

Projet de mathématiques
:: Licence - Semestre 6 ::
Compression d'un signal par les ondelettes de Haar

Gaëtan DELERS et Vincent NOLOT

24 mai 2008

Table des matières

I. Introduction	4
1 Présentation du projet	5
2 Théorie des ondelettes	5
2.1 Historique	5
2.2 Limites de la Transformée de Fourier	5
2.3 Théorie générale	6
2.4 Multirésolutions	7
II. Généralités mathématiques	8
3 Présentation du cadre de travail	9
3.1 Définitions et Théorèmes importants	9
4 Explications sur la modélisation d'un signal en 1D	11
4.1 Réception et analyse du signal	11
4.2 Principe de la base de Haar	11
4.3 Compression et reconstruction du signal	11
5 Application : modélisation théorique d'un signal en 1D	12
5.1 Réception du signal	12
5.2 Construction de la base de Haar	14
5.3 Décomposition sur la base de Haar	16
5.4 Algorithme de calcul des $d_{j,k}$	16
5.5 Principe de compression	17
5.6 Algorithme de reconstruction	17
III. Application des ondelettes	19
6 Exemples en dimension 1	20
6.1 Compression d'un signal en 1D	20
6.2 Débruitage d'un signal en 1D	20
7 Exemples en dimension 2	20
7.1 Compression d'un signal en 2D	20
7.2 Débruitage d'un signal en 2D	21
IV. Conclusion et ouverture	23

8 Conclusion et ouverture	24
8.1 Des ondelettes plus modernes	24
8.2 ... encore peu utilisées	25
8.3 ... et optimisables.	25
9 Synthèse	25
A Bibliographie	27
B Procédures Maple	27
B.1 DecompCanonique	27
B.2 DecompOndelettes	27
B.3 Compression	28
B.4 Reconstruction	28

Première Partie
Introduction

1 Présentation du projet

2 Théorie des ondelettes

→ Dans le cadre de l'unité d'enseignement "Projet" de 3ème année de licence de mathématiques nous avons voulu utiliser des outils d'Analyse Fonctionnelle. Après la lecture d'un ouvrage, il nous est apparu intéressant de nous pencher sur la théorie des ondelettes. Notre principale motivation est que ce domaine mathématique trouve de nombreuses applications dans la vie quotidienne. M. Batude nous a conseillé de nous orienter vers les ondelettes de Haar, qui sont parmi les plus simples.

→ Nous allons au cours de ce compte-rendu vous présenter d'une part les outils mathématiques nécessaires au développement des ondelettes (pour bien comprendre le fondement de la théorie), puis d'autre part vous expliquer le déroulement de l'algorithme que nous avons réalisé sous Maple. Pour finir nous réaliserons une ouverture sur les autres méthodes et leurs performances respectives dans les mêmes domaines d'applications que les ondelettes.

2.1 Historique

→ L'analyse par ondelettes a été introduite au début des années 1960, dans un contexte d'analyse du signal et d'exploration pétrolière. Il s'agissait à l'époque de donner une représentation des signaux permettant de faire apparaître simultanément des informations temporelles (localisation dans le temps, durée) et fréquentielles, facilitant par là l'identification des caractéristiques physiques de la source du signal. Les ondelettes n'ont depuis, cessé de se développer et de trouver de nouveaux champs d'application.

2.2 Limites de la Transformée de Fourier

→ On dit que l'analyse d'un signal par les ondelettes est similaire à l'analyse de Fourier puisqu'elle transforme aussi un signal en ses principaux constituants afin de l'analyser. Cependant, la transformée de Fourier ne suffit plus ici. Elle est en effet incapable de détecter quelles portions du signal varient lentement ou rapidement, or nous avons besoin de distinguer les zones riches en information de celles qui ne le sont pas.

→ La DWT (*Discret Wavelet Transform*) est une méthode de compression, à base d'ondelettes, utilisée par le format émergent (en passe de devenir un standard international) JPEG 2000, successeur du JPEG, qui lui est à base de DCT (*Discrete Cosine Transform*). Grâce à l'utilisation des ondelettes, ce format bénéficie d'une compression 50 à 100 fois supérieure à son ancêtre JPEG, tout en conservant une bien meilleure définition des détails sur l'image finale.

→ Cas des images Il est avéré que les ondelettes apportent de bien meilleurs résultats que l'algorithme le plus utilisé actuellement en imagerie multimédia pour la compression : la DCT. En effet, contrairement à la DCT, la DWT s'applique à la totalité de l'image et non pas à des blocs de pixels ce qui permet d'éviter l'apparition de carrés

uniformes lorsque le taux de compression est élevé. De plus, l'utilisation d'une ondelette réversible permet une compression sans perte de données, ce qui n'était pas possible avec le format JPEG par exemple. Voici une comparaison de ces deux méthodes de compression.

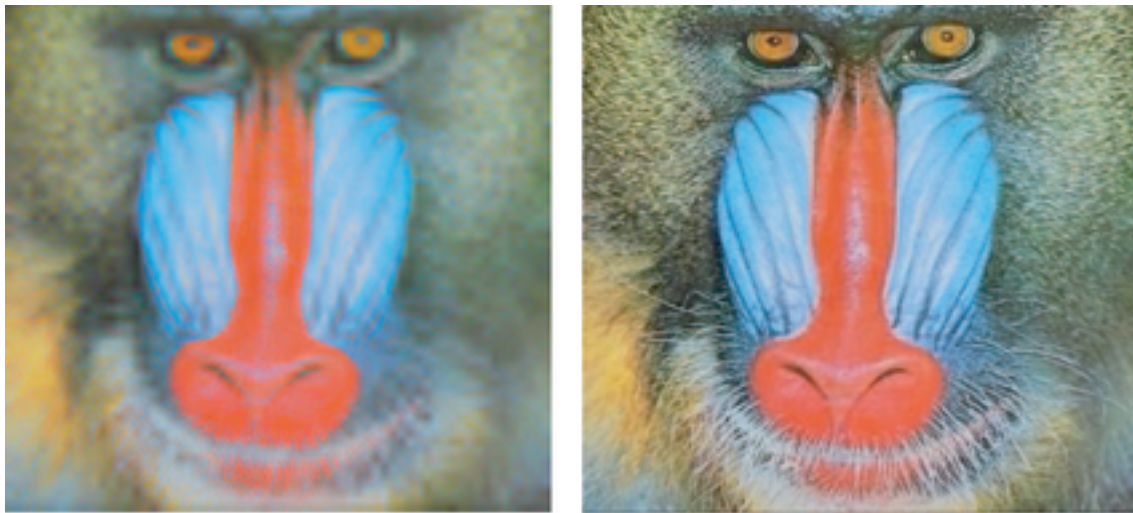


Image compressée avec JPEG (DCT)
(Transformée de Fourier)

Image compressée avec JPEG2000 (DWT)
(Transformée par Ondelettes)

→ Si on modifie une valeur d'un coefficient de la transformée de Fourier d'un signal tout le signal va se déformer. Maintenant si on fait la même chose mais en passant par la transformée en ondelettes on ne va dégrader qu'une partie du signal.

2.3 Théorie générale

Définition 2.1 Une ondelette est un objet mathématique utilisé pour le traitement du signal.

Définition 2.2 La transformée en ondelettes est un opérateur linéaire, invariant par translation et par dilatation. Quelle que soit l'échelle et quel que soit l'endroit, l'analyse du signal se fait avec la même fonction (l'ondelette mère).

→ L'idée est de construire une représentation qui fait à la fois apparaître des informations temporelles et fréquentielles. Ce type de représentation peut par exemple aider à représenter la structure physique de la structure géologique à l'origine du signal. Dans ce type de graphique les variations de fréquence vont représenter ces fameuses zones, contrastées ou pas. Selon le type de zone détectée, la compression ne sera pas appliquée de la même façon.

→ Avec les ondelettes, le signal est découpé en différents morceaux qui sont des versions translatées et dilatées d'une même fonction (l'*ondelette mère*). Il en résulte une superposition d'ondelettes décalées et dilatées qui ne diffèrent entre elles que par leur taille. On obtient une *transformée en ondelettes*, fonction composée de deux variables : le temps et la fréquence.

→ Le point fort de cette technique est que les ondelettes s'adaptent en fonction des caractéristiques recherchées : hautes fréquences (l'ondelette est très fine) ou basses fréquences (l'ondelette s'étire). On parlera alors de *multirésolution*.

2.4 Multirésolutions

→ Une multirésolution est une famille de sous-espaces fermés de $L^2(\mathbb{R})$ particulière qui est un outil de construction de bases d'ondelettes. A partir d'une multirésolution choisie selon les caractéristiques du signal traité, on pourra travailler avec différentes ondelettes : ondelettes de Haar (dans notre cas), ondelettes de Daubechies.

→ L'une des raisons essentielles du succès rencontré par les méthodes basées sur la transformation de Fourier tient dans l'existence d'algorithmes rapides de calcul qui lui sont associés (la fameuse FFT : *Fast Fourier Transform*). Or, il s'est avéré que les transformations en ondelettes discrètes, pour peu que l'ondelette soit convenablement choisie, sont naturellement associées à des algorithmes qui peuvent être encore plus efficaces que les algorithmes de FFT.

Les coefficients d'ondelettes d'un signal sont obtenus à partir d'une série de lissages du signal, à des résolutions de plus en plus grossières. La remarque clé est que ces lissages peuvent être effectués de façon récursive, en utilisant de façon systématique un unique opérateur de lissage.

Deuxième Partie
Généralités mathématiques

3 Présentation du cadre de travail

3.1 Définitions et Théorèmes importants

Rappels On rappelle qu'un ensemble E est un \mathbb{K} -espace vectoriel s'il existe deux opérations $+$ et \cdot telles que :

- $(E, +)$ est un groupe abélien
- \cdot est une loi externe telle que $\forall \lambda \in \mathbb{K}$, si $u \in E$ alors $\lambda.u \in E$
- On a les propriétés suivantes :

1 est un neutre à gauche : $\forall u \in E, 1.u = u$
 \cdot est distributive à gauche par rapport à $+$ (de E) : $\forall \lambda \in \mathbb{K}$ et $u, v \in E$ on a $\lambda.(u+v) = \lambda.u + \lambda.v$
 \cdot est exo-distributive à droite par rapport à $+$ (de \mathbb{K}) : $\forall \lambda, \mu \in \mathbb{K}, \forall u \in E$ on a $(\lambda+\mu).u = \lambda.u + \mu.u$.
 \cdot est exo-associative par rapport à x (de \mathbb{K}) : $\forall \lambda, \mu \in \mathbb{K}, \forall u \in E$, on a $(\lambda.\mu).u = \lambda.(.\mu.u)$.

Définition 3.1 (*Produit scalaire*) Une application bilinéaire $P : E \times E \longrightarrow \mathbb{K}$ est un produit scalaire sur E si elle vérifie :

- $\forall (u, v) \in E^2, P(u, v) = P(v, u)$ (commutativité)
- $\forall (u, v, w) \in E^3, P(u, v+w) = P(u, v) + P(u, w)$ (associativité)
- $\forall (\lambda, u, v) \in \mathbb{K} \times E^2, P(\lambda.u, v) = P(u, \lambda.v) = \lambda P(u, v)$
- $\forall u \in E^2, P(u, u) \leq 0$ et $P(u, u) = 0 \Leftrightarrow u = 0$ (définie positive)

De plus si $(E, +, \cdot)$ est muni d'un produit scalaire, alors E est dit préhilbertien.

Définition 3.2 (*Base hilbertienne*) Si H est un espace Hilbertien de dimension $n \in \mathbb{N}^*$. On dit que $\mathcal{B} = (e_1, \dots, e_n)$ est une base (système libre et générateur de H) hilbertienne ou (de Hilbert) si :

- $\forall i \neq j \in \{1, \dots, n\}$ on a $\langle e_i, e_j \rangle = 0$ (propriété d'orthogonalité)
- $\forall i \in \{1, \dots, n\}$ on a $\langle e_i, e_i \rangle = \|e_i\|^2 = 1$ (propriété de normalité)
- $\forall x \in H, \exists (\lambda)_{i \in \{1, \dots, n\}}$ telle que $\sum_{i=1}^n \lambda_i . e_i = x$.

Définition 3.3 (*Semi-norme*) Une semi-norme sur un \mathbb{K} -espace vectoriel E est une application $N : E \longrightarrow \mathbb{R}^+$ vérifiant :

- $\forall (\lambda, x) \in \mathbb{K} \times E, N(\lambda.x) = |\lambda|.N(x)$ (homogénéité)
- $\forall (x, y) \in E^2, N(x + y) \neq N(x) + N(y)$ (sous-additivité)

Remarque(s) : N devient une *norme* si en plus de ces axiomes, elle respecte la propriété de séparation : $\forall x \in E, N(x) = 0 \Rightarrow x = 0_E$.

→ Après ces rappels élémentaires, nous énonçons maintenant les définitions qui nous seront utiles tout au long de notre présentation.

Définition 3.4 (Espace \mathcal{L}^p) Soient E un ensemble, τ une tribu sur E et μ une mesure sur τ . Pour $p \in [1, +\infty[$, on désigne par $\mathcal{L}^p(E, \tau, \mu)$ l'ensemble des applications $f : E \rightarrow \mathbb{C}$ mesurables telles que

$$\int_E |f|^p d\mu < +\infty$$

On définit alors pour $f \in \mathcal{L}^p(E, \tau, \mu)$ une semi-norme telle que

$$\|f\|_p = \left(\int_E |f|^p d\mu \right)^{\frac{1}{p}}$$

Remarque(s) : Soit $p \in [1, +\infty]$. Sur $\mathcal{L}^p(E, \tau, \mu)$, la relation " $f = g \ \mu - p.p$ " ($\#$) est une relation d'équivalence :

- $f = f \ \mu - p.p.$ (réflexivité)
- $f = g \ \mu - p.p.$ et $g = h \ \mu - p.p. \Rightarrow f = h \ \mu - p.p.$ (transitivité)
- $f = g \ \mu - p.p. \Rightarrow g = f \ \mu - p.p.$ (symétrie)

→ Pour $f \in \mathcal{L}^p(E, \tau, \mu)$, on note $\tilde{f} = \{g \in \mathcal{L}^p(E, \tau, \mu); f = g \ \mu - p.p.\}$ la classe d'équivalence de f . On se placera dans le cas où $E = \mathbb{R}$, $\tau = \mathcal{L}(\mathbb{R})$ la tribu de Lebesgue sur \mathbb{R} et $\mu = \lambda$ la mesure de Lebesgue.

Définition 3.5 (Espace L^p) On désigne par $L^p(\mathbb{R}, \mathcal{L}(\mathbb{R}), \lambda)$ l'ensemble des classes d'équivalences de $\mathcal{L}^p(\mathbb{R}, \mathcal{L}(\mathbb{R}), \lambda)$ pour ($\#$). On a donc :

$$L^p(\mathbb{R}, \mathcal{L}(\mathbb{R}), \lambda) = \{\tilde{f}; f \in \mathcal{L}^p(\mathbb{R}, \mathcal{L}(\mathbb{R}), \lambda)\}$$

THÉORÈME 3.1 L'espace $L^2([0, 1], \mathcal{L}(\mathbb{R}), \lambda)$ est un espace Euclidien pour le produit scalaire

$$(f, g) \mapsto \langle f|g \rangle = \int_0^1 fg d\mu \tag{1}$$

→ Nous travaillerons dorénavant dans l'espace $L^2([0, 1]) = L^2([0, 1], \mathcal{L}(\mathbb{R}), \lambda)$.

4 Explications sur la modélisation d'un signal en 1D

4.1 Réception et analyse du signal

→ Au départ, on reçoit un signal sous la forme d'une application f continue en fonction du temps t . C'est un signal à une dimension (1D). On regardera par la suite le signal sur l'intervalle $I = [0, 1]$. Pour modéliser $f(t)$, on commence par analyser la fonction sur des petits morceaux d'intervalles : il s'agit de discrétiser l'intervalle I en une subdivision σ_n à n parties régulières.

→ On se donne ensuite la base canonique $\{\varphi_{n,k}\}_{k=0}^{2^n-1}$, de l'espace vectoriel des fonctions en escalier associée à σ_n de I , espace que l'on notera V_n . On pourra alors décomposer la fonction f sur cette base orthonormée et associer ainsi à chaque partie de σ_n , un coefficient correspondant à f , qui sera noté $c_{n,k}$.

Remarque(s) : La famille $\{V_j\}_{j=0}$, forme la multi-résolution de Haar.

4.2 Principe de la base de Haar

→ A partir de la base de V_n précédemment définie, on va construire des nouvelles fonctions $\psi_{n,k}$ appelées fonctions de Haar (voir le paragraphe 3.2), de telle façon que $\{\varphi_{0,0}, \psi_{n,k}\}$ soit une base Hilbertienne et orthonormée de $L^2([0, 1])$. Il nous suffira alors de changer de base pour les coefficients $c_{n,k}$ initialement calculés dans la base canonique. Nous obtiendrons des coefficients $d_{n,k}$.

→ De plus, cette décomposition aura pour avantage, de par la forme des fonctions de Haar $\psi_{n,k}$, de garder une grande partie des variations fréquentielles du signal.

4.3 Compression et reconstruction du signal

→ Une fois décomposé, on procédera à la compression du signal, ce qui signifie mettre à 0 les coefficients que l'on considérera négligeable (c'est-à-dire dont la valeur absolue est en-dessous de 0.05 par exemple).

→ Enfin, il ne restera plus qu'à recomposer le signal f dans la base canonique : on obtiendra alors le graphe d'une fonction que l'on pourra comparer au signal de départ. On remarquera que plus la subdivision de découpe initiale σ_n est fine sur $[0, 1]$, plus la recombinaison se rapprochera de f . On aura cependant compressé les coefficients, ce qui rendra le signal recomposé moins "lourd".

5 Application : modélisation théorique d'un signal en 1D

5.1 Réception du signal

→ On pose $I = [0, 1[$ et on décompose I en développement diadique en considérant pour $n \in \mathbb{N}$ (que restera fixé pour le reste de notre énoncé) et $0 \leq k \leq 2^n - 1$, les intervalles

$$I_{n,k} = \left[\frac{k}{2^n}, \frac{k+1}{2^n} \right[$$

Ainsi on obtient l'union disjointe

$$I = \bigsqcup_{k=0}^{2^n-1} I_{n,k}$$

Avec cette décomposition, on peut considérer les fonctions $\varphi_{n,k} : I_{n,k} \longrightarrow \{0, 2^{\frac{n}{2}}\}$ telles que :

$$\varphi_{n,k} = 2^{\frac{n}{2}} \mathcal{X}_{I_{n,k}}$$

où \mathcal{X} est la fonction indicatrice habituelle.

Ainsi définie, il est évident que la famille $\{\varphi_{n,k}; k = 0, \dots, 2^n - 1\}$ est un système libre donc forme une base sur l'espace qu'elles engendrent, qu'on notera par la suite $V_n \subset L^2([0, 1])$.

Nous pouvons vérifier que cette base est *orthonormée*.

Démonstration :

■ Vérifions les deux propriétés d'une base.

– Pour la norme de $L^2([0, 1])$ induite sur V_n on a pour la mesure de Lebesgue λ :

$$\|\varphi_{n,k}\| = \int_I |\varphi_{n,k}|^2 d\lambda = \int_I (2^{\frac{n}{2}})^2 \mathcal{X}_{I_{n,k}} d\lambda = \int_{\frac{k}{2^n}}^{\frac{k+1}{2^n}} 2^n d\lambda = 2^n \left(\frac{k+1}{n} - \frac{k}{n} \right) = 1$$

Cela signifie que les $\varphi_{n,k}$ sont de norme 1.

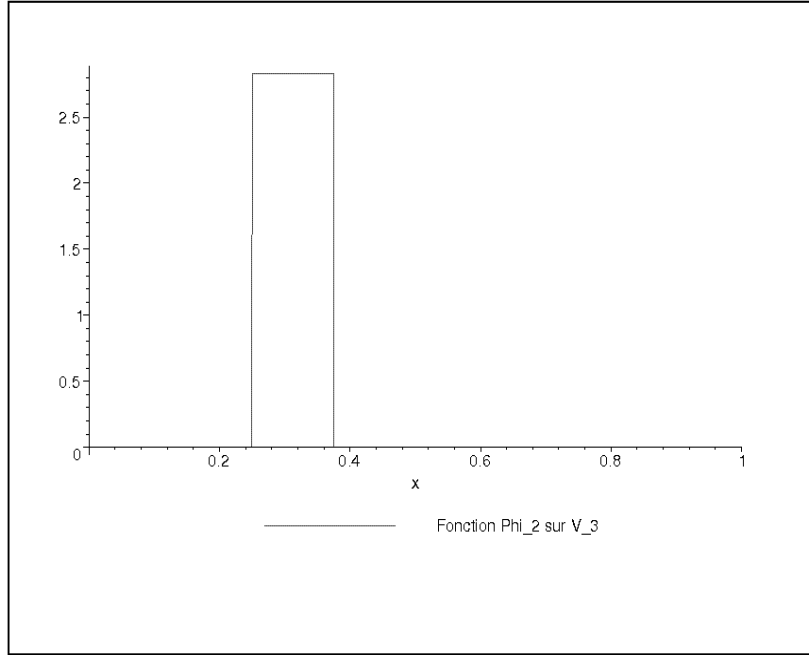
– Le produit scalaire induit par cette norme est donné par (1). On obtient donc pour $k_1 \neq k_2$ et $k_1, k_2 = 0, \dots, 2^n - 1$:

$$\langle \varphi_{n,k_1} | \varphi_{n,k_2} \rangle = \int_I \varphi_{n,k_1} \varphi_{n,k_2} d\lambda = \int_I 2^{\frac{n}{2}} \mathcal{X}_{I_{n,k_1}} 2^{\frac{n}{2}} \mathcal{X}_{I_{n,k_2}} d\lambda = 0$$

si $k_1 \neq k_2$ car I_{n,k_1} et I_{n,k_2} sont disjoints. Cela signifie que les $\varphi_{n,k}$ sont orthogonaux (par rapport à k).

■

→ Voici le graphe d'une fonction $\varphi_{n,k}$ ainsi définie :

**Remarque(s) :**

- On a $\dim(V_n) = 2^n$
- $\varphi_{0,0} = 2^{\frac{0}{2}} \mathcal{X}_{I_{0,0}} = \mathcal{X}_I$
- Si $f \in V_n$ alors $f = \sum_{k=0}^{2^n-1} \langle f | \varphi_{n,k} \rangle \varphi_{n,k}$
- On a $I_{n+1,2k} \sqcup I_{n+1,2k+1} = I_{n,k}$ et $V_n \subset V_{n+1} \forall n \in \mathbb{N}^*$

On peut aussi écrire d'après la définition $\forall n \in \mathbb{N}^*, \forall k \in \mathbb{N}; k = 0, \dots, 2^n - 1$:

$$\varphi_{n,k} = \frac{1}{\sqrt{2}}(\varphi_{n+1,2k} + \varphi_{n+1,2k+1}) \quad (2)$$

et on note alors

$$c_{n,k} = \langle f | \varphi_{n,k} \rangle = \int_I f(t) \cdot \varphi_{n,k}(t) d\lambda(t) = \frac{1}{\sqrt{2}}(c_{n+1,2k} + c_{n+1,2k+1}) \quad (3)$$

Cf Annexe A.1 (Procédure de décomposition sur la base canonique)

5.2 Construction de la base de Haar

→ On a vu que $V_n \subset V_{n+1}$. Cette inclusion est stricte, on peut donc considérer W_n le supplémentaire orthogonal de V_n dans V_{n+1} , c'est-à-dire : $V_{n+1} = V_n \oplus W_n$. On a alors

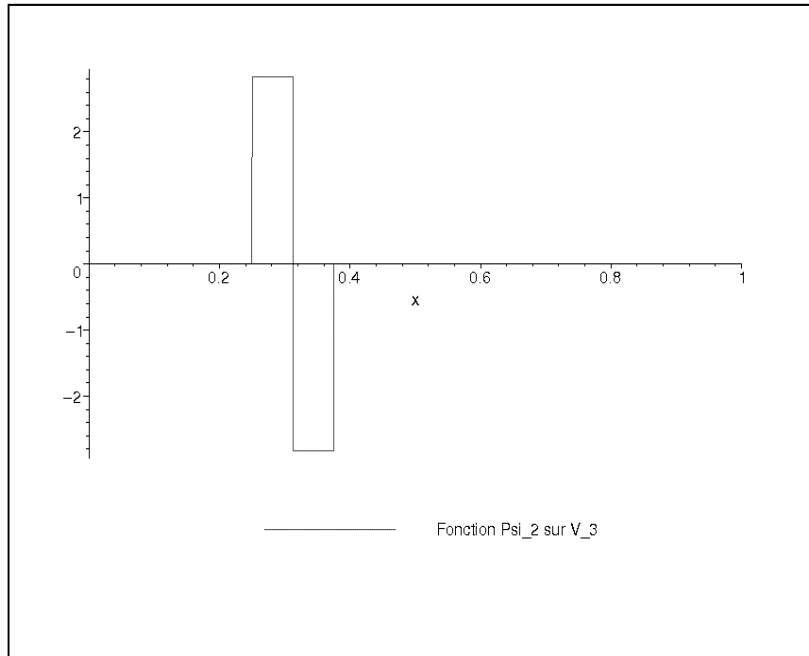
$$V_n = V_0 \oplus_{j=1}^{n-1} W_j$$

Comme $\dim(V_{n+1}) = 2^{n+1}$, on a que $\dim(W_n) = 2^n = \dim(V_n)$. On définit maintenant des fonctions de $\psi_{n,k} : I_{n,k} \rightarrow]-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}[$ pour $k = 0, \dots, 2^n - 1$ telles que

$$\psi_{n,k} = \frac{1}{\sqrt{2}}(\varphi_{n+1,2k} - \varphi_{n+1,2k+1}) \quad (4)$$

Par définition la famille des $\{\psi_{n,k}, k = 0, \dots, 2^n - 1\}$ est une *base orthonormée* de W_n . Ainsi la nouvelle base $\mathcal{H}_n = \{\varphi_{0,0}\} \cup_{j=0}^{n-1} \{\psi_{j,k}; k = 0, \dots, 2^j - 1\}$ est dite *base de Haar* de V_n .

→ Voici le graphe d'une fonction $\psi_{n,k}$ ainsi définie :



On vérifie facilement que $\text{Card}(\mathcal{H}_n) = 1 + \sum_{j=0}^{n-1} 2^j = 1 + 2^n - 1 = 2^n = \dim(V_n)$ et que \mathcal{H}_n est une *base orthonormée* de V_n .

Démonstration :

■ Vérifions les deux propriétés d'orthonormalité.

– On sait déjà que $\varphi_{0,0}$ est de norme 1. Pour $k = 1, \dots, 2^n - 1$ on a :

$$\begin{aligned}
 \|\psi_{n,k}\| &= \int_I |\psi_{n,k}|^2 d\lambda \\
 &= \frac{1}{2} \int_I |\varphi_{n+1,2k} - \varphi_{n+1,2k+1}|^2 d\lambda \\
 &= \frac{1}{2} \left(\int_I |\varphi_{n+1,2k}|^2 d\lambda + \int_I |\varphi_{n+1,2k+1}|^2 d\lambda - 2 \int_I \varphi_{n+1,2k} \varphi_{n+1,2k+1} d\lambda \right) \\
 &= \frac{1}{2} (1 + 1 + 0) \\
 &= 1
 \end{aligned}$$

On en déduit que les $\psi_{n,k}$ et $\varphi_{0,0}$ éléments de \mathcal{H}_n sont de norme 1.

– Pour $k_1 \neq k_2$ et $k_1, k_2 = 1, \dots, 2^n - 1$ on a :

$$\begin{aligned}
 \langle \psi_{n,k_1} | \psi_{n,k_2} \rangle &= \int_I \psi_{n,k_1} \psi_{n,k_2} d\lambda \\
 &= \frac{1}{2} \int_I (\varphi_{n+1,2k_1} - \varphi_{n+1,2k_1+1})(\varphi_{n+1,2k_2} - \varphi_{n+1,2k_2+1}) d\lambda \\
 &= \frac{1}{2} \left(\int_I \varphi_{n+1,2k_1} \varphi_{n+1,2k_2} d\lambda - \int_I \varphi_{n+1,2k_1} \varphi_{n+1,2k_2+1} d\lambda \right. \\
 &\quad \left. - \int_I \varphi_{n+1,2k_1+1} \varphi_{n+1,2k_2} d\lambda + \int_I \varphi_{n+1,2k_1+1} \varphi_{n+1,2k_2+1} d\lambda \right) \\
 &= 0
 \end{aligned}$$

De plus pour $k = 0, \dots, 2^n - 1$ on a :

$$\begin{aligned}
 \langle \varphi_{0,0} | \psi_{n,k} \rangle &= \int_I \varphi_{0,0} \psi_{n,k} d\lambda \\
 &= \frac{1}{\sqrt{2}} \int_I \varphi_{0,0} (\varphi_{n+1,2k} - \varphi_{n+1,2k+1}) d\lambda \\
 &= \frac{1}{\sqrt{2}} \left(\int_I \mathcal{X}_I \varphi_{n+1,2k} d\lambda - \int_I \mathcal{X}_I \varphi_{n+1,2k+1} d\lambda \right) \\
 &= \frac{1}{\sqrt{2}} \left(\int_{\frac{2k}{2^{n+1}}}^{\frac{2k+1}{2^{n+1}}} 2^{\frac{n+1}{2}} d\lambda - \int_{\frac{2k+1}{2^{n+1}}}^{\frac{2k+2}{2^{n+1}}} 2^{\frac{n+1}{2}} d\lambda \right) \\
 &= \frac{1}{\sqrt{2}} \left(2^{\frac{n+1}{2}} \left(\frac{1}{2^{n+1}} \right) - 2^{\frac{n+1}{2}} \left(\frac{1}{2^{n+1}} \right) \right) \\
 &= 0
 \end{aligned}$$

On en déduit que les $\psi_{n,k}$ et $\varphi_{0,0}$ sont orthogonaux. D'où le résultat.

■

Remarque(s) : La *base de Haar* est une base essentielle dans la décomposition car elle permet de mesurer les variations fréquentielles du fait des formes des fonctions de Haar.

5.3 Décomposition sur la base de Haar

→ On regarde notre signal $f(t)$. Si $f \in V_n$ alors sa décomposition sur la base de Haar est :

$$f = c_{0,0}\varphi_{0,0} + \sum_{j=1}^{n-1} \sum_{k=0}^{2^j-1} d_{j,k}\psi_{j,k} \quad (5)$$

où les coefficients sont donnés (d'après (2)) par :

$$d_{j,k} = \langle f | \psi_{j,k} \rangle = \frac{1}{\sqrt{2}}(c_{j+1,2k} - c_{j+1,2k+1}) \quad (6)$$

pour $j = 1, \dots, n$ et $k = 0, \dots, 2^j - 1$. Il faut maintenant expliciter comment calculer les coefficients $d_{j,k}$, c'est l'objet du prochain paragraphe.

Cf Annexe A.2 (Procédure de décomposition sur la base de Haar)

5.4 Algorithme de calcul des $d_{j,k}$

→ On se donne au début une fonction $f : I \rightarrow \mathbb{R}$ qui représente le signal. On va calculer les coefficients de f dans la *base de Haar* à partir de ses coefficients dans la base canonique.

Notons $L[n] = [c_{n,0}, \dots, c_{n,2^n-1}]$ la liste des coefficients $c_{n,k}$ rangés dans cet ordre (n étant fixé depuis le début).

▷ Etape 1 : On commence par calculer les coefficients $c_{n-1,k}$ puis $d_{n-1,k}$ pour $k = 0, \dots, 2^{n-1} - 1$ en utilisant respectivement (3) et (6). On replace alors ces coefficients dans la liste $L[n]$ dans l'ordre des k donnés. On obtient alors

$$L[n] = [L[n-1], [d_{n-1,0}, \dots, d_{n-1,2^{n-1}-1}]]$$

et on note $D[n-1] = [d_{n-1,0}, \dots, d_{n-1,2^{n-1}-1}]$. Cela nous rappelle la décomposition $V_n = V_{n-1} \oplus W_{n-1}$. Ainsi on obtient de (5) :

$$f = \sum_{k=0}^{2^{n-1}-1} c_{n-1,k}\varphi_{n-1,k} + \sum_{k=0}^{2^{n-1}-1} d_{n-1,k}\psi_{n-1,k}$$

▷ Etape 2 : On recommence la décomposition avec maintenant la liste $L[n-1]$. On retrouve ainsi $V_n = V_{n-2} \oplus W_{n-2} \oplus W_{n-1}$.

▷ Etape n-(j-1) : On calcule les coefficients $c_{j-1,k}$ puis $d_{j-1,k}$ pour $k = 0, \dots, 2^{j-1} - 1$ de la même façon avec (3) et (6). On reporte ces coefficients dans la liste $L[j]$ dans l'ordre. On obtient alors

$$L[n] = [L[j-1], D[j-1], D[j], \dots, D[n-1]]$$

avec les mêmes notations que précédemment. Et on retrouve une fois de plus la décomposition $V_n = V_{n-j} \oplus_{k=n-j}^{n-1} W_k$.

▷ Etape n : Ici $j = 1$. On reconstruit les seuls coefficients $c_{0,0}$ et $d_{0,0}$. On obtient la liste $L[n] = [c_{0,0}, L[0], \dots, L[n-1]]$. C'est en fait la liste des composantes de f sur la base \mathcal{H}_n .

5.5 Principe de compression

La compression consiste à mettre à 0 tous les coefficients $d_{n,k}$ qui seront considérés comme négligeables. On effectue cette compression *avant la reconstruction*.

Cf Annexe A.3 (Procédure de compression)

5.6 Algorithme de reconstruction

Intéressons-nous maintenant à la réciproque : on se donne une fonction f à composantes dans la *base de Haar* et on veut les recalculer dans la base canonique.

En utilisant (3) et (6), on peut écrire les coefficients suivants, pour $k = 0, \dots, 2^{j-1} - 1$:

$$c_{j,2k} = \frac{1}{\sqrt{2}}(c_{j-1,k} + d_{j-1,k})c_{j,2k+1} = 1\sqrt{2}(c_{j-1,k} - d_{j-1,k})$$

▷ Etape 1 : On calcule $c_{1,0}$ et $c_{1,1}$ en fonction de $c_{0,0}$ et $d_{0,0}$ et on les retranscrit dans la liste de la façon suivante :

$$L[n] = [c_{1,0}, c_{1,1}, D[1], \dots, D[j], \dots, D[n-1]]$$

▷ Etape j : On calcule $c_{j,2k}$ et $c_{j,2k+1}$ en fonction $c_{j-1,k}$ (calculé à l'étape $j-1$ et rangé dans la liste) et de $d_{j-1,k}$ pour $k = 0, \dots, 2^{j-1} - 1$. On les range ensuite dans la liste de telle façon à ce qu'on ait :

$$L[n] = [c_{j,0}, \dots, c_{j,2^j-1}, D[j], D[j+1], \dots, D[n-1]]$$

▷ Etape n : On calcule $c_{n,2k}$ et $c_{n,2k+1}$ en fonction de $c_{n-1,k}$ et $d_{n-1,k}$ pour $k =$

$0, \dots, 2^{n-1} - 1$. On les range de la même façon pour obtenir la liste des 2^n composantes de f sur base canonique :

$$L[n] = [c_{n,0}, \dots, c_{n,2^n-1}]$$

Cf Annexe A.4 (Procédure de reconstruction)

Troisième Partie
Application des ondelettes

6 Exemples en dimension 1

6.1 Compression d'un signal en 1D

- Compression d'un MP3.

→ Le principe de la *compression* d'un MP3 est le suivant : la décomposition en ondelettes du signal nous permet de séparer les basses fréquences et les hautes fréquences. Les basses étant inaudibles pour l'homme, on les détruit et on reconstruit le signal à partir des coefficients restants. On obtient une musique compressée qui produit le même son pour l'homme.

Musique au format .wav	Musique au format .mp3
Taille : environ 40 Mo	Taille : environ 3 Mo

6.2 Débruitage d'un signal en 1D

- Enlever les parasites d'une conversation brouillée.
- Radar, satellites, scanner, échographie.
- Rallonger ou réduire le tempo d'un signal audio.

→ Les ondelettes fournissent une méthode assez simple pour extraire les signaux du bruit blanc (bruit existant à toutes les fréquences). La difficulté du *débruitage* réside dans le fait que le bruit peut aussi bien être de haute fréquence (analyses médicales), que de basse fréquence (bruit produit par un sous-marin), ou encore des deux natures à la fois (images bruitées).

On transforme le signal en ondelettes, on élimine, à toutes les résolutions, les coefficients supérieurs à un certain seuil. Ensuite, on reconstruit le signal avec les coefficients restants.

7 Exemples en dimension 2

7.1 Compression d'un signal en 2D

- Compression d'une image (au même titre que le format JPEG).

→ Le passage à l'espace de dimension 2 peut être pensé d'une manière très simple : Si on a une image avec $N \times N$ pixels, où N est une puissance de 2, on revient à la définition des ondelettes en dimension 1 et on procède en deux étapes :

A. On transforme en ondelettes chaque ligne de l'image, 2 par 2 : on obtient une approximation de l'image et des détails.

B. On prend l'approximation de l'image et on transforme de nouveau en ondelettes, en procédant cette fois sur les colonnes.

Finalement on obtient une image d'approximations et 3 images de détails.

→ Voici un exemple d'images décomposées en ondelettes, puis reconstruites à partir des coefficients restants (ceux qui n'ont pas été détruits).



Compression avec 90%
des coefficients détruits



Compression avec 97%
des coefficients détruits

→ On rappelle que la transformée en ondelettes est à la base de l'algorithme de JPEG2000. L'essentiel de l'information d'un signal se situe dans ses structures irrégulières. Pour le traitement d'images, les contours sont caractérisés par la discontinuité de l'intensité.

→ La compression par ondelettes consiste à considérer les zones d'une image contenant de fortes variations du contraste comme des hautes fréquences, le reste de l'image étant de la basse fréquence. On extrait les hautes fréquences (qui sont les contours des objets) et on les garde telles quelles par une compression non destructrice, le reste étant réduit de façon destructrice puis compressé à nouveau par ondelettes. On constate en comparant les deux images précédentes, que lorsqu'on pousse la compression, on commence par perdre des détails là où les variations de contraste sont faibles.

7.2 Débruitage d'un signal en 2D

– Lisser une image.

→ Lisser une image revient à la rendre floue, c'est-à-dire à en diminuer la résolution. En fait on peut interpréter la reconstruction d'une image par les ondelettes de cette

façon : "l'image, à sa résolution la plus grande, est égale à la somme d'une version floue, et des détails apparaissant à des échelles différentes, c'est-à-dire à des résolutions différentes".

→ Voici un exemple de deux images : la première, originale avec une résolution assez haute et la seconde, lissée à une résolution beaucoup plus faible à partir de la première.

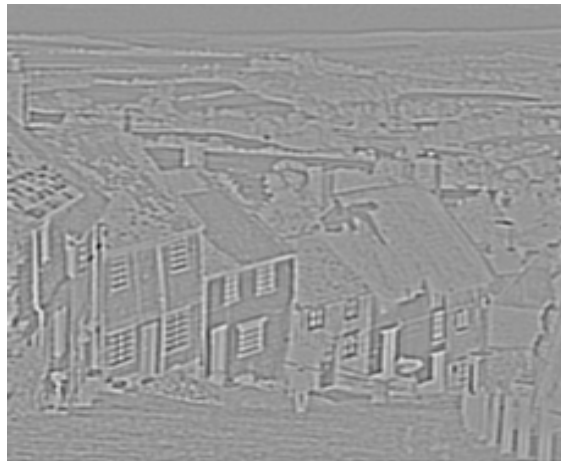


Image originale



Image lissée à résolution plus faible

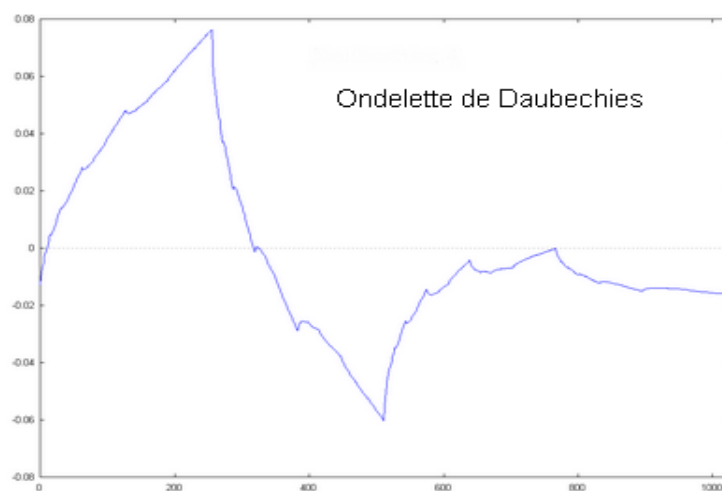
→ La différence entre les deux images correspond aux coefficients d'ondelettes : les détails présents dans la première image et absents dans la seconde.



Différence des deux images

Dernière Partie

Conclusion et ouverture



8 Conclusion et ouverture

8.1 Des ondelettes plus modernes ...

→ L'analyse en ondelettes a été étudiée pour le cas de Haar, bien qu'il en existe d'autres plus puissantes comme celles de Daubechies. La théorie reste la même pour ces ondelettes qui généralisent celles de Haar. On peut aussi trouver les ondelettes de Shannon qui sont moins répandues mais plus performantes selon leur domaine d'application. En effet la décomposition d'un signal particulier sera d'autant plus efficace si l'ondelette est correctement choisie en fonction.



Figure 6.11 – Image originale.

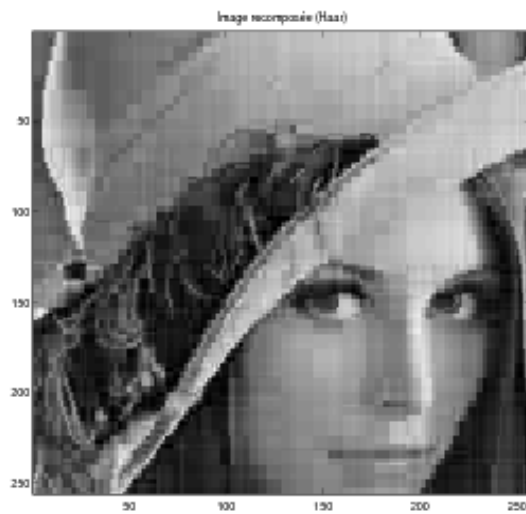


Figure 6.12 – Image reconstituée (Haar).



Figure 6.13 – Image reconstituée (Schauder).



Figure 6.14 – Image reconstituée (Daubechies).

8.2 ... encore peu utilisées ...

→ Les ondelettes bien que très efficaces dans des domaines comme la compression, ne sont pourtant encore que peu utilisées. Par exemple dans l'imagerie la plupart des appareils numériques et des sites internet utilisent encore le format JPEG. Cela s'explique en partie par le fait que le JPEG2000 est un format propriétaire (payant), contrairement au JPEG. Or aujourd'hui on imagine mal payer pour compresser une image alors que le JPEG offre malgré tout des résultats visuels convenables pour des taux de compression non excessifs.

8.3 ... et optimisables.

→ En dépit du grand nombre de succès qu'elle a pu rencontrer et de la diversité des applications qu'elle a trouvées, l'analyse par ondelettes est loin de nous donner une réponse universelle et finale au problème de représentation et codage des signaux. Cependant, elle fournit souvent une réponse qui n'est pas très éloignée de la réponse optimale, ce qui, combiné à la grande simplicité de l'outil et sa grande efficacité algorithmique (rapide), explique son succès grandissant. En revanche, dans certains cas bien précis, il est nécessaire de rechercher des techniques permettant de s'approcher plus encore de l'optimalité, tout en essayant de conserver les atouts des ondelettes.

→ En effet les ondelettes peuvent être vues comme un ensemble de signaux élémentaires, suffisant pour exprimer n'importe quel signal. En d'autres termes, elles fournissent un *dictionnaire de signaux élémentaires*, à partir duquel l'on peut reconstruire tout signal. Une des solutions possibles est donc de "répertoire", toutes les "méthodes efficaces" dans l'analyse d'un signal donné, et pour celui-ci de déterminer celle qui est la plus appropriée. Pour ce faire on utilise des ondelettes généralisées (appelés paquets d'ondelettes), qui contiennent non seulement les ondelettes usuelles, mais aussi une multitude d'autres procédés, tous aussi performants au niveau de l'algorithme. Cette méthodologie appelée *décomposition adaptative d'un signal* est actuellement mise à l'étude par les grandes compagnies pétrolières pour la compression des signaux récoltés lors des campagnes de mesure en prospection pétrolière.

9 Synthèse

→ Ce paragraphe termine notre compte-rendu. Dès le choix du sujet fixé, nous avons beaucoup fait de recherches sur Internet et lu un bon nombre d'articles portant sur la *Théorie des Ondelettes*. Nous nous sommes aperçu que le sujet était souvent traité. En général les articles étaient incomplets, parfois trop formels et parfois sans explications. C'est pourquoi nous avons décidé de présenter ici un rapport clair, réorganisé et précis sur la transformée d'un signal à l'aide des ondelettes de Haar. Le document de J-C. Cortet nous aura été d'une aide précieuse : nous l'avons décortiqué pour récupérer tout ce qui nous semblait nécessaire à notre présentation.

→ Nous nous sommes lancés immédiatement dans le compte-rendu et en particulier la partie *Généralités Mathématiques*. Il nous a donc fallu trier et rassembler les données essentielles parmi tous les articles qu'on avait recueillis. Ce projet nous a alors forcé à bien nous organiser : notre forum Internet (<http://objectifagreg.free.fr>) créé pour l'entre-aide sur les études mathématiques, nous a ainsi bien été utile. En effet, nous n'avions que rarement l'occasion de nous voir en dehors des horaires à l'université. Le forum nous permettait donc de nous tenir informés sur le recensement de tous les liens, les articles et les rédactions de parties du projet, qu'on écrivait parfois chacun de son côté. Le forum supportant le langage TeX, nous maîtrisons tous deux aujourd'hui l'écriture et le logiciel LaTeX.

→ Comme nous l'avons dit dans la présentation, en choisissant ce thème nous voulions aborder ces mathématiques en pensant à leurs applications dans la vie de tous les jours, tout en utilisant des outils d'analyse fonctionnelle. En effet quel utilisateur d'un ordinateur n'a jamais rencontré une image au format JPEG ou JPEG2000 ? Qui de nos jours n'a jamais écouté une musique au format MP3 ?

→ En plus du perfectionnement avec le logiciel Maple, le projet nous a permis de nous organiser et de travailler efficacement. Nous avons travaillé seuls, avec notre unique outil, qui n'est pas des moindres : Internet. Nous aurions aimé avoir plus de temps, pour s'attaquer aux algorithmes de compression en 2 dimensions plutôt que de se contenter d'exemples. Mais nous avons fait le choix de bien maîtriser les techniques en 1 dimension et en particulier l'algorithme de compression. Peut-être que le projet de L3 aurait intérêt, pour les étudiants qui souhaitent faire quelque chose de conséquent, à s'étendre sur les deux semestres ?

A Bibliographie

- Polycopié et fichier Maple fournis par Mr. Batude
- <http://fr.wikipedia.org>
- <http://www-ljk.imag.fr/membres/Valerie.Perrier/PUBLI/Cours3-VP.pdf>
- <http://www-igm.univ-mlv.fr/LabInfo/rapportsInternes/2007/03.pdf>
- <http://www.tsi.enst.fr/tsi/enseignement/ressources/mti/donoho/index.html>

B Procédures Maple

B.1 DecompCanonique

```
DecompCanonique:=proc(n, f, b)
  global i, k, phi, c:
  # On définit la base canonique des phi(n,k) de V_j
  phi:=(n, k, x)->piecewise(x>=k/(2^n) and x<(k+1)/(2^n), 2^(n/2), 0):
  # On calcule alors les composantes c(n,k) de f sur la base de V_n
  for k from 0 to 2^n-1 do # 2^n est la dimension de V_n donc le nombre de phi(n,k)
    c[n, k]:=int(f(x)*phi(n, k, x), x=0..1):
  od:
  if (b=1) then # Si on veut tracer le signal
    plot(sum(c[n, i]*phi(n, i, x), i=0..2^n-1), x=0..1, style=line, thickness=[2], title=
"Représentation de la décomposition canonique du signal", titlefont=[HELVETICA, BOLD, 10]);
  fi:
end:
```

B.2 DecompOndelettes

```
DecompOndelettes:=proc(n, c)
  local d, j, k:
  for k from 0 to 2^(n-1)-1 do # On calcule les composantes d(n-1,k) de f
    d[n-1, k]:=(c[n, 2*k]-c[n, 2*k+1])*(1/sqrt(2)): # sur la base des fonctions de Haar de W_{n-1}
  od:
  plot(sum(d[n-1, j]*phi(n-1, j, x), j=0..2^(n-1)-1), x=0..1, style=line, thickness=[2], title=
"Visualisation des composantes sur la base des ondelettes", titlefont=[HELVETICA, BOLD, 10]);
end:
```

B.3 Compression

```

Compression:=proc(n,tol,c,d)
  local coeffNuls,coeffNonNuls,j,k:
  coeffNuls:=0: coeffNonNuls:=0:
  DecompOndelettes(n,c,d):
  # on calcule toutes les composantes d[j,k]
  for j from n by -1 to 1 do
    for k from 0 to 2^(j-1)-1 do
      d[j-1,k]:=(c[j,2*k]-c[j,2*k+1])*(1/sqrt(2)):
      c[j-1,k]:=(c[j,2*k]+c[j,2*k+1])*(1/sqrt(2)):
    od:
  od:
  # on met à 0 les coefficients négligeables
  for j from 0 to n-1 do
    for k from 0 to 2^j-1 do
      if (abs(evalf(d[j,k])) < tol) then
        d[j,k]:=0:
        coeffNuls:=coeffNuls+1:
      else
        coeffNonNuls:=coeffNonNuls+1:
      fi:
    od:
  od:
  print(coeffNuls);
  print(coeffNonNuls);
end:

```

B.4 Reconstruction

```

Recomposition:=proc(n,c,d)
  global i,j,k,tab,graphe:
  for j from 1 to n-1 do
    for k from 0 to 2^j-1 do
      c[j,2*k]:=(c[j-1,k]+d[j-1,k])*(1/sqrt(2)):
      c[j,2*k+1]:=(c[j-1,k]-d[j-1,k])*(1/sqrt(2)):
    od:
  od:
  plot([sum(c[n,i]*phi(n,i,x),i=0..2^n-1),f(x)],x=0..1,
  color=[blue,red],style=[line,line],legend=["Signal f de départ",
  "Recomposition de f"],thickness=[2,1],title="Comparaison de la
  recomposition finale et du signal de départ",titlefont=[HELVETICA,BOLD,10]);
end:

```