalentes :*

1. *A est unitaire*
2. *A est inversible et $A^{-1} = A^\dagger \in \mathcal{B}(\mathcal{H})$*
3. *$A^\dagger A = AA^\dagger = id_\mathcal{H}$*
4. *$A^\dagger$ est unitaire.*

Preuve.

1. $\implies$ 2. : on sait que si A est unitaire, alors A est injectif, i.e. inversible sur son image, mais, par définition, A est surjectif, donc il est bijectif et inversible sur tout $\mathcal{H}$. Pour démontrer que $A^{-1} = A^\dagger$ on va écrire :

$$\langle A^\dagger Ax, y\rangle = \langle Ax, Ay\rangle = \langle x, y\rangle, \quad \forall x, y \in \mathcal{H}$$

donc $A^\dagger A = id_\mathcal{H}$ (ce qui montre que $A^\dagger$ est l'inverse gauche de A) et alors :

$$A^\dagger = A^\dagger(AA^{-1}) = (A^\dagger A)A^{-1} = id_\mathcal{H} A^{-1} = A^{-1}$$

i.e. $A^\dagger = A^{-1}$. Il reste seulement à prouver que $A^\dagger = A^{-1}$ est borné : comme A est surjectif, $\exists y \in \mathcal{H}$ tel que $x = Ay$, et, par unitarité : $\|x\| = \|Ay\| = \|y\|$, i.e. $\|x\| = \|y\|$ et alors :

$$\|A^\dagger x\| = \|A^{-1}x\| = \|A^{-1}Ay\| = \|y\| = \|x\| \quad \forall x \in \mathcal{H},$$

ce qui implique que $\|A^{-1}\| = \|A^\dagger\| = \sup_{\|x\|=1} \|x\| = 1$.

2. $\implies$ 3. :

$$A^\dagger = A^{-1} \implies A^\dagger A = A^{-1}A = id_\mathcal{H},$$

et aussi

$$A^\dagger = A^{-1} \implies AA^\dagger = AA^{-1} = id_\mathcal{H}.$$

3. $\implies$ 4. : de $AA^\dagger = id_\mathcal{H}$ on obtient : $\langle x, AA^\dagger y\rangle = \langle x, y\rangle\ \forall x, y \in \mathcal{H}$, mais $\langle x, AA^\dagger y\rangle = \langle A^\dagger x, A^\dagger y\rangle$, donc $\langle A^\dagger x, A^\dagger y\rangle = \langle x, y\rangle\ \forall x, y \in \mathcal{H}$, i.e. $A^\dagger$ est isométrique. Il reste à prouver que $A^\dagger$ est surjectif, pour cela on utilise l'autre identité, $A^\dagger A = id_\mathcal{H}$, qui implique

$$A^\dagger(Ay) = (A^\dagger A)y = id_\mathcal{H}(y) = y, \forall y \in \mathcal{H},$$

i.e. $\forall y \in \mathcal{H}\ \exists \xi = Ay \in \mathcal{H}$ tel que $A^\dagger \xi = y$, i.e. $A^\dagger$ est surjectif.

$\boxed{\mathbf{1.} \implies \mathbf{4.} \implies \mathbf{1.}}$: nous avons démontré que, donné un opérateur unitaire (quelconque), son adjoint est unitaire aussi. Donc, si on fait l'hypothèse que $A^\dagger$ soit un opérateur unitaire, alors $A^{\dagger\dagger}$ est unitaire, mais $A^{\dagger\dagger} = A$, donc $A^\dagger$ unitaire implique A unitaire. □

Ce résultat a comme conséquence qu'on peut étudier l'unitarité d'un opérateur A via celle de son adjoint, qui, parfois, peut être une tâche plus simple.

Nous soulignons dans le corollaire suivant le résultat sur la norme de l'adjoint et de l'inverse d'un opérateur unitaire.

Corollaire 6.8.1 *Si $A \in \mathcal{B}(\mathcal{H})$ est unitaire, alors $\|A\| = \|A^\dagger\| = \|A^{-1}\| = 1$.*

Écrivons avec $\mathcal{U}(\mathcal{H})$ l'ensemble des opérateurs unitaires sur un espace de Hilbert $\mathcal{H}$. Si $A, B \in \mathcal{U}(\mathcal{H})$, alors, par calcul direct, on peut vérifier que $AB \in \mathcal{U}(\mathcal{H})$. Le théorème qu'on vient de démontrer nous permet de dire que si $A \in \mathcal{U}(\mathcal{H})$ alors $A^{-1} = A^\dagger \in \mathcal{U}(\mathcal{H})$, i.e. $\mathcal{U}(\mathcal{H})$ vérifie les axiomes de groupe par rapport à la loi de composition d'opérateurs.

Déf. 6.8.2 *On appelle $\mathcal{U}(\mathcal{H})$ le* ***groupe unitaire de $\mathcal{H}$.*** *$\mathcal{U}(\mathcal{H})$ coïncide avec le* ***groupe des automorphismes de $\mathcal{H}$:*** *$Aut(\mathcal{H})$.*

Allons voir des applications de la caractérisation des opérateurs unitaires.

Soit $\mathcal{H} = L^2(X, \mathcal{A}, \mu)$ et $g \in L^\infty(X, \mathcal{A}, \mu)$, on a vu que l'opérateur de multiplication par g défini par :

$$\begin{array}{rccc} M_g : & L^2(X,\mathcal{A},\mu) & \longrightarrow & L^2(X,\mathcal{A},\mu) \\ & f & \longmapsto & M_g f = f \cdot g, \end{array}$$

où $f \cdot g(x) = f(x)g(x)\ \forall x \in X$ est linéaire et borné. De plus, on sait que $M_g^\dagger = M_{\bar{g}}$, donc $M_g M_g^\dagger = M_{g\bar{g}} = M_{|g|^2}$, et alors $M_g M_g^\dagger = id_{\mathcal{H}}$ si et seulement si $M_{|g|^2} = id_{\mathcal{H}}$, mais cela est équivalent à demander $|g|^2 = 1$, i.e. que dans la classe d'équivalence de g il y a au moins un représentant, qu'on écrira encore avec g par simplicité, de la forme $g(x) = e^{ih(x)}$, avec $h : X \to \mathbb{R}$ mesurable.

Allons appliquer te théorème qu'on vient de démontrer pour vérifier que : $\forall (u_n)_{n\in\mathbb{N}}$ s.o.n.c de $\mathcal{H}$, l'opérateur U défini comme ça

$$Uu_n = (-i)^n u_n$$

est un opérateur unitaire. Utilisons la caractérisation algébrique $U^\dagger U = UU^\dagger = id_{\mathcal{H}}$: d'un côté, par définition,

$$\langle Uu_n, u_n\rangle = \langle u_n, U^\dagger u_n\rangle, \tag{6.25}$$

de l'autre côté

$$\langle Uu_n, u_n\rangle = \langle (-i)^n u_n, u_n\rangle = (-i)^n \langle u_n, u_n\rangle = \langle u_n, \overline{(-i)^n} u_n\rangle \underset{(6.25)}{=} \langle u_n, U^\dagger u_n\rangle \quad \forall n \in \mathbb{N},$$

donc $U^\dagger u_n = \overline{(-i)^n} u_n$ et alors

$$U^\dagger U u_n = U^\dagger (U u_n) = \overline{(-i)^n} U u_n = \overline{(-i)^n}(-i)^n u_n = |(-i)^n|^2 u_n = u_n$$

et

$$UU^\dagger u_n = U(U^\dagger u_n) = U\overline{(-i)^n}u_n = (-i)^n\overline{(-i)^n}u_n = |(-i)^n|^2 u_n = u_n$$

comme $(u_n)_{n\in\mathbb{N}}$ est s.o.n.c, le fait que $U^\dagger U = UU^\dagger$ soient l'identité sur tout élément u_n peut être étendu à tout $\mathcal{H}$, en fait, pour tout $x \in \mathcal{H}$ ça vaut $x = \sum_{n\in\mathbb{N}}\langle x, u_n\rangle u_n$ et, par la linéarité et la continuité de $U^\dagger U$ et de $U^\dagger U$ on peut écrire :

$$U^\dagger U x = \sum_{n\in\mathbb{N}}\langle x, u_n\rangle U^\dagger U u_n = \sum_{n\in\mathbb{N}}\langle x, u_n\rangle u_n = x,$$

et pareil pour $UU^\dagger x$, i.e. $U^\dagger U = UU^\dagger = id_{\mathcal{H}}$ et ceci montre que U est unitaire.

Comme en dimension finie, les opérateurs unitaires permettent d'établir une relation d'équivalence entre opérateurs, formalisée dans la définition suivante.

Déf. 6.8.3 *Deux opérateurs $A, B \in \mathcal{B}(\mathcal{H})$ sont unitairement équivalents s'il existe un opérateur unitaire $U \in \mathcal{B}(\mathcal{H})$ tel que : $A = UBU^{-1}$.*

Nous n'avons pas l'opportunité de discuter des propriétés de l'équivalence unitaire, nous nous limitons à dire que l'équivalence unitaire préserve les propriétés des opérateurs, comme la continuité, l'inversibilité, le fait d'être auto-adjoint et les propriétés spectrales.

6.8.2 Relation entre opérateurs isométriques et unitaires et systèmes orthonormés

La dernière propriété des opérateurs isométriques et unitaires qu'on veut discuter est leur interaction avec les s.o.n. et les s.o.n.c. des espaces de Hilbert. En dimension finie, les opérateur isométriques et unitaires coïncident et ils transforment bases orthonormales en bases orthonormales. En dimension infinie ceci n'est plus vrai en général.

Théorème 6.8.6 (Caractérisation géométrique opérateurs isométriques) *$A \in \mathcal{B}(\mathcal{H})$ est isométrique si et seulement s'il transforme tout s.o.n.c $(u_n)_{n\in\mathbb{N}}$ de $\mathcal{H}$ en un s.o.n $(Au_n)_{n\in\mathbb{N}}$.*

Preuve.

$\boxed{\Rightarrow}$: pour tout s.o.n.c $(u_n)_{n\in\mathbb{N}}$ de $\mathcal{H}$, grâce à l'isométrie de A on peut écrire

$$\langle Au_n, Au_m\rangle = \langle u_n, u_m\rangle = \delta_{n,m} \qquad \forall n, m \in \mathbb{N},$$

donc $(Au_n)_{n\in\mathbb{N}}$ est s.o.n. de $\mathcal{H}$.

$\boxed{\Leftarrow}$: soit $A \in \mathcal{B}(\mathcal{H})$, $(u_n)_{n\in\mathbb{N}}$ est s.o.n.c de $\mathcal{H}$ et $(Au_n)_{n\in\mathbb{N}}$ un s.o.n de $\mathcal{H}$. Il faut prouver que A est isométrique.

D'un côté, le fait que $(u_n)_{n\in\mathbb{N}}$ soit un s.o.n.c implique que, pour tout $x \in \mathcal{H}$ ça vaut la décomposition en série de Fourier généralisée $x = \sum_{n\in\mathbb{N}}\langle x, u_n\rangle u_n$ et l'identité de Plancherel $\|x\|^2 = \sum_{n\in\mathbb{N}}|\langle x, u_n\rangle|^2$.

De l'autre côté, grâce à la continuité de A, nous pouvons écrire $Ax = \sum_{n\in\mathbb{N}} \langle x, u_n\rangle Au_n$, de plus, l'hypothèse que $(Au_n)_{n\in\mathbb{N}}$ soit s.o.n. de $\mathcal{H}$ nous permet d'invoquer la deuxième partie du théorème de Fischer-Riesz 5.5.3 pour dire que [9]

$$\|Ax\|^2 = \sum_{n\in\mathbb{N}} |\langle x, u_n\rangle|^2,$$

i.e. $\|Ax\|^2 = \|x\|^2$, donc A est isométrique. □

Pour les opérateurs unitaires ça vaut un résultat encore plus fort.

Théorème 6.8.7 (Caractérisation géométrique des opérateurs unitaires) *$A \in \mathcal{B}(\mathcal{H})$ est unitaire si et seulement s'il transforme tout s.o.n.c $(u_n)_{n\in\mathbb{N}}$ de $\mathcal{H}$ en un s.o.n.c $(Au_n)_{n\in\mathbb{N}}$.*

Preuve.

$\boxed{\Rightarrow}$: A unitaire est isométrique, donc, grâce au théorème précédent, $(Au_n)_{n\in\mathbb{N}}$ est un s.o.n. Il faut seulement démontrer que $(Au_n)_{n\in\mathbb{N}}$ est complet. Pour cela, on utilise une des propriétés qui caractérisent une base hilbertienne : comme on l'a vu dans le point **2**. du théorème 5.5.4, si $\langle x, Au_n\rangle = 0\ \forall n \in \mathbb{N}$ implique $x = 0_{\mathcal{H}}$ alors $(Au_n)_{n\in\mathbb{N}}$ est un s.o.n.c. Comme A est surjectif, il existe $y \in \mathcal{H}$ tel que $x = Ay$, donc la condition $\langle x, Au_n\rangle = 0\ \forall n \in \mathbb{N}$ devient :

$$\forall n \in \mathbb{N}\ :\ \langle Ay, Au_n\rangle \underset{A \text{ unitaire}}{=} \langle y, u_n\rangle = 0,$$

mais comme $(u_n)_{n\in\mathbb{N}}$ est un s.o.n.c de $\mathcal{H}$, $y = 0_{\mathcal{H}}$, ce qui implique $x = A0_{\mathcal{H}} = 0_{\mathcal{H}}$, et alors $(Au_n)_{n\in\mathbb{N}}$ est s.o.n.c de $\mathcal{H}$.

$\boxed{\Leftarrow}$: grâce au théorème précédent, on peut garantir que, si A transforme tout s.o.n.c $(u_n)_{n\in\mathbb{N}}$ de $\mathcal{H}$ en un s.o.n.c $(Au_n)_{n\in\mathbb{N}}$ de $\mathcal{H}$, alors A est, au moins, isométrique, donc ça reste seulement à démontrer la surjectivité.

Pour cela, on a vu que l'image d'un opérateur isométrique est fermé, i.e., par linéarité, $\mathrm{Im}(A) = \overline{\mathrm{span}(Au_n,\ n \in \mathbb{N})} = \mathcal{H}$, car $(Au_n)_{n\in\mathbb{N}}$ est s.o.n.c par hypothèse, donc A est surjectif, ce qui implique que A est unitaire. □

Terminons la section avec un simple exercice qui met en relation les opérateurs unitaires et les projecteurs orthogonaux.

Exercice 6.7

Soit $\mathcal{H}$ un espace de Hilbert. Montrer que les propriétés suivantes sont équivalentes. Suggestion : montrer que 1. $\iff$ 2. et 2. $\iff$ 3.

1. $A \in \mathcal{B}(\mathcal{H})$ est auto-adjoint et unitaire.
2. L'opérateur $P = \frac{1}{2}(A + id_{\mathcal{H}})$ est un projecteur orthogonal.

9. Explicitement, la deuxième partie du théorème de Fischer-Riesz dit que, donné un s.o.n $(v_n)_{n\in\mathbb{N}}$ dans un espace de Hilbert $\mathcal{H}$, si la série $\sum_{n\in\mathbb{N}} k_n v_n$ converge vers $y \in \mathcal{H}$, alors ça vaut $\|y\|^2 = \sum_{n\in\mathbb{N}} |k_n|^2$, dans notre cas : $y = Ax$, $v_n = Au_n$ et $k_n = \langle x, u_n\rangle$.

3. Il existe deux sous-espaces vectoriels fermés et orthogonaux $\mathcal{H}_1$ et $\mathcal{H}_2$ de $\mathcal{H}$ tels que $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$ et il existe un opérateur $A \in \mathcal{B}(\mathcal{H})$ tel que, pour tout $x = x_1 + x_2$, $x_i \in \mathcal{H}_i$, ça vaut : $Ax = x_1 - x_2$.

Solution de l'exercice 6.7

Commençons avec l'équivalence 1. $\iff$ 2.

1. $\implies$ 2. : par hypothèse, A est auto-adjoint, i.e. $A = A^\dagger$, et unitaire, i.e. $A^\dagger = A^{-1}$, mais alors $A = A^{-1}$ et donc $A^2 = AA = AA^{-1} = id_{\mathcal{H}}$. Utilisons ça pour montrer que $P = \frac{1}{2}(A + id_{\mathcal{H}})$ est auto-adjoint et idempotent, ce qui impliquera qu'il est un projecteur orthogonal :

$$P^\dagger = \frac{1}{2}(A + id_{\mathcal{H}})^\dagger \underset{\text{linéairité de } \dagger}{=} \frac{1}{2}(A^\dagger + id_{\mathcal{H}}^\dagger) = \frac{1}{2}(A + id_{\mathcal{H}}) = P$$

$$P^2 = \frac{1}{4}(A^2 + 2A + id_{\mathcal{H}}) = \frac{1}{4}(id_{\mathcal{H}} + 2A + id_{\mathcal{H}}) = \frac{1}{2}(A + id_{\mathcal{H}}) = P.$$

2. $\implies$ 1. : si ça vaut 2. alors on peut écrire $A = 2P - id_{\mathcal{H}}$ avec P projecteur orthogonal, démontrons que A est auto-adjoint et unitaire :

$$A^\dagger = 2P^\dagger - id_{\mathcal{H}}^\dagger = 2P - id_h = A$$

$$A^\dagger A = A^\dagger A = (2P - id_{\mathcal{H}})^2 = 4P^2 - 4P + id_h = 4P - 4P + id_h = id_h \implies A^\dagger = A^{-1}.$$

Analysons maintenant l'équivalence 2. $\iff$ 3.

2. $\implies$ 3. : si ça vaut 2. alors on sait que $\mathcal{H} = \mathrm{Im}(P) \oplus \ker(P)$, donc $\mathcal{H}_1 = \mathrm{Im}(P)$ et $\mathcal{H}_2 = \ker(P)$. De plus, écrit $\mathcal{H} \ni x = x_1 + x_2$, $x_1 \in \mathrm{Im}(P)$ et $x_2 \in \ker(P)$,

$$Px = x_1 \text{ et } P(x) = \frac{1}{2}(A + id_{\mathcal{H}})(x) = \frac{1}{2}(Ax + x_1 + x_2),$$

i.e. $x_1 = \frac{1}{2}(Ax + x_1 + x_2)$, et alors $Ax = 2x_1 - x_1 - x_2 = x_1 - x_2$.

3. $\implies$ 2. : supposons que 3. soit vérifiée. $P(x) = \frac{1}{2}(Ax + x) = \frac{1}{2}(x_1 - x_2 + x_1 + x_2) = x_1$ pour tout $x \in \mathcal{H}$, donc P, par définition, est le projecteur orthogonal $P_{\mathcal{H}_1}$ grâce à l'hypothèse que $\mathcal{H}_1$ et $\mathcal{H}_2$ sont orthogonaux et fermés.

□

6.9 La transformée de Fourier sur $\mathcal{S}(\mathbb{R}^n)$, $L^1(\mathbb{R}^n)$ et $L^2(\mathbb{R}^n)$

La transformée de Fourier sur $L^2(\mathbb{R}^n)$ est l'exemple le plus important d'opérateur unitaire sur $L^2(\mathbb{R}^n)$ par rapport à ses applications dans la physique théorique, la théorie des équations différentielles et le traitement des signaux, pour ne citer que trois champs d'application.

Malgré ça, la construction de cet opérateur n'est pas simple car $L^2(\mathbb{R}^n)$ n'est pas l'espace le plus naturel pour la transformée de Fourier. L'environnement le plus adapté à la transformée de Fourier s'avère être l'espace de Schwarz.

Plusieurs constructions de la transformée de Fourier sur $L^2(\mathbb{R}^n)$ sont disponibles dans la littérature, la plus courante, celle que l'on va adopter ici, consiste à définir la transformée de Fourier sur l'espace de Schwarz, pour pouvoir souligner ses propriétés remarquables, et après à opérer une extension à $L^2(\mathbb{R}^n)$ via une procédure limite. À côté de ce résultat, nous montrerons aussi une formule explicite qui passe par l'utilisation de la base hilbertienne de Hermite de $L^2(\mathbb{R}^n)$.

6.9.1 Un espace invariante par la transformée de Fourier : l'espace de Schwarz

Allons d'abord définir la transformée de Fourier sur l'espace de Schwarz $\mathcal{S}(\mathbb{R}^n)$ pour $n = 1, 2$ et après on va généraliser la définition pour n arbitraire (fini).

Déf. 6.9.1 *La transformée de Fourier sur $\mathcal{S}(\mathbb{R})$ est l'opérateur linéaire suivant*

$$\begin{array}{rccl} \hat{F}: & \mathcal{S}(\mathbb{R}) & \longrightarrow & \mathcal{S}(\mathbb{R}) \\ & f & \longmapsto & \hat{F}(f) = \hat{f},\ \text{où : } \hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) e^{-i\omega x} dm(x), \end{array}$$

où m est la mesure de Lebesgue sur $\mathbb{R}$ et $\omega \in \mathbb{R}$. L'anti-transformée de Fourier, ou transformée de Fourier inverse sur $\mathcal{S}(\mathbb{R})$ est l'opérateur linéaire suivant

$$\begin{array}{rccl} \check{F}: & \mathcal{S}(\mathbb{R}) & \longrightarrow & \mathcal{S}(\mathbb{R}) \\ & f & \longmapsto & \check{F}(f) = \check{f},\ \text{où : } \check{f}(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(\omega) e^{i\omega x} dm(\omega). \end{array}$$

Plus en général, la transformée de Fourier sur $\mathcal{S}(\mathbb{R}^n)$ est l'opérateur linéaire suivant

$$\begin{array}{rccl} \hat{F}: & \mathcal{S}(\mathbb{R}^n) & \longrightarrow & \mathcal{S}(\mathbb{R}^n) \\ & f & \longmapsto & \hat{F}(f) = \hat{f},\ \text{où : } \hat{f}(\omega) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(x) e^{-i\langle \omega, x\rangle} dm(x), \end{array}$$

où m est la mesure de Lebesgue sur $\mathbb{R}^n$, $\omega \in \mathbb{R}^n$ et $\langle \omega, x\rangle = \sum_{k=1}^{n} \omega_1 x_i$ est le produit scalaire Euclidien en $\mathbb{R}^n$. L'anti-transformée de Fourier, ou transformée de Fourier inverse sur $\mathcal{S}(\mathbb{R}^n)$ est l'opérateur linéaire suivant

$$\begin{array}{rccl} \check{F}: & \mathcal{S}(\mathbb{R}^n) & \longrightarrow & \mathcal{S}(\mathbb{R}^n) \\ & f & \longmapsto & \check{F}(f) = \check{f},\ \text{où : } \check{f}(x) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(\omega) e^{i\langle \omega, x\rangle} dm(\omega). \end{array}$$

Pour vérifier que les définitions sont bien posées, il faut vérifier que les intégrales existent et que $\hat{f}$ et $\check{f}$ sont des fonctions à décroissance rapide. L'existence des intégrales est immédiate si on considère que $\mathcal{S}(\mathbb{R}^n) \subset L^1(\mathbb{R}^n)$ et donc

$$\int_{\mathbb{R}^n} \left| f(x) e^{-i\langle \omega, x\rangle} \right| dm(x) = \int_{\mathbb{R}^n} |f(x)|\, dm(x) < +\infty,$$

et pareil pour l'anti-transformée de Fourier. Le fait que $\hat{f}$ et $\check{f}$ soient des fonctions à décroissance rapide peut être vérifié en itérant la dérivation sous signe d'intégrale et le théorème d'intégration par partie.

Dans la table suivante on résume les plus importantes propriétés de la transformée de Fourier pour une fonction $f \in \mathcal{S}(\mathbb{R})$, $a, b, c \in \mathbb{R}$, $a \neq 0$.
Observations importantes :

- $\hat{F}$ transforme le produit pour une constante en la division par la même constante (à un coefficient près) ;
- $\hat{F}$, comme la DFT, transforme la translation de la variable initiale en le produit par un exponentiel complexe ;

Fonction originale $f \in \mathcal{S}(\mathbb{R})$	Transformée de Fourier $\hat{f} \in \mathcal{S}(\mathbb{R})$
$f(ax)$	$\frac{1}{\lvert a\rvert}\hat{f}\left(\frac{\omega}{a}\right)$
$f(x-b)$	$e^{-i\omega b}\hat{f}(\omega)$
$f(ax-b)$	$\frac{e^{-i\omega\frac{b}{a}}}{\lvert a\rvert}\hat{f}\left(\frac{\omega}{a}\right)$
$e^{icx}f(x)$	$\hat{f}(\omega-c)$
$f'(x)$	$i\omega\hat{f}(\omega)$
$f''(x)$	$-\omega^2\hat{f}(\omega)$
$\frac{d^n f}{dx^n}$	$(i\omega)^n\hat{f}(\omega)$
$(-ix)^n f(x)$	$\frac{d^n\hat{f}}{d\omega^n}(\omega)$
$e^{-\frac{x^2}{2}}$	$e^{-\frac{\omega^2}{2}}$
$e^{-c^2x^2}$	$\frac{1}{c\sqrt{2}}e^{-\frac{\omega^2}{4c^2}}$

— $\hat{F}$ transforme la dérivation n-ième en le produit par une puissance de $i\omega$. Cette propriété est très importante pour transformer les équations différentielles en équations algébriques ;

— $\hat{F}$ transforme une Gaussienne d'écart-type 1 en une Gaussienne d'écart-type 1. Si la Gaussienne n'a pas écart-type 1, alors $\hat{F}$ **opère une inversion de l'écart-type** : une Gaussienne avec un écart-type petit, i.e. avec valeurs bien localisées autour de sa valeur moyenne, est transformée par $\hat{F}$ en une Gaussienne avec un écart-type grand, i.e. avec valeurs bien étalées par rapport à la moyenne, et vice-versa.

On veut démontrer la propriété : $\widehat{f'}(\omega) = i\omega\hat{f}(\omega)$.

Preuve. On commence par observer que pour $f : \mathbb{R} \to \mathbb{C}$, $f \in \mathcal{S}(\mathbb{R})$, alors $f(x) \underset{|x|\to+\infty}{\longrightarrow} 0$. On écrit la transformée de Fourier de f' :

$$\begin{aligned}
\widehat{f'}(\omega) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} f'(x)e^{-i\omega x}dx = \text{(par parties)} \\
&= \frac{1}{\sqrt{2\pi}}\left[f(x)e^{-i\omega x}\right]_{-\infty}^{+\infty} - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} f(x)(-i\omega)e^{-i\omega x}dx \\
&= 0 - (-i\omega)\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} f(x)e^{-i\omega x}dx \\
&= i\omega\hat{f}(\omega).
\end{aligned}$$

□

Le fait que la Gaussienne d'écart-type 1 soit invariante par rapport à la transformée de Fourier n'est pas un résultat intuitif, donc c'est utile de le démontrer. On va avoir besoin de deux lemmes.

Lemme 6.9.1 *Ça vaut :*

$$\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}dx = \sqrt{2\pi}.$$

Preuve. On écrit $I = \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx$, alors :

$$I^2 = \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx \cdot \int_{-\infty}^{+\infty} e^{-\frac{y^2}{2}} dy \underset{\text{(th. Fubini)}}{=} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(x^2+y^2)} dxdy.$$

Si on passe en coordonnés polaires (ρ, ϑ), $\rho \in [0, +\infty)$, $\vartheta \in [0, 2\pi)$ et on rappelle que le Jacobien en coordonnés polaires est ρ, on obtient :

$$\begin{aligned} I^2 &= \int_0^{+\infty} \int_0^{2\pi} e^{-\frac{\rho^2}{2}} \rho \, d\rho d\vartheta = \int_0^{2\pi} d\vartheta \int_0^{+\infty} e^{-\frac{\rho^2}{2}} \rho \, d\rho \\ &= 2\pi \left[-e^{-\frac{\rho^2}{2}} \right]_0^{+\infty} = 2\pi \left[-e^{-\infty} + e^0 \right] = 2\pi. \end{aligned}$$

Donc $I = \sqrt{2\pi}$. □

Lemme 6.9.2 *Ça vaut :*

$$\int_{-\infty}^{+\infty} e^{-\frac{(x+i\omega)^2}{2}} dx = \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx.$$

La preuve utilise le calcul des résidus de l'analyse complexe.

Maintenant on peut démontrer que :

$$\widehat{e^{-\frac{x^2}{2}}}(\omega) = e^{-\frac{\omega^2}{2}}.$$

Preuve.

$$\begin{aligned} \widehat{e^{-\frac{x^2}{2}}}(\omega) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} e^{-i\omega x} dx \\ &\underset{\cdot e^{-\frac{\omega^2}{2}} \cdot e^{\frac{\omega^2}{2}}}{=} \frac{e^{-\frac{\omega^2}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} e^{-i\omega x} e^{\frac{\omega^2}{2}} dx \\ &= \frac{e^{-\frac{\omega^2}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2+2i\omega x-\omega^2}{2}} dx \\ &= \frac{e^{-\frac{\omega^2}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{(x+i\omega)^2}{2}} dx \\ &\underset{\text{Lemme 6.9.2}}{=} \frac{e^{-\frac{\omega^2}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx \\ &\underset{\text{Lemme 6.9.1}}{=} \frac{e^{-\frac{\omega^2}{2}}}{\sqrt{2\pi}} \sqrt{2\pi} = e^{-\frac{\omega^2}{2}}. \end{aligned}$$

□

Pour démontrer la propriété d'inversion de l'écart-type, i.e. que $e^{-c^2x^2} \underset{\hat{F}}{\mapsto} \frac{1}{c\sqrt{2}} e^{-\frac{\omega^2}{4c^2}}$, nous allons utiliser une technique alternative à celle précédente (qui, bien sûr, pourrait être également utilisée).

Cette technique passe par la résolution d'une équation différentielle : en fait, si $f(x) = e^{-c^2x^2}$, alors $f'(x) = -2c^2xf(x)$, donc $f' + 2c^2xf = 0$ et, compte tenu des propriétés $f'(x) \underset{\hat{F}}{\mapsto} i\omega\hat{f}(\omega)$, $-ixf(x) \underset{\hat{F}}{\mapsto} \hat{f}'(\omega)$ et du fait que $2c^2xf = i2c^2(-ixf)$, en appliquant $\hat{F}$ aux deux côtés de l'équation différentielle précédente on arrive à écrire :

$$i\omega\hat{f}(\omega) + i2c^2\hat{f}'(\omega) = 0 \iff \omega\hat{f}(\omega) + 2c^2\hat{f}'(\omega) = 0. \tag{6.26}$$

Nous avons obtenu une *équation différentielle en variables séparables*[10] *par rapport à* $\hat{f}$. La technique canonique de résolution de ce type d'équation différentielle passe, d'abord, par la recherche de solutions constantes $\hat{f}(\omega) = C \in \mathbb{R}\ \forall\omega \in \mathbb{R}$, ce qui implique $\hat{f}'(\omega) = 0\ \forall\omega \in \mathbb{R}$, donc (6.26) devient $\omega\hat{f}(\omega) = 0$ qui peut être vérifiée pour toute $\omega \in R$ seulement quand $\hat{f}(\omega) \equiv 0$, donc la seule solutions constante de l'équation différentielle (6.26) peut être la fonction identiquement nulle. Néanmoins, cette fonction n'est pas coherente avec le fait que $\hat{f}(0) \neq 0$, en fait :

$$\begin{aligned} f(0) &\underset{\text{déf. de } \hat{f}(0)\,!}{=} \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} f(x)e^{-i0x}dx \\ &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} f(x)dx = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-c^2x^2}dx \\ &= \frac{1}{\sqrt{2\pi}c\sqrt{2}}\int_{\mathbb{R}} e^{-y^2/2}dy \underset{\text{Lemme (6.9.1)}}{=} \frac{1}{\sqrt{2\pi}c\sqrt{2}}\sqrt{2\pi} = \frac{1}{c\sqrt{2}}. \end{aligned}$$

donc, $\hat{f}(0) \equiv 0$ n'est pas une solution de (6.26). Supposons alors $\hat{f}(\omega) \neq 0$ et allons chercher les solutions non constantes de (6.26) avec la technique de séparation des variables, i.e. écrivons l'équations comme ceci :

$$\frac{\hat{f}'(\omega)}{\hat{f}(\omega)} = -\frac{\omega}{2c^2},$$

par intégrations des deux côtés : $\log|\hat{f}(\omega)| = -\frac{\omega^2}{4c^2} + \log C$, $C > 0$, où $\log C$ est la constante arbitraire issue de l'intégration, elle est écrite comme ça car, si on prend l'exponentiel aux deux côtés, on obtient

$$|\hat{f}(\omega)| = e^{-\frac{\omega^2}{4c^2}+\log C} = Ce^{-\frac{\omega^2}{4c^2}} \qquad \hat{f}(\omega) = \pm Ce^{-\frac{\omega^2}{4c^2}},$$

i.e. $\hat{f}(\omega) = Ke^{-\frac{\omega^2}{4c^2}}$, $K \in \mathbb{R}\backslash\{0\}$. Il suffit maintenant d'observer que $K = \hat{f}(0) = \frac{1}{c\sqrt{2}}$, comme vu ci-dessus, pour arriver à la solution $\hat{f}(\omega) = \frac{1}{c\sqrt{2}}e^{-\frac{\omega^2}{4c^2}}$.

Les propriétés de la transformée de Fourier définie sur $\mathcal{S}(\mathbb{R}^n)$, que l'on va résumer dans le tableau suivante (où $c \in \mathbb{R}$, $c \neq 0$, $a, b \in \mathbb{R}^n$, $k \in \{1, \ldots, n\}$), suivent directement de celles du cas $n = 1$ avec des changements plutôt évidents dans la technique de démonstration et, notamment, avec l'application du théorème de Fubini pour le calcul d'intégrales multiples.

Terminons cette section en citant le résultat qui justifie l'intérêt privilégié que l'on réserve à l'espace de Schwarz dans la théorie de la transformée de Fourier (et qui justifie aussi le nom de $\check{F}$).

10. Nous rappelons qu'une équation différentielle par rapport à la fonction $y(t)$ est dite en variables séparables si elle peut être écrite comme $y'(t) = f(y(t)) \cdot g(t)$, où f et g sont deux fonctions continues.

Fonction originale $f \in \mathcal{S}(\mathbb{R}^n)$	Transformée de Fourier $\hat{f} \in \mathcal{S}(\mathbb{R}^n)$
$f(cx)$	$\frac{1}{\lvert c\rvert}\hat{f}\left(\frac{\omega}{c}\right)$
$f(x-b)$	$e^{-i\langle\omega,b\rangle}\hat{f}(\omega)$
$f(cx-b)$	$\frac{e^{-i\omega\frac{b}{c}}}{\lvert c\rvert}\hat{f}\left(\frac{\omega}{c}\right)$
$e^{i\langle a,x\rangle}f(x)$	$\hat{f}(\omega-a)$
$\partial_{x_k}f(x)$	$i\omega_k\hat{f}(\omega)$
$\partial^2_{x_k}f(x)$	$-\omega_k^2\hat{f}(\omega)$
$\partial^n_{x_k}f(x)$	$(i\omega_k)^n\hat{f}(\omega)$
$(-ix_k)^n f(x)$	$\partial^n_{x_k}\hat{f}(\omega)$
$e^{-\frac{\lVert x\rVert^2}{2}}$	$e^{-\frac{\lVert\omega\rVert^2}{2}}$
$e^{-c^2\lVert x\rVert^2}$	$\frac{1}{c\sqrt{2}}e^{-\frac{\lVert\omega\rVert^2}{4c^2}}$

Théorème 6.9.1 *La transformation $\hat{F}$ est un isomorphisme linéaire de $\mathcal{S}(\mathbb{R}^n)$ en lui même et sa transformation inverse est $\check{F}$: $\check{F} = \hat{F}^{-1}$. De plus, si on interprète $f \in \mathcal{S}(\mathbb{R}^n)$ comme une fonction de $L^2(\mathbb{R}^n)$, alors : $\lVert f\rVert = \lVert\hat{f}\rVert$ $\forall f \in \mathcal{S}(\mathbb{R}^n) \subset L^2(\mathbb{R}^n)$.*

Donc, l'espace de Schwarz est invariant par rapport à l'application de la transformée de Fourier $\hat{F}$, qui a une formule intégrale explicite, qui possède une inverse donnée par $\check{F}$ et qui conserve la norme des fonctions à décroissance rapide quand on les interprète comme des éléments de $L^2(\mathbb{R}^n)$. Dans aucun autre espace fonctionnel de dimension infinie la transformée de Fourier a toutes ces propriétés à la fois.

En fait, on va voir tout de suite que, en $L^1(\mathbb{R})$ on doit renoncer à l'invariance de l'espace, et donc à l'inversibilité de l'opérateur linéaire sur tout l'espace, tandis que en $L^2(\mathbb{R})$ on doit renoncer à la formule intégrale explicite.

6.9.2 Extension de la transformée de Fourier de $\mathcal{S}(\mathbb{R}^n)$ à $L^1(\mathbb{R}^n)$: le théorème de Riemann-Lebesgue

Les éléments de l'espace de Schwarz sont des fonctions trop régulières pour être exhaustives, surtout par rapport aux applications.

Il est donc impératif de s'intéresser à l'extension de la transformée de Fourier à des espaces de fonctions moins régulières comme $L^1(\mathbb{R})$ et $L^2(\mathbb{R})$. Dans cette section on va considérer $L^1(\mathbb{R})$: par rapport à cet espace, il vaut le célèbre résultat qui suit.

Théorème 6.9.2 (Théorème de Riemann-Lebesgue) *L'opérateur $\hat{F}$ de la section 6.9.1 peut être étendu d'une manière unique à l'opérateur linéaire **injectif** et **continu** définit de la manière suivante*

$$\begin{array}{rccl} \hat{F}_1: & L^1(\mathbb{R}^n) & \longrightarrow & \mathcal{C}_\infty(\mathbb{R}^n) \\ & f & \longmapsto & \hat{F}_1(f) =,\ \text{où : } \hat{F}_1 f(\omega) = \frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n} f(x)e^{-i\langle\omega,x\rangle}dm(x). \end{array}$$

Le même énoncé vaut pour l'extension de $\check{F}$ à $L^1(\mathbb{R}^n)$ avec la formule intégrale correspondante, i.e.

$$\begin{array}{rccl} \check{F}_1: & L^1(\mathbb{R}^n) & \longrightarrow & \mathcal{C}_\infty(\mathbb{R}^n) \\ & f & \longmapsto & \check{F}_1(f) =,\ \text{où : } \check{F}_1 f(x) = \frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n} f(\omega)e^{i\langle\omega,x\rangle}dm(\omega). \end{array}$$

On rappelle que $\mathcal{C}_\infty(\mathbb{R}^n)$ est l'espace des fonctions définies et continues sur $\mathbb{R}^n$ qui tendent vers 0 à l'infini muni de la norme $\|f\|_\infty = \sup_{x\in\mathbb{R}^n} |f(x)|$.

Observations :

— Le théorème de Riemann-Lebesgue garantit que la formule intégrale de la transformée de Fourier reste valide pour les éléments de $L^1(\mathbb{R}^n)$, ce qui est très important car les fonctions absolument intégrables au sens de Lebesgue sont énormément plus répandues dans les applications que celles à décroissance rapide.

— L'injectivité de F_1 permet son inversion sur l'image $F_1(L^1(\mathbb{R}^n)) \subset \mathcal{C}_\infty(\mathbb{R}^n)$, mais pas sur $L^1(\mathbb{R}^n)$. Un contre-exemple classique quand $n = 1$ est fourni par la fonction indicatrice de l'intervalle $[-1, 1]$ en $\mathbb{R}$, i.e. $\chi_{[-1,1]}$: elle appartient à $L^1(\mathbb{R})$, mais, par calcul direct, on obtient

$$\hat{F}_1(\chi_{[-1,1]})(\omega) = \sqrt{\frac{2}{\pi}}\frac{\sin\omega}{\omega}, \tag{6.27}$$

qui, bien sûr, appartient à $\mathcal{C}_\infty(\mathbb{R})$, mais elle n'appartient pas à $L^1(\mathbb{R})$...en fait elle appartient à $L^2(\mathbb{R})$... Par conséquent, $\check{F}_1$, qui est défini sur tout $L^1(\mathbb{R})$, n'est pas l'inverse de $\hat{F}_1$.

6.9.3 Extension de la transformée de Fourier à un opérateur unitaire sur $L^2(\mathbb{R}^n)$: la transformée de Fourier Plancherel

La technique classique pour étendre $\hat{F}$ à $L^2(\mathbb{R}^n)$ consiste en l'utilisation de ce théorème fondamentale de l'analyse fonctionnelle, qu'on énonce et démontre après la définition d'extension d'un opérateur linéaire.

Déf. 6.9.2 (Extension d'un opérateur linéaire) *Soient E, V, W des espaces vectoriels sur le même corps $\mathbb{K}$ et soit E un sous-espace vectoriel de V. Soit $A : E \to W$ un opérateur linéaire. On dit que l'opérateur linéaire $B : V \to W$ est une extension de A sur V si la restriction de B à E coïncide avec A, i.e. si $Ax = Bx$ $\forall x \in E$.*

Théorème 6.9.3 (Théorème d'extension d'un opérateur linéaire borné) *Soient E et F deux espaces vectoriels normés, avec F espace de Banach. Soit $A : D_A \subseteq E \to F$ un opérateur linéaire borné, où D_A est un sous-espace vectoriel de E. Alors, il existe seulement un opérateur linéaire $\overline{A}$ avec les propriétés suivantes :*

1. *le domaine de $\overline{A}$ est l'adhérence de D_A en E : $D_{\overline{A}} = \overline{D_A}$;*
2. *$\overline{A}$ est continu : $A \in \mathcal{B}(\overline{D_A}, F)$;*
3. *$\|\overline{A}\| = \|A\|$.*

Cet opérateur est défini comme ça : soit $(x_n)_{n\in\mathbb{N}} \subset D_A$ une suite quelconque d'éléments qui converge vers $x \in \overline{D_A}$, alors

$$\begin{array}{rccl} \overline{A}: & \overline{D_A} \subseteq E & \longrightarrow & F \\ & x & \longmapsto & \boxed{\overline{A}x = \lim\limits_{n\to+\infty} Ax_n}. \end{array}$$

Preuve. Soit x un élément quelconque de $D_{\overline{A}} = \overline{D_A}$, alors, par définition, il existe une suite $(x_n)_{n\in\mathbb{N}} \subset D_A$ telle que $x = \lim\limits_{n\to+\infty} x_n$. $(x_n)_{n\in\mathbb{N}}$, en tant que suite convergente, est de Cauchy et, vu que A est continu, aussi la suite $(Ax_n)_{n\in\mathbb{N}} \subset F$ est de Cauchy, grâce au théorème 6.2.5.

Comme F est un espace de Banach, il existe $y = \lim\limits_{n\to+\infty} Ax_n$, par conséquent, l'opérateur $\overline{A} : \overline{D_A} \to F$, $\overline{A}x = \lim\limits_{n\to+\infty} Ax_n$ est bien défini et linéaire, car il est défini via l'opération de limite, qui est linéaire.

De plus, $\overline{A}$ ne dépend pas de la suite qui converge vers x, en fait, si $(x'_n)_{n\in\mathbb{N}} \subset E$ est une autre suite telle que $x = \lim\limits_{n\to+\infty} x'_n$, alors

$$\begin{aligned}
\| \lim_{n\to+\infty} Ax_n - \lim_{n\to+\infty} Ax'_n \| &= \lim_{n\to+\infty} \|Ax_n - Ax'_n\| \quad \text{(Continuité } \|\ \|\text{)} \\
&= \lim_{n\to+\infty} \|A(x_n - x'_n)\| \leqslant \lim_{n\to+\infty} \|A\|\|x_n - x'_n\| \quad (A \text{ borné}) \\
&= \|A\| \lim_{n\to+\infty} \|x_n - x'_n\| = \|A\| \, \| \lim_{n\to+\infty} (x_n - x'_n)\| \quad \text{(Continuité } \|\ \|\text{)} \\
&= \|A\| \, \| \lim_{n\to+\infty} x_n - \lim_{n\to+\infty} x'_n\| = \|A\|\|x - x\| = 0.
\end{aligned}$$

Évidemment, tout $x \in D_A$ peut être identifié comme la limite de la suite constante $x_n = x$ $\forall n \in \mathbb{N}$, donc, vu l'indépendance de la définition de $\overline{A}$ par rapport à la suite choisie, $\overline{A}x = Ax$ $\forall x \in D_A$, i.e. la restriction de $\overline{A}$ à D_A est A et, inversement, $\overline{A}$ est une extension de A sur $\overline{D_A}$.

Le fait que $\overline{A}$ soit un opérateur borné sur $\overline{D_A}$ peut être vérifié en considérant la limite de l'inégalité $\|Ax_n\| \leqslant \|A\|\|x_n\|$, en fait, la limite préserve la relation d'ordre, i.e.

$$\lim_{n\to+\infty} \|Ax_n\| \leqslant \lim_{n\to+\infty} \|A\|\|x_n\|$$

et, par la continuité de la norme on a

$$\| \lim_{n\to+\infty} Ax_n\| \leqslant \|A\| \, \| \lim_{n\to+\infty} x_n\| \iff \|\overline{A}x\| \leqslant \|A\|\|x\|,$$

pour tout $x \in \overline{D_A}$, i.e. $\overline{A}$ est borné, i.e. continu.

Démontrons maintenant que toute autre extension de A à $\overline{D_A}$ coïncide avec $\overline{A}$. Soit B une autre extension borné de A sur $\overline{D_A}$, alors, pour tout $x \in \overline{D_A}$, il existe une suite $(x_n)_{n\in\mathbb{N}} \subset D_A$, telle que $x = \lim\limits_{n\to+\infty} x_n$ et, par *définition de* $\overline{A}$ et grâce à la *continuité de* B on a

$$\overline{A}x - Bx = \lim_{n\to+\infty} Ax_n - B\left(\lim_{n\to+\infty} x_n\right) = \lim_{n\to+\infty} Ax_n - \lim_{n\to+\infty} Bx_n = \lim_{n\to+\infty} (Ax_n - Bx_n).$$

Pour tout $n \in \mathbb{N}$, $x_n \in D_A$ et, comme B est une extension de A, par définition $Bx_n = Ax_n$ $\forall n \in \mathbb{N}$, alors $Ax_n - Bx_n = 0$ $\forall n \in \mathbb{N}$ et donc $\overline{A}x - Bx = \lim\limits_{n\to+\infty} (Ax_n - Bx_n) = \lim\limits_{n\to+\infty} 0 = 0$, i.e. $\overline{A} = B$.

Il nous reste à démontrer seulement que l'extension est isométrique, i.e. que $\|\overline{A}\| = \|A\|$. Nous avons déjà vu que $\left\|\overline{A}x\right\| \leqslant \|A\| \, \|x\|$ pour tout $x \in \overline{D_A}$, donc

$$\left\|\overline{A}\right\| = \sup_{\|x\|=1} \left\|\overline{A}x\right\| \leqslant \sup_{\|x\|=1} \|A\| \, \|x\| = \|A\|,$$

alors, si on démontre que $\left\|\overline{A}\right\| \geqslant \|A\|$, on aura prouvé l'isométrie de l'extension. Ceci est immédiat si on considère la définition de norme opératorielle suivante :

$$\|\overline{A}\| = \sup\left\{ \frac{\left\|\overline{A}x\right\|}{\|x\|},\ x \in \overline{D_A}\backslash\{0_E\} \right\} \geqslant \sup\left\{ \frac{\|Ax\|}{\|x\|},\ x \in D_A\backslash\{0_E\} \right\} = \|A\|,$$

vu que $\overline{A}x = Ax\ \forall x \in D_A$ et que $D_A \subseteq \overline{D_A}$, d'où $\left\|\overline{A}\right\| = \|A\|$ et le théorème est complètement démontré. □

Grâce au théorème d'extension et au fait que $\overline{\mathcal{S}(\mathbb{R}^n)} = L^2(\mathbb{R}^n)$, nous pouvons étendre la transformée de Fourier de l'espace de Schwarz à $L^2(\mathbb{R}^n)$ via la formule limite du théorème d'extension, comme formalisé dans le résultat suivant.

Théorème 6.9.4 *Les opérateurs $\hat{F}$ et $\check{F}$ qui définissent la transformée et l'antitransformée de Fourier sur $\mathcal{S}(\mathbb{R}^n)$, respectivement, peuvent être étendus d'une manière unique à deux opérateurs* ***unitaires*** *F et $\tilde{F}$ sur $L^2(\mathbb{R}^n)$, de plus $\tilde{F} = F^{-1}$.*

L'opérateur F est dit ***transformée de Fourier-Plancherel*** *et il est défini comme ça : soit $(f_n)_{n\in\mathbb{N}} \subset \mathcal{S}(\mathbb{R}^n)$ une suite quelconque d'éléments de $\mathcal{S}(\mathbb{R}^n)$ qui convergent vers $f \in L^2(\mathbb{R}^n)$, alors*

$$\begin{array}{rccc} F: & L^2(\mathbb{R}^n) & \longrightarrow & L^2(\mathbb{R}^n) \\ & f & \longmapsto & F(f) = \lim\limits_{n\to+\infty} \hat{f}_n, \end{array}$$

et :

$$\begin{array}{rccc} F^{-1}: & L^2(\mathbb{R}^n) & \longrightarrow & L^2(\mathbb{R}^n) \\ & f & \longmapsto & F(f) = \lim\limits_{n\to+\infty} \check{f}_n, \end{array}$$

Donc, la transformée de Fourier-Plancherel F sur $L^2(\mathbb{R}^n)$ a les propriétés (essentielles !) d'être un opérateur linéaire unitaire (et donc continu) et d'avoir l'opérateur unitaire F^{-1} comme inverse.

Ce qui rend moins naturel d'examiner la transformée de Fourier dans $L^2(\mathbb{R}^n)$ par rapport à $\mathcal{S}(\mathbb{R}^n)$ est la manque d'une formule intégrale valide pour tous les éléments de $L^2(\mathbb{R}^n)$. Une solution partielle à ce problème est fourni par le théorème suivant.

Théorème 6.9.5 *Si $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, alors $F = \hat{F}_1$ et , i.e. pour les fonctions qui appartiennent à $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ la formule intégrale de la transformée de Fourier reste valide.*

Heureusement, comme on l'a vu dans la section 4.4.4, parmi les fonctions de $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ on a les fonctions bornées de $L^1(\mathbb{R}^n)$ et celles de $L^2(\mathbb{R}^n)$ qui s'annulent hors d'un sous-ensemble compact, que l'on rencontre souvent dans les applications.

6.9.4 Relation entre transformée de Fourier-Plancherel et base hilbertienne de Hermite

Une base hilbertienne très importante de $L^2(\mathbb{R})$ est celle de Hermite, définie comme ça :

$$u_n(x) = \frac{(-1)^n}{\sqrt{2^n n!\sqrt{\pi}}} e^{\frac{1}{2}x^2} \frac{d^n}{dx^n} e^{-x^2}, \quad x \in \mathbb{R},\ n \in \mathbb{N}.$$

Il est possible de démontrer que les fonctions u_n sont à décroissance rapide, donc leur transformée de Fourier est obtenue via l'application de la formule intégrale, i.e.

$$\begin{aligned} Fu_n(\omega) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} u_n(x) e^{-i\omega x} dm(x) \\ &= \frac{1}{\sqrt{2\pi}} \frac{(-1)^n}{\sqrt{2^n n!\sqrt{\pi}}} \int_{\mathbb{R}} e^{-i\omega x + \frac{1}{2}x^2} \frac{d^n}{dx^n} e^{-x^2} dm(x). \end{aligned}$$

Avec des simples manipulations algébriques, il est possible de démontrer que :

$$Fu_n = (-i)^n u_n, \qquad n \in \mathbb{N},$$

et donc F coïncide avec l'opérateur unitaire introduit dans la section 6.8.1. Grâce à la continuité de F, on peut écrire :

$$\boxed{Ff = \sum_{n\in\mathbb{N}} (-i)^n \langle f, u_n\rangle u_n}, \qquad \forall f \in L^2(\mathbb{R}),\ (u_n)_{n\in\mathbb{N}} \ : \ \text{s.o.n.c de Hermite.}$$

6.9.5 La transformée de Fourier et la convolution

Nous dédions une section à part aux propriétés de la transformée de Fourier par rapport à la convolution, vu l'importance et l'utilité (théorique et appliquée) de ces propriétés dans les mathématiques. Une référence très complète par rapport ce sujet est offerte par le livre [10].

Nous allons définir la convolution et, après avoir discuté ses propriétés basiques, nous allons prouver la propriété la plus connue et plus importante de la transformée de Fourier en $L^1(\mathbb{R}^n)$ par rapport à la convolution : le fait que la transformée de Fourier modifie le produit de convolution en le produit ponctuel des transformées (à un coefficient près).

Déf. 6.9.3 *Soient $f, g : \mathbb{R}^n \to \mathbb{R}$, la convolution entre f et g est la fonction $f * g$ définie par :*

$$(f * g)(x) = \int_{\mathbb{R}^n} f(x-y)g(y)dm(y), \qquad \forall x \in \mathbb{R}^n,$$

toutefois que l'intégrale existe dans le sens de Lebesgue.

Théorème 6.9.6 (Propriétés basiques de la convolution) *Les propriétés suivantes sont vraies :*

1. *Si $f \in L^1(\mathbb{R}^n)$ et $g \in L^\infty(\mathbb{R}^n)$ ou vice-versa, alors la convolution est bien définie ;*
2. *Si $f, g \in L^2(\mathbb{R}^n)$, alors la convolution est bien définie et, en général, est un élément de $L^\infty(\mathbb{R}^n)$;*
3. *Si $f, g \in L^1(\mathbb{R}^n)$, alors la convolution est bien définie et appartient à $L^1(\mathbb{R}^n)$, qui devient une algèbre de Banach par rapport à la convolution ;*
4. *Si le produit de convolution est défini, alors :*
   — *$f * (\alpha g + \beta h) = \alpha f * g + \beta f * h$ (linéarité) ;*
   — *$f * g = g * f$ (commutativité) ;*
   — *$f * (g * h) = (f * g) * h$ (associativité).*

Preuve. Nous allons prouver seulement les deux premières propriétés, la preuve des propriétés suivantes est laissé comme exercice.

1. Si $f \in L^1(\mathbb{R}^n)$ et $g \in L^\infty(\mathbb{R}^n)$, alors

$$\int_{\mathbb{R}^n} |f(x-y)g(y)|dm(y) \leqslant \|g\|_\infty \int_{\mathbb{R}^n} |f(x-y)|dm(y) = \|g\|_\infty \|f\|_1,$$

grâce à l'invariance par translations de la mesure de Lebesgue.

2. Si $f, g \in L^2(\mathbb{R}^n)$, alors, par l'inégalité de Hölder (4.19) :

$$\int_{\mathbb{R}^n} |f(x-y)g(y)|dm(y) \leqslant \left(\int_{\mathbb{R}^n} |f(x-y)|^2 dm(y)\right)^{1/2} \left(\int_{\mathbb{R}^n} |g(y)|^2 dm(y)\right)^{1/2} = \|f\|_2 \|g\|_2,$$

encore une fois grâce à l'invariance par translations de la mesure de Lebesgue.

□

Théorème 6.9.7 (Convolution et transformée de Fourier en L^1) *Soient $f, g \in L^1(\mathbb{R}^n)$, alors la transformée de Fourier vérifie cette propriété :*

$$\widehat{f * g} = (2\pi)^{n/2} \hat{f} \cdot \hat{g}.$$

Preuve. Nous écrivons tout simplement la définition de convolution et de transformée de Fourier et nous appliquons deux fois le théorème de Fubini, avec une petit astuce algébrique au milieu :

$$\begin{aligned}
\widehat{(f * g)}(\omega) &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} f(x-y)g(y)dm(y)\right) e^{-i\langle\omega,x\rangle} dm(x) \\
&= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x-y)g(y)e^{-i\langle\omega,x\rangle} dm(x)dm(y) \quad \text{(Fubini)} \\
&= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x-y)g(y)e^{-i\langle\omega,x-y+y\rangle} dm(x)dm(y) \quad \text{(Astuce algébrique)} \\
&= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x-y)e^{-i\langle\omega,x-y\rangle} g(y)e^{-i\langle\omega,y\rangle} dm(x)dm(y) \\
&= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(x-y)e^{-i\langle\omega,x-y\rangle} dm(x) \int_{\mathbb{R}^n} g(y)e^{-i\langle\omega,y\rangle} dm(y) \quad \text{(Fubini)} \\
&= (t = x-y,\ dm(t) = dm(x-y)) \\
&= (2\pi)^{n/2} \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(t)e^{-i\langle\omega,t\rangle} dm(t) \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} g(y)e^{-i\langle\omega,y\rangle} dm(y) \\
&= (2\pi)^{n/2} \hat{f}(\omega) \cdot \hat{g}(\omega), \qquad \forall \omega \in \mathbb{R}^n.
\end{aligned}$$

□

Si nous inversons la transformée de Fourier sur l'image $F_1(L^1(\mathbb{R}^n)) \subset \mathcal{C}_\infty(\mathbb{R}^n)$, alors nous trouvons $f * g = (2\pi)^{n/2}(\hat{f} \cdot \hat{g})^\vee$, que l'on trouve souvent écrite sous la forme suivante

$$(f \cdot g)^\vee = (2\pi)^{-n/2} \check{f} * \check{g}. \tag{6.28}$$

Nous allons utiliser cette formule dans la section 6.11.

La convolution est une opération stationnaire, i.e. elle commute avec la translation, comme dans le cas discret. En fait, si on fixe $s \in \mathbb{R}$ et $g \in L^1(\mathbb{R})$, alors on peut définir l'opérateur de translation à droite R_s et l'opérateur de convolution avec g, T_g, comme cela :

$$R_s f(t) = f(t-s), \quad T_g f(t) = (f * g)(t) = \int_{\mathbb{R}} f(t-x)g(x)dx,$$

alors, pour tout $t \in \mathbb{R}$:

$$R_s T_g f(t) = T_g f(t-s) = \int_{\mathbb{R}} f(t-s-x)g(x)dx = \int_{\mathbb{R}} R_s f(t-x)g(x)dx = T_g R_s f(t).$$

Nous avons déjà vu dans le cas discret (cfr. section 2.9.6) que l'opération de convolution avec la fonction Gaussienne produit le floutage d'un signal. Nous avons la possibilité de comprendre ce résultat sous un autre point de vu. Pour cela, on commence par considérer la « fonction impulsion » suivante :

$$I_\varepsilon(t) = \begin{cases} \frac{1}{\varepsilon} & 0 < t < \varepsilon \\ 0 & \text{autrement} \end{cases}$$

si $f \in L^1(\mathbb{R})$, alors

$$(f * I_\varepsilon)(t) = \int_{\mathbb{R}} f(t-x)I_\varepsilon(x)dx = \frac{1}{\varepsilon}\int_0^\varepsilon f(t-x)dx,$$

si maintenant on opère le changement de variable $u = t - x$, on obtient $du = -dx$ et les extrêmes inférieur et supérieur de l'intégrale par rapport à la nouvelle variable u deviennent t et $t - \varepsilon$, alors

$$(f * I_\varepsilon)(t) = -\frac{1}{\varepsilon}\int_t^{t-\varepsilon} f(u)du = \frac{1}{\varepsilon}\int_{t-\varepsilon}^t f(u)du = \langle f \rangle_{[t,t+\varepsilon]},$$

i.e. la moyenne de f dans l'intervalle $[t-\varepsilon, t]$, qui a taille ε.

Une Gaussienne $G_{\mu,\sigma}$ de moyenne μ et écart-type σ, est une version « lisse » de l'impulsion I_ε, qui tend vers 0 rapidement hors de l'intervalle $[\mu - \sigma, \mu + \sigma]$, donc :

$$\boxed{f * G_{\mu,\sigma} \simeq \text{moyenne locale de } f \text{ en } [\mu - \sigma, \mu + \sigma]}.$$

Dans la section 2.9.6 on avait vu que l'explication du floutage dans le domaine fréquentiel était le fait que le multiplicateur de Fourier correspondant à la convolution avec la Gaussienne était un filtre passe bas. Ici on a l'explication de l'effet de floutage dans le domaine originale du signal f : en fait, après la convolution avec une Gaussienne, chaque valeur de f en t est replacée par une approximation de la moyenne locale des valeurs de f, le paramètre qui gère la localité est l'écart-type de la Gaussienne.

Une autre propriété de la convolution, cruciale dans les applications à la théorie des équations différentielles, est la possibilité de « faire tomber » la dérivée d'une convolution sur un des deux facteurs selon notre nécessité, comme montré par le résultat suivant.

Théorème 6.9.8 *Soient $f \in \mathcal{C}(\mathbb{R}^n)$ avec dérivées partielles bornés et $g \in L^1(\mathbb{R}^n)$, alors $f * g \in \mathcal{C}(\mathbb{R}^n)$ et :*

$$\partial_{x_k}(f * g) = (\partial_{x_k} f) * g, \qquad \forall k = 1, \ldots, n.$$

De la même manière, si $g \in \mathcal{C}(\mathbb{R}^n)$ avec dérivées partielles bornées et $f \in L^1(\mathbb{R}^n)$, alors

$$\partial_{x_k}(f * g) = f * (\partial_{x_k} g), \qquad \forall k = 1, \ldots, n.$$

Preuve. Les hypothèses du théorème garantissent la possibilité de passer la dérivation sous le signe d'intégrale, donc $\forall k = 1, \ldots, n$:

$$\partial_{x_k}\left(\int_{\mathbb{R}^n} f(x-y)g(y)dm(y)\right) = \int_{\mathbb{R}^n} \partial_{x_k}\left(f(x-y)g(y)\right)dm(y) = \int_{\mathbb{R}^n} \left(\partial_{x_k} f(x-y)\right)g(y)dm(y),$$

vu que seulement f dépend de x, i.e. $\partial_{x_k}(f * g) = (\partial_{x_k} f) * g$. La deuxième formule est une conséquence de la propriété commutative de la convolution, qui permet d'échanger les rôles de f et de g. □

6.9.6 Convolution et transformée de Fourier en L^2 : le problème de la localisation de la transformée de Fourier

Une généralisation de l'équation (6.27) permet de souligner une limite importante de la transformée de Fourier. Pour formaliser cette affirmation nous avons besoin de citer le résultat suivant [10], qui dit que la transformée de Fourier du produit de des éléments de $L^2(\mathbb{R}^n)$ est proportionnel à la convolution de leurs transformées de Fourier.

Théorème 6.9.9 *Si $f, g \in L^2(\mathbb{R}^n)$, alors*

$$\boxed{\widehat{f \cdot g} = (2\pi)^{-n/2}\hat{f} * \hat{g}}\ .$$

Imaginons d'être intéressés par le spectre de $f \in L^2(\mathbb{R})$ seulement dans le voisinage d'une valeur de t_0. Avec une translation on peut toujours supposer que t_0 soit $t = 0$. L'idée la plus simple, mais (comme on le verra de suite) *incorrecte*, pour localiser l'analyse du spectre de $f(t)$ consiste à le « tronquer », i.e. à le multiplier par la fonction créneau de taille $2T$

$$\chi(t) = \begin{cases} 1 & \text{si } |t| \leqslant T \\ 0 & \text{autrement,} \end{cases}$$

où $2T$ est la taille du voisinage d'intérêt. Vu que $\chi \in L^2(\mathbb{R})$, grâce [11] au théorème 6.9.9, la transformée de Fourier du **signal tronqué** $\tilde{f}(t) = f(t)\chi(t)$ est $\hat{\tilde{f}}(\omega) = 1/\sqrt{2\pi}\hat{f}(\omega) * \hat{\chi}(\omega)$, où

$$\hat{\chi}(\omega) = \frac{1}{\sqrt{2\pi}}\sqrt{\frac{2}{\pi}}T\frac{\sin(\omega T)}{\omega T} = \frac{T}{\pi}\mathrm{sinc}(\omega T),$$

où la fonction sinc, $\mathbb{R} \ni t \mapsto \mathrm{sinc}(t) = \frac{\sin t}{t}$, est dite **sinus cardinal**. Donc,

$$\hat{\tilde{f}}(\omega) = \frac{T}{\pi}\left(\hat{f}(\omega) * \mathrm{sinc}(\omega T)\right),$$

i.e. **le spectre du signal tronqué est proportionnel à la convolution entre le spectre du signal initial et le sinus cardinal de** ωT.

Cela montre que nous ne pouvons obtenir des informations localisées précises sur le signal original tout simplement en le tronquant. Ceci est une des difficultés liées à la localisation de l'analyse en fréquence d'un signal dans l'analyse de Fourier. La théorie des ondelettes [9], développée surtout à la fin des année 1980, donne des outils très puissants pour contourner ce problème.

11. Cet argument n'est pas valide si $f \in L^1(\mathbb{R})$ car la formule du théorème 6.9.9, dans ce cas, serait valide seulement si $\hat{f}$ et $\hat{\chi}$ appartiennent à $L^1(\mathbb{R})$, mais, comme on l'a vu dans la section 6.9.2, $\hat{\chi} \notin L^1(\mathbb{R})$.

6.10 Le théorème d'échantillonnage de Shannon, Nyquist et Whittaker

Le théorème de Shannon, Nyquist et Whittaker, l'un des plus importants de la théorie des signaux, dit que, lorsqu'une fonction f a un spectre borné, comme précisé dans la définition suivante, nous pouvons la reconstruire à partir d'un ensemble discret d'échantillons.

Déf. 6.10.1 *On dit que la fonction $f : \mathbb{R} \to \mathbb{C}$ est un **signal continue à bande limitée** s'il existe $\Omega \in \mathbb{R}^+$ tel que :*

$$\hat{f}(\omega) = 0 \qquad \forall |\omega| > \Omega.$$

Le concept de signal à bande limitée est de grande importance dans les applications, en fait, même si f n'est pas à bande limitée, les senseurs humains le sont toujours !

Par exemple, le système visuel humain n'est pas capable de percevoir les fréquences d'une onde électromagnétique comme lumière quand la fréquence d'oscillation est inférieure à 400 THz et supérieure à 800 THz, T=« Tera=10^{12} ». Le système auditif humain peut percevoir les sons comme ondes de variation de la pression de l'aire, quand elles ont une fréquence entre 20Hz et 20KHz, K=« Kilo=10^3 ».

Les signaux visuels et auditifs passés au cerveau pour être interprétés sont, donc, deux exemples extrêmement importantes de signaux à bande limitée.

Théorème 6.10.1 (Théorème d'échantillonnage de Shannon-Nyquist-Whittaker [12]) *Soient :*

- *$f : \mathbb{R} \to \mathbb{C}$ un signal à bande limitée : $\exists \Omega \in \mathbb{R}^+$ tel que $\hat{f}(\omega) = 0 \quad \forall |\omega| > \Omega$.*
- *$\hat{f}$: continue et $\mathcal{C}^1(\mathbb{R})$ par morceaux.*

Alors, f est complètement déterminée par ses échantillons dans les points $t_n = \frac{\pi}{\Omega} n$, $n \in \mathbb{Z}$, en fait ça vaut la formule suivante :

$$\boxed{f(t) = \sum_{n \in \mathbb{Z}} f\left(\frac{\pi}{\Omega} n\right) \operatorname{sinc}(\Omega t - \pi n)}, \tag{6.29}$$

et la convergence de la série est uniforme.

Il y a plusieurs preuves du théorème d'échantillonnage, notamment une qui utilise la formule de sommation de Poisson (Pithiviers, 1781 − Paris, 1840), ici on propose une preuve alternative à la méthode de Poisson que l'on peut trouver dans [4], page 118.

Preuve. On va utiliser la série et la transformée de Fourier de $\hat{f}$. Pour cela, on doit interpréter $\hat{f}$ comme une fonction 2Ω-périodique quand on écrit sa série de Fourier et comme une fonction à support limité en $[-\Omega, \Omega]$ quand on calcule sa transformée de Fourier. Donc, on souligne que pour reconstruire f on utilise une extension en copies périodiques de son spectre.

Grâce aux hypothèses, $\hat{f} \in L^2[-\Omega, \Omega]$ et donc on peut développer $\hat{f}$ en série de Fourier :

$$\hat{f}(\omega) = \sum_{k \in \mathbb{Z}} c_k e^{i \frac{2\pi\omega k}{2\Omega}} = \sum_{k \in \mathbb{Z}} c_k e^{i \frac{\pi\omega k}{\Omega}}, \tag{6.30}$$

avec :

$$\begin{aligned} c_k &= \frac{1}{2\Omega}\int_{-\Omega}^{\Omega} \hat{f}(\omega) e^{-i\frac{\pi\omega k}{\Omega}} d\omega \\ &\underset{\hat{f}(\omega)=0\ \forall|\omega|>\Omega}{=} \frac{\sqrt{2\pi}}{2\Omega}\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \hat{f}(\omega) e^{i(\frac{-\pi}{\Omega}k)\omega} d\omega \\ &= \frac{\sqrt{2\pi}}{2\Omega}\check{\hat{f}}\left(-\frac{\pi}{\Omega}k\right) = \frac{\sqrt{2\pi}}{2\Omega} f\left(-\frac{\pi}{\Omega}k\right), \end{aligned}$$

où dans la dernière étape on a utilisé la définition d'anti-transformée de Fourier de $\hat{f}$, i.e. f, calculée en $-\frac{\pi}{\Omega}k$ et on a inclut en c_k le facteur de normalisation de la série.

Donc on peut réécrire la série de Fourier (6.30) comme ceci :

$$\hat{f}(\omega) = \sum_{k\in\mathbb{Z}} \frac{\sqrt{2\pi}}{2\Omega} f\left(-\frac{\pi}{\Omega}k\right) e^{i\frac{\pi\omega k}{\Omega}} \underset{(n=-k\Leftrightarrow k=-n)}{=} \sum_{n\in\mathbb{Z}} \frac{\sqrt{2\pi}}{2\Omega} f\left(\frac{\pi}{\Omega}n\right) e^{-i\frac{\pi\omega n}{\Omega}},$$

et cette série *converge uniformément* car $\hat{f}$ est continue et $\mathcal{C}^1$ par morceaux.

On calcule $f(t)$ via transformée de Fourier inverse de $\hat{f}(\omega)$:

$$\begin{aligned} f(t) &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \hat{f}(\omega) e^{i\omega t} d\omega \\ &\underset{\hat{f}(\omega)=0\ \forall|\omega|>\Omega}{=} \frac{1}{\sqrt{2\pi}}\int_{-\Omega}^{\Omega} \hat{f}(\omega) e^{i\omega t} d\omega \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\Omega}^{\Omega} \sum_{n\in\mathbb{Z}} \frac{\sqrt{2\pi}}{2\Omega} f\left(\frac{\pi}{\Omega}n\right) e^{-i\frac{\pi\omega n}{\Omega}} e^{i\omega t} d\omega \\ &= \sum_{n\in\mathbb{Z}} \frac{1}{2\Omega} f\left(\frac{\pi}{\Omega}n\right) \int_{-\Omega}^{\Omega} e^{i\omega\frac{t\Omega-\pi n}{\Omega}} d\omega, \end{aligned} \tag{6.31}$$

où, dans la dernière étape, on a pu échanger la série et l'intégrale grâce au fait que la série converge uniformément. On s'intéresse maintenant à l'intégrale :

$$\int_{-\Omega}^{\Omega} e^{i\omega\frac{t\Omega-\pi n}{\Omega}} d\omega = \int_{-\Omega}^{\Omega} \cos\left(\omega\frac{t\Omega-\pi n}{\Omega}\right) d\omega + i\int_{-\Omega}^{\Omega} \sin\left(\omega\frac{t\Omega-\pi n}{\Omega}\right) d\omega,$$

la deuxième intégrale est nulle car la fonction sinus est impaire et le domaine d'intégration est symétrique, par contre la fonction cosinus est paire et donc on obtient :

$$\begin{aligned} \int_{-\Omega}^{\Omega} e^{i\omega(\frac{t\Omega-\pi n}{\Omega})} d\omega &= 2\int_{0}^{\Omega} \cos\left(\frac{t\Omega-\pi n}{\Omega}\right) d\omega = 2\left[\frac{\sin\left(\omega\frac{t\Omega-\pi n}{\Omega}\right)}{\frac{t\Omega-\pi n}{\Omega}}\right]_0^{\Omega} \\ &= 2\Omega\frac{\sin\left(\Omega\frac{t\Omega-\pi n}{\Omega}\right)}{t\Omega-\pi n} - 0 = 2\Omega\frac{\sin\left(t\Omega-\pi n\right)}{t\Omega-\pi n}. \end{aligned}$$

En introduisant ce résultat dans l'expression de $f(t)$ en (6.31), on obtient :

$$f(t) = \sum_{n\in\mathbb{Z}} \frac{2\Omega}{2\Omega} f\left(\frac{\pi}{\Omega}n\right) \frac{\sin\left(t\Omega-\pi n\right)}{t\Omega-\pi n} = \sum_{n\in\mathbb{Z}} f\left(\frac{\pi}{\Omega}n\right) \mathrm{sinc}\left(t\Omega-\pi n\right),$$

et, comme déjà souligné, la convergence de la série est uniforme. □

6.10.1 La fréquence de Nyquist : « aliasing » et « oversampling »

Vu que la fonction sinus cardinal est toujours la même pour chaque signal f, ce qui caractérise univoquement f est la suite d'échantillons $f\left(\frac{\pi}{\Omega}n\right)$.

On observe que la période d'échantillonnage des valeurs de f qui apparait dans la formule du théorème est : $T = \frac{\pi}{\Omega}$, donc la fréquence d'échantillonnage de la formule du théorème, dite **fréquence de Nyquist** est écrite comme ν_N, est $\nu_N = \frac{1}{T} = \frac{\Omega}{\pi}$.

On veut maintenant comparer la fréquence de Nyquist avec la fréquence maximale du signal f : on rappelle qu'on est parti de l'hypothèse que f soit un signal à bande limitée avec pulsation maximale Ω, alors la fréquence maximale $\nu_{\max}$ de f est définie par la relation $\Omega = 2\pi\nu_{\max}$, i.e. $\nu_{\max} = \frac{\Omega}{2\pi}$.

Si on compare la fréquence d'échantillonnage de Nyquist ν_N avec la fréquence maximale $\nu_{\max}$ du signal f on obtient : $\nu_N = 2\nu_{\max}$, i.e. **le théorème d'échantillonnage vaut si et seulement si la fréquence d'échantillonnage est (minimum) deux fois la fréquence maximale du signal f !**

Si on échantillonne avec une fréquence inférieure à celle de Nyquist, alors on obtient un phénomène connu avec le nom de **aliasing**, qui correspond à des erreurs dans la reconstruction du signal. Les erreurs sont dues au fait que, comme on l'a vu dans la preuve, on doit considérer une extension périodique du spectre de f, la fréquence de Nyquist ν_N est la fréquence (minimale) correcte pour pouvoir reconstruire f et pour éviter de « déborder » dans une période adjacente du spectre ! Une fréquence d'échantillonnage inférieure introduit des informations parasites qui viennent des périodes spectres adjacentes à gauche et à droite.

Pour terminer, on observe que le terme général de la série du théorème converge vers 0 comme $\frac{1}{n}$ quand $n \to +\infty$, qui est une convergence plutôt lente. Pour augmenter la vitesse de convergence, par exemple à $\frac{1}{n^2}$ on peut augmenter la fréquence d'échantillonnage avec une technique connue comme sur-échantillonnage (**oversampling**).

6.11 Application de la transformée de Fourier à la résolution d'équations différentielles en dérivées ordinaires et partielles

Le comportement de la transformée de Fourier en relation aux dérivées permet d'obtenir des techniques très efficaces pour résoudre certains types d'équations différentielles. Pour comprendre l'idée générale, on commence avec un exemple donné par une équation différentielle ordinaire (EDO).

6.11.1 Un exemple de EDO résolue avec la transformée de Fourier

Soient $y, g : \mathbb{R} \to \mathbb{R}$, $y, g \in L^1(\mathbb{R})$, y deux fois différentiable, on considère l'EDO suivante :

$$y''(t) - y(t) = -g(t) \qquad \forall t \in \mathbb{R}.$$

Si on applique la transformée de Fourier aux deux côtés, grâce à sa linéarité on peut écrire :

$$\widehat{y''}(\omega) - \hat{y}(\omega) = -\hat{g}(\omega),$$

i.e.

$$-\omega^2\hat{y}(\omega) - \hat{y}(\omega) = -\hat{g}(\omega) \quad \iff \quad (1+\omega^2)\hat{y}(\omega) = \hat{g}(\omega),$$

i.e.

$$\hat{y}(\omega) = \frac{1}{1+\omega^2} \cdot \hat{g}(\omega) \qquad \text{(Solution dans le domain fréquentiel).}$$

Grâce aux propriétés de la transformée de Fourier, on a rendu l'EDO une équation algébrique dans le domaine fréquentiel ! Si on connaît la transformée de Fourier de g, alors on a résolu l'EDO dans l'espace de Fourier.

Néanmoins, l'EDO initiale a été formulée par rapport à la variable t, donc il faut revenir à la représentation originale en appliquant la transformée de Fourier inverse aux deux côtés de la dernière équation en utilisant la propriété (6.28) :

$$(\hat{y}(\omega))^{\vee}(t) = y(t) = \left[\frac{1}{1+\omega^2} \cdot \hat{g}(\omega)\right]^{\vee}(t) = \frac{1}{\sqrt{2\pi}}\left(\left(\frac{1}{1+\omega^2}\right)^{\vee}(t) * g(t)\right). \tag{6.32}$$

Par calcul direct il est possible de vérifier que :

$$\widehat{e^{-a|t|}}(\omega) = \sqrt{\frac{2}{\pi}}\frac{a}{a^2+\omega^2},$$

donc, si on considère $a = 1$:

$$\sqrt{\frac{\pi}{2}}\widehat{e^{-|t|}}(\omega) = \frac{1}{1+\omega^2},$$

et alors :

$$y(t) = \frac{1}{\sqrt{2\pi}}\sqrt{\frac{\pi}{2}}e^{-|t|} * g(t),$$

i.e.

$$y(t) = \frac{1}{2}\int_{-\infty}^{+\infty} e^{-|t-s|}g(s)ds = \frac{1}{2}\int_{-\infty}^{+\infty} g(t-s)e^{-|s|}ds.$$

Si on sait calculer l'intégrale (ceci dépend de l'expression analytique de g), alors on peut déterminer explicitement $y(t)$, sinon on peut en tout cas approximer sa valeur.

Les étapes de la résolution de une EDO via transformée de Fourier sont donc :

1. Transformer l'EDO dans l'espace fréquentiel en appliquant la transformée de Fourier aux deux côtés de l'équation ;
2. Résoudre l'EDO algébrique dans l'espace de Fourier ;
3. Appliquer la transformée de Fourier inverse pour obtenir la solution de l'EDO dans la représentation originale ;
4. Typiquement, la solution dans l'espace de Fourier est donné par un produit, alors la solution dans la représentation originale est écrite avec une convolution.

Pour pouvoir utiliser cette technique, les coefficients des dérivées doivent être constantes et les fonctions doivent être intégrables.

6.11.2 Transformée de Fourier et EDPs

La technique de la transformée de Fourier est encore plus efficace quand on l'applique aux EDP : équations différentielles en dérivées partielles. Pour fixer les idées, on considère seulement des fonctions du type suivant : $u = u(t, x)$ ou $u = u(t, x, y, z)$, où t est la coordonné temporelle et x ou (x, y, z) sont coordonnées spatiales 1D et 3D, respectivement. On suppose toujours implicitement que $u \in L^1(\mathbb{R}^2)$ ou $u \in L^1(\mathbb{R}^4)$, respectivement, et, bien sûr, que u soit dérivable un nombre suffisant de fois pour pouvoir écrire l'EDP correspondante. Dans des cas moins réguliers, d'autres types de solutions doivent être cherchées, c'est aussi à cause de ça que plusieurs types de convergence doivent être considérées. Dans l'appendice C nous discutons un aperçu de ce sujet.

Pour simplifier la notation, on écrit :

$$\frac{\partial u}{\partial x} = u_x,\ \frac{\partial^2 u}{\partial x^2} = u_{xx},\ \frac{\partial u}{\partial t} = u_t,\ \ldots$$

Les propriétés de la transformée de Fourier par rapport aux dérivées partielles sont les suivantes :

— Si la variable d'intégration de la transformée de Fourier est x, alors

$$\widehat{u_x}(t,\omega) = i\omega\hat{u}(t,\omega),\ \ \widehat{u_{xx}}(t,\omega) = -\omega^2\hat{u}(t,\omega),$$

$$\widehat{u_t}(t,\omega) = \frac{\partial}{\partial t}\hat{u}(t,\omega),\ \ \widehat{u_{tt}}(t,\omega) = \frac{\partial^2}{\partial t^2}\hat{u}(t,\omega),$$

les deux premières formules sont attendues, pour obtenir les autres deux il faut observer que, comme $u \in L^1(\mathbb{R}^2)$, alors on peut échanger l'ordre de dérivation et d'intégration :

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\frac{\partial u(t,x)}{\partial t}e^{-i\omega x}dx = \frac{\partial}{\partial t}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}u(t,x)\,e^{-i\omega x}dx = \frac{\partial}{\partial t}\hat{u}(t,\omega),$$

et pareil pour u_{tt}.

— Si la variable d'intégration de la transformée de Fourier est t, alors

$$\widehat{u_t}(x,\omega) = i\omega\hat{u}(x,\omega),\ \ \widehat{u_{tt}}(x,\omega) = -\omega^2\hat{u}(x,\omega),$$

$$\widehat{u_x}(x,\omega) = \frac{\partial}{\partial x}\hat{u}(x,\omega),\ \ \widehat{u_{xx}}(x,\omega) = \frac{\partial^2}{\partial x^2}\hat{u}(x,\omega).$$

— Ces considération peuvent être étendues à $u(t, x, y, z)$.

6.11.3 Résolution de l'EDP de la propagation de la chaleur avec la transformée de Fourier

On considère le problème de Cauchy pour $u \in \mathcal{C}^2(\mathbb{R}^2) \cap L^1(\mathbb{R}^2)$ et $\varphi \in \mathcal{C}^2(\mathbb{R}) \cap L^1(\mathbb{R})$ défini par :

$$\begin{cases} u_t = \alpha^2 u_{xx} & \forall x \in (-\infty, +\infty), \forall t \in (0, +\infty),\ \alpha \in \mathbb{R}^+ \\ u(0,x) = \varphi(x) & \forall x \in (-\infty, +\infty), t = 0 \end{cases}$$

— $u\,(t,x)$: température d'une barre 1D à l'instant t et dans le point x ;

— $u_t(t,x)$: vitesse de changement de la température à l'instant t et dans le point x ;

— $u_{xx}(t,x)$: concavité du profile de la température à l'instant t et dans le point x (la dérivée seconde est par rapport à la variable spatiale, donc ça serait incorrect d'interpréter u_{xx} comme une accélération) ;

— $\varphi(x)$: concavité initiale du profile de la température dans le point x.

Si on écrit la dérivée seconde discrète (avec un pas Δx) par rapport à x, on peut comprendre qu'elle définit la comparaison entre la température en x à l'instant t et celle des voisins au même instant, en fait :

$$\begin{aligned} u_{xx}(t,x) &\simeq \frac{u(t,x+\Delta x)-2u(t,x)+u(t,x-\Delta x)}{(\Delta x)^2} \\ &= \frac{2}{(\Delta x)^2}\left[\underbrace{\frac{u(t,x+\Delta x)+u(t,x-\Delta x)}{2}}_{\text{température moyenne voisins}} - u(t,x)\right] \end{aligned}$$

Donc, l'équation $u_t = \alpha^2 u_{xx}$ dit que :

— si $u(t,x)$ est inférieure à la température moyenne des voisins, alors $u_{xx} > 0$ et donc $u_t(t,x)(= \alpha^2 u_{xx}) > 0$ et cela signifie que la température en x augmente dans le temps : les voisins cèdent un peu de leur chaleur à x pour arriver à l'équilibre thermique ;

— dans le cas contraire, $u_t(t,x)(= \alpha^2 u_{xx}) < 0$ et la température en x baisse dans le temps : x cède un peu de sa chaleur aux voisins pour arriver à l'équilibre thermique.

— La constante positive α^2 est une caractéristique du matériel dite « coefficient de diffusion thermique », plus α^2 est grand, plus rapidement la barre arrive à l'équilibre thermique.

L'équation de la chaleur est utilisée dans beaucoup d'autres domaines, par exemple en traitement d'images pour lisser les imperfections et en mathématiques financières dans le modèle de Black-Scholes-Merton.

Pour résoudre l'équation de la chaleur, on calcule la transformée de Fourier (en intégrant par rapport à la variable x) des deux côtés :

$$u_t(t,x) = \alpha^2 u_{xx}(t,x) \quad \overset{\widehat{}}{\longrightarrow} \quad \frac{\partial}{\partial t}\widehat{u}(t,\omega) = -\alpha^2\omega^2\widehat{u}(t,\omega).$$

La condition initiale dans l'espace de Fourier devient : $\widehat{u}(0,\omega) = \widehat{\varphi}(\omega)$. On a donc transformé la EDP dans une EDO :

$$\begin{cases} u_t(t,x) = \alpha u_{xx}(t,x) \\ u(0,x) = \varphi(x) \end{cases} \quad \overset{\widehat{}}{\longrightarrow} \begin{cases} \frac{\partial}{\partial t}\widehat{u}(t,\omega) = -\alpha^2\omega^2\widehat{u}(t,\omega) \\ \widehat{u}(0,\omega) = \widehat{\varphi}(\omega), \end{cases}$$

en fait ω est une constante par rapport à la variable t, donc l'équation $\frac{\partial}{\partial t}\widehat{u}(t,\omega) = -\alpha^2\omega^2\widehat{u}(t,\omega)$ est ordinaire.

On rappelle que la solution du problème de Cauchy :

$$\begin{cases} y' = -ky \\ y(0) = y_0 \end{cases}$$

est : $y(t) = y_0 e^{-kt}$ et donc, dans notre cas :

$$\widehat{u}\,(t,\omega) = \widehat{\varphi}(\omega) \cdot e^{-\alpha^2\omega^2 t} = \widehat{\varphi}(\omega) \cdot e^{-(\alpha^2 t)\omega^2} \quad \text{(Solution dans l'espace de Fourier).}$$

On doit maintenant appliquer la transformée de Fourier inverse pour revenir à la solution dans la représentation originale. Grâce à l'éq. (6.28) on obtient :

$$\begin{aligned} u(t,x) &= \left(\widehat{\varphi}(\omega) \cdot e^{-(\alpha^2 t)\omega^2}\right)^{\vee}(t,x) \\ &= \frac{1}{\sqrt{2\pi}}\check{\widehat{\varphi}}(x) * \left(e^{-(\alpha^2 t)\omega^2}\right)^{\vee}(t,x), \end{aligned}$$

mais $\check{\widehat{\varphi}}(x) = \varphi(x)$ et $e^{-(\alpha^2 t)\omega^2}$ est une Gaussienne par rapport à ω, donc on peut utiliser la propriété :

$$\widehat{e^{-c^2x^2}}(\omega) = \frac{1}{c\sqrt{2}} e^{-\frac{\omega^2}{4c^2}} \quad \Longleftrightarrow \quad c\sqrt{2}\widehat{e^{-c^2x^2}}(\omega) = e^{-\frac{1}{4c^2}\omega^2},$$

qui dans notre cas ça donne : $\frac{1}{4c^2} = \alpha^2 t$, or $c^2 = \frac{1}{4\alpha^2 t}$ et alors $c = \frac{1}{2\alpha\sqrt{t}}$ (physiquement, ça a du sens de considérer seulement la détermination positive de la racine, car la détermination négative change du signe à la température). On peut finalement écrire :

$$\left(e^{-(\alpha^2 t)\omega^2}\right)^{\vee}(t,x) = \frac{\sqrt{2}}{2\alpha\sqrt{t}} e^{-\frac{x^2}{4\alpha^2 t}} = \frac{1}{\alpha\sqrt{2t}} e^{-\frac{x^2}{4\alpha^2 t}}$$

et alors la solution de l'EDP de propagation de la chaleur est :

$$\boxed{u(t,x) = \frac{1}{\alpha\sqrt{4\pi t}} \int_{-\infty}^{+\infty} e^{-\frac{(x-y)^2}{4\alpha^2 t}} \varphi(y) dy}.$$

Pour certaines expressions de $\varphi(x)$ on peut intégrer exactement et obtenir une expression analytique pour $u(t,x)$, plus en général, on peut seulement approximer $u(t,x)$.

C'est intéressant d'observer que, comme l'expression standard d'une Gaussienne est :

$$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

alors $\sigma = \alpha\sqrt{2t}$, i.e. $\sigma^2 = 2\alpha^2 t$: la variance de la Gaussienne qui apparait dans la solution de l'EDP de la propagation de la chaleur n'est pas fixe, mais augmente linéairement avec le temps t.

Cela veut dire que le support de la Gaussienne devient de plus en plus large au fur et à mesure que le temps avance, ce que est parfaitement cohérente avec le fait que, quand $t \to +\infty$, la barre arrive à l'équilibre thermique, et donc la température est uniforme partout.

Ce qu'on vient de démontrer nous permet de comprendre encore plus en profondeur la technique classique de convolution avec une Gaussienne en traitement de signaux, qu'on a vu, par exemple, dans l'opération de floutage d'une image numérique.

En fait, si on identifie $\varphi(y)$ avec l'intensité originale d'un pixel quelconque y dans une image numérique, et $u(t,x)$ avec l'intensité de l'image flouté à l'instant t et dans un pixel fixé de position x, on peut interpréter la convolution d'une image avec une Gaussienne comme une étape du processus d'échange d'intensité entre x et ses voisins, où cet échange suit la même loi de la propagation de la chaleur.

De plus, comme la propagation de la chaleur est un *processus irréversible*, on ne peut pas penser de pouvoir inverser de manière simple l'effet de floutage obtenu par convolution avec une Gaussienne.

Une dernière observation relative à la dimension spatiale du problème : l'application de la technique qu'on vient de voir nécessite que x soit variable entre et $-\infty$ et $+\infty$, pour résoudre des problèmes avec x variable dans un intervalle borné il faut utiliser d'autres techniques : la transformée de Fourier sinus et cosinus et la transformée de Laplace.

6.12 Résumé du chapitre 6

- Les opérateurs linéaires entre espaces vectoriels normés sont continus dans un point quelconque si et seulement s'ils sont continus partout si et seulement s'ils sont bornés. Tous les opérateurs linéaires définis sur un espace vectoriel de dimension finie sont continus (et donc bornés), ceci n'est plus vrai, en général, l'espace de définition de l'opérateur n'a pas une dimension finie, comme le montre l'exemple classique de l'opérateur de dérivation défini sur l'espace vectoriel engendré par le s.o.n.c de Fourier sur $L^2[0, 2\pi]$.

- Pour les opérateurs linéaires bornés on peut définir une norme, avec 4 définitions équivalentes, qui rend l'ensemble $\mathcal{B}(V, W)$ des opérateurs linéaires bornés entre les espaces vectoriels normés V et W un espace vectoriel normé lui-même. En particulier, si $V = W$, alors on peut définir la composition des opérateurs définit un produit en $\mathcal{B}(V)$ par rapport auquel $\mathcal{B}(V)$ devient une algèbre associative normée avec unité (l'opérateur identité). De plus, si W est complet, alors $\mathcal{B}(V, W)$ est complet, en particulier, si $V = W = \mathcal{H}$, espace de Hilbert, alors $\mathcal{B}(\mathcal{H})$ est une algèbre de Banach avec unité, i.e. une algèbre normée associative, avec unité, complète et telle que $\|AB\| \leqslant \|A\| \, \|B\|$ $\forall A, B \in \mathcal{B}(\mathcal{H})$.

- Le noyau d'un opérateur borné est toujours un sous-espace vectoriel fermé du domaine de l'opérateur. Si le noyau est réduit au seul vecteur nul, alors l'opérateur est inversible, mais son inverse peut ne pas être borné. Il existe une caractérisation très simple, et très utile, de l'inversibilité bornée d'un opérateur $A : V \to W$, non nécessairement borné : l'existence de $\mu > 0$ telle que $\|Ax\| \geqslant \mu\|x\|$. Si V est un espace de Banach et cette condition est vérifiée, alors $\mathrm{Im}(A)$, l'espace image de A, est fermé. Dans les applications on pourra utiliser la fermeture de $\ker(A)$ (A continu) et de $\mathrm{Im}(A)$ (sous les hypothèses ci-dessus) pour caractériser un sous-espace fermé : il suffit de démontrer qu'il coïncide avec le noyau ou l'image d'un opérateur linéaire qui satisfait des hypothèses opportunes.

- Le dual d'un espace vectoriel quelconque V sur le corps $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ est l'espace vectoriel V^* des fonctionnelles linéaires définies sur l'espace vectoriel lui-même. Si l'espace est normé alors il est naturel de demander la compatibilité avec la structure topologique engendrée par la norme, i.e. que les fonctionnelles soient continues, alors $V^* = \mathcal{B}(V, \mathbb{K})$. Comme $\mathbb{K}$ est complet, V^* est toujours complet, même quand V ne l'est pas. Dans le cas d'un espace de Hilbert $\mathcal{H}$, le théorème de représentation de Riesz nous assure que $\mathcal{H}$ et $\mathcal{H}^*$ sont isomorphes grâce à la transformation qui associe à chaque $x \in \mathcal{H}$ la fonctionnelle T_x qui implémente le produit scalaire, i.e. $T_x(y) = \langle y, x \rangle$ $\forall y \in \mathcal{H}$. Ce théorème nous donne la possibilité de définir l'adjoint $A^\dagger$ de tout opérateur

$A \in \mathcal{B}(\mathcal{H})$ via la relation $\langle A^\dagger x, y\rangle = \langle x, Ay\rangle \ \forall x, y \in \mathcal{H}$. Si $A = A^\dagger$ on dit que A est auto-adjoint. Deux exemples d'opérateurs auto-adjoints sont $A^\dagger A$ et $AA^\dagger$.

- L'adjoint d'un opérateur linéaire borné est un opérateur très important dans la théorie et dans les applications. Nous pouvons avoir un aperçu de son importance grâce au théorème qui donne la caractérisation d'un opérateur de projection orthogonale sur un espace de Hilbert : $A \in \mathcal{B}(\mathcal{H})$ est un projecteur orthogonale, sur $\mathrm{Im}(A)$, si et seulement si A est auto-adjoint $A = A^\dagger$ et idempotent $A^2 = A$. Ce résultat permet, par exemple, de démontrer que les opérateurs de multiplication sur $L^2(\mathbb{R}^n)$ sont des projecteurs orthogonaux si et seulement s'ils multiplient par la fonction indicatrice d'un sous-ensemble mesurable de $\mathbb{R}^n$. Il existe une représentation géométrique très importante des projecteurs orthogonaux : $A \in \mathcal{B}(\mathcal{H})$ est un projecteur orthogonale si et seulement s'il existe un s.o.n $(u_n)_{n\in\mathbb{N}}$ dans $\mathcal{H}$ tel que $Ax = \sum\limits_{n\in\mathbb{N}} \langle x, u_n\rangle u_n, \ \forall x \in \mathcal{H}$. On voit que cette explicitation du projecteur est l'extension, en dimension finie, de la réalisation des projecteurs orthogonaux en dimension finie.

- L'adjoint intervient aussi dans l'analyse des opérateurs isométriques et unitaires. Un opérateur $A \in \mathcal{B}(\mathcal{H})$ est isométrique s'il préserve la norme (ou, d'une manière équivalente, le produit scalaire), un opérateur unitaire est un opérateur isométrique et surjectif. Les deux catégories d'opérateurs ont une norme unitaire. Le lien entre opérateurs isométriques et projecteurs orthogonaux est donné par le résultat suivant : si $A \in \mathcal{B}(\mathcal{H})$ est isométrique, alors $AA^\dagger$ est un projecteur orthogonal. Si $A \in \mathcal{B}(\mathcal{H})$ est isométrique, alors $\mathrm{Im}(A) = \mathrm{Im}(AA^\dagger)$ et, vu que $AA^\dagger$ est un projecteur orthogonal (car A est supposé être isométrique), $\mathrm{Im}(AA^\dagger)$ est fermée, par conséquent, l'espace image d'un opérateur isométrique est toujours fermé. Vu que $\ker(A^\dagger) = \mathrm{Im}\ (A)^\perp$, si A est isométrique mais non surjectif, alors $\mathrm{Im}(A) \neq \mathcal{H}$, donc $\mathrm{Im}\ (A)^\perp \neq \{0_\mathcal{H}\}$ et alors $A^\dagger$ n'est pas inversible. Par contre, avec le même argument, on voit que si A est unitaire, alors $A^\dagger$ est inversible.

- Comme pour les opérateurs de projection orthogonale, on peut donner une caractérisation algébrique des opérateurs isométriques et unitaires via l'adjoint : $A \in \mathcal{B}(\mathcal{H})$ est isométrique si et seulement si $A^\dagger A = id_\mathcal{H}$, tandis que $A \in \mathcal{B}(\mathcal{H})$ est unitaire si et seulement si $A^\dagger A = AA^\dagger = id_\mathcal{H}$, dans ce dernier cas A est inversible et $A^{-1} = A^\dagger$, de plus, on peut démontrer que A est unitaire si et seulement si $A^\dagger$ l'est. Ce résultat a comme conséquence qu'on peut étudier l'unitarité d'un opérateur A via celle de son adjoint, qui, parfois, peut être une tache plus simple. Venons-en à la réalisation géométrique des opérateurs isométriques et unitaires : $A \in \mathcal{B}(\mathcal{H})$ est isométrique si et seulement s'il transforme des bases hilbertiennes en s.o.n, tandis que $A \in \mathcal{B}(\mathcal{H})$ est unitaire si et seulement s'il transforme des bases hilbertiennes en bases hilbertiennes. Ceci montre une différence importante par rapport à la dimension finie.

- La transformée de Fourier, $f(x) \mapsto \hat{f}(\omega) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(x) e^{-i\langle\omega,x\rangle}\, dx$ est un outil universellement utilisé en mathématiques pures et appliquées. L'espace le plus naturel pour définir cette transformation est celui de Schwarz, dans lequel la transformée de Fourier a la formule intégrale écrite ci-dessus, et elle est un isomorphisme isométrique par rapport à la norme héritée par $L^2(\mathbb{R}^n)$. Si on veut étendre la transformée à un espace de fonctions moins régulières, par exemple à $L^1(\mathbb{R}^n)$ ou $L^2(\mathbb{R}^n)$, on doit renoncer

à quelques propriétés. Sur $L^1(\mathbb{R}^n)$ on renonce à l'inversibilité sur l'espace entier, car l'image est $\mathcal{C}_\infty(\mathbb{R}^n)$, mais on garde la formule intégrale, sur $L^2(\mathbb{R}^n)$ on perd la formule intégrale, remplacée par une formule limite, mais on garde le caractère isomorphe de la transformation, en fait l'extension de la transformée de Fourier sur $L^2(\mathbb{R}^n)$ définit un opérateur unitaire $F \in \mathcal{B}(L^2(\mathbb{R}^n))$. Une formule explicite pour cet opérateur unitaire est permise grâce à la base hilbertienne de Hermite : $Ff = \sum_{n\in\mathbb{N}} (-i)^n \langle f, u_n \rangle u_n$. Pour terminer, on souligne que, à une constante près, la transformée de Fourier de la convolution de deux fonctions de $L^1(\mathbb{R}^n)$ est le produit ponctuel des transformées.

- Nous avons terminé avec le théorème d'échantillonnage de Shannon-Nyquist-Whittaker, qui permet de reconstruire un signal à bande fréquentielle bornée à partir d'un ensemble suffisamment dense, mais fini, d'échantillons du signal et en présentant des applications de la transformée de Fourier à la résolution d'équations différentielles, notamment, à celle de la chaleur, en étudiant celle sur laquelle Fourier a développé sa théorie.

Appendices

Annexe A

L'espace vectoriel quotient

Le concept d'espace vectoriel quotient est essentiel en mathématique et, selon l'opinion de l'auteur, il est parfois traité d'une manière trop superficielle dans les livres d'algèbre linéaire. C'est pour ça que l'on va dédier une annexe à la définition et à l'interprétation de l'espace vectoriel quotient.

Déf. A.0.1 *Une relation d'équivalence $\sim$ définie sur un espace vectoriel V (de dimension quelconque) sur le corps $\mathbb{K}$ est dite* ***compatible*** *avec la structure linéaire de V si :*

$$v \sim v',\ w \sim w' \quad \Longrightarrow \quad \alpha v + \beta w \sim \alpha v' + \beta w', \qquad \forall \alpha, \beta \in \mathbb{K}.$$

Le cas échéant, la classe d'équivalence de 0, le vecteur nul de V, est un espace vectoriel Z (vu qu'elle est stable par rapport aux combinaisons linéaires et contient 0) qui est dit ***noyau*** *de la relation d'équivalence.*

Comme cas particulier de la définition nous pouvons considérer $w = w' = v'$ et $\alpha = 1$, $\beta = -1$, ce qui implique :

$$v \sim v' \quad \Longrightarrow \quad v - v' \sim 0 \iff v - v' \in Z.$$

Vice-versa, si $v - v' \in Z$, i.e. $v - v' \sim 0 \iff v - v' \sim v' - v'$ et du fait que $v' \sim v'$, on obtient, grâce à la compatibilité de $\sim$ avec la structure linéaire de V : $v - v' + v' \sim v' - v' + v'$, i.e. $v \sim v'$.

En résumé : $v \sim v' \iff v - v' \in Z$, ce qui nous dit qu'une relation d'équivalence compatible avec la structure linéaire d'un espace vectoriel est univoquement déterminée par son noyau, qui est un sous-espace vectoriel de V. Cette observation nous permet de reverser le processus : donné un sous-espace vectoriel quelconque W de V, si l'on définit

$$\boxed{v \sim_W v' \iff v - v' \in W} \qquad \forall v, v' \in V,$$

alors $\sim$ **est une relation d'équivalence en V compatible avec sa structure linéaire et avec noyau W !**

Par symétrie, $v \sim_W v' \iff v' \sim_W v \iff v' - v \in W$ et ça veut dire qu'il existe $w \in W$ tel que $v' - v = w$, i.e. $v' = v + w$, donc, la classe d'équivalence $[v]_W$ à laquelle $v \in V$ appartient est le sous-ensemble de V donné par

$$\boxed{[v]_W = v + W = \{v + w \ : \ w \in W\}},$$

que l'on appelle une **sous-variété linéaire** et que l'on interprète géométriquement comme la translation du sous-espace W à travers du vecteur v.

Nous observons que si $v \in W$, alors la translation de W via v, par linéarité, ne modifie pas W. W, en tant que sous-espace vectoriel de V, contient le 0, donc, si $v \notin W$ alors la classe d'équivalence $v + W$ ne contient pas le 0 et donc elle ne peut pas être un sous-espace vectoriel de V.

Le lemme suivant est indispensable pour la définition d'espace vectoriel quotient et il sera utilisé souvent dans la suite.

Lemme A.0.1 *Soit V un espace vectoriel quelconque, $v_1, v_2 \in V$ et W un sous-espace vectoriel de V. Alors, l'égalité $v_1 + W = v_2 + W$ vaut si et seulement si $W_1 = W_2 \equiv W$ et $v_1 - v_2 \in W$.*

Preuve.

$\boxed{\Leftarrow}$: soit $v_1 - v_2 \equiv w_0 \in W = W_1 = W_2$, alors $v_2 = v_1 - w_0$ et :

$$v_1 + W = \{v_1 + w \ : \ w \in W\}, \quad v_2 + W = \{v_1 + w - w_0 \ : \ w \in W\},$$

mais, évidemment, $W = \{w \ : \ w \in W\} = \{w - w_0 \ : \ w \in W\}$, donc $v_1 + W = v_2 + W$.

$\boxed{\Rightarrow}$: vice-versa, soit $v_1 + W_1 = v_2 + W_2$, alors, par définition de sous-variété linéaire $v_1 - v_2 + W_1 = v_2 - v_2 + W_2 = W_2$, donc, si $w_0 = v_1 - v_2$, on obtient $w_0 + W_1 = W_2$. Comme W_2 est un sous-espace vectoriel de V, il contient 0, mais alors aussi $w_0 + W_1$ contient 0, i.e. $-w_0 \in W_1$ et alors $w_0 \in W_1$ car aussi W_1 est un sous-espace vectoriel. Mais alors, opérer la translation des vecteurs de W_1 avec w_0, qui est un vecteur de W_1, ne change par le sous-espace, i.e. $w_0 + W_1 = W_1$, mais comme $w_0 + W_1 = W_2$, on arrive à $W_1 = W_2 = W$ et $w_0 = v_1 - v_2 \in W_1 = W$. Ce qui termine la preuve de l'implication directe, et donc du lemme. □

L'interprétation du lemme est la suivante : **chaque sous-variété linéaire est univoquement déterminée par un seul sous-espace W, dont la sous-variété est la translation**. De plus, aussi **le vecteur qui opère la translation est univoquement déterminé**, à moins de la somme avec un vecteur de W.

Nous pouvons finalement donner la définition d'espace vectoriel quotient et prouver que la définition est bien posée.

Déf. A.0.2 (Espace vectoriel quotient (ou facteur)) *Soit V un espace vectoriel quelconque et W un sous-espace vectoriel de V. L'**espace vectoriel quotient** V/W est l'ensemble de toutes les sous-variétés linéaires de V qui sont des translations de W, muni des opérations linéaires suivantes :*

$$(v_1 + W) + (v_2 + W) = (v_1 + v_2) + W, \qquad \forall v_1, v_2 \in V$$

$$\alpha(v + W) = \alpha v + W, \qquad \forall v \in V, \forall \alpha \in \mathbb{K}.$$

Allons vérifier que ces opérations sont bien définies et elles font de V/W un espace vectoriel sur $\mathbb{K}$.

Le lecteur pourra vérifier que les axiomes d'espace vectoriel descendent immédiatement des propriétés d'espace vectoriel vérifiées par V. Par contre, les deux propriétés suivantes ne sont pas totalement triviales :

a) si $v_1 + W = v_1' + W$ et $v_2 + W = v_2' + W$, alors $v_1 + v_2 + W = v_1' + v_2' + W$;

b) si $v_1 + W = v_1' + W$, alors $\alpha v_1 + W = \alpha v_1' + W$,

$\forall v_1, v_2, v_1', v_2' \in V$ et $\forall \alpha \in \mathbb{K}$.

Commençons par prouver la validité de **a)** : le lemme A.0.1 nous garantit que $v_1 - v_1' \equiv w_1 \in W$ et que $v_2 - v_2' \equiv w_2 \in W$, donc

$$(v_1 + v_2) + W = (v_1' + v_2') + \underbrace{(w_1 + w_2) + W}_{=W} = (v_1' + v_2') + W.$$

Pour démontrer la validité de **b)** il faut observer tout simplement que, si nous écrivons $v_1 - v_1' \equiv w \in W$, alors $\alpha(v_1 - v_1') = \alpha w \in W$, et alors le lemme A.0.1 nous garantit que $\alpha v_1 + W = \alpha v_1' + W$.

Une question naturelle que nous pouvons nous poser maintenant est : quelle est la dimension de V/W ? Est-elle associée à celle de V et de W ? Pour répondre à cette question, nous allons analyser des résultats qui nous feront mieux comprendre la signification géométrique de l'espace vectoriel quotient.

Commençons avec un résultat préliminaire.

Lemme A.0.2 *Soit V un espace vectoriel quelconque, W un sous-espace vectoriel de V et H un sous-espace de V supplémentaire à W, i.e. tel que $W \cap H = \{0\}$ et $V = W \oplus H$. Alors, pour tout vecteur $v \in V$ qui opère une translation de W, il existe seulement un vecteur $h_v \in H \cap (v + W)$. Ce vecteur est celui qui intervient dans l'écriture (unique) de v dans la somme directe suivante : $v = w_v + h_v$.*

Preuve. Commençons à démontrer qu'il existe un vecteur h_v qui appartient à H et à $v + W$. Grâce à l'hypothèses $V = W \oplus H$, tout vecteur $v \in V$ peut être écrit d'une manière unique comme $v = w_v + h_v$, $w_v \in W$ et $h_v \in H$, il faut démontrer que $h_v \in v + W$.

Pour cela, considérons maintenant un vecteur $v' = w_{v'} + h_{v'} \in V$, $w_{v'} \in W$ et $h_{v'} \in H$, qui appartient à la même classe d'équivalence de v, i.e. tel que $v + W = v' + W$, alors, encore grâce au lemme A.0.1, $v' - v \in W$, i.e. $w_{v'} + h_{v'} - w_v - h_v \in W$, i.e. $w_{v'} - w_v + h_{v'} - h_v \in W$. Mais $w_{v'} - w_v \in W$ et $h_{v'} - h_v \in H$, la seule possibilité que leur somme reste en W est que $h_{v'} - h_v = 0$ (vu que $W \cap H = \{0\}$!).

Donc $h_{v'} = h_v$ et alors $v + W \ni v' = w_{v'} + h_v$, i.e. $w_{v'} + h_v \in v + W$ mais, encore grâce au lemme A.0.1, nous savons que la somme d'un vecteur qui appartient à $v + W$ avec un vecteur de W ne fait pas sortit de la classe d'équivalence $v + W$, donc $h_v \in v + W$. Comme $h_v \in H$ et $h_v \in v + W$, alors $h_v \in H \cap (v + W)$.

Vice-versa, si $h \in H \cap (v + W)$, en particulier $h \in v + W$, i.e. $h \sim_W v$, i.e. $\exists v \in V$ et $\tilde{w}_v \in W$ tel que $h = \tilde{w}_v + v$, i.e. $v = w_v + h$, où $w_v = -\tilde{w}_v$, c'est-à-dire, $h = h_v$. □

Théorème A.0.1 *Si W est un sous-espace de l'espace vectoriel V qui admet un supplémentaire H en V, alors H est isomorphe à V/W :*

$$V/W \simeq H, \qquad V = W \oplus H.$$

Preuve. L'unicité du vecteur h_v que l'on vient d'établir avec le lemme A.0.2 nous permet de construire la correspondance bijective et intrinsèquement linéaire qui associe à une sous-variété linéaire arbitraire $v + W$ de V la composante en H d'un représentant quelconque $v \in v + W$, i.e.

$$\begin{array}{rcl} V/W & \longrightarrow & H \\ v + W & \longmapsto & h_v, \text{ tel que : } v = w_v + h_v. \end{array}$$

est un isomorphisme linéaire. □

Observons que, donné un *sous-espace vectoriel fermé* W d'un espace de Hilbert $\mathcal{H}$, le théorème de la projection orthogonale 5.4.1 nous garantit qu'il existe toujours sont supplémentaire : le complément orthogonale $W^\perp$, donc, dans ce cas :

$$\boxed{\mathcal{H}/W \simeq W^\perp},$$

i.e. l'espace vectoriel quotient d'un espace de Hilbert sur un sous-espace vectoriel fermé W est isomorphe au complément orthogonale de W. Ce résultat nous permet aussi de déterminer la dimension de V/W en fonction de celle de V et de W en dimension finie. En fait, dans ce cas, $\dim(V) = \dim(W)+\dim(H)$ et $\dim(H) = \dim(V/W)$, alors $\dim(V/W) = \dim(V)-\dim(W)$.

Corollaire A.0.1 (Dimension de V/W) *Soit V un espace vectoriel de dimension finie et W un sous-espace vectoriel de V, alors :*

$$\boxed{\dim(V/W) = \dim(V) - \dim(W)}.$$

Des nombreux problèmes en mathématique pure et appliquées amènent à considérer situations où V et W ont une dimension infinie, tandis que V/W a une dimension finie. Dans ce cas, on appelle dim(V/W) la codimension de W en V et on écrit codim(V/W).

Une fois déterminée la dimension de V/W et l'isomorphisme linéaire avec H, nous obtenons d'une manière pratiquement immédiate le corollaire suivant sur les bases de V/W en dimension finie.

Corollaire A.0.2 (Bases de V/W) *Soit V un espace vectoriel de dimension finie n et W un sous-espace vectoriel de V, alors les sous-variétés linéaires $(e_i + W)_{i=1}^n \subset V/W$ sont une base de l'espace vectoriel quotient V/W si et seulement si les représentants $(e_i)_{i=1}^n \subset V$ forment une base pour un sous-espace supplémentaire de W en V.*

Nous observons que, évidemment, le zéro de V/W est la sous-variété linéaire qui contient le 0 de V, i.e. $0_V + W \equiv W$ **est le zéro de** $V + W$.

Terminons l'analyse de V/W en considérant la projection naturelle de V sur V/W :

$$\begin{array}{rccl} \pi : & V & \longrightarrow & V/W \\ & v & \longmapsto & \pi(v) = v + W. \end{array}$$

Les propriétés de π sont décrites ci-dessous :

— π *est surjective* : ceci découle du fait que chaque élément de V/W est représenté par un vecteur de V ;

— les fibres de π, i.e. les contres-images des éléments de V/W via π, sont les éléments de V/W interprétés comme sous-variétés de V, en fait :

$$\pi^{-1}([v_0]) = \{v \in V \ : \ v + W = v_0 + W\}$$

mais l'égalité entre ensembles $v + W = v_0 + W$ est vérifiée seulement pour $v = v_0 + W$, donc :

$$\pi^{-1}([v_0]) = v_0 + W,$$

où d'abord nous interprétons $[v_0]$ comme la classe d'équivalence qui correspond à l'élément de V/W identifié par $v_0 \in V$ et après comme $v_0 + W$, vu comme un sous-ensemble de V ;

— π est une application linéaire, grâce à la bonne définition de V/W ;

— le noyau de π est W : $\ker(\pi) = W$, en fait, grâce au Lemme A.0.1 $v_0 + W = W$ si et seulement si $v_0 \in W$.

Annexe B

L'opérateur transposé (ou dual) d'un opérateur linéaire

À tout opérateur linéaire $A : V \to W$, où V et W sont deux espaces vectoriels de dimension finie, on peut associer univoquement un opérateur linéaire A^t qui est dit l'**opérateur transposé** ou **dual** de A et qui est définit comme ça :

$$\begin{array}{rccl} A^t : & W^* & \longrightarrow & V^* \\ & \varphi & \longmapsto & A^t\varphi = \varphi \circ A, \end{array}$$

i.e. :

$$\begin{array}{rccl} A^t\varphi : & V & \longrightarrow & \mathbb{K} \\ & v & \longmapsto & A^t\varphi(v) = \varphi(Av), \end{array}$$

la définition est naturelle car nous utilisons seulement A, les élément fournis par les espaces vectoriels mêmes et rien d'autre !

Si on utilise la notation canonique pour exprimer l'action d'une fonctionnelle linéaire, on peut réécrire la forme de dualité $A^t\varphi(v) = \varphi(Av)$ comme

$$\langle A^t\varphi, v\rangle = \langle \varphi, Av\rangle. \tag{B.1}$$

La bonne position de la définition, i.e. la linéarité de la fonctionnelle $A^t\varphi$ est garantie du fait que, une fois fixé φ, la forme $v \mapsto \varphi(Av)$ est linéaire car composition d'applications linéaires. Nous pouvons démontrer l'unicité de la définition comme ça : soient A_1^t et A_2^t deux opérateurs transposés tels que $A_1^t\varphi(v) = \varphi(Av) = A_2^t\varphi(v)$, i.e. $(A_1^t - A_2^t)\varphi(v) = 0$. Alors, si on fixe d'une manière arbitraire $\varphi \in V^*$ et qu'on laisse v libre de varier en V, il est évident que l'équation $(A_1^t - A_2^t)\varphi(v) = 0$ nous dit que $(A_1^t - A_2^t)\varphi$ est la fonctionnelle identiquement nulle. Comme ceci vaut pour tout $\varphi \in V^*$, ce qui implique que $A_1^t - A_2^t = 0$, i.e. $A_1^t = A_2^t$.

Supposons maintenant que V et W soient deux espaces de Banach de dimension finie. Dans ce cas, la définition est encore valide si, pour chaque $A \in \mathcal{B}(V, W)$ l'opérateur transposé défini ci-dessus est continu, i.e. $A^t \in \mathcal{B}(W^*, V^*)$ et si $A^t\varphi$ est une fonctionnelle linéaire bornée sur V toutefois que φ est une fonctionnelle linéaire bornée sur W.

Allons vérifier ces propriétés :

— $A^t\varphi$ est une fonctionnelle linéaire bornée sur V $\forall \varphi \in W^*$: la linéarité est évidente par définition, ça reste à démontrer que $A^t\varphi$ est borné

$$\begin{aligned}\|A^t\varphi\| = \sup_{\|v\|=1} \|(A^t\varphi)v\| &\underset{\text{déf de } A^t}{=} \sup_{\|v\|=1} \|\varphi(Av)\| \underset{\varphi \text{ borné}}{\leqslant} \sup_{\|v\|=1} \|\varphi\|\|Av\| \\ &\underset{A \text{ borné}}{\leqslant} \sup_{\|v\|=1} \|\varphi\|\|A\|\|v\| = \|\varphi\|\|A\| < +\infty.\end{aligned}$$

— $A^t \in \mathcal{B}(W^*, V^*)$:

$$\|A^t\| = \sup_{\|\varphi\|=1} \|A^t\varphi\| \underset{\text{déf de } A^t}{=} \sup_{\|\varphi\|=1} \|\varphi \circ A\| \underset{\varphi \text{ borné}}{\leqslant} \sup_{\|\varphi\|=1} \|\varphi\|\|A\| = \|A\| \underset{A \in \mathcal{B}(V,W)}{<} +\infty.$$

Si $V = W = \mathcal{H}$, où $\mathcal{H}$ est un espaces de Hilbert, alors l'isomorphisme de Riesz $T : \mathcal{H} \to \mathcal{H}^*$, $\mathcal{H} \ni x \mapsto T(x) = T_x$, où $T_x(y) = \langle y, x\rangle$ $\forall y \in \mathcal{H}$ est associé avec l'opérateur adjoint qu'on a défini dans la section 6.4 via l'expression

$$A^\dagger = T^{-1}A^tT.$$

Annexe C

La convergence uniforme, forte et faible

La volonté de donner un sens à la convergence de suites définies par l'action d'opérateurs « sauvages », comme le sont souvent les opérateurs différentiels, amène d'une manière naturelle à l'introduction de formes différentes de définition de convergence par rapport à celles qu'on a examiné jusqu'ici. Ceci peut être fait dans les espaces de Banach, de Hilbert et dans l'algèbre de Banach $\mathcal{B}(\mathcal{H})$.

Pour avoir un exemple concret, allons examiner la situation suivante : soit $(u_n)_{n\in\mathbb{N}}$ une base hilbertienne quelconque d'un espace de Hilbert $\mathcal{H}$. Pour tout $n \in \mathbb{N}$ définissons l'opérateur linéaire

$$\begin{array}{rccl} A_n: & \mathcal{H} & \longrightarrow & \mathcal{H} \\ & x & \longmapsto & A_n x = \sum\limits_{m=1}^{n} \langle x, u_m\rangle u_m. \end{array}$$

Grâce à la caractérisation géométrique des opérateurs de projection (cfr. théorème 6.7.2), on sait que A_n est l'opérateur de projection orthogonale sur le sous-espace vectoriel de $\mathcal{H}$ engendré par $u_1, \ldots, u_n$: $S_n = \text{span}(u_1, \ldots, u_n)$.

Comme tout $x \in \mathcal{H}$ peut être écrit comme $x = \sum\limits_{n=1}^{\infty} \langle x, u_n\rangle u_n$, on est amené à penser que la suite de projecteurs $(A_n)_{n\in N}$ converge vers $id_{\mathcal{H}}$ lorsque $n \to +\infty$.

Néanmoins, comme $S_n \subset S_{n'}$ $\forall n < n'$, on sait que $A_{n'} - A_n$ est le projecteur relatif à $\text{span}(u_{n+1}, u_{n+2}, \ldots, u_{n'}) \neq \{0_{\mathcal{H}}\}$. Nous avons vu que tout projecteur orthogonale sur un sous-espace non banal a norme unitaire, i.e. $\|A_{n'} - A_n\| = 1$ $\forall n < n'$, donc la suite $(A_n)_{n\in\mathbb{N}}$ n'est pas une suite de Cauchy en $\mathcal{B}(\mathcal{H})$ par rapport à la norme opératorielle et donc elle ne peut pas être convergente !

En revanche, si on considère la suite $(A_n x)_{n\in\mathbb{N}}$ en $\mathcal{H}$, elle converge vers x pour tout $x \in \mathcal{H}$, grâce au fait que $(u_n)_{n\in\mathbb{N}}$ est une base hilbertienne. Ceci montre l'opportunité de définir une forme de convergence alternative (qu'on verra être la convergence forte dans l'algèbre de Banach $\mathcal{B}(\mathcal{H})$) pour pouvoir donner un sens rigoureux à l'idée intuitive que la suite $(A_n)_{n\in N}$ converge vers $id_{\mathcal{H}}$.

Des exemples similaires peuvent être faits dans les espaces de Banach et de Hilbert, c'est pour ça qu'on a décidé de diviser la présentation des formes alternatives de convergence dans des sections séparées selon l'espace examiné.

C.1 Convergence forte et faible dans un espace de Banach

Soit $(V, \| \ \|)$ un espace de Banach. Par définition, une suite $(x_n)_{n\in\mathbb{N}} \subset V$ converge vers $x \in V$ si $\|x_n - x\| \xrightarrow[n\to+\infty]{} 0$. On peut définir un autre type de convergence en V en utilisant les fonctionnelles linéaires continues de son espace dual V^*.

Déf. C.1.1 (La convergence faible dans un espace de Banach) *Soit V un espace de Banach. La suite $(x_n)_{n\in\mathbb{N}} \subset V$* ***converge faiblement*** *vers $x \in V$ si, pour tout $\varphi \in V^*$,*

$$\varphi(x_n) \xrightarrow[n\to+\infty]{} \varphi(x),$$

où ici la convergence est celle des suites de scalaires en $\mathbb{K}$. x est dit la ***limite faible*** *de la suite $(x_n)_{n\in\mathbb{N}}$ et nous écrivons $x_n \xrightarrow[n\to+\infty]{w} x$, où w est utilisé pour « weak », qui veut dire faible en anglais.*

On observe que, pour tout $\varphi \in V^*$ et $x \in V$, $\|\varphi(x)\| \leqslant \|\varphi\| \, \|x\|$, donc, si $\|x_n - x\| \xrightarrow[n\to+\infty]{} 0$, alors

$$\|\varphi(x_n) - \varphi(x)\| = \|\varphi(x_n - x)\| \leqslant \|\varphi\| \, \|x_n - x\| \xrightarrow[n\to+\infty]{} 0,$$

i.e. la convergence usuelle implique celle faible. Pour cette raison, la convergence usuelle dans un espace de Banach est appelée aussi **convergence forte**.

Des contre-exemples montrent que l'inverse n'est pas vrai, en général. Donc, dans un espace de Banach, la topologie définie par la convergence faible a moins d'ouverts que celle définie par la convergence forte.

C.2 Convergence forte et faible dans un espace de Hilbert

Un espace de Hilbert $\mathcal{H}$ est aussi un espace de Banach, donc la définition de convergence forte et faible qu'on a donné ci-dessus s'applique aussi à un espace de Hilbert. Néanmoins, grâce au théorème de représentation de Riesz, on sait que l'action de toute fonctionnelle linéaire continue sur $\mathcal{H}$ peut être identifiée avec un produit scalaire. C'est pour ça qu'on peut donner une définition équivalente, mais plus explicite pour les calculs, de convergence faible dans un espace de Hilbert.

Déf. C.2.1 (La convergence faible dans un espace de Hilbert) *Soit $\mathcal{H}$ un espace de Hilbert. La suite $(x_n)_{n\in\mathbb{N}} \subset \mathcal{H}$* ***converge faiblement*** *vers $x \in \mathcal{H}$ si, pour tout $y \in \mathcal{H}$,*

$$\langle y, x_n \rangle \xrightarrow[n\to+\infty]{} \langle y, x \rangle,$$

Comme pour les espaces de Banach, x est dit la ***limite faible*** *de la suite $(x_n)_{n\in\mathbb{N}}$ et on écrit $x_n \xrightarrow[n\to+\infty]{w} x$.*

On peut montrer un contre-exemple très simple pour vérifier que la convergence faible n'implique pas, en général, la convergence forte en un espace de Hilbert.

Soit $y \in \mathcal{H}$ quelconque et $x_n = u_n \ \forall n \in \mathbb{N}$, où $(u_n)_{n\in\mathbb{N}}$ est un s.o.n quelconque de $\mathcal{H}$. Grâce à l'inégalité de Bessel $\sum\limits_{n\in\mathbb{N}} |\langle y, u_n \rangle|^2 \leqslant \|y\|^2$, la série est convergente et donc son terme général tend vers 0.

Comme $\mathcal{H}$ est complet, toute série absolument convergente est convergente, donc $\sum_{n\in\mathbb{N}} \langle y, u_n\rangle^2$ est convergente et alors $\langle y, u_n\rangle^2 \xrightarrow[n\to+\infty]{} 0$, mais ça vaut si et seulement si $\langle y, u_n\rangle \xrightarrow[n\to+\infty]{} 0$ pour tout $y \in \mathcal{H}$.

Par conséquent, **tout s.o.n $(u_n)_{n\in\mathbb{N}}$ d'un espace de Hilbert converge faiblement vers 0**.

Néanmoins, on sait que la distance entre deux éléments quelconques d'un s.o.n est $\sqrt{2}$: $\|u_n - u_m\| = \sqrt{2}\ \forall n, m \in \mathbb{N}$, donc $(u_n)_{n\in\mathbb{N}}$ ne vérifie pas la condition de Cauchy et alors elle ne peut pas être fortement convergente.

C.3 Convergence uniforme, forte et faible dans l'algèbre de Banach $\mathcal{B}(\mathcal{H})$

Dans l'algèbre de Banach $(\mathcal{B}(\mathcal{H}), \|\ \|)$, où $\mathcal{H}$ est un espace de Hilbert quelconque et $\|\ \|$ est la norme opératorielle, on a la possibilité de définir trois convergences différentes pour une suite d'opérateurs $(A_n)_{n\in\mathbb{N}} \subset \mathcal{B}(\mathcal{H})$.

Déf. C.3.1 *On utilisera les lettres u, s et w pour : uniform, strong et weak, i.e. uniforme, forte et faible, en anglais. Soit $(A_n)_{n\in\mathbb{N}} \subset \mathcal{B}(\mathcal{H})$ une suite d'opérateurs linéaires bornés sur l'espace de Hilbert $\mathcal{H}$ et soit $A \in \mathcal{B}(\mathcal{H})$.*

— ***Convergence uniforme*** *:* (*la convergence usuelle, en norme opératorielle*)

$$A_n \xrightarrow[n\to+\infty]{u} A \iff \|A_n - A\| \xrightarrow[n\to+\infty]{} 0.$$

— ***Convergence forte*** *:*

$$A_n \xrightarrow[n\to+\infty]{s} A \iff A_n x \xrightarrow[n\to+\infty]{} Ax \iff \|A_n x - Ax\|_{\mathcal{H}} \xrightarrow[n\to+\infty]{} 0 \quad \forall x \in \mathcal{H}.$$

— ***Convergence faible*** *:*

$$A_n \xrightarrow[n\to+\infty]{w} A \iff \langle y, A_n x\rangle \xrightarrow[n\to+\infty]{} \langle y, Ax\rangle \quad \forall x, y \in \mathcal{H}.$$

Comme on l'a vu dans la motivation initiale de ce chapitre, pour toute base hilbertienne $(u_m)_{m\in\mathbb{N}}$, la suite

$$\begin{array}{rccc} A_n : & \mathcal{H} & \longrightarrow & \mathcal{H} \\ & x & \longmapsto & A_n x = \sum_{m=1}^{n} \langle x, u_m\rangle u_m, \end{array}$$

ne converge pas uniformément vers $id_{\mathcal{H}}$. Néanmoins, elle converge fortement vers l'opérateur identité, vu que, par la continuité de la norme, on a

$$\lim_{n\to+\infty} \|A_n x - id_{\mathcal{H}}(x)\|_{\mathcal{H}} = \| \lim_{n\to+\infty} A_n x - x\|_{\mathcal{H}} = \| \sum_{m\in\mathbb{N}} \langle x, u_m\rangle u_m - x\|_{\mathcal{H}} = 0,$$

vu que $id_{\mathcal{H}}(x)$ ne dépend pas de n et grâce à la formule de décomposition en série de Fourier généralisée.

Il est possible de démontrer que en $\mathcal{B}(\mathcal{H})$, la convergence uniforme implique celle forte, qui implique celle faible. L'exemple ci-dessus montre que la convergence forte n'implique pas celle uniforme. Avec d'autres contre-exemples, on peut montrer que la convergence faible en $\mathcal{B}(\mathcal{H})$ n'implique pas la forte.

Index

L^∞, 139
L^p, 130
V/W, 281
$\mathbb{C}^*$-algèbre, 229
$\mathbb{K}^{\mathbb{N}}$, 133
$\ell^2(\mathbb{Z}_N)$, 34
ℓ^∞, 140
ℓ^p, 133
$\mathcal{D}(\Omega) = \mathcal{C}_c^\infty(\Omega)$, 148
σ-algèbre, 98
$\mathcal{B}(V, W)$, 201
$\mathcal{B}(\mathcal{H})$, 203
$\mathcal{S}(\mathbb{R})$, 149
$\mathcal{S}(\mathbb{R}^n)$, 149
$\mathcal{D}(\mathbb{R})$, 148

Adhérence, 107
Algèbre
    de Banach, 204
    sur un corps, 202

Base
    hilbertienne, 171
    hilbertienne de Fourier de L^2, 177
    hilbertienne de Hermite, 264
    orthogonale, 17
    orthonormale, 17
    orthonormale de Fourier de $\ell^2(\mathbb{Z}_N)$, 40
Bipolaire, 162
Borélien, 99
Borne supérieure essentielle, 140

Codomaine d'un opérateur linéaire, 194
Coefficients
    de Fourier en $\ell^2(\mathbb{Z}_N)$, 42
    de Fourier généralisés, 168
Commutateur, 247
Complément orthogonale, 152
Continuité des opérations fondamentales des espaces pré-hilbertiens, 111
Contraction, 126
Convergence
    d'une suite d'opérateurs bornés, 201
    faible, 288
    forte, 288
    uniforme, 289
Convolution, 66, 265

Densité d'une partie, 109
Dimension orthogonale d'un espace de Hilbert, 175
Dual, 213

Ensemble
    mesurable, 98
    fermé, 107
    ouvert, 107
Enveloppe convexe fermé, 162
Equivalence des topologies en dimension finie, 116
Espace
    de Banach, 119
    de Hilbert, 119
    de Hilbert séparable, 166
    de Schwartz, 149
    métrique complét, 118
    mesurable, 98
    pré-Hilbertien complexe, 9
    pré-Hilbertien réel, 7
    vectoriel métrique, 107
    vectoriel normé, 11
    vectoriel quotient, 281
    vectoriel topologique, 116
Expansion en série de Fourier généralisée, 172

Exponentiel matriciel, 125
Extension d'un opérateur linéaire, 262

Famille orthogonale de vecteurs, 15
Fermeture, 107
FFT - Fast Fourier Transform, 50
Fonction
  essentiellement bornée, 139
  1-lipschitzienne, 236
  étagée (ou simple), 147
  continue entre espaces métriques, 109
  indicatrice (ou caractéristique), 100
  mesurable, 100
  test, 148
Fonctionnelle linéaire, 213
Forme, 7
  bilinéaire, 7
  bilinéaire ou sesquilinéaire bornée, 218
  coercive, 224
  définie, 8, 9
  elliptique, 224
  hermitienne, 9
  positive, 8, 9
  quadratique, 218
  sesquilinéaire, 9
  symétrique, 8
Formule
  d'analyse, 52
  de polarisation, 14
  de synthèse, 51
  du parallélogramme, 13
Fréquence de Nyquist, 57

Gram-Schmidt - Algorithme
    d'orthonormalisation, 23

Harmonique
  d'ordre supérieur, 52, 182
  fondamentale, 52, 182
Homogénéité de la norme, 11

Identité
  de Parseval, 172
  de Parseval en dimension finie, 24
  de Plancherel, 172
  de Plancherel en dimension finie, 24
Image d'un opérateur linéaire, 194
Impulsion unitaire, 53

Inégalité
  de Bessel, 168
  de Cauchy-Schwarz, 11
  de Hölder pour les intégrales, 132
  de Hölder pour les séries, 134
  de Minkowski pour les intégrales, 131
  de Minkowski pour les séries, 133
  triangulaire, 11
Intégrale à la Lebesgue d'une fonction, 102
Isométrie, 176
Isomorphisme entre espaces de Hilbert, 175

Lemme
  de Fatou, 104
  Riemann-Lebesgue, 187

Méthodes aux éléments finis, 227
Matrice de Sylvester, 48
Mesure, 99
  σ-finie, 99
  de Borel, 102
  de Borel régulière, 102
  de comptage, 133
  finie, 99
multi-indice, 148
Multiplicateur de Fourier, 59

Norme, 11
  d'une forme bilinéaire ou sesquilinéaire
    bornée, 218
  de Frobenius d'une matrice, 125
  hilbertienne, 11
  opératorielle, 199

Opérateur
  adjoint, 227
  auto-adjoint (ou hermitien), 229
  d'intégration, 194
  de dérivation, 194
  de multiplication en $\ell^2(\mathbb{Z}_N)$, 58
  de multiplication en L^2, 242
  de projection (oblique), 234
  de projection orthogonale en espaces
    de Hilbert, 235
  de rotation, 250
  de translation, 250
  de translation en $\ell^2(\mathbb{Z}_N)$, 61
  identité, 194

inverse, 209
isométrique, 249
linéaire borné, 196
linéaire continu, 196
nul, 194
transposé, 285
unitaire, 249

Polaire, 162
Presque partout (p.p), 100
Produit
d'opérateurs bornés, 202
scalaire Euclidien complexe, 10
scalaire Euclidien réel, 8
Produit de Kronecker, 86
Projection orthogonale en dimension finie, 20

Série
absolutement convergente en norme, 109
convergente en norme, 109
de Fourier réelle en L^2, 181
Signal à bande limitée, 269
Spectre
d'amplitude, 53
de phase, 53
de puissance, 53
Suite
bornée, 119
bornée dans un espace métrique, 119
convergente en norme, 108
de Cauchy, 118
extraite (ou sous-suite), 120
Support d'une fonction, 148
Symbole de Kronecker, 15
Système
orthonormé (s.o.n), 166
orthonormé complet (s.o.n.c), 171

Théorème
d'échantillonnage, 269
d'extension d'un opérateur linéaire borné, 262
de Banach-Schauder (ou de l'application ouverte), 212
de caractérisation de la complétude des espaces normés via les séries, 123
de caracterisation d'une norme hilbertienne, 114
de Carnot, 13
de complétion d'un espace métrique non complet, 121
de convergence dominée, 104
de convergence monotone, 104
de décomposition sur une base orthonormale, 24
de Fischer-Riesz, 169
de l'inverse continu, 212
de la projection orthogonale dans un espace de Hilbert, 164
de Lax-Milgram, 225
de point fixe de Banach, 126
de projection sur un convexe fermé, 155
de Pythagore généralisé, 15
de représentation de Riesz, 213
de Riemann-Lebesgue, 261
de Riesz-Fisher (complétude des espaces L^p), 134
Topologie
métrique, 108
séparée, 108
Transformée
de Fourier discrète (DFT), 43
de Fourier discrète inverse (IDFT), 44
de Fourier-Plancherel sur $L^2(\mathbb{R}^n)$, 264
Tribu, 98
de Borel, 99
engendrée, 98

Vecteur résidu, 20
Voisinage ouvert, 107

Bibliographie

[1] M.C. Abbati and R. Cirelli. *Metodi matematici per la fisica - Operatori lineari negli spazi di Hilbert*. Città studi, 1997.

[2] R. Bartle. *The elements of integration*. Wiley and sons, 1966.

[3] S.K. Berberian. *Introduction to Hilbert spaces*. Oxford University Press, 1961.

[4] A. Boggess and F.J. Narcowich. *A first course in wavelets with Fourier analysis*. John Wiley & Sons, 2015.

[5] M. Briane and G. Pagès. *Théorie de l'intégration - Cours et exercices*. Vuibert, 1998.

[6] L. Debnath and P. Mikusinski. *Introduction to Hilbert spaces with applications*. Academic press, 2005.

[7] N. Dunford and J.T. Schwartz. *Linear operators, Part 1*. Wiley Interscience, 1958.

[8] N. El Hage Hassan. *Topologie générale et espaces normés : Cours et exercices corrigés*. Dunod, 2011.

[9] M. W. Frazier. *Introduction to wavelets through linear algebra*. Springer, 2001.

[10] C. Gasquet and P. Witomski. *Fourier analysis and applications : filtering, numerical computation, wavelets*, volume 30. Springer Science & Business Media, 2013.

[11] V. Moretti. *Spectral theory and quantum mechanics*, volume 64. Springer, 2013.

[12] K. Saxe. *Beginning functional analysis*. Springer, 2000.

[13] D. Sondaz. *Bien maîtriser les mathématiques : Limites, applications continues, espaces complets*. Cépaduès, 2010.

[14] A. Vretblad. *Fourier Analysis and Its Applications*. Springer,, 2003.

[15] K. Yosida. *Functional analysis*. Springer-Verlag Berlin Heidelberg, 1995.