

# 1 Bases de Haar

Dans ce chapitre nous allons construire et décrire une base Hilbertienne de  $L^2(\mathbb{R})$  à l'aide de fonctions constantes par morceaux à support compact. Avant de regarder plus en détail le cas de  $L^2(\mathbb{R})$  nous allons étudier le cas spécifique de  $L^2([0, 1])$ .

On commencera ce chapitre par une feuille de TD introductive qui présente les principales propriétés des ondelettes de Haar.

## 1.1 Bases de Haar sur les fonctions nulles en dehors de $[0, 1]$

1. Définition des fonctions  $\Phi_{j,k}$ .

**Définition 1. Définition des fonctions  $\Phi_{j,k}$**

On note  $\Phi$  la fonction définie sur  $\mathbb{R}$  indicatrice de l'intervalle  $[0, 1[$  :

$$\Phi(x) = \begin{cases} 1 & \text{si } x \in [0, 1[ \\ 0 & \text{sinon} \end{cases}$$

Pour tout  $j \leq 0$  et pour tout  $k \in \llbracket 0 : 2^{-j} - 1 \rrbracket$  on définit la fonction  $\Phi_{j,k}$  par la fonction définie sur  $\mathbb{R}$  par

$$\Phi_{j,k}(x) = 2^{-\frac{j}{2}} \Phi(2^{-j}x - k)$$

On note  $V_j = \text{Vect}((\Phi_{j,k})_{0 \leq k \leq 2^{-j}-1})$  l'espace vectoriel engendré par les fonctions  $\Phi_{j,k}$  pour l'ensemble des valeurs de  $k$

**Remarque :** Exprimer  $\Phi_{-1,0}$  et  $\Phi_{-1,1}$ .

**Remarque :** Pour tout  $j \leq 0$ , l'espace  $V_j$  est l'espace des fonctions constantes par morceaux sur des intervalles de la forme  $[k2^j, (k+1)2^j[$ .

On munit l'espace vectoriel des fonctions définies sur  $[0, 1[$  à valeurs réelles telles que

$$\int_0^1 |f(t)|^2 dt < +\infty$$

que l'on note  $L^2([0, 1])$  du produit scalaire suivant :

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

**Remarque :** Muni de ce produit scalaire l'espace  $L^2([0, 1])$  est un espace pré-hilbertien.

**Proposition 1. Propriétés des espaces  $V_j$**

Pour tout  $j \leq 0$

- (a) la dimension de  $V_j$  est  $2^{-j}$ ,
- (b)  $V_j \subset V_{j-1}$ ,
- (c) la famille  $(\Phi_{j,k})_{0 \leq k \leq 2^{-j}-1}$  est une base orthonormée de  $V_j$

De plus, pour tout  $\varepsilon > 0$  et pour toute fonction  $f$  continue sur  $[0, 1[$  (et donc pour toute fonction  $f \in L^2([0, 1[)$ ) il existe un  $j_0 \leq 0$  et une fonction  $g \in V_{j_0}$  telle que  $\|f - g\| \leq \varepsilon$ .  
On dit que l'union  $\bigcup_{j \leq 0} V_j$  des espaces  $V_j$  est dense dans  $L^2([0, 1[)$ .

**Remarque :** Comme pour tout  $j \leq 0$ ,  $V_j \subset V_{j-1}$ , pour tout  $k \in \llbracket 0 : 2^{-j} - 1 \rrbracket$ , la fonction  $\Phi_{j,k}$  appartient à  $V_{j-1}$  et s'exprime donc comme combinaison linéaire des  $(\Phi_{j,l})_{0 \leq l \leq 2^{j+1}-1}$ .

Exemples : Exprimer  $\Phi = \Phi_{0,0}$  en fonction de  $\Phi_{-1,0}$  et  $\Phi_{-1,1}$ , puis  $\Phi_{-1,0}$  en fonction des  $\Phi_{-2,k}$  pour  $k \in \llbracket 0 : 3 \rrbracket$ .

On peut remarquer le résultat suivant :

**Proposition 2.** Lien entre les fonctions  $\Phi_{j,k}$

Pour tout couple  $(j, k)$ ,

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

## 2. Les espaces $W_j$

Comme les espaces  $V_j$  sont inclus les uns dans les autres on peut définir le supplémentaire orthogonal de  $V_j$  dans  $V_{j-1}$ .

**Définition 2. Définition des espaces  $W_j$**

Pour tout  $j \leq 0$ , on note  $W_j$  le supplémentaire orthogonal de  $V_j$  dans  $V_{j-1}$ .

**Proposition 3.** Pour tout  $j \leq 0$ ,

- (a) la dimension de l'espace  $W_j$  est  $2^{-j}$ .
- (b) les espaces  $W_j$  sont les espaces des fonctions  $f$  constantes sur les intervalles de la forme  $[k2^{j-1}, (k+1)2^{j-1}[$  et telles que  $f(2k2^{j-1}) = -f((2k+1)2^{j-1})$ .
- (c) les espaces  $W_j$  sont engendrés par les fonctions  $(\Psi_{j,k})_{0 \leq k \leq 2^{-j-1}-1}$  définies par

$$\Psi_{j,k}(x) = \begin{cases} -2^{-\frac{j}{2}} & \text{si } x \in [k2^{j-1}, (k+1)2^{j-1}[ \\ 2^{-\frac{j}{2}} & \text{si } x \in [(k+1)2^{j-1}, (k+2)2^{j-1}[ \\ 0 & \text{sinon} \end{cases} \quad (3)$$

On appelle ces fonctions  $\Psi_{j,k}$ , les ondelettes de Haar et la fonction  $\Psi_{0,0}$  ondelette mère.

Toutes les ondelettes peuvent être construites par dilatation et translation de l'ondelette mère.

**Remarque :** Par définition de  $W_j$  on a

$$V_j = W_{j+1} \oplus V_{j+1} \quad (4)$$

En décomposant  $V_{j+1}$  on déduit la décomposition suivante de  $V_j$  :

$$V_j = W_{j+1} \oplus W_{j+2} \oplus V_{j+2} \quad (5)$$

et en itérant on peut écrire pour tout  $j < 0$

$$V_j = W_{j+1} \oplus W_{j+2} \oplus W_{j+3} \oplus \cdots \oplus W_0 \oplus V_0 = \bigoplus_{j < l \leq 0} W_l \oplus V_0 \quad (6)$$

Ainsi on peut construire différentes bases de  $V_j$  :

- Il y a bien entendu la famille  $(\Phi_{j,k})_{0 \leq k \leq 2^{-j}-1}$ ,
- En utilisant la première décomposition (??), on obtient une base par union des bases de  $V_{j+1}$  et  $W_{j+1}$  ainsi  $(\Phi_{j+1,k})_{0 \leq k \leq 2^{-j-1}-1} \cup (\Psi_{j+1,k})_{0 \leq k \leq 2^{-j-1}-1}$  est une base orthonormée de  $V_j$ .

— De même, en utilisant la décomposition de (??) on obtient une base orthonormée de  $V_j$  en faisant l'union des bases orthonormées des espaces  $(W_l)_{j < l \leq 0}$  et de la base de  $V_0$ .

Ainsi  $\bigcup_{j < l \leq 0} (\Psi_{l,k})_{0 \leq k \leq 2^{-l-1}} \cup \Phi_{0,0}$  est une base orthonormée de  $V_j$ .

Comme  $W_j \subset V_{j-1}$ , pour tout  $k$  on peut écrire  $\Psi_{j,k}$  comme une combinaison linéaire des  $\Phi_{j-1,l}$ . plus précisément on a

**Proposition 4.** *Lien entre les fonctions  $\Psi_{j,k}$  et  $\Phi_{j-1,l}$*

*Pour tout couple  $(j,k)$ ,*

$$\Psi_{j,k} = -\frac{\sqrt{2}}{2}\Phi_{j-1,2k} + \frac{\sqrt{2}}{2}\Phi_{j-1,2k+1}$$

## 1.2 Bases de Haar sur $\mathbb{R}$

3. Définitions de la fonction d'échelle de Haar  $\Phi$ , des espaces d'approximation  $V_j$  et des fonctions  $\Phi_{j,k}$ .

**Définition 3. Définition de la fonction  $\Phi$  de Haar et des espaces  $V_j$**

*On définit la fonction d'échelle de Haar comme la fonction définie sur  $\mathbb{R}$  par*

$$\Phi(x) = \begin{cases} 1 & \text{si } x \in [0, 1[ \\ 0 & \text{sinon} \end{cases}$$

*On définit pour tout  $k \in \mathbb{Z}$  la translatée  $\Phi_k$  de  $\Phi$  par*

$$\forall x \in \mathbb{R}, \quad \Phi_k(x) = \Phi_{x-k}.$$

*On note également  $V_0$  l'espace vectoriel engendré par toutes les translatées de  $\Phi$  :*

$$V_0 = \text{Vect}((\Phi_k)_{k \in \mathbb{Z}}) \quad (7)$$

*Cet espace est l'espace des fonctions à support compact, constantes sur des intervalles de la forme  $[k, k+1[$ .*

**Définition 4.** *Dans toute la suite on considère le produit scalaire réel défini par*

$$\langle f, g \rangle = \int_{\mathbb{R}} f(t)g(t)dt$$

*et on munit l'espace vectoriel des fonctions de  $L^2(\mathbb{R})$  de la norme associée à ce produit scalaire.*

**Remarque :** Pour ce produit scalaire, la famille  $(\Phi_k)_{k \in \mathbb{Z}}$  forme une base orthonormée de  $V_0$ .

**Remarque :** On définit de la même manière que dans le cas précédent les fonctions  $\Phi_{j,k}$  et les espaces  $V_j$ . La grande différence avec le cas précédent est que tous ces espaces  $V_j$  sont maintenant de dimension infinie.

**Remarque :** On définit également de la même manière la fonction  $\Psi$ , les fonctions  $\Psi_{j,k}$  et les espaces  $W_j$ .

Les espaces  $(V_j)_{j \in \mathbb{Z}}$  ainsi construits constituent ce que l'on appelle une approximation multi-résolution :

4. Définition d'une multi-résolution.

**Définition 5.** *Une suite  $(V_j)_{j \in \mathbb{Z}}$  de sous espaces fermés de  $L^2(\mathbb{R})$  est une approximation multi-résolution si elle vérifie les 6 propriétés suivantes :*

$$\forall (j,k) \in \mathbb{Z}^2, \quad f(t) \in V_j \Leftrightarrow f(t - 2^j k) \in V_j \quad (8)$$

$$\forall j \in \mathbb{Z}, \quad V_{j+1} \subset V_j \quad (9)$$

$$\forall j \in \mathbb{Z}, \quad f(t) \in V_j \Leftrightarrow f\left(\frac{t}{2}\right) \in V_{j+1} \quad (10)$$

$$\lim_{j \rightarrow +\infty} V_j = \bigcap_{j=-\infty}^{+\infty} V_j = \{0\} \quad (11)$$

$$\lim_{j \rightarrow -\infty} V_j = \text{Adhérence} \left( \bigcup_{j=-\infty}^{+\infty} V_j \right) = L^2(\mathbb{R}) \quad (12)$$

Il existe  $\Phi$  tel que  $\{\Phi(t-n)\}_{n \in \mathbb{N}}$  est une base hilbertienne de  $V_0$ .

Nous verrons qu'il est possible de construire des approximations multi-résolution à l'aide d'autres fonctions que les fonctions de Haar.

5. Filtres  $h_n$  et  $g_n$  et changements de bases sur  $V_j$ .

#### Exercice

(a) Justifier qu'il existe une unique suite  $(h_n)_{n \in \mathbb{Z}}$  telle que

$$\Phi_{1,0}(t) = \sum_{n=-\infty}^{+\infty} h_n \Phi(t-n). \quad (13)$$

En utilisant (??) déterminer explicitement la suite  $(h_n)_{n \in \mathbb{Z}}$ .

(b) En calculant la transformée de Fourier des deux membres de l'égalité ci dessus, établir la relation suivante :

$$\forall \omega \in \mathbb{R}, \quad \hat{\Phi}(2\omega) = \frac{1}{\sqrt{2}} \hat{h}(\omega) \hat{\Phi}(\omega) \quad (14)$$

(c) En déduire que pour tout  $P \in \mathbb{N}$

$$\hat{\Phi}(\omega) = \left( \prod_{p=1}^P \frac{\hat{h}(2^{-p}\omega)}{\sqrt{2}} \right) \hat{\Phi}(2^{-P}\omega). \quad (15)$$

On peut montrer en effectuant un passage à la limite que

$$\hat{\Phi}(\omega) = \left( \prod_{p=1}^{+\infty} \frac{\hat{h}(2^{-p}\omega)}{\sqrt{2}} \right) \hat{\Phi}(0). \quad (16)$$

(d) Justifier qu'il existe une suite  $(g_n)_{n \in \mathbb{N}}$  que l'on explicitera telle que

$$\Psi_{1,0} = \sum_{n=-\infty}^{+\infty} g_n \Phi_n \quad (17)$$

En déduire la relation suivante :

$$\forall \omega \in \mathbb{R}, \quad \hat{\Psi}(2\omega) = \frac{1}{\sqrt{2}} \hat{g}(\omega) \hat{\Phi}(\omega) \quad (18)$$

Nous verrons que l'on peut définir d'autres ondelettes associées à des suites  $(h_n)_{n \in \mathbb{N}}$  et  $(g_n)_{n \in \mathbb{N}}$ . Les relations précédentes permettent de définir les fonctions  $\Phi$  et  $\Psi$  uniquement en fonction des suites  $(h_n)_{n \in \mathbb{N}}$  et  $(g_n)_{n \in \mathbb{N}}$  qui sont elles mêmes liées par une relation simple.

6. Algorithme de transformée en ondelettes.

L'algorithme de transformée en ondelettes est un algorithme purement discret et exact. Il calcule exactement le produit scalaire d'une fonction avec des ondelettes qui nécessite pourtant en théorie un calcul d'intégrale. Cet algorithme repose sur le fait qu'on considère que les valeurs que l'on traite, ne sont pas des valeurs en des réels  $x_n$  d'une fonction  $f$  mais des produits scalaires entre une fonction  $f$  et des fonctions  $\Phi_{j,k}$  pour une certaine valeur de  $j$ . On utilise ensuite les relations qui existent entre les fonctions  $\Phi_{j,k}$  pour calculer les produits scalaires  $\langle f, \Phi_{l,k} \rangle$  et  $\langle f, \Psi_{l,k} \rangle$  les uns en fonction des autres.

Cette hypothèse qui peut sembler surprenante n'est en fait pas totalement déraisonnable dans la mesure où des produits scalaires  $\langle f, \Phi_{j,k} \rangle$  sont en fait des moyennes locales de la fonction  $f$ . Ainsi on interprète les nombres que l'on considère comme les coefficients de la projection de  $f$  sur un espace  $V_j$  dans la base des  $(\Phi_{j,k})_k$ . L'algorithme de transformée en ondelettes consiste ainsi en un changement de bases orthonormé de  $V_j$ .

Ainsi, suivant le choix de l'ondelette que l'on fait, la suite de nombres que l'on considère a un sens différent.

**Exercice**

- (a) En utilisant la relation (??) portant sur la fonction  $\Phi$  établir un lien entre  $\langle \Phi_{j,k}, f \rangle$ ,  $\langle \Phi_{j-1,2k}, f \rangle$  et  $\langle \Phi_{j-1,2k+1}, f \rangle$  où  $f \in L^2(\mathbb{R})$ .
- (b) Etablir une relation analogue entre  $\langle \Psi_{j,k}, f \rangle$ ,  $\langle \Phi_{j-1,2k}, f \rangle$  et  $\langle \Phi_{j-1,2k+1}, f \rangle$ .
- (c) La famille  $(\Phi_{j,k})_{k \in \mathbb{Z}}$  forme une base hilbertienne de  $V_j$ , de même la famille  $(\Phi_{j+1,k})_{k \in \mathbb{Z}}$  forme une base hilbertienne de  $V_{j+1}$  et la famille  $(\Psi_{j+1,k})_{k \in \mathbb{Z}}$  forme une base hilbertienne de  $W_{j+1}$ . Rappelons que

$$V_j = V_{j+1} \oplus W_{j+1}. \quad (19)$$

Déduire des questions précédentes une manière d'exprimer les coefficients de décomposition de  $f$  dans les bases  $(\Phi_{j+1,k})_{k \in \mathbb{Z}}$  et  $(\Psi_{j+1,k})_{k \in \mathbb{Z}}$  à partir de ceux dans  $(\Phi_{j,k})_{k \in \mathbb{Z}}$ .

- (d) Soit  $f \in V_j$ , déduire des questions précédentes un algorithme itératif de décomposition qui permet de calculer rapidement une décomposition de  $f$  dans  $\bigoplus_{i=j+1}^l W_i + V_l$  pour un  $l > j$  à partir des coefficients dans  $V_j$ .

On effectuera dans un premier temps le changement de base associé à la décomposition  $V_j = W_{j+1} \oplus V_{j+1}$  puis celui associé à la décomposition  $V_{j+1} = W_{j+2} \oplus V_{j+2}$ .

## 2 Bases hilbertiennes (orthogonales) d'ondelettes

Dans cette partie, nous allons introduire différentes bases orthogonales d'ondelettes, toutes construites selon la même architecture que la base de Haar vue précédemment. La base d'ondelettes est définie de manière univoque par la suite  $(h_n)_{n \in \mathbb{N}}$  qui peut être choisie autrement que pour la base de Haar. A partir de cette suite, on peut définir la fonction d'échelle  $\Phi$ , l'ondelette mère  $\Psi$  ainsi que les espaces  $V_j$  et  $W_j$ . On peut aussi définir la fonction d'échelle  $\Phi$  ou l'approximation multi-résolution  $(V_j)_{j \in \mathbb{Z}}$  et définir les autres éléments à partir de celle-ci. C'est ce dernier choix que nous faisons pour la présentation mais d'autres sont possibles.

1. Définition d'une fonction d'échelle  $\Phi$  associée à une multi-résolution.

**Définition 6.** Soit  $(V_j)_{j \in \mathbb{Z}}$  une approximation multi-résolution. On appelle fonction d'échelle, toute fonction  $\Phi$  telle que la famille des  $(\Phi_n)_{n \in \mathbb{Z}}$  forme une BON de  $V_0$ .

**Remarque :** Comme La fonction  $\Phi_{1,0} \in V_0$ , il existe une suite  $(h_n)_{n \in \mathbb{N}}$  telle que

$$\Phi_{1,0}(t) = \sum_{n=-\infty}^{+\infty} h_n \Phi(t-n). \quad (20)$$

Du fait que  $\Phi$  est une fonction d'échelle, la suite  $(h_n)_{n \in \mathbb{Z}}$  vérifie un certain nombre de propriétés. Le théorème suivant est dû à Mallat et Meyer.

**Théorème 1.** Soit  $\Phi \in L^2(\mathbb{R})$  une fonction d'échelle, alors la transformée de Fourier de la suite  $(h_n)_{n \in \mathbb{N}}$  associée vérifie la relation suivante :

$$\forall \omega \in \mathbb{R}, \quad |\hat{h}(\omega)|^2 + |\hat{h}(\omega + \pi)|^2 = 2 \quad (21)$$

et

$$\hat{h}(0) = \sqrt{2}. \quad (22)$$

Inversement si  $\hat{h}$  est  $2\pi$ -périodique et continument différentiable au voisinage de 0 et qu'elle vérifie les deux relations ci dessous et si de plus

$$\inf_{\omega \in [-\pi/2, \pi/2]} |\hat{h}(\omega)| > 0 \quad (23)$$

alors la fonction  $\Phi$  définie par

$$\Phi(\omega) = \left( \prod_{p=1}^{+\infty} \frac{\hat{h}(2^{-p}\omega)}{\sqrt{2}} \right). \quad (24)$$

est la transformée de Fourier d'une fonction d'échelle de  $L^2(\mathbb{R})$

Ce théorème indique qu'il est possible de construire une grande variété de BON d'ondelettes, dont Haar n'est qu'un exemple.

**Exercice** Montrer que la suite  $(h_n)_{n \in \mathbb{Z}}$  associée à la base de Haar vérifie bien les hypothèses du théorème.

**Exercice**

(a) Montrer que la fonction  $\hat{h}$  définie par

$$\forall \omega \in [-\pi, \pi], \quad \hat{h}(\omega) = \sqrt{2} \chi_{[-\pi/2, \pi/2]}(\omega) \quad (25)$$

vérifie bien les hypothèses du théorème.

(b) En déduire que la fonction d'échelle associée vérifie

$$\forall \omega \in \mathbb{R}, \quad \hat{\Phi}(\omega) = \chi_{[-\pi, \pi]}(\omega) \quad (26)$$

En déduire une expression de  $\Phi$ .

**Exercice**

On note  $(g_n)_{n \in \mathbb{Z}}$  la suite de  $\ell^2(\mathbb{Z})$  définie par

$$g_n = (-1)^{1-n} h_{1-n}.$$

(a) Exprimer  $\hat{g}$  en fonction de  $\hat{h}$ .

(b) Montrer que pour tout  $\omega \in \mathbb{R}$

$$|\hat{h}(\omega)|^2 + |\hat{g}(\omega)|^2 = 2 \text{ et que} \quad (27)$$

On définit la fonction  $\Psi$  par

$$\hat{\Psi}(2\omega) = \frac{1}{\sqrt{2}} \hat{g}(\omega) \hat{\Phi}(\omega)$$

La fonction  $\Psi$  est appelée ondelette.

(c) Montrer que

$$\Psi(x) = \sum_n g_n \sqrt{2} \Phi(2x - n) = \sum_n g_n \Phi_{-1,n}(x).$$

(d) Calculer  $\hat{g}(0)$  puis  $\int_{x \in \mathbb{R}} \Psi(x) dx$ .

(e) Montrer que pour tout  $k \in \mathbb{Z}$

$$\langle \Psi(\cdot), \Psi(\cdot - k) \rangle = \sum_n h_n h_{n-2k}$$

(f) Montrer que pour tout  $k \in \mathbb{Z}$

$$\langle \Phi(\cdot), \Phi(\cdot - k) \rangle = \sum_n h_n h_{n-2k}$$

En déduire que les  $(\Psi(\cdot - k))_{k \in \mathbb{Z}}$  forment une base orthonormée de l'espace qu'elles engendrent.

2. Décomposition de  $V_j$  en somme d'espaces orthogonaux  $W_l$ .

Comme pour les ondelettes de Haar, on définit les fonctions  $\Phi_{j,k}$  :

$$\Phi_{j,k}(x) = 2^{-\frac{j}{2}} \Phi(2^{-j}x - k) \quad (28)$$

et les ondelettes  $\Psi_{j,k}$

$$\Psi_{j,k}(x) = 2^{-\frac{j}{2}} \Psi(2^{-j}x - k) \quad (29)$$

Et on définit de même les espaces  $V_j$  et  $W_j$  par

$$\forall j \in \mathbb{Z}, \quad V_j = \text{Vect}(\Phi_{j,k})_{k \in \mathbb{Z}} \text{ et } W_j = \text{Vect}(\Psi_{j,k})_{k \in \mathbb{Z}}. \quad (30)$$

Ces fonctions forment ainsi des bases hilbertiennes des fonctions qu'elles engendrent. On a ainsi

$$V_j = V_{j+1} \oplus W_{j+1}. \quad (31)$$

et pour tout  $l \geq j$

$$V_j = \bigoplus_{k=l}^{j+1} W_k + V_l \quad (32)$$

Comme  $(V_j)_{j \in \mathbb{Z}}$  est une approximation multi-résolution,  $\bigcup_{j \in \mathbb{Z}} V_j = L^2(\mathbb{R})$  et ainsi pour toute fonction  $f \in L^2(\mathbb{R})$  on a la formule de reconstruction suivante :

$$f = \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} \langle f, \Psi_{j,k} \rangle \Psi_{j,k} \quad (33)$$

3. Notion de transformée en ondelettes orthogonale.

Comme nous avons utilisé la relation (??) pour établir un algorithme de transformée rapide en ondelettes on peut utiliser les relations

$$\Phi(t) = \sum_{n \in \mathbb{Z}} h_n \Phi_{-1,n}(t) = \sum_{n \in \mathbb{Z}} h_n 2^{\frac{1}{2}} \Phi(2t - n) \quad \text{et} \quad \Psi(t) = \sum_{n \in \mathbb{Z}} g_n \Phi_{-1,n}(t) = \sum_{n \in \mathbb{Z}} g_n 2^{\frac{1}{2}} \Phi(2t - n) \quad (34)$$

pour établir un algorithme de transformée en ondelettes.

En effet en utilisant la définition des fonctions  $\Phi_{j,k}$  et  $\Psi_{j,k}$  données plus haut on déduit des égalités précédentes que pour tout couple  $(j, k) \in \mathbb{Z}^2$

$$\Phi_{j,k} = \sum_{n \in \mathbb{Z}} h_n \Phi_{j-1,n+2k} \quad (35)$$

et de même

$$\Psi_{j,k} = \sum_{n \in \mathbb{Z}} g_n \Phi_{j-1,n+2k} \quad (36)$$

Comme pour Haar, on exploite la décomposition  $V_{j-1} = V_j \oplus W_j$  de manière récursive et le fait que pour toute fonction  $f$  on a d'une part

$$\langle f, \Phi_{j,k} \rangle = \sum_{n \in \mathbb{Z}} h_n \langle f, \Phi_{j-1,n+2k} \rangle \quad (37)$$

et d'autre part

$$\langle f, \Psi_{j,k} \rangle = \sum_{n \in \mathbb{Z}} g_n \langle f, \Phi_{j-1,n+2k} \rangle \quad (38)$$

Ce qui permet de calculer les coefficients de décomposition d'une fonction  $f$  sur les espaces  $V_j$  et  $W_j$  de manière très rapide à partir de ceux sur  $V_{j-1}$ . Dans la pratique on calcule chacun des jeux de coefficients par une convolution et un sous-échantillonnage d'un facteur 2. On itère ensuite le procédé autant de fois qu'on le souhaite sur les espaces  $V_l$  résiduels en conservant inchangés les coefficients sur les espaces  $W_l$ .

C'est l'algorithme de transformée rapide en ondelettes

4. Moments nuls, régularité et décroissance des coefficients d'ondelettes. On commencera ce chapitre par une feuille de TD consacrées aux moments nuls des bases d'ondelettes orthogonales.

**Définition 7.** On dit qu'une ondelette  $\Psi$  admet  $p$  moments nuls si pour tout  $k < p$  on a

$$\int_{t \in \mathbb{R}} \Psi(t) t^k dt = 0. \quad (39)$$

**Questions :**

- (a) Montrer que l'ondelette de Haar admet 1 moment nul.
- (b) Justifier que pour tout choix de  $h$  vérifiant les hypothèses de Mallat, on a nécessairement  $\hat{g}(0) = 0$ . En déduire que  $\int_{t \in \mathbb{R}} \Psi(t) dt = 0$  et donc que toute ondelette possède au moins un moment nul.

**Remarque :** Si l'ondelette  $\Psi$  admet  $p$  moments nuls alors l'ondelette est orthogonale à tous les polynômes de degré strictement inférieur à  $p$ .

**Définition 8.** On dit qu'une fonction  $f$  est à support compact s'il existe  $K > 0$  tel que  $f$  est nulle en dehors de l'intervalle  $[-K, K]$ .

On dit qu'une suite  $(h_n)_{n \in \mathbb{Z}}$  est à support compact s'il existe  $n_0 \in \mathbb{N}$  tel que  $|n| \geq n_0 \implies h_n = 0$ . Le support de  $(h_n)_{n \in \mathbb{Z}}$  est ainsi le plus petit intervalle  $[N_1, N_2]$  tel que  $h$  est nul en dehors de cet intervalle.



**Proposition 5.** Si  $(h_n)_{n \in \mathbb{Z}}$  est à support compact, alors la fonction d'échelle associée est à support compact et de même support que  $(h_n)_{n \in \mathbb{Z}}$ . L'ondelette  $\Psi$  associée est également à support compact.

5. Approximation en ondelettes.

**Proposition 6.** Si  $f \in L^2([0, 1])$  est une fonction,  $C^m$  par morceaux et admettant un nombre fini de discontinuités alors si  $\Psi$  est une ondelette à support compact ayant  $p > m$  moments nuls alors il existe  $C > 0$  tel que l'erreur d'approximation non-linéaire  $\epsilon_N^l(f)$  à  $N$  termes de  $f$  dans la base d'ondelettes  $(\Psi_{j,k})_{(j,k) \in \mathbb{Z}^2}$  vérifie

$$\epsilon_N^n(f) \leq \frac{C}{n^{2m}} \quad (40)$$

**Idée de la démonstration :**

- (a) Pour démontrer cette décroissance, on analyse la décroissance des coefficients d'ondelettes réordonnés par ordre décroissant de valeur absolue.
- (b) Les coefficients d'ondelettes peuvent se classer en deux catégories :
  - i. Ceux qui sont associés à des ondelettes dont le support intersecte une singularité. Comme l'ondelette est à support compact, il existe un entier  $K$  tel qu'à chaque échelle  $j$  le nombre de tels coefficients est borné par  $k$ . Comme  $f$  est bornée on a

$$|\langle f, \Psi_{j,k} \rangle| \leq \|f\|_\infty \|\Psi_{j,k}\|_1 = 2^{\frac{j}{2}} \|f\|_\infty \|\Psi\|_1 \quad (41)$$

Ainsi ces coefficients décroissent rapidement grâce au fait qu'ils sont très peu nombreux par échelle et qu'ils décroissent en fonction de l'échelle.

- ii. Ceux qui sont associés à des ondelettes dont le support n'intersecte pas une discontinuité. Pour ceux là, on utilise le fait qu'il existe un réel  $C$  tel que sur le support  $[a, b]$  de l'ondelette, la fonction  $f$  peut être approchée par un polynôme  $P$  de degré  $n$  tel que

$$\forall x \in I, \quad |f(x) - P(x)| \leq C(x-a)^n. \quad (42)$$

On peut prendre par exemple le polynôme de Taylor au point  $a$ .

On utilise ensuite le fait que

$$|\langle f, \Psi_{j,k} \rangle| \leq |\langle f - P, \Psi_{j,k} \rangle| + |\langle P, \Psi_{j,k} \rangle| = |\langle f - P, \Psi_{j,k} \rangle| \quad (43)$$

car  $\Psi$  et donc  $\Psi_{j,k}$  est orthogonale aux polynômes dont le degré est inférieur ou égal à  $p-1 \geq n$ .

On observe ensuite que la taille du support d'une ondelette  $\Psi_{j,k}$  est borné par  $K2^j$  où  $K$  est la taille du support de  $\Psi$ .

On a ainsi

$$|\langle f - P, \Psi_{j,k} \rangle| \leq \|f - P\|_\infty \|\Psi_{j,k}\|_1 \leq C(K2^j)^n 2^{\frac{j}{2}} \|\Psi\|_1 = CK^n \|\Psi\|_1 2^{\frac{(2n+1)j}{2}} \quad (44)$$

On a ainsi

$$|\langle f, \Psi_{j,k} \rangle|^2 \leq C2^{(2n+1)j} \quad (45)$$

On conclut en observant qu'il y a au plus  $2^{-j}$  coefficients par échelle de ce type (c'est le nombre total de coefficients d'ondelettes à cette échelle car la fonction  $f$  est définie sur l'intervalle  $[0, 1]$ ).

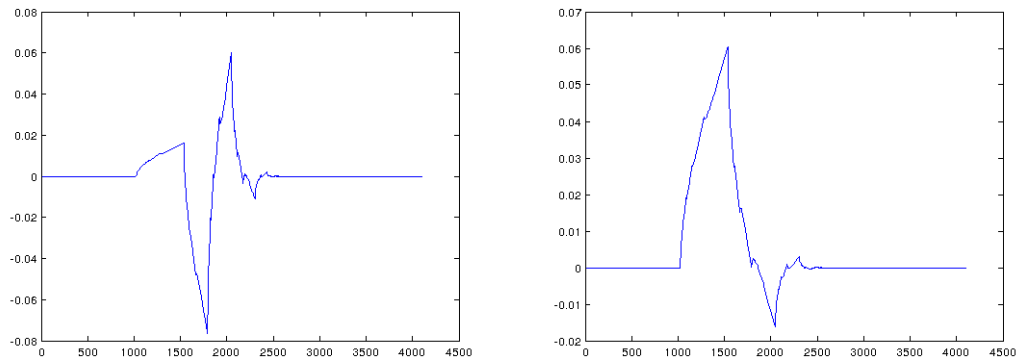


FIGURE 1 – Ondelette  $\Psi$  à gauche et fonction d'échelle  $\Psi$  à droite dites de *Daubechies* de paramètre 4 avec 2 moments nuls.

- (c) Une fois qu'on a contrôlé la décroissance des deux types de coefficients on remarque que ce sont les seconds qui sont dominants, et on peut conclure.

On peut observer que cette décroissance est la même que ce qu'on peut avoir pour des fonctions  $C^n$  et  $2\pi$ -périodique dans une base de Fourier.

La localisation des ondelettes permet d'assurer la même décroissance des coefficients avec ou en l'absence de discontinuité, ce qui est un avantage décisif par rapport à la base des exponentielles complexes.

**Remarque :** On peut remarquer qu'en terme d'approximation linéaire, les ondelettes n'apportent aucun avantage par rapport à d'autres bases comme Fourier par exemple.

## 6. Exemples des ondelettes de Daubechies et de Coifman.

Le filtre  $(h_n)_{n \in \mathbb{Z}}$  associé à l'ondelette de Daubechies de paramètre 4 est défini par

$$h_0 = a, \quad h_1 = a(2 - b), \quad h_2 = a(1 - 2b), \quad h_3 = -ab \quad \text{avec} \quad a = \frac{\sqrt{2}}{4(\sqrt{3} - 1)} \text{ et } b = 2 - \sqrt{3} \quad (46)$$

On peut vérifier que cette ondelette a bien deux moments nuls, en effet  $\hat{h}(\pi) = 0$  et  $\hat{h}'(\pi) = 0$ . C'est l'ondelette de support minimal qui a cette propriété.

Cette propriété de support implique qu'en tout point  $x \in \mathbb{R}$  et pour une échelle  $j$ , il existe au plus 3 ondelettes dont le support intersecte  $x$ , comme l'indique la figure 2. C'est la seule ondelette ayant au moins deux moments nuls à vérifier cette propriété.

On peut aussi mentionner les Ondelettes de Coifman dont le support est légèrement plus grand que celui des ondelettes de Daubechies.

## 7. On termine le chapitre par une feuille de TD matlab sur l'approximation linéaire et non-linéaire en ondelettes.

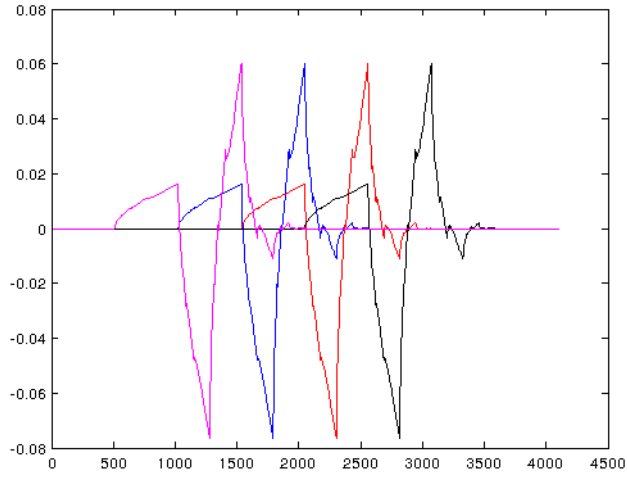


FIGURE 2 – Quatre Ondelettes  $\Psi$  de *Daubechies* 4 de même échelle  $j$  pour des valeurs de  $k$  consécutives.

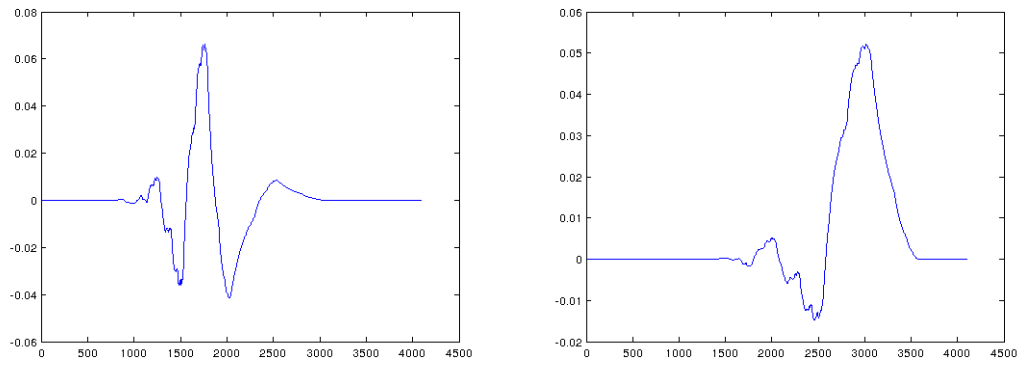


FIGURE 3 – Ondelette  $\Psi$  à gauche et fonction d'échelle  $\Psi$  à droite dites de *Coifman* ou *Coiflet* de paramètre 1 avec 2 moments nuls.

### 3 Ondelettes 2D

#### 1. Définitions :

Une base orthonormée d'ondelettes séparables de  $L^2(\mathbb{R}^2)$  est construite à partir de produits séparables de dilatées et de translatées d'une fonction d'échelle  $\Phi$  et d'une ondelette  $\Psi$ . La fonction d'échelle est associée à une approximation multirésolution  $\{V_j\}_{j \in \mathbb{Z}}$  de  $L^2(\mathbb{R})$ .

Ainsi  $V_j^2$  l'espace d'approximation d'échelle  $j$  est l'espace vectoriel engendré par les fonctions de la forme

$$\Phi_{j,k_1,k_2}(x,y) = \Phi_{j,k_1}(x)\Phi_{j,k_2}(y) \quad (47)$$

Comme dans le cas 1D, on a les inclusions  $V_j^2 \subset V_{j-1}^2$  et on peut définir l'espace  $W_j$  comme le supplémentaire orthogonal de  $V_j^2$  dans  $V_{j-1}^2$  :

$$V_{j-1}^2 = V_j^2 \oplus W_j^2 \quad (48)$$

Les espaces  $W_j$  peuvent chacun s'écrire sous la forme d'une somme orthogonale de trois espaces :

$$W_j = W_j^v \oplus W_j^h \oplus W_j^d \quad (49)$$

Chacun de ces espaces correspond à un espace de détail d'une direction particulière.

(a)  $W_j^v$  est espace des détails verticaux, il est engendré par les fonctions

$$\Psi_{j,k_1,k_2}^v(x,y) = \Psi_{j,k_1}(x)\Phi_{j,k_2}(y)$$

pour  $(k_1, k_2) \in \mathbb{Z}^2$ .

(b)  $W_j^h$  est espace des détails horizontaux, il est engendré par les fonctions

$$\Psi_{j,k_1,k_2}^h(x,y) = \Phi_{j,k_1}(x)\Psi_{j,k_2}(y)$$

pour  $(k_1, k_2) \in \mathbb{Z}^2$ .

(c)  $W_j^d$  est espace des détails diagonaux, il est engendré par les fonctions

$$\Psi_{j,k_1,k_2}^d(x,y) = \Psi_{j,k_1}(x)\Psi_{j,k_2}(y)$$

pour  $(k_1, k_2) \in \mathbb{Z}^2$ .

Ainsi les bases d'ondelettes 2D sont des produits dits *tensoriels* de bases d'ondelettes 1D.

**Remarque :** Contrairement au cas 1D on n'a pas de décroissance de l'erreur d'approximation non linéaire comparable pour les fonctions régulières et les fonctions régulières par morceaux.

Cependant on peut montrer que la décroissance de cette erreur d'approximation non linéaire est meilleure que celle qu'on obtient en bases de Fourier 2D. Les coefficients de large amplitudes sont concentrés le long des courbes de singularité.

Dans la pratique, les ondelettes sont très efficaces pour coder les images et sont utilisées pour la compression, le débruitage, la reconstruction et bien d'autres applications.

#### 2. Algorithme de Transformée en ondelettes 2D.

L'algorithme de transformée en ondelettes 2D consiste à alterner les étapes de la transformée en ondelettes 1D sur les lignes et les colonnes. (Voir feuille de TD associée).

#### 3. Applications.



FIGURE 4 – Lenna à gauche et sa transformée en ondelettes de Daubechies 4 à droite. Essayez de repérer où se situe les différents espaces de détails.

## 4 Algorithmes itératifs pour la résolution de systèmes linéaires

Dans cette partie, nous allons considérer des méthodes itératives pour déterminer une solution exacte ou approchée d'un problème linéaire de la forme

$$y = Ax \quad (50)$$

où  $A$  est une application linéaire supposée connue de l'espace  $\mathbb{R}^N$  dans l'espace  $\mathbb{R}^n$  qu'on pourra identifier à une matrice ayant  $n$  lignes et  $N$  colonnes,  $y \in \mathbb{R}^n$  est un vecteur connu qu'on appellera aussi *observations* et  $x \in \mathbb{R}^N$  un vecteur inconnu que l'on cherche à déterminer. On s'intéressera particulièrement aux situations où  $N > n$ .

Un grand nombre de problèmes de traitement des signaux ou des images peuvent s'exprimer ainsi. L'opérateur  $A$  peut être

- Un opérateur de convolution, dans ce cas  $y$  est un signal ou une image flou,
- Un opérateur de masquage,  $y$  est alors une image avec des pixels manquants,
- Une transformée de Radon ou un opérateur associé à une IRM,
- Un opérateur d'échantillonnage compressé (Compressed Sensing),
- Une union de deux bases orthogonales, pour faire de la séparation de sources,
- Des mesures acoustiques ou issues de radars, lidar, etc ...

On ne traitera pas ici le cas de mesures bruitées mais dans la plupart des problèmes pratiques on doit même supposer que

$$y = Ax + b \quad (51)$$

où  $b$  est une perturbation sur les mesures, déterministe ou aléatoire. On supposera donc dans un premier temps que  $b = 0$ , c'est à dire que les observations ne sont pas corrompues.

Comme  $N > n$ , il peut exister une infinité de solutions au problème (??). Nous allons proposer ici plusieurs méthodes itératives qui permettent d'en déterminer une.

## 4.1 Le Matching Pursuit (MP) et ses variantes, OMP et OLS

### 4.1.1 Le Matching Pursuit (MP)

Le Matching Pursuit est un algorithme itératif qui a pour but de rechercher une solution exacte ou approchée du problème (??) relativement parcimonieuse, c'est à dire qui admet un faible nombre de composantes non nulles. Nous verrons que dans un grand nombre de problèmes, les *bonnes solutions* au problème indéterminé qu'on se pose sont parcimonieuses. Le Matching Pursuit est un moyen de chercher de telles solutions. Cet algorithme a été proposé par Mallat et Zhang en 1993. Le même Stéphane Mallat que les ondelettes et des Data Sciences...

L'idée de l'algorithme est de partir du vecteur nul  $x_0 = 0$  et d'ajouter séquentiellement des composantes au vecteur  $x$  de manière à ce que  $Ax$  se rapproche de  $y$ . Pour cela on considère la version matricielle de l'opérateur  $A$ , c'est à dire qu'on identifie l'opérateur à une matrice et on considère la matrice comme une collection de  $N$  colonnes  $(a_i)_{i \leq N}$  qui sont toutes des éléments de  $R^n$ .

Résoudre (??) c'est chercher à écrire  $y$  comme une combinaison linéaire des  $(a_i)_{i \leq N}$ , les coefficients  $x(i)$  de  $x$  étant les coefficients de la combinaison linéaire :

$$y = Ax \Leftrightarrow y = \sum_{i=1}^N x(i) a_i \quad (52)$$

On appellera *atomes* les colonnes  $(a_i)_{i \leq N}$  de  $A$ .

La philosophie de Matching Pursuit (MP) est de chercher de manière itérative les atomes les plus pertinents dans cette décomposition. A chaque étape  $k$  de l'algorithme on cherche ainsi l'indice  $i_k$  qui va enrichir la représentation de manière la plus efficace. On parle d'algorithme glouton ou greedy. A chaque étape, on dispose d'un ensemble d'atomes (éventuellement vide au début de la première itération) et on se pose la question de savoir si on devait choisir un unique atome supplémentaire, lequel on choisirait et avec quel coefficient ?

L'algorithme de MP consiste à sélectionner la composante qui est la plus corrélée avec le résidu, celui qui ressemble le plus au résidu. Le résidu à l'étape  $k$  étant défini par  $r_k = y - Ax_{k-1}$ . On définit la valeur du coefficient de manière à ce que le nouveau résidu  $r_{k+1}$  soit orthogonal à l'atome sélectionné  $a_{i_k}$ . L'algorithme consiste à projeter le résidu sur l'atome qui lui ressemble le plus.

Comme  $x_0$  est le vecteur nul, le résidu utilisé pour effectuer la première étape est  $r_1 = y$ . Il n'existe pas un unique critère d'arrêt de l'algorithme. La plupart du temps on se fixe soit un nombre d'étapes  $K$  ou une norme sur le résidu  $y - Ax$ .

L'algorithme peut ainsi se décrire de la manière suivante :

1.  $x_0 = 0, r_1 = y$ .
2. Pour  $k = 1$  à  $K$ 
  - (a)  $x_k = x_{k-1}$
  - (b)  $i_k = \arg \max_{i \leq N} |\langle a_i, r_k \rangle|$ ,
  - (c)  $x_k(i_k) = x_{k-1}(i_k) + \frac{\langle r_k, a_{i_k} \rangle}{\|a_{i_k}\|^2}$
  - (d)  $r_{k+1} = y - Ax_k$

Question : Montrer qu'on a bien :  $r_{k+1} \perp a_{i_k}$  et justifier le choix de la mise à jour de  $x_k(i_k)$ .

En déduire que

$$\|r_{k+1}\|^2 = \|r_k\|^2 - |\langle r_k, a_{i_k} \rangle|^2 \quad (53)$$

et ainsi que la norme du résidu décroît.

Faire un DESSIN au tableau pour illustrer l'algorithme sur une matrice  $2 \times 2$ .

On peut montrer que si  $y$  appartient à l'image de  $A$  alors la norme de ce résidu tend vers 0 quand le nombre d'étapes  $K$  tend vers  $+\infty$ .

Si les atomes sont suffisamment décorrelés les uns des autres et qu'il existe une décomposition réellement parcimonieuse de  $y$  dans cette famille alors le MP est capable de la retrouver. Plus précisément, on peut définir la cohérence d'une famille d'atomes normalisés :

**Définition 9.** Soit  $(a_i)_{i \leq N}$  une famille de vecteurs tels que pour tout  $i \leq N$ ,  $\|a_i\|_2 = 1$ , on appelle cohérence de la famille le maximum de la corrélation entre deux vecteurs distincts :

$$C = \max_{i \neq j} |\langle a_i, a_j \rangle| \quad (54)$$

On définit la cohérence  $C(A)$  d'une matrice  $A$  comme la cohérence de la famille formée des colonnes de  $A$ . On peut montrer que si les atomes sont des vecteurs de dimension  $n$  et que  $A$  est formée de plus de  $N = n + 1$  vecteurs alors la cohérence de  $A$  est au moins de l'ordre de  $\frac{1}{\sqrt{n}}$ . Or comme on va le voir tout de suite, plus cette cohérence est faible, plus le MP est performant pour retrouver les vecteurs parcimonieux. On a alors le théorème suivant :

**Théorème 2.** Soit  $x$  un vecteur de  $\mathbb{R}^N$  et  $A$  une matrice ayant  $n$  lignes et  $N$  colonnes normalisées et  $y = Ax$ . Si le nombre de composantes non nulles  $k$  de  $x$  vérifie

$$k < \frac{1}{2} \left(1 + \frac{1}{C}\right) \quad (55)$$

alors le MP ne sélectionnera à chaque étape que des atomes actifs associés à  $x$  c'est-à-dire associés à des indices des coefficients non nuls de  $x$ .

Dans la pratique, cette condition est très restrictive et le MP permet de trouver des solutions approchées du problème (??) même si cette condition n'est pas vérifiée.

On peut noter toutefois que l'algorithme converge rarement en temps fini et qu'il peut sélectionner au fil des étapes plusieurs fois la même composante comme on peut s'en rendre compte sur un simple système  $2 \times 2$ .

On peut utiliser simplement le MP pour résoudre les trois problèmes suivants :

- La déconvolution,
- L'inpainting,
- La séparation de sources.

Il est possible de modifier le MP pour qu'il converge plus vite vers une solution parcimonieuse mais ces accélérations ont un coût en terme calculatoire.

#### 4.1.2 Orthogonal Matching Pursuit (OMP)

L'OMP est une variante du MP qui sélectionne les atomes un par un en utilisant le même critère de maximum de corrélation mais qui recalcule à chaque étape toutes les composantes de  $x$  de manière à ce que le résidu soit orthogonal non pas au dernier atome sélectionné mais à la totalité des atomes sélectionnés depuis la première étape. Ainsi à l'étape  $k$ , le vecteur  $x_k$  est choisi tel que  $Ax_k$  est la projection des observations  $y$  sur l'espace vectoriel engendré par les atomes  $(a_i)_{i \in I_k}$  sélectionnés.

On rappelle que si  $(a_i)_{i \in I}$  est une famille libre de vecteurs de  $\mathbb{R}^n$  formant une matrice  $A_I$  et si  $y$  est vecteur de  $\mathbb{R}^n$  alors la projection orthogonale de  $y$  sur l'espace vectoriel engendré par les  $(a_i)_{i \in I}$  est obtenue par la formule suivante  $P(y) = A_I(A_I^T A_I)^{-1} A_I^T y$  et que les coefficients de décomposition de cette projection  $P(y)$  dans la base des  $(a_i)_{i \in I}$  est donnée par  $x = (A_I^T A_I)^{-1} A_I^T y$ .

L'algorithme OMP peut ainsi se décrire de la manière suivante :

1.  $I_0 = \emptyset$ ,  $x_0 = 0$ ,  $r_1 = y$ .

2. Pour  $k = 1$  à  $K$

- (a)  $i_k = \arg \max_{i \leq N} |\langle a_i, r_k \rangle|$ ,
- (b)  $I_k = I_{k-1} \cup i_k$
- (c)  $x_k = 0$ ,  $x_k(I_k) = (A_{I_k}^t A_{I_k})^{-1} A_{I_k}^t y$
- (d)  $r_{k+1} = y - Ax_k$

La différence majeure entre MP et OMP est que *OMP* nécessite une inversion de matrice, la matrice  $A_{I_k}^t A_{I_k}$  qui est une matrice carré  $k \times k$ . Cette inversion peut s'avérer coûteuse quand  $k$  devient grand. En contrepartie, OMP est nettement plus rapide et ne sélectionne jamais deux fois le même atome. On a le théorème suivant sur la reconstruction de vecteurs parcimonieux :

**Théorème 3.** Soit  $x$  un vecteur de  $\mathbb{R}^N$  et  $A$  une matrice ayant  $n$  lignes et  $N$  colonnes normalisées et  $y = Ax$ . Si le nombre de composantes non nulles  $k$  de  $x$  vérifie

$$k < \frac{1}{2} \left(1 + \frac{1}{C}\right) \quad (56)$$

alors le OMP retrouve exactement  $x$  en exactement  $k$  étapes.

Le choix de l'un ou l'autre algorithme dépend de la complexité de calculer ces inversions de matrices.

#### 4.1.3 Orthogonal Least Square algorithm (OLS)

L'OLS est une autre variante du Matching Pursuit, plus rapide au sens où il nécessite souvent moins d'étapes que le MP ou l'OMP mais chaque étape est plus coûteuse que pour chacun des précédents algorithmes. En effet l'idée de l'algorithme est de sélectionner à chaque étape l'atome qui minimise l'erreur de projection orthogonale, c'est à dire la distance entre  $y$  et l'espace vectoriel engendré par les atomes actifs. Cette fois ci, il n'est pas possible de sélectionner le nouvel atome par un simple calcul de produit scalaire, il faut calculer pour chaque nouvel indice  $i$  possible ( $N-k+1$ ) à l'étape  $k$ , la projection associée sur l'espace engendré par les  $(a_i)_{i \in I}$  avec  $I = I_{k-1} \cup i$  (en utilisant la formule donnée dans la section précédente) et sélectionner l'atome qui minimise l'erreur quadratique associé. L'algorithme se parallélise facilement mais est souvent trop coûteux. On lui préférera donc souvent OMP ou plus simplement MP. Il existe aussi des théorèmes de convergence d'OLS sous contraintes de parcimonie. Par exemple le théorème précédent est également valable pour OLS.

## 4.2 Méthodes des projections itérées.

La méthodes des projections itérées est une méthode de résolution de système linéaire qui consiste à considérer la résolution d'un système linéaire comme la recherche d'un point de l'intersection d'hyperplans affines. En effet, on peut associer chaque ligne du système à un hyperplan de  $\mathbb{R}^N$  où  $N$  est le nombre de colonnes de  $A$ .

**Exemple :**

$$y = Ax \quad \text{avec} \quad A = \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} \quad \text{et} \quad y = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad (57)$$

L'algorithme consiste alors à partir d'un  $x_0$  quelconque, à itérer les projections sur chacun de ces hyperplans affines.

**Question** Si on note  $a_i$  cette fois ci la  $i$ ème ligne de la matrice  $A$ , déterminer la formule de mise à jour de  $x$  par projection sur l'hyperplan affine associé. On peut remarquer qu'une telle projection modifie a priori toutes les composantes de  $x$  à la fois. Ce n'est donc pas un algorithme qui permet d'obtenir une solution parcimonieuse. Il existe différentes manières de *tourner* entre les hyperplans, on peut les sélectionner



une par un dans un ordre préalablement fixé ou faire une sélection aléatoire.

Comme dans les algorithmes précédents, il n'y a pas de règle universelle pour arrêter l'algorithme. On peut se fixer une précision, ou un nombre de fois on traite chaque ligne par exemple.

Même si chaque étape fait intervenir une projection comme c'est le cas dans les méthodes énoncées précédemment, elle en diffère pour plusieurs raisons et ne renvoie donc a priori pas le même résultat quand le système admet une infinité de solutions.

- L'algorithme ne cherchant pas à trouver de solution parcimonieuse, on peut le faire débiter en n'importe quel point de l'espace  $\mathbb{R}^N$ .
- Les projections considérées se font dans l'espace de départ, on ne cherche pas à minimiser la norme d'un résidu dans l'espace d'arrivée.
- On ne sélectionne pas à chaque étape, un meilleur espace sur lequel projeter. On projette  $x \in \mathbb{R}^N$  sur un hyperplan de  $\mathbb{R}^N$  préalablement choisi, et pas  $r \in \mathbb{R}^n$  sur un espace de dimension 1 choisi de manière optimale.

**Exemple sur**

$$y = Ax \quad \text{avec} \quad A = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 3 & -1 \end{pmatrix} \quad \text{et} \quad y = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad (58)$$

Si les lignes sont orthogonales par blocks, on peut effectuer les projections sur des lignes orthogonales d'un seul coup. C'est cette méthode qui est utilisée pour effectuer l'inversion de l'opérateur de tomographie.

**Exemple sur un tableau  $3 \times 3$**