

Notes de Cours

Analyse hilbertienne et de Fourier

Année 2007-2008

B. Maurey

Janvier 2008

Sommaire

Chapitre 1. Transformation de Fourier	1
1.1. Transformée de Fourier des fonctions intégrables	1
1.2. Fourier et convolution	9
1.3. Fourier dans $L^2(\mathbb{R})$	12
1.4. Compléments sur la convolution et sur Fourier	18
1.5. Fourier et dérivation	20
Chapitre 2. Espaces de Hilbert	29
2.1. Produit scalaire, orthogonalité, bases hilbertiennes	29
2.2. La base hilbertienne des exponentielles complexes	41
2.3. Projection orthogonale et applications	47
2.4. Opérateur adjoint hilbertien	55
Chapitre 3. Séries de Fourier	59
3.1. Séries de Fourier dans L^2	59
3.2. Convergence ponctuelle des séries de Fourier	65
3.3. Variations, applications des séries de Fourier	75
3.4. Transformation de Fourier et séries de Fourier	77
Annexe : rappels d'intégration	83
Index alphabétique	88
Index des notations	91

1. Transformation de Fourier

1.1. Transformée de Fourier des fonctions intégrables

Si f est une fonction intégrable sur \mathbb{R} , réelle ou complexe, on pose

$$\forall y \in \mathbb{R}, \quad \widehat{f}(y) = \int_{\mathbb{R}} f(x) e^{-ixy} dx.$$

La fonction \widehat{f} est la *transformée de Fourier* de la fonction f ; on notera que $\widehat{f}(0)$ est l'intégrale de f . La fonction \widehat{f} est bornée sur \mathbb{R} , puisqu'on a pour tout y

$$|\widehat{f}(y)| = \left| \int_{\mathbb{R}} f(x) e^{-ixy} dx \right| \leq \int_{\mathbb{R}} |f(x) e^{-ixy}| dx = \int_{\mathbb{R}} |f(x)| dx;$$

cette dernière expression est la *norme de f dans $L^1(\mathbb{R})$* ,

$$\|f\|_1 := \int_{\mathbb{R}} |f(x)| dx.$$

On utilisera aussi dans ce chapitre (et dans la suite du cours) la norme de l'espace $L^p(\mathbb{R})$, qui est définie pour $1 \leq p < +\infty$ par

$$\|f\|_p := \left(\int_{\mathbb{R}} |f(x)|^p dx \right)^{1/p}.$$

À vrai dire, seuls les cas $p = 1$ et $p = 2$ seront vraiment importants pour nous, ainsi que le cas $p = +\infty$ qui correspond à l'espace $L^\infty(\mathbb{R})$ des fonctions mesurables *bornées* sur \mathbb{R} . En utilisant la norme ^(a) de L^∞ , la majoration de la transformée de Fourier s'écrit

$$(M) \quad \|\widehat{f}\|_\infty \leq \|f\|_1.$$

Il est clair que l'application $f \rightarrow \widehat{f}$ est linéaire de $L^1(\mathbb{R})$ dans $L^\infty(\mathbb{R})$, et elle est continue, de norme ≤ 1 d'après l'inégalité (M).

On sait aussi que pour toute $f \in L^1(\mathbb{R})$, la fonction \widehat{f} est continue sur \mathbb{R} et tend vers 0 à l'infini (lemme de Riemann-Lebesgue); la continuité de \widehat{f} est une conséquence facile du théorème de continuité des intégrales à paramètre de la théorie de Lebesgue. La convergence de \widehat{f} vers 0 à l'infini est un peu plus délicate à établir; on rappellera la preuve dans la proposition 1.1.1.

On peut donc préciser : la transformation de Fourier est une application linéaire de $L^1(\mathbb{R})$ dans l'espace $C_0(\mathbb{R})$ des fonctions continues sur \mathbb{R} qui tendent vers 0 à l'infini, muni de la norme uniforme. La transformation $f \rightarrow \widehat{f}$ est continue et de norme ≤ 1 de $L^1(\mathbb{R})$ dans $C_0(\mathbb{R})$.

Premiers exemples

1. Si A est un sous-ensemble d'un ensemble X , on notera $\mathbf{1}_A$ la fonction indicatrice de A , définie sur X par : $\mathbf{1}_A(x) = 1$ si $x \in A$, et $\mathbf{1}_A(x) = 0$ si $x \notin A$. Si $f = \mathbf{1}_{[a,b]}$ est l'indicatrice de l'intervalle $[a, b]$, on a

$$(1) \quad \widehat{\mathbf{1}_{[a,b]}}(y) = \int_a^b e^{-ity} dt = \frac{e^{-ia y} - e^{-ib y}}{iy} \quad \text{pour } y \neq 0, \quad \widehat{\mathbf{1}_{[a,b]}}(0) = b - a;$$

en particulier pour un intervalle symétrique $[-b, b]$, la transformée de Fourier de la fonction $f = \mathbf{1}_{[-b,b]}$ est la fonction

$$y \neq 0 \rightarrow \frac{2 \sin(by)}{y}, \quad 0 \rightarrow 2b.$$

On peut voir que cette fonction n'est pas Lebesgue-intégrable sur \mathbb{R} : la transformation de Fourier ne transforme pas en général les fonctions intégrables en fonctions intégrables.

Par combinaison linéaire de fonctions indicatrices d'intervalles on obtient les *fonctions en escalier sur \mathbb{R}* , et on peut calculer leur transformée de Fourier par linéarité, à partir du calcul des lignes précédentes ; la fonction g est en escalier sur \mathbb{R} s'il existe des intervalles bornés $[a_{j-1}, a_j]$ et des valeurs c_j , $j = 1, \dots, N$ tels que

$$g = \sum_{j=1}^N c_j \mathbf{1}_{[a_{j-1}, a_j]}.$$

Il revient au même de dire qu'il existe des nombres $b_0 < b_1 < \dots < b_m$ tels que g soit constante sur chaque intervalle ouvert $]b_{j-1}, b_j[$, $j = 1, \dots, m$ et $g(x) = 0$ quand $x \leq b_0$ et $x \geq b_m$. À partir de la première expression pour la fonction en escalier g on obtient pour $y \neq 0$

$$(2) \quad \widehat{g}(y) = \sum_{j=1}^N c_j \frac{e^{-ia_{j-1}y} - e^{-ia_j y}}{iy},$$

et il en résulte que

$$|\widehat{g}(y)| \leq \left(\sum_{j=1}^N |c_j| \right) \frac{2}{|y|},$$

donc $|\widehat{g}(y)|$ est $O(|y|^{-1})$ à l'infini, en particulier \widehat{g} tend vers 0 à l'infini. On est prêt pour rappeler la preuve de la propriété de Riemann-Lebesgue.

Proposition 1.1.1. *Pour toute fonction $f \in L^1(\mathbb{R})$, la transformée de Fourier \widehat{f} tend vers 0 à l'infini.*

Preuve. — Dans le cours d'Intégration (voir aussi le théorème A.11 de l'annexe), il est montré que toute fonction $f \in L^1(\mathbb{R})$ peut être approchée en norme L^1 par des fonctions en escalier : si $\varepsilon > 0$ est donné, on peut trouver g en escalier telle que $\|f - g\|_1 < \varepsilon/2$; pour $|y|$ assez grand, on aura $|\widehat{g}(y)| < \varepsilon/2$ d'après ce qui précède, et par la relation (M) on aura pour tout y l'inégalité $|\widehat{f}(y) - \widehat{g}(y)| \leq \|f - g\|_1 < \varepsilon/2$. Pour $|y|$ assez grand, on a donc $|\widehat{f}(y)| < \varepsilon$.

2. Disons qu'une fonction g est de classe C^1 sur un intervalle fermé $[a, b]$ si elle admet une dérivée continue sur l'ouvert $]a, b[$ qui se prolonge par continuité au fermé $[a, b]$; on sait alors par la règle de l'Hospital que les limites en a et en b de la dérivée sont les dérivées à droite en a et à gauche en b de la fonction g . On sait aussi que

$$g(b) - g(a) = \int_a^b g'(x) dx.$$

Considérons maintenant une fonction f continue sur un intervalle fermé borné $[a, b]$ et de classe C^1 par morceaux, ce qui signifie qu'il existe une subdivision $a_0 = a < a_1 < a_2 < \dots < a_N = b$ de l'intervalle $[a, b]$ telle que f soit de classe C^1 sur chaque intervalle fermé $[a_{j-1}, a_j]$, $j = 1, \dots, N$. La fonction dérivée f' n'est pas définie aux points a_0, a_1, \dots, a_N , mais elle est presque partout définie, mesurable bornée (**b**): elle définit donc un élément $f' \in L^1([a, b])$. Pour chaque $j = 1, \dots, N$ on a

$$f(a_j) - f(a_{j-1}) = \int_{a_{j-1}}^{a_j} f'(x) dx.$$

En additionnant toutes les valeurs de $j = 1$ à $j = N$, on obtient

$$f(b) - f(a) = \int_a^b f'(x) dx.$$

Énonçons ce résultat qui nous servira plus tard.

Lemme 1.1.2. *Si f est continue et de classe C^1 par morceaux sur un intervalle fermé borné $[a, b]$, on a*

$$f(b) - f(a) = \int_a^b f'(x) dx,$$

où la fonction dérivée f' est définie sur $[a, b]$ privé d'un nombre fini de points, et définit une classe de fonction mesurable bornée.

Considérons ensuite une fonction f continue sur \mathbb{R} , nulle en dehors d'un intervalle borné $[a, b]$ et de classe C^1 par morceaux sur l'intervalle $[a, b]$. La fonction dérivée f' est maintenant définie sur \mathbb{R} privé d'un nombre fini de points, et elle est nulle hors de $[a, b]$, donc $f' \in L^1(\mathbb{R})$. Comme $f(x)$ est nul pour $x \leq a$ et $x \geq b$, on obtient

$$0 = f(b) - f(a) = \int_a^b f'(x) dx = \int_{\mathbb{R}} f'(x) dx.$$

Pour calculer la transformée de Fourier de f , fixons y et considérons la fonction auxiliaire $g : x \rightarrow f(x) e^{-ixy}$; la fonction g est elle aussi continue de classe C^1 par morceaux, nulle en dehors de $[a, b]$; on peut donc lui appliquer l'égalité précédente

$$0 = \int_{\mathbb{R}} g'(x) dx = \int_{\mathbb{R}} f'(x) e^{-ixy} dx - iy \int_{\mathbb{R}} f(x) e^{-ixy} dx.$$

Il en résulte que

$$\forall y \in \mathbb{R}, \quad iy \hat{f}(y) = \hat{f}'(y).$$

3. Considérons maintenant une fonction f continue sur \mathbb{R} et affine par morceaux, nulle en dehors d'un intervalle borné; comme cette famille de fonctions interviendra plusieurs fois dans la suite, on va poser une définition.

Définition 1.1.3. On désignera par $\text{Aff}(a, b)$ le sous-espace vectoriel de l'espace des fonctions continues sur l'intervalle fermé borné $[a, b]$, formé des fonctions f pour lesquelles il existe une subdivision $a_0 = a < a_1 < \dots < a_N = b$ de $[a, b]$ telle que f soit affine sur chaque intervalle $[a_{j-1}, a_j]$, $j = 1, \dots, N$. On dira que f est *continue et affine par morceaux* sur $[a, b]$.

On désignera par $\text{Aff}_c(\mathbb{R})$ le sous-espace vectoriel de l'espace des fonctions continues sur \mathbb{R} , formé des fonctions f pour lesquelles il existe un intervalle fermé borné $[a, b]$ tel que f soit nulle en dehors de $[a, b]$ et que la restriction de f à $[a, b]$ soit continue et affine par morceaux sur $[a, b]$.

On dit qu'une fonction continue f est à *support compact* quand il existe un compact K tel que $f(x) = 0$ pour tout $x \notin K$. Ainsi, les fonctions de $\text{Aff}_c(\mathbb{R})$ sont les fonctions continues à support compact et affines par morceaux sur \mathbb{R} .

Soit $f \in \text{Aff}_c(\mathbb{R})$, affine de pente p_j sur chaque intervalle $[a_{j-1}, a_j]$, $j = 1, \dots, N$ d'une subdivision de $[a, b]$, et $f(x) = 0$ quand $x \leq a$ ou $x \geq b$. La classe de la dérivée f' définie presque partout est donc

$$f' = \sum_{j=1}^N p_j \mathbf{1}_{[a_{j-1}, a_j]}.$$

D'après l'exemple précédent on sait que $iy\widehat{f}(y) = \widehat{f}'(y)$, et \widehat{f}' se calcule à partir du résultat (2) de l'exemple 1 ; on obtient pour tout $y \neq 0$

$$\widehat{f}(y) = \frac{\widehat{f}'(y)}{iy} = \sum_{j=1}^N p_j \frac{e^{-ia_j y} - e^{-ia_{j-1} y}}{y^2}.$$

On voit en particulier que pour tout $y \neq 0$,

$$|\widehat{f}(y)| \leq \left(\sum_{j=1}^N |p_j| \right) \frac{2}{y^2},$$

ce qui montre le résultat suivant, qu'on utilisera plus loin.

Lemme 1.1.4. Si f appartient à l'espace $\text{Aff}_c(\mathbb{R})$, alors \widehat{f} est intégrable sur \mathbb{R} .

4. Pour la fonction $f : t \in \mathbb{R} \rightarrow e^{-|t|}$ on trouve

$$\begin{aligned} \widehat{f}(y) &= \int_{\mathbb{R}} e^{-|t|-ity} dt = \int_0^{+\infty} e^{-t-ity} dt + \int_0^{+\infty} e^{-s+isy} ds \\ &= \frac{1}{1+iy} + \frac{1}{1-iy} = \frac{2}{1+y^2}. \end{aligned}$$

Signaux simples périodiques, signaux plus complexes

Dans ce paragraphe, la variable t représentera le temps. Un signal simple (par exemple l'onde porteuse émise par une station de radio et reçue par une antenne) se représente en sinus ou cosinus, sous la forme $t \rightarrow \sin(\omega t)$ ou $t \rightarrow \cos(\omega t)$, ou bien avec une phase φ , qui traduit un décalage dans le temps

$$\sin(\omega t + \varphi) = \cos(\varphi) \sin(\omega t) + \sin(\varphi) \cos(\omega t);$$

l'exponentielle complexe sera plus agréable mathématiquement, car elle transforme les sommes en produits : les deux fonctions $e^{i\omega t}$ et $e^{-i\omega t}$ permettent de reconstituer par combinaisons linéaires les fonctions précédentes.

Si le signal est $t \rightarrow \sin(2\pi Ft)$, le nombre F est la *fréquence* du signal, c'est-à-dire le nombre des oscillations de la fonction sur un intervalle de longueur 1 (dans le contexte de la radio : un intervalle de 1 seconde ; une fréquence FM de 100 MHz correspond à 100 millions d'oscillations par seconde) ; on utilise aussi l'écriture $\sin(\omega t)$, et $\omega = 2\pi F$ est alors appelé la *pulsation*.

On peut envisager un signal plus complexe qui soit un mélange de signaux de fréquences différentes,

$$t \rightarrow \sum_{j=1}^N \varphi_j e^{it\xi_j};$$

plus particulièrement, il pourrait être de la forme

$$t \rightarrow \sum_{j=1}^N (\xi_j - \xi_{j-1}) \varphi(\xi_j) e^{it\xi_j}$$

où φ est une fonction continue sur $[a, b]$ et $\xi_0 = a < \xi_1 < \dots < \xi_N = b$ une subdivision de l'intervalle $[a, b]$. Les signaux pratiquement étudiés dans ce cours seront de la forme limite suivante, qui est obtenue quand le pas de la subdivision tend vers 0,

$$f(t) = \int_{\mathbb{R}} e^{it\xi} \varphi(\xi) d\xi,$$

où on fera des hypothèses qui permettront un traitement mathématique raisonnable du problème.

Décryptage du spectre des fréquences

Théorème 1.1.5. *Si φ est continue, bornée et intégrable sur \mathbb{R} , et si la fonction signal f définie par*

$$\forall t \in \mathbb{R}, \quad f(t) = \int_{\mathbb{R}} e^{it\xi} \varphi(\xi) d\xi$$

est intégrable sur \mathbb{R} , alors

$$\forall \xi \in \mathbb{R}, \quad \varphi(\xi) = \frac{1}{2\pi} \widehat{f}(\xi).$$

La transformée de Fourier permet donc d'analyser la contribution au signal f des différentes fréquences/pulsations ξ ; cette contribution est donnée, dans la formule ci-dessus qui définit f , par la fonction φ , qui est en principe connue de l'émetteur de l'émission, mais pas du récepteur qui ne voit que le signal f . La transformation de Fourier permet au récepteur de retrouver cette information φ sur la composition du signal.

On commence par montrer un cas particulier du théorème.

Lemme 1.1.6. Si ψ est continue, bornée et intégrable sur \mathbb{R} , et si la fonction g définie par

$$\forall t \in \mathbb{R}, \quad g(t) = \int_{\mathbb{R}} e^{it\xi} \psi(\xi) d\xi$$

est intégrable sur \mathbb{R} , alors

$$\widehat{g}(0) = 2\pi \psi(0).$$

Preuve. — Si on attaque tout droit, on écrit

$$\widehat{g}(0) = \int_{\mathbb{R}} g(t) dt = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{it\xi} \psi(\xi) d\xi \right) dt;$$

en admettant même qu'on puisse intervertir les intégrales, on se retrouve avec

$$\int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{it\xi} dt \right) \psi(\xi) d\xi$$

où l'intégrale intérieure n'a pas de sens évident. Il faut donc ruser, d'une façon ou d'une autre. On va remplacer $\int_{\mathbb{R}} g(t) dt$ par $\int_{\mathbb{R}} g(t)k(\varepsilon t) dt$, où k sera continue sur \mathbb{R} et $k(0) = 1$; on fera tendre $\varepsilon > 0$ vers 0, de sorte que $k(\varepsilon t) \rightarrow k(0) = 1$ nous redonne à la limite l'expression $\int_{\mathbb{R}} g(t) dt$ que nous voulons calculer.

Précisons les hypothèses : on suppose que la fonction k est continue bornée intégrable sur \mathbb{R} et que sa transformée de Fourier \widehat{k} est intégrable sur \mathbb{R} . Pour tout $\varepsilon > 0$, puisque les fonctions $t \rightarrow k(\varepsilon t)$ et ψ sont intégrables, on peut écrire d'après le théorème de Fubini que

$$\int_{\mathbb{R}} g(t)k(\varepsilon t) dt = \int_{\mathbb{R}^2} \psi(\xi)k(\varepsilon t) e^{it\xi} dt d\xi;$$

ensuite, en posant $s = \varepsilon t$ et $v = \varepsilon^{-1}\xi$, on obtient $dt d\xi = ds dv$ par changement de variable et on prolonge la chaîne d'égalités

$$\int_{\mathbb{R}^2} \psi(\xi)k(\varepsilon t) e^{it\xi} dt d\xi = \int_{\mathbb{R}^2} \psi(\varepsilon v)k(s) e^{isv} ds dv = \int_{\mathbb{R}} \psi(\varepsilon v)\widehat{k}(-v) dv.$$

Puisque k et ψ sont continues bornées, g et \widehat{k} intégrables, on obtient par deux applications du théorème de convergence dominée, quand $\varepsilon \rightarrow 0$

$$\int_{\mathbb{R}} g(t)k(\varepsilon t) dt \rightarrow k(0) \int_{\mathbb{R}} g(t) dt \quad \text{et} \quad \int_{\mathbb{R}} \psi(\varepsilon v)\widehat{k}(-v) dv \rightarrow \psi(0) \int_{\mathbb{R}} \widehat{k}(-v) dv.$$

Il en résulte que

$$k(0) \int_{\mathbb{R}} g(t) dt = \psi(0) \int_{\mathbb{R}} \widehat{k}(-v) dv.$$

Pour finir, on applique ce qui précède avec la fonction $k(t) = e^{-|t|}$ de l'exemple 4 (on n'est pas obligé, mais puisqu'on a fait les calculs, autant s'en servir). Alors $k(0) = 1$ et on a vu que $\widehat{k}(v) = 2/(1 + v^2)$, donc

$$\widehat{g}(0) = \int_{\mathbb{R}} g(t) dt = \psi(0) \int_{\mathbb{R}} \frac{2}{1 + v^2} dv = 2\pi \psi(0),$$

et le lemme est démontré.

Démonstration du théorème 1.1.5. — Pour pouvoir montrer ce théorème à partir du lemme 1.1.6, on fixe un ω quelconque et on effectue un décalage en fréquence en posant

$$\forall \xi \in \mathbb{R}, \quad \psi(\xi) = \varphi(\xi + \omega).$$

La fonction ψ est continue, bornée et intégrable d'après l'hypothèse faite sur φ dans l'énoncé du théorème. Pour appliquer le lemme 1.1.6, posons

$$g(t) = \int_{\mathbb{R}} e^{it\xi} \psi(\xi) d\xi = \int_{\mathbb{R}} e^{it\xi} \varphi(\xi + \omega) d\xi = \int_{\mathbb{R}} e^{it(v-\omega)} \varphi(v) dv = e^{-it\omega} f(t);$$

ainsi $|g(t)| = |f(t)|$, la fonction g est intégrable d'après l'hypothèse sur f , donc en appliquant le lemme 1.1.6, on sait que $\widehat{g}(0) = 2\pi\psi(0)$ et

$$\widehat{f}(\omega) = \int_{\mathbb{R}} e^{-it\omega} f(t) dt = \int_{\mathbb{R}} g(t) dt = \widehat{g}(0) = 2\pi\psi(0) = 2\pi\varphi(\omega).$$

Première formule d'inversion de Fourier

Pour ne pas nous répéter sans cesse dans la suite, on introduira une notation pour la classe des fonctions qui ont permis de démontrer le théorème 1.1.5 ; il ne s'agit pas d'une notation consacrée qu'on trouve dans les livres, simplement quelque chose «entre nous».

Définition 1.1.7. On désignera par X l'ensemble des fonctions g réelles ou complexes définies sur \mathbb{R} , qui sont continues, bornées, intégrables et telles que la transformée de Fourier \widehat{g} soit intégrable sur \mathbb{R} .

On voit facilement que X est un espace vectoriel de fonctions, et X est contenu dans $L^2(\mathbb{R})$: si la fonction f est élément de X , elle est à la fois intégrable et bornée, et il en résulte que $|f(t)|^2 \leq \|f\|_{\infty} |f(t)|$ est intégrable. D'après le lemme 1.1.4, on sait que si f est continue à support compact et affine par morceaux, sa transformée \widehat{f} est intégrable sur \mathbb{R} , et comme f est continue à support compact, elle est bornée et intégrable ; tout ceci justifie le résultat qui suit.

Lemme 1.1.8. *L'espace X contient l'espace $\text{Aff}_c(\mathbb{R})$ des fonctions continues à support compact et affines par morceaux sur \mathbb{R} .*

Comme l'espace $\text{Aff}_c(\mathbb{R})$ est dense dans $L^p(\mathbb{R})$, pour tout p tel que $1 \leq p < +\infty$ (théorème A.11), on voit que X est dense dans $L^2(\mathbb{R})$; on utilisera cette densité pour étendre la transformation de Fourier, depuis l'espace X jusqu'à l'espace $L^2(\mathbb{R})$.

Théorème 1.1.9. *Si φ est une fonction de l'espace X , on a*

$$\forall x \in \mathbb{R}, \quad \varphi(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{ixy} \widehat{\varphi}(y) dy.$$

Preuve. — Pour montrer le théorème 1.1.9, il suffit de changer de point de vue dans le théorème 1.1.5, en oubliant le sens physique des variables t et ξ , et en considérant la formule du théorème 1.1.5 d'un point de vue purement mathématique : si φ est continue, bornée et intégrable, et si on sait que $f(y) = \int_{\mathbb{R}} e^{ixy} \varphi(x) dx$ est intégrable, le théorème 1.1.5 nous dit que

$$\forall x \in \mathbb{R}, \quad \widehat{f}(x) = 2\pi\varphi(x).$$

Or on voit que

$$f(y) = \int_{\mathbb{R}} e^{ixy} \varphi(x) dx = \int_{\mathbb{R}} e^{-ix(-y)} \varphi(x) dx = \widehat{\varphi}(-y)$$

ce qui garantit que f est intégrable, puisque $\widehat{\varphi}$ est intégrable d'après l'hypothèse $\varphi \in X$; on a donc bien par application du théorème 1.1.5

$$2\pi \varphi(x) = \widehat{f}(x) = \int_{\mathbb{R}} e^{-ixy} f(y) dy = \int_{\mathbb{R}} e^{-ixy} \widehat{\varphi}(-y) dy = \int_{\mathbb{R}} e^{ixu} \widehat{\varphi}(u) du ;$$

c'est le résultat annoncé.

Première estimation en norme L^2

Proposition 1.1.10. *Si la fonction g appartient à l'espace X de la définition 1.1.7, sa transformée de Fourier \widehat{g} est dans $L^2(\mathbb{R})$ et*

$$\int_{\mathbb{R}} |\widehat{g}(y)|^2 dy = 2\pi \int_{\mathbb{R}} |g(x)|^2 dx.$$

Preuve. — On a pour toute fonction $g \in X$, en prenant le complexe conjugué de l'égalité du théorème 1.1.9

$$\forall x \in \mathbb{R}, \quad 2\pi \overline{g(x)} = \overline{\int_{\mathbb{R}} e^{ixy} \widehat{g}(y) dy} = \int_{\mathbb{R}} e^{-ixy} \overline{\widehat{g}(y)} dy ;$$

ensuite, en appliquant le théorème de Fubini

$$2\pi \int_{\mathbb{R}} g(x) \overline{g(x)} dx = \int_{\mathbb{R}^2} g(x) e^{-ixy} \overline{\widehat{g}(y)} dx dy = \int_{\mathbb{R}} \widehat{g}(y) \overline{\widehat{g}(y)} dy$$

ce qui donne le résultat annoncé.

On vient d'établir pour toute fonction $g \in X$ un cas particulier de *l'identité de Parseval*,

$$(3) \quad \|\widehat{g}\|_2 = \sqrt{2\pi} \|g\|_2.$$

Cette propriété nous servira un peu plus loin pour prolonger la transformation de Fourier à l'espace $L^2(\mathbb{R})$ (au paragraphe 1.3). On utilisera l'équation précédente et la densité de l'espace X dans l'espace $L^2(\mathbb{R})$.

1.2. Fourier et convolution

Signaux retardés

Si un signal f est transmis avec un retard $s > 0$, le signal retardé est $f_s : t \rightarrow f(t - s)$; on a

$$\widehat{f}_s(\xi) = e^{-is\xi} \widehat{f}(\xi).$$

On peut envisager un signal qui soit composé d'un mélange de signaux retardés f_s , avec une mesure de mélange $g(s) ds$; on obtient alors le signal composé

$$t \rightarrow \int_{\mathbb{R}} f(t - s)g(s) ds.$$

On voit apparaître la *convolution* des fonctions f et g , qui a été définie dans le cours d'intégration. On va rappeler les propriétés de cette opération.

Proposition 1.2.1. *Si $f \in L^1(\mathbb{R})$ et $g \in L^1(\mathbb{R})$, la fonction $s \in \mathbb{R} \rightarrow f(x - s)g(s)$ est intégrable pour presque tout $x \in \mathbb{R}$, et on peut donc poser pour presque tout $x \in \mathbb{R}$*

$$(4) \quad (f * g)(x) = \int_{\mathbb{R}} f(x - s)g(s) ds = (g * f)(x) ;$$

on obtient de cette façon une classe de fonctions $f * g \in L^1(\mathbb{R})$ et

$$\|f * g\|_1 \leq \|f\|_1 \|g\|_1, \quad \int_{\mathbb{R}} (f * g)(x) dx = \left(\int_{\mathbb{R}} f(x) dx \right) \left(\int_{\mathbb{R}} g(x) dx \right).$$

Preuve. — Supposons d'abord f et g mesurables ≥ 0 ; en admettant la valeur $+\infty$, on peut toujours écrire

$$\forall x \in \mathbb{R}, \quad (f * g)(x) = \int_{\mathbb{R}} f(x - s)g(s) ds = \int_{\mathbb{R}} f(u)g(x - u) du = (g * f)(x).$$

On a par Fubini positif

$$\begin{aligned} \int_{\mathbb{R}} (f * g)(x) dx &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x - s)g(s) ds \right) dx = \int_{\mathbb{R}^2} f(x - s)g(s) ds dx = \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x - s) dx \right) g(s) ds = \left(\int_{\mathbb{R}} f(x) dx \right) \left(\int_{\mathbb{R}} g(s) ds \right). \end{aligned}$$

Si f, g sont intégrables positives, on a donc

$$(5) \quad \int_{\mathbb{R}} (f * g)(x) dx = \left(\int_{\mathbb{R}} f(x) dx \right) \left(\int_{\mathbb{R}} g(x) dx \right) < +\infty,$$

ce qui implique en particulier que $(f * g)(x)$ est fini pour presque tout x . Si $f, g \in L^1(\mathbb{R})$, les fonctions $|f| : x \rightarrow |f(x)|$ et $|g|$ sont intégrables positives ; l'équation (5) appliquée à $|f|$ et $|g|$ montre que l'intégrale de $|f| * |g|$ est finie, donc l'ensemble N des $x \in \mathbb{R}$ tels que

$$(|f| * |g|)(x) = \int_{\mathbb{R}} |f(x - s)| |g(s)| ds = +\infty$$

est négligeable ; on voit que pour tout $x \notin \mathbb{N}$, la fonction

$$s \rightarrow f(x-s)g(s)$$

est intégrable sur \mathbb{R} , ce qui permet de poser

$$\forall x \notin \mathbb{N}, \quad (f * g)(x) = \int_{\mathbb{R}} f(x-s)g(s) \, ds.$$

Si on veut, on peut poser $(f * g)(x) = 0$ pour $x \in \mathbb{N}$, et on obtient de cette façon une fonction mesurable $f * g$ définie partout sur \mathbb{R} (l'affirmation de mesurabilité fait partie de l'énoncé du théorème de Fubini). On a ensuite

$$|(f * g)(x)| = \left| \int_{\mathbb{R}} f(x-s)g(s) \, ds \right| \leq \int_{\mathbb{R}} |f(x-s)||g(s)| \, ds = (|f| * |g|)(x)$$

et en utilisant (5) on obtient

$$\|f * g\|_1 = \int_{\mathbb{R}} |f * g| \leq \int_{\mathbb{R}} (|f| * |g|) = \left(\int_{\mathbb{R}} |f| \right) \left(\int_{\mathbb{R}} |g| \right) = \|f\|_1 \|g\|_1.$$

Pour l'intégrale de $f * g$, on reprend le calcul du début de la preuve, mais en utilisant maintenant Fubini général au lieu de Fubini positif.

Exemple. Convolution avec une fonction indicatrice d'intervalle

Si $\tau > 0$ et si on introduit la fonction $K_\tau = \tau^{-1}\mathbf{1}_{[0,\tau]}$, on voit que pour toute $f \in L^1(\mathbb{R})$ et tout t , on a

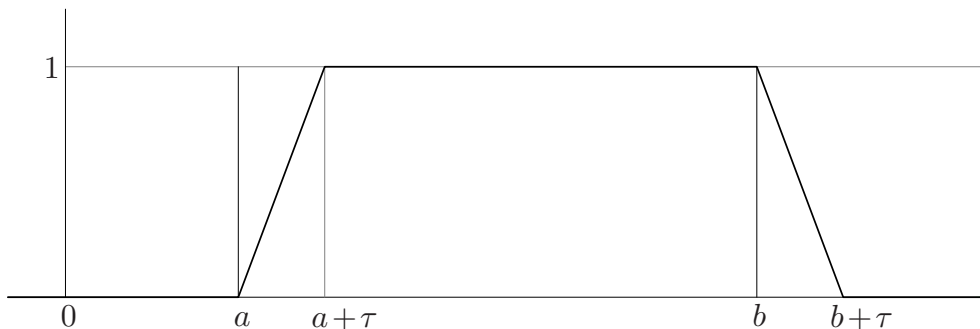
$$(f * K_\tau)(t) = \frac{1}{\tau} \int_{t-\tau}^t f(s) \, ds.$$

Il en résulte que $f * K_\tau$ est continue, dérivable si f est continue. La fonction $f * K_\tau$ est bornée par $\tau^{-1}\|f\|_1$ et elle est intégrable par la proposition précédente ; en résumé :

(6) si $f \in L^1(\mathbb{R})$, la fonction $f * K_\tau$ est continue bornée et intégrable.

Si $f = \mathbf{1}_{[a,b]}$ avec $b - a > \tau$, on voit (c) que $\mathbf{1}_{[a,b]} * K_\tau$ est nulle en dehors de $[a, b + \tau]$, égale à 1 dans $[a + \tau, b]$, avec un raccord linéaire entre les deux cas,

$$\begin{aligned} (\mathbf{1}_{[a,b]} * K_\tau)(x) &= (x - a)/\tau \quad \text{si } a \leq x \leq a + \tau, \\ (\mathbf{1}_{[a,b]} * K_\tau)(x) &= (b + \tau - x)/\tau \quad \text{si } b \leq x \leq b + \tau. \end{aligned}$$



Graphe de $\mathbf{1}_{[a,b]} * K_\tau$

En particulier, la valeur absolue de la différence entre $\mathbf{1}_{[a,b]}$ et $\mathbf{1}_{[a,b]} * K_\tau$ est majorée par la somme $\mathbf{1}_{[a,a+\tau]} + \mathbf{1}_{[b,b+\tau]}$, fonction dont la norme est petite quand $\tau \rightarrow 0$. Plus précisément, pour tout p tel que $1 \leq p < +\infty$, on peut majorer la norme L^p de la différence

$$\begin{aligned} \|\mathbf{1}_{[a,b]} - \mathbf{1}_{[a,b]} * K_\tau\|_p^p &= \int_{\mathbb{R}} \left| \mathbf{1}_{[a,b]}(t) - (\mathbf{1}_{[a,b]} * K_\tau)(t) \right|^p dt \\ &\leq \int_{\mathbb{R}} \left| \mathbf{1}_{[a,a+\tau]}(t) + \mathbf{1}_{[b,b+\tau]}(t) \right|^p dt = 2\tau. \end{aligned}$$

ce qui montre que

$$(7) \quad \|\mathbf{1}_{[a,b]} - \mathbf{1}_{[a,b]} * K_\tau\|_p \leq (2\tau)^{1/p}.$$

Si on fait un peu plus attention, on peut arriver à majorer par $\tau^{1/p}$ (au lieu de $(2\tau)^{1/p}$). Dans le cas $0 < b - a \leq \tau$, on peut voir que la majoration (7) reste juste (mais pas le dessin), mais en fait on a dans ce cas $\|\mathbf{1}_{[a,b]} - \mathbf{1}_{[a,b]} * K_\tau\|_p \leq (2b - 2a)^{1/p}$, ce qui est mieux.

On va montrer une propriété essentielle qui concerne les rapports entre Fourier et convolution ; elle est importante pour les applications, en particulier pour la notion de *filtre* que nous verrons plus loin.

Théorème 1.2.2. *Si $f, g \in L^1(\mathbb{R})$, on a*

$$\widehat{f * g} = \widehat{f} \widehat{g}.$$

Preuve. — On a

$$\begin{aligned} (\widehat{f * g})(y) &= \int_{\mathbb{R}} (f * g)(x) e^{-ixy} dx = \\ &= \int_{\mathbb{R}^2} f(x-s)g(s) e^{-ixy} dx ds = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x-s) e^{-i(x-s)y} g(s) e^{-isy} ds \right) dx \end{aligned}$$

qui apparaît comme l'intégrale de $f_1 * g_1$, où on aurait posé $f_1(u) = e^{-iuy} f(u)$ et $g_1(u) = e^{-iuy} g(u)$; d'après les rappels sur la convolution de la proposition 1.2.1,

$$\begin{aligned} (\widehat{f * g})(y) &= \int_{\mathbb{R}} f_1 * g_1 = \left(\int_{\mathbb{R}} f_1 \right) \left(\int_{\mathbb{R}} g_1 \right) = \\ &= \left(\int_{\mathbb{R}} e^{-iuy} f(u) du \right) \left(\int_{\mathbb{R}} e^{-iuy} g(u) du \right) = \widehat{f}(y) \widehat{g}(y). \end{aligned}$$

Filtres

Supposons que g soit une fonction dont la transformée de Fourier « ressemble » à l'indicatrice d'un intervalle de la forme $[-2\pi a, 2\pi a]$, par exemple supposons que \widehat{g} soit égale à 1 sur l'intervalle $[-2\pi a, 2\pi a]$, nulle en dehors de $[-2\pi(a + \tau), 2\pi(a + \tau)]$ avec $\tau > 0$ petit, affine sur les deux intervalles $[2\pi a, 2\pi(a + \tau)]$ et $[-2\pi(a + \tau), -2\pi a]$, et continue ; si on forme le signal obtenu par moyenne des translatés f_s de f avec la mesure $g(s) ds$, on voit que l'effet de la convolution avec g est de supprimer dans le signal $f * g$ les fréquences du signal f qui sont au-delà de la valeur $a + \tau$, en gardant intacte la partie du

signal qui contient les fréquences plus petites que a ; en effet, on a vu que la transformée de Fourier de $f * g$ indique le poids des différentes fréquences dans le signal total ; or cette transformée est la fonction $\widehat{f\hat{g}}$, dans laquelle on a (à peu de chose près) annulé ce qui se passe en dehors de $[-2\pi a, 2\pi a]$ sans modifier l'intérieur. Il existe des dispositifs physiques (en particulier des circuits électroniques) dont l'action sur un signal d'entrée f ressemble à ce qui est décrit dans les lignes précédentes. On appelle ces dispositifs des *filtres*, et par extension on nommera ainsi des fonctions g qui pourraient *théoriquement* avoir cet effet.

Pour qu'un filtre soit réalisable, le minimum est qu'il soit obtenu en mélangeant des *retardés* du signal f , c'est-à-dire qu'on limite la mesure $g(s) ds$ aux valeurs $s > 0$, ce qui veut dire qu'on suppose que la fonction g est nulle sur $]-\infty, 0]$. Il est en effet impossible d'utiliser des *anticipés* du signal, correspondant à des valeurs futures qui sont normalement inconnues ! On pourra voir en exercice que le filtre mathématique g proposé au début du paragraphe précédent *n'est pas* physiquement réalisable.

Dans d'autres contextes, la convolution peut être utilisée pour faire par exemple une moyenne *spatiale* de traduits d'un phénomène, au lieu de la moyenne temporelle que nous avons envisagée. Dans un tel cas, il n'y aura en général aucune raison de limiter le support de g au côté positif. On en verra un exemple plus loin à propos de l'équation de la chaleur.

Calcul de la transformée de Fourier de K_τ

On voit que $\widehat{K}_\tau(0) = 1$, et pour $\xi \neq 0$ on a

$$(8) \quad \widehat{K}_\tau(\xi) = \frac{1}{\tau} \int_0^\tau e^{-i\xi t} dt = \frac{1 - e^{-i\tau\xi}}{i\tau\xi} = e^{-i\tau\xi/2} \frac{\sin(\tau\xi/2)}{\tau\xi/2}.$$

Si on pense à K_τ comme à un filtre (un filtre mathématique, pas un filtre réaliste) agissant sur le signal f pour donner le «signal filtré» $f * K_\tau$, on voit que les pulsations $\xi \neq 0$ telles que $\sin(\tau\xi/2) = 0$ seront «tuées» par la convolution avec K_τ , ce qui correspond à $\tau\xi \in 2\pi\mathbb{Z}$, et ξ non nul ; en terme de fréquence $F = \xi/(2\pi)$, la première fréquence annulée est $F = \tau^{-1}$, et les autres sont les multiples de τ^{-1} . Si on veut écouter une radio qui émet sur $100 \text{ MHz} = 10^8 \text{ Hz}$, il ne faut sûrement pas commencer par faire la moyenne de son signal sur un intervalle de temps de $\tau = 10^{-8} \text{ s}$!

1.3. Fourier dans $L^2(\mathbb{R})$

On a montré à la proposition 1.1.10 que la transformation de Fourier, agissant sur le petit sous-espace vectoriel $X \subset L^1(\mathbb{R})$ de la définition 1.1.7, possède des propriétés intéressantes pour la norme L^2 . On va maintenant exploiter à fond ce résultat. On va commencer par un résultat très simple mais utile, qu'on aurait pu déduire des calculs de convolution avec la fonction K_τ , mais qu'il est plus raisonnable de démontrer directement, sans parler de convolution. Il s'agit d'approcher les fonctions en escalier par des fonctions de l'espace $\text{Aff}_c(\mathbb{R})$ des fonctions affines par morceaux, ce qui est bien évident sur un dessin !

Lemme 1.3.1. *Si g est une fonction en escalier sur \mathbb{R} , nulle en dehors de $[a, b]$, il existe pour tout p tel que $1 \leq p < +\infty$ et tout $\varepsilon > 0$ une fonction $\varphi \in \text{Aff}_c(\mathbb{R})$ nulle en dehors de $[a, b]$ et telle que*

$$\|\varphi - g\|_p < \varepsilon.$$

Preuve. — On commence par établir le cas où g est une fonction indicatrice d'intervalle, $g = \chi = \mathbf{1}_{[u,v]}$. On considère ε tel que $0 < \varepsilon < (v - u)/2$, et on introduit la fonction affine par morceaux h , nulle en dehors de $[u, v]$ telle que $h(u) = h(v) = 0$, et telle que $h(u + \varepsilon) = h(v - \varepsilon) = 1$.



La fonction h est affine sur les trois intervalles $[u, u + \varepsilon]$, $[u + \varepsilon, v - \varepsilon]$ et $[v - \varepsilon, v]$. Comme χ et h sont nulles en dehors de $[u, v]$, on a

$$\int_{\mathbb{R}} |h(x) - \chi(x)|^p dx = \int_u^v |h(x) - \chi(x)|^p dx.$$

Sur l'intervalle $[u, v]$, la fonction $h - \chi$ est affine par morceaux : elle décroît de 1 à 0 entre u et $u + \varepsilon$, elle est nulle entre $u + \varepsilon$ et $v - \varepsilon$, et enfin elle croît de 0 à 1 entre $v - \varepsilon$ et v ; on voit **(d)** donc que

$$\|h - \chi\|_p^p = \int_u^v |h(x) - \chi(x)|^p dx \leq \int_u^{u+\varepsilon} 1^p dx + \int_{v-\varepsilon}^v 1^p dx = 2\varepsilon.$$

Passons au cas général. On peut exprimer la fonction g sous la forme d'une combinaison linéaire d'indicatrices d'intervalles $g = \sum_{j=1}^N c_j \chi_j$ où χ_j est de la forme $\mathbf{1}_{[a_j, b_j]}$ avec $a \leq a_j \leq b_j \leq b$. Pour chaque $j = 1, \dots, N$ on peut trouver une fonction h_j affine par morceaux, continue, nulle hors de $[a_j, b_j]$ et telle que

$$\|h_j - \chi_j\|_p < \frac{\varepsilon}{1 + \sum_{j=1}^N |c_j|};$$

la fonction $\varphi = \sum_{j=1}^N c_j h_j$ est nulle hors de $[a, b]$, continue, affine par morceaux et d'après l'inégalité triangulaire on a

$$\|\varphi - g\|_p \leq \sum_{j=1}^N |c_j| \|h_j - \chi_j\|_p < \varepsilon.$$

Lemme 1.3.2. *Pour tout p tel que $1 \leq p < +\infty$ et toute fonction $f \in L^p(\mathbb{R})$ nulle en dehors d'un intervalle $[a, b]$, il existe une suite $(\varphi_n) \subset \text{Aff}_c(\mathbb{R})$ de fonctions nulles hors de $[a, b]$ telle que*

$$\|f - \varphi_n\|_p \rightarrow 0.$$

Preuve. — Puisque l'espace des fonctions en escalier est dense dans $L^p(\mathbb{R})$ (voir le théorème A.11 de l'annexe), il existe une fonction g en escalier telle que

$$\|g - f\|_p < \varepsilon/2.$$

On peut supposer que g est nulle hors de $[a, b]$, en la remplaçant par $g_1 = \mathbf{1}_{[a,b]}g$; en effet, g_1 est encore en escalier et puisque f est nulle hors de $[a, b]$, on a

$$\|g_1 - f\|_p^p = \int_a^b |g(x) - f(x)|^p dx \leq \int_{\mathbb{R}} |g(x) - f(x)|^p dx = \|g - f\|_p^p.$$

Par le lemme 1.3.1, on peut trouver une fonction $\varphi \in \text{Aff}_c(\mathbb{R})$, nulle en dehors de $[a, b]$ et telle que $\|\varphi - g\|_p < \varepsilon/2$. On a donc

$$\|f - \varphi\|_p < \varepsilon.$$

Il ne reste plus qu'à appliquer ce qui précède à une suite (ε_n) de valeurs tendant vers 0.

Corollaire 1.3.3. *L'espace X de la définition 1.1.7 est dense dans $L^2(\mathbb{R})$.*

Preuve. — On sait d'après le lemme 1.1.8 que l'espace $\text{Aff}_c(\mathbb{R})$ est contenu dans X . Pour toute $f \in L^2(\mathbb{R})$ et tout $\varepsilon > 0$, on sait d'après le lemme A.9 que pour n assez grand, on a $\|f - \mathbf{1}_{[-n,n]}f\|_2 < \varepsilon/2$. D'après le lemme précédent, il existe $\varphi \in \text{Aff}_c(\mathbb{R})$ telle que $\|\varphi - \mathbf{1}_{[-n,n]}f\|_2 < \varepsilon/2$; on a donc trouvé une fonction $\varphi \in X$ telle que $\|\varphi - f\|_2 < \varepsilon$.

Proposition 1.3.4. *Si $f \in L^2(\mathbb{R})$ est nulle en dehors d'un intervalle borné de \mathbb{R} , alors f est intégrable, \widehat{f} est de carré intégrable et on a l'égalité de Parseval*

$$\|\widehat{f}\|_2 = \sqrt{2\pi} \|f\|_2.$$

Preuve. — On suppose que $f \in L^2(\mathbb{R})$ est nulle en dehors d'un intervalle borné $[-a, a]$. Dans ce cas f est aussi dans $L^1(\mathbb{R})$ d'après le lemme A.7 de l'annexe, et la fonction \widehat{f} est bien définie. D'après le lemme 1.3.2 on sait qu'il existe une suite (φ_n) de fonctions de X qui converge vers f dans L^2 , et qui sont nulles en dehors de $[-a, a]$; ceci implique qu'elles tendent vers f au sens de L^1 également, puisqu'on sait alors d'après le lemme A.7 que $\|f - \varphi_n\|_1 \leq \sqrt{2a} \|f - \varphi_n\|_2$. Il en résulte que $\widehat{\varphi_n}$ tend uniformément vers \widehat{f} , puisque d'après (M)

$$\|\widehat{f} - \widehat{\varphi_n}\|_\infty \leq \|f - \varphi_n\|_1 \rightarrow 0.$$

Par ailleurs la suite (φ_n) est de Cauchy dans L^2 puisqu'elle converge vers f , et la relation de Parseval

$$\|\widehat{\varphi_n} - \widehat{\varphi_m}\|_2 = \sqrt{2\pi} \|\varphi_n - \varphi_m\|_2$$

obtenue à la proposition 1.1.10 pour les fonctions de l'espace X montre que la suite $(\widehat{\varphi_n})$ est elle aussi de Cauchy dans L^2 . Il en résulte que la suite $(\widehat{\varphi_n})$ converge dans l'espace complet $L^2(\mathbb{R})$ vers une fonction h . Puisque $\widehat{\varphi_n}$ tend uniformément vers \widehat{f} , cette fonction h ne peut être que \widehat{f} , d'après le corollaire A.4 de l'annexe. Finalement $\widehat{\varphi_n}$ tend vers \widehat{f} dans L^2 et

$$\|\widehat{f}\|_2 = \lim_n \|\widehat{\varphi_n}\|_2 = \sqrt{2\pi} \lim_n \|\varphi_n\|_2 = \sqrt{2\pi} \|f\|_2.$$

Théorème et définition 1.3.5. Si $f \in L^2(\mathbb{R})$ on pose $f_n = \mathbf{1}_{[-n,n]}f$, pour tout entier $n \geq 1$; alors f_n est intégrable, la transformée de Fourier \widehat{f}_n est de carré intégrable et la suite (\widehat{f}_n) converge dans $L^2(\mathbb{R})$; par définition la transformée de Fourier $\mathcal{F}f$ de la fonction $f \in L^2(\mathbb{R})$ est la limite de cette suite,

$$\mathcal{F}f = \lim_{L^2} \widehat{f}_n.$$

Si de plus $f \in L^1(\mathbb{R})$, on a $\mathcal{F}f = \widehat{f}$ presque partout.

L'application $f \rightarrow \mathcal{F}f$ est une application linéaire continue de $L^2(\mathbb{R})$ dans $L^2(\mathbb{R})$, et on a l'identité de Parseval

$$\forall f \in L^2(\mathbb{R}), \quad \|\mathcal{F}f\|_2 = \sqrt{2\pi} \|f\|_2.$$

Preuve. — On voit d'abord par le lemme A.9 que $f_n = \mathbf{1}_{[-n,n]}f$ tend vers f dans L^2 , donc (f_n) est de Cauchy dans L^2 . On a vu à la proposition 1.3.4 que pour toute fonction $h \in L^2(\mathbb{R})$ nulle en dehors d'un intervalle borné, on a $\widehat{h} \in L^2(\mathbb{R})$ et

$$\|\widehat{h}\|_2 = \sqrt{2\pi} \|h\|_2.$$

Cette relation s'applique aux fonctions $f_n - f_m$, qui sont nulles hors d'un intervalle borné,

$$\|\widehat{f}_n - \widehat{f}_m\|_2 = \sqrt{2\pi} \|f_n - f_m\|_2;$$

puisque (f_n) est de Cauchy dans L^2 , on voit que la suite (\widehat{f}_n) est elle aussi de Cauchy dans L^2 , et comme $L^2(\mathbb{R})$ est complet cette suite est convergente et on peut poser (e)

$$\mathcal{F}f = \lim_{L^2} \widehat{f}_n.$$

On obtient l'égalité de Parseval en disant que la norme de la limite est la limite des normes et en rappelant que $\|\widehat{f}_n\|_2 = \sqrt{2\pi} \|f_n\|_2$ pour tout n , par la proposition 1.3.4. Montrons que l'application \mathcal{F} est linéaire : si g est une autre fonction de $L^2(\mathbb{R})$ on associera à la combinaison linéaire $af + bg \in L^2(\mathbb{R})$ la suite

$$(af + bg)_n = \mathbf{1}_{[-n,n]}(af + bg) = af_n + bg_n,$$

et par l'addition des limites dans un espace vectoriel normé et la linéarité de la transformation de Fourier sur $L^1(\mathbb{R})$

$$\mathcal{F}(af + bg) = \lim_n (af_n + bg_n) = a \lim_n \widehat{f}_n + b \lim_n \widehat{g}_n = a \mathcal{F}f + b \mathcal{F}g.$$

Enfin, si f est à la fois dans $L^1(\mathbb{R})$ et dans $L^2(\mathbb{R})$, la suite (f_n) tend aussi vers f pour la norme de L^1 (lemme A.9), ce qui entraîne que \widehat{f}_n tend vers \widehat{f} uniformément. Comme (\widehat{f}_n) tend vers $\mathcal{F}f$ en norme L^2 , on conclut que $\mathcal{F}f = \widehat{f}$ presque partout (corollaire A.4 de l'annexe).

Intégrales semi-convergentes et Fourier

Proposition 1.3.6. Soit (a_n) une suite de réels qui tend vers $-\infty$ et soit (b_n) une suite de réels qui tend vers $+\infty$; si $f \in L^2(\mathbb{R})$ et si la limite

$$F(y) = \lim_n \int_{a_n}^{b_n} e^{-ixy} f(x) dx$$

existe pour presque tout $y \in \mathbb{R}$, alors $\mathcal{F}f = F$ presque partout. En particulier, si la fonction f est dans $L^2(\mathbb{R})$ et si l'intégrale généralisée

$$F(y) = \int_{-\infty}^{+\infty} e^{-ixy} f(x) dx$$

est semi-convergente pour presque tout $y \in \mathbb{R}$, alors $\mathcal{F}f = \widehat{F}$ presque partout.

Preuve. — Si on pose $f_n = \mathbf{1}_{[a_n, b_n]} f$, on a

$$\int_{a_n}^{b_n} e^{-ixy} f(x) dx = \widehat{f}_n(y),$$

donc l'hypothèse de la proposition est que (\widehat{f}_n) converge vers F presque partout ; mais comme on sait que (f_n) tend vers f dans L^2 (lemme A.9), on déduit que (\widehat{f}_n) tend vers $\mathcal{F}f$ dans L^2 d'après le théorème 1.3.5 ; il en résulte que $\mathcal{F}f = F$ presque partout (corollaire A.4).

Inversion de Fourier dans $L^2(\mathbb{R})$

Désignons par σ l'application qui associe à chaque fonction f sur \mathbb{R} la fonction σf qui est définie sur \mathbb{R} en posant $(\sigma f)(x) = f(-x)$ pour tout x . Il est clair que σ est linéaire, isométrique sur $L^p(\mathbb{R})$:

$$\forall f \in L^p(\mathbb{R}), \quad \|\sigma f\|_p = \|f\|_p ;$$

de plus, il est évident que $\sigma \circ \sigma = \text{Id}$.

Sur l'espace X de la définition 1.1.7, on a $\mathcal{F} \circ \mathcal{F} = 2\pi \sigma$ et $\mathcal{F} \circ \sigma = \sigma \circ \mathcal{F}$: en effet, les fonctions $f \in X$ sont dans $L^2(\mathbb{R})$ et dans $L^1(\mathbb{R})$, donc $\mathcal{F}f = \widehat{f}$ d'après le théorème 1.3.5 ; si $f \in X$ on sait (théorème 1.1.9) que pour tout $x \in \mathbb{R}$, la valeur de $2\pi(\sigma f)(x)$ est égale à

$$2\pi f(-x) = \int_{\mathbb{R}} e^{i(-x)y} \widehat{f}(y) dy = \int_{\mathbb{R}} e^{-ixy} \widehat{f}(y) dy = (\mathcal{F}\widehat{f})(x) = (\mathcal{F} \circ \mathcal{F})(f)(x),$$

donc $2\pi(\sigma f) = (\mathcal{F} \circ \mathcal{F})(f)$, et par ailleurs

$$(\mathcal{F} \circ \sigma)(f)(y) = \mathcal{F}(\sigma f)(y) = \int_{\mathbb{R}} e^{-ixy} f(-x) dx = \int_{\mathbb{R}} e^{-iu(-y)} f(u) du = (\mathcal{F}f)(-y),$$

qui est égal à $(\sigma \circ \mathcal{F})(f)(y)$. Grâce à la densité de l'espace vectoriel X dans $L^2(\mathbb{R})$ (corollaire 1.3.3) on déduit le résultat qui suit.

Théorème 1.3.7. Pour la transformation de Fourier \mathcal{F} agissant de $L^2(\mathbb{R})$ dans $L^2(\mathbb{R})$ on a

$$\mathcal{F} \circ \mathcal{F} = 2\pi \sigma, \quad \mathcal{F} \circ \sigma = \sigma \circ \mathcal{F}$$

donc \mathcal{F} est inversible et

$$\mathcal{F}^{-1} = \frac{1}{2\pi} \mathcal{F} \circ \sigma = \frac{1}{2\pi} \sigma \circ \mathcal{F}.$$

En particulier, si la fonction g est à la fois dans $L^2(\mathbb{R})$ et dans $L^1(\mathbb{R})$, on peut définir (un représentant de) $\mathcal{F}^{-1}g \in L^2(\mathbb{R})$ en posant

$$\forall x \in \mathbb{R}, \quad (\mathcal{F}^{-1}g)(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{ixy} g(y) dy.$$

Preuve. — L'application $T = \mathcal{F} \circ \mathcal{F} - 2\pi\sigma$ est continue de $L^2(\mathbb{R})$ dans $L^2(\mathbb{R})$, donc l'ensemble

$$Y = \{f \in L^2(\mathbb{R}) : Tf = 0\}$$

est un sous-ensemble fermé de $L^2(\mathbb{R})$. Mais on a dit qu'il contient X , donc Y contient aussi l'adhérence \overline{X} , qui est égale à $L^2(\mathbb{R})$ puisque X est dense (corollaire 1.3.3), donc $Y = L^2(\mathbb{R})$ et $T = 0$, ce qu'il fallait démontrer pour le premier point. La preuve du second est analogue. Pour trouver l'inverse de \mathcal{F} à partir des premières conclusions, on remarque que

$$(\mathcal{F} \circ \sigma) \circ \mathcal{F} = \mathcal{F} \circ (\mathcal{F} \circ \sigma) = \mathcal{F} \circ \mathcal{F} \circ \sigma = 2\pi \sigma \circ \sigma = 2\pi \text{Id}_{L^2}.$$

Pour obtenir la dernière conclusion, on se rappelle que dans le cas où $g \in L^2(\mathbb{R})$ est aussi intégrable, la transformée $\mathcal{F}g$ peut se calculer comme \widehat{g} (théorème 1.3.5), ce qui entraîne que $(\mathcal{F}^{-1}g)(x) = (2\pi)^{-1}(\sigma\widehat{g})(x) = (2\pi)^{-1}\widehat{g}(-x)$, comme annoncé.

Exemple de la fonction $y \rightarrow \sin(y)/y$

Puisque $g(y) = \sin(y)/y$ est la transformée de Fourier de $f = \frac{1}{2} \mathbf{1}_{[-1,1]}$ et que g est paire, on a $g = g \circ \sigma$ et l'inversion de Fourier dans L^2 nous dit que

$$\mathbf{1}_{[-1,1]} = \frac{1}{\pi} \mathcal{F}g$$

en tant que classe de fonctions. Par ailleurs on va montrer que

$$(*) \quad \int_{-n}^n e^{-ixy} g(y) dy = \int_{-n}^n \cos(xy) \sin(y) \frac{dy}{y}$$

tend vers une limite qu'on va identifier, pour tout $x \in \mathbb{R}$; on sait (f) que

$$V = \lim_{n \rightarrow +\infty} \int_0^n \frac{\sin y}{y} dy = \int_0^{+\infty} \frac{\sin y}{y} dy = \frac{1}{2} \int_{-\infty}^{+\infty} \frac{\sin y}{y} dy$$

existe, et il en résulte par changement de variable que

$$\lim_{n \rightarrow +\infty} \int_{-n}^n \frac{\sin(ay)}{y} dy$$

est égale à $2V$ pour tout $a > 0$ et à $-2V$ pour tout $a < 0$. En utilisant les formules d'addition des sinus, on voit que

$$2 \int_{-n}^n \cos(xy) \sin(y) \frac{dy}{y} = \int_{-n}^n \sin[(1+x)y] \frac{dy}{y} + \int_{-n}^n \sin[(1-x)y] \frac{dy}{y}$$

tend vers $4V$ lorsque $|x| < 1$, et vers 0 quand $|x| > 1$ (il y a aussi convergence quand $x = \pm 1$, la limite est $2V$ dans ce cas mais nous ne l'utiliserons pas). Il en résulte d'après la relation (*) et la proposition 1.3.6 que $\mathcal{F}g = 2V \mathbf{1}_{[-1,1]}$ comme classe de fonctions, et on a dit que $\mathcal{F}g = \pi \mathbf{1}_{[-1,1]}$. On déduit donc de l'inversion de Fourier que $V = \pi/2$, c'est-à-dire qu'on a trouvé la valeur V de l'intégrale,

$$\int_0^{+\infty} \frac{\sin y}{y} dy = \frac{\pi}{2}.$$

On peut noter que

$$\int_0^{+\infty} \frac{\sin^2 y}{y^2} dy = \frac{\pi}{2};$$

en effet, on sait que $g(y) = \sin(y)/y$ est la transformée de Fourier de $f(x) = \frac{1}{2} \mathbf{1}_{[-1,1]}(x)$, et on obtient donc que

$$\int_{\mathbb{R}} g^2(y) dy = 2\pi \int_{\mathbb{R}} f^2(x) dx = \frac{2\pi}{4} \int_{-1}^1 dx = \pi$$

par l'identité de Parseval.

1.4. Compléments sur la convolution et sur Fourier

Convolutions $L^p * L^1$

Proposition 1.4.1. *Si $f \in L^p(\mathbb{R})$ et $g \in L^1(\mathbb{R})$, avec $1 \leq p \leq +\infty$, alors pour presque tout x la fonction $y \rightarrow f(x-y)g(y)$ est intégrable sur \mathbb{R} , et la convolée $f * g$ peut donc être définie par l'équation (4); de plus, $f * g \in L^p(\mathbb{R})$ et*

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

Donnons la preuve du cas $p = 2$. On suppose d'abord f et g mesurables positives; avec Cauchy-Schwarz (g),

$$\left(\int_{\mathbb{R}} f(x-y)g(y) dy \right)^2 = \left(\int_{\mathbb{R}} f(x-y)\sqrt{g(y)}\sqrt{g(y)} dy \right)^2 \leq$$

$$\left(\int_{\mathbb{R}} f^2(x-y)g(y) dy \right) \left(\int_{\mathbb{R}} g(y) dy \right) = \|g\|_1 \int_{\mathbb{R}} f^2(x-y)g(y) dy = \|g\|_1 (f^2 * g)(x)$$

avec f^2 et g intégrables; d'après la proposition 1.2.1,

$$\int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x-y)g(y) dy \right)^2 dx \leq \|g\|_1 \|f^2\|_1 \|g\|_1 = \|g\|_1^2 \|f\|_2^2.$$

Quand f et g ne sont pas positives, on considère d'abord les fonctions $|f|$ et $|g|$. Par ce qui précède, on obtient pour commencer que $I(x) = \int |f(x-y)||g(y)| dy$ est fini pour presque tout x puisqu'on a vu que $\int I(x)^2 dx < +\infty$. Ceci permet de justifier l'existence de l'intégrale qui définit la convolution $f * g$: pour presque tout x , on peut poser

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

La majoration de norme est facile, car l'inégalité

$$\left| \int_{\mathbb{R}} f(x-y)g(y) dy \right|^2 \leq \left(\int_{\mathbb{R}} |f(x-y)||g(y)| dy \right)^2$$

nous ramène au cas positif déjà traité.

Le cas $p = +\infty$ est facile (exercice) et doit se traiter à part. Le cas $p = 1$ est déjà connu (proposition 1.2.1), et le cas $1 < p < +\infty$, $p \neq 2$ se traite en remplaçant Cauchy-Schwarz par Hölder (introduire l'exposant q conjugué de p , écrire $g(y) = g(y)^{1/p} g(y)^{1/q}$ et adapter ce qui précède).

Approximation de l'unité dans $L^p(\mathbb{R})$

On va énoncer un cas très particulier d'une méthode générale appelée *approximation par convolution*. On en verra un autre exemple au chapitre suivant avec le lemme 2.2.2.

Théorème 1.4.2. *Si $f \in L^p(\mathbb{R})$, avec $1 \leq p < +\infty$, la convolée $f * K_\tau$ converge vers f dans L^p lorsque $\tau \rightarrow 0$.*

Preuve. — Soit $\varepsilon > 0$ donné ; d'après le cours d'intégration (voir aussi le théorème A.11 de l'annexe), on peut trouver une fonction en escalier φ telle que l'on ait $\|f - \varphi\|_p < \varepsilon/3$. On a vu à l'équation (7) que $\|h - h * K_\tau\|_p < (2\tau)^{1/p}$ lorsque h est une indicatrice d'intervalle borné, donc pour τ assez petit, on aura par linéarité $\|\varphi - \varphi * K_\tau\|_p < \varepsilon/3$; d'après la proposition 1.4.1 on a aussi

$$\|f * K_\tau - \varphi * K_\tau\|_p = \|(f - \varphi) * K_\tau\|_p \leq \|f - \varphi\|_p \|K_\tau\|_1 = \|f - \varphi\|_p < \varepsilon/3.$$

Finalement pour $\tau > 0$ assez petit on a $\|f - f * K_\tau\|_p < \varepsilon$ par l'inégalité triangulaire.

Corollaire. *Si f et \widehat{f} sont dans $L^1(\mathbb{R})$, on a pour presque tout x*

$$2\pi f(x) = \int_{\mathbb{R}} e^{ixy} \widehat{f}(y) dy.$$

Preuve. — Si f vérifie ces hypothèses, on voit que $f * K_\tau$ est dans l'espace X de la définition 1.1.7 : lorsque $f \in L^1$, on a toujours que $f * K_\tau$ est continue bornée intégrable par le résumé (6), et de plus ici $f * \widehat{K}_\tau = \widehat{f} \widehat{K}_\tau$ est bornée en module par $|\widehat{f}|$, donc intégrable d'après l'hypothèse du corollaire. Il en résulte que $f * K_\tau \in X$, donc $f * K_\tau$ vérifie la formule de Fourier inverse ponctuelle du théorème 1.1.9

$$2\pi (f * K_\tau)(x) = \int_{\mathbb{R}} e^{itx} \widehat{f * K_\tau}(t) dt = \int_{\mathbb{R}} e^{itx} \widehat{K}_\tau(t) \widehat{f}(t) dt.$$

Quand τ tend vers 0, $\widehat{K}_\tau(t)$ tend vers 1 en restant majoré par 1 (voir l'équation (8)), donc l'intégrale

$$2\pi (f * K_\tau)(x) = \int_{\mathbb{R}} e^{itx} \widehat{K}_\tau(t) \widehat{f}(t) dt$$

tend vers

$$F(x) = \int_{\mathbb{R}} e^{itx} \widehat{f}(t) dt$$

par Lebesgue dominé, mais $2\pi f * K_\tau$ tend vers $2\pi f$ dans L^1 par le théorème précédent. Le résultat $F = 2\pi f$ presque partout en découle par le corollaire A.4 de l'annexe.

Remarque. Ce corollaire donne une amélioration très légère au théorème 1.1.9. Dans ce théorème, on suppose en plus que f est continue bornée; ici, on montre que cette hypothèse est superflue : si on sait seulement que f et \widehat{f} sont dans $L^1(\mathbb{R})$, on déduit du corollaire que f admet un représentant continu et borné. Cette amélioration a un intérêt théorique (par exemple pour obtenir le corollaire qui suit, qui est important) : mais dans la pratique, si on rencontre explicitement une fonction telle que f , on devrait bien s'apercevoir qu'elle est continue et bornée!

Corollaire. *La transformation de Fourier est injective sur $L^1(\mathbb{R})$.*

En effet, si $f \in L^1(\mathbb{R})$ et $\widehat{f} = 0$, on a bien f et \widehat{f} intégrables et le corollaire précédent s'applique. Comme $\widehat{f} = 0$, la formule d'inversion donne $f = 0$.

1.5. Fourier et dérivation

Proposition 1.5.1. *Si f et $x \rightarrow xf(x)$ sont intégrables sur \mathbb{R} , la transformée de Fourier \widehat{f} est continûment dérivable sur \mathbb{R} et sa dérivée est obtenue par dérivation sous l'intégrale,*

$$(\widehat{f})'(y) = -i \int_{\mathbb{R}} e^{-ixy} xf(x) dx.$$

Si f est de classe C^1 par morceaux sur \mathbb{R} , avec f et f' intégrables sur \mathbb{R} , alors

$$\widehat{f}'(y) = iy \widehat{f}(y).$$

Preuve. — Le premier résultat est obtenu par dérivation sous l'intégrale. Pour le second, on écrit grâce au lemme 1.1.2

$$(f' * K_\tau)(x) = \frac{1}{\tau} \int_{x-\tau}^x f'(u) du = \frac{f(x) - f(x-\tau)}{\tau};$$

comme $f' \in L^1(\mathbb{R})$, on sait que $f' * K_\tau$ tend dans L^1 vers f' quand $\tau \rightarrow 0$ d'après le théorème 1.4.2, donc $f' * \widehat{K}_\tau$ tend uniformément vers \widehat{f}' par la relation (M); d'après la partie droite de l'équation précédente, la transformée de Fourier de $f' * K_\tau$ est égale à

$$y \rightarrow \frac{1 - e^{-i\tau y}}{\tau} \widehat{f}(y)$$

qui tend simplement vers $y \rightarrow iy \widehat{f}(y)$ quand $\tau \rightarrow 0$, d'où le résultat $\widehat{f}'(y) = iy \widehat{f}(y)$.

Exemple de la densité gaussienne

La fonction f définie sur \mathbb{R} par

$$f(x) = e^{-x^2/2}$$

est intégrable, et $x \rightarrow xf(x)$ est intégrable aussi : il en résulte que la fonction g définie par $g(y) = \widehat{f}(y)$ est de classe C^1 . On obtient par la proposition ci-dessus, suivie d'une intégration par parties

$$g'(y) = -i \int_{\mathbb{R}} x e^{-x^2/2} e^{-ixy} dx = i \left(\int_{\mathbb{R}} iy e^{-x^2/2} e^{-ixy} dx \right) = -y g(y).$$

La résolution de l'équation différentielle $g'(y) = -y g(y)$, compte-tenu de l'égalité

$$g(0) = \widehat{f}(0) = \int_{\mathbb{R}} f(x) dx = \int_{\mathbb{R}} e^{-x^2/2} dx = \sqrt{2\pi}$$

donne

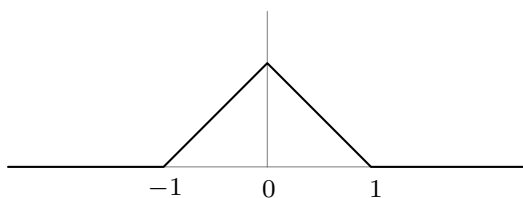
$$\widehat{f}(y) = g(y) = \sqrt{2\pi} e^{-y^2/2} = \sqrt{2\pi} f(y).$$

On peut observer que la fonction f est un vecteur propre de la transformation de Fourier, correspondant à la valeur propre $\sqrt{2\pi}$.

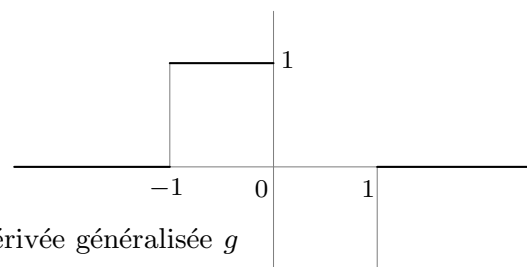
On dira que la fonction g , intégrable sur tout intervalle borné de \mathbb{R} , est la *dérivée généralisée* de la fonction continue G si on a pour tous $u \leq v$

$$G(v) - G(u) = \int_u^v g(s) ds.$$

C'est le cas par exemple si G est continue, linéaire par morceaux, avec $G = 0$ hors de $[-1, 1]$, $G(0) = 1$ et G linéaire sur $[-1, 0]$ et $[0, 1]$. La dérivée généralisée g de cette fonction G est égale à $g = \mathbf{1}_{[-1,0]} - \mathbf{1}_{[0,1]}$ dans ce cas.



Graphique de G



Dérivée généralisée g

Lorsque G est de classe C^1 , la dérivée généralisée g est simplement la dérivée usuelle G' . La première partie de la proposition 1.5.2 correspond au deuxième cas de la proposition 1.5.1 : lorsque G est de classe C^1 avec G, G' à la fois dans L^1 et L^2 , les deux énoncés peuvent être appliqués à la fonction G .

Proposition 1.5.2. *On suppose que g et G sont deux fonctions de $L^2(\mathbb{R})$ telles que G soit continue sur \mathbb{R} et*

$$G(v) - G(u) = \int_u^v g(s) \, ds$$

pour tous $u \leq v$; on a alors pour presque tout y

$$(\mathcal{F}g)(y) = iy(\mathcal{F}G)(y).$$

Inversement, si $G \in L^2(\mathbb{R})$ est continue et si $y \rightarrow y(\mathcal{F}G)(y)$ est de carré intégrable, il existe une fonction $g \in L^2(\mathbb{R})$ telle que

$$G(v) - G(u) = \int_u^v g(s) \, ds$$

pour tous $u \leq v$.

Preuve. — La première partie se fait comme dans le cas de L^1 (preuve de la proposition 1.5.1), en disant ici que $g * K_\tau$ tend vers g dans L^2 (théorème 1.4.2), et en exploitant la définition de la dérivée généralisée,

$$(g * K_\tau)(x) = \frac{1}{\tau} \int_{x-\tau}^x g(s) \, ds = \frac{G(x) - G(x-\tau)}{\tau};$$

de plus, on vérifie que la transformée de Fourier de $x \rightarrow G(x-\tau)$ est $\xi \rightarrow e^{-i\tau\xi}(\mathcal{F}G)(\xi)$, comme dans le cas de $L^1(\mathbb{R})$ (revenir à la définition 1.3.5).

Pour la seconde partie : d'après la preuve du résultat sur la dérivabilité de la transformée de Fourier des fonctions de L^1 on voit que la fonction G_n définie sur \mathbb{R} par

$$G_n(x) = \frac{1}{2\pi} \int_{-n}^n e^{ixy}(\mathcal{F}G)(y) \, dy$$

est continûment dérivable, avec dérivée $g_n = G'_n$ exprimée par

$$g_n(x) = \frac{1}{2\pi} \int_{-n}^n iy e^{ixy}(\mathcal{F}G)(y) \, dy$$

ce qui implique que pour tous $u \leq v$, $G_n(v) - G_n(u) = \int_u^v g_n(s) \, ds$. De plus, si on pose

$$H_n(y) = \mathbf{1}_{[-n,n]}(y)(\mathcal{F}G)(y), \quad h_n(y) = \mathbf{1}_{[-n,n]}(y)iy(\mathcal{F}G)(y),$$

on voit que $G_n = \mathcal{F}^{-1}H_n$ et $g_n = \mathcal{F}^{-1}h_n$ (théorème 1.3.7); d'après le lemme A.9 les suites (H_n) et (h_n) tendent dans L^2 vers $\mathcal{F}G$ et $h : y \rightarrow iy(\mathcal{F}G)(y)$ respectivement. Puisque \mathcal{F}^{-1} est continue de L^2 dans L^2 , il en résulte que G_n tend dans L^2 vers $\mathcal{F}^{-1}\mathcal{F}G = G$, et g_n tend dans L_2 vers la fonction $g = \mathcal{F}^{-1}h$. Puisque G_n tend vers G dans L^2 , on peut trouver une sous-suite (G_{n_k}) telle que $G(u) = \lim_k G_{n_k}(u)$ pour presque tout u . Si on choisit un point u_0 tel que $G(u_0) = \lim_k G_{n_k}(u_0)$, on obtiendra **(h)** pour tout $v \in \mathbb{R}$

$$\Psi(v) := G(u_0) + \int_{u_0}^v g(s) \, ds = \lim_k \left(G_{n_k}(u_0) + \int_{u_0}^v g_{n_k}(s) \, ds \right) = \lim_k G_{n_k}(v).$$

Il en résulte que la fonction continue Ψ définie à la ligne précédente est égale à G presque partout, donc partout (théorème A.5). Ceci termine la preuve : on a bien trouvé une fonction $g \in L^2(\mathbb{R})$ telle que

$$G(v) - G(u) = \Psi(v) - \Psi(u) = \int_u^v g(s) \, ds$$

pour tous $u < v$.

Corollaire 1.5.3. On suppose que G est de classe C^1 sur \mathbb{R} , que G et G' sont dans $L^2(\mathbb{R})$; alors, pour presque tout y

$$(\mathcal{F}(G'))(y) = iy(\mathcal{F}G)(y).$$

Comme on l'a déjà dit, cet énoncé correspond, pour L^2 , à la deuxième partie de la proposition 1.5.1, qui s'appliquait à L^1 .

Corollaire 1.5.4. On suppose que G et $H : x \rightarrow xG(x)$ sont deux fonctions de $L^2(\mathbb{R})$; alors $\mathcal{F}G$ admet une dérivée généralisée égale à

$$-i\mathcal{F}H \in L^2(\mathbb{R}).$$

C'est la deuxième partie de la proposition 1.5.2, lue à l'envers (Fourier au lieu de Fourier inverse). Cet énoncé correspond, pour L^2 , à la première partie de la proposition 1.5.1.

Équation de la chaleur sur la droite

On suppose donnée une fonction v sur \mathbb{R} où $v(x)$ représente la température au point x d'une barre rectiligne infinie, à l'instant $t = 0$. Si la barre est homogène, la théorie physique prévoit que l'évolution de la température au cours du temps $t \geq 0$ est régie par l'équation de la chaleur

$$\frac{\partial u}{\partial t} = \kappa \frac{\partial^2 u}{\partial x^2}$$

où κ est une constante dépendant de la barre, et où $u(x, t)$ représente la température au point x à l'instant $t \geq 0$. On cherche donc une fonction de deux variables $u(x, t)$ telle que $u(x, 0) = v(x)$, et telle que pour tous $t > 0$ et $x \in \mathbb{R}$, on ait

$$\frac{\partial}{\partial t} u(x, t) = \kappa \frac{\partial^2}{\partial x^2} u(x, t).$$

En utilisant les rapports entre Fourier et dérivation, on peut transformer cette équation aux dérivées partielles en une équation différentielle ordinaire, moyennant quelques hypothèses optimistes : on va supposer qu'on peut trouver une solution telle que $x \rightarrow u(x, t)$ soit de classe C^2 , intégrable ainsi que ses deux dérivées en x , pour tout $t > 0$. Posons alors

$$w(x, t) = w_t(\xi) = \int_{\mathbb{R}} e^{-i\xi x} u(x, t) dx.$$

Pour tout $t > 0$ fixé, w_t est la transformée de Fourier de la fonction u_t définie par $u_t(x) = u(x, t)$. En utilisant les théorèmes précédents, on verra que la transformée de Fourier de $u_t'' = \frac{\partial^2 u}{\partial x^2}$ est égale à

$$\widehat{u_t''}(\xi) = -\xi^2 w_t(\xi).$$

Si on peut dériver w_t par rapport à t sous l'intégrale, on obtient

$$\frac{\partial w_t}{\partial t}(\xi) = \int_{\mathbb{R}} e^{-i\xi x} \frac{\partial u}{\partial t}(x, t) dx$$

qui est égal d'après l'équation de la chaleur à

$$\kappa \int_{\mathbb{R}} e^{-i\xi x} \frac{\partial^2 u}{\partial x^2}(x, t) dx = \kappa \widehat{u_t''}(\xi) = -\kappa \xi^2 w_t(\xi).$$

On a ainsi transformé l'équation de la chaleur en l'équation différentielle ordinaire (on raisonne pour ξ fixé)

$$\frac{\partial w}{\partial t}(\xi, t) = -\kappa \xi^2 w(\xi, t).$$

Cette équation différentielle se résout facilement, pour chaque ξ fixé : on obtient

$$w_t(\xi) = w(\xi, t) = e^{-\kappa \xi^2 t} w(\xi, 0) = e^{-\kappa \xi^2 t} \widehat{v}(\xi).$$

On voit apparaître des densités gaussiennes (et leurs transformées de Fourier) : la fonction w_t , qui est la transformée de Fourier de u_t , apparaît comme le produit de la transformée de Fourier $\xi \rightarrow e^{-\kappa \xi^2 t}$ d'une certaine densité gaussienne g_t avec la transformée de Fourier de v . Par l'injectivité de Fourier, on en déduit que la fonction $x \rightarrow u(x, t)$ est la convolution de la donnée initiale v avec la densité gaussienne g_t dont la transformée de Fourier est $\widehat{g}_t(\xi) = e^{-\kappa \xi^2 t}$, c'est-à-dire

$$g_t(x) = \frac{e^{-x^2/(4\kappa t)}}{\sqrt{4\pi \kappa t}},$$

et pour tout $t > 0$,

$$u(x, t) = (v * g_t)(x) = \int_{\mathbb{R}} v(x - y) \frac{e^{-y^2/(4\kappa t)}}{\sqrt{4\pi \kappa t}} dy.$$

Fourier multi-dimensionnel

Si $x = (x_1, \dots, x_d)$ et $y = (y_1, \dots, y_d)$ sont deux éléments de \mathbb{R}^d , on introduit leur produit scalaire

$$x \cdot y = \sum_{j=1}^d x_j y_j$$

et si f est intégrable sur \mathbb{R}^d on pose

$$\forall y \in \mathbb{R}^d, \quad \widehat{f}(y) = \int_{\mathbb{R}^d} e^{-i x \cdot y} f(x) dx.$$

Dans le cas d'une fonction « décomposée » de la forme

$$f(x_1, x_2, \dots, x_d) = f_1(x_1) f_2(x_2) \dots f_d(x_d),$$

le théorème de Fubini donne

$$\widehat{f}(y_1, y_2, \dots, y_d) = \widehat{f}_1(y_1) \widehat{f}_2(y_2) \dots \widehat{f}_d(y_d).$$

Ainsi, la densité gaussienne 2-dimensionnelle $g(x_1, x_2) = e^{-x_1^2/2 - x_2^2/2}$ admet pour transformée de Fourier

$$(y_1, y_2) \rightarrow \sqrt{2\pi} e^{-y_1^2/2} \sqrt{2\pi} e^{-y_2^2/2} = 2\pi e^{-y_1^2/2 - y_2^2/2} = 2\pi g(y_1, y_2).$$

Fourier dans $L^2(\mathbb{R}^d)$

On peut définir la transformation de Fourier \mathcal{F} sur l'espace $L^2(\mathbb{R}^d)$, et l'égalité de Parseval devient, pour toute fonction $f \in L^2(\mathbb{R}^d)$

$$\int_{\mathbb{R}^d} |\mathcal{F}f(y)|^2 dy = (2\pi)^d \int_{\mathbb{R}^d} |f(x)|^2 dx.$$

Indiquons un chemin possible pour la preuve, lorsque $d = 2$ pour faire simple. Pour une fonction v sur \mathbb{R}^2 de la forme $v(x_1, x_2) = u_1(x_1)u_2(x_2)$, avec $u_1, u_2 \in X$, on aura la formule d'inversion

$$\begin{aligned} \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \widehat{v}(y_1, y_2) dy_1 dy_2 &= \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \widehat{u}_1(y_1) \widehat{u}_2(y_2) e^{i(x_1 y_1 + x_2 y_2)} dy_1 dy_2 \\ &= \left(\frac{1}{2\pi} \int_{\mathbb{R}} \widehat{u}_1(y_1) e^{i x_1 y_1} dy_1 \right) \left(\frac{1}{2\pi} \int_{\mathbb{R}} \widehat{u}_2(y_2) e^{i x_2 y_2} dy_2 \right) = u_1(x_1) u_2(x_2) = v(x_1, x_2); \end{aligned}$$

cette formule d'inversion s'étend évidemment aux combinaisons linéaires w de fonctions de cette forme. On en déduira la relation de Parseval pour w ,

$$(9) \quad \int_{\mathbb{R}^2} |\widehat{w}(y)|^2 dy = (2\pi)^2 \int_{\mathbb{R}^2} |w(x)|^2 dx,$$

exactement comme on était passé de la formule d'inversion du théorème 1.1.9 à l'énoncé de la proposition 1.1.10.

Ensuite, considérons l'indicatrice χ d'un rectangle $Q = [a_1, b_1] \times [a_2, b_2]$ dans \mathbb{R}^2 ; on peut écrire $\chi(x_1, x_2) = \chi_1(x_1)\chi_2(x_2)$, où χ_j est l'indicatrice du segment $[a_j, b_j]$, $j = 1, 2$. D'après le lemme 1.3.2, on peut approcher χ_j dans $L^2(\mathbb{R})$ par une fonction $u_j \in X$ nulle hors de $[a_j, b_j]$, et il est facile de voir que la fonction χ est alors approchée dans $L^2(\mathbb{R}^2)$ par la fonction produit $v(x_1, x_2) = u_1(x_1)u_2(x_2)$, qui est nulle hors de Q . On peut donc approcher les fonctions en escalier sur \mathbb{R}^2 , c'est-à-dire les combinaisons linéaires d'indicatrices de rectangles, par des fonctions telles que w , qui vérifient la relation de Parseval (9). Pour continuer, on vérifie que toute fonction $f \in L^2(\mathbb{R}^2)$ à support dans un rectangle compact Q peut être approchée par des combinaisons linéaires telles que w , nulles hors de Q , et on en déduit la relation de Parseval pour f . Pour finir, on passe à $L^2(\mathbb{R}^2)$ comme on l'a fait en dimension 1.

Notes du chapitre 1

(a) Un élément f de $L^\infty(\mathbb{R})$ est une classe de fonctions mesurables qui admet des représentants bornés. La quantité $\|f\|_\infty$ est le plus petit nombre M tel que f admette un représentant f_1 (une vraie fonction mesurable) qui vérifie $|f_1(x)| \leq M$ pour tout $x \in \mathbb{R}$. Si on considère un représentant quelconque f_2 de la classe f , l'ensemble

$$\{x \in \mathbb{R} : |f_2(x)| > m\}$$

dépend du représentant, mais sa mesure (de Lebesgue) est la même pour tous les représentants de f . On peut dire que M est le min des nombres réels m tels que la mesure de l'ensemble précédent soit nulle (pour la mesure de Lebesgue sur \mathbb{R}). À partir d'un représentant quelconque f_2 , on peut obtenir f_1 borné par M en posant $f_1(x) = f_2(x)$ quand $|f_2(x)| \leq M$ et $f_1(x) = 0$ (par exemple) sinon.

(b) Pour chacun des intervalles ouverts $]a_{j-1}, a_j[$, la fonction f' se prolonge par continuité à l'intervalle fermé correspondant, donc f' est bornée. Ici il est facile de montrer que f' est mesurable, car d'après l'hypothèse f' est définie et continue sur $\mathbb{R} \setminus \{a_0, a_1, \dots, a_N\}$, donc l'image inverse de tout ouvert de \mathbb{R} est ouverte, donc mesurable. De façon générale, une fonction dérivée f' définie presque partout est mesurable, car elle est limite simple presque partout d'une suite de fonctions continues, donc mesurables : en effet, la suite des fonctions continues

$$x \rightarrow n(f(x + 1/n) - f(x))$$

tend vers la valeur $f'(x)$ en tout point x où la dérivée existe.

(c) On fera les calculs dans le cas $\tau < b - a$; posons

$$I(x) = \tau(\mathbf{1}_{[a,b]} * K_\tau)(x) = \int_{\mathbb{R}} \mathbf{1}_{[a,b]}(x-t) \mathbf{1}_{[0,\tau]}(t) dt.$$

La quantité à intégrer

$$q(t, x) = \mathbf{1}_{[a,b]}(x-t) \mathbf{1}_{[0,\tau]}(t)$$

est non nulle si et seulement si $a \leq x-t \leq b$ et $0 \leq t \leq \tau$, ce qui se résume en $t \leq \beta(x) = \min(x-a, \tau)$ et $t \geq \alpha(x) = \max(x-b, 0)$. Quand $\beta(x) < \alpha(x)$, la fonction $t \rightarrow q(t, x)$ est identiquement nulle, donc $I(x) = 0$; sinon, la valeur $I(x)$ de l'intégrale est la longueur $\beta(x) - \alpha(x)$ de l'intervalle $[\alpha(x), \beta(x)]$.

L'expression de $\beta(x)$ « change de définition » quand $x = a + \tau$, et celle de $\alpha(x)$ change quand $x = b$; résumons la situation dans un tableau, selon l'appartenance de x à l'un des trois intervalles $I_1 =]-\infty, a + \tau[$, $I_2 =]a + \tau, b[$ ou $I_3 =]b, +\infty[$,

	I ₁	I ₂	I ₃
$\beta(x)$: $x - a$;	τ ;	τ
$\alpha(x)$: 0 ;	0 ;	$x - b$
$\beta(x) - \alpha(x)$: $x - a$;	τ ;	$b + \tau - x$

On obtient deux nouveaux points remarquables $x = a$ et $x = b + \tau$, où la différence $\beta(x) - \alpha(x)$ change de signe. Comme $I(x) = \max(0, \beta(x) - \alpha(x))$, on obtient finalement le tableau suivant,

	I' ₁	I'' ₁	I ₂	I' ₃	I'' ₃
I(x)	: 0 ;	$x - a$;	τ ;	$b + \tau - x$;	0

où on a subdivisé les intervalles I_1 et I_3 en utilisant les points a et $b + \tau$. Finalement, il y a quatre points de « changement de formule », qui sont $a < a + \tau < b < b + \tau$ et cinq intervalles successifs dans lesquels $I(x)$ prend les différentes expressions $I(x) = 0$, $I(x) = x - a$, $I(x) = \tau$, $I(x) = b + \tau - x$ et $I(x) = 0$.

(d) La formule exacte, qui est sans importance pour nous, serait

$$\begin{aligned} \|h - \chi\|_p^p &= \int_u^{u+\varepsilon} (1 - (x-u)/\varepsilon)^p dx + \int_{v-\varepsilon}^v (1 + (x-v)/\varepsilon)^p dx \\ &= 2 \int_0^\varepsilon (t/\varepsilon)^p dt = 2\varepsilon \int_0^1 s^p ds = \frac{2\varepsilon}{p+1} \leq \varepsilon. \end{aligned}$$

(e) Au lieu de tout faire à la main, on aurait pu appliquer un théorème du cours de topologie : l'espace E des fonctions à support borné est dense dans $L^2(\mathbb{R})$ et l'application $f \rightarrow \widehat{f}$ est uniformément continue de E dans l'espace complet $L^2(\mathbb{R})$. On peut donc la prolonger à $L^2(\mathbb{R})$, par le théorème de prolongement des applications uniformément continues.

(f) Pour tout $a > 0$, on trouve par intégration par parties

$$\int_0^a \frac{\sin x}{x} dx = \left[\frac{1 - \cos x}{x} \right]_0^a + \int_0^a \frac{1 - \cos x}{x^2} dx = \frac{1 - \cos a}{a} + \int_0^a \frac{1 - \cos x}{x^2} dx.$$

Quand a tend vers l'infini, le premier terme tend vers 0 et l'intégrale en $1/x^2$ est absolument convergente à l'infini, donc

$$\lim_{a \rightarrow +\infty} \int_0^a \frac{\sin x}{x} dx = \int_0^{+\infty} \frac{1 - \cos x}{x^2} dx.$$

(g) Si f et g sont mesurables positives on a toujours

$$\int fg \leq \left(\int f^2 \right)^{1/2} \left(\int g^2 \right)^{1/2}$$

à condition d'admettre la valeur $+\infty$ et d'interpréter $(+\infty)^{1/2}$ comme $+\infty$. En effet, si le produit dans le membre de droite est infini, l'inégalité est évidente ; si le produit est fini non nul, alors f et g sont de carré intégrable et c'est Cauchy-Schwarz ; si le produit est nul, l'un des deux facteurs est nul ; si $\int f^2 = 0$ on déduit que f est nulle presque partout, donc $fg = 0$ presque partout, $\int fg = 0$ et l'inégalité est vérifiée dans ce cas ; c'est pareil si $\int g^2 = 0$.

(h) Si une suite (g_n) tend vers g dans $L^2(\mathbb{R})$, alors pour tous $a < b$

$$\int_a^b g(t) dt = \lim_n \int_a^b g_n(t) dt.$$

En effet par Cauchy-Schwarz

$$\begin{aligned} \left| \int_a^b (g(t) - g_n(t)) dt \right| &= \left| \int_{\mathbb{R}} \mathbf{1}_{[a,b]}(t) (g(t) - g_n(t)) dt \right| \\ &\leq \left(\int_a^b 1^2 dt \right)^{1/2} \left(\int_a^b |g(t) - g_n(t)|^2 dt \right)^{1/2} \rightarrow 0. \end{aligned}$$

2. Espaces de Hilbert

2.1. Produit scalaire, orthogonalité, bases hilbertiennes

On considère un espace vectoriel E sur $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} . On appelle *forme hermitienne* sur l'espace E une application $(x, y) \in E \times E \rightarrow \varphi(x, y) \in \mathbb{K}$ telle que

- l'application $x \in E \rightarrow \varphi(x, y)$ est \mathbb{K} -linéaire pour tout $y \in E$ fixé ;
- pour tous $x, y \in E$, on a $\varphi(y, x) = \overline{\varphi(x, y)}$;

Lorsque $\mathbb{K} = \mathbb{R}$, les barres de conjugaison sont inutiles, et on obtient dans ce cas une *forme bilinéaire symétrique* ; toutefois, considérer les deux cas \mathbb{R} et \mathbb{C} ensemble nous évitera de devoir tout répéter.

On remarque que

$$(1) \quad \varphi(x, x) = \overline{\varphi(x, x)}$$

est réel pour tout vecteur $x \in E$.

On appelle *produit scalaire* sur E une application $(x, y) \in E \times E \rightarrow \langle x, y \rangle \in \mathbb{K}$ telle que

- l'application $x \in E \rightarrow \langle x, y \rangle$ est \mathbb{K} -linéaire pour tout $y \in E$ fixé ;
- pour tous $x, y \in E$, on a $\langle y, x \rangle = \overline{\langle x, y \rangle}$;
- le nombre $\langle x, x \rangle$, qui est réel d'après l'équation (1), vérifie $\langle x, x \rangle \geq 0$ pour tout $x \in E$, et de plus $\langle x, x \rangle > 0$ pour tout vecteur x non nul.

Autrement dit, un produit scalaire est une forme hermitienne qui vérifie $\varphi(x, x) > 0$ pour tout $x \neq 0$. On dira *semi-produit scalaire* ^(a) lorsqu'on supposera seulement que $\langle x, x \rangle \geq 0$ pour tout vecteur x , mais en permettant que $\langle x, x \rangle$ puisse être nul pour certains vecteurs $x \in E$ non nuls.

Premières propriétés des formes hermitiennes et des semi-produits scalaires

On suppose donnée une forme hermitienne φ sur E . Pour tout vecteur $x \in E$ fixé, l'application $y \rightarrow \varphi(x, y)$ est *additive*,

$$\varphi(x, y_1 + y_2) = \overline{\varphi(y_1 + y_2, x)} = \overline{\varphi(y_1, x) + \varphi(y_2, x)} = \varphi(x, y_1) + \varphi(x, y_2).$$

Si x, y sont deux vecteurs de E et λ un scalaire, on a ^(b)

$$\varphi(\lambda x, y) = \lambda \varphi(x, y) \quad \text{mais} \quad \varphi(x, \lambda y) = \overline{\varphi(\lambda y, x)} = \overline{\lambda} \varphi(x, y).$$

L'application $y \rightarrow \varphi(x, y)$ n'est donc pas linéaire dans le cas complexe, car l'image du vecteur λy est $\overline{\lambda} \varphi(x, y)$ (et pas $\lambda \varphi(x, y)$). On dit qu'une telle application f du \mathbb{K} -espace vectoriel E dans un autre \mathbb{K} -espace vectoriel, qui vérifie $f(\lambda v + w) = \overline{\lambda} f(v) + f(w)$ pour tout scalaire λ et tous vecteurs $v, w \in E$, est une *application antilinéaire*. Ainsi, pour tout $x \in E$ fixé,

l'application $y \rightarrow \varphi(x, y)$ est antilinéaire de E dans \mathbb{K} .

Lorsque $\mathbb{K} = \mathbb{R}$, toutes les barres de conjugaison sont inutiles, et il n'y a pas de différence entre linéaire et antilinéaire. Si x et y sont deux vecteurs de E , on obtient que

$$(2) \quad \varphi(x + y, x + y) = \varphi(x, x) + 2 \operatorname{Re} \varphi(x, y) + \varphi(y, y).$$

En remplaçant y par $-y$, on aura $\varphi(x - y, x - y) = \varphi(x, x) - 2 \operatorname{Re} \varphi(x, y) + \varphi(y, y)$ et en additionnant les deux on trouve

$$\varphi(x + y, x + y) + \varphi(x - y, x - y) = 2\varphi(x, x) + 2\varphi(y, y).$$

Dans le cas d'un produit scalaire, la relation précédente est appelée *identité du parallélogramme*^(c) ou bien *relation de la médiane*

$$(3) \quad \langle x + y, x + y \rangle + \langle x - y, x - y \rangle = 2\langle x, x \rangle + 2\langle y, y \rangle.$$

En faisant la différence au lieu de la somme on obtient la formule

$$(4) \quad 4 \operatorname{Re} \varphi(x, y) = \varphi(x + y, x + y) - \varphi(x - y, x - y);$$

cette formule, dite *formule de polarisation*, montre qu'on peut reconstituer $\varphi(x, y)$ pour de deux vecteurs quelconques à partir de tous les carrés $\varphi(v, v)$, pour v variant dans l'espace vectoriel E . Dans le cas complexe, on pourra récupérer $\varphi(x, y)$ lui-même en combinant^(d) la formule précédente avec celle qu'on peut écrire pour x et iy .

Proposition 2.1.1 (Cauchy-Schwarz). *Soit E un \mathbb{K} -espace vectoriel muni d'un semi-produit scalaire; pour tous les vecteurs x, y de E on a*

$$|\langle x, y \rangle| \leq \sqrt{\langle x, x \rangle} \sqrt{\langle y, y \rangle}.$$

Preuve. — Écrivons le nombre complexe $\langle x, y \rangle$ sous forme polaire $\langle x, y \rangle = |\langle x, y \rangle| e^{i\theta}$, pour un certain nombre réel θ , de sorte que $\langle e^{-i\theta} x, y \rangle = e^{-i\theta} \langle x, y \rangle = |\langle x, y \rangle|$. Choisissons deux nombres réels $\lambda > \sqrt{\langle y, y \rangle}$ et $\mu > \sqrt{\langle x, x \rangle}$. On écrit pour le vecteur $z = \lambda e^{-i\theta} x - \mu y$

$$\begin{aligned} 0 &\leq \langle z, z \rangle = \langle \lambda e^{-i\theta} x - \mu y, \lambda e^{-i\theta} x - \mu y \rangle \\ &= \lambda^2 \langle x, x \rangle - 2\lambda\mu \operatorname{Re} \langle e^{-i\theta} x, y \rangle + \mu^2 \langle y, y \rangle = \lambda^2 \langle x, x \rangle - 2\lambda\mu |\langle x, y \rangle| + \mu^2 \langle y, y \rangle. \end{aligned}$$

Il en résulte que

$$2\lambda\mu |\langle x, y \rangle| \leq \lambda^2 \langle x, x \rangle + \mu^2 \langle y, y \rangle < 2\lambda^2 \mu^2,$$

d'où $|\langle x, y \rangle| < \lambda\mu$ par simplification, puisque $\lambda\mu > 0$. Pour finir, on fait tendre λ vers $\sqrt{\langle y, y \rangle}$ et μ vers $\sqrt{\langle x, x \rangle}$.

Semi-norme déduite d'un semi-produit scalaire

Soit E un \mathbb{K} -espace vectoriel muni d'un semi-produit scalaire; l'application

$$x \in E \rightarrow \|x\| = \sqrt{\langle x, x \rangle}$$

est une *semi-norme* sur E , c'est-à-dire qu'elle vérifie

$$\|x\| \geq 0, \quad \|\lambda x\| = |\lambda| \|x\| \quad \text{et} \quad \|x + y\| \leq \|x\| + \|y\|$$

pour tous $x, y \in E$ et $\lambda \in \mathbb{K}$. Le dernier point (*l'inégalité triangulaire*) résulte de Cauchy-Schwarz,

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle = \langle x, x \rangle + 2 \operatorname{Re} \langle x, y \rangle + \langle y, y \rangle \leq \langle x, x \rangle + 2|\langle x, y \rangle| + \langle y, y \rangle \\ &\leq \langle x, x \rangle + 2\sqrt{\langle x, x \rangle} \sqrt{\langle y, y \rangle} + \langle y, y \rangle = (\sqrt{\langle x, x \rangle} + \sqrt{\langle y, y \rangle})^2 = (\|x\| + \|y\|)^2. \end{aligned}$$

Lorsque E est muni d'un produit scalaire, l'hypothèse $\|x\| = 0$ implique $\langle x, x \rangle = 0$, donc $x = 0$ par définition d'un produit scalaire, et la semi-norme est en fait une *norme* sur l'espace vectoriel E .

Exemples. Sur \mathbb{C}^n on définit le produit scalaire

$$\langle z, w \rangle = \sum_{i=1}^n z_i \overline{w_i}$$

où $z = (z_1, \dots, z_n)$ et $w = (w_1, \dots, w_n)$, avec $z_i, w_j \in \mathbb{C}$.

Sur l'espace $E = \mathcal{L}^2(\Omega, \mu)$ on peut définir le semi-produit scalaire

$$\langle f, g \rangle = \int_{\Omega} f(\omega) \overline{g(\omega)} d\mu(\omega);$$

il lui correspond la semi-norme

$$\|f\|_2 = \left(\int_{\Omega} |f(\omega)|^2 d\mu(\omega) \right)^{1/2}.$$

Si f est négligeable sans être la fonction nulle, on aura $\|f\|_2 = 0$ sans que le vecteur f soit le vecteur nul de E : cette semi-norme n'est en général pas une norme.

L'espace vectoriel $E = \ell^2(\mathbb{N})$ formé des toutes les suites scalaires $\mathbf{x} = (x_n)_{n \geq 0}$ telles que $\sum_n |x_n|^2 < +\infty$ peut être muni du produit scalaire

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{n=0}^{+\infty} x_n \overline{y_n};$$

il lui correspond la norme

$$\|\mathbf{x}\|_2 = \left(\sum_{n=0}^{+\infty} |x_n|^2 \right)^{1/2}.$$

Soit E un \mathbb{K} -espace vectoriel muni d'un semi-produit scalaire; on dit que deux vecteurs $u, v \in E$ sont *orthogonaux* quand $\langle u, v \rangle = 0$, et on note $x \perp y$; on voit que cette relation est symétrique : $\langle v, u \rangle = \overline{\langle u, v \rangle} = \overline{0} = 0$.

Remarque 2.1.2. Pour chaque vecteur $y \in E$ fixé, l'application $x \rightarrow \langle x, y \rangle$ est linéaire; il en résulte que l'ensemble

$$y^\perp = \{x \in E : \langle x, y \rangle = 0\}$$

des vecteurs orthogonaux à y est un sous-espace vectoriel : si y est orthogonal à des vecteurs a_1, \dots, a_n de E , alors y est aussi orthogonal à toutes les combinaisons linéaires de ces vecteurs.

Si A est un sous-ensemble de E , on dit que le vecteur x est *orthogonal* à A si x est orthogonal à tous les vecteurs $a \in A$. Ceci signifie que A est contenu dans le sous-espace vectoriel x^\perp , par conséquent $\text{Vect}(A)$, qui est *le plus petit* sous-espace vectoriel contenant A , est aussi contenu dans x^\perp . Il en résulte que x est orthogonal à l'espace vectoriel $\text{Vect}(A)$ engendré par A :

$$(x \perp A) \Rightarrow (x \perp \text{Vect } A).$$

Proposition 2.1.3 (Pythagore). Soit E un espace muni d'un semi-produit scalaire ; si u_1, \dots, u_n sont des vecteurs de l'espace E , deux à deux orthogonaux, on a pour la semi-norme associée

$$\left\| \sum_{j=1}^n u_j \right\|^2 = \sum_{j=1}^n \|u_j\|^2.$$

Preuve. — Si u et v sont orthogonaux,

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

Ensuite, on montre la proposition par récurrence sur $n \geq 2$: on pose $v = \sum_{j=1}^n u_j$, on remarque que u_{n+1} est orthogonal à v (linéarité du produit scalaire) et

$$\left\| \sum_{j=1}^{n+1} u_j \right\|^2 = \|v + u_{n+1}\|^2 = \|v\|^2 + \|u_{n+1}\|^2 = \left(\sum_{j=1}^n \|u_j\|^2 \right) + \|u_{n+1}\|^2.$$

Définition. Un espace de Hilbert est un \mathbb{K} -espace vectoriel E muni d'un produit scalaire, et tel que E soit complet pour la norme associée à ce produit scalaire, $\|x\| = \sqrt{\langle x, x \rangle}$. On peut récrire l'inégalité de Cauchy-Schwarz en utilisant la norme,

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

Exemple. Rappelons la définition de l'espace $L^2(\Omega, \mathcal{A}, \mu)$ (bien noter que la lettre «L» dans le symbole L^2 est droite maintenant). Un élément $\tilde{f} \in L^2(\Omega, \mu)$ n'est pas une fonction, mais une *classe de fonctions* : si f est une fonction mesurable sur (Ω, \mathcal{A}) à valeurs dans \mathbb{K} , on lui associe l'ensemble \tilde{f} de toutes les fonctions mesurables f_1 à valeurs dans \mathbb{K} telles que $f_1 = f$ μ -presque partout. Si l'un des éléments d'une classe \tilde{f} est intégrable, alors tous les autres éléments de la classe sont intégrables et ont la même intégrale, ce qui permet d'employer la notation $\int_{\Omega} \tilde{f} d\mu$ et de parler de *classe intégrable*. Les carrés f^2 des éléments d'une classe \tilde{f} sont dans une même classe, qu'on peut raisonnablement noter \tilde{f}^2 et appeler le carré de la classe \tilde{f} .

L'espace $L^2(\Omega, \mathcal{A}, \mu)$ est l'espace vectoriel des classes de carré intégrable. Le vecteur nul $\tilde{0}$ de cet espace vectoriel est la *classe nulle*, c'est-à-dire l'ensemble des fonctions μ -négligeables à valeurs dans \mathbb{K} . L'espace $E = L^2(\Omega, \mathcal{A}, \mu)$, muni du produit scalaire^(e)

$$\langle \tilde{f}, \tilde{g} \rangle = \int_{\Omega} f(\omega) \overline{g(\omega)} d\mu(\omega)$$

où f, g sont des représentants quelconques des classes $\tilde{f}, \tilde{g} \in L^2$, et muni de la norme $\|\tilde{f}\|_2 = \langle \tilde{f}, \tilde{f} \rangle^{1/2}$, est un espace de Hilbert. Dans la suite on ne mentionnera plus les classes ; on fera «comme si» \tilde{f} et \tilde{g} étaient des vraies fonctions, pour toute question ne dépendant pas du représentant choisi, comme le calcul des intégrales par exemple.

L'espace $\ell^2(\mathbb{N})$ est un cas particulier du précédent ; c'est donc un espace de Hilbert. On peut construire cet exemple sur d'autres ensembles d'indices que \mathbb{N} , en particulier on utilisera $\ell^2(\mathbb{Z})$ dans la théorie des séries de Fourier.

Corollaire. Soit E un espace de Hilbert ; des vecteurs (u_k) , $k = 1, \dots, n$, orthogonaux et non nuls sont linéairement indépendants.

Preuve. — En effet, si $\sum_{k=1}^n c_k u_k = 0$, on aura par la proposition 2.1.3

$$\left\| \sum_{k=1}^n c_k u_k \right\|^2 = \sum_{k=1}^n \|c_k u_k\|^2 = \sum_{k=1}^n |c_k|^2 \|u_k\|^2 = 0$$

donc $|c_k| \|u_k\| = 0$ pour tout k , et $c_k = 0$ puisque $\|u_k\| \neq 0$ pour $k = 1, \dots, n$. Ainsi, la seule combinaison linéaire de ces vecteurs qui soit nulle est celle dont tous les coefficients c_k sont nuls : le système de vecteurs est libre.

Si les vecteurs e_1, \dots, e_n sont *orthonormés*, c'est-à-dire (f) que $\langle e_k, e_\ell \rangle = \delta_{k,\ell}$ pour tous $k, \ell = 1, \dots, n$, on a

$$\left\| \sum_{k=1}^n c_k e_k \right\|^2 = \sum_{k=1}^n |c_k|^2.$$

De plus, si $F = \text{Vect}(e_1, \dots, e_n)$ et si z est un vecteur de F , les coordonnées de z dans la base (e_1, \dots, e_n) de F sont données par les produits scalaires $\langle z, e_k \rangle$, $k = 1, \dots, n$

$$z = \sum_{k=1}^n \langle z, e_k \rangle e_k, \quad \|z\|^2 = \sum_{k=1}^n |\langle z, e_k \rangle|^2.$$

En effet, si $z \in F$, il existe des coefficients c_k tels que $z = \sum_{k=1}^n c_k e_k$; pour tout indice $j = 1, \dots, n$ on a

$$\langle z, e_j \rangle = \sum_{k=1}^n c_k \langle e_k, e_j \rangle = c_j.$$

Formes linéaires continues

Soit H un espace de Hilbert ; pour tout vecteur $v \in H$ fixé, l'application

$$\ell_v : x \in H \rightarrow \langle x, v \rangle$$

est une forme linéaire continue sur H . En effet, par Cauchy-Schwarz, on a

$$|\ell_v(x)| = |\langle x, v \rangle| \leq \|v\| \|x\|.$$

Comme on le sait, ceci implique la continuité de ℓ , puisqu'alors $|\ell_v(x_1) - \ell_v(x_2)| = |\ell_v(x_1 - x_2)| \leq \|v\| \|x_1 - x_2\|$ pour $x_1, x_2 \in H$. L'orthogonal du vecteur v ,

$$v^\perp = \{x \in H : x \perp v\}$$

est un sous-espace vectoriel (noyau de ℓ_v), et il est fermé (continuité de ℓ_v).

Le *dual topologique* E' d'un espace normé E est l'espace vectoriel des formes linéaires continues sur E . Il est normé de la façon suivante : si ℓ est une forme linéaire continue sur E , on pose

$$\|\ell\| = \sup\{|\ell(x)| : x \in E, \|x\| \leq 1\}.$$

On a alors $|\ell(x)| \leq \|\ell\| \|x\|$ pour tout $x \in E$; en fait $\|\ell\|$ est le plus petit nombre C tel que l'on ait $|\ell(x)| \leq C \|x\|$ pour tout vecteur $x \in E$. Muni de cette norme, l'espace E' est complet (8).

Revenons au cas de l'espace de Hilbert H . On vient de voir par Cauchy-Schwarz que $|\ell_v(x)| \leq \|v\| \|x\|$ pour tout x , donc $\|\ell_v\| \leq \|v\|$; en appliquant ℓ_v au vecteur v lui-même, on obtient

$$\|v\|^2 = \langle v, v \rangle = \ell_v(v) \leq \|\ell_v\| \|v\|,$$

ce qui implique $\|v\| \leq \|\ell_v\|$ quand $v \neq 0$. On en déduit que $\|\ell_v\| = \|v\|$ pour tout vecteur $v \in H$. On montrera plus loin que toute forme linéaire continue sur un espace de Hilbert est de la forme ℓ_v , pour un certain vecteur $v \in H$ (théorème 2.3.8).

On a vu que si x est orthogonal à un sous-ensemble $A \subset E$, alors $\text{Vect}(A)$ est contenu dans le sous-espace vectoriel x^\perp , noyau de la forme linéaire ℓ_x ; puisque ℓ_x est continue, son noyau x^\perp est fermé, et s'il contient $\text{Vect}(A)$, il contient aussi son adhérence :

$$(5) \quad (x \perp A) \Rightarrow (x \perp \overline{\text{Vect}(A)}).$$

Projection orthogonale sur un sous-espace de dimension finie

Considérons dans un espace de Hilbert H un sous-espace vectoriel $F = \text{Vect}(e_1, \dots, e_n)$ engendré par une suite orthonormée finie e_1, \dots, e_n : pour tout vecteur $x \in H$, posons

$$P_F x = \sum_{k=1}^n \langle x, e_k \rangle e_k.$$

Lemme 2.1.4. *Pour tout $x \in H$, le point $P_F x$ est dans F et $x - P_F x$ est orthogonal à F . Le point $P_F x$ est le point de F le plus proche de x ; en particulier, $P_F y = y$ pour tout $y \in F$. De plus $\|P_F x\| \leq \|x\|$, c'est-à-dire que*

$$\forall x \in H, \quad \sum_{k=1}^n |\langle x, e_k \rangle|^2 \leq \|x\|^2.$$

L'application $x \rightarrow P_F x$ est linéaire continue de H dans H .

On dit que $P_F x$ est la *projection orthogonale* de x sur le sous-espace vectoriel F .

Preuve. — Posons $y = \sum_{k=1}^n \langle x, e_k \rangle e_k$; pour tout $j = 1, \dots, n$ on a

$$\langle y, e_j \rangle = \left\langle \sum_{k=1}^n \langle x, e_k \rangle e_k, e_j \right\rangle = \sum_{k=1}^n \langle x, e_k \rangle \langle e_k, e_j \rangle = \langle x, e_j \rangle.$$

On voit ainsi que $\langle e_j, x - y \rangle = 0$ pour tout $j = 1, \dots, n$; puisque $x - y$ est orthogonal aux vecteurs $(e_j)_{j=1}^n$, il est orthogonal à l'espace vectoriel engendré par ces vecteurs (remarque 2.1.2), ce qui montre que $x - y$ est orthogonal à $F = \text{Vect}(e_j : j = 1, \dots, n)$: on a bien vérifié que $y \in F$ et $x - y \perp F$. Si z est un vecteur de F quelconque, la différence $y - z$ est encore dans F , donc orthogonale à $x - y$ et par Pythagore

$$\|x - z\|^2 = \|(x - y) + (y - z)\|^2 = \|x - y\|^2 + \|y - z\|^2 \geq \|x - y\|^2.$$

En reprenant la ligne précédente avec $z = 0$,

$$\|x\|^2 = \|(x - y) + y\|^2 = \|x - y\|^2 + \|y\|^2 \geq \|y\|^2.$$

Il est clair sur les formules de définition que P_F est linéaire, et l'inégalité $\|P_F x\| \leq \|x\|$ implique la continuité de P_F .

Séries de vecteurs dans un espace vectoriel normé

Si $(u_k)_{k \geq 0}$ est une suite de vecteurs d'un espace normé E , on peut définir les *sommes partielles*

$$U_n = \sum_{k=0}^n u_k \in E;$$

par définition, la série de vecteurs $\sum u_k$ est dite *convergente dans* E quand il existe un vecteur $s \in E$ tel que

$$\lim_n \|s - U_n\|_E = 0,$$

c'est-à-dire que $s = \lim_n U_n$ dans E . Ce vecteur s (qui est unique) est appelé *la somme de la série*. On pose alors

$$\sum_{k=0}^{+\infty} u_k = s = \lim_n U_n.$$

Exemple. Soit $f \in L^p([0, \infty])$, $1 \leq p < +\infty$; pour chaque entier $k \geq 0$ définissons la fonction $u_k = \mathbf{1}_{[k, k+1[} f$; la série $\sum u_k$ converge dans $L^p([0, \infty])$ et

$$f = \sum_{k=0}^{+\infty} u_k.$$

Si T est une application linéaire continue de E dans un espace normé F , on déduit de la linéarité que l'image TU_n de la somme partielle U_n est égale à la somme partielle de la série $\sum Tu_k$ des images, et comme T est continue, TU_n converge dans F vers Ts ; ainsi :

si la série de vecteurs $\sum u_k$ converge dans E et si T est linéaire continue de E dans un espace normé F , la série $\sum Tu_k$ converge dans F et

$$T\left(\sum_{k=0}^{+\infty} u_k\right) = \sum_{k=0}^{+\infty} Tu_k.$$

Proposition 2.1.5. *Soit $(u_k)_{k \geq 0}$ une suite de vecteurs deux à deux orthogonaux dans un espace de Hilbert H ; la série $\sum u_k$ converge dans H si et seulement si $\sum \|u_k\|^2 < +\infty$, et dans ce cas*

$$\left\| \sum_{k=0}^{+\infty} u_k \right\|^2 = \sum_{k=0}^{+\infty} \|u_k\|^2.$$

Preuve. — Posons pour tout entier $n \geq 0$

$$U_n = u_0 + \cdots + u_n, \quad V_n = \|U_n\|^2 = \|u_0\|^2 + \cdots + \|u_n\|^2.$$

Puisque H est complet, la série de vecteurs $\sum u_k$ converge dans H si et seulement si la suite des sommes partielles (U_n) est de Cauchy dans H . Pour tous les entiers $m < n$ on a $U_n - U_m = u_{m+1} + \cdots + u_n$, donc par Pythagore

$$\|U_n - U_m\|^2 = \sum_{k=m+1}^n \|u_k\|^2 = V_n - V_m.$$

Il en résulte que la suite de vecteurs (U_n) est de Cauchy si et seulement si la suite numérique (V_n) est de Cauchy, c'est-à-dire si et seulement si $\sum \|u_k\|^2 < +\infty$. Dans le cas où la série converge, on obtient

$$\left\| \sum_{k=0}^{+\infty} u_k \right\|^2 = \left\| \lim_n U_n \right\|^2 = \lim_n \|U_n\|^2 = \lim_n \sum_{k=0}^n \|u_k\|^2 = \sum_{k=0}^{+\infty} \|u_k\|^2.$$

Le sous-espace vectoriel $\text{Vect}(v_i : i \in I)$ engendré par une famille $(v_i)_{i \in I}$ de vecteurs de H est l'ensemble de toutes les combinaisons linéaires de ces vecteurs ; par définition, une combinaison linéaire $w \in \text{Vect}(v_i : i \in I)$ fait intervenir un ensemble fini $J \subset I$ et des coefficients scalaires $(c_j)_{j \in J}$,

$$w = \sum_{j \in J} c_j v_j.$$

On voit facilement que $\text{Vect}(v_i : i \in I)$ est le plus petit sous-espace vectoriel de H contenant tous les vecteurs v_i . Si $I = \mathbb{N}$, on voit que $\text{Vect}(v_n : n \geq 0)$ est la réunion des espaces de dimension finie $F_N = \text{Vect}(v_k : 0 \leq k \leq N)$, lorsque N varie dans \mathbb{N} .

On peut montrer facilement que l'adhérence d'un sous-espace vectoriel V est encore un sous-espace vectoriel. Le sous-espace vectoriel fermé engendré par la famille $(v_i)_{i \in I}$ de vecteurs de H est l'adhérence de $\text{Vect}(v_i : i \in I)$: c'est le plus petit sous-espace vectoriel fermé contenant tous les vecteurs v_i , $i \in I$.

Proposition 2.1.6. Soit $(e_k)_{k \geq 0}$ une suite orthonormée infinie de vecteurs dans un espace de Hilbert H ; pour tout vecteur $x \in H$, la série de vecteurs

$$(*) \quad \sum_{k=0}^{+\infty} \langle x, e_k \rangle e_k$$

est convergente dans H . Désignons par F le sous-espace vectoriel fermé engendré par la suite $(e_k)_{k \geq 0}$; la somme $y = \sum_{k=0}^{+\infty} \langle x, e_k \rangle e_k$ de la série précédente est un vecteur de F , et $x - y$ est orthogonal à F . On dit encore que ce vecteur y est la projection orthogonale de x sur F .

De plus, tout vecteur z du sous-espace vectoriel F peut s'écrire

$$z = \sum_{k=0}^{+\infty} \langle z, e_k \rangle e_k.$$

Preuve. — Pour tout entier $n \geq 0$, posons $F_n = \text{Vect}(e_0, \dots, e_n)$; on sait par le lemme 2.1.4 que le vecteur $y_n = \sum_{k=0}^n \langle x, e_k \rangle e_k \in F$ est la projection orthogonale de x sur F_n , et que

$$\sum_{k=0}^n |\langle x, e_k \rangle|^2 = \|y_n\|^2 \leq \|x\|^2.$$

Il en résulte que la série numérique $\sum |\langle x, e_k \rangle|^2$ converge, ce qui entraîne par la proposition 2.1.5 la convergence de la série $(*)$ de vecteurs orthogonaux ; la suite (y_n) des sommes partielles converge vers la somme de la série y ; ce vecteur limite y est dans F , puisque F est fermé. Pour $m \leq n$ on a $e_m \in F_n$ donc $\langle x - y_n, e_m \rangle = 0$. Le produit scalaire avec e_m étant continu, on obtient

$$\langle x - y, e_m \rangle = \lim_n \langle x - y_n, e_m \rangle = 0.$$

Ceci montre que $x - y$ est orthogonal à tous les vecteurs e_m ; comme l'orthogonal $(x - y)^\perp$ est un sous-espace vectoriel, il contient d'abord toutes les combinaisons linéaires des e_m , et comme il est fermé, il contient aussi les limites de combinaisons linéaires, c'est-à-dire finalement tous les vecteurs de F . Ainsi $F \subset (x - y)^\perp$, ce qui signifie que $x - y \perp F$.

Prenons maintenant $z \in F$ et considérons le vecteur $y = \sum_{k=0}^{+\infty} \langle z, e_k \rangle e_k$; d'après ce qui précède, on a $y \in F$ et $z - y \perp F$; mais puisque $z - y \in F$, il en résulte que

$$\langle z - y, z - y \rangle = 0,$$

par conséquent $z = y = \sum_{k=0}^{+\infty} \langle z, e_k \rangle e_k$, comme annoncé.

Exemple 2.1.7. Les fonctions $e_n(t) = e^{int}$, $n \in \mathbb{Z}$, sont un exemple (très important) de suite orthonormée dans l'espace $H = L^2([0, 2\pi], dt/(2\pi))$. En effet, on a pour $m \neq n$

$$\langle e_m, e_n \rangle = \int_0^{2\pi} e_m(t) \overline{e_n(t)} \frac{dt}{2\pi} = \int_0^{2\pi} e^{i(m-n)t} \frac{dt}{2\pi} = \left[\frac{e^{i(m-n)t}}{i(m-n)} \right]_{t=0}^{2\pi}$$

qui est nul parce que ces fonctions sont 2π -périodiques. Pour toute fonction $f \in H$, les deux séries de vecteurs

$$\sum_{n=0}^{+\infty} \langle f, e_n \rangle e_n, \quad \sum_{n=1}^{+\infty} \langle f, e_{-n} \rangle e_{-n}$$

convergent dans H . On montrera plus loin que

$$f = \sum_{n \in \mathbb{Z}} \langle f, e_n \rangle e_n := \sum_{n=0}^{+\infty} \langle f, e_n \rangle e_n + \sum_{n=1}^{+\infty} \langle f, e_{-n} \rangle e_{-n}.$$

Définition. Une *base hilbertienne* d'un espace de Hilbert H est une famille de vecteurs $(e_i)_{i \in I}$ indexée par un ensemble I , et telle que

- les vecteurs sont de norme un, et deux à deux orthogonaux ;
- l'espace vectoriel engendré $\text{Vect}(e_i, i \in I)$ est dense dans H .

On ne travaillera pas ici avec les espaces de Hilbert qui ont une base hilbertienne indexée par un ensemble I non dénombrable. Mentionnons que *tout espace de Hilbert admet des bases hilbertiennes*, mais nous ne le prouverons que dans le cas d'un espace de Hilbert *séparable* ; dans ce cas, la base peut être indexée par un ensemble I fini ou dénombrable. Si I est infini dénombrable, si $(e_i)_{i \in I}$ est une base hilbertienne de H et si $(i_n)_{n \geq 0}$ est une énumération quelconque des éléments de I , on a pour tout $x \in H$

$$x = \sum_{n=0}^{+\infty} \langle x, e_{i_n} \rangle e_{i_n}, \quad \|x\|^2 = \sum_{n=0}^{+\infty} |\langle x, e_{i_n} \rangle|^2.$$

Pour la première égalité, il suffit d'appliquer la proposition 2.1.6, puisque le sous-espace fermé engendré par la suite (e_{i_n}) est ici égal à l'espace H tout entier. La deuxième égalité résulte de la proposition 2.1.5.

L'exemple le plus simple de base hilbertienne est la *base canonique* de l'espace $E = \ell^2(\mathbb{N})$: pour chaque entier $n \geq 0$ désignons par \mathbf{e}_n le vecteur de E dont toutes les composantes sont nulles sauf la n ième, qui vaut 1,

$$\mathbf{e}_n = (\delta_{n,k})_{k \geq 0}.$$

Expression du produit scalaire dans une base hilbertienne infinie dénombrable

Si l'espace de Hilbert H admet une base hilbertienne infinie dénombrable $(e_n)_{n \geq 0}$, si $x, y \in H$, le vecteur x est limite dans H de la suite

$$x_N = \sum_{k=0}^N \langle x, e_k \rangle e_k ;$$

puisque la forme linéaire ℓ_y est continue sur H , on obtient

$$\langle x, y \rangle = \ell_y(x) = \lim_N \langle x_N, y \rangle = \lim_N \sum_{k=0}^N \langle x, e_k \rangle \langle e_k, y \rangle = \lim_N \sum_{k=0}^N \langle x, e_k \rangle \overline{\langle y, e_k \rangle}$$

c'est-à-dire que

$$(S) \quad \langle x, y \rangle = \sum_{k=0}^{+\infty} \langle x, e_k \rangle \overline{\langle y, e_k \rangle}.$$

Exemple de base hilbertienne : le système de Haar

Pour chaque intervalle borné $I = [a, b]$ posons $m = (a + b)/2$, puis $I_- = [a, m[$ et $I_+ = [m, b]$; on décrit une famille infinie d'intervalles par récurrence : la famille \mathcal{G}_0 est formée du seul intervalle $[0, 1[$; si \mathcal{G}_n est définie, la famille \mathcal{G}_{n+1} est formée de tous les intervalles I_-, I_+ lorsque I varie dans \mathcal{G}_n .

On voit que \mathcal{G}_n est formée de 2^n intervalles disjoints qui recouvrent $[0, 1[$. On considère la famille de fonctions formée de $h_\emptyset = 1$, puis de toutes les fonctions

$$h_I = |I|^{-1/2} (\mathbf{1}_{I_-} - \mathbf{1}_{I_+})$$

où I varie dans toutes les familles \mathcal{G}_n , $n \geq 0$ et $|I|$ dénote la longueur de l'intervalle I . Cette famille de fonctions est une base hilbertienne de l'espace de Hilbert $L^2(0, 1)$. À cet effet, on vérifie d'abord que ces fonctions (h_I) sont deux à deux orthogonales (exercice). Ensuite, l'espace

$$V_{n+1} = \text{Vect}(\mathbf{1}_I : I \in \mathcal{G}_{n+1})$$

est de dimension 2^{n+1} , et il contient les fonctions h_\emptyset et les h_I , pour I dans la réunion des \mathcal{G}_k , $0 \leq k \leq n$, qui sont indépendantes car orthogonales; le nombre de ces fonctions est

$$1 + 1 + 2 + 4 + \dots + 2^n = 2^{n+1},$$

donc elles forment une base de V_{n+1} , et il en résulte que

$$V_{n+1} = \text{Vect}(h_\emptyset, h_I : I \in \mathcal{G}_k, 0 \leq k \leq n).$$

On voit donc que l'espace engendré par la famille de toutes les fonctions h_I contient tous les espaces V_n ; il est clair (**h**) que toute fonction en escalier peut être approchée dans $L^2([0, 1])$ par une fonction de la réunion des V_n . Il en résulte que la famille (h_I) est une base hilbertienne de $L^2([0, 1])$.

Gram-Schmidt et bases hilbertiennes

Le procédé de Gram-Schmidt est utile en dimension finie comme en dimension infinie. Dans la proposition qui suit, on va utiliser une convention de notation qui permet de traiter les deux cas en même temps, au prix d'une petite perte de lisibilité.

Proposition 2.1.8. *Désignons par N un entier ≥ 0 fini ou bien $N = +\infty$. Si on a une suite $(v_n)_{0 \leq n < N}$ de vecteurs linéairement indépendants dans un espace de Hilbert H , il existe une suite orthonormée $(f_n)_{0 \leq n < N}$ telle que*

$$\text{Vect}(f_k : 0 \leq k \leq n) = \text{Vect}(v_k : 0 \leq k \leq n)$$

pour tout entier n tel que $0 \leq n < N$.

Preuve. — Le vecteur v_0 est non nul, puisqu'il fait partie d'un système libre ; on introduit pour commencer le vecteur de norme 1

$$f_0 = \|v_0\|^{-1}v_0,$$

et on a bien que $\text{Vect}(f_0) = \text{Vect}(v_0)$. Supposons que f_0, \dots, f_n aient été déterminés et que

$$\text{Vect}(v_0, \dots, v_n) = F_n = \text{Vect}(f_0, \dots, f_n);$$

la projection orthogonale P_n sur F_n est donnée pour tout vecteur $x \in H$ par

$$P_n x = \sum_{k=0}^n \langle x, f_k \rangle f_k.$$

Si $n + 1 = N$, c'est fini : on a traité tous les vecteurs disponibles. Sinon, $n + 1 < N$ et puisque les vecteurs (v_j) sont indépendants, on sait que $v_{n+1} \notin \text{Vect}(v_0, \dots, v_n) = F_n$, donc $y_{n+1} = P_n v_{n+1} \neq v_{n+1}$; posons

$$f_{n+1} = \|v_{n+1} - y_{n+1}\|^{-1}(v_{n+1} - y_{n+1}).$$

Ce vecteur de norme 1 est orthogonal à F_n , donc à f_0, \dots, f_n ce qui montre que les vecteurs f_0, \dots, f_{n+1} sont orthonormés. Par ailleurs, $f_{n+1} \in F_{n+1} = \text{Vect}(v_0, \dots, v_{n+1})$ puisque $y_{n+1} \in F_n \subset F_{n+1}$ et $v_{n+1} \in F_{n+1}$; les $n + 2$ vecteurs f_0, \dots, f_{n+1} sont libres, et ils sont tous dans l'espace F_{n+1} , qui est de dimension $n + 2$. Ils forment donc une base de F_{n+1} et par conséquent

$$\text{Vect}(v_0, \dots, v_{n+1}) = \text{Vect}(f_0, \dots, f_{n+1}).$$

On a ainsi démontré la possibilité de la récurrence.

Corollaire 2.1.9. *Si F est un sous-espace vectoriel de dimension finie de H , il admet des bases orthonormées ; le sous-espace F est fermé ⁽ⁱ⁾ dans H .*

Preuve. — Si $N = \dim F$, on peut trouver une base $(v_n)_{0 \leq n < N}$ pour F ; en appliquant la proposition 2.1.8, on peut remplacer cette base par une base orthonormée f_0, \dots, f_{N-1} . La projection orthogonale P_F définie par

$$\forall x \in H, \quad P_F x = \sum_{k=0}^{N-1} \langle x, f_k \rangle f_k$$

est continue, et F est exactement l'ensemble des vecteurs $x \in H$ tels que $P_F x = x$, donc $F = \ker(\text{Id}_H - P_F)$ est fermé.

Définition : espaces séparables. Un espace métrique (X, d) est *séparable* s'il existe une suite $(x_n)_{n \geq 0}$ de points de X qui est dense dans X : pour tout $x \in X$ et tout $\varepsilon > 0$, il existe un entier n tel que $d(x, x_n) < \varepsilon$.

Il est connu que les espaces \mathbb{R} , \mathbb{C} ou \mathbb{R}^m , pour tout entier $m \geq 2$, sont séparables, mais on va le revoir plus loin. On peut montrer que l'espace $L^2([a, b])$ est séparable. On déduit de Gram-Schmidt le résultat qui suit.

Corollaire. Si H est un espace de Hilbert de dimension infinie et séparable, il existe une base hilbertienne $(e_n)_{n \geq 0}$ pour H , indexée par l'ensemble \mathbb{N} des entiers ≥ 0 .

Preuve. — Par définition il existe une suite dense (x_k) dans H ; on peut (j) construire par récurrence une sous-suite $v_n = x_{k_n}$ formée de vecteurs indépendants et telle que $x_j \in \text{Vect}(v_0, \dots, v_n)$ pour tout $j \leq k_n$. D'après Gram-Schmidt, il existe une suite orthonormée $(e_j)_{j \geq 0}$ telle que $\text{Vect}(e_j, j \geq 0)$ contienne tous les $\text{Vect}(v_0, \dots, v_n)$ et en particulier tous les vecteurs $(x_k)_{k \geq 0}$. Il en résulte que $\text{Vect}(e_j, j \geq 0)$ est dense dans H , donc $(e_j)_{j \geq 0}$ est une base hilbertienne, par définition.

Réciproque : si H admet une base hilbertienne finie ou dénombrable, il est séparable.

Preuve. — Supposons pour simplifier que $\mathbb{K} = \mathbb{R}$, et supposons d'abord que la base hilbertienne soit finie (l'espace H est de dimension finie dans ce cas), disons (e_1, \dots, e_N) . On considère pour tout $n \geq 1$ l'ensemble A_n des vecteurs z de la forme

$$z = \sum_{k=1}^N c_k e_k$$

où $|c_k| \leq n$ pour $1 \leq k \leq N$; il est clair que $A_n \subset A_{n+1}$ pour tout n et que la réunion des A_n est égale à H . Considérons ensuite, pour chaque entier $n \geq N$, le sous-ensemble B_n des points z de A_n dont les coordonnées c_k sont astreintes à être des multiples entiers $j n^{-2}$, $j = -n^3, -n^3 + 1, \dots, n^3$ de n^{-2} ; l'approximation d'un élément quelconque a de A_n par un élément de B_n sera possible avec une erreur $\leq N n^{-2} \leq 1/n$ (remplacer chaque coordonnée de a par le réel $j n^{-2}$ le plus proche, avec une erreur $\leq n^{-2}$). L'ensemble B_n est fini, de cardinal $(2n^3 + 1)^N$; si on place dans une liste $(d_k)_{k \geq 0}$ tous les éléments de B_N , suivis de tous les éléments de B_{N+1} , B_{N+2} , etc., il est clair que la suite (d_k) sera dense dans H .

Si la dimension de H est infinie, soit $(e_k)_{k \geq 0}$ la base hilbertienne de H , donnée par hypothèse ; d'après la première étape on peut trouver pour tout entier k un ensemble dénombrable $D_k \subset H_k = \text{Vect}(e_0, \dots, e_k)$ qui soit dense dans H_k ; la réunion $D = \bigcup_k D_k$ est encore dénombrable, et elle est dense dans H : si x est un vecteur de H , on sait que x est la limite dans H des vecteurs

$$x_k = \sum_{j=0}^k \langle x, e_j \rangle e_j \in H_k,$$

lorsque $k \rightarrow +\infty$; on peut donc trouver k tel que $\|x - x_k\| < \varepsilon/2$. On peut ensuite trouver un élément d de D_k , donc de D , tel que $\|x_k - d\| < \varepsilon/2$ et finalement $\|x - d\| < \varepsilon$, ce qui montre que D est dense dans H .

2.2. La base hilbertienne des exponentielles complexes

Pour chaque entier $n \in \mathbb{Z}$, définissons la fonction e_n par

$$\forall x \in \mathbb{R}, \quad e_n(x) = e^{inx}.$$

Ces fonctions sont 2π -périodiques ; on les considérera aussi comme des fonctions sur l'intervalle $[0, 2\pi]$, qu'on munira de la mesure $d\mu(x) = (2\pi)^{-1}dx$, multiple de la mesure de Lebesgue. Pour cette mesure μ , l'intervalle $[0, 2\pi]$ est de mesure 1, et on a vu à l'exemple 2.1.7 que la famille $(e_n)_{n \in \mathbb{Z}}$ est orthonormée dans $L^2(\mu)$. On va maintenant montrer que cette famille est une base hilbertienne de l'espace $L^2([0, 2\pi], dx/(2\pi))$. Avant de se lancer dans la preuve, on va faire quelques rappels et introduire de nouveaux éléments utiles.

Norme uniforme, convergence uniforme

Si $C([a, b])$ désigne l'espace vectoriel des fonctions continues sur l'intervalle fermé borné $[a, b]$, on le munit de la *norme uniforme* $\|f\|_u$ définie pour toute fonction continue f par

$$\|f\|_u = \max\{|f(t)| : t \in [a, b]\}$$

(d'après le théorème A.5 de l'annexe, cette norme est égale à $\|f\|_\infty$, la norme induite par l'espace $L^\infty([a, b])$, où $[a, b]$ est muni de la mesure de Lebesgue). Dire qu'une suite (f_n) de fonctions continues converge uniformément sur $[a, b]$ vers une fonction f peut s'exprimer au moyen de la norme uniforme,

$$\lim_n \|f_n - f\|_u = 0.$$

Quand f est une fonction continue 2π -périodique sur \mathbb{R} , son maximum sur \mathbb{R} est identique à son maximum sur n'importe quel intervalle $[a, b]$ de longueur 2π , par exemple égal au maximum de f sur l'intervalle $[a, b] = [0, 2\pi]$.

Intégrale sur une période. Si f est continue 2π -périodique, la dérivée en x de

$$\Phi : x \rightarrow \int_x^{x+2\pi} f(t) dt$$

est nulle, puisqu'elle vaut $f(x+2\pi) - f(x) = 0$, ce qui montre que l'intégrale sur une période ne dépend pas de l'intervalle de longueur 2π choisi. Le même résultat est vrai aussi pour une fonction 2π -périodique intégrable sur $[0, 2\pi]$, en découpant l'intégrale : si $0 \leq x < 2\pi$ on peut écrire, d'abord pour une fonction mesurable positive f (qui permet de raisonner avec des intégrales infinies)

$$\begin{aligned} \Phi(x) &= \int_x^{x+2\pi} f(t) dt = \int_x^{2\pi} f(t) dt + \int_{2\pi}^{x+2\pi} f(t) dt \\ &= \int_x^{2\pi} f(t) dt + \int_0^x f(t) dt = \int_0^{2\pi} f(t) dt = \Phi(0); \end{aligned}$$

pour x quelconque, on peut trouver x_1 de la forme $x + 2k\pi$, $k \in \mathbb{Z}$, tel que $0 \leq x_1 < 2\pi$; par la périodicité de f on voit facilement que $\Phi(x) = \Phi(x_1)$, et on sait que $\Phi(x_1) = \Phi(0)$

par ce qui précède. Si f est intégrable, on reprend les calculs, et on se sert, pour justifier l'existence des intégrales, du cas de $|f|$ déjà traité.

La *convolution périodique* de f, g , fonctions 2π -périodiques intégrables sur $[0, 2\pi]$, sera définie par

$$\forall x \in \mathbb{R}, \quad (f \underset{\text{per}}{*} g)(x) = \int_a^{a+2\pi} f(x-t)g(t) \frac{dt}{2\pi},$$

où la valeur de l'intégrale ne dépend pas de a , d'après la remarque précédente. Le fait que l'intégrale ait un sens pour presque tout x résulte du cas de la convolution sur \mathbb{R} , traité au chapitre 1 : en effet, si on se limite à x dans un intervalle borné $[u, v]$, la formule précédente coïncide, au facteur 2π près, avec la convolution sur \mathbb{R} de deux fonctions intégrables sur \mathbb{R} , à savoir $\mathbf{1}_{[u-a-2\pi, v-a]}$ f et $\mathbf{1}_{[a, a+2\pi]}$ g . Le fait que la valeur de la convolution périodique ne dépend pas de l'intervalle de longueur 2π choisi permet de montrer par changement de variable que

$$(f \underset{\text{per}}{*} g)(x) = (g \underset{\text{per}}{*} f)(x).$$

Théorème 2.2.1. *Les fonctions $(e_n)_{n \in \mathbb{Z}}$ définies par $e_n(t) = e^{int}$ forment une base hilbertienne de l'espace de Hilbert $L^2([0, 2\pi], dx/(2\pi))$.*

La preuve sera un peu longue. On sait déjà que $(e_n)_{n \in \mathbb{Z}}$ est une suite orthonormée, il reste à prouver qu'elle engendre un sous-espace vectoriel qui est dense dans $L^2([0, 2\pi])$. Ce sous-espace vectoriel est l'espace des *polynômes trigonométriques*, qui contient toutes les fonctions P de la forme

$$P = \sum_{k=-N}^N c_k e_k,$$

où N varie dans \mathbb{N} . On va en fait établir au théorème 2.2.5 un résultat plus fort, qui donne une densité des polynômes trigonométriques dans l'espace des fonctions continues 2π -périodiques (densité pour la norme uniforme). Comme la mesure de $[0, 2\pi]$ est finie, cette information permettra de conclure, comme on le verra (nettement) plus loin.

L'ensemble des polynômes trigonométriques est un sous-espace vectoriel, et on note aussi que le produit de deux polynômes trigonométriques est encore un polynôme trigonométrique.

Lemme 2.2.2. *Il existe une suite (F_n) de polynômes trigonométriques vérifiant les trois propriétés suivantes : pour tout $n \geq 0$, la fonction F_n est réelle ≥ 0 , on a*

$$\int_{-\pi}^{\pi} F_n(t) \frac{dt}{2\pi} = 1$$

et pour tout δ tel que $0 < \delta \leq \pi$,

$$\lim_n \int_{\{\delta < |t| \leq \pi\}} F_n(t) \frac{dt}{2\pi} = 0.$$

Autrement dit, la «masse» de F_n (limitée à l'intervalle $[-\pi, \pi]$) se concentre à l'origine quand $n \rightarrow +\infty$; c'est le phénomène *d'approximation de l'unité*.

Preuve. — Pour chaque entier $n \geq 1$, considérons le polynôme trigonométrique

$$f_n = e_0 + \cdots + e_{n-1},$$

qu'on peut calculer explicitement comme somme d'une progression géométrique,

$$f_n(x) = 1 + e^{ix} + e^{2ix} + \cdots + e^{(n-1)ix} = \frac{e^{inx} - 1}{e^{ix} - 1} = \frac{e^{inx/2} \sin(nx/2)}{e^{ix/2} \sin(x/2)};$$

la dernière expression est commode pour trouver le module $|f_n(x)|$; la fonction complexe conjuguée $\overline{f_n} = \sum_{k=0}^{n-1} e_{-k}$ est aussi un polynôme trigonométrique; le carré du module de f_n est un autre polynôme trigonométrique, égal au produit $f_n \overline{f_n}$. D'après ce qui précède, on a aussi

$$|f_n(x)|^2 = \left(\frac{\sin(nx/2)}{\sin(x/2)} \right)^2.$$

L'expression initiale de f_n comme somme des n fonctions orthogonales e_0, \dots, e_{n-1} de norme 1 montre que

$$\int_0^{2\pi} |f_n(x)|^2 \frac{dx}{2\pi} = n.$$

Pour $n \geq 1$, on posera

$$F_n(x) = \frac{|f_n(x)|^2}{n} = \frac{1}{n} \left(\frac{\sin(nx/2)}{\sin(x/2)} \right)^2.$$

Le polynôme trigonométrique F_n est une fonction réelle ≥ 0 , paire, d'intégrale 1 sur chaque période (pour la mesure normalisée). On voit facilement que la suite (F_n) vérifie la troisième condition du lemme : en effet, pour tout $\delta > 0$ fixé, plus petit que π , on a

$$\int_{\{\delta < |x| < \pi\}} F_n(x) \frac{dx}{2\pi} = 2 \int_{\delta}^{\pi} F_n(x) \frac{dx}{2\pi} \leq \frac{1}{n\pi} \int_{\delta}^{\pi} \frac{dx}{\sin(x/2)^2} \leq \frac{1}{n \sin(\delta/2)^2}$$

quantité qui tend vers 0 quand $n \rightarrow +\infty$.

Remarque. Il y a de nombreuses façons d'obtenir des suites (F_n) qui vérifient les conditions données au lemme précédent; le plus souvent, elles proviennent du calcul explicite d'une certaine formule (ci-dessus, la somme d'une progression géométrique). La suite précédente (F_n) est un exemple classique : la fonction $K_n = F_{n-1}$ définie pour tout $n \geq 0$ est un polynôme trigonométrique appelé *noyau de Fejér*; on en reparlera dans le chapitre sur les séries de Fourier.

Un autre exemple est donné par le *noyau de Poisson* : pour $0 < r < 1$ on pose

$$\begin{aligned} \forall \theta \in \mathbb{R}, \quad P_r(\theta) &= \sum_{k \in \mathbb{Z}} r^{|k|} e^{ik\theta} = \sum_{k=0}^{+\infty} r^k e^{ik\theta} + \sum_{k=1}^{+\infty} r^k e^{-ik\theta} \\ &= \frac{1}{1 - r e^{i\theta}} + \frac{r e^{-i\theta}}{1 - r e^{-i\theta}} = \frac{1 - r^2}{|1 - r e^{i\theta}|^2} = \frac{1 - r^2}{1 + r^2 - 2r \cos(\theta)}. \end{aligned}$$

On obtiendra une suite du type (F_n) en considérant (P_{r_n}) pour une suite $r_n \rightarrow 1$. Il ne s'agit pas à proprement parler d'un polynôme trigonométrique, mais d'une *série* trigonométrique. On pourrait aussi utiliser les fonctions $g_r(\theta) = c_r e^{r \cos \theta}$ pour $r \rightarrow +\infty$, c_r étant choisi pour que l'intégrale de g_r soit égale à 1.

Remarquons avant de continuer que pour toute fonction 2π -périodique g intégrable sur chaque période, et pour tout entier relatif $n \in \mathbb{Z}$, la convolée $g *_{\text{per}} e_n$ est un multiple de e_n :

$$(g *_{\text{per}} e_n)(x) = \int_0^{2\pi} g(t) e_n(x-t) \frac{dt}{2\pi} = \int_0^{2\pi} g(t) e^{in(x-t)} \frac{dt}{2\pi} = e^{inx} \int_0^{2\pi} g(t) e^{-int} \frac{dt}{2\pi}.$$

On a donc pour tout $n \in \mathbb{Z}$

$$(6) \quad g *_{\text{per}} e_n = c_n(g) e_n,$$

où on pose

$$c_n(g) = \int_{-\pi}^{\pi} g(t) e^{-int} \frac{dt}{2\pi};$$

ce nombre $c_n(g)$ est le n ième coefficient de Fourier complexe de la fonction 2π -périodique g . Il est clair par linéarité que si $F = \sum_{k=-n}^n a_k e_k$ est un polynôme trigonométrique, alors

$$g *_{\text{per}} F = \sum_{k=-n}^n a_k c_k(g) e_k$$

est encore un polynôme trigonométrique. Énonçons le résultat obtenu.

Lemme 2.2.3. *Si F est un polynôme trigonométrique et g une fonction 2π -périodique intégrable sur chaque période, la convolée périodique $g *_{\text{per}} F$ est un polynôme trigonométrique.*

Proposition 2.2.4. *Pour toute suite (F_n) vérifiant les conditions du lemme 2.2.2 et pour toute fonction continue 2π -périodique g , la suite $g *_{\text{per}} F_n$ est formée de polynômes trigonométriques et tend uniformément vers g .*

Preuve. — On se donne $\varepsilon > 0$ et on trouve, parce que g est ^(k) uniformément continue sur \mathbb{R} , un réel $\delta > 0$ tel que $|g(x) - g(y)| < \varepsilon/2$ dès que $|x - y| < \delta$; on peut supposer $0 < \delta < \pi$ et on choisit ensuite n_0 assez grand pour que

$$\int_{\{\delta < |t| < \pi\}} F_n(t) \frac{dt}{2\pi} < \frac{\varepsilon}{1 + 4\|g\|_u}$$

pour tout $n \geq n_0$; on écrit

$$g(x) - (g *_{\text{per}} F_n)(x) = \int_{-\pi}^{\pi} (g(x) - g(x-t)) F_n(t) \frac{dt}{2\pi};$$

on découpe l'intégrale selon que $|t| < \delta$ ou bien $|t| \geq \delta$; dans le premier cas, on a $|g(x) - g(x-t)| < \varepsilon/2$ par le choix de δ , et dans le deuxième cas on majore la différence par $2\|g\|_u$; on obtient ainsi, pour tout x réel

$$\begin{aligned} |g(x) - (g *_{\text{per}} F_n)(x)| &\leq \frac{\varepsilon}{2} \int_{-\pi}^{\pi} F_n(t) \frac{dt}{2\pi} + 2\|g\|_u \int_{\{\delta \leq |t| \leq \pi\}} F_n(t) \frac{dt}{2\pi} \\ &\leq \frac{\varepsilon}{2} + 2\|g\|_u \frac{\varepsilon}{1 + 4\|g\|_u} < \varepsilon, \end{aligned}$$

donc $\|g - g *_{\text{per}} F_n\|_u < \varepsilon$.

Théorème 2.2.5 (Weierstrass périodique). *Si g est une fonction continue 2π -périodique sur \mathbb{R} , on peut l'approcher uniformément par des polynômes trigonométriques.*

Preuve. — C'est une conséquence immédiate de ce qui précède : d'après le lemme 2.2.2, il existe une approximation de l'unité (F_n) formée de polynômes trigonométriques ; d'après la proposition 2.2.4, les polynômes trigonométriques $g *_{\text{per}} F_n$ convergent uniformément vers g .

Fin de la démonstration du théorème 2.2.1. — On doit prouver que toute fonction $f \in L^2([0, 2\pi])$ peut être approchée en norme L^2 par un polynôme trigonométrique. Par le lemme 1.3.2, on peut déjà trouver une fonction $g \in \text{Aff}_c(\mathbb{R})$, nulle hors de $[0, 2\pi]$ et telle que

$$\int_0^{2\pi} |f(x) - g(x)|^2 \frac{dx}{2\pi} < \varepsilon^2/4.$$

Considérons la fonction 2π -périodique \tilde{g} qui coïncide avec g sur $[0, 2\pi[$. Comme on a $g(0) = g(2\pi) = 0$, la fonction \tilde{g} est continue sur \mathbb{R} . D'après le théorème précédent, il existe un polynôme trigonométrique P tel que pour tout $x \in [0, 2\pi]$, on ait

$$|g(x) - P(x)| = |\tilde{g}(x) - P(x)| < \varepsilon/2,$$

donc

$$\int_0^{2\pi} |P(x) - g(x)|^2 \frac{dx}{2\pi} < \varepsilon^2/4.$$

Ainsi, on a $\|f - g\|_2 < \varepsilon/2$ et $\|g - P\|_2 < \varepsilon/2$, donc $\|f - P\|_2 < \varepsilon$.

On a donc prouvé que le système trigonométrique est une base hilbertienne : pour toute fonction $f \in L^2([0, 2\pi])$, on a dans l'espace L^2

$$f = \sum_{n \in \mathbb{Z}} \langle f, e_n \rangle e_n$$

où

$$\langle f, e_n \rangle = \int_0^{2\pi} f(t) e^{-int} \frac{dt}{2\pi} = c_n(f).$$

La série indexée par \mathbb{Z} peut sembler troublante, on va la commenter. Tout d'abord, on sait d'après la proposition 2.1.6 que les deux séries « ordinaires »

$$\sum_{n=0}^{+\infty} \langle f, e_n \rangle e_n, \quad \sum_{n=1}^{+\infty} \langle f, e_{-n} \rangle e_{-n} = \sum_{-\infty}^{-1} \langle f, e_n \rangle e_n$$

convergent dans L^2 . De plus, si on énumère tous les indices dans \mathbb{Z} , par exemple sous la forme $0, -1, 1, -2, 2, \dots$ c'est-à-dire que $n_{2k} = k$ et $n_{2k-1} = -k$, on sait aussi que

$$f = \sum_{k=0}^{+\infty} \langle f, e_{n_k} \rangle e_{n_k}.$$

Mais la somme partielle d'indice $2p$ est égale à

$$\sum_{k=0}^{2p} \langle f, e_{n_k} \rangle e_{n_k} = \sum_{n=-p}^{-1} \langle f, e_n \rangle e_n + \sum_{n=0}^p \langle f, e_n \rangle e_n,$$

qui converge quand $p \rightarrow +\infty$ vers la somme des deux séries « ordinaires » ; on peut donc écrire pour se rassurer

$$f = \sum_{n \in \mathbb{Z}} \langle f, e_n \rangle e_n = \sum_{n=-\infty}^{-1} \langle f, e_n \rangle e_n + \sum_{n=0}^{+\infty} \langle f, e_n \rangle e_n.$$

Le fait que $f = \sum_{n \in \mathbb{Z}} \langle f, e_n \rangle e_n$ au sens de l'espace L^2 n'indique pas s'il y a égalité pour des valeurs de x , ni même si la série numérique $\sum_{n \in \mathbb{Z}} c_n(f) e^{inx}$ converge pour des valeurs de x . C'est vrai d'après le théorème suivant, qui est un résultat très difficile datant du milieu des années 1960 (l'article du mathématicien suédois Lennart Carleson est paru en 1966).

Théorème de Carleson : pour toute fonction $f \in L^2([0, 2\pi])$ et pour presque tout x , la série de Fourier

$$\sum_{k=-\infty}^{+\infty} c_k(f) e^{ikx}$$

converge et sa somme est égale à $f(x)$.

Quand nous disons que la série converge, cela signifie que les deux séries $\sum_{k=0}^{+\infty} c_k(f) e^{ikx}$ et $\sum_{k=0}^{+\infty} c_{-k}(f) e^{-ikx}$ convergent.

Exemple : développement en série de Fourier de la fonction périodique f définie par $f(x) = 1 - |x|/\pi$ lorsque $|x| \leq \pi$. Pour $n \neq 0$, on voit en utilisant la parité de f que

$$\begin{aligned} c_n(f) &= \int_{-\pi}^{\pi} f(x) e^{-inx} \frac{dx}{2\pi} = \int_{-\pi}^{\pi} f(x) \cos(nx) \frac{dx}{2\pi} = \frac{1}{\pi} \int_0^{\pi} (1 - x/\pi) \cos(nx) dx \\ &= \frac{1}{\pi} \left(\frac{1}{n} \left[(1 - x/\pi) \sin(nx) \right]_0^{\pi} + \frac{1}{\pi n} \int_0^{\pi} \sin(nx) dx \right) = \frac{1 - \cos(n\pi)}{\pi^2 n^2}, \end{aligned}$$

qui est égal à 0 pour n pair non nul, et à $2/(\pi^2 n^2)$ pour n impair. On voit aussi que $c_0(f) = 1/2$. La série de Fourier de f est donc

$$\frac{1}{2} + \sum_{n \in \mathbb{Z}} \frac{1 - \cos(n\pi)}{\pi^2 n^2} e^{inx} = \frac{1}{2} + \frac{4}{\pi^2} \sum_{k=0}^{+\infty} \frac{\cos((2k+1)x)}{(2k+1)^2}.$$

La série précédente est normalement convergente, et sa somme $g(x)$ définit donc une fonction continue ; comme f et g sont continues et doivent être égales presque partout (puisqu'elles représentent la même classe dans L^2) il en résulte ⁽¹⁾ que $f(x) = g(x)$ pour tout x . La valeur en $x = 0$ ou $x = \pi$ donne $\sum_{k=0}^{+\infty} (2k+1)^{-2} = \pi^2/8$, ce qui implique l'égalité classique

$$\sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

2.3. Projection orthogonale et applications

Lemme 2.3.1. Soient H un espace de Hilbert sur \mathbb{K} , F un \mathbb{K} -sous-espace vectoriel de H , $x \in H$ et $y \in F$; alors $x - y$ est orthogonal à F si et seulement si y est le point de F le plus proche de x ,

$$\|x - y\| = d(x, F) = \min\{\|x - z\| : z \in F\}.$$

Le point $y \in F$ qui minimise la distance de x aux points de F est unique (s'il existe ; dans ce cas, il est appelé projection orthogonale de x sur F).

Preuve. — Supposons d'abord que $x - y \perp F$; pour tout vecteur $z \in F$, on voit que $y - z \in F$, donc $y - z$ est orthogonal à $x - y$ et

$$\|x - z\|^2 = \|(x - y) + (y - z)\|^2 = \|x - y\|^2 + \|y - z\|^2 \geq \|x - y\|^2,$$

donc y est le point de F le plus proche de x . Supposons inversement que y soit le point de F le plus proche de x ; si v est un vecteur de F , le vecteur $y + tv$ est dans F pour tout réel t , donc

$$\|x - (y + tv)\|^2 \geq \|x - y\|^2$$

pour tout t ; la fonction

$$t \rightarrow \|x - (y + tv)\|^2 = \|x - y\|^2 - 2t \operatorname{Re}\langle x - y, v \rangle + t^2 \|v\|^2$$

atteint son minimum en $t = 0$, donc sa dérivée en $t = 0$ est nulle, ce qui donne

$$\operatorname{Re}\langle x - y, v \rangle = 0$$

pour tout vecteur $v \in F$; ceci est suffisant quand $\mathbb{K} = \mathbb{R}$, mais quand $\mathbb{K} = \mathbb{C}$ il faut faire un pas de plus : comme F est alors un \mathbb{C} -sous-espace vectoriel de H , on peut choisir θ réel de façon que si $v_1 = e^{i\theta} v$, on ait $\langle x - y, v_1 \rangle \in \mathbb{R}$. On a encore $v_1 \in F$: la première partie du raisonnement montre que $\langle x - y, v_1 \rangle = 0$, donc $\langle x - y, v \rangle = 0$, pour tout vecteur $v \in F$, ce qui signifie que $x - y$ est orthogonal à F .

Le point $y \in F$ qui minimise la distance est unique : si y' était un autre point de F tel que $x - y'$ soit orthogonal à F , on aurait par différence $y - y' \perp F$, et comme $y - y' \in F$, le vecteur $y - y'$, orthogonal à lui-même, serait nul.

Remarque. On pourrait prouver l'existence de la projection orthogonale sur un sous-espace vectoriel fermé par une succession de petits bricolages, mais on verra un argument simple et général plus loin. Voici comment bricoler : considérons d'abord un sous-espace vectoriel F fermé, séparable et de dimension infinie. D'après le corollaire de Gram-Schmidt, on peut trouver une base hilbertienne $(f_n)_{n \geq 0}$ pour F . La proposition 2.1.6 donne l'existence de la projection orthogonale sur F dans ce cas. On peut ensuite prouver ainsi l'existence de la projection de x sur un sous-espace vectoriel fermé F quelconque (c'est-à-dire même si F n'est pas séparable) : on peut trouver une suite $(y_n) \subset F$ telle que $\|x - y_n\|$ tende vers la distance de x à F

$$d = d(x, F) = \inf\{\|x - z\| : z \in F\} ;$$

le sous-espace vectoriel fermé $G \subset F$ engendré par la suite (y_n) vérifie $d(x, G) = d$ et il possède une base hilbertienne finie ou dénombrable d'après Gram-Schmidt. Le point $g \in G$ le plus proche de x existe donc, d'après ce qui précède, mais comme $\|x - g\| = d$, le point g est en même temps le point de F le plus proche de x .

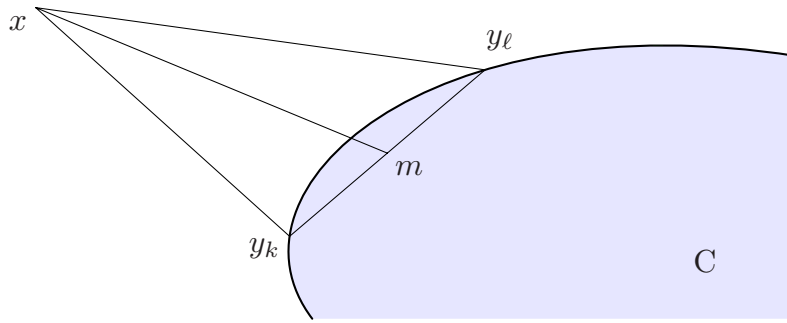
Projection en général

Considérons plus généralement un sous-ensemble convexe fermé non vide C de l'espace de Hilbert H , et un point $x \in H$. On peut toujours trouver une suite $(y_n) \subset C$ telle que $d(x, y_n) = \|x - y_n\|$ tende vers

$$d := d(x, C) = \inf\{\|x - z\| : z \in C\}.$$

On va montrer que cette suite (y_n) est de Cauchy. On pose $m = \frac{1}{2}(y_k + y_\ell)$, le milieu du segment qui joint les points y_k et y_ℓ ; ce point m est dans C , d'après la convexité de C . Posons de plus $u = x - y_k$ et $v = x - y_\ell$. On obtient par la relation du parallélogramme (3)

$$4\|x - m\|^2 + \|y_k - y_\ell\|^2 = 2(\|x - y_k\|^2 + \|x - y_\ell\|^2).$$



Comme $m \in C$, on a $\|x - m\| \geq d$ et

$$\frac{1}{2}\|y_k - y_\ell\|^2 \leq \|x - y_k\|^2 + \|x - y_\ell\|^2 - 2d^2$$

qui tend vers 0 quand $k, \ell \rightarrow +\infty$. La suite de Cauchy (y_n) converge dans l'espace complet H vers un vecteur y , qui est dans C parce que C est fermé. On a de plus

$$\|x - y\| = \lim_n \|x - y_n\| = d,$$

et on a ainsi montré l'existence d'un point $y \in C$ qui réalise la plus courte distance de x à un point de C . L'unicité résulte ici de la relation du parallélogramme : si y, y' sont deux points de C tels que

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

on peut appliquer les inégalités ci-dessus en prenant $y_k = y$ et $y_\ell = y'$; alors

$$\frac{1}{2}\|y - y'\|^2 = \frac{1}{2}\|y_k - y_\ell\|^2 \leq \|x - y_k\|^2 + \|x - y_\ell\|^2 - 2d^2 = d^2 + d^2 - 2d^2 = 0.$$

On a donc montré le théorème qui suit.

Théorème 2.3.2. Soient H un espace de Hilbert et C un sous-ensemble convexe, fermé et non vide de H ; pour tout vecteur $x \in H$, il existe un élément y de C unique qui est le point de C le plus proche de x ,

$$\|x - y\| = \min\{\|x - z\| : z \in C\}.$$

Linéarité de la projection sur un sous-espace vectoriel fermé

Dans le cas où on projette sur un sous-espace vectoriel fermé F , la projection P_F est linéaire : si les points $x_1, x_2 \in H$ ont pour projections $y_1 = P_F x_1$ et $y_2 = P_F x_2 \in F$, on voit facilement que le vecteur

$$a_1 x_1 + a_2 x_2 - (a_1 y_1 + a_2 y_2) = a_1 (x_1 - y_1) + a_2 (x_2 - y_2)$$

est orthogonal à F , ce qui montre que $a_1 y_1 + a_2 y_2 \in F$ est la projection orthogonale de $a_1 x_1 + a_2 x_2$ sur F (lemme 2.3.1). La projection P_F est donc linéaire,

$$P_F(a_1 x_1 + a_2 x_2) = a_1 y_1 + a_2 y_2 = a_1 P_F x_1 + a_2 P_F x_2.$$

Notons encore que par Pythagore, $\|x\|^2 = \|x - P_F x\|^2 + \|P_F x\|^2 \geq \|P_F x\|^2$.

Théorème 2.3.3. *Soit F un sous-espace vectoriel fermé de l'espace de Hilbert H ; l'application P_F de projection orthogonale de H sur F est linéaire, et*

$$\forall x \in H, \quad \|P_F x\| \leq \|x\|.$$

Exemples.

1. Partition finie, moyennes. Considérons un espace de probabilité (Ω, \mathcal{A}, P) . Supposons que A_1, \dots, A_N soit une partition de Ω en ensembles de la tribu \mathcal{A} , tels que $P(A_j) > 0$ pour tout $j = 1, \dots, N$; le sous-espace F de dimension N de $L^2(\Omega, \mathcal{A}, P)$ formé des fonctions qui sont constantes sur chaque ensemble de la partition admet pour base orthonormée les fonctions

$$f_j = P(A_j)^{-1/2} \mathbf{1}_{A_j}, \quad j = 1, \dots, N.$$

La projection orthogonale de $f \in L^2(\Omega, \mathcal{A}, P)$ sur F est donnée par

$$P_F f = \sum_{j=1}^N \langle f, f_j \rangle f_j = \sum_{j=1}^N \left(\frac{1}{P(A_j)} \int_{A_j} f \, dP \right) \mathbf{1}_{A_j} ;$$

la fonction $P_F f$ est constante sur chaque ensemble A_j , et sa valeur sur A_j est la *moyenne de f sur A_j* . Cette projection P_F est un cas (très) particulier de la notion d'*espérance conditionnelle*, importante en probabilité ; l'exemple suivant est un autre cas d'*espérance conditionnelle*.

2. Dans l'espace de Hilbert $H = L^2([0, 1]^2)$ on considère le sous-espace vectoriel F formé des fonctions ne dépendant que de la variable x : la fonction g appartient à F s'il existe une fonction $G(x)$ d'une seule variable, de carré intégrable sur $[0, 1]$, et telle que $g(x, y) = G(x)$ pour presque tout couple $(x, y) \in [0, 1]^2$. Ce sous-espace F est fermé (exercice). La projection orthogonale sur F d'une fonction $f \in H$ est donnée par

$$(P_F f)(x, y) = G(x) = \int_0^1 f(x, y) \, dy.$$

Décomposition en sous-espaces vectoriels orthogonaux

Pour toute partie $A \subset H$, on définit l'orthogonal A^\perp de cette partie,

$$A^\perp = \{x \in H : x \perp A\},$$

où la notation $x \perp A$ signifie que x est orthogonal à tous les éléments de A ; si A est vide, il est naturel de poser $A^\perp = H$. On voit que A^\perp est toujours un sous-espace vectoriel fermé de H , pour toute partie $A \subset H$, puisque A^\perp est l'intersection des sous-espaces vectoriels fermés $a^\perp = \ker \ell_a$, où a varie dans l'ensemble A .

Il est clair que $A^\perp \supset B^\perp$ lorsque $A \subset B$; en particulier $H^\perp = \{0\}$ est le plus petit orthogonal, correspondant à la plus grande partie possible, $A = H$. On a

$$(7) \quad A^\perp = (\overline{\text{Vect } A})^\perp.$$

Puisque $A \subset \overline{\text{Vect } A}$, il est clair que $A^\perp \supset (\overline{\text{Vect } A})^\perp$. Inversement, si $y \in A^\perp$, c'est-à-dire que y est orthogonal à A , on a vu à l'équation (5) que $y \in (\overline{\text{Vect } A})^\perp$, et la vérification est finie.

Lemme 2.3.4. *Soit H un espace de Hilbert; on suppose que l'espace H est égal à la somme vectorielle $F + G$, où F et G sont deux sous-espaces vectoriels orthogonaux (tout vecteur $f \in F$ est orthogonal à tout vecteur $g \in G$); alors H admet la décomposition en somme directe*

$$H = F \oplus G,$$

on a $G = F^\perp$, $F = G^\perp$, les sous-espaces F et G sont fermés, les deux projections $f + g \rightarrow f \in F$ et $f + g \rightarrow g \in G$ de la somme directe sont les projections orthogonales de H sur F et G respectivement.

Preuve. — Pour commencer, si $x \in F \cap G$, alors x élément de F est orthogonal à lui-même, élément de G , donc $x = 0$; puisque $F \cap G = \{0\}$ et $H = F + G$, l'espace H est somme directe de F et G . Par hypothèse on a $G \subset F^\perp$; inversement, si $v \in H$ est orthogonal à F , écrivons $v = f + g$, avec $f \in F$ et $g \in G$; on a

$$0 = \langle v, f \rangle = \langle f + g, f \rangle = \langle f, f \rangle + \langle f, g \rangle = \langle f, f \rangle,$$

ce qui montre que $f = 0$, donc $v = g$ est dans G et on a montré que $G = F^\perp$. On montre de la même façon que $F = G^\perp$, et il en résulte que F et G sont fermés.

Si $x = f + g$, avec $f \in F$ et $g \in G$, le vecteur f est dans F , et $x - f = g$ est orthogonal à F , donc f est la projection orthogonale de x sur F . De même, g est la projection orthogonale de x sur G .

Remarque. Le lemme peut s'appliquer quand H est la somme de trois (ou plus) sous-espaces orthogonaux : si $H = F_1 + F_2 + F_3$ est la somme vectorielle de trois sous-espaces vectoriels F_1, F_2 et F_3 deux à deux orthogonaux, alors

$$H = F_1 + (F_2 + F_3)$$

avec F_1 orthogonal à $F_2 + F_3$, donc F_1 est fermé d'après le lemme, ainsi que $F_2 + F_3$ et on a $H = F_1 \oplus (F_2 + F_3)$; de même les autres sous-espaces sont fermés, et de plus $H = F_1 \oplus F_2 \oplus F_3$. En effet, le sous-espace fermé $F_2 + F_3$ peut être considéré comme un espace de Hilbert auquel appliquer le lemme 2.3.4, donc $F_2 + F_3 = F_2 \oplus F_3$ et finalement $H = F_1 \oplus (F_2 \oplus F_3) = F_1 \oplus F_2 \oplus F_3$.

Théorème 2.3.5. Soit F un sous-espace vectoriel fermé d'un espace de Hilbert H ; l'espace H admet la décomposition en somme directe orthogonale

$$H = F \oplus F^\perp.$$

Les deux projecteurs de la somme directe sont les projections orthogonales sur les sous-espaces F et F^\perp . On a de plus

$$(F^\perp)^\perp = F.$$

Preuve. — Pour tout $x \in H$ on a

$$x = P_F(x) + (x - P_F(x))$$

avec $P_F x \in F$ et $x - P_F x \perp F$; ceci montre que $H = F + F^\perp$. D'après le lemme 2.3.4, l'espace H est somme directe des deux sous-espaces vectoriels orthogonaux F et $G = F^\perp$, les deux projections de la somme directe sont les projections orthogonales sur les facteurs, et de plus F est l'orthogonal de $G = F^\perp$.

On notera que la projection orthogonale sur F^\perp est égale à $\text{Id} - P_F$.

Proposition 2.3.6 : critère de densité. Pour qu'une partie A de H engendre un sous-espace vectoriel $\text{Vect}(A)$ dense dans H , il faut et il suffit que 0 soit le seul vecteur orthogonal à l'ensemble A , c'est-à-dire que

$$A^\perp = \{0\}.$$

Preuve. — Si $F = \overline{\text{Vect}(A)}$ est différent de H , on voit, par exemple par le théorème 2.3.5, que F^\perp n'est pas réduit à 0 , donc on peut trouver un vecteur v non nul orthogonal à F , en particulier orthogonal à tous les éléments de A . Inversement, si $\text{Vect}(A)$ est dense dans H , on a en utilisant l'équation (7)

$$A^\perp = (\overline{\text{Vect} A})^\perp = H^\perp = \{0\}.$$

Théorème de Weierstrass

Proposition 2.3.7. Pour tout réels $a < b$, les monômes $f_n(x) = x^n$, $n \in \mathbb{N}$, engendrent un sous-espace vectoriel dense dans $L^2([a, b])$.

Preuve. — Dans le cas contraire on pourrait, par la proposition 2.3.6, trouver une fonction $g \in L^2([a, b])$ non nulle qui serait orthogonale à tous les monômes,

$$\int_a^b g(x) x^n dx = 0$$

pour tout $n \geq 0$. Considérons la fonction \tilde{g} sur \mathbb{R} qui est égale à g dans $[a, b]$ et à 0 en dehors, et la transformée de Fourier de \tilde{g}

$$G(y) = \int_{\mathbb{R}} \tilde{g}(x) e^{-ixy} dx = \int_a^b g(x) e^{-ixy} dx.$$

On va montrer que $G(y) = 0$ pour tout y : fixons y , notons que pour tout $x \in [a, b]$

$$g(x) e^{-ixy} = \sum_{k=0}^{+\infty} g(x) \frac{(-ixy)^k}{k!}.$$

La série des modules de la série précédente est convergente, et sa somme est intégrable :

$$\sum_{k=0}^{+\infty} \frac{|xy|^k}{k!} |g(x)| = e^{|xy|} |g(x)|;$$

cette fonction $x \rightarrow e^{|xy|} |g(x)|$ est intégrable sur l'intervalle borné $[a, b]$ (lemme A.7, annexe). Il en résulte par le théorème A.2 qu'il est possible d'invertir la série et l'intégrale,

$$G(y) = \int_a^b g(x) e^{-ixy} dx = \sum_{k=0}^{+\infty} \frac{(-iy)^k}{k!} \int_a^b g(x) x^k dx = 0.$$

On en déduit par l'injectivité de Fourier (qui résulte de Parseval, puisque $\tilde{g} \in L^2(\mathbb{R})$) que $\tilde{g} = 0$, donc $g = 0$, contrairement à notre hypothèse initiale.

Lemme. *Pour toute fonction continue F sur un intervalle fermé borné $[a, b] \subset \mathbb{R}$ et pour tout $\varepsilon > 0$, il existe une fonction $G \in \text{Aff}(a, b)$ telle que $|F(x) - G(x)| < \varepsilon$ pour tout $x \in [a, b]$, c'est-à-dire que $\|F - G\|_u < \varepsilon$.*

Preuve. — Soient F une fonction continue sur $[a, b]$ et $\varepsilon > 0$; par continuité uniforme, on peut trouver $\delta > 0$ tel que $|F(y) - F(x)| < \varepsilon/2$, dès que $|y - x| < \delta$. Considérons une subdivision $a = x_0 < x_1 < \dots < x_N = b$ de l'intervalle $[a, b]$, de pas $< \delta$, c'est-à-dire que $x_j - x_{j-1} < \delta$ pour tout $j = 1, \dots, N$. Introduisons la fonction G continue et affine sur chaque intervalle $[x_j, x_{j-1}]$, $j = 1, \dots, N$, telle que $G(x_j) = F(x_j)$ pour $j = 0, \dots, N$. Si x est dans l'intervalle $[x_{j-1}, x_j]$, on a $x = (1 - t)x_{j-1} + tx_j$ pour un certain $t \in [0, 1]$, donc $G(x) = (1 - t)G(x_{j-1}) + tG(x_j)$ et

$$|G(x) - G(x_{j-1})| = t |G(x_j) - G(x_{j-1})| \leq |F(x_j) - F(x_{j-1})|$$

qui est $< \varepsilon/2$ par le choix de δ ; on a aussi $|F(x) - F(x_{j-1})| < \varepsilon/2$, $G(x_{j-1}) = F(x_{j-1})$, donc $|F(x) - G(x)| < \varepsilon$. Ceci est valable pour tout x de $[a, b]$, donc $\|F - G\|_u < \varepsilon$.

Théorème (Weierstrass). *Pour tout intervalle fermé borné $[a, b] \subset \mathbb{R}$, les monômes $f_n(x) = x^n$, $n \in \mathbb{N}$, engendrent un sous-espace vectoriel dense dans $C([a, b])$; autrement dit, l'espace des fonctions polynomiales sur $[a, b]$ est dense dans $C([a, b])$.*

Preuve. — Soient F une fonction continue sur $[a, b]$ et $\varepsilon > 0$; par le lemme précédent, il existe une fonction $G \in \text{Aff}(a, b)$ telle que $\|F - G\|_u < \varepsilon$. Soit g la dérivée de G, qui est définie sauf en un nombre fini de points; la classe de g est une fonction en escalier, donc $g \in L^2([a, b])$. D'après la proposition 2.3.7, on peut pour tout $\varepsilon > 0$ trouver une fonction polynomiale p sur $[a, b]$ telle que $\|p - g\|_2 < \varepsilon/\sqrt{b - a}$; alors

$$P(x) = G(a) + \int_a^x p(t) dt$$

est une fonction polynomiale, et on sait par le lemme 1.1.2 que

$$G(x) = G(a) + \int_a^x g(t) dt;$$

par Cauchy-Schwarz on a pour tout x de $[a, b]$

$$|G(x) - P(x)| = \left| \int_a^x (g(t) - p(t)) dt \right| \leq \int_a^b |g(t) - p(t)| dt \leq \sqrt{b - a} \|g - p\|_2 < \varepsilon,$$

ce qui démontre la possibilité d'approcher à 2ε près la fonction F par une fonction polynomiale P, uniformément sur $[a, b]$.

Polynômes orthogonaux

La suite des monômes est linéairement indépendante dans $L^2([a, b])$, et elle engendre un espace vectoriel dense dans $L^2([a, b])$. Par Gram-Schmidt, on pourra fabriquer une base hilbertienne de $L^2([a, b])$ formée de fonctions polynomiales. On peut en fait écrire des formules explicites, par exemple dans le cas de l'intervalle $[-1, 1]$; la fonction polynomiale $P_n(x) = D^n(1 - x^2)^n$ (dérivée n ème de $x \rightarrow (1 - x^2)^n$; on pose $P_0 = 1$) est de degré n , et on montre par intégration par parties que les polynômes $(P_n)_{n \geq 0}$ sont deux à deux orthogonaux. Il en résulte que P_n est proportionnel au n ème vecteur obtenu par la méthode de Gram-Schmidt appliquée à la suite des monômes.

Dual (topologique) d'un espace de Hilbert H

On a vu que pour tout vecteur $v \in H$, l'application $\ell_v : x \in H \rightarrow \langle x, v \rangle$ est une forme linéaire continue sur H . On va voir une réciproque.

Théorème 2.3.8. *Toute forme linéaire continue ℓ sur un espace de Hilbert H est de la forme $\ell = \ell_v$ pour un certain vecteur $v \in H$, c'est-à-dire que*

$$\forall x \in H, \quad \ell(x) = \langle x, v \rangle;$$

ce vecteur v est unique.

En d'autres termes, l'application $v \rightarrow \ell_v$ est une bijection de l'espace de Hilbert H sur son dual topologique H' . *Attention !* cette application n'est pas linéaire dans le cas complexe, car l'image du vecteur λv est la forme linéaire $\bar{\lambda} \ell_v$ (et pas $\lambda \ell_v$); l'application $v \rightarrow \ell_v$ est une bijection *antilinéaire* de l'espace de Hilbert H sur son dual topologique H' .

Preuve. — Montrons l'unicité : si v_1 et v_2 étaient deux vecteurs de H tels que $\ell_{v_1} = \ell_{v_2}$, on aurait $\langle x, v_1 \rangle = \ell(x) = \langle x, v_2 \rangle$ pour tout vecteur x , donc $\langle x, v_1 - v_2 \rangle = 0$; en appliquant ceci à $x = v_1 - v_2$, on déduit que $v_1 - v_2 = 0$. Pour démontrer l'existence on prouvera un lemme en apparence plus général.

Lemme. *Soient E un \mathbb{K} -sous-espace vectoriel de l'espace de Hilbert H , C un nombre réel et $\ell : E \rightarrow \mathbb{K}$ une forme linéaire sur E telle que $|\ell(x)| \leq C \|x\|$ pour tout $x \in E$; il existe un vecteur $v \in H$ tel que*

$$\forall x \in E, \quad \ell(x) = \langle x, v \rangle.$$

Preuve du lemme. — Si la forme linéaire ℓ est nulle sur E , il suffit de prendre le vecteur $v = 0$ pour la représenter; dans le cas contraire, on peut trouver un vecteur $x_0 \in E$ tel que $\ell(x_0) = 1$. Désignons par F l'adhérence dans H du noyau $\ker \ell$; il est clair que x_0 n'est pas dans F , puisque pour tout $z \in \ker \ell$ on a

$$1 = \ell(x_0) = \ell(x_0 - z) \leq C \|x_0 - z\|,$$

ce qui montre que la distance de x_0 à $\ker \ell$ est $\geq 1/C > 0$. Puisque x_0 n'appartient pas au sous-espace vectoriel fermé F , la projection $P_F x_0$ de x_0 sur F est différente de x_0 , donc $w = x_0 - P_F x_0$ est un vecteur non nul orthogonal à F . On a donc $\langle P_F x_0, w \rangle = 0$ et

$$0 < \langle w, w \rangle = \langle w + P_F x_0, w \rangle = \langle x_0, w \rangle.$$

Pour tout vecteur $x \in E$, remarquons que

$$x = (x - \ell(x) x_0) + \ell(x) x_0 = y + \ell(x) x_0$$

et $y = x - \ell(x) x_0 \in \ker \ell$; en effet, $\ell(y) = \ell(x) - \ell(x) \ell(x_0) = 0$; puisque $y \in \ker \ell \subset F$, le vecteur y est orthogonal à w et

$$\langle x, w \rangle = \langle y + \ell(x) x_0, w \rangle = \langle y, w \rangle + \ell(x) \langle x_0, w \rangle = \ell(x) \langle w, w \rangle.$$

On vient ainsi de montrer que

$$\ell(x) = \langle w, w \rangle^{-1} \langle x, w \rangle = \langle x, \langle w, w \rangle^{-1} w \rangle$$

pour tout $x \in E$. On voit donc que la forme linéaire ℓ est représentée par le produit scalaire avec le vecteur $v = \langle w, w \rangle^{-1} w$.

Exercice. Pour simplifier un tout petit peu on prendra $\mathbb{K} = \mathbb{R}$ dans cet exercice. On dit qu'une fonction réelle φ définie sur un intervalle I de la droite réelle est *C-lipschitzienne* si pour tous $x, y \in I$ on a

$$|\varphi(x) - \varphi(y)| \leq C |x - y|.$$

Avec une fonction 1-lipschitzienne φ sur $[0, 1]$ on fabrique une forme linéaire ℓ sur l'espace vectoriel \mathcal{E} des fonctions en escalier réelles de la façon suivante : si $h = \sum_{j=0}^{n-1} c_j \mathbf{1}_{[x_j, x_{j+1}[}$, où $0 = x_0 < x_1 < \dots < x_n = 1$, on pose

$$\ell(h) = \sum_{j=0}^{n-1} c_j (\varphi(x_{j+1}) - \varphi(x_j)).$$

On a

$$(8) \quad |\ell(h)| \leq \sum_{j=0}^{n-1} |c_j| |\varphi(x_{j+1}) - \varphi(x_j)| \leq \sum_{j=0}^{n-1} |c_j| (x_{j+1} - x_j) = \|h\|_1 \leq \|h\|_2.$$

D'après le lemme précédent appliqué à $E = \mathcal{E} \subset L^2([0, 1])$, la forme linéaire ℓ peut être représentée par le produit scalaire avec une fonction réelle $f \in L^2([0, 1])$: pour toute fonction en escalier h ,

$$\ell(h) = \langle h, f \rangle = \int_0^1 h(t) f(t) dt.$$

En particulier, lorsque $h = \mathbf{1}_{[0, x[}$, on trouve que

$$\varphi(x) - \varphi(0) = \ell(\mathbf{1}_{[0, x[}) = \int_0^x f(t) dt$$

pour tout $x \in [0, 1]$; on peut ensuite (m) montrer que f est dans $L^\infty([0, 1])$, et donc : toute fonction lipschitzienne est « primitive » d'une fonction mesurable bornée (et inversement, évidemment). Le théorème de représentation du dual de L^2 a permis de faire apparaître une fonction f qui n'était pas du tout visible au départ !

2.4. Opérateur adjoint hilbertien

Si H_1, H_2 sont deux espaces de Hilbert et T une application linéaire continue de H_1 dans H_2 , on introduit une application linéaire T^* de H_2 dans H_1 de la façon suivante : si y est un vecteur de H_2 , l'application

$$x \in H_1 \rightarrow \langle Tx, y \rangle$$

est une forme linéaire continue sur H_1 ; il existe donc un vecteur $T^*y \in H_1$ qui représente cette forme linéaire. On a donc

$$\forall x \in H_1, \forall y \in H_2, \quad \langle Tx, y \rangle = \langle x, T^*y \rangle,$$

et cette propriété caractérise T^* . On voit ainsi que pour tout scalaire λ , l'adjoint de λId_H est $\bar{\lambda} \text{Id}_H$ et il est clair avec cette caractérisation que

$$(T^*)^* = T.$$

On montre assez facilement que $\|T^*\| = \|T\|$. Si T_1 et T_2 sont composables, on a $(T_2 T_1)^* = T_1^* T_2^*$ (bien noter l'inversion).

Exemple. Considérons la transformation de Fourier \mathcal{F} , agissant de $L^2(\mathbb{R})$ dans $L^2(\mathbb{R})$. Essayons de trouver la formule pour l'adjoint \mathcal{F}^* , agissant pour commencer sur le sous-espace dense $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. Si $f, g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, on doit avoir

$$\langle f, \mathcal{F}^*g \rangle = \langle \mathcal{F}f, g \rangle = \int_{\mathbb{R}} (\mathcal{F}f)(y) \overline{g(y)} dy = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x) e^{-ixy} dx \right) \overline{g(y)} dy.$$

D'après Fubini, on voit que

$$\langle f, \mathcal{F}^*g \rangle = \int_{\mathbb{R}} f(x) \overline{\left(\int_{\mathbb{R}} g(y) e^{ixy} dy \right)} dx,$$

qui est le produit scalaire de f avec la fonction

$$h : x \rightarrow \int_{\mathbb{R}} g(y) e^{ixy} dy;$$

on a donc $\langle f, \mathcal{F}^*g \rangle = \langle f, h \rangle$: la différence $h - \mathcal{F}^*g$ est orthogonale à toutes les fonctions f du sous-espace dense $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ de H , donc elle est nulle. Il en résulte que \mathcal{F}^*g est égale à la fonction h . On voit ainsi que l'adjoint \mathcal{F}^* peut être défini sur $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ par une formule de même type que celle qui définit la transformée de Fourier : pour toute fonction $g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, et pour tout x on a

$$(\mathcal{F}^*g)(x) = \int_{\mathbb{R}} g(y) e^{ixy} dy.$$

Une application linéaire $T \in \mathcal{L}(H_1, H_2)$ est une *isométrie* si $\|Tx\| = \|x\|$ pour tout $x \in H_1$, ce qui s'écrit aussi

$$\langle T^*Tx, x \rangle = \langle Tx, Tx \rangle = \|Tx\|^2 = \|x\|^2 = \langle x, x \rangle;$$

on voit ainsi que la forme hermitienne

$$\varphi(x, y) = \langle T^*Tx - x, y \rangle$$

est nulle sur la diagonale ; il en résulte par polarisation (voir la relation (4)) qu'elle est nulle pour tous x, y . On en déduit que $T^*Tx = x$ pour tout $x \in H_1$. Il est facile de remonter les égalités.

Lemme. L'application $T \in \mathcal{L}(H_1, H_2)$ est isométrique si et seulement si T^*T est l'identité de H_1 .

Dans le cas $H_1 = H_2 = H$, on dit que $U \in \mathcal{L}(H)$ est *unitaire* si U est isométrique et bijective. Alors U est inversible, donc son inverse ne peut être que U^* : U est unitaire si et seulement si

$$U^*U = UU^* = \text{Id}_H.$$

Exemple. Si \mathcal{F} désigne la transformation de Fourier agissant sur $L^2(\mathbb{R})$, on sait que $U = \frac{1}{\sqrt{2\pi}}\mathcal{F}$ est unitaire. Le fait que U^* soit l'inverse de U correspond à ce qu'on a dit au chapitre 1, théorème 1.3.7, pour le calcul de \mathcal{F}^{-1} .

On dit que l'application $T \in \mathcal{L}(H)$ est *hermitienne* quand $T^* = T$. Un exemple fondamental est celui des projections orthogonales. Si F est un sous-espace vectoriel de l'espace H , on a pour tous $x, y \in H$

$$\langle P_F x, y \rangle = \langle P_F x, P_F y \rangle$$

puisque $P_F x$, élément de F , est orthogonal au vecteur différence $y - P_F y$; maintenant, la situation est symétrique en x et y et

$$\langle P_F x, y \rangle = \langle P_F x, P_F y \rangle = \langle x, P_F y \rangle,$$

ce qui montre que P_F est son propre adjoint. Dans une certaine mesure, les combinaisons linéaires à coefficients réels de projections orthogonales jouent par rapport aux hermitiens généraux le rôle des fonctions en escalier par rapport aux fonctions mesurables bornées.

Si $T \in \mathcal{L}(H)$ est un endomorphisme quelconque, on voit que $A = T^*T$ est hermitien, puisque $A^* = T^*(T^*)^* = T^*T$. De plus A est *positif*, ce qui signifie que

$$\langle Ax, x \rangle = \langle T^*Tx, x \rangle = \langle Tx, Tx \rangle = \|Tx\|^2 \geq 0.$$

Notes du chapitre 2

(a) Le terme *semi-produit scalaire* n'est pas classique. Mais la plupart des gens exige qu'un produit scalaire sur l'espace vectoriel E vérifie la propriété suivante : si $x \neq 0_E$, alors $\langle x, x \rangle \neq 0$. Quand nous supposons seulement $\langle x, x \rangle \geq 0$, nous disons *semi-produit scalaire* comme on dit *semi-norme* ; ceci n'est utilisé qu'au début du chapitre.

(b) Le choix du côté qui est linéaire dans la fonction de deux variables $(x, y) \rightarrow \langle x, y \rangle$ (pour nous, c'est le gauche) n'est pas universel. Dans les milieux proches de la Physique, on fait souvent le choix opposé à celui qui est fait dans ces notes de cours.

(c) En géométrie plane élémentaire, cette relation signifie que dans un parallélogramme $ABCD$, la somme des carrés des longueurs des deux diagonales AC et BD est égale à la somme des carrés des longueurs des quatre côtés AB, BC, CD et DA : prendre le vecteur x de la relation (3) égal au vecteur \overrightarrow{AB} et le vecteur y égal au vecteur \overrightarrow{BC} .

(d) En appliquant à x et iy le résultat obtenu, on a

$$4 \operatorname{Im} \varphi(x, y) = 4 \operatorname{Re}(-i\varphi(x, y)) = 4 \operatorname{Re} \varphi(x, iy) = \varphi(x + iy, x + iy) - \varphi(x - iy, x - iy),$$

ce qui permet d'obtenir $4\varphi(x, y)$ en combinant $\operatorname{Re} \varphi(x, y) + i \operatorname{Im} \varphi(x, y)$.

(e) Il s'agit maintenant d'un vrai produit scalaire, c'est-à-dire que la relation $\langle \tilde{f}, \tilde{f} \rangle = 0$ implique que $\tilde{f} = \tilde{0}$, le zéro de l'espace vectoriel $L^2(\Omega, \mathcal{A}, \mu)$. En effet, pour n'importe quel représentant f de \tilde{f} , on a

$$\langle \tilde{f}, \tilde{f} \rangle = \int_{\Omega} f(\omega) \overline{f(\omega)} d\mu(\omega) = \int_{\Omega} |f(\omega)|^2 d\mu(\omega);$$

quand l'intégrale d'une fonction mesurable positive telle que $|f|^2$ est nulle, cela entraîne qu'elle est nulle presque partout, donc la fonction $\mathbf{0}$ est dans la classe de f , donc $\tilde{f} = \tilde{0}$.

(f) Le symbole de Kronecker $\delta_{i,j}$ est égal à 1 quand $i = j$ et 0 quand $i \neq j$. Commentaire personnel, sans aucune valeur historique garantie : c'est sans doute un ennemi de Kronecker (il en avait) qui a fait en sorte que le nom de Kronecker reste associé à une chose aussi triviale !

(g) Si (ℓ_n) est une suite de Cauchy dans E' , on sait que pour tout $\varepsilon > 0$ il existe N tel que $\|\ell_m - \ell_n\| < \varepsilon$ lorsque $m, n \geq N$. Pour tout vecteur $x \in H$, on a

$$|\ell_m(x) - \ell_n(x)| \leq \|\ell_m - \ell_n\| \|x\| \leq \varepsilon \|x\|$$

lorsque $m, n \geq N$, ce qui permet de voir que la suite scalaire $(\ell_n(x))$ est de Cauchy, donc convergente puisque le corps des scalaires est complet. Si on pose $\ell(x) = \lim_n \ell_n(x)$, pour tout $x \in H$, on montre facilement que ℓ est linéaire, et on obtient en passant à la limite quand $n \rightarrow +\infty$

$$|\ell_m(x) - \ell(x)| \leq \varepsilon \|x\|$$

pour tout x , qui montre que $\|\ell - \ell_m\| \leq \varepsilon$ quand $m \geq N$, donc ℓ est la limite dans E' de la suite (ℓ_n) . On a ainsi montré que E' est complet. On montre de la même façon que l'espace $\mathcal{L}(E, F)$ des applications linéaires continues d'un espace normé E dans un espace normé complet F est complet.

(h) On montre facilement par récurrence que les fonctions de V_n engendrent toutes les indicatrices des intervalles de la forme $[0, j2^{-n}[$, $j = 1, \dots, 2^n$; si on prend $[0, b[$ avec $0 \leq b \leq 1$, on peut trouver une suite $j_n 2^{-n}$ qui converge vers b ; il en résulte que $\mathbf{1}_{[0, j_n 2^{-n}[}$ tend vers $\mathbf{1}_{[0, b[}$ dans $L^2(0, 1)$. Le sous-espace fermé F engendré par le système de Haar contient donc les indicatrices de tous les intervalles $[0, b[$, donc celles de tous les $[a, b[$, donc toutes les fonctions en escalier : il en résulte que $F = L^2(0, 1)$.

(i) La plupart des étudiants a déjà appris qu'un sous-espace vectoriel F de dimension finie est toujours fermé dans un espace normé E : la preuve passe par l'équivalence des normes en dimension finie, qui permet de voir que F est complet pour la norme induite par E , donc fermé dans E . Dans le cas d'un espace de Hilbert, la démonstration par projection orthogonale est plutôt plus courte et plus naturelle.

(j) Puisque H est de dimension infinie, il existe au moins un vecteur $w \neq 0$ dans H , et puisque (x_k) est dense, il existe un entier k tel que $\|w - x_k\| < \frac{1}{2}\|w\|$, ce qui implique que x_k n'est pas nul. On définit k_0 comme le plus petit entier k tel que $x_k \neq 0$ et on pose $v_0 = x_{k_0}$.

Si k_0, \dots, k_{n-1} ont été définis et si on a posé $v_j = x_{k_j}$ pour $0 \leq j < n$, on dit que H , qui est de dimension infinie, ne peut être égal à $F = \text{Vect}(v_0, \dots, v_{n-1})$; le sous-espace F est de dimension finie, donc fermé dans H (corollaire 2.1.9). Si w est un vecteur de H qui n'est pas dans le fermé F , on peut trouver une boule ouverte B qui contient w et ne rencontre pas F ; comme (x_k) est dense, il existe des vecteurs x_k qui sont dans B , donc pas dans F . On peut alors définir k_n comme le plus petit entier k tel que $x_k \notin F$ et on pose $v_n = x_{k_n}$.

Pour tous les entiers $j < k_n$ on a $x_j \in F \subset \text{Vect}(v_0, \dots, v_n)$, et pour $j = k_n$ on a aussi $x_{k_n} = v_n \in \text{Vect}(v_0, \dots, v_n)$. Les vecteurs (v_n) sont indépendants puisque $v_0 \neq 0$ et $v_n \notin \text{Vect}(v_0, \dots, v_{n-1})$, pour tout $n \geq 1$.

(k) Considérons la fonction g sur un intervalle de deux périodes, par exemple $[0, 4\pi]$; sur ce compact, on sait que la fonction g est uniformément continue, donc il existe $\delta > 0$ tel que tel que $|g(x) - g(y)| < \varepsilon/2$ dès que $x, y \in [0, 4\pi]$ vérifient $|x - y| < \delta$; on peut choisir $\delta < 2\pi$; si x_1, y_1 sont deux points quelconques de \mathbb{R} tels que $|x_1 - y_1| < \delta < 2\pi$, on peut toujours trouver un entier k tel que $x = x_1 - 2k\pi$ et $y = y_1 - 2k\pi$ soient tous les deux dans $[0, 4\pi]$: si $x_1 \leq y_1$ par exemple, on choisit k tel que $x = x_1 - 2k\pi \in [0, 2\pi[$; alors si $y = y_1 - 2k\pi$, on a $0 \leq y - x = y_1 - x_1 < 2\pi$ donc

$$0 \leq x \leq y \leq x + 2\pi < 4\pi.$$

On aura $|x - y| = |x_1 - y_1| < \delta$, et $|g(x_1) - g(y_1)| = |g(x) - g(y)| < \varepsilon/2$ par périodicité.

(l) Il s'agit de voir qu'une fonction continue h nulle presque partout sur $[0, 2\pi]$ est en fait nulle *partout*: voir le théorème A.5 de l'annexe.

(m) Considérons, pour un réel $c > 1$ quelconque, l'ensemble mesurable

$$A_c = \{x \in [0, 1] : |f(x)| > c\};$$

on peut trouver une suite (h_n) de fonctions en escalier qui tendent dans L^1 et presque partout vers la fonction $\mathbf{1}_{A_c} \text{sign } f$ (annexe, théorème A.11 et théorème A.3); comme la fonction limite est à valeurs dans $[-1, 1]$, on peut supposer que $|h_n| \leq 1$. D'après l'inégalité (8), on sait que $|\ell(h_m) - \ell(h_n)| \leq \|h_m - h_n\|_1$ tend vers 0. La suite numérique $(\ell(h_n))$ est de Cauchy, donc convergente vers une limite $\xi \in \mathbb{R}$. On a d'après (8)

$$|\ell(h_n)| \leq \|h_n\|_1 \rightarrow \|\mathbf{1}_{A_c} \text{sign } f\|_1 = \int_0^1 \mathbf{1}_{A_c}(t) dt = \lambda(A_c),$$

la mesure de Lebesgue de l'ensemble A_c ; on a donc $|\xi| \leq \lambda(A_c)$. Par ailleurs, la suite $(h_n f)$ tend presque partout vers la limite $(\mathbf{1}_{A_c} \text{sign } f) f = \mathbf{1}_{A_c} |f|$, en étant majorée par la fonction intégrable fixe $|f| \in L^2([0, 1])$, donc

$$c \lambda(A_c) = c \int_0^1 \mathbf{1}_{A_c}(t) dt \leq \int_0^1 \mathbf{1}_{A_c}(t) |f(t)| dt = \lim_n \int_0^1 h_n(t) f(t) dt = \lim_n \ell(h_n) = \xi,$$

et il en résulte que $c \lambda(A_c) \leq \lambda(A_c)$, mais $c > 1$ donc $\lambda(A_c) = 0$; ainsi, l'ensemble A_c est négligeable pour tout $c > 1$, et la fonction $|f|$ est donc bornée par 1 presque partout.

3. Séries de Fourier

3.1. Séries de Fourier dans L^2

Dans la plus grande partie de ce chapitre, on travaillera avec des fonctions ^(a) réelles ou complexes 2π -périodiques sur \mathbb{R} . On munira les intervalles de longueur 2π de la mesure $d\mu(x) = dx/(2\pi)$; on utilisera cette mesure μ pour définir les trois notions suivantes : les *normes L^p périodiques*,

$$\|f\|_p = \left(\int_a^{a+2\pi} |f(x)|^p \frac{dx}{2\pi} \right)^{1/p},$$

lorsque f est 2π -périodique et $|f|^p$ intégrable sur chaque période, $1 \leq p < +\infty$; le résultat ne dépend pas du choix de a ; comme la mesure μ est de masse 1 sur chaque période, on peut comparer les normes L^p : elles sont croissantes avec p , en particulier,

$$\|f\|_1 \leq \|f\|_2 \leq \|f\|_\infty;$$

la deuxième notion est le *produit scalaire périodique*

$$\langle f, g \rangle = \int_a^{a+2\pi} f(x) \overline{g(x)} \frac{dx}{2\pi},$$

lorsque f, g sont 2π -périodiques et $|f|^2, |g|^2$ intégrables sur chaque période; et pour finir la convolution périodique (ce sera la seule qu'on verra dans ce chapitre; on laissera tomber la notation plus lourde $f \underset{\text{per}}{*} g$ utilisée dans les chapitres précédents) : on posera

$$(f * g)(y) = \int_a^{a+2\pi} f(y-x)g(x) \frac{dx}{2\pi},$$

lorsque f, g sont 2π -périodiques et intégrables sur chaque période. On passera souvent d'une fonction f définie sur une seule période, par exemple f définie sur $[-\pi, \pi]$, à la fonction \tilde{f} définie sur \mathbb{R} , 2π -périodique et telle que $\tilde{f}(x) = f(x)$ pour tout $x \in [-\pi, \pi]$.

On a introduit au chapitre 2 les fonctions $(e_n)_{n \in \mathbb{Z}}$ qui sont définies par $e_n(x) = e^{inx}$. Lorsque $f \in L^1([0, 2\pi], \mu)$, on peut considérer les coefficients de Fourier de f , pour tout $n \in \mathbb{Z}$,

$$c_n(f) = \int_0^{2\pi} f(x) e^{-inx} \frac{dx}{2\pi},$$

et pour $n \geq 0$, les *sommes de Fourier*

$$S_n f = \sum_{k=-n}^n c_k(f) e_k.$$

Dans l'espace de Hilbert $L^2([0, 2\pi], \mu)$, la suite $(e_n)_{n \in \mathbb{Z}}$ est orthonormée,

$$\langle e_m, e_n \rangle = \int_0^{2\pi} e^{imx} e^{-inx} \frac{dx}{2\pi} = \int_0^{2\pi} e^{i(m-n)x} \frac{dx}{2\pi} = \delta_{m,n}$$

(symbole de Kronecker) et on a montré au chapitre 2 que cette suite de fonctions est une base hilbertienne de $L^2([0, 2\pi], \mu)$. Quand $f \in L^2([0, 2\pi], \mu)$, on peut écrire aussi

$$c_n(f) = \langle f, e_n \rangle, \quad S_n f = \sum_{k=-n}^n \langle f, e_k \rangle e_k,$$

et on voit ainsi que $S_n f$ est la projection orthogonale de f sur $\text{Vect}(e_k : |k| \leq n)$. Puisque que les fonctions $(e^{inx})_{n \in \mathbb{Z}}$ forment une base hilbertienne de $L^2([0, 2\pi], \mu)$, on en déduit que

$$\|f - S_n f\|_2^2 = \int_0^{2\pi} \left| f(x) - (S_n f)(x) \right|^2 \frac{dx}{2\pi}$$

tend vers 0 quand $n \rightarrow +\infty$, c'est-à-dire qu'on a une représentation de toute fonction $f \in L^2([0, 2\pi], \mu)$ par sa série de Fourier

$$f = \sum_{n \in \mathbb{Z}} \langle f, e_n \rangle e_n,$$

la somme de la série étant prise au sens de l'espace L^2 . Ceci ne permet pas d'affirmer que la série numérique

$$\sum_{n \in \mathbb{Z}} c_n(f) e^{inx}$$

converge vers $f(x)$ quand on donne une valeur particulière de x , mais on va voir un cas simple où cela est possible.

Proposition 3.1.1. *Si f est une fonction continue 2π -périodique et si*

$$\sum_{n \in \mathbb{Z}} |c_n(f)| < +\infty,$$

alors $f(x)$ est égal pour tout x à la somme de la série de Fourier,

$$\forall x \in \mathbb{R}, \quad f(x) = \sum_{n \in \mathbb{Z}} c_n(f) e^{inx}.$$

Preuve. — D'après l'hypothèse de la proposition, on peut poser pour tout $x \in \mathbb{R}$

$$g(x) = \sum_{n \in \mathbb{Z}} c_n(f) e^{inx};$$

la série de fonctions ci-dessus est normalement convergente, donc g est continue. La suite des fonctions continues $(S_n f)$ définies par

$$(S_n f)(x) = \sum_{k=-n}^n c_k(f) e^{ikx}$$

converge uniformément vers g , et la même suite $(S_n f)$ est convergente dans $L^2([0, 2\pi])$ vers la fonction f . Il en résulte que $g = f$ partout, par les résultats d'intégration qui sont

rappelés dans l'annexe : corollaire A.4 et théorème A.5. Puisque $f(x) = g(x)$ pour tout x , on peut bien affirmer que

$$\forall x \in \mathbb{R}, \quad f(x) = \sum_{n \in \mathbb{Z}} c_n(f) e^{inx}.$$

Exercice. La fonction de Bessel J_0 , qui est définie par la formule

$$\forall x \in \mathbb{R}, \quad J_0(x) = \int_{-\pi}^{\pi} e^{ix \sin \theta} \frac{d\theta}{2\pi}$$

est développable en série entière $\sum_{n=0}^{+\infty} a_n x^n$, on va identifier les coefficients (a_n) de cette série entière. Pour x réel fixé considérons la fonction 2π -périodique f_x définie par

$$\forall \theta \in \mathbb{R}, \quad f_x(\theta) = \exp\left(\frac{x}{2} e^{i\theta}\right).$$

On voit que

$$\overline{f_{-x}(\theta)} = \overline{\exp\left(-\frac{x}{2} e^{i\theta}\right)} = \exp\left(-\frac{x}{2} e^{-i\theta}\right) = \exp\left(-\frac{x}{2} e^{-i\theta}\right)$$

donc $f_x(\theta) \overline{f_{-x}(\theta)} = e^{ix \sin \theta}$ et

$$\forall x \in \mathbb{R}, \quad J_0(x) = \int_{-\pi}^{\pi} f_x(\theta) \overline{f_{-x}(\theta)} \frac{d\theta}{2\pi} = \langle f_x, f_{-x} \rangle.$$

En développant l'exponentielle dans f_x , on voit que

$$f_x(\theta) = \sum_{n=0}^{+\infty} \frac{x^n}{2^n n!} e^{in\theta}.$$

Cette série converge normalement, donc uniformément, donc en norme L^2 : on a

$$f_x = \sum_{n=0}^{+\infty} \frac{x^n}{2^n n!} e_n$$

au sens de L^2 , ce qui permet, par l'unicité des coefficients dans une base hilbertienne, de montrer que le développement ci-dessus est le développement de Fourier de f_x ,

$$c_n(f_x) = \langle f_x, e_n \rangle = \frac{x^n}{2^n n!}$$

pour tout $n \geq 0$, et $c_n(f_x) = 0$ pour $n < 0$. De même, en remplaçant x par $-x$

$$c_n(f_{-x}) = \frac{(-x)^n}{2^n n!} \text{ pour } n \geq 0, \text{ et } c_n(f_{-x}) = 0 \text{ pour } n < 0.$$

On rappelle le calcul du produit scalaire dans une base orthonormée : si v, w sont deux vecteurs d'un espace de Hilbert qui a une base hilbertienne $(e_n)_{n \geq 0}$, on a (formule (S) du chapitre 2)

$$\langle v, w \rangle = \sum_{n=0}^{+\infty} \langle v, e_n \rangle \overline{\langle w, e_n \rangle}.$$

Ici on utilise la base hilbertienne $e_n(x) = e^{inx}$, $n \in \mathbb{Z}$. Alors

$$J_0(x) = \langle f_x, f_{-x} \rangle = \sum_{n \in \mathbb{Z}} c_n(f_x) \overline{c_n(f_{-x})} = \sum_{n=0}^{+\infty} \frac{x^n}{2^n n!} \frac{(-x)^n}{2^n n!} = \sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n}}{2^{2n} (n!)^2}.$$

On peut introduire les autres fonctions de Bessel (J_n) d'indice n entier en disant que $J_n(x)$ est le coefficient de Fourier d'indice n de la fonction 2π -périodique $g_x : \theta \rightarrow e^{ix \sin \theta}$. On va calculer par exemple

$$J_1(x) = c_1(g_x) = \int_{-\pi}^{\pi} e^{ix \sin(\theta)} e^{-i\theta} \frac{d\theta}{2\pi} = \int_{-\pi}^{\pi} f_x(\theta) \overline{f_{-x}(\theta) e_1(x)} \frac{d\theta}{2\pi} = \langle f_x e_{-1}, f_{-x} \rangle.$$

On a maintenant

$$f_x(\theta) e_{-1}(\theta) = \left(\sum_{n=0}^{+\infty} \frac{x^n}{2^n n!} e^{in\theta} \right) e^{-i\theta} = \sum_{m=-1}^{+\infty} \frac{x^{m+1}}{2^{m+1} (m+1)!} e^{im\theta}.$$

On voit donc que les coefficients de Fourier de $f_x e_{-1}$ sont égaux à

$$c_m(f_x e_{-1}) = \frac{x^{m+1}}{2^{m+1} (m+1)!} \text{ pour } m \geq -1,$$

et $c_m(f_x e_{-1}) = 0$ pour tout $m < -1$; donc

$$\begin{aligned} J_1(x) &= \langle f_x e_{-1}, f_{-x} \rangle = \sum_{m \in \mathbb{Z}} c_m(f_x e_{-1}) \overline{c_m(f_{-x})} = \sum_{m \geq 0} c_m(f_x e_{-1}) \overline{c_m(f_{-x})} \\ &= \sum_{m=0}^{+\infty} \frac{x^{m+1}}{2^{m+1} (m+1)!} \frac{(-x)^m}{2^m m!} = \sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n+1}}{2^{2n+1} n! (n+1)!}. \end{aligned}$$

On pourra trouver de la même façon pour tout entier $k \geq 0$

$$\forall x \in \mathbb{R}, \quad J_k(x) = \sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n+k}}{2^{2n+k} n! (n+k)!}.$$

La preuve précédente permet de voir que ces séries entières convergent pour tout x réel : elles sont donc de rayon de convergence infini, ce qu'on retrouverait facilement par les critères usuels pour les séries entières (critère de d'Alembert par exemple).

Fonction de classe C^1 par morceaux

Théorème 3.1.2. *Soit f continue 2π -périodique, et de classe C^1 par morceaux ; alors les coefficients de Fourier de f sont absolument sommables, et $f(x)$ est égal pour tout x à la somme de la série de Fourier,*

$$\forall x \in \mathbb{R}, \quad f(x) = \sum_{n \in \mathbb{Z}} c_n(f) e^{inx}.$$

Preuve. — Posons $g(t) = f(t) e^{-int}$ pour un $n \in \mathbb{Z}$ quelconque ; on voit que g est 2π -périodique continue et de classe C^1 par morceaux sur $[0, 2\pi]$, donc d'après le lemme 1.1.2 on a

$$0 = g(2\pi) - g(0) = \int_0^{2\pi} (f'(t) e^{-int} - in f(t) e^{-int}) dt.$$

On en déduit que

$$\forall n \in \mathbb{Z}, \quad c_n(f') = in c_n(f).$$

La fonction f' est (b) bornée, donc de carré intégrable sur chaque période. Il en résulte que $\sum_{n \in \mathbb{Z}} |c_n(f')|^2 < +\infty$, et par Cauchy-Schwarz dans $\ell^2(\mathbb{N})$

$$\sum_{n=1}^{+\infty} |c_n(f)| = \sum_{n=1}^{+\infty} \frac{1}{n} |c_n(f')| \leq \left(\sum_{n \geq 1} \frac{1}{n^2} \right)^{1/2} \left(\sum_{n \geq 1} |c_n(f')|^2 \right)^{1/2} < +\infty.$$

On procède de même pour $n < 0$ et on obtient ainsi que $\sum_{n \in \mathbb{Z}} |c_n(f)| < +\infty$. Il suffit ensuite d'appliquer la proposition 3.1.1.

Exemple. Pour chaque x fixé, considérons la fonction g_x définie par

$$\forall \theta \in \mathbb{R}, \quad g_x(\theta) = e^{ix \sin \theta}.$$

Il est clair que cette fonction g_x est 2π -périodique, de classe C^1 (en fait de classe C^∞), donc ses coefficients de Fourier sont absolument sommables et $g_x(\theta)$ est la somme de la série de Fourier. On a défini les fonctions de Bessel (J_n) d'indice entier $n \in \mathbb{Z}$ en considérant les coefficients de Fourier de la fonction g_x ,

$$J_n(x) = \int_0^{2\pi} e^{ix \sin \theta - in\theta} \frac{d\theta}{2\pi} = c_n(g_x).$$

Il est facile de montrer que J_n est, comme J_0 , la somme d'une série entière. Comme la fonction g_x est de classe C^1 on sait que son développement de Fourier est absolument convergent,

$$\sum_{n \in \mathbb{Z}} |c_n(g_x)| = \sum_{n \in \mathbb{Z}} |J_n(x)| < +\infty,$$

et on a pour tout x et tout θ

$$(1) \quad e^{ix \sin \theta} = \sum_{n \in \mathbb{Z}} J_n(x) e^{in\theta}.$$

On peut retrouver les développements en série entière des fonctions (J_n) par un argument légèrement différent de celui qu'on a employé pour J_0 . On commence de la même façon,

$$e^{ix \sin \theta} = \exp\left(x(e^{i\theta} - e^{-i\theta})/2\right) = \exp(x e^{i\theta}/2) \exp(-x e^{-i\theta}/2),$$

et on développe comme avant en utilisant la série entière de e^z , $z \in \mathbb{C}$

$$e^{ix \sin \theta} = \left(\sum_{m=0}^{+\infty} \frac{(x/2)^m}{m!} e^{im\theta} \right) \left(\sum_{k=0}^{+\infty} \frac{(-x/2)^k}{k!} e^{-ik\theta} \right);$$

mais maintenant, on va procéder comme dans le théorème de Fubini : on a deux séries absolument convergentes, ce qui justifie les regroupements et interversions qui suivent ; en posant $n = m - k$, on trouve

$$e^{ix \sin \theta} = \sum_{m,k=0}^{+\infty} (-1)^k \frac{(x/2)^{m+k}}{m! k!} e^{i(m-k)\theta} = \sum_{n \in \mathbb{Z}} \left(\sum_{k=\max(0,-n)}^{+\infty} (-1)^k \frac{(x/2)^{2k+n}}{k! (k+n)!} \right) e^{in\theta}.$$

Pour $n \geq 0$, on trouve par identification

$$J_n(x) = \sum_{k=0}^{+\infty} (-1)^k \frac{(x/2)^{2k+n}}{k! (k+n)!}.$$

Pour $n < 0$, on trouve en posant $m = n + k$

$$J_n(x) = \sum_{m=0}^{+\infty} (-1)^{m-n} \frac{(x/2)^{2m-n}}{m! (m-n)!} = (-1)^n \sum_{m=0}^{+\infty} (-1)^m \frac{(x/2)^{2m+|n|}}{m! (m+|n|)!} = (-1)^n J_{-n}(x).$$

Bessel et FM

En modulation de fréquence, le signal $s_m(t)$ (musical par exemple ; on le supposera à valeurs réelles) qu'on veut transmettre est d'abord intégré (à partir d'une certaine origine de temps, disons $t = 0$)

$$S_m(t) = \int_0^t s_m(u) du$$

et il est transmis en modifiant un signal porteur $s_p(t) = e^{i\omega_p t}$ de la façon suivante : c'est la fréquence du signal porteur qui est modifiée, en introduisant le signal *modulé en fréquence*

$$(2) \quad t \rightarrow \exp(i\omega_p t + i\kappa S_m(t)),$$

où κ est un coefficient convenablement choisi. Cette expression est en général très difficile à étudier mathématiquement. Examinons un cas très particulier, celui d'une émission musicale assez ennuyeuse qui transmettrait un son constant,

$$s_m : t \rightarrow \cos(\omega_m t)$$

où ω_m correspond à la fréquence du son musical, très inférieure à la fréquence $\omega_p/(2\pi)$ de la porteuse (par exemple : $\omega_m/(2\pi) = 440$ Hz, $\omega_p/(2\pi) = 101.1$ MHz). Dans ce cas on a $S_m(t) = \omega_m^{-1} \sin(\omega_m t)$ et on obtient pour le signal (2) l'expression

$$t \rightarrow \exp(i\omega_p t + ik \sin(\omega_m t)) = e^{i\omega_p t} e^{ik \sin(\omega_m t)},$$

où on a posé $k = \kappa \omega_m^{-1}$. On reconnaît l'expression qu'on a développée avec les fonctions de Bessel à l'équation (1),

$$e^{i\omega_p t} \sum_{n \in \mathbb{Z}} J_n(k) e^{in\omega_m t} = \sum_{n \in \mathbb{Z}} J_n(k) e^{i(\omega_p + n\omega_m)t}.$$

On voit que le signal (2) est une combinaison de signaux périodiques (théoriquement infinie, mais les coefficients $J_n(k)$ décroissent assez vite); le signal (2) occupe les fréquences $(\omega_p + n\omega_m)/(2\pi)$, $n \in \mathbb{Z}$; dans la pratique, il est important de pouvoir limiter les valeurs de n vraiment utilisées, pour rester à l'intérieur de la plage de fréquences attribuée à la station!

Revenons sur le terme «modulation de fréquence». Dans l'équation de la porteuse, l'expression est $e^{if(t)}$ avec $f(t) = \omega_p t$ et on obtient la pulsation ω_p de la porteuse en dérivant f ,

$$f'(t) = \omega_p.$$

Dans l'équation du signal (2), on a

$$f(t) = \omega_p t + k \sin(\omega_m t)$$

dont la dérivée est

$$f'(t) = \omega_p + \kappa \cos(\omega_m t);$$

cette expression représente la *pulsation instantanée* : si on découpe un intervalle de temps d'une seconde en dix mille fractions d'un dix millième de seconde, le nombre des oscillations du signal (2) n'est pas le même sur tous les petits intervalles de temps; dans certains de ces petits intervalles I , on aura $\cos(\omega_m t) \simeq 1$ pour tout $t \in I$, alors que $\cos(\omega_m t) \simeq -1$ pour d'autres intervalles; on peut dire que la «fréquence varie» entre les valeurs $(\omega_p + \kappa)/(2\pi)$ et $(\omega_p - \kappa)/(2\pi)$: la fréquence est «modulée», et on retrouve l'information musicale ω_m en examinant la variation de la fréquence.

3.2. Convergence ponctuelle des séries de Fourier

On va s'intéresser à des aspects de convergence ponctuelle de la série de Fourier d'une fonction f : la question de base est de savoir si en un point x donné, la série de Fourier $\sum_{n \in \mathbb{Z}} c_n(f) e^{inx}$ converge, et si la somme de la série est bien égale à la valeur $f(x)$.

Lemme 3.2.1 de Riemann-Lebesgue. *Si f est une fonction intégrable sur $[0, 2\pi]$, on a*

$$\lim_{|n| \rightarrow +\infty} c_n(f) = 0.$$

Preuve. — On applique le lemme de Riemann-Lebesgue (proposition 1.1.1) connu pour la transformation de Fourier : on sait que $\widehat{g}(y)$ tend vers 0 lorsque $|y| \rightarrow +\infty$, pour toute fonction g intégrable sur \mathbb{R} . Si f est une fonction intégrable sur $[0, 2\pi]$ et si g est la fonction sur \mathbb{R} qui est égale à f dans $[0, 2\pi]$ et qui est nulle en dehors de $[0, 2\pi]$, on voit que pour tout $n \in \mathbb{Z}$ on a

$$\widehat{g}(n) = \int_{\mathbb{R}} g(x) e^{-inx} dx = \int_0^{2\pi} f(x) e^{-inx} dx = 2\pi c_n(f);$$

il en résulte que $\lim_{|n| \rightarrow +\infty} c_n(f) = 0$.

Théorème 3.2.2. Soient f une fonction 2π -périodique mesurable, ℓ une valeur scalaire et x_0 un point tels que

$$\int_{-\pi}^{\pi} \left| \frac{f(x_0 - t) - \ell}{t} \right| dt < +\infty;$$

alors la série de Fourier de f converge au point x_0 et

$$\sum_{n \in \mathbb{Z}} c_n(f) e^{inx_0} = \ell.$$

Preuve. — L'hypothèse du théorème entraîne que la fonction f est intégrable sur la période $[x_0 - \pi, x_0 + \pi]$, puisque

$$|f(x_0 - t) - \ell| \leq \pi \left| \frac{f(x_0 - t) - \ell}{t} \right|$$

lorsque $|t| \leq \pi$, ce qui permet de considérer ses coefficients de Fourier $c_n(f)$. On a pour tous les entiers $m, n \geq 0$

$$\begin{aligned} \sum_{k=-m}^n c_k(f) e^{ikx_0} &= \sum_{k=-m}^n \left(\int_{-\pi}^{\pi} f(s) e^{-iks} \frac{ds}{2\pi} \right) e^{ikx_0} \\ &= \sum_{k=-m}^n \int_{-\pi}^{\pi} f(s) e^{ik(x_0-s)} \frac{ds}{2\pi} = \int_{-\pi}^{\pi} \left(\sum_{k=-m}^n e^{ikt} \right) f(x_0 - t) \frac{dt}{2\pi}; \end{aligned}$$

comme l'entier $k = 0$ est entre $-m$ et n on a aussi

$$\int_{-\pi}^{\pi} \left(\sum_{k=-m}^n e^{ikt} \right) \frac{dt}{2\pi} = \sum_{k=-m}^n \int_{-\pi}^{\pi} e^{ikt} \frac{dt}{2\pi} = 1.$$

Il en résulte que

$$(3) \quad \left(\sum_{k=-m}^n c_k(f) e^{ikx_0} \right) - \ell = \int_{-\pi}^{\pi} \left(\sum_{k=-m}^n e^{ikt} \right) (f(x_0 - t) - \ell) \frac{dt}{2\pi}.$$

Un calcul de somme de progression géométrique ^(c) donne

$$\sum_{k=-m}^n e^{ikt} = \frac{e^{-imt} - e^{i(n+1)t}}{1 - e^{it}}.$$

Introduisons la fonction périodique g définie pour $t \notin 2\pi\mathbb{Z}$ par

$$g(t) = \frac{f(x_0 - t) - \ell}{1 - e^{it}};$$

quand $|t| \leq \pi$, on a ^(d)

$$|1 - e^{it}| = 2 |\sin(t/2)| \geq \frac{2|t|}{\pi},$$

ce qui montre que l'hypothèse du théorème implique que g est intégrable sur $[-\pi, \pi]$. L'équation (3) devient

$$\begin{aligned} \left(\sum_{k=-m}^n c_k(f) e^{ikx_0} \right) - \ell &= \int_{-\pi}^{\pi} (e^{-imt} - e^{i(n+1)t}) \frac{f(x_0 - t) - \ell}{1 - e^{it}} \frac{dt}{2\pi} \\ &= \int_{-\pi}^{\pi} g(t) e^{-imt} \frac{dt}{2\pi} - \int_{-\pi}^{\pi} g(t) e^{i(n+1)t} \frac{dt}{2\pi} = c_m(g) - c_{-n-1}(g). \end{aligned}$$

Comme la fonction $t \rightarrow g(t)$ est intégrable sur $[-\pi, \pi]$, ses coefficients de Fourier tendent vers 0 d'après le lemme 3.2.1. Si on fait tendre n tout seul vers $+\infty$, en gardant par exemple $m = 0$, on constate que

$$\left(\sum_{k=0}^n c_k(f) e^{ikx_0} \right) - \ell$$

tend vers $c_0(g)$, ce qui montre que la partie positive de la série de Fourier de f au point x_0 converge ; on voit de même que la partie négative converge, et quand on fait tendre m et n vers $+\infty$ on obtient le résultat annoncé,

$$\ell = \sum_{k=-\infty}^{+\infty} c_k(f) e^{ikx_0}.$$

Corollaire 3.2.3. *Soit f une fonction 2π -périodique et intégrable sur chaque période ; en tout point x_0 tel que*

$$\frac{f(x) - f(x_0)}{x - x_0}$$

reste borné pour x dans un voisinage de x_0 , la série de Fourier de f converge et

$$f(x_0) = \sum_{n \in \mathbb{Z}} c_n(f) e^{inx_0}.$$

Preuve. — On applique le théorème précédent avec $\ell = f(x_0)$. Par hypothèse il existe M et $\varepsilon > 0$ tels que

$$\left| \frac{f(x_0 - t) - \ell}{t} \right| = \left| \frac{f(x_0 - t) - f(x_0)}{t} \right| \leq M$$

lorsque $|t| < \varepsilon$. On a donc

$$\begin{aligned} \int_{-\pi}^{\pi} \left| \frac{f(x_0 - t) - \ell}{t} \right| dt &\leq \int_{-\varepsilon}^{\varepsilon} \left| \frac{f(x_0 - t) - f(x_0)}{t} \right| dt + \int_{\varepsilon \leq |t| \leq \pi} \left| \frac{f(x_0 - t) - f(x_0)}{\varepsilon} \right| dt \\ &\leq 2M\varepsilon + \varepsilon^{-1} \left(2\pi |f(x_0)| + \int_{-\pi}^{\pi} |f(x_0 - t)| dt \right) < +\infty ; \end{aligned}$$

le résultat du corollaire 3.2.3 est donc obtenu en appliquant le théorème 3.2.2.

Exemple 3.2.4. Soit f la fonction impaire 2π -périodique définie par $f(x) = 1 - x/\pi$ lorsque $0 < x \leq \pi$; pour tout point x tel que $0 < x < \pi$, on a

$$(4) \quad 1 - \frac{x}{\pi} = f(x) = \frac{2}{\pi} \sum_{n=1}^{+\infty} \frac{\sin(nx)}{n}.$$

On va appliquer le corollaire 3.2.3 avec $x_0 = x$. Si t est assez petit pour que $0 < x \pm t < \pi$, c'est-à-dire si $|t| < \min(x, \pi - x)$, on a

$$|f(x+t) - f(x)| = \left| (1 - (x+t)/\pi) - (1 - x/\pi) \right| = |t|/\pi;$$

il en résulte que la fonction $t \rightarrow (f(x+t) - f(x))/t$ est bornée pour t dans un voisinage de 0 et le corollaire précédent s'applique : on a donc

$$(5) \quad 1 - \frac{x}{\pi} = \sum_{n \in \mathbb{Z}} c_n(f) e^{inx}.$$

Puisque f est impaire, $c_0(f) = 0$ et pour $n \neq 0$

$$2\pi c_n(f) = -i \int_{-\pi}^{\pi} f(t) \sin(nt) dt = -2i \int_0^{\pi} (1 - t/\pi) \sin(nt) dt;$$

par intégration par parties

$$\pi c_n(f) = i \left(\left[(1 - t/\pi) \frac{\cos(nt)}{n} \right]_0^{\pi} + \int_0^{\pi} \frac{\cos(nt)}{\pi n} dt \right) = -\frac{i}{n}.$$

En regroupant $n > 0$ et $-n$,

$$c_n(f) e^{inx} + c_{-n}(f) e^{-inx} = \frac{-i}{n\pi} e^{inx} + \frac{i}{n\pi} e^{-inx} = \frac{2}{n\pi} \sin(nx).$$

On voit donc que la série de Fourier complexe de (5) se transforme en la relation (4). En intégrant la relation (4) en prenant les primitives nulles en 0 on obtient pour $0 \leq x \leq \pi$

$$x - \frac{x^2}{2\pi} = \frac{2}{\pi} \sum_{n=1}^{+\infty} \frac{1 - \cos(nx)}{n^2}.$$

La valeur en $x = \pi$ permet de retrouver la relation $\sum_{n=1}^{+\infty} n^{-2} = \pi^2/6$, d'où

$$\sum_{n=1}^{+\infty} \frac{\cos(nx)}{n^2} = \frac{x^2}{4} - \frac{\pi x}{2} + \frac{\pi^2}{6}.$$

En continuant d'intégrer, on peut trouver les valeurs de $\sum n^{-4}$, $\sum n^{-6}$, $\sum n^{-8}$, etc., qui sont des valeurs particulières de la fonction zêta de Riemann

$$\forall s > 1, \quad \zeta(s) = \sum_{n=1}^{+\infty} \frac{1}{n^s}.$$

Corollaire 3.2.5. Soit f une fonction 2π -périodique et lipschitzienne; en tout point x_0 , la série de Fourier de f converge et (e)

$$f(x_0) = \sum_{n \in \mathbb{Z}} c_n(f) e^{inx_0}.$$

Preuve. — C'est évidemment un cas particulier du corollaire 3.2.3, puisqu'ici la fonction

$$x \rightarrow \frac{f(x) - f(x_0)}{x - x_0}$$

est bornée pour $x \in \mathbb{R}$ par la constante de Lipschitz de f , pour tout x_0 .

On définit les *sommes de Fourier* d'une fonction 2π -périodique f , intégrable sur chaque période, en procédant à la sommation de façon symétrique :

$$(S_n f)(x) = \sum_{k=-n}^n c_k(f) e^{ikx}.$$

Corollaire 3.2.6 (théorème de Dirichlet-Dini). Soient f une fonction 2π -périodique mesurable, ℓ_+, ℓ_- deux valeurs scalaires et x_0 un point tels que

$$\int_0^\pi \frac{|f(x_0 - t) - \ell_-|}{t} dt + \int_0^\pi \frac{|f(x_0 + t) - \ell_+|}{t} dt < +\infty;$$

alors les sommes de Fourier de f au point x_0 convergent et

$$\lim_n (S_n f)(x_0) = \frac{\ell_+ + \ell_-}{2}.$$

Preuve. — Pour simplifier, prenons $x_0 = 0$ et introduisons la fonction F définie par $F(t) = f(-t) + f(t)$, qui est 2π -périodique sur \mathbb{R} . Cette fonction F est paire et elle vérifie l'hypothèse du théorème 3.2.2 au point 0, avec la valeur $\ell = \ell_+ + \ell_-$: l'intégrale

$$\begin{aligned} \int_{-\pi}^\pi \left| \frac{F(t) - \ell}{t} \right| dt &= 2 \int_0^\pi \frac{|F(t) - \ell|}{t} dt = 2 \int_0^\pi \frac{|f(-t) + f(t) - \ell_- - \ell_+|}{t} dt \\ &\leq 2 \int_0^\pi \frac{|f(-t) - \ell_-|}{t} dt + 2 \int_0^\pi \frac{|f(t) - \ell_+|}{t} dt < +\infty \end{aligned}$$

est finie par hypothèse, donc la série de Fourier de F au point 0 converge vers ℓ d'après le théorème 3.2.2. Mais

$$\begin{aligned} c_k(F) &= \int_{-\pi}^\pi F(t) e^{-ikt} \frac{dt}{2\pi} = \int_{-\pi}^\pi f(-t) e^{-ikt} \frac{dt}{2\pi} + \int_{-\pi}^\pi f(t) e^{-ikt} \frac{dt}{2\pi} \\ &= \int_{-\pi}^\pi f(u) e^{iku} \frac{du}{2\pi} + \int_{-\pi}^\pi f(u) e^{-iku} \frac{du}{2\pi} = c_k(f) + c_{-k}(f). \end{aligned}$$

En faisant la somme de $k = -n$ à n , chaque terme $c_k(f)$ sera sommé deux fois, donc

$$\sum_{k=-n}^n c_k(F) = 2(S_n f)(0)$$

qui converge vers ℓ quand $n \rightarrow +\infty$, d'après le théorème 3.2.2.

Remarque. Dans la situation du théorème de Dirichlet on ne peut pas dire sans précaution que la série de Fourier converge au point x_0 : en effet, il est possible que la série considérée d'un seul côté

$$\sum_{n=0}^{+\infty} c_n(f) e^{inx_0}$$

soit divergente, mais la série symétrisée

$$c_0(f) + \sum_{n=1}^{+\infty} (c_n(f) e^{inx_0} + c_{-n}(f) e^{-inx_0}),$$

elle, est bien convergente.

Exemple. Le corollaire précédent s'applique à la fonction de l'exemple 3.2.4, mais le résultat est un peu décevant : on a $f(x) = 1 - x/\pi$ lorsque $0 < x \leq \pi$; les limites à droite et à gauche au point $x = 0$ sont 1 et -1 , leur demi-somme est nulle et par conséquent la valeur de la série de Fourier

$$\frac{2}{\pi} \sum_{n=1}^{+\infty} \frac{\sin(nx)}{n}$$

en $x = 0$ est égale à

$$0 = \frac{\ell_+ + \ell_-}{2}$$

ce qui n'est pas très malin !

Un meilleur exemple est fourni par la fonction f définie par $f(x) = e^{ax}$ pour $|x| < \pi$, où $a \in \mathbb{C}$ n'est pas élément de $i\mathbb{Z}$. On a pour tout $n \in \mathbb{Z}$

$$c_n(f) = \int_{-\pi}^{\pi} e^{ax-inx} \frac{dx}{2\pi} = (-1)^n \frac{e^{a\pi} - e^{-a\pi}}{2\pi(a - in)}.$$

Le théorème de Dirichlet appliqué à la fonction f au point $x = \pi$ donne des relations intéressantes (exercice) : on obtient une représentation en série de la demi-somme des limites à droite et à gauche,

$$\frac{\ell_+ + \ell_-}{2} = \frac{e^{a\pi} + e^{-a\pi}}{2} = \operatorname{ch}(a\pi) = \frac{\operatorname{sh}(a\pi)}{\pi} \lim_n \sum_{k=-n}^n \frac{1}{a - ik}.$$

Pour $a \notin i\mathbb{Z}$, on en déduit que

$$\pi \operatorname{coth}(a\pi) = \lim_n \left(\sum_{k=-n}^n \frac{1}{a - ik} \right) = \frac{1}{a} + \sum_{k=1}^{+\infty} \frac{2a}{a^2 + k^2}.$$

Si on pose $a = ix$, $x \notin \mathbb{Z}$, on obtient (f) un développement fameux, dû à Euler

$$\pi \operatorname{cotan}(\pi x) = \lim_n \left(\sum_{k=-n}^n \frac{1}{x - k} \right) = \frac{1}{x} + \sum_{k=1}^{+\infty} \frac{2x}{x^2 - k^2}.$$

Continuons avec Euler : on observe que $x \rightarrow \ln((\sin \pi x)/(\pi x))$ admet pour dérivée $\pi \operatorname{cotan}(\pi x) - 1/x$, et se prolonge par 0 quand $x \rightarrow 0$; on va donc intégrer l'égalité

$$\pi \operatorname{cotan}(\pi x) - \frac{1}{x} = \sum_{k=1}^{+\infty} \frac{2x}{x^2 - k^2} = \sum_{k=1}^{+\infty} \frac{-2x}{k^2 - x^2}$$

terme à terme, en prenant pour chaque terme de la série sa primitive sur $[0, 1[$ nulle en 0, à savoir $\ln(k^2 - x^2) - \ln(k^2) = \ln(1 - x^2/k^2)$. On obtient

$$\ln\left(\frac{\sin(\pi x)}{\pi x}\right) = \sum_{k=1}^{+\infty} \ln\left(1 - \frac{x^2}{k^2}\right).$$

En prenant l'exponentielle, on obtient la représentation de sin en *produit infini* (g)

$$\frac{\sin(\pi x)}{\pi x} = \prod_{k=1}^{+\infty} \left(1 - \frac{x^2}{k^2}\right).$$

Noyau de Dirichlet

Si on note $e_n(x) = e^{inx}$, on sait que la convolution périodique avec la fonction e_n d'une fonction 2π -périodique f , intégrable sur chaque période, est égale à

$$f * e_n = c_n(f) e_n$$

pour tout $n \in \mathbb{Z}$; par conséquent, la somme $S_n f = \sum_{k=-n}^n c_k(f) e_k$ est obtenue (voir le lemme 2.2.3 et la formule (6) avant ce lemme) par convolution de f avec la fonction

$$D_n = \sum_{k=-n}^n e_k.$$

Cette fonction D_n est appelée *noyau de Dirichlet*. Il est clair que

$$\int_0^{2\pi} D_n(x) \frac{dx}{2\pi} = 1$$

puisque les e_k sont d'intégrale nulle pour tout $k \neq 0$, alors que la fonction $e_0 = \mathbf{1}$ donne la valeur 1. Calculons explicitement la valeur de $D_n(x)$

$$2i \sin(x/2) D_n(x) = \sum_{k=-n}^n (e^{i(k+\frac{1}{2})x} - e^{i(k-\frac{1}{2})x}) = e^{i(n+\frac{1}{2})x} - e^{-i(n+\frac{1}{2})x}$$

donc

$$D_n(x) = \frac{\sin[(n + \frac{1}{2})x]}{\sin(x/2)}.$$

Noyau de Fejér

Il existe des fonctions continues 2π -périodiques f dont la série de Fourier diverge en certains points : pour un certain point t_0 , la suite $(S_n f)(t_0)$ ne converge pas (par translation, on peut se ramener au cas où $t_0 = 0$). Il n'est pas facile d'exhiber ce phénomène désagréable ; cela a été fait vers 1880 (c'est-à-dire longtemps après les débuts de la théorie des séries de Fourier), et peu de temps après, Fejér a trouvé un moyen pour atténuer cette difficulté, moyen que nous allons expliquer maintenant : quand une suite numérique (x_n) est convergente vers une limite ℓ , la suite des sommes de Cesàro, qui sont les moyennes de la suite (x_n) ,

$$y_n = \frac{1}{n+1}(x_0 + \cdots + x_n)$$

converge vers la même limite ℓ . D'un autre côté, il existe des suites (x_n) qui ne convergent pas, mais telles que la suite des sommes de Cesàro (y_n) converge : on dit alors que (x_n) converge au sens de Cesàro. Par exemple, la suite $x_n = (-1)^n$ n'est pas convergente, mais $|y_n| \leq 1/(n+1)$, donc (y_n) tend vers 0.

Lorsque la suite des sommes de Fourier $(S_n f)(t_0)$ ne converge pas, on peut se demander si elle converge au sens plus faible de Cesàro ; Fejér a montré que c'est vrai si f est continue. On introduit donc les *sommes de Fejér* $(\sigma_n f)$, qui sont les moyennes des sommes de Fourier,

$$(\sigma_n f)(x) = \frac{1}{n+1} \left((S_0 f)(x) + (S_1 f)(x) + \cdots + (S_n f)(x) \right).$$

Comme on a vu que

$$(\mathbf{S}_k f)(x) = (f * \mathbf{D}_k)(x)$$

(convolution périodique), on déduit

$$(\sigma_n f)(x) = \frac{1}{n+1} \left((f * \mathbf{D}_0)(x) + (f * \mathbf{D}_1)(x) + \cdots + (f * \mathbf{D}_n)(x) \right) = (f * \mathbf{K}_n)(x)$$

où on a posé

$$\mathbf{K}_n = \frac{1}{n+1} \sum_{k=0}^n \mathbf{D}_k,$$

fonction appelée *noyau de Fejér*. Comme l'intégrale de chaque \mathbf{D}_k est égale à 1, il en résulte que

$$\int_0^{2\pi} \mathbf{K}_n(x) \frac{dx}{2\pi} = 1.$$

On va calculer \mathbf{K}_n avec une somme de sinus :

$$\begin{aligned} (n+1)\mathbf{K}_n(x) \sin(x/2) &= \sum_{k=0}^n \sin(kx + x/2) \\ &= \operatorname{Im} \left(\sum_{k=0}^n e^{i(kx + x/2)} \right) = \operatorname{Im} \left(e^{ix/2} \frac{e^{i(n+1)x} - 1}{e^{ix} - 1} \right) = \operatorname{Im} \left(\frac{e^{i(n+1)x} - 1}{2i \sin(x/2)} \right) \\ &= \operatorname{Re} \left(\frac{1 - e^{i(n+1)x}}{2 \sin(x/2)} \right) = \frac{1 - \cos(nx + x)}{2 \sin(x/2)} = \frac{\sin^2(nx/2 + x/2)}{\sin(x/2)}. \end{aligned}$$

On obtient ainsi

$$\mathbf{K}_n(x) = \frac{1}{n+1} \frac{\sin^2 \left[\frac{n+1}{2} x \right]}{\sin^2(x/2)}.$$

On remarque que \mathbf{K}_n est une fonction positive, donc la norme L^1 de \mathbf{K}_n est égale à son intégrale, $\|\mathbf{K}_n\|_1 = 1$. On en déduit une propriété importante

$$(6) \quad \|\sigma_n f\|_\infty \leq \|f\|_\infty.$$

En effet, pour tout x

$$|(\sigma_n f)(x)| \leq \int_{-\pi}^{\pi} |f(x-t)| \mathbf{K}_n(t) \frac{dt}{2\pi} \leq \|f\|_\infty \int_{-\pi}^{\pi} \mathbf{K}_n(t) \frac{dt}{2\pi} = \|f\|_\infty.$$

On peut aussi remarquer que

$$\mathbf{K}_n = \sum_{k=-n}^n \left(1 - \frac{|k|}{n+1} \right) e_k$$

donc

$$\sigma_n f = \sum_{k=-n}^n \left(1 - \frac{|k|}{n+1} \right) c_k(f) e_k.$$

Théorème de Fejér. Si f est continue 2π -périodique, les sommes de Fejér $(\sigma_n f)$ convergent uniformément vers f . Si $f \in L^p([0, 2\pi])$ avec $1 \leq p < \infty$, la suite $(\sigma_n f)$ converge vers f en norme L^p .

Preuve. — Le cas de la convergence uniforme a été traité au chapitre précédent, à la proposition 2.2.4 ; on en a déduit que les polynômes trigonométriques sont denses dans tous les $L^p([0, 2\pi])$, quand $1 \leq p < \infty$. Traitons maintenant le cas de $f \in L^p$; pour tout polynôme trigonométrique

$$g = \sum_{k=-N}^N c_k e_k,$$

il est clair que $S_n g = g$ pour $n \geq N$; il en résulte que la suite des moyennes $\sigma_n g$ tend vers g , puisque

$$\|\sigma_n g - g\|_p = \frac{1}{n+1} \left\| \sum_{k=0}^{N-1} (S_k g - g) \right\|_p = \frac{C(N, g)}{n+1}$$

qui tend vers 0 quand n tend vers l'infini. Si $f \in L^p$, on peut l'approcher par un polynôme trigonométrique g , et d'après (h) la proposition 1.4.1

$$\|\sigma_n f - \sigma_n g\|_p = \|\mathbf{K}_n * (f - g)\|_p \leq \|\mathbf{K}_n\|_1 \|f - g\|_p < \varepsilon/3.$$

Pour n assez grand, on a $\|\sigma_n g - g\|_p < \varepsilon/3$ et le résultat en découle par l'inégalité triangulaire.

Remarque. On peut montrer que si $1 < p < +\infty$ et $f \in L^p$, les sommes de Fourier $(S_n f)$ convergent vers f dans L^p . On l'a vu si $p = 2$, mais les autres cas ne sont pas évidents.

Pour $p = 1$ et $f \in L^1$, les sommes de Fourier ne convergent pas toujours vers f dans L^1 , mais les sommes de Fejér convergent toujours. Pour $p = +\infty$, le résultat est faux : si f est une fonction mesurable bornée dont la classe ne contient (i) aucune fonction continue, les fonctions continues $(\sigma_n f)$ ne peuvent pas converger en norme L^∞ vers f , sinon elle formeraient une suite de Cauchy en norme uniforme (théorème A.5 de l'annexe), donc convergente vers une fonction continue g , qui doit être presque partout égale à f , ce que nous avons exclu.

Fonctions à valeurs réelles : base de sinus et cosinus

On va exprimer la somme de Fourier

$$S_n f = \sum_{k=-n}^n c_k(f) e_k$$

avec les fonctions réelles \cos et \sin . Pour chaque $k > 0$, regroupons les deux termes correspondant à k et $-k$,

$$c_k e^{ikx} + c_{-k} e^{-ikx} = (c_k + c_{-k}) \cos(kx) + i(c_k - c_{-k}) \sin(kx)$$

et écrivons le résultat sous la forme traditionnelle $a_k \cos(kx) + b_k \sin(kx)$, où

$$a_k = c_k + c_{-k} = \int_{-\pi}^{\pi} f(x) (e^{-ikx} + e^{ikx}) \frac{dx}{2\pi} = \int_{-\pi}^{\pi} f(x) (2 \cos(kx)) \frac{dx}{2\pi}$$

et

$$b_k = i(c_k - c_{-k}) = i \int_{-\pi}^{\pi} f(x)(e^{-ikx} - e^{ikx}) \frac{dx}{2\pi} = \int_{-\pi}^{\pi} f(x)(2 \sin(kx)) \frac{dx}{2\pi}.$$

On a donc pour tout $k \geq 1$

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(kx) dx, \quad b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(kx) dx.$$

Il est utile de se rappeler que

$$(7) \quad a_k = c_k(f) + c_{-k}(f), \quad b_k = ic_k(f) - ic_{-k}(f).$$

Par raison de cohérence, on pose aussi

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) dx.$$

De cette façon on obtient pour tout entier $n \geq 0$

$$(S_n f)(x) = \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos(kx) + b_k \sin(kx)).$$

Il est clair d'après les formules qui précèdent que les coefficients a_k, b_k sont réels lorsque f est réelle. De plus, les b_k sont nuls pour une fonction paire, et les a_k sont nuls pour une fonction impaire.

Proposition 3.2.7. *Les fonctions $\mathbf{1}, x \rightarrow \sqrt{2} \cos(kx), x \rightarrow \sqrt{2} \sin(kx)$ pour $k = 1, 2, \dots$ forment une base hilbertienne de l'espace réel $L^2_{\mathbb{R}}([0, 2\pi])$.*

Preuve. — Les formules d'addition de trigonométrie permettent de montrer que la suite est orthonormée. Si $f \in L^2([0, 2\pi])$ est une fonction réelle, on sait d'après le cas complexe que f est limite de la suite $(S_n(f))$; mais on a vu que $S_n f$ est une combinaison linéaire à coefficients réels des fonctions proposées. On a donc bien une base hilbertienne de l'espace réel $L^2_{\mathbb{R}}([0, 2\pi])$.

Exemple. Considérons la fonction $f = \mathbf{1}_{[-\varepsilon, \varepsilon]}$ où $0 < \varepsilon < \pi$. Cette fonction est paire, donc tous les coefficients b_k sont nuls. Pour tout $k \geq 1$,

$$a_k = \frac{2}{\pi} \int_0^{\pi} f(t) \cos(kt) dt = \frac{2}{\pi} \int_0^{\varepsilon} \cos(kt) dt = \frac{2 \sin(k\varepsilon)}{k\pi}.$$

On voit que $a_0 = \varepsilon/\pi$, et le théorème de Dirichlet au point $x = \varepsilon$ donne

$$\frac{1}{2} = \frac{\varepsilon}{\pi} + \sum_{k=1}^{+\infty} \frac{2 \sin(k\varepsilon) \cos(k\varepsilon)}{k\pi} = \frac{\varepsilon}{\pi} + \sum_{k=1}^{+\infty} \frac{\sin(2k\varepsilon)}{k\pi};$$

si on pose $y = 2\varepsilon$, on voit que si $0 < y < 2\pi$,

$$\sum_{k=1}^{+\infty} \frac{\sin(ky)}{k} = \frac{1}{2}(\pi - y).$$

On a retrouvé la formule de l'exemple 3.2.4.

3.3. Variations, applications des séries de Fourier

Base de sinus

À chaque fonction $f \in L^2([0, \pi])$ on associe la fonction impaire $F \in L^2([-\pi, \pi])$ qui est égale à f sur $]0, \pi]$. La fonction impaire F se représente comme série de sinus, convergente dans $L^2([-\pi, \pi])$

$$\lim_n \int_{-\pi}^{\pi} \left| F(x) - \sum_{k=1}^n b_k \sin(kx) \right|^2 dx = 0.$$

Il en résulte immédiatement, en limitant l'intégrale à $[0, \pi]$, que

$$\lim_n \int_0^{\pi} \left| f(x) - \sum_{k=1}^n b_k \sin(kx) \right|^2 dx = 0.$$

Par ailleurs

$$\int_{-\pi}^{\pi} \sin(kx) \sin(\ell x) dx = 2 \int_0^{\pi} \sin(kx) \sin(\ell x) dx$$

ce qui montre que les fonctions $\sin(kx)$, restreintes à $[0, \pi]$, sont orthogonales. Si on choisit de définir la norme par

$$\|f\|_2^2 = \int_0^{\pi} |f(x)|^2 \frac{dx}{\pi}$$

on voit que les fonctions $x \rightarrow \sqrt{2} \sin(kx)$, $k = 1, 2, \dots$ forment une base hilbertienne de l'espace (réel) $L^2([0, \pi])$.

Si f est continue sur $[0, \pi]$, de classe C^1 par morceaux, avec $f(0) = f(\pi) = 0$, alors la fonction impaire F sur $[-\pi, \pi]$ vérifie les mêmes hypothèses, donc ses coefficients de Fourier complexes ($c_n(F)$) sont absolument sommables (théorème 3.1.2), et les coefficients réels (b_k) sont sommables aussi d'après la relation (7); l'égalité

$$F(x) = \sum_{k=1}^{+\infty} b_k \sin(kx)$$

est vraie pour tout x dans $[-\pi, \pi]$; en particulier pour tout x dans $[0, \pi]$, on a

$$f(x) = \sum_{k=1}^{+\infty} b_k \sin(kx).$$

Cette représentation est bien adaptée à la description du phénomène des cordes vibrantes.

Cordes vibrantes, équation des ondes

On veut étudier le mouvement d'une corde vibrante (guitare, violon par exemple), qu'on suppose de longueur π pour simplifier un peu l'écriture des formules. On repère chaque point P de la corde par un nombre réel x entre 0 et π , la quantité x représentant la distance entre le point P et l'origine de la corde associée à la valeur 0. On désigne par

$f(x, t)$ l'écart du point x de la corde, $0 \leq x \leq \pi$, à l'instant t , par rapport à sa position au repos. La corde étant fixée aux deux bouts, on doit supposer que

$$f(0, t) = f(\pi, t) = 0$$

pour tout temps $t \geq 0$. La théorie physique (essentiellement, l'équation fondamentale de la dynamique) conduit à chercher une fonction $f(x, t)$ qui vérifie l'équation des ondes

$$\frac{\partial^2 f}{\partial t^2} = \kappa \frac{\partial^2 f}{\partial x^2}$$

où κ est un paramètre > 0 dépendant de la tension T de la corde et de sa masse m par unité de longueur,

$$\kappa = \frac{T}{m}.$$

Des solutions particulières sont données par

$$u_n(x, t) = \cos(n\sqrt{\kappa}t) \sin(nx),$$

pour tout $n \geq 1$. Si la position de la corde au temps 0 est donnée par une fonction g , nulle en 0 et en π , de classe C^1 par morceaux, on sait qu'on peut représenter g par

$$\forall x \in [0, \pi], \quad g(x) = \sum_{n=1}^{+\infty} b_n \sin(nx),$$

où $\sum |b_n| < +\infty$; pour arranger nos affaires mathématiques, supposons même que $\sum n^2 |b_n| < +\infty$, et considérons

$$f(x, t) = \sum_{n=1}^{+\infty} b_n \cos(n\sqrt{\kappa}t) \sin(nx).$$

Pour $t = 0$, on a bien $f(x, 0) = g(x)$, la position initiale donnée, et nos hypothèses nous permettent de dériver deux fois terme à terme, pour vérifier que f est bien solution de l'équation des ondes, correspondant à la donnée initiale $f(x, 0) = g(x)$. La question de l'unicité de la solution est une question trop délicate pour ce cours.

On peut traiter de façon analogue un cas particulier de l'équation de la chaleur

$$\frac{\partial f}{\partial t} = \kappa \frac{\partial^2 f}{\partial x^2},$$

à savoir celui d'une barre homogène, toujours de longueur π pour simplifier, et dont les deux extrémités sont maintenues à température 0. Ici $f(x, t)$ représente la température au point x à l'instant t . Si la température à l'instant $t = 0$ est donnée par la même fonction g que ci-dessus, on posera

$$f(x, t) = \sum_{n=1}^{+\infty} b_n e^{-n^2 \kappa t} \sin(nx).$$

Les exponentielles négatives permettent à la série et aux séries dérivées de converger plus facilement que dans le cas des ondes : l'hypothèse $\sum |b_n| < +\infty$ suffit pour donner un sens à nos dérivations de séries de fonctions.

Changement de période

On veut pouvoir travailler aussi avec des fonctions de période $T > 0$ quelconque. Pour le faire, il suffit de ramener le problème à la période 2π par changement de variable, d'appliquer les résultats connus dans ce cas et de revenir. Il est tout de même bon de retenir quelques formules qui s'appliquent au cas de la période T .

Tout d'abord, on travaille avec la mesure dx/T qui donne la masse 1 à tous les intervalles-périodes tels que $[0, T]$ ou $[-T/2, T/2]$. Ensuite, on introduit pour tout $n \in \mathbb{Z}$ l'exponentielle

$$e_{n,T} : x \rightarrow e^{i2\pi nx/T}$$

qui est de période T . Les fonctions $(e_{n,T})_{n \in \mathbb{Z}}$ forment une base hilbertienne de $L^2([0, T])$. Si f est continue, T -périodique et de classe C^1 par morceaux, on sait d'après le cas 2π -périodique que les coefficients dans la base sont absolument sommables, et on a pour tout x

$$f(x) = \sum_{n \in \mathbb{Z}} \left(\frac{1}{T} \int_{-T/2}^{T/2} f(t) e^{-i2\pi nt/T} dt \right) e^{i2\pi nx/T}.$$

3.4. Transformation de Fourier et séries de Fourier

Des séries de Fourier à la transformation de Fourier

Supposons que g soit de classe C^2 à support compact sur \mathbb{R} . Pour T assez grand, la fonction g sera nulle en dehors de l'intervalle $[-T/2, T/2]$; si on regarde g comme la restriction à $[-T/2, T/2]$ d'une fonction T -périodique f sur \mathbb{R} , on aura pour tout x tel que $|x| \leq T/2$

$$g(x) = f(x) = \sum_{n \in \mathbb{Z}} \left(\frac{1}{T} \int_{-T/2}^{T/2} f(t) e^{-i2\pi nt/T} dt \right) e^{i2\pi nx/T}.$$

Comme g est nulle hors de $[-T/2, T/2]$, l'intégrale de f sur cet intervalle est aussi l'intégrale de g sur \mathbb{R} et

$$g(x) = \sum_{n \in \mathbb{Z}} \frac{1}{T} \widehat{g}(2\pi n/T) e^{i2\pi nx/T}.$$

Posant $h = 2\pi/T$, il vient

$$g(x) = \frac{1}{2\pi} \sum_{n \in \mathbb{Z}} h \widehat{g}(nh) e^{ixnh}.$$

Il s'agit d'une somme de Riemann généralisée pour la fonction

$$\varphi : y \rightarrow \frac{1}{2\pi} \widehat{g}(y) e^{ixy},$$

correspondant à un découpage de \mathbb{R} en petits intervalles de longueur h . Grâce à l'hypothèse que g est de classe C^2 sur \mathbb{R} , on peut montrer que $\widehat{g}(y)$ est majoré par $C(1+y^2)^{-1}$, et il en résulte que ces sommes de Riemann tendent vers l'intégrale de $-\infty$ à $+\infty$ de la fonction φ lorsque $h \rightarrow 0$, ce qui donne

$$g(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{g}(y) e^{ixy} dy.$$

C'est la formule d'inversion de Fourier, qui était à la base de la théorie L^2 de l'intégrale de Fourier. Ainsi dans cette approche l'intégrale de Fourier apparaît comme limite des séries de Fourier de période T lorsque $T \rightarrow +\infty$.

Signal périodique tronqué

Si on envisage un signal T -périodique $f(t)$ non nul, on ne peut pas considérer sa transformée de Fourier, car f n'est ni intégrable ni de carré intégrable, les seuls cas que nous savons traiter. Néanmoins, si on considère une fonction $\theta(t)$ qui soit à support compact, mais égale à 1 sur un long intervalle, on pourra appliquer la transformation de Fourier à la fonction $t \rightarrow \theta(t)f(t)$, en nous disant que cela nous donnera peut-être une idée de ce que devrait être la transformée de Fourier de f , si elle existait. Supposons pour simplifier que f soit de classe C^1 , ce qui implique que la série de Fourier est normalement convergente,

$$f(t) = \sum_{n \in \mathbb{Z}} a_n e^{i2\pi nt/T}, \quad \sum_{n \in \mathbb{Z}} |a_n| < +\infty.$$

Comme fonction θ , prenons d'abord une fonction φ continue à support compact, égale à 1 sur un voisinage de 0, puis $\theta(t) = \varphi(ct)$, $c > 0$ petit ; quand c devient très petit, la fonction θ est égale à 1 sur un intervalle de plus en plus grand autour de 0. D'un autre côté,

$$\widehat{\theta}(\xi) = c^{-1} \widehat{\varphi}(c^{-1}\xi)$$

se concentre en 0 : le graphe de $\widehat{\theta}$ se présente essentiellement comme une raie verticale de largeur $\sim c$ et de hauteur $\sim 1/c$, située à l'abscisse $\xi = 0$, et l'intégrale de $\widehat{\theta}$ vaut $2\pi \theta(0) = 2\pi$. On aura

$$\begin{aligned} (\widehat{\theta f})(\xi) &= \int_{\mathbb{R}} \theta(t) \left(\sum_{n \in \mathbb{Z}} a_n e^{i2\pi nt/T} \right) e^{-i\xi t} dt = \sum_{n \in \mathbb{Z}} a_n \int_{\mathbb{R}} \theta(t) e^{-i(\xi - 2\pi n/T)t} dt \\ &= \sum_{n \in \mathbb{Z}} a_n \widehat{\theta}(\xi - 2\pi n/T). \end{aligned}$$

Le spectre de θf est donc formé de raies situées autour des abscisses multiples entiers de $2\pi/T$, et dont l'altitude est proportionnelle au coefficient de Fourier a_n de la fonction f .

Formule de Poisson

On considère une fonction F sur \mathbb{R} , suffisamment régulière et suffisamment décroissante à l'infini, par exemple :

la fonction F est de classe C^2 sur \mathbb{R} , avec F'' intégrable ; de plus, il existe une constante C et $\alpha > 1$ tels que $|F(x)| \leq C(1 + |x|)^{-\alpha}$ pour tout $x \in \mathbb{R}$.

On se donne $T > 0$ et on pose

$$\forall x \in \mathbb{R}, \quad f(x) = \sum_{k \in \mathbb{Z}} F(x + kT).$$

On obtient une fonction continue sur \mathbb{R} , qui est T -périodique. La continuité de f se justifie ainsi : pour $y \in [0, T]$ et $k \in \mathbb{Z}$, posons $u_k(y) = F(y + kT)$; alors

$$|u_k(y)| \leq C(1 + |y + kT|)^{-\alpha} \leq C_1(2T + |y + kT|)^{-\alpha} \leq C_1(|k| + 1)^{-\alpha} T^{-\alpha}$$

ce qui donne une série majorante convergente pour la série de fonctions : il y a convergence normale de la série de fonctions $\sum u_k$, donc continuité de la somme f . Calculons les coefficients de Fourier de la fonction T -périodique f ,

$$c_{n,T}(f) = \int_0^T f(x) e^{-i2\pi nx/T} \frac{dx}{T} = \int_0^T \left(\sum_{k \in \mathbb{Z}} F(x + kT) \right) e^{-i2\pi nx/T} \frac{dx}{T}$$

$$\begin{aligned}
&= \sum_{k \in \mathbb{Z}} \int_0^T F(x + kT) e^{-i2\pi nx/T} \frac{dx}{T} = \sum_{k \in \mathbb{Z}} \int_{kT}^{kT+T} F(s) e^{-i2\pi ns/T} \frac{ds}{T} \\
&= \int_{\mathbb{R}} F(x) e^{-i2\pi nx/T} \frac{dx}{T} = \frac{\widehat{F}(2\pi n/T)}{T}.
\end{aligned}$$

Si $\sum |\widehat{F}(2\pi n/T)| < +\infty$, on sait que $\sum_{n \in \mathbb{Z}} |c_{n,T}(f)| < +\infty$, donc f est égale en tout point à la somme de sa série de Fourier d'après la proposition 3.1.1. On a donc

$$f(x) = \sum_{n \in \mathbb{Z}} c_{n,T}(f) e^{i2\pi nx/T}.$$

On obtient ainsi la formule de Poisson

$$(P) \quad \sum_{k \in \mathbb{Z}} F(x + kT) = \frac{1}{T} \sum_{n \in \mathbb{Z}} \widehat{F}(2\pi n/T) e^{i2\pi nx/T}.$$

Justifions la convergence de la série des coefficients de Fourier : d'après l'hypothèse de majoration de F , on sait que F est intégrable, et F'' est intégrable par hypothèse. Les transformées de Fourier \widehat{F} et $\widehat{F}''(\xi) = -\xi^2 \widehat{F}(\xi)$ sont donc bornées, donc $(1 + \xi^2) \widehat{F}(\xi)$ est borné sur \mathbb{R} . Puisqu'il existe une constante M telle que

$$|\widehat{F}(\xi)| \leq \frac{M}{1 + \xi^2}$$

on déduit que $\sum_{n \in \mathbb{Z}} |\widehat{F}(2\pi n/T)| < +\infty$.

Échantillonnage, formule de Shannon

Étant donné un signal $G(t)$ dépendant du temps, on cherche à pouvoir le coder en remplaçant l'information continue, quand le temps t prend toutes les valeurs réelles, par la connaissance des valeurs de G en une suite discrète de temps, qui sont les multiples kT , k entier, d'une valeur $T > 0$ petite : c'est l'échantillonnage du signal.

Pour pouvoir fonctionner, on supposera que le signal G est à *spectre limité* : on fera l'hypothèse que la transformée de Fourier \widehat{G} est nulle en dehors d'un certain intervalle borné $[-M, M]$. On supposera en plus que G vérifie les hypothèses de la formule de Poisson du paragraphe précédent.

Considérons $\lambda \in [-M, M]$ et posons

$$\forall t \in \mathbb{R}, \quad F(t) = G(t) e^{-i\lambda t};$$

on a $\widehat{F}(\xi) = \widehat{G}(\xi + \lambda)$. La formule de Poisson (P), appliquée à F avec $x = 0$, donne

$$\sum_{k \in \mathbb{Z}} G(kT) e^{-i\lambda kT} = \frac{1}{T} \sum_{n \in \mathbb{Z}} \widehat{G}(2\pi n/T + \lambda).$$

Si on suppose que

$$\frac{2\pi}{T} \geq 2M$$

et puisque $|\lambda| \leq M$, on voit facilement qu'il n'y a qu'un terme non nul dans la somme de droite, celui qui correspond à $n = 0$: en effet si $n \neq 0$, on a

$$|2\pi n/T + \lambda| \geq 2\pi/T - |\lambda| \geq 2M - M = M \quad \text{donc} \quad \widehat{G}(2\pi n/T + \lambda) = 0,$$

et par conséquent

$$\sum_{k \in \mathbb{Z}} G(kT) e^{-i\lambda kT} = \frac{1}{T} \widehat{G}(\lambda);$$

on voit déjà que \widehat{G} est complètement connue à partir des valeurs $G(kT)$, $k \in \mathbb{Z}$, puisque ces valeurs permettent de calculer \widehat{G} sur l'intervalle $[-M, M]$, et que $\widehat{G} = 0$ par hypothèse en dehors de cet intervalle. Si on veut aller plus loin, on écrit pour tout $\lambda \in \mathbb{R}$

$$\widehat{G}(\lambda) = \widehat{G}(\lambda) \mathbf{1}_{[-M, M]}(\lambda) = T \sum_{k \in \mathbb{Z}} G(kT) \mathbf{1}_{[-M, M]}(\lambda) e^{-i\lambda kT},$$

et par Fourier inverse

$$\begin{aligned} G(t) &= \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{G}(\lambda) e^{i\lambda t} d\lambda = \frac{T}{2\pi} \sum_{k \in \mathbb{Z}} G(kT) \int_{-M}^M e^{i\lambda(t-kT)} d\lambda \\ &= \frac{T}{\pi} \sum_{k \in \mathbb{Z}} G(kT) \frac{\sin(M(t-kT))}{t-kT}. \end{aligned}$$

Comme on a choisi $\pi/T \geq M$, la transformée de Fourier \widehat{G} est nulle $[-\pi/T, \pi/T]$, donc on peut aussi bien prendre $M = \pi/T$, ce qui donne finalement la *formule de Shannon*

$$\forall t \in \mathbb{R}, \quad G(t) = \sum_{k \in \mathbb{Z}} G(kT) \frac{\sin(\pi(t-kT)/T)}{\pi(t-kT)/T}.$$

La condition de validité est que la *fréquence d'échantillonnage* $1/T$ soit $\geq 2M/(2\pi)$, donc plus que le *double* de la fréquence maximale $M/(2\pi)$ attendue dans le signal G .

Dans le cas de CD musicaux, on ne cherche pas à transmettre de fréquences supérieures à 20KHz, car ces fréquences ne sont pas audibles pour un humain. Dans la pratique, on rencontre souvent dans ce domaine musical une fréquence d'échantillonnage $1/T = 44\text{KHz}$, qui signifie que le signal est numérisé 44000 fois par seconde. Par ailleurs, pour garantir la validité de la méthode, il faut commencer par assurer au préalable que le signal ne contient pas de fréquence supérieure à 20KHz : dans ce but on utilise un filtre passe-bas (que nous avons évoqué au chapitre 1).

Notes du chapitre 3

(a) En fait souvent ce seront des *classes périodiques* : pour une fonction mesurable f définie sur \mathbb{R} , on montre facilement (parce que la mesure de Lebesgue est invariante par translation) que la classe de la fonction translatée $\tau_{2\pi}f$ définie par

$$\forall x \in \mathbb{R}, \quad (\tau_{2\pi}f)(x) = f(x - 2\pi)$$

ne dépend que de la classe de f ; ainsi, on peut définir la translatée d'une classe \widetilde{f} et on dit que la classe \widetilde{f} est 2π -périodique si elle coïncide, en tant que classe, avec la classe translatée $\tau_{2\pi}\widetilde{f}$. Cela revient à dire la chose suivante : pour tout représentant f de la classe \widetilde{f} , l'ensemble

$$\{x \in \mathbb{R} : f(x + 2\pi) \neq f(x)\}$$

est Lebesgue négligeable ; on peut, si on veut, choisir un représentant mesurable, partout défini et qui soit rigoureusement 2π -périodique (exercice).

(b) Voir la discussion avant le lemme 1.1.2 : la fonction f' n'est pas définie partout, mais elle est définie presque partout, mesurable et bornée.

(c) Si on observe bien les destructions de termes, on voit que

$$\begin{aligned} & (e^{-imt} + e^{-imt+it} + \dots + 1 + \dots + e^{imt})(1 - e^{it}) \\ &= e^{-imt} + e^{-imt+it} + \dots + 1 + \dots + e^{imt} \\ & \quad - e^{-imt+it} - \dots - 1 - \dots - e^{imt} - e^{imt+it} \\ &= e^{-imt} - e^{imt+it}. \end{aligned}$$

(d) La fonction sinus est concave sur l'intervalle $[0, \pi/2]$, puisque sa dérivée seconde $-\sin$ est négative ou nulle sur cet intervalle. Sur l'intervalle $[0, \pi/2]$, la fonction sin est donc au dessus de la fonction affine qui coïncide avec sin en 0 et $\pi/2$, à savoir

$$x \rightarrow \frac{2}{\pi} x;$$

on a donc $\sin(x) \geq 2x/\pi$ pour $0 \leq x \leq \pi/2$.

(e) Pour faire fonctionner la démonstration du corollaire 3.2.5 directement à partir du théorème 3.2.2, il suffirait que la fonction vérifie une *condition de Hölder* d'un ordre $\alpha > 0$, c'est-à-dire qu'il existe une constante C telle que

$$|f(x) - f(y)| \leq C|x - y|^\alpha$$

pour tous x, y . Notons qu'une restriction plus forte que la seule continuité de f est nécessaire pour que le corollaire 3.2.5 soit exact : il existe des fonctions continues telles qu'en certains points, la série de Fourier soit divergente.

(f) En utilisant les séries entières, on voit facilement que

$$\operatorname{ch}(ix) = \sum_{k=0}^{+\infty} \frac{(ix)^{2k}}{(2k)!} = \cos x$$

et de même,

$$\operatorname{sh}(ix) = \sum_{k=0}^{+\infty} \frac{(ix)^{2k+1}}{(2k+1)!} = i \sin x.$$

(g) Étant donnée une suite $u_0, u_1, \dots, u_n, \dots$ de réels > 0 , on dit que le *produit infini* $\prod u_k$ est convergent si

$$\Pi_n = \prod_{k=0}^n u_k$$

tend vers une limite *non nulle* quand $n \rightarrow +\infty$, et on pose

$$\prod_{k=0}^{+\infty} u_k = \lim_n \Pi_n.$$

En prenant le logarithme des Π_n , on se ramène facilement à l'étude d'une série.

(h) À vrai dire, dans cette proposition 1.4.1 il s'agissait de convolution sur \mathbb{R} , et pas de convolution périodique. Le lecteur courageux pourra se convaincre que la preuve donnée pour la proposition 1.4.1 marche aussi ici.

(i) Prenons par exemple f nulle sur $] -\pi, 0[$ et égale à 1 sur $]0, \pi[$; la fonction f est continue sur l'intervalle $I =]0, \pi[$, qui n'est pas réduit à un point. S'il existait une fonction continue f_1 équivalente à f , la différence $f_1 - f$ serait continue sur I et presque partout nulle sur I . D'après le théorème A.5, la différence serait nulle en tout point de I , donc $f_1 = 1$ en tout point de $]0, \pi[$; par le même argument on voit qu'on aurait $f_1 = 0$ en tout point de $] -\pi, 0[$, ce qui est impossible si f_1 est continue en 0.

Annexe : rappels d'intégration

Quand on développe la théorie de l'intégrale de Lebesgue des fonctions positives sur un espace mesuré $(\Omega, \mathcal{A}, \mu)$, on admet les fonctions mesurables à valeurs dans $[0, +\infty]$, c'est-à-dire qui peuvent prendre la valeur $+\infty$; en conséquence, l'intégrale peut aussi être égale à $+\infty$. Si (u_n) est une suite de fonctions mesurables positives sur $(\Omega, \mathcal{A}, \mu)$, on peut toujours considérer la fonction mesurable S définie par

$$S(\omega) = \sum_{n=0}^{+\infty} u_n(\omega)$$

où la somme de la série numérique vaut $+\infty$ lorsqu'elle n'est pas convergente au sens usuel.

Théorème A.1. *Si (u_n) est une suite de fonctions mesurables positives sur un espace mesuré $(\Omega, \mathcal{A}, \mu)$, on a*

$$\int_{\Omega} \left(\sum_{n=0}^{+\infty} u_n(\omega) \right) d\mu(\omega) = \sum_{n=0}^{+\infty} \int_{\Omega} u_n(\omega) d\mu(\omega).$$

Preuve. — Pour chaque entier n , on a par l'additivité de l'intégrale positive

$$\int_{\Omega} \left(\sum_{k=0}^n u_k(\omega) \right) d\mu(\omega) = \sum_{k=0}^n \int_{\Omega} u_k(\omega) d\mu(\omega).$$

La suite $s_n = \sum_{k=0}^n u_k$ tend en croissant vers $\sum_{k=0}^{+\infty} u_k$ (valeur $+\infty$ admise), et on a par le théorème de convergence monotone

$$\begin{aligned} \int_{\Omega} \left(\sum_{k=0}^{+\infty} u_k(\omega) \right) d\mu(\omega) &= \lim_n \int_{\Omega} s_n(\omega) d\mu(\omega) \\ &= \lim_n \sum_{k=0}^n \int_{\Omega} u_k(\omega) d\mu(\omega) = \sum_{k=0}^{+\infty} \int_{\Omega} u_k(\omega) d\mu(\omega). \end{aligned}$$

///

Le théorème suivant donne un critère assez efficace pour pouvoir intervertir une intégrale et une série.

Théorème A.2. *Si (u_n) est une suite de fonctions mesurables sur un espace mesuré $(\Omega, \mathcal{A}, \mu)$, telle que la fonction $\sum |u_n|$ soit μ -intégrable, on a*

$$\int_{\Omega} \left(\sum_{n=0}^{+\infty} u_n(\omega) \right) d\mu(\omega) = \sum_{n=0}^{+\infty} \int_{\Omega} u_n(\omega) d\mu(\omega).$$

Preuve. — Définissons une fonction V mesurable à valeurs dans $[0, +\infty]$ (valeur $+\infty$ admise) en posant

$$\forall \omega \in \Omega, \quad V(\omega) = \sum_{k=0}^{+\infty} |u_k(\omega)|.$$

Par hypothèse, V est intégrable ; il en résulte que $V(\omega) < +\infty$ pour presque tout ω , c'est-à-dire que la série $\sum_{k=0}^{+\infty} |u_k(\omega)|$ est absolument convergente pour presque tout ω . Si on pose $s_n = \sum_{k=0}^n u_k$, on voit que $|s_n| \leq V$ (la fonction V est un majorant intégrable indépendant de n) et (s_n) tend presque partout vers

$$S(\omega) = \sum_{k=0}^{+\infty} u_k(\omega).$$

Par le théorème de convergence dominée, on conclut.

///

Théorème A.3. *Si (f_n) est une suite de fonctions mesurables, si $1 \leq p < +\infty$ et si $\int_{\Omega} |f_n(\omega) - f(\omega)|^p d\mu(\omega)$ tend vers 0, il existe une sous-suite (f_{n_k}) telle que $f_{n_k}(\omega)$ tende vers $f(\omega)$ pour presque tout $\omega \in \Omega$.*

Preuve. — En effet, puisque $\int_{\Omega} |f_n(\omega) - f(\omega)|^p d\mu(\omega)$ tend vers 0, on peut, par récurrence, trouver pour tout entier $k \geq 0$ un entier $n_k > n_{k-1}$ tel que

$$\int_{\Omega} |f_{n_k}(\omega) - f(\omega)|^p d\mu(\omega) < 2^{-k};$$

on en déduit, en appliquant le théorème A.1, que

$$\int_{\Omega} \left(\sum_{k \geq 0} |f_{n_k}(\omega) - f(\omega)|^p \right) d\mu(\omega) = \sum_{k \geq 0} \int_{\Omega} |f_{n_k}(\omega) - f(\omega)|^p d\mu(\omega) < \sum_{k \geq 0} 2^{-k} = 2 < \infty.$$

Il en résulte que $\sum_{k \geq 0} |f_{n_k}(\omega) - f(\omega)|^p$ est fini μ -presque partout, et en particulier, le terme général de cette série tend vers 0 pour presque tout ω , donc

$$|f_{n_k}(\omega) - f(\omega)|^p \rightarrow 0, \quad \text{c'est-à-dire} \quad f_{n_k}(\omega) \rightarrow f(\omega)$$

pour presque tout ω , comme annoncé.

///

Corollaire A.4. *Si une suite de fonctions (u_n) tend vers u dans l'espace $L^p(\Omega, \mu)$, avec $1 \leq p < \infty$ et si (u_n) tend simplement μ -presque partout vers une autre fonction v , alors $u = v$ μ -presque partout.*

Preuve. — Il existe d'après le théorème précédent une sous-suite (n_k) et un ensemble μ -négligeable N tels que $u_{n_k}(\omega)$ tende vers $u(\omega)$ pour tout $\omega \notin N$; par hypothèse il existe un ensemble négligeable M tel que $u_n(\omega)$ tende vers $v(\omega)$ pour tout $\omega \notin M$; il en résulte que $u(\omega) = v(\omega)$ pour tout $\omega \notin N \cup M$, c'est-à-dire que $u = v$ μ -presque partout.

///

Théorème A.5. *Si $a < b$, si f est une fonction continue sur $[a, b]$, sa norme uniforme $\|f\|_u$ est égale à sa norme $\|f\|_{\infty}$ dans l'espace $L^{\infty}([a, b], dx)$. Si I est un intervalle de \mathbb{R} non réduit à un point, et si f est une fonction continue sur I et nulle en presque tout point de I (pour la mesure de Lebesgue), alors f est nulle en tout point de I ; si deux fonctions continues f, g sont égales en presque tout point de I , elles sont égales partout.*

Preuve. — Soit f continue sur $[a, b]$, et posons $M = \|f\|_u$; alors $|f(x)| \leq M$ pour tout x , donc l'ensemble $\{x : |f(x)| > M\}$ est de mesure nulle, puisqu'il est vide! Il en résulte que $\|f\|_\infty \leq \|f\|_u$. Inversement, si on a $m < M$, l'ensemble $V = \{x \in [a, b] : |f(x)| > m\}$ est un ouvert non vide de $[a, b]$; il existe donc un intervalle $]c, d[$ avec $a \leq c < d \leq b$ contenu dans V . Mais cet intervalle est de mesure de Lebesgue $d - c > 0$. La mesure de V est donc non nulle, et on en déduit que $\|f\|_\infty \geq m$, ce qui permet de terminer la preuve de ce premier point.

Si f continue sur I n'est pas nulle en tout point de I , on peut trouver par continuité un intervalle $[a, b]$ avec $a < b$, contenu dans I et sur lequel f est non nulle; puisque $[a, b]$ est de mesure > 0 , la fonction f n'est pas presque partout nulle. L'affirmation sur l'égalité de f et g en résulte immédiatement, en considérant la fonction $f - g$.

///

Théorème A.6. *Le produit de deux fonctions f, g de $L^2(\Omega, \mu)$ est μ -intégrable, et*

$$\|fg\|_1 \leq \|f\|_2 \|g\|_2.$$

Preuve. — Il suffit d'appliquer l'inégalité de Hölder du cours d'Intégration : lorsque $p, q \geq 1$ vérifient $1/p + 1/q = 1$, on a

$$\int_{\Omega} |f(\omega)g(\omega)| \, d\mu(\omega) \leq \left(\int_{\Omega} |f(\omega)|^p \, d\mu(\omega) \right)^{1/p} \left(\int_{\Omega} |g(\omega)|^q \, d\mu(\omega) \right)^{1/q}.$$

En prenant $p = q = 2$, on obtient le résultat voulu.

///

Lemme A.7. *Si $\varphi \in L^2(\Omega, \mu)$ est nulle hors d'un ensemble mesurable A de mesure finie, alors $\varphi \in L^1(\Omega, \mu)$ et on a*

$$\|\varphi\|_1 \leq \sqrt{\mu(A)} \|\varphi\|_2.$$

Preuve. — Puisque φ est nulle hors de A , on a $\varphi = \mathbf{1}_A \varphi$, donc par Cauchy-Schwarz (c'est-à-dire Hölder pour $p = 2$)

$$\begin{aligned} \int_{\Omega} |\varphi(\omega)| \, d\mu(\omega) &= \int_{\Omega} \mathbf{1}_A(\omega) |\varphi(\omega)| \, d\mu(\omega) \leq \\ &\leq \left(\int_{\Omega} \mathbf{1}_A(\omega)^2 \, d\mu(\omega) \right)^{1/2} \left(\int_{\Omega} |\varphi(\omega)|^2 \, d\mu(\omega) \right)^{1/2} = \sqrt{\mu(A)} \|\varphi\|_2. \end{aligned}$$

///

Lemme A.8. *Soient $(\Omega, \mathcal{A}, \mu)$ un espace mesuré et p un réel tel que $1 \leq p < +\infty$; si une suite (f_n) de fonctions mesurables sur Ω tend μ -presque partout vers une fonction f en étant dominée par $g \in L^p(\Omega, \mu)$, c'est-à-dire qu'on a pour μ -presque tout $\omega \in \Omega$*

$$|f_n(\omega)| \leq g(\omega)$$

pour tout entier n , alors $f \in L^p(\Omega, \mu)$ et la suite (f_n) tend vers f dans $L^p(\Omega, \mu)$.

Preuve. — Puisque $(f_n(\omega))$ tend vers $f(\omega)$ presque partout, la fonction f est mesurable et l'hypothèse de domination entraîne que $|f(\omega)| \leq g(\omega)$ pour presque tout ω ; il en résulte déjà que f appartient à $L^p(\Omega, \mu)$; posons

$$g_n(\omega) = |f_n(\omega) - f(\omega)|^p;$$

cette suite de fonctions (g_n) tend vers 0 presque partout en étant dominée par la fonction intégrable $(2g)^p$; d'après le théorème de convergence dominée,

$$\|f_n - f\|_p^p = \int_{\Omega} |f_n - f|^p \, d\mu = \int_{\Omega} g_n \, d\mu \rightarrow 0.$$

///

Lemme A.9. Soit (a_n) une suite de réels qui tend vers $-\infty$ et soit (b_n) une suite de réels qui tend vers $+\infty$; pour tout p tel que $1 \leq p < +\infty$ et toute fonction $f \in L^p(\mathbb{R})$, la suite $f_n = \mathbf{1}_{[a_n, b_n]} f$ tend vers f dans $L^p(\mathbb{R})$.

Preuve. — La suite (f_n) tend simplement vers la fonction f en étant dominée par la fonction $g = |f|$ qui est dans $L^p(\mathbb{R})$: le résultat découle du lemme précédent. ///

Rappelons qu'une tribu de parties d'un ensemble X est une famille \mathcal{A} de parties de X qui contient l'ensemble vide, qui est stable par passage au complémentaire et par réunion dénombrable. La tribu borélienne de \mathbb{R} est la plus petite tribu de parties de \mathbb{R} qui contienne tous les intervalles ouverts. Il en résulte qu'elle contient tous les ouverts, donc aussi tous les fermés ; par combinaison de l'opération de passage au complémentaire et des opérations dénombrables de réunion, on obtient la plupart des sous-ensembles de la droite réelle \mathbb{R} qui apparaissent dans un cours d'Analyse.

Un sous-ensemble N de \mathbb{R} , même non borélien, est dit *Lebesgue-négligeable* s'il existe un borélien B , de mesure de Lebesgue nulle, tel que $N \subset B$; un sous-ensemble X de \mathbb{R} est *Lebesgue-mesurable* si on peut l'écrire sous la forme $X = A \cup N$, où A est borélien et N Lebesgue-négligeable. On peut construire une théorie mathématique aussi cohérente que celle qu'on utilise ordinairement, dans laquelle on pose pour axiome que *tout sous-ensemble de \mathbb{R} est Lebesgue-mesurable* (Robert Solovay, 1970 ; il faut retoucher l'axiome du choix, qui justement, permet de trouver des ensembles non mesurables).

Théorème A.10. On considère un sous-espace vectoriel F de l'espace réel $L^p(\mathbb{R})$, pour un p tel que $1 \leq p < +\infty$. On suppose que F vérifie les deux propriétés suivantes :

- si $f, g \in F$, la fonction $\sup(f, g)$ est dans F ;
- pour tout intervalle ouvert borné $I \subset \mathbb{R}$ et pour tout $\varepsilon > 0$, il existe $f \in F$ telle que $\|f - \mathbf{1}_I\|_p < \varepsilon$.

Il en résulte que F est dense dans $L^p(\mathbb{R})$.

Preuve. — Désignons par F_p l'adhérence de F dans $L^p(\mathbb{R})$; il est assez facile de montrer que F_p est un sous-espace vectoriel. De plus, si on a $f_1, g_1 \in F_p$ et $\varepsilon > 0$, il existe $f, g \in F$ tels que $\|f_1 - f\|_p < \varepsilon$, $\|g_1 - g\|_p < \varepsilon$; on note que

$$|\sup(f_1, g_1) - \sup(f, g)| \leq |f_1 - f| + |g_1 - g|$$

et il en résulte que

$$\|\sup(f_1, g_1) - \sup(f, g)\|_p \leq 2\varepsilon,$$

ce qui montre que $\sup(f_1, g_1) \in F_p$. La deuxième hypothèse sur F signifie que F_p contient l'indicatrice $\mathbf{1}_I$, pour tout intervalle ouvert borné I .

Fixons un intervalle ouvert borné I et considérons la famille \mathcal{A}_I de tous les ensembles boréliens $A \in \mathcal{B}_{\mathbb{R}}$ tels que $\mathbf{1}_{A \cap I}$ appartienne à F_p . Pour tout intervalle ouvert J , borné ou non, on a $J \in \mathcal{A}_I$ puisque $J \cap I$ est un intervalle ouvert borné. Montrons que \mathcal{A}_I est une tribu : puisque \mathbb{R} est un intervalle ouvert, on a $\mathbb{R} \in \mathcal{A}_I$; si $A, B \in \mathcal{A}_I$, on voit que

$$\mathbf{1}_{(A \cup B) \cap I} = \sup(\mathbf{1}_{A \cap I}, \mathbf{1}_{B \cap I}) \in F_p,$$

ce qui montre que $A \cup B$ appartient à \mathcal{A}_I . Pour le complémentaire A^c de $A \in \mathcal{A}_I$, on a

$$\mathbf{1}_{A^c \cap I} = \mathbf{1}_I - \mathbf{1}_{A \cap I} \in F_p$$

puisque F_p est un sous-espace vectoriel. Enfin, si (A_n) est une suite croissante d'éléments de \mathcal{A}_I et si on pose $A = \bigcup_n A_n$, la suite $\mathbf{1}_{A_n \cap I}$ tend simplement vers $\mathbf{1}_{A \cap I}$ en étant dominée par $\mathbf{1}_I \in L^p(\mathbb{R})$; on en déduit que $\|\mathbf{1}_{A_n \cap I} - \mathbf{1}_{A \cap I}\|_p$ tend vers 0, et on voit ainsi que $\mathbf{1}_{A \cap I}$ est limite d'une suite dans F_p , donc $\mathbf{1}_{A \cap I} \in F_p$ et $A \in \mathcal{A}_I$. Il en résulte que \mathcal{A}_I est une tribu, contenue dans la tribu borélienne et contenant tous les intervalles ouverts : on a donc $\mathcal{A}_I = \mathcal{B}_{\mathbb{R}}$, et ceci pour tout I . Ceci signifie que pour tout borélien borné A , on a $\mathbf{1}_A \in F_p$: on peut en effet trouver un intervalle ouvert borné I tel que $A \subset I$, donc $\mathbf{1}_A = \mathbf{1}_{A \cap I} \in F_p$. Si $f \in L^p(\mathbb{R})$ est étagée et nulle en dehors d'un borné, on peut écrire

$$f = \sum_{j=1}^N c_j \mathbf{1}_{A_j}$$

où les A_j sont des ensembles boréliens bornés ; on a donc $\mathbf{1}_{A_j} \in F_p$ pour chaque indice $j = 1, \dots, N$ et il en résulte que $f \in F_p$.

Si $f \in L^p(\mathbb{R})$ est une fonction réelle positive, il existe d'après le lemme A.9 une fonction $g \in L^p(\mathbb{R})$, positive et nulle hors d'un borné, telle que $\|f - g\|_p < \varepsilon$; la fonction g est limite simple d'une suite croissante (g_n) de fonctions étagées qui converge simplement vers g , telles que $0 \leq g_n \leq g$; la suite (g_n) tend simplement vers g en étant dominée par la fonction $g \in L^p(\mathbb{R})$; d'après le lemme A.8, la suite (g_n) tend vers g en norme L^p ; puisque g_n est étagée nulle en dehors d'un borné, on sait que $g_n \in F_p$. Il en résulte que $g \in F_p$; puisque F_p est fermé et que f est aussi proche qu'on veut d'un élément $g \in F_p$, on conclut que $f \in F_p$.

Si $f \in L^p(\mathbb{R})$ est une fonction réelle, on peut écrire $f = f_+ - f_-$, où f_+ et f_- sont positives ; on conclut que $f \in F_p$, ce qui prouve que $F_p = L^p(\mathbb{R})$ et termine la preuve.

///

Théorème A.11. *Les trois sous-espaces suivants sont denses dans $L^p(\mathbb{R})$, pour tout p tel que $1 \leq p < +\infty$:*

- l'espace des fonctions continues à support compact ;
- l'espace des fonctions continues affines par morceaux, à support compact ;
- l'espace des fonctions en escalier.

Preuve. — Bien entendu, le deuxième exemple implique le premier. Dans les trois cas, on voit que le sup de deux fonctions de l'espace F considéré reste dans l'espace (le sup de deux fonctions continues est continu ; le sup de deux fonctions affines est affine par morceaux) ; si $I =]a, b[$, la suite (f_n) des fonctions affines par morceaux définies par

$$f_n(x) = n \int_{x-1/n}^x \mathbf{1}_{[a, b-1/n]}(t) dt$$

tend simplement vers $\mathbf{1}_I$ en étant dominée par $\mathbf{1}_I \in L^p(\mathbb{R})$, donc $\mathbf{1}_I$ est limite en norme L^p d'une suite $(f_n) \subset F$ (voir aussi le lemme 1.3.1). Dans le cas de l'espace des fonctions en escalier, il suffit de prendre $f_n = \mathbf{1}_I$ pour tout n . D'après le théorème précédent, toute fonction réelle $f \in L^p(\mathbb{R})$ est adhérente à F .

Si $f \in L^p(\mathbb{R})$ est une fonction complexe, on peut écrire $f = \operatorname{Re} f + i \operatorname{Im} f$; pour approcher f , il suffit d'approcher séparément les deux fonctions réelles $\operatorname{Re} f$ et $\operatorname{Im} f$.

///

Index alphabétique

adjoint hilbertien (opérateur)	55
affine par morceaux (fonction continue et)	4, 7, 13, 87
antilinéaire (application)	29, 53
application antilinéaire	29, 53
application hermitienne	56
application isométrique	55
application unitaire	56
approximation par convolution	19
base de sinus ou de cosinus	73, 75
base hilbertienne	37
Bessel (fonction de)	61, 62, 63
C^1 par morceaux (fonction)	3, 62
Carleson (théorème de)	46
Cauchy-Schwarz (inégalité de)	30
Cesàro (sommés de)	71
chaleur (équation de la)	23, 76
classe de fonctions	32
coefficient de Fourier	44, 59
condition de Hölder	81
convergence au sens de Cesàro	71
convergence ponctuelle des séries de Fourier	65
convexe (ensemble)	48
convolution de fonctions intégrables	9
convolution périodique	42
cordes vibrantes	75
densité des fonctions en escalier	87
densité gaussienne	21
dérivée généralisée	21
directe (somme)	50, 51
Dirichlet (noyau de)	71
Dirichlet-Dini (théorème de)	69
dual topologique	33
dual topologique d'un espace de Hilbert	53
échantillonnage	79
engendré (sous-espace vectoriel)	36
ensemble convexe	48
équation aux dérivées partielles	23
équation de la chaleur	23, 76
équation des ondes	75
escalier (fonction en)	2
espace séparable	40
espace vectoriel engendré	36
espérance conditionnelle	49
Fejér (noyau de)	43, 71
Fejér (sommés de)	71
filtre	11, 80
fonction continue à support compact	4
fonction continue et affine par morceaux	4, 7, 13, 87
fonction de Bessel	61, 62, 63
fonction de classe C^1 par morceaux	3, 62

fonction en escalier	2
fonctions en escalier (densité des)	87
fonction indicatrice	2
fonction lipschitzienne	54, 68
fonction zêta	68
forme hermitienne	29
formule de Poisson	78
formule de polarisation	30
formule de Shannon	79
Fourier dans L^2	12
Fourier (transformée d'une fonction intégrable)	1
fréquence d'échantillonnage	80
fréquence d'un signal périodique	5
fréquence (modulation de)	64
gaussienne (densité)	21
généralisée (dérivée)	21
Gram-Schmidt (procédé de)	39
Haar (système de)	38
hermitienne (application linéaire)	56
hermitienne (forme)	29
Hölder (condition de)	81
identité de Parseval	8, 14, 15
identité du parallélogramme	30
indicatrice (fonction)	2
inégalité de Cauchy-Schwarz	30
inégalité triangulaire	30
injectivité de la transformation de Fourier	20
inversion de la transformation de Fourier	7, 16
isométrique (application)	55
Kronecker (symbole de)	33, 57, 60
lemme de Riemann-Lebesgue	1, 2, 65
lipschitzienne (fonction)	54, 68
médiane (relation de la)	30
modulation de fréquence	64
norme uniforme	41
noyau de Dirichlet	71
noyau de Fejér	43, 71
noyau de Poisson	43
ondes (équation des)	75
opérateur adjoint hilbertien	55
orthogonal d'une partie	50
orthogonale (projection)	34, 47, 48
orthogonaux (vecteurs)	31
orthonormée (suite de vecteurs)	33
parallélogramme (identité du)	30
Parseval (identité de)	8, 14, 15
Poisson (formule de)	78
Poisson (noyau de)	43
polarisation (formule de)	30
polynôme trigonométrique	42
procédé de Gram-Schmidt	39
produit infini	70
produit scalaire	29

projection orthogonale	34, 47, 48
pulsation d'un signal périodique	5
Pythagore (théorème de)	32
relation de la médiane	30
retardé (signal)	9
Riemann-Lebesgue (lemme de)	1, 65
semi-norme	30
séparable (espace)	40
série de vecteurs	35
Shannon (théorème de)	79
signal périodique	5
signal retardé	9
somme d'une série de vecteurs	35
somme directe	50, 51
somme directe orthogonale	51
sommes de Cesàro	71
sommes de Fejér	71
sommes de Fourier	59, 69
sous-espace vectoriel engendré	36
sous-espace vectoriel fermé engendré	36
suite orthonormée	33
support compact (fonction continue à)	4
symbole de Kronecker	33, 57, 60
système de Haar	38
théorème de Carleson	46
théorème de Dirichlet-Dini	69
théorème de Pythagore	32
théorème de Shannon	79
théorème de Weierstrass	51, 52
théorème de Weierstrass périodique	45
transformée de Fourier d'une fonction de L^2	12
transformée de Fourier d'une fonction intégrable	1
triangulaire (inégalité)	30
trigonométrique (polynôme)	42
uniforme (norme)	41
unitaire (application)	56
vecteurs orthogonaux	31
Weierstrass (théorème de)	51, 52
Weierstrass périodique (théorème de)	45
zêta (fonction)	68

Index des notations

$\mathbf{1}_A$	fonction indicatrice du sous-ensemble A	2
A^\perp	orthogonal de la partie A d'un espace de Hilbert	50
$\text{Aff}(a, b)$	fonctions continues et affines par morceaux sur $[a, b]$	4
$\text{Aff}_c(\mathbb{R})$	fonctions continues à support compact et affines par morceaux sur \mathbb{R}	4
a_n, b_n	coefficients de Fourier réels de la fonction f , pour $n \geq 0$ et $n \geq 1$	73
$c_n(f)$	coefficients de Fourier complexes de la fonction f , $n \in \mathbb{Z}$	59
$C([a, b])$	espace des fonctions continues sur $[a, b]$	41
$\delta_{i,j}$	symbole de Kronecker, égal à 0 ou 1 selon que $i \neq j$ ou non	33
$D_n(x)$	noyau de Dirichlet	71
e_n	fonction $t \rightarrow e^{int}$ de la base exponentielle, pour $n \in \mathbb{Z}$	41
E'	dual topologique de l'espace normé E	33
\mathcal{F}	transformation de Fourier sur $L^2(\mathbb{R})$	15
$F \oplus G$	somme directe orthogonale	50
\widehat{f}	transformée de Fourier de la fonction f	1
$f * g$	convolution des fonctions f et g	9
$f *_{\text{per}} g$	convolution périodique des fonctions f et g	42
J_0	fonction de Bessel d'indice 0	61
K_τ	une fonction sur \mathbb{R} utilisée dans le cours	10
$L^1(\mathbb{R})$	espace des (classes de) fonctions intégrables sur \mathbb{R}	1
$L^p(\mathbb{R})$	espace des (classes de) fonctions de puissance p ième intégrable	1
$L^\infty(\mathbb{R})$	espace des (classes de) fonctions mesurables bornées	1
$\ell_2(\mathbb{N})$	espace des suites scalaires de carré sommable	31
P_F	projection orthogonale sur le sous-espace vectoriel F	34
$\sum_{k=0}^{+\infty} u_k$	somme d'une série de vecteurs dans un espace vectoriel normé	35
$S_n(f)$	sommes de Fourier de la fonction f , $n \geq 0$	59
$\sigma_n(f)$	sommes de Fejér de la fonction f , $n \geq 0$	71
T^*	adjoint hilbertien de l'application $T \in \mathcal{L}(H)$	55
$\text{Vect}(A)$	sous-espace vectoriel engendré par la partie A	31
X	un espace de fonctions sur \mathbb{R} utilisé dans le cours	7
x^\perp	sous-espace orthogonal au vecteur x d'un espace de Hilbert	31
$x \perp y$	orthogonalité des vecteurs x et y d'un espace de Hilbert	31
$\langle x, y \rangle$	produit scalaire des vecteurs x et y	29
$\zeta(s)$	fonction zêta de Riemann	68
$\ f\ _p$	norme des fonctions dans les espaces L^p , pour $1 \leq p < +\infty$	1
$\ f\ _\infty$	norme de la fonction f dans l'espace L^∞	25
$\ f\ _u$	norme uniforme d'une fonction continue f	41
$\ x\ $	norme (ou semi-norme) déduite d'un produit scalaire	30