

(B01-public)

Résumé : Nous allons donner un bref aperçu de la théorie mathématique des ondelettes qui décompose des fonctions dans des bases hilbertiennes bien choisies. On applique cette théorie au traitement du signal.

Mots clefs : Analyse hilbertienne. Approximation.

- *Il est rappelé que le jury n'exige pas une compréhension exhaustive du texte. La présentation, bien que totalement libre, doit être organisée et le jury apprécie qu'un plan soit annoncé en préliminaire. L'exposé doit être construit en évitant la paraphrase et mettant en lumière les connaissances, à partir des éléments du texte. Il doit contenir des illustrations informatiques réalisées sur ordinateur, ou, à défaut, des propositions de telles illustrations. Des pistes de réflexion, indicatives et largement indépendantes les unes des autres, vous sont proposées en fin de texte.*

1. Introduction

Les signaux numériques traités par les ordinateurs sont des vecteurs $u \in \mathbb{R}^n$ avec n très grand. Par exemple un signal musical d'une durée d'une seconde est numérisé dans un vecteur à $n = 44100$ coordonnées. On peut choisir de stocker u ou de le transmettre à l'aide de ses coordonnées $u[k]$, $1 \leq k \leq n$ de décomposition dans la base canonique de \mathbb{R}^n , ou en utilisant une autre base \mathcal{B} de \mathbb{R}^n . En effet en choisissant bien \mathcal{B} , on peut espérer que beaucoup des coefficients de décomposition de u dans \mathcal{B} seront négligeables, et on les mettra tout simplement à zéro. C'est une manière d'effectuer une opération de « compression » du signal. Toute information sur u qui nous permet de comprendre quelle base \mathcal{B} serait adéquate pour ces opérations est importante à exploiter.

Or dans un certain nombre de cas, les signaux numériques proviennent de la discrétisation sur un intervalle $[0, T]$ d'une fonction $f : [0, T] \rightarrow \mathbb{R}$. On obtient ainsi avec $f : [0, T] \rightarrow \mathbb{R}$ le vecteur u tel que $u[k] = f(kT/n)$ pour $0 \leq k \leq n - 1$. Dans toute la suite on prendra $T = 1$ pour ne pas alourdir les notations.

Utiliser le fait que u provienne d'une fonction de la variable continue $f : t \mapsto f(t)$ dans $L^2([0, 1])$ permet de faciliter l'analyse mathématique en utilisant une base $\tilde{\mathcal{B}}$ de $L^2([0, 1])$, si elle existe, dont la base \mathcal{B} est une version discrète. En particulier si on sait que f a peu de coefficients non nuls de décomposition dans la base $\tilde{\mathcal{B}}$, on pourra espérer en déduire que u aura beaucoup de coefficients de décomposition dans la base \mathcal{B} qui seront proches de 0.

Un premier choix possible pour $\tilde{\mathcal{B}}$ est celui de la base hilbertienne des exponentielles complexes de période 1 $\{e_k : t \mapsto e^{2i\pi kt}, k \in \mathbb{Z}\}$, permettant d'effectuer la décomposition d'une fonction $f \in L^2([0, 1])$ à l'aide de ses coefficients de Fourier $c_k(f) = \int_0^1 f(t)e_{-k}(t)dt$,

$k \in \mathbb{Z}$. Ce choix, si il a des avantages, a aussi des inconvénients. D'abord la représentation de f à l'aide des coefficients de Fourier ne reflète pas l'information locale présente dans la fonction : si celle-ci est nulle sur un long intervalle, et non nulle ensuite, les coefficients de Fourier prendront en compte la partie non nulle de la fonction, et seront en général non nuls. De plus, dans le cas d'une fonction f discontinue (par exemple $t \mapsto H(t) = \mathbf{1}_{[1/3,1]}(t)$) on met en évidence que le fait de remplacer la fonction par sa série de Fourier d'ordre N $t \mapsto S_N(f)(t) = \sum_{k=-N}^N c_k(f) e^{2i\pi kt}$, introduit des oscillations hautes fréquences parasites, que l'on appelle « phénomène de Gibbs » en mathématiques. Celles-ci conduisent à des distortions gênantes dans la pratique.

Pour remédier à ces problèmes on construit une base \mathcal{B} et sa version continue $\tilde{\mathcal{B}}$ qui est composée de fonctions bien localisées (qui donneront donc des coefficients nuls dans tout intervalle où la fonction est nulle), et permettent de contourner le phénomène de Gibbs. Dans ce texte, notre but est d'étudier la construction de la base d'ondelettes de Haar, qui formera une base hilbertienne de $L^2([0,1])$.

On notera $\langle f_1, f_2 \rangle = \int_{[0,1]} f_1(t) \overline{f_2(t)} dt$ le produit scalaire sur l'espace hilbertien $L^2([0,1])$.

2. Une suite de fonctions approximantes de $f \in L^2([0,1])$

Soit $f \in L^2([0,1])$. L'idée naturelle pour approcher une fonction f est de considérer des fonctions constantes par morceaux. On utilise que f est une fonction $L^2([0,1]) \subset L^1([0,1])$.

Définition 1. Dans toute la suite on posera, pour tout $(j, k) \in \mathbb{Z}^2$, $I_{j,k} = [k2^{-j}, (k+1)2^{-j}]$. Pour tout $j \geq 0$, pour tout $t \in [0,1]$, posons

$$(1) \quad P_j(f)(t) = \sum_{k=0}^{2^j-1} c_{j,k}(f) 2^{j/2} \mathbf{1}_{I_{j,k}}(t), \quad \text{où } c_{j,k}(f) = 2^{j/2} \int_{I_{j,k}} f(t) dt.$$

Pour tout couple d'entiers (j, k) , $\mathbf{1}_{I_{j,k}}$ est la fonction indicatrice de l'intervalle dyadique $I_{j,k}$. La somme définissant $P_j(f)$ dans (1) est prise sur les k tels que $I_{j,k} \subset [0,1]$.

Par construction, pour tout $j \geq 0$, $P_j(f)$ est constante sur chaque intervalle $I_{j,k}$. De plus,

$$(2) \quad \text{pour tout } (j, k), \quad \int_{I_{j,k}} f(t) dt = \int_{I_{j,k}} P_j(f)(t) dt \quad \text{et donc} \quad \int_{[0,1]} f(t) dt = \int_{[0,1]} P_j(f)(t) dt.$$

Posons $\varphi(t) = \mathbf{1}_{[0,1]}(t)$, et pour tout couple d'entiers (j, k) , $\varphi_{j,k} : t \mapsto 2^{j/2} \varphi(2^j t - k)$.

Proposition 2. Pour tout $j \geq 0$, la famille de fonctions $\{\varphi_{j,k}\}_{k \in \{0,1,\dots,2^j-1\}}$ forme une base orthonormale de l'espace fonctionnel V_j défini par

$$(3) \quad V_j = \{f \in L^2([0,1]) : \text{pour tout } k \in \{0,1,\dots,2^j-1\}, f|_{I_{j,k}} \text{ est constante}\}.$$

Pour toute $f \in L^2([0,1])$, $P_j(f)$ est la projection orthogonale de f sur V_j et s'écrit

$$(4) \quad \text{pour tout } t \in [0,1], \quad P_j(f)(t) = \sum_{k=0}^{2^j-1} \langle f, \varphi_{j,k} \rangle \varphi_{j,k}(t) = \sum_{k=0}^{2^j-1} c_{j,k}(f) \varphi_{j,k}(t).$$

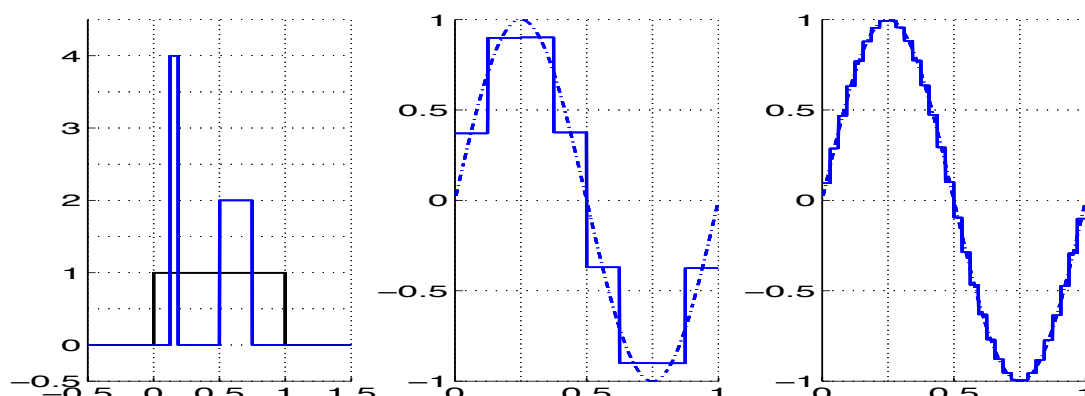


FIGURE 1. **Gauche** : La fonction φ et les fonctions $\varphi_{2,2}$ et $\varphi_{4,2}$. **Centre** : La fonction $f(t) = \sin(2\pi t)$ et son approximation $P_3(f)$. **Droite** : f et $P_5(f)$.

La suite d'espace de fonctions $(V_j)_{j \geq 0}$ est une suite croissante : pour tout $j \geq 0$, $V_j \subset V_{j+1}$. Pour le voir, on peut constater par exemple que

$$(5) \quad \text{pour tout couple } (j, k), \quad \varphi_{j,k}(t) = \frac{1}{\sqrt{2}} (\varphi_{j+1,2k}(t) + \varphi_{j+1,2k+1}(t)).$$

Enfin, les fonctions constantes par morceaux sont denses dans $L^2([0, 1])$ pour la norme naturelle dans cet espace. Cela entraîne $\lim_{j \rightarrow +\infty} \|f - P_j(f)\|_{L^2([0,1])} = 0$, ou $\lim_{j \rightarrow +\infty} P_j(f) = f$ dans $L^2([0, 1])$, ou de manière équivalente $\bigcup_{j \geq 0} V_j = L^2([0, 1])$.

3. L'ondelette de Haar dans $L^2([0, 1])$

Déduisons des $(\varphi_{j,k})$ une base hilbertienne de $L^2([0, 1])$. Posons, pour $f \in L^2([0, 1])$ et $j \geq 0$,

$$(6) \quad Q_j(f) = P_{j+1}(f) - P_j(f) \in V_{j+1}.$$

Proposition 3. Pour tous $j \geq 0$ et $k \in \{0, 1, \dots, 2^j - 1\}$, on a $\int_{I_{j,k}} Q_j(f)(t) dt = 0$, et cette fonction vérifie $Q_j(f)|_{I_{j+1,2k}} = -Q_j(f)|_{I_{j+1,2k+1}}$.

Démonstration. Le premier point résulte d'un calcul simple. Pour le second point, remarquons que $I_{j,k} = I_{j+1,2k} \cup I_{j+1,2k+1}$. Comme $Q_j(f) \in V_{j+1}$, $Q_j(f)$ est constante sur $I_{j+1,2k}$ et $I_{j+1,2k+1}$. Or $0 = \int_{I_{j,k}} Q_j(f)(t) dt = \int_{I_{j+1,2k}} Q_j(f)(t) dt + \int_{I_{j+1,2k+1}} Q_j(f)(t) dt$. D'où la conclusion. \square

Fixons donc $j \geq 0$. Par la proposition 3, pour déterminer $Q_j(f)$, il suffit de calculer la valeur de $Q_j(f)$ sur chaque intervalle $I_{j+1,2k}$, pour $k \in \{0, 1, \dots, 2^j - 1\}$. Soit $k \in \{0, 1, \dots, 2^j - 1\}$ et $t \in I_{j+1,2k} \subset I_{j,k}$. En utilisant que $I_{j,k} = I_{j+1,2k} \cup I_{j+1,2k+1}$, on réécrit

$$(7) \quad Q_j(f)(t) = P_{j+1}(f)(t) - P_j(f)(t) = 2^{j/2} \int_{I_{j,k}} f(s) 2^{j/2} (\mathbf{1}_{I_{j+1,2k}}(s) - \mathbf{1}_{I_{j+1,2k+1}}(s)) ds.$$

Cette dernière égalité est remarquable :
appelons *ondelette-mère* la fonction $\psi : t \mapsto \mathbf{1}_{[0,1/2]}(t) - \mathbf{1}_{[1/2,1]}(t)$, et définissons, pour tout couple d'entiers (j, k) , l'ondelette $\psi_{j,k} : t \mapsto 2^{j/2} \psi(2^j t - k)$. Les ondelettes $(\psi_{j,k})$ sont des versions translatées et dilatées de l'ondelette-mère ψ . Ces fonctions sont appelées *ondelettes* car elles sont d'intégrale nulle et à support compact. Intuitivement, chaque ondelette $\psi_{j,k}$ oscille à la fréquence 2^j sur son support $I_{j,k}$. Elle est donc localisée (exactement) en temps et (approximativement) en fréquence.

Proposition 4. Pour tout $j \geq 0$, pour tout $t \in [0, 1]$,

$$(8) \quad Q_j(f)(t) = \sum_{k=0}^{2^j-1} d_{j,k}(f) \psi_{j,k}(t), \quad \text{où } d_{j,k}(f) = \langle f, \psi_{j,k} \rangle = 2^{j/2} \int_{I_{j,k}} f(s) \psi(2^j s - k) ds.$$

Démonstration. De (7), on déduit que sur l'intervalle $I_{j+1,2k}$, la fonction $t \mapsto Q_j(f)(t)$ est constante et vaut exactement $2^{j/2} \langle f, \psi_{j,k} \rangle = d_{j,k}(f) \psi_{j,k}$ (car la fonction $\psi_{j,k}$ est elle-même constante et égale à $2^{j/2}$ sur $I_{j+1,2k}$). De la proposition 3, on trouve que sur l'intervalle $I_{j+1,2k+1}$, $Q_j(f)$ est constante et vaut $-2^{j/2} \langle f, \psi_{j,k} \rangle$, qui est encore égal à $d_{j,k}(f) \psi_{j,k}$. \square

On a alors la propriété suivante :

Théorème 5. La famille de fonctions formée par $(\varphi, (\psi_{j,k})_{j \geq 0, k \in \{0,1,\dots,2^j-1\}})$ forme une base hilbertienne de fonctions de $L^2([0, 1])$.

Ainsi, toute fonction $f \in L^2([0, 1])$ se décompose en (l'égalité ayant lieu dans $L^2([0, 1])$)

$$(9) \quad f(t) = c_{0,0}(f) \varphi(t) + \sum_{j=0}^{+\infty} \sum_{k=0}^{2^j-1} d_{j,k}(f) \psi_{j,k}(t).$$

4. Étude numérique

4.1. Programmation pour un signal sur l'intervalle $[0, 1]$

Prenons $f \in L^2([0, 1])$. On rappelle que les $c_{j,k}(f)$ et les coefficients d'ondelettes $d_{j,k}(f)$ sont donnés par les formules (1) et (8). On appellera j la *génération* du coefficient $d_{j,k}(f)$, et $k2^{-j}$ sa *position*. Une fois les $c_{j,k}(f)$ trouvés, les coefficients d'ondelettes $d_{j,k}(f)$ se déduisent facilement. En effet, les formules liant les coefficients $c_{j,k}(f)$ et $d_{j,k}(f)$ sont les suivantes :

$$(10) \quad \text{pour tout } (j, k), \text{ on a } \begin{cases} c_{j,k}(f) &= \frac{1}{\sqrt{2}}(c_{j+1,2k}(f) + c_{j+1,2k+1}(f)), \\ d_{j,k}(f) &= \sqrt{2} c_{j+1,2k}(f) - c_{j,k}(f). \end{cases}$$

Avec un signal numérique $u = \left(f\left(\frac{k}{2^J}\right) \right)_{0 \leq k \leq 2^J-1}$ dans \mathbb{R}^{2^J} ($J \in \mathbb{N}^*$, on pourra prendre $J = 10$), on calcule $c_{j,k}$ pour $0 \leq k \leq 2^j - 1$, puis les coefficients $c_{j,k}$ et $d_{j,k}$ avec (10) pour $0 \leq j \leq J - 1$ et $0 \leq k \leq 2^j - 1$. Si nécessaire on pose $c_{j,k+2^j} = c_{j,k}$ et $d_{j,k+2^j} = d_{j,k}$. Dans le cas où c'est pertinent on peut prendre $c_{J,k} = \frac{1}{2^{J/2}} f\left(\frac{k}{2^J}\right)$.

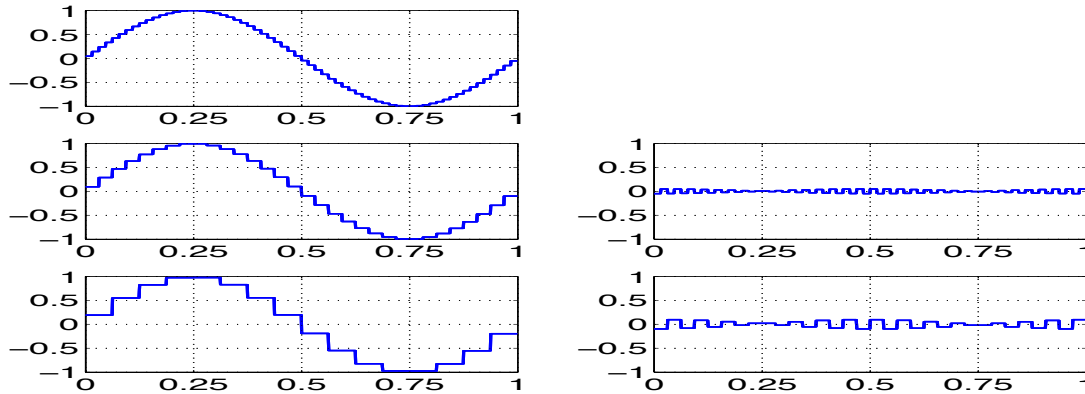


FIGURE 2. **Gauche, de haut en bas :** Les fonctions $P_6(f)$, $P_5(f)$, $P_4(f)$. **Droite, de haut en bas :** Les fonctions $Q_5(f) = P_6(f) - P_5(f)$ et $Q_4(f)$.

4.2. Quelques heuristiques

Les coefficients d'ondelettes sont en quelque sorte les coordonnées de la fonction f dans la base d'ondelettes : si l'ondelette est fixée, la connaissance dans L^2 de la fonction f est équivalente à la connaissance de ses coefficients $c_{0,0}(f)$ et $d_{j,k}(f)$.

On pourra se convaincre sur plusieurs exemples qu'en pratique, il existe un petit nombre de coefficients d'ondelettes très grands (en valeur absolue) par rapport aux autres. On pourra étudier les cas d'une fonction de Heaviside $H : t \mapsto \mathbf{1}_{[1/3,1]}(t)$ ou d'un créneau $C : t \mapsto \mathbf{1}_{[1/3,2/3]}(t)$, pour lesquels, à chaque génération j , seulement un nombre borné (plus petit que 4) de coefficients $d_{j,k}(f)$ est non nul. On pourra s'intéresser également à la fonction $x \mapsto \sqrt{x}$, dont les coefficients d'ondelettes $d_{j,k}(f)$ les plus gros se concentrent autour de $k = 0$.

Il est alors naturel de se demander quelle est l'erreur commise lorsque l'on considère, au lieu de f , la fonction construite à partir des coefficients d'ondelettes de f , mais dont on ne retient qu'un nombre fini de coefficients, les plus gros. L'intuition nous dit que l'on obtient ainsi une bonne approximation de f et l'idée de « compression » du signal f apparaît alors clairement.

5. Propriétés des coefficients d'ondelettes : Application à la compression

Pour simplifier, on suppose dans cette partie que $\int_{[0,1]} f(t) dt = 0$, autrement dit $c_{0,0}(f) = 0$. Pour $p \geq 1$, on notera L^p au lieu de $L^p([0,1])$, et $\|f\|_p$ la norme dans L^p .

D'après l'égalité de Parseval, pour toute fonction $f \in L^2$, on utilise (9) pour obtenir

$$(11) \quad \|f\|_{L^2}^2 = \int_{[0,1]} |f(t)|^2 dt = \sum_{j \geq 0} \sum_{k=0}^{2^j-1} |d_{j,k}(f)|^2.$$

L'énergie de la fonction est égale à la somme des carrés des coefficients d'ondelettes. En particulier, les coefficients $d_{j,k}(f)$ convergent vers 0 lorsque $j \rightarrow +\infty$, indépendamment de k .

On peut s'interroger sur la vitesse de convergence des coefficients vers 0 en fonction de f . Rappelons que f est höldérienne d'exposant $0 < \alpha < 1$ sur $[0, 1]$ (noté $f \in C^\alpha$) si

$$(12) \quad \text{il existe } C > 0 \text{ telle que pour } (x, y) \in [0, 1]^2, \quad x \neq y \Rightarrow |f(x) - f(y)| \leq C|x - y|^\alpha.$$

Théorème 6. Soit $f \in C^\alpha$ pour un certain $0 < \alpha < 1$. Alors il existe une constante $C' > 0$ telle que pour tout couple (j, k) tel que $j \geq 1$ et $k \in \{0, 1, \dots, 2^j - 1\}$, on a

$$(13) \quad |d_{j,k}(f)| \leq C' 2^{-j(\alpha+1/2)}.$$

Réciproquement, si (13) est vérifiée pour tous les couples (j, k) , alors $f \in C^\alpha$.

Pour démontrer le sens direct du théorème 6, on pourra utiliser que pour tout couple (j, k) , $\int_{[0,1]} f(k2^{-j}) \psi_{j,k}(t) dt = 0$. Pour le sens réciproque, lorsque x et y sont fixés dans $[0, 1]$, on pourra introduire l'unique entier J_0 tel que $2^{-J_0} \leq |x - y| < 2^{-J_0+1}$, décomposer $f(x) - f(y)$ grâce à (9) et couper alors la somme suivant $j \leq J_0$ et $j > J_0$.

Ainsi, plus la fonction f est régulière, plus les coefficients d'ondelettes décroissent vite vers 0.

Proposition 7. Soit $f \in C^\alpha$ pour un certain $0 < \alpha < 1$. Alors il existe une constante $C > 0$ telle que pour tout entier $J \geq 1$, $\|f(t) - P_J(f)(t)\|_{L^2} \leq C 2^{-J\alpha}$.

Cette proposition résulte du théorème 6 et de l'égalité de Parseval (11).

Suggestions et pistes de réflexion

► Les pistes de réflexion suivantes ne sont qu'indicatives et il n'est pas obligatoire de les suivre. Vous pouvez choisir d'étudier, ou non, certains des points proposés, de façon plus ou moins approfondie, mais aussi toute autre question à votre initiative. Vos investigations comporteront une partie traitée sur ordinateur et, si possible, des représentations graphiques de vos résultats. À défaut, si vos illustrations informatiques n'ont pas abouti, il est conseillé d'expliquer ce que vous auriez souhaité mettre en œuvre.

- On pourra compléter les diverses affirmations et preuves des sections 2 et 3 (notamment le théorème 1).
- On pourra tracer des exemples de fonctions $\varphi_{j,k}$ et d'ondelettes $\psi_{j,k}$ pour différentes valeurs de j et k , et calculer les coefficients $c_{j,k}(f)$ et $d_{j,k}(f)$ ainsi que les approximations $P_j(f)$ et $Q_j(f)$ pour la fonction $t \mapsto \sin(2\pi t)$ sur $[0, 1]$ et pour d'autres fonctions.
- Dans la partie 4, on pourra illustrer numériquement les assertions de la section 4.2.
- On pourra démontrer le théorème 6 et la proposition 7.
- On pourra vérifier numériquement sur une fonction C^α puis C^1 les résultats de la partie 5.