

Traitement numérique et compression de données
Master professionnel de Mathématiques

Alain Yger

22 septembre 2008

Table des matières

1	L’outil Fourier	1
1.1	La transformée de Fourier discrète d’une signal digital ou d’une image	1
1.1.1	Définition de la transformation ; notion de dft	1
1.1.2	Aspects algorithmiques de la dft ; algorithmes de fft	3
1.1.3	Transformation de Fourier et corrélation cyclique : Fourier comme outil de calcul	8
1.1.4	Pourquoi l’opération de convolution est-elle si essentielle ?	9
1.1.5	Une réalité “physique” : le principe d’incertitude	11
1.2	Filtrage et perception “visuelle” d’une information	13
1.2.1	Le filtrage	13
1.2.2	La perception “visuelle” d’une information	14
1.2.3	L’algorithme pyramidal	16
1.2.4	Les idées soutendant jpeg repensées	19
1.2.5	Un exemple traité via l’algorithme pyramidal	20
2	Critère d’entropie ; analyse-compression des signaux ou des images	23
2.1	La notion d’entropie dans le cadre Hilbertien	23
2.1.1	Dimension théorique d’une information	25
2.2	Le <i>splitting lemma</i> (lemme de décomposition) dans le cadre dyadique	26
2.3	L’algorithme “ <i>Split and Merge</i> ” de V. Wickerhauser	29
2.3.1	Le cadre des signaux digitaux	29
2.3.2	Le cas des images digitales	32
2.3.3	Implémentation algorithmique en 1D	32
2.3.4	La généralisation au cadre p -adique	37
3	La lecture temps-échelles d’une information	43
3.1	L’analyse temps-échelles ; pourquoi faire ?	43
3.2	La décomposition de Franklin d’un signal digital	44
3.3	L’instrument d’analyse : l’analyse multi-résolution dyadique	47
3.4	Analyse temps-échelles I. De l’information à son “résumé”.	49
3.5	La mère de l’analyse multi-résolution	52
3.6	Analyse temps-échelles II : de l’information à ses détails	54
3.7	La décomposition traitée sur un exemple	56

Chapitre 1

L'outil Fourier

1.1 La transformée de Fourier discrète d'un signal digital ou d'une image

1.1.1 Définition de la transformation ; notion de dft

Un signal digital (au sens de l'analyse) de longueur $N \in \mathbb{N}^*$ est par définition un vecteur $(s(1), \dots, s(N))$ (pour nous un vecteur ligne) de nombres complexes ; une image digitale (toujours au sens de l'analyse) de taille (N_1, N_2) avec $N_1, N_2 \in \mathbb{N}^*$, est par définition la donnée d'une matrice

$$[I(k_1, k_2)]_{\substack{1 \leq k_1 \leq N_1 \\ 1 \leq k_2 \leq N_2}}$$

(k_1 indice de ligne, k_2 indice de colonne, pour suivre la syntaxe d'un logiciel de calcul scientifique tel **MATLAB** ou **SciLaB**).

Dans une image digitale, le nombre $I(k_1, k_2)$ peut représenter une quantité scalaire (entière, rationnelle, réelle ou complexe) ou alors (comme c'est le cas le plus souvent) un code de couleurs. Dans le cas où $I(k_1, k_2)$ représente une quantité scalaire (par exemple un nombre réel), on peut visualiser l'image I par un graphique en trois dimensions. Sur la figure 1.1 ci dessous, on a par exemple, on a "visualisé" en trois dimension une image réelle I (ici une carte en relief) : La représentation de la même image transcrite en code de couleur fournit l'image digitale représentée sur la figure 1.2.

Si N est fixé, soit $W_N := \exp(-2i\pi/N)$; si N_1 et N_2 sont fixés, on considère de même $W_{N_1} := \exp(-2i\pi/N_1)$ et $W_{N_2} := \exp(-2i\pi/N_2)$.

La transformation de Fourier (monodimensionnelle) discrète d'ordre N (dite aussi **dft** pour "*Discrete Fourier Transform*" d'ordre N) transforme le signal de longueur N qu'est $s = (s(1), \dots, s(N))$ en le signal digital (aussi de longueur N) \hat{s} défini par

$$\hat{s}(k+1) = \sum_{l=0}^{N-1} s(l+1) W_N^{kl}, \quad k = 0, \dots, N-1.$$

La transformation de Fourier (bidimensionnelle) discrète d'ordre (N_1, N_2) (dite aussi **dft2** pour "*Discrete Fourier Transform 2-dimensional*" de bi-ordre (N_1, N_2)) trans-

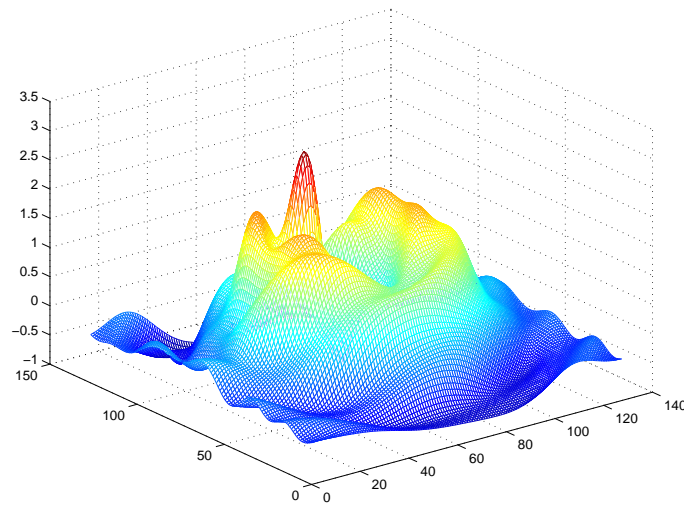


FIG. 1.1 – Une image digitale réelle en 3D

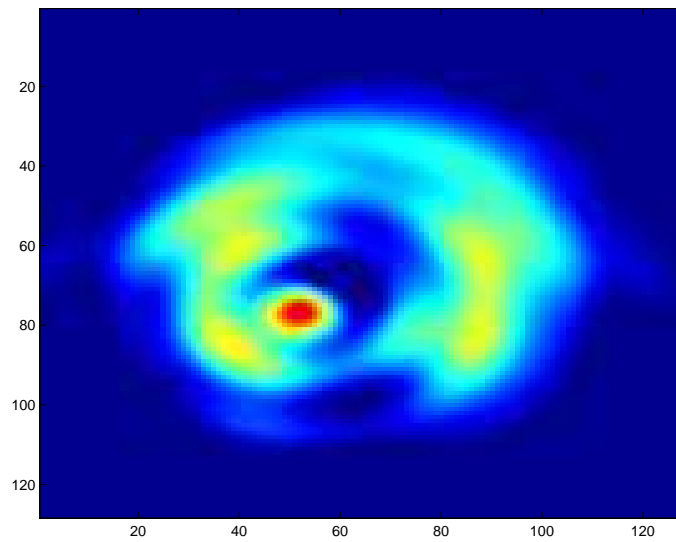


FIG. 1.2 – La même image digitale transcrite en codes de couleurs

forme l'image digitale de taille (N_1, N_2)

$$I := \left[I(k_1, k_2) \right]_{\substack{1 \leq k_1 \leq N_1 \\ 1 \leq k_2 \leq N_2}}$$

en l'image digitale (aussi de taille (N_1, N_2)) \hat{I} définie par

$$\begin{aligned} \hat{I}(k_1 + 1, k_2 + 1) &= \sum_{l_1=0}^{N_1-1} \sum_{l_2=0}^{N_2-1} I(l_1 + 1, l_2 + 1) W_{N_1}^{k_1 l_1} W_{N_2}^{k_2 l_2} \\ k_1 &= 0, \dots, N_1 - 1, \quad k_2 = 0, \dots, N_2 - 1; \end{aligned}$$

cette transformation bidimensionnelle s'opère en deux temps via une transformation monodimensionnelle d'ordre N_1 appliquée sur les colonnes de la matrice I , puis ensuite une seconde transformation monodimensionnelle d'ordre N_2 , cette fois sur les lignes de la matrice transformée.

Comme la matrice

$$\mathcal{W}_N = \left[W_N^{(k-1)(l-1)} \right]_{1 \leq k, l \leq N}$$

vérifie

$$\mathcal{W}_N \bullet \overline{\mathcal{W}}_N = N \mathbf{I}_N,$$

où \mathbf{I}_N est la matrice identité d'ordre N , on a les formules d'inversion à la fois pour les transformées mono ou bi-dimensionnelles.

- dans le cas monodimensionnel, on récupère s en appliquant à \hat{s} la transformation de Fourier discrète inverse (**idft** d'ordre n pour “Inverse Discrete Fourier Transform” d'ordre n) qui transforme le signal digital \hat{s} de taille N en le signal s (aussi de taille N) donné par

$$s(k + 1) = \frac{1}{N} \sum_{l=0}^{N-1} \hat{s}(l + 1) \overline{W_N^{kl}}, \quad k = 0, \dots, N - 1.$$

- dans le cas bidimensionnel, on récupère I en appliquant à \hat{I} la transformation de Fourier discrète inverse (**idft2** de bi-ordre (N_1, N_2) pour “Inverse Discrete Fourier Transform 2-dimensional” de bi-ordre (N_1, N_2)) qui transforme l'image digitale \hat{I} de taille (N_1, N_2) en l'image digitale I (aussi de taille (N_1, N_2)) donnée par

$$\begin{aligned} I(k_1 + 1, k_2 + 1) &= \frac{1}{N_1 N_2} \sum_{l_1=0}^{N_1-1} \sum_{l_2=0}^{N_2-1} I(l_1 + 1, l_2 + 1) \overline{W_{N_1}^{k_1 l_1}} \overline{W_{N_2}^{k_2 l_2}} \\ k_1 &= 0, \dots, N_1 - 1, \quad k_2 = 0, \dots, N_2 - 1. \end{aligned}$$

1.1.2 Aspects algorithmiques de la dft ; algorithmes de fft

C'est à Cooley et Tuckey (1966) que revient l'idée du calcul algorithmique de l'action de la **dft** lorsque l'ordre N est une puissance de 2, $N = 2^k$. Le nombre de multiplications nécessité par la multiplication d'un vecteur colonne de longueur 2^k par la matrice

$$\mathcal{W}_{2^k}[s]$$

(ce calcul correspond précisément à l'action de la DFT d'ordre 2^k sur le vecteur ligne s correspondant) se trouve ainsi réduit à $kN/2$ au lieu de ce qu'il devait être *a priori*, c'est à dire N^2 . C'est cette réduction drastique qui rend aujourd'hui possible l'usage intensif de la transformation de Fourier discrète, non seulement en analyse et traitement du signal, mais dans des pans entiers des mathématiques appliquées ou de l'informatique. L'algorithme de Cooley-Tuckey marque le début de ce que l'on pourrait qualifier de "révolution numérique".

L'idée clef de Cooley-Tuckey est de remarquer que, si N est une puissance de 2, $N = 2^k$, alors $N/2$ aussi ($N/2 = 2^{k-1}$), et par conséquent

$$W_N^{N/2} = e^{-i\pi} = -1.$$

C'est sur la cellule de calcul élémentaire dite *cellule papillon* que s'articule toute l'architecture de l'algorithme de Cooley-Tuckey ; il s'agit d'une cellule nécessitant deux additions (et aucune multiplication, hormis la multiplication triviale par 1) et transformant le vecteur (z_1, z_2) en le vecteur

$$\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix},$$

cellule que l'on peut représenter (c'est de là que vient la terminologie) comme sur la figure 2.1

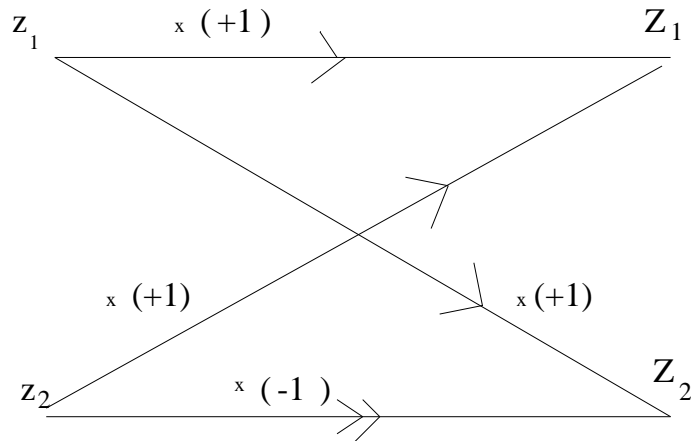
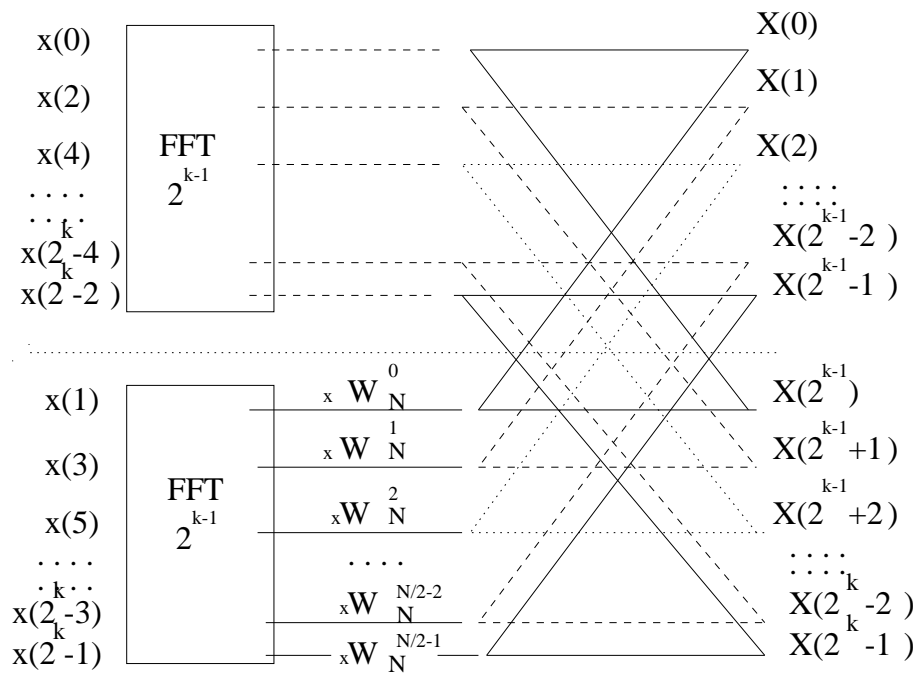


FIG. 1.3 – Cellule papillon

Cette opération correspond d'ailleurs à l'opération de transformée de Fourier discrète lorsque $N = 2$. Exprimons sur un diagramme (figure 2.2) l'architecture de l'algorithme correspondant à $N = 2^k$ en fonction de celle correspondant à $N = 2^{k-1}$.

FIG. 1.4 – *Algorithme de Cooley Tuckey* $N/2 = 2^{k-1} \rightarrow N = 2^k$

On voit que l'algorithme doit commencer par une phase de ré-agencement des entrées (qui se traduit de fait par un renversement des bits dans l'écriture de l'indice d'entrée en base 2) ; ensuite les mécanismes s'enchainent comme indiqué sur le diagramme ci-dessus et la consommation au niveau des opérations arithmétiques significatives est, comme on le voit immédiatement, de

$$\frac{N}{2} + 2\frac{N}{4} + 4\frac{N}{8} + \dots + 2^{k-1}\frac{N}{2^k} = \frac{kN}{2}$$

comme annoncé. Cet algorithme est implémenté dans l'environnement **MATLAB** sous l'une ou l'autre des commandes

```
>>X=fft(x);
>>X=fft(x,2^k);
```

(on parle de "Fast Fourier Transform", d'où la terminologie **fft**). Dans le second cas, le signal est soit tronqué, soit complété trivialement par des zéros suivant que sa longueur est supérieure ou inférieure ou égale à 2^k . La transformation inverse est un algorithme exactement du même type, mis à part le fait que W_N est remplacé par \bar{W}_N et qu'il y a une division finale par N . Sous l'environnement **MATLAB**, ce sont l'une ou l'autre des commandes

```
>>X=ifft(x);
>>X=ifft(x,2^k);
```

qui réalisent l'inverse de la transformée de Fourier discrète lorsque N est une puissance de 2. Les versions bidimensionnelles sont données par les commandes

```
>>A=fft2(I);
>>A=fft2(I,2^(k1),2^(k2));
>>I=ifft2(A);
>>I=ifft2(A,2^(k1),2^(k2));
```


Notons que les transformations tronquées ne sont pas inverses l'une de l'autre (les tailles ne sont pas respectées), tandis que l'on a bien toujours

```
>>X=ifft(fft(X));
>>A=ifft2(fft2(A));
```

Après Cooley-Tuckey, d'autres méthodes, inspirées de résultats simples d'arithmétique dans \mathbf{Z} , ont été introduites pour construire des algorithmes rapides pour la **dft** d'ordre un entier N qui ne soit plus *a priori* une puissance de 2; des logiciels de calcul tels **MATLAB** ou **SciLaB** utilisent ces algorithmes, dus essentiellement à B. Gold et S. Winograd (dans les années 1975-1980), et que nous présenterons succinctement ici.

C'est un résultat bien connu d'arithmétique (dit *théorème chinois* ou théorème des restes chinois) qui nous permet d'affirmer qu'étant donnés p nombres entiers N_1, \dots, N_p strictement positifs, premiers entre eux deux à deux, il existe un isomorphisme d'anneaux

$$\mathbf{Z}/(N_1 \cdots N_p)\mathbf{Z} \xrightarrow{\Phi} \mathbf{Z}/N_1\mathbf{Z} \times \cdots \times \mathbf{Z}/N_p\mathbf{Z}.$$

Cet isomorphisme Φ est d'ailleurs très facile à matérialiser : c'est l'application qui à la classe de k modulo $N_1 \cdots N_p$ associe l'élément

$$(\dot{k}_1, \dots, \dot{k}_p),$$

où \dot{k}_l est la classe (modulo N_l) de k .

Plaçons nous pour simplifier dans le cas $p = 2$. Si k et l sont deux entiers entre $(0, N_1 N_2 - 1)$, il leur correspond d'après le théorème chinois, et de manière unique, des couples $(k_1, k_2), (l_1, l_2)$ d'éléments de $\{0, \dots, N_1 - 1\} \times \{0, \dots, N_2 - 1\}$. D'autre part, l'algorithme d'Euclide conduit, si N_1 et N_2 sont premiers entre eux, au calcul de deux entiers U_1 et U_2 (eux non uniques, tandis que leurs classes respectives modulo N_2 et N_1 le sont) tels que

$$1 = U_1 N_1 + U_2 N_2.$$

On voit que le restes respectifs de $U_1 N_1$ modulo N_1 et N_2 sont $(0, 1)$ tandis que ceux de $U_2 N_2$ sont $(1, 0)$. On a donc

$$kl \equiv k_1 l_1 U_2 N_2 + k_2 l_2 U_1 N_1 \pmod{N_1 N_2},$$

ce qui implique que l'on puisse écrire

$$W_{N_1 N_2}^{kl} = W_N^{k_1 l_1 U_2 N_2 + k_2 l_2 U_1 N_1} = (W_{N_1}^{U_2})^{k_1 l_1} (W_{N_2}^{U_1})^{k_2 l_2},$$

et par conséquent réécrire différemment les formules

$$\hat{s}(k+1) = \sum_{l=0}^{N_1 N_2 - 1} s(l+1) W_{N_1 N_2}^{kl}, \quad k = 0, \dots, N_1 N_2 - 1$$

qui décrivent l'action de la transformation de Fourier discrète d'ordre $N_1 N_2$, sous la forme

$$\hat{s}((k_1, k_2) + 1) = \sum_{l_1=0}^{N_1-1} \sum_{l_2=0}^{N_2-1} s((l_1, l_2) + 1) (W_{N_1}^{U_2})^{k_1 l_1} (W_{N_2}^{U_1})^{k_2 l_2},$$

pour $(k_1, k_2) \in \{0, \dots, N_1 - 1\} \times \{0, \dots, N_2 - 1\}$, où nous avons joué sur la correspondance biunivoque $k \leftrightarrow (k_1, k_2)$ et $l \leftrightarrow (l_1, l_2)$. Cette idée judicieuse permet d'enchaîner les mécanismes de **dft** correspondant à N_1 et N_2 suivant le mécanisme algorithmique suivant :

- On effectue une permutation des lignes et des colonnes de la matrice $[W_N^{kl}]_{0 \leq k, l \leq N-1}$ en remplaçant l'ordre naturel sur $\{0, \dots, N-1\}$ par l'ordre lexicographique imposé par la correspondance $l \leftrightarrow (l_1, l_2)$, où l_1 est la classe modulo N_1 et l_2 la classe modulo N_2 . Réaliser ces permutations sur les lignes et les colonnes de la matrice revient à effectuer (comme d'ailleurs dans l'algorithme de Cooley-Tuckey) un ré-agencement des entrées et des sorties.
- Une fois ce travail fait, l'aspect de la matrice $\mathcal{W}_{N_1 N_2}$ est celui d'une matrice $N_1 N_2 \times N_1 N_2$ se présentant sous la forme d'une matrice constituée de N_1^2 blocs de taille (N_2, N_2) , le bloc (r, s) où $0 \leq r, s \leq N_1 - 1$ correspondant à la matrice \mathcal{W}_{N_2} (dans laquelle W_{N_2} se trouve remplacé par $W_{N_2}^{U_1}$) multipliée par le coefficient en position (r, s) de la matrice \mathcal{W}_{N_1} (dans laquelle W_{N_1} aurait été remplacé par $W_{N_1}^{U_2}$). C'est cette architecture que présente la figure 2.3.

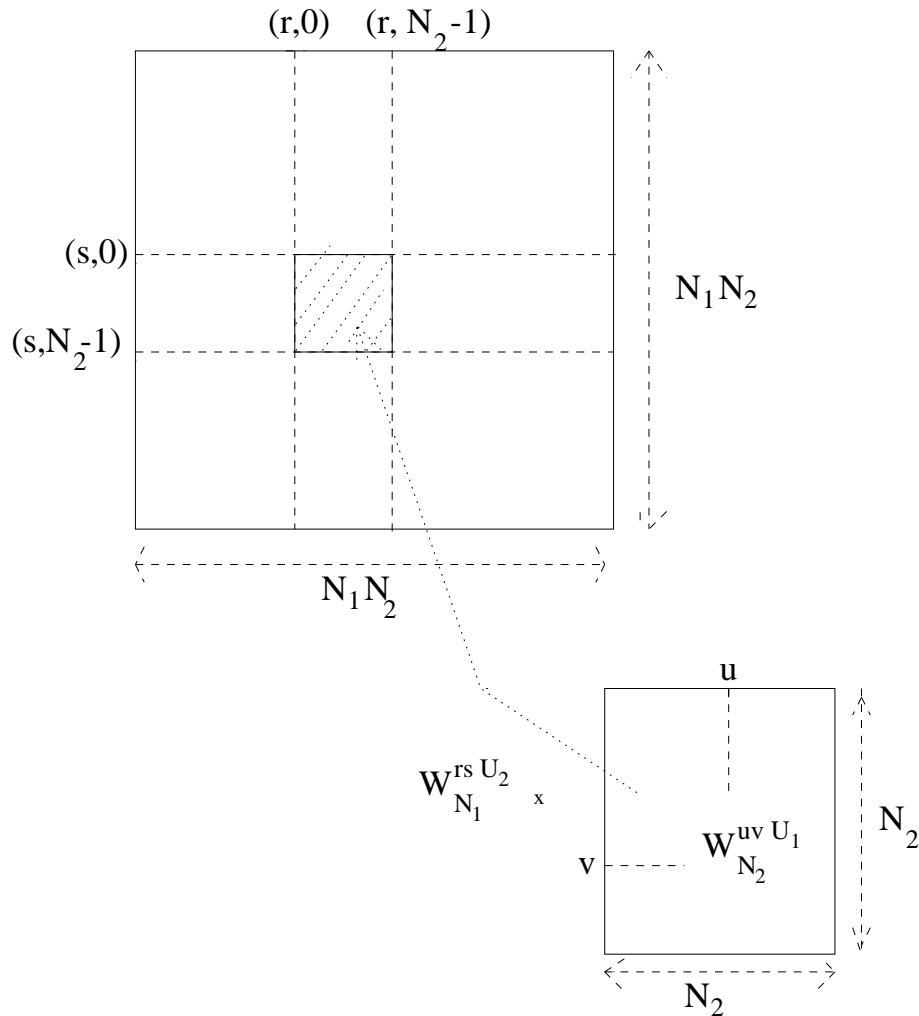


FIG. 1.5 – Enchaînement de Gold-Winograd (N_1, N_2)

Il est alors immédiat de voir comment les mécanismes s'enchaînent, une multiplication au niveau de l'algorithme de **dft** pour N_1 générant au niveau des calculs

enchaînés une prise de **dft** relative à la dimension N_2 . Si le mécanisme de **dft** relatif à N_1 consomme α_1 additions et μ_1 multiplications (attention ! il faut prendre garde à compter les multiplications triviales par $+1$ ou -1 telles celles qui apparaissent dans l'algorithme de **fft** pour les puissances de 2) et celui relatif à N_2 α_2 additions et μ_2 multiplications (même remarque au niveau des précautions de comptage), l'algorithme enchaîné va consommer $\mu_1\mu_2$ multiplications et $\alpha_2\mu_1 + \alpha_1\mu_2$ additions.

C'est sur ce modèle d'enchaînement que sont construits les mécanismes de transformation de Fourier rapide (on dit encore **fft**) qu'utilise l'environnement **MATLAB** lorsque N n'est plus une puissance de 2 mais (par exemple) un produit de puissances de petits nombres premiers. Un gain ne serait-ce que d'une multiplication dans le mécanisme de **dft** d'ordre p avec p premier s'avère donc très appréciable au niveau de la rapidité des calculs dans les enchaînements du type Gold-Winograd.

1.1.3 Transformation de Fourier et corrélation cyclique : Fourier comme outil de calcul

Signalons enfin que c'est S. Winograd qui a remarqué que la complexité algorithmique de l'action de \mathcal{W}_N était exactement la même que celle de l'opération de *corrélation cyclique d'ordre N* , à savoir : étant donnés deux polynômes

$$P(X) = x_0 + x_1X + \cdots + x_{N-1}X^{N-1}, \quad Q(X) = y_0 + y_1X + \cdots + y_{N-1}X^{N-1},$$

calculer les coefficients z_0, \dots, z_{N-1} du reste

$$[PQ : X^N - 1] = z_0 + z_1X + \cdots + z_{N-1}X^{N-1}$$

de PQ dans la division euclidienne par $X^N - 1$. Si y est supposé prolongé à \mathbf{Z} comme une suite périodique de période N , on voit que l'on a les formules

$$z_k = \sum_{q=0}^{N-1} x_q y_{k-q}, \quad k = 0, \dots, N-1,$$

et l'on voit apparaître ici pour la première fois dans un cadre discret et périodique l'opération mathématique qui soutendra l'aspect opérationnel de la transformation de Fourier, celle de *convolution* (ou de *filtrage*). Dans cette optique, nous avons la proposition immédiate, mais qui s'avèrera très importante du point de vue de l'utilisation future de la transformation de Fourier discrète :

Proposition 1.1 *Si $(x(1), \dots, x(N))$ et $(y(1), \dots, y(N))$ sont deux signaux digitaux de longueur N et*

$$\sum_{l=0}^{N-1} z(l+1)X^l = \left(\sum_{l=0}^{N-1} x(l+1)X^l \right) \left(\sum_{l=0}^{N-1} y(l+1)X^l \right) \bmod X^N - 1,$$

alors, pour tout k entre 0 et $N-1$, on a

$$\hat{z}(k+1) = \hat{x}(k+1)\hat{y}(k+1).$$

On a donc, si $z = (z(1), \dots, z(N))$ est le signal digital défini par

$$z(l+1) = \sum_{q=0}^{N-1} x(q+1)y(l+1-q), \quad l = 0, \dots, N-1,$$

où \tilde{y} est le prolongement de y par N -périodicité,

$$z = \text{idft}_N \left((\hat{x}(k+1)\hat{y}(k+1))_{k=0}^{N-1} \right).$$

Preuve. Elle est très simple, il suffit juste de remarquer que les deux polynômes

$$\sum_{l=0}^{N-1} z_{l+1} X^l$$

et

$$\left(\sum_{l=0}^{N-1} x(l+1) X^l \right) \left(\sum_{l=0}^{N-1} y(l+1) X^l \right)$$

coïncident pour $X = W_N^k$, $k = 0, \dots, N-1$. \diamond

On peut montrer aussi que, si le polynôme $X^N - 1$ se factorise en k facteurs de degré strictement positif dans $\mathbf{Q}[X]$, il existe toujours un algorithme réalisant la corrélation cyclique d'ordre N et consommant $2N - k$ multiplications, les multiplications par les rationnels n'étant pas comptées. C'est encore une application du lemme chinois pour plus de détails sur tous ces aspects algorithmiques (mais néanmoins importants) intervenant en analyse ou traitement de l'information.

L'outil Fourier permet tout autant de traiter la corrélation cyclique des images digitales que celle des signaux digitaux.

1.1.4 Pourquoi l'opération de convolution est-elle si essentielle ?

Soit L un appareil dont les paramètres restent immuables dans le temps ; L peut être réalisé (comme sur le schéma de la figure 1.6 ci-dessous, dit montage en boucle) *via* la mise en parallèle, en série ou en boucle (comme c'est le cas sur la figure) de cellules électriques (à base de condensateurs, résistances, bobines) ou mécaniques (à base de ressorts par exemple).

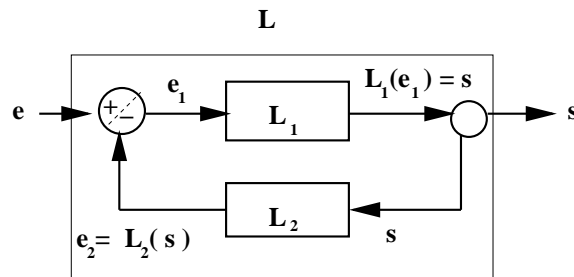


FIG. 1.6 – Schéma de boîte noire

En revanche, l'orgue acoustique constitué par le conduit vocal d'un locuteur ne saurait être assimilé à un tel appareil, puisque les paramètres de cet "orgue" se déforment bien sûr au cours du temps (lorsque le locuteur prononce une séquence de phonèmes).

Supposons que l'appareil réponde à l'impulsion e_0 définie comme la suite digitale

$$e_0 = (e_{0,n})_{n \in \mathbb{Z}}$$

définie par

$$e_{0,n} = \begin{cases} 1 & \text{si } n = 0 \\ 0 & \text{si } n \neq 0 \end{cases}$$

en renvoyant en sortie la suite digitale

$$h = (h_n)_{n \in \mathbb{Z}}$$

(cette suite h est appelée *réponse impulsionnelle* de l'appareil). Si l'appareil L agit de manière linéaire sur les entrées (on dit alors que c'est une *boîte noire* ou encore un *filtre*), il doit répondre à une suite $e = (e_n)_{n \in \mathbb{Z}}$ en renvoyant en sortie la suite $(s_n)_{n \in \mathbb{Z}} = L(e)$, où

$$s_n = \sum_{k \in \mathbb{Z}} e_k h_{n-k} = \sum_{k \in \mathbb{Z}} e_{n-k} h_k, \quad n \in \mathbb{Z}$$

(au moins formellement, on ne se préoccupe pas des convergences, toutes les suites ici sont supposées ne contenir qu'un nombre fini de digits non nuls).

En effet, on peut écrire

$$e = \sum_{k \in \mathbb{Z}} e_k e_0(\cdot - k),$$

où la suite $e_0(\cdot - k)$ est la suite dont le terme général vaut 1 si $n = k$ et 0 sinon (c'est l'impulsion e_0 décalée dans le temps de k); on a donc, puisque l'appareil L est censé agir de manière linéaire

$$L(e) = \sum_{k \in \mathbb{Z}} e_k L(e_0(\cdot - k))$$

et, du fait que les paramètres de l'appareil restent immuables dans le temps,

$$L(e_0(\cdot - k)) = (h_{n-k})_{n \in \mathbb{Z}}$$

(puisque $L(e_0) = (h_n)_{n \in \mathbb{Z}}$). Ceci justifie donc notre affirmation.

Ainsi l'opération de convolution discrète

$$(e, h) \rightarrow e * h$$

avec

$$(e * h)_n := \sum_{k \in \mathbb{Z}} e_k h_{n-k} = \sum_{k \in \mathbb{Z}} h_k e_{n-k}$$

(formellement) est-elle si importante du point de vue pratique! Le mérite de la transformation de Fourier comme outil de calcul est de transformer (on l'a vu dans la sous-section précédente) l'opération de convolution en une opération de multiplication à la fois plus maniable et moins coûteuse au niveau du temps de calcul. Le fait de disposer d'algorithmes rapides pour la transformation de Fourier et son inverse permet le traitement de la convolution discrète *via* un aller-retour par Fourier.

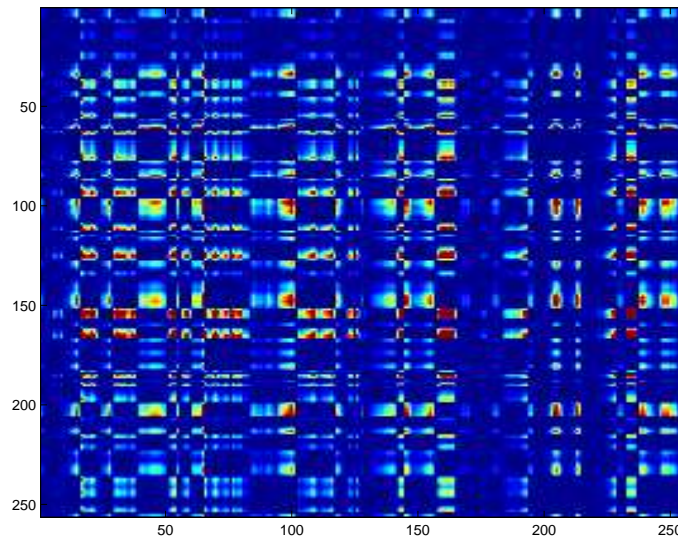


FIG. 1.7 – Une image digitale I avec à la fois trame périodique et “accidents”

1.1.5 Une réalité “physique” : le principe d’incertitude

Voici une seconde remarque importante : l’opération de prise de transformation de Fourier diffuse les objets localisés et localise les objets diffus. Par exemple, sur les figures 1.7 et 1.8, nous avons affiché une image présentant à la fois des structures régulières et périodiques (que l’on pourrait assimiler à une trame) en même temps que des “accidents” correspondant à des pics d’intensité bien localisés. La prise de spectre transforme ces pics localisés en structures diffuses et par contre transforme la trame de l’image en une structure bien localisée dans l’espace de Fourier. C’est ce que montrent les deux figures 1.7 et 1.8.

Ce qu’indique cette remarque trouve son explication dans le cadre de la physique. La transformation de Fourier (ou prise de spectre) correspondant au mécanisme physique de diffraction, elle ne localise pas les objets ; pire, plus une information est localisée, plus son spectre est diffus ; de même, plus le spectre est diffus, mieux l’information se trouve localisée.

On peut voir cela de manière plus “mathématique” : du point de vue mathématique, la transformée de Fourier transforme les objets en leurs objets duaux ; par exemple, si Λ est un réseau de \mathbf{R}^2 construit à partir d’une base (e_1, e_2) de vecteurs (de matrice A dans la base canonique), la transformation de Fourier 2D transforme ce réseau en le réseau dual (e_1^*, e_2^*) (de matrice ${}^t[A^{-1}]$ dans la base duale de la base canonique). Cette propriété est exploitée en cristallographie en dimension 3, où l’on sait comment un réseau cristallin est diffracté à travers un prisme.

Ce principe d’incertitude nous assure qu’autant l’analyse de Fourier (c’est-à-dire la prise de spectre) s’avère un bon outil pour dégager les *structures cohérentes* (voir périodiques) d’un signal ou d’une image, autant les détails (ou les accidents) de ce signal ou cette image se trouvent dilués dans l’espace des fréquences après la prise de spectre et par conséquent rendus indiscernables. La comparaison avec l’amphi de cours est intéressante : la prise de spectre permettra indubitablement de repérer (et d’isoler) la structure cohérente des rangées de tables ainsi que la répartition des étudiants, mais non les détails par exemple vestimentaires de chacun d’eux pris

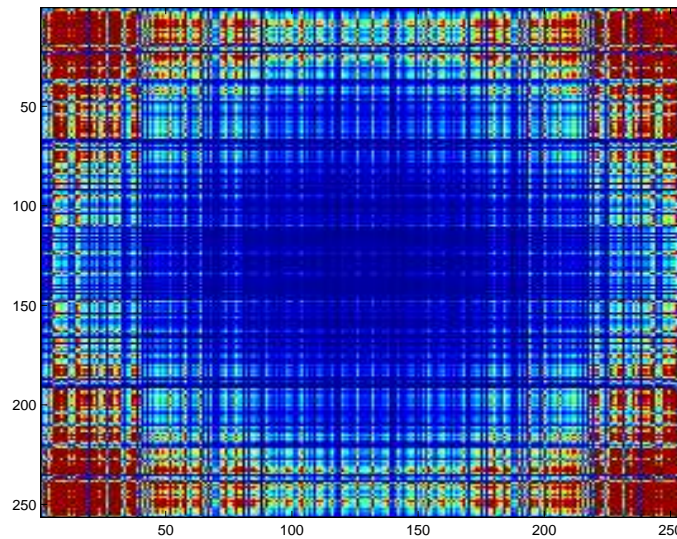
FIG. 1.8 – Le module de la transformée de Fourier de I 

FIG. 1.9 – Le codage musical en temps et en fréquences

individuellement. De fait, on sait depuis le moyen-âge que le codage des signaux avec deux paramètres, le temps t et la fréquence ω , s'avère beaucoup plus riche que le codage avec un seul paramètre (à savoir la fréquence) ; c'est le codage musical que matérialisent les partitions orchestrales comme sur la figure 1.9. La lecture du spectre d'un signal (fonction du seul paramètre ω) nous indique simplement comment, statistiquement, une fréquence ω donnée se révèle importante dans la décomposition d'un signal ; elle ne nous indique nullement à quels instants, pendant quels laps de temps, telle fréquence (ou telle autre) est importante. La lecture en temps et en fréquences répond, elle, à ces exigences plus précises et constitue donc une première parade au principe d'incertitude d'Heisenberg. On exploitera aussi cette idée aux fins du traitement des signaux ou des images en cryptographie. Le principe d'incertitude soutend d'ailleurs les idées sur lesquelles se fondent les versions originelles d'algorithmes de compression tels **jpeg** : on prend une image, on en calcule la transformée de Fourier (ou de fait la transformée en cosinus, qui correspond du point de vue complexité à une transformation équivalente) bloc par bloc (en général des blocs 8×8), ce qui nous donne une nouvelle image (ou plutôt une mosaïque d'images 8×8) dans le plan des fréquences. On plaque une grille contre cette nouvelle image et l'on fait une compression de données dans chaque "case" de la grille, puis on revoie l'image ainsi comprimée par Fourier inverse toujours bloc par bloc. La compression se base sur le principe selon lequel l'oeil est moins sensible aux

composantes hautes fréquences de l’image, on privilégie donc les composantes “bases fréquences” dans les diverses cases. Comme les détails de l’image (dans chaque case) ont été rendus diffus par Fourier, la compression dans les diverses cases se trouve ne pas trop affecter les informations devenues maintenant redondantes. On reviendra dans ce cours sur des versions “modernes” de **jpeg** mêlant les idées empruntées à l’outil Fourier et à l’analyse multi-résolution.

1.2 Filtrage et perception “visuelle” d’une information

1.2.1 Le filtrage

L’opération de *convolution discrète* (par une suite $(h_n)_{n \in \mathbb{Z}}$, que l’on supposera ne contenir qu’un nombre fini de termes non nuls) transforme, on vient de le voir, une information d’entrée $(e_n)_{n \in \mathbb{Z}}$ en l’information de sortie $(s_n)_{n \in \mathbb{Z}}$, où

$$s_n = \sum_{k \in \mathbb{Z}} e_k h_{n-k} = \sum_{k \in \mathbb{Z}} h_k e_{n-k}.$$

On peut imaginer aussi cette opération “cyclique” comme au paragraphe 1.1.4 ci-dessus en concevant les signaux e, h, s donnés sur $\{0, \dots, N-1\}$ et prolongés ensuite par périodicité à \mathbb{Z} tout entier ; ce point de vue est un peu différent et l’on parle alors de *convolution cyclique* (ou parfois de *corrélacion cyclique*) ; notons que ce qualificatif de “corrélacion” est un peu équivoque ici car la corrélacion de deux informations discrètes $(s_n^{(1)})_{n \in \mathbb{Z}}$ et $(s_n^{(2)})_{n \in \mathbb{Z}}$ constituées de suites de nombres complexes est par définition le produit scalaire (au sens de l’énergie) dans $l^2(\mathbb{Z})$, soit

$$\langle s^{(1)}, s^{(2)} \rangle := \sum_{n \in \mathbb{Z}} s_n^{(1)} \overline{s_n^{(2)}};$$

la *fonction d’autocorrélacion* d’une information $(s_n)_{n \in \mathbb{Z}}$ est, en revanche, la fonction

$$k \in \mathbb{Z} \rightarrow \sum_{n \in \mathbb{Z}} s_{n+k} \overline{s_n}$$

et c’est ici que l’on retrouve l’opération de convolution (d’où la terminologie un peu abusive “corrélacion cyclique”). On verra plus loin dans ce cours combien les notions de corrélacion (et de fonction d’autocorrélacion) sont importantes en traitement de l’information.

La convolution discrète est aussi appelée *filtrage* et l’aspect “outil” de la transformation de Fourier est le fait qu’elle transforme la convolution en multiplication. Si e et h sont deux suites indexées par \mathbb{Z} (et dont tous les termes sont nuls sauf un nombre fini), le produit des deux fonctions 2π périodiques (ce sont même ici des polynômes trigonométriques)

$$\omega \rightarrow \sum_{k \in \mathbb{Z}} e_k e^{-ik\omega}$$

et

$$\omega \rightarrow \sum_{k \in \mathbb{Z}} h_k e^{-ik\omega}$$

(dites “spectres” de e et h) est égal à la fonction

$$\omega \rightarrow \sum_{k \in \mathbb{Z}} s_k e^{-ik\omega},$$

où $s = e * h$ est par définition la convolée discrète de e et h . Dans le cas de la convolution cyclique, le spectre de l'entrée $e = (e(0), \dots, e(N-1))$ n'est évalué qu'aux points $0, 2\pi/N, \dots, (N-1)\pi/N$ et la suite de ces évaluations constitue la transformée de Fourier discrète de e ; le produit des transformées de Fourier discrètes des deux suites périodiques (de période N) e et h est, on l'a vu à la section 1.1.4, la transformée de Fourier discrète de la convolée périodique des suites e et h .

Lorsque l'on pense au mot filtrage, on pense à l'élimination du bruit dans une information : ceci n'est pas toujours le cas ; il existe en effet des *filtres intégrateurs* tels

$$(e_n)_{n \in \mathbb{Z}} \rightarrow \left(\frac{e_n + e_{n+1}}{2} \right)_{n \in \mathbb{Z}}$$

qui ont cette propriété (notons que la somme des h_n vaut $1/2 + 1/2 = 1$ dans ce cas), mais aussi des *filtres dérivateurs* comme

$$(e_n)_{n \in \mathbb{Z}} \rightarrow (e_n - e_{n+1})_{n \in \mathbb{Z}}$$

qui, eux, ont tendance à mettre en valeur les irrégularités d'une information au lieu de les lisser comme le fait un filtre intégrateur.

Si l'on pense au cadre continu, la dérivation crée des oscillations additionnelles à partir d'un événement assimilable à une impulsion. Sur la figure 1.10 par exemple, on a représenté pour mémoire les deux premières dérivées d'une gaussienne (notons que si la gaussienne joue un rôle si important comme modèle de particule élémentaire, c'est qu'elle réalise le meilleur compromis au niveau du principe d'incertitude d'Heisenberg : la fonction $t \rightarrow \exp(-t^2/2)$ est en effet un vecteur propre (correspondant à la valeur propre $\sqrt{2\pi}$ de la prise de spectre continu qui à un signal d'énergie finie s dans l'espace des temps associe le signal

$$\omega \rightarrow \int_{\mathbb{R}} s(t) e^{-i\omega t} dt$$

(l'intégrale converge au sens de l'énergie) dans l'espace des fréquences. Plus on drève la gaussienne, plus on crée d'oscillations ; cette remarque banale nous sera importante par la suite.

1.2.2 La perception “visuelle” d'une information

Toute information physique, qu'elle dépende d'un paramètre (celui du temps, auquel cas on parle de signal) ou de plusieurs (auquel cas on parle d'image) doit être quantifiée de manière digitale aux fins d'être traitée numériquement. Le mécanisme de quantification numérique nécessitant une discrétisation de l'espace des temps ou du support bi-dimensionnel sur lequel se trouve tracée l'image, on a le plus souvent affaire à un mécanisme de “stockage” local d'information en vue d'une moyennisation locale.

Or il se trouve que l'une des phases du mécanisme de la vision rétinienne consiste précisément en cette phase de “stockage” d'informations : se forme au creux de la rétine une image correspondant à une moyennisation locale de l'information.

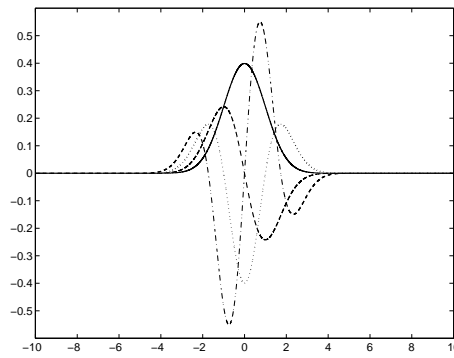


FIG. 1.10 – Les dérivées successives d’une gaussienne

Mais le mécanisme de la vision se double d’une seconde opération, celle qui consiste en la redistribution (effectuée cette fois au niveau du support cérébral) des moyennes ainsi stockées. Il s’agit là d’une opération tout à fait analogue à celle qui s’effectue au terme de la transmission d’une image prise par un satellite, “compressée” au niveau du satellite de manière à ce que la quantité de données à transmettre soit minimale, puis recomposée sur terre à partir d’une redistribution de ces données. Cet algorithme (car il s’agit bien d’un algorithme) est appelé *algorithme pyramidal*. Le diagramme ci-dessous illustre ce mécanisme :

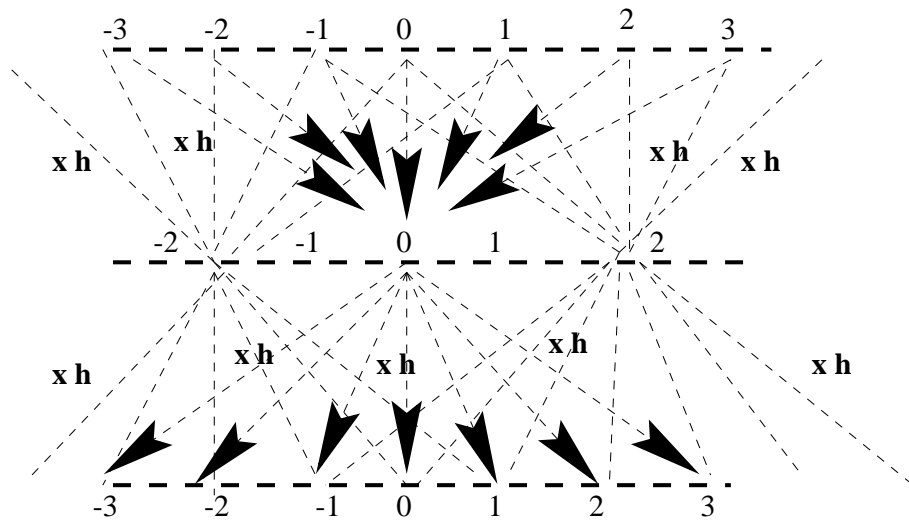


FIG. 1.11 – Le mécanisme de la vision

À chaque cran de l’algorithme pyramidal, la version récomposée à partir de la redistribution de la version stockée apparaît comme une version “brouillée” de l’information. La différence entre cette version brouillée $VB[s]$ et la version originelle s représente les détails que l’œil perçoit à ce niveau de l’algorithme. Bien sûr, la version moyennisée (celle où la quantité d’information se trouve divisée par deux) peut être traitée à nouveau et l’on obtient ainsi une nouveau signal correspondant à des détails, mais à un niveau de lecture correspondant à la moitié du niveau de détails obtenu précédemment.

Nous allons présenter de manière moins heuristique ces idées dans la section suivante.

1.2.3 L'algorithme pyramidal

Cette section constitue les prémisses de l'analyse temps-échelles que nous étudierons plus en détail au chapitre suivant. Le filtrage passe-bas (ou “moyennisant”) a pour effet, on l'a vu dans la section 1.2.1, de gommer les composantes du signal dont le module de la pulsation dépasse un seuil de coupure donné, et ce en effectuant ce que l'on pourrait qualifier de procédé de *moyennisation*. Le mécanisme de la vision rétinienne, s'il réalise bien un traitement similaire, l'effectue en la doublant d'un mécanisme de *décimation*. L'information moyennisée est stockée tous les deux noeuds du maillage temporel (elle devient alors un *résumé* du signal), puis redistribuée sur le maillage original pour générer une version brouillée (ou “blurred version”) dont la différence avec le signal de départ constitue les *détails* que le mécanisme de la vision a, de part son fonctionnement, gommé dans un premier temps.

Ce mécanisme est simple à modéliser du point de vue mathématique sur l'espace $l^2(\mathbb{Z})$ des signaux discrets d'énergie finie (la modélisation serait en tout point identique sur l'espace des images discrètes d'énergie finie). On se donne une suite “moyennisante” $(h(n))_{n \in \mathbb{Z}}$ avec $h(-n) = h(n)$ pour tout n , $h(0) \neq 0$, $h(n) = 0$ si $|n| \geq M$ et deux conditions supplémentaires ; la première est

$$\sum_{n \in \mathbb{Z}} h(n) = 1$$

(c'est la condition *sine qua non* pour une suite d'être une suite moyennisante, on verra d'ailleurs plus tard que si l'on pense en termes d'énergie, il est plus judicieux de remplacer 1 par $\sqrt{2}$) ; la seconde est :

$$\sum_{n \in \mathbb{Z}} h(2n) = \sum_{n \in \mathbb{Z}} h(2n+1) = \frac{1}{2},$$

ceci signifiant que dans le processus de redistribution, la somme des paramètres de pondération au noeud n du maillage, soit

$$\sum_{k \in \mathbb{Z}} h(n-2k)$$

est indépendante du noeud. Comme modèle pratique, nous utiliserons dans ce cours à titre d'exemple les suites du type $(h_a(n))_n$, où

$$\begin{cases} h_a(n) = 0, & |n| > 2 \\ h_a(0) = a \\ h_a(1) = h_a(-1) = 1/4 \\ h_a(2) = h_a(-2) = (1-2a)/4 \end{cases}$$

a désignant un paramètre entre 0 et 1. On peut imaginer d'autres suites, l'important étant de respecter les clauses ci-dessus.

L'opérateur qui décime et moyennise peut être considéré comme l'opérateur $R = R^{(h)}$ de $l^2(\mathbb{Z})$ dans $l^2(\mathbb{Z})$

$$R^{(h)} : s = s^{(0)} = (s_n)_{n \in \mathbb{Z}} \mapsto \left(\sum_{\nu \in \mathbb{Z}} s_\nu h(\nu - 2n) \right)_{n \in \mathbb{Z}}.$$

C’est bien le composé de deux opérateurs, le premier étant un opérateur de convolution discrète avec la suite paire h ,

$$\mathcal{L} = \mathcal{L}_h : (s_n)_{n \in \mathbb{Z}} \mapsto \left(\sum_{\nu \in \mathbb{Z}} s_\nu h(n - \nu) \right)_{n \in \mathbb{Z}} = \left(\sum_{\nu \in \mathbb{Z}} s_\nu h(\nu - n) \right)_{n \in \mathbb{Z}}$$

le second étant l’algorithme de décimation qui consiste à “sauter” un point sur deux en lisant l’information

$$\text{Dec} : (\sigma_n)_{n \in \mathbb{Z}} \mapsto (\sigma_{2n})_{n \in \mathbb{Z}}.$$

Le signal $R[(s_n)_n] = R_1[(s_n)_n] = (s_n^{(1)})_{n \in \mathbb{Z}}$ constitue le *résumé* du signal de départ. Il faut se souvenir que la quantité d’information est divisée par deux et que si l’on implémente cet algorithme comme nous l’avons fait sur l’espace des signaux digitaux de longueur $2^q + 1$, $q \in \mathbb{N}$, nous générons une application

$$R_{(q)}^{(q)} : \mathbb{C}^{2^q+1} \mapsto \mathbb{C}^{2^{q-1}+1}.$$

Nous avons implémenté ce premier algorithme comme suit :

```
>function f=rpyramid(s,a);
%Decimation + filtrage
%avec un filtre
%h(0)=a, h(1)=h(-1)=1/4
%h(2)=h(-2)=1/4-a/2
%h(k)=0 pour abs(k) plus grand que 3.
%pour une entree de longueur 2^q+1,
%l'algorithme genere une sortie
%de longueur 2^(q-1)+1
```

La redistribution des valeurs du résumé sur le maillage temporel original (c’est à dire avant décimation) se fait suivant l’algorithme

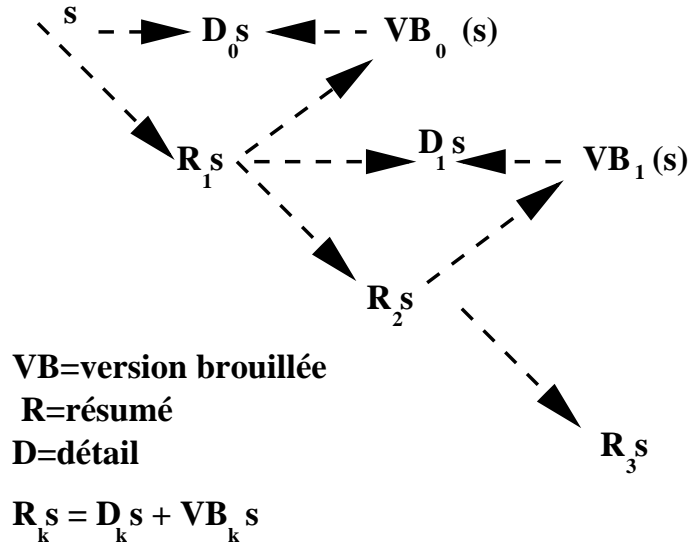
$$R^* : (s_n^{(1)})_{n \in \mathbb{Z}} \mapsto \left(\sum_{\nu \in \mathbb{Z}} s_\nu^{(1)} h(n - 2\nu) \right)_{n \in \mathbb{Z}}.$$

et l’on vérifie immédiatement que c’est bien l’adjoint de R dont l’action apparaît ici, ce qui justifie la notation utilisée. Cette fois, il faut se souvenir que la quantité d’information est multipliée par deux (car on redistribue le résumé aux noeuds du maillage originel) et que par conséquent, si l’on implémente cet algorithme comme nous l’avons fait sur l’espace des signaux digitaux de longueur $2^q + 1$, $q \in \mathbb{N}$, nous générons une application

$$R_{(q)}^{(h)*} : \mathbb{C}^{2^q+1} \mapsto \mathbb{C}^{2^{q+1}+1}.$$

Nous avons implémenté ce second algorithme comme suit :

```
>function f=dpymid(s,a);
%phase de redistribution
%pour l'algorithme a pyramide
%avec un filtre
%h(0)=a, h(1)=h(-1)=1/4
%h(2)=h(-2)=1/4-a/2
%h(k)=0 pour abs(k) plus grand que 3.
```

FIG. 1.12 – *Algorithme pyramidal*

```

%(a entre 0 et 1)
%pour une entree de longueur 2^q+1,
%l'algorithme genere une sortie
%de longueur 2^(q+1)+1

```

Au vu du facteur $1/2$ qui apparait comme la norme de l'opérateur de décimation, c'est le signal

$$VB[s^{(0)}] = 2R^* \circ R[s^{(0)}]$$

(ico $s^{(0)}$ désigne la suite $(s_n)_n$ traitée) qu'il faut considérer comme une version brouillée du signal $s = s^{(0)}$ (c'est d'ailleurs $2R_{(q)}^*$ qui se trouve implémenté dans notre programme **rpyramid** ci-dessus). Cette version brouillée est donnée par les formules

$$VB[s^{(0)}]_n = 2 \sum_{\nu_1=-\infty}^{\infty} \sum_{\nu_2=-\infty}^{\infty} h(n - 2\nu_1)h(\nu_2 - 2\nu_1)s_{\nu_2}^{(0)}.$$

Les détails qui ont échappé à ce brouillage, sont données comme le signal

$$D_0[s] := s^{(0)} - VB[s^{(0)}].$$

Il est maintenant naturel d'imaginer que cet algorithme puisse être itéré : une fois les détails de la génération 0 (en l'occurrence D_0) isolés, nous pouvons répéter l'analyse au niveau du résumé $R_1[s]$ du signal s , qui devient le signal de départ pour la génération 1. Sur la figure 1.12, nous avons présenté le synopsis de l'algorithme à la génération 0 (et l'on y retrouve la terminologie selon laquelle cet algorithme classique développé par Burt-Adelson mais inspiré des travaux de Marr sur la vision, est appelé *algorithme pyramidal* ou *algorithme à pyramide*). Le mécanisme de reconstruction du signal originel à partir de sa décomposition est donné par les formules récurrentes

$$R_k[s] = D_k[s] + 2R^*[R_{k+1}[s]].$$

Nous verrons dans les sections suivantes que l’opérateur R (pourvu que les paramètres h_n soient bien choisis, ce qui malheureusement nous obligera à des choix impliquant des nombres irrationnels, donc demandant nécessairement à être approchés) peut être couplé avec un opérateur D de prise de détails (au lieu de prise de résumé) de manière à ce que, sur $l^2(\mathbb{Z})$, on ait l’identité (que l’on pourrait qualifier de Bézout) entre opérateurs

$$\text{Id} = R^* \circ R + D^* \circ D;$$

l’opérateur D correspondra à une opération de filtrage combinée avec une opération de décimation, le filtrage s’effectuant cette fois selon un mécanisme accentuant les détails, soit avec un filtre dont les paramètres g_n , $n \in \mathbb{Z}$, satisfont

$$\sum_{n \in \mathbb{Z}} g_n = 0$$

au lieu de

$$\sum_{n \in \mathbb{Z}} h_n = 1$$

qui était la condition imposée aux paramètres du filtre impliqué dans la réalisation de R . De plus, la décomposition de l’information s en $R^*R[s]$ et $D^*D[s]$ sera une décomposition orthogonale (notons que cela peut être un avantage, mais aussi un inconvénient car imposer le découpage en objets orthogonaux fragilise souvent les diverses composantes !)

1.2.4 Les idées soutendant jpeg repensées

Supposons que l’on dispose d’opérateurs R et D de $l^2(\mathbb{Z})$ dans lui-même ayant les propriétés imposées dans la sous-section 1.2.3 précédente.

Une image digitale I (disons $2^k \times 2^k$) génère de fait 4 images de taille $2^{k-1} \times 2^{k-1}$:

- la première de ces images (notée $\mathbf{R}[I]$) s’obtient en faisant agir R dans les deux directions (on filtre et on décime) ; c’est une version “résumée” de l’image ;
- la seconde (notée $\mathbf{D}_{\text{vert}}[I]$), s’obtient en faisant agir R dans la direction verticale et D dans la direction horizontale, mettant ainsi plus en évidence les accidents “verticaux” de l’image ;
- la troisième (notée $\mathbf{D}_{\text{horiz}}[I]$) s’obtient en faisant agir R dans la direction horizontale et D dans la direction verticale, mettant ainsi plus en évidence les accidents “horizontaux” de l’image ;
- la quatrième enfin (notée $\mathbf{D}_{\text{obl}}[I]$) s’obtient en faisant agir D dans la direction horizontale et D dans la direction verticale, mettant ainsi plus en évidence les accidents “obliques” de l’image.

Sur le diagramme 1.13, nous avons représenté la manière dont peuvent être présentées ces 4 images $2^{k-1} \times 2^{k-1}$. Le mécanisme peut être répété (ou non) sur chacune des 4 images, un critère décisionnel convenable devant être défini nous disant si nous devons découper ou non. On voit donc ainsi diverses “décompositions” de l’image possible, décompositions au terme desquelles on pourra faire de la compression (comme dans **jpeg**) ; ces décompositions prennent en compte les paramètres de position, de fréquence, et d’échelle (la prise de spectre prenait en compte le paramètre fréquentiel). En ce sens, elles sont beaucoup plus riches et sont susceptibles

de générer des algorithmes de traitement d'images intéressants aux fins des questions de cryptologie nous occupant dans ce cours. Nous y reviendrons plus précisément au chapitre suivant.

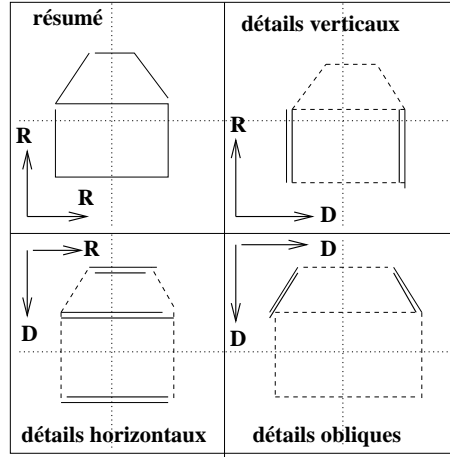


FIG. 1.13 – la décomposition d'images inspirée de l'algorithme de Burt-Adelson

1.2.5 Un exemple traité via l'algorithme pyramidal

Nous avons décomposé ici le signal

$$t \mapsto s(t) = \sin(40\pi t) + \cos(5.25\pi t(t - 3, 75)),$$

le temps variant entre $t = 0$ et $t = 10.24$ avec un pas $\tau = 0.01$; nous avons veillé à ce que la contrainte

$$|\omega_{\text{inst}}| \leq \frac{\pi}{\tau} = 100\pi$$

(condition de Nyquist) soit ici respectée, de manière à ne pas se trouver confronté à des problèmes de sous-échantillonnage. Le signal est représenté sur la figure 1.14 (en haut) est représentée sur la partie inférieure de cette même figure.

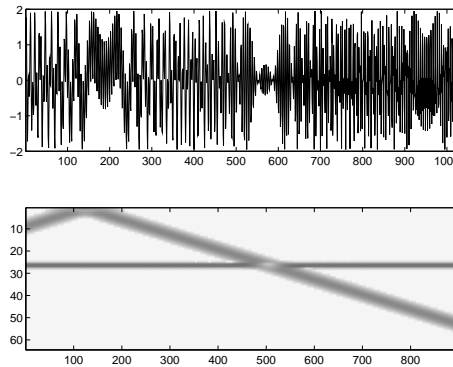
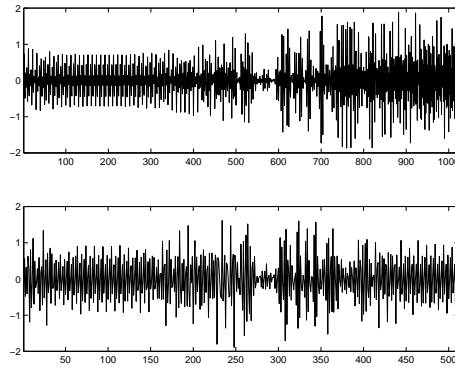
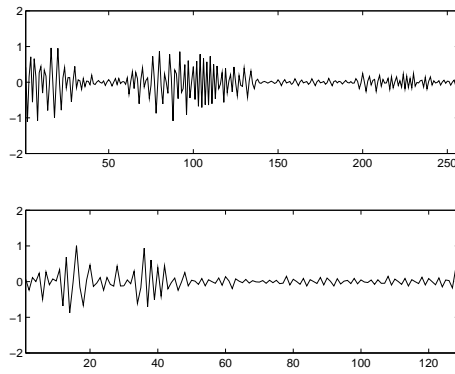
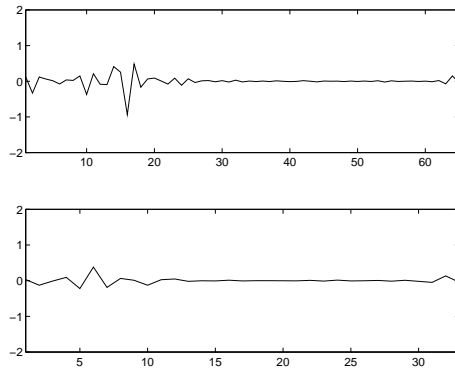


FIG. 1.14 – le signal traité et son image temps-fréquences

Les différents détails $d_1, d_2, d_3, d_4, d_5, d_6$ de la décomposition via l'algorithme pyramidal avec $h = h_a$ et $a = 0,6$ sont représentés sur les figures 1.15, 1.16, 1.17.

FIG. 1.15 – d_1, d_2 FIG. 1.16 – d_3, d_4 FIG. 1.17 – d_5, d_6

On voit que la composante stationnaire du signal (ici $t \mapsto \sin(40\pi t)$) passe dans les détails d_1 et d_2 (avec les hautes fréquences du “chirp” $t \mapsto \cos(5.25\pi t(t - 3.75))$) tandis que l’évolution vers les basses-fréquences de ce même “chirp” cette fois isolé se lit sur la suite des détails d_3 à d_6 .

Chapitre 2

Critère d'entropie ; analyse-compression des signaux ou des images

2.1 La notion d'entropie dans le cadre Hilbertien

L'entropie en thermodynamique est un indicateur quantitatif du chaos : minimiser l'entropie, c'est chercher à coder le signal en fonction de sa cohérence ; c'est aussi cette idée qui préside au maniement de la notion d'entropie dans le traitement numérique des informations digitales (1D ou 2D) tel que nous le proposons ici. La notion d'entropie est une notion d'analyse hilbertienne tout aussi essentielle que l'est celle d'énergie (dont la forme polarisée est, rappelons le, le produit scalaire de corrélation, par exemple dans le cadre de $l^2(\mathbb{Z})$,

$$\langle (s_n)_n, (\tilde{s}_n)_n \rangle := \sum_{n \in \mathbb{Z}} s_n \overline{\tilde{s}_n}.$$

C'est d'ailleurs essentiellement le cadre hilbertien de $l^2(\mathbb{Z})$ (ou $l^2(\mathbb{Z} \times \mathbb{Z})$), espace des signaux ou des images d'énergie finie (ou les pendants continus que sont respectivement $L^2(\mathbb{R}, dx)$, et $L^2(\mathbb{R}^2, dxdy)$) qui nous servira ici de modèle, étant entendu que signaux ou images seront toujours supposés ne contenir qu'un nombre fini de pixels $n \in \mathbb{Z}$ ou $(k_1, k_2) \in \mathbb{Z}^2$ affectés d'une valeur (ou d'un code de couleurs) non nulle.

Si $(s_n)_n$ une suite de réels tendant vers 0 à l'infini, $-\log_2 |s_n|^2$ correspond au nombre de bits nécessaires au codage de l'information $|s_n|^2$ et l'on peut voir la quantité

$$-\sum_{n \in \mathbb{Z}} |s_n|^2 \log_2 |s_n|^2$$

comme une entropie ; chercher à la minimiser revient aussi à faire en sorte que soit "le mieux organisée possible" la suite $(s_n)_{n \in \mathbb{Z}}$ (de préférence, peu de coefficients, mais significatifs, plutôt que beaucoup de coefficients, eux petits). D'ailleurs, si la suite $(s_n)_n$ est une suite composée simplement de zéros et de 1, l'exponentielle de la quantité

$$-\sum_{n \in \mathbb{Z}} |s_n|^2 \log_2 |s_n|^2$$

(avec la convention $0 \times \log 0 = 0$) figure exactement le nombre de 1 présents dans la suite en question.

Définition 2.1 Soit H un espace de Hilbert (séparable) et $(H_\iota)_\iota$ une décomposition (indexée forcément par un ensemble I dénombrable) orthogonale de cet espace

$$H = \bigoplus_{\iota \in I}^\perp H_\iota.$$

L'entropie d'un élément s de H relativement à cette décomposition est par définition :

$$e(s; (H_\iota)_\iota) := - \sum_{\iota} \|\text{pr}_{H_\iota}[s]\|^2 \log \|\text{pr}_{H_\iota}[s]\|^2.$$

La proposition suivante est une re-écriture immédiate du théorème de Pythagore ; c'est ce que l'on appelle l'équation de Shannon :

Proposition 2.1 Soient $H = H_1 \bigoplus^\perp H_2$ une décomposition orthogonale de H et $\mathcal{H}_1 = (H_{1,\iota_1})_{\iota_1 \in I_1}$ et $\mathcal{H}_2 = (H_{2,\iota_2})_{\iota_2 \in I_2}$ des décompositions orthogonales respectives de H_1 et H_2 :

$$H_1 = \bigoplus_{\iota_1 \in I_1}^\perp H_{1,\iota_1}, \quad H_2 = \bigoplus_{\iota_2 \in I_2}^\perp H_{2,\iota_2}.$$

Soit \mathcal{H} la décomposition de H obtenue en concaténant les décompositions \mathcal{H}_1 et \mathcal{H}_2 ; alors, si s est un signal normalisé (de norme 1) de H , on a

$$e(s; \mathcal{H}) = e(s; \{H_1, H_2\}) + \|s_1\|^2 e(s_1/\|s_1\|; \mathcal{H}_1) + \|s_2\|^2 e(s_2/\|s_2\|; \mathcal{H}_2),$$

si $s_j = \text{pr}_{H_j}[s]$, $j = 1, 2$.

Le corollaire de cette proposition immédiate est ce que l'on appelle le *critère de Shannon* ; c'est ce critère qui présidera à nos méthodes de minimisation de l'entropie et justifiera l'utilisation dans le cadre hilbertien de l'entropie que l'on vient de définir (dite *entropie de Shannon*), même si le recours à d'autres entropies, plus en relation avec ce que l'on prétend extraire de notre information, est envisageable, quand bien même le critère de Shannon ne serait plus satisfait. Notons toutefois qu'un produit scalaire induit un cadre hilbertien, et, par voie de conséquence, une entropie de Shannon adéquate.

Voici en tout cas le critère de Shannon :

Proposition 2.2 Si $H = H_1 \bigoplus^\perp H_2$, $s = s_1 + s_2 \in H$, de norme 1 ($s_j = \text{pr}_{H_j}[s]$, $j = 1, 2$), et si les décompositions orthogonales \mathcal{H}_1 et \mathcal{H}_2 de H_1 et H_2 sont telles que $e(s_1/\|s_1\|; \mathcal{H}_1)$ soit minimale (pour toutes les décompositions orthogonales possibles de H_1) ainsi que $e(s_2/\|s_2\|; \mathcal{H}_2)$ (cette fois pour toutes les décompositions orthogonales possibles de H_2), alors $e(s; \mathcal{H})$, où $\mathcal{H} := \{\mathcal{H}_1, \mathcal{H}_2\}$, est minimale (pour toutes les décompositions de H).

De fait, nous utiliserons aussi la version de l'entropie non normalisée, alors que le critère de Shannon nous indique que l'entropie d'un vecteur s se décomposant sous la forme

$$s = \sum_{k \in \mathbb{Z}} x_k$$

devrait être (pour que notre notion se plie à ce critère)

$$e_{\text{norm}}(s; \text{décomposition}) = - \sum_k \frac{\|x_k\|^2}{\|s\|^2} \log \frac{\|x_k\|^2}{\|s\|^2}.$$

Qu'à cela ne tienne, l'entropie définie par

$$e(s; \text{décomposition}) = - \sum_k \|x_k\|^2 \log \|x_k\|^2$$

(ou encore $L^2 \log L^2$ -norme) que nous utiliserons est telle que si s et s' sont deux vecteurs de même norme, alors

$$e(s; \text{décomp}) < e(s'; \text{décomp}) \iff e_{\text{norm}}(s; \text{décomp}) < e_{\text{norm}}(s'; \text{décomp});$$

trouver la base dans laquelle la décomposition de s est telle que cette entropie (non normalisée) soit minimale reviendra de fait à trouver la base relativement à laquelle l'entropie (elle normalisée) de s est minimale, et c'est là la raison pour laquelle nous utiliserons cette $L^2 \log L^2$ norme plutôt que l'entropie normalisée dans nos exemples à venir.

2.1.1 Dimension théorique d'une information

La convexité de la fonction $t \in]0, \infty[\mapsto t \log t$ implique que la dimension théorique d'une suite $(x_n)_{n \in \mathbb{N}}$ dont au plus N coefficients sont non nuls est comprise entre 1 et N : si $M > 0$ est en effet le nombre de x_k non nuls, on a, de par l'inégalité de convexité,

$$\log \left[\frac{1}{M} \sum_k \frac{|x_k|^2}{\|x\|^2} \right] \times \frac{\sum_k |x_k|^2}{M \|x\|^2} = - \frac{\log M}{M} \leq \frac{1}{M} \sum_k \frac{|x_k|^2}{\|x\|^2} \log \frac{|x_k|^2}{\|x\|^2},$$

ce qui implique que la dimension théorique de la suite $(x_n)_n$ est dans ce cas majorée par $\exp(\log M) = M$.

De plus, si $(x_n)_n$ et (x'_n) sont deux suites de N nombres réels que l'on peut réarranger en des suites décroissantes $(y_n)_n$ et $(y'_n)_n$ de manière à ce que, pour tout $n = 1, \dots, N$, le total cumulé de la suite réordonnée $(y_n)_n$

$$T_n := \sum_{k=1}^N y_k$$

domine le total cumulé

$$T'_n := \sum_{k=1}^N y'_k$$

de la suite réordonnée $(y'_n)_n$, alors la dimension théorique de $(x_n)_n$ est plus petite que celle de $(x'_n)_{n \in \mathbb{N}}$. L'entropie de Shannon est donc une bonne mesure de l'efficacité de concentration d'une décomposition. Elle permet en effet de comparer les suites par leur taux de décroissance, autrement dit, le taux avec lequel les éléments deviennent négligeables s'ils sont réarrangés dans l'ordre décroissant.

Il faut noter aussi que bien souvent en cryptographie (et en particulier dans les techniques récentes développées en *watermarking*, nous y reviendrons), il arrive que

l'on utilise des entropies *ad hoc* ne satisfaisant pas *a priori* le critère de Shannon, mais conduisant à la recherche d'une "meilleure base" dans laquelle se décompose une information. La "clef" devient alors le critère d'entropie lui-même conduisant à l'arbre de sélection de cette meilleure base *via* le lemme de décomposition que nous introduisons dans la section suivante.

2.2 Le *splitting lemma* (lemme de décomposition) dans le cadre dyadique

Soit m_0, m_1 deux fonctions définies presque partout sur \mathbb{R} , 2π -périodiques, d'énergie finie sur $[0, 2\pi]$ (c'est-à-dire des éléments de $L^2(\mathbb{R}/2\pi\mathbb{Z})$), telles que la matrice

$$M := \begin{pmatrix} m_0(\omega) & m_1(\omega) \\ m_0(\omega + \pi) & m_1(\omega + \pi) \end{pmatrix}$$

soit, pour presque tout ω , une matrice unitaire (les colonnes sont de norme 1 et les lignes sont orthogonales pour le produit scalaire usuel dans \mathbb{C}^2). Par exemple, si m_0 est un élément de $L^2(\mathbb{R}/2\pi\mathbb{Z})$ tel que

$$|m_0(\omega)|^2 + |m_0(\omega + \pi)|^2 \equiv 1 \quad \text{p.p.}, \quad (2.1)$$

le couple réalisé par les fonctions

$$\begin{aligned} m_0 &: \omega \mapsto m_0(\omega) \\ m_1 &: \omega \mapsto e^{-i(\omega+\pi)} \overline{m_0(\omega + \pi)} \end{aligned}$$

est un couple de fonctions qui convient ; notons toutefois que la contrainte sur m_0 est assez drastique : la fonction

$$\omega \mapsto \frac{1 + e^{-i\omega}}{2} = e^{\frac{-i\omega}{2}} \cos(\omega/2)$$

y satisfait ; en revanche, pour N fixé, on ne peut trouver qu'un seul polynôme trigonométrique

$$m_0^{(N)} = \left(\frac{1 + e^{-i\omega}}{2} \right)^N (1 + \gamma_{N,1} e^{-i\omega} + \cdots + \gamma_{N,N-1} e^{-i(N-1)\omega})$$

qui y satisfasse. De fait, nous nous contenterons de travailler avec des exemples de fonctions m_0 du type polynôme trigonométrique comme ci-dessus ; il faut noter que les exigences portant sur les coefficients (le fait que de tels polynômes trigonométriques soient en relation avec des identités trigonométriques très précises, sur le modèle de $\cos^2 + \sin^2 \equiv 1$) exclut que ces coefficients puissent être des nombres rationnels. Le côté "numérique" et non "arithmétique" de ces outils est ainsi évident. Cela peut bien sûr s'avérer un handicap sur des terrains comme la cryptographie où les aspects arithmétiques jouent un rôle majeur.

La fonction m_0 admet un développement en série de Fourier

$$m_0(\omega) = \frac{1}{\sqrt{2}} \sum_{n \in \mathbb{Z}} h(n) e^{-in\omega},$$

où $(h(n))_{n \in \mathbb{Z}} \in l^2(\mathbb{Z})$; la fonction m_1 qui lui est couplée admet alors le développement

$$m_1(\omega) = e^{-i(\omega+\pi)} \overline{m_0(\omega + \pi)} = \sum_{n \in \mathbb{Z}} (-1)^n \overline{h(1-n)} e^{-in\omega} = \sum_{n \in \mathbb{Z}} g(n) e^{-in\omega}$$

avec

$$g(n) = (-1)^n \overline{h(1-n)}, \quad n \in \mathbb{Z}.$$

On impose aussi la contrainte que la suite $(h(n))_{n \in \mathbb{Z}}$ soit dans $l^1(\mathbb{Z})$ et que

$$\sum_{n \in \mathbb{Z}} h(n) = \sqrt{2},$$

ce qui assure que les fonctions m_0 et m_1 soient des fonctions 2π -périodiques continues, avec de plus $m_0(0) = 0$ et $m_1(0) = 1$; de fait, on est obligé d'imposer une contrainte plus forte relativement aux convergences, à savoir qu'il existe $\epsilon > 0$ tel que l'on ait

$$\sum_{n \in \mathbb{Z}} |h(n)| |n|^\epsilon < +\infty; \quad (2.2)$$

pareille contrainte est bien sûr satisfaite dès que m_0 se trouve être un polynôme trigonométrique (comme c'est le cas si $m_0 = m_0^{(N)}$).

Du point de vue de l'analyse de Fourier, la convolution d'un signal discret $(s_n)_n$ avec la suite $(h(-n))_{n \in \mathbb{Z}}$ correspond à l'action sur le signal d'un filtrage passe-bas (les basses-fréquences sont préservées puisque $m_0(0) = 1$), tandis que la convolution du même signal avec la suite $(g(-n))_{n \in \mathbb{Z}}$ correspond à l'action sur ce même signal d'un filtrage passe-bande (les fréquences voisines de 0 sont coupées puisque $m_1(0) = 0$).

Ce n'est cependant pas ce point de vue du filtrage que nous allons privilégier, mais plutôt l'idée du filtrage couplé avec la décimation présente dans l'algorithme pyramidal.

Le mécanisme de la vision nous suggère de penser une information digitale $(s_n)_{n \in \mathbb{Z}}$ (disons pour fixer les idées dans $l^2(\mathbb{Z})$) comme le signal "continu" (et appartenant lui, à $L^2(\mathbb{R})$)

$$s : t \mapsto \sum_{n \in \mathbb{Z}} s_n \varphi(t - n),$$

où φ est un "atome temporel" bien localisé autour de $[0, 1]$, d'énergie finie, et tel que ses translatés $(\varphi(\cdot) - n)_{n \in \mathbb{Z}}$ constituent un système orthonormé dans $L^2(\mathbb{R})$ (ce qui rend "orthogonales" les diverses informations visuelles s_n , $n \in \mathbb{Z}$).

Notons que la donnée de la fonction m_0 suggère (tout au moins formellement), la construction d'un tel atome φ , celui que l'on construit, grâce à l'équation fonctionnelle

$$\widehat{\varphi}_0(2\omega) = m_0(\omega) \widehat{\varphi}_0(\omega)$$

sous la contrainte $\widehat{\varphi}_0(0) \neq 0$, c'est-à-dire, l'atome φ donné via son spectre (si toutefois le produit infini converge) par

$$\widehat{\varphi}_0(\omega) = \prod_{j=1}^{+\infty} m_0(2^{-j}\omega); \quad (2.3)$$

rappelons que le spectre d'un signal φ de $L^2(\mathbb{R}, dt)$ est donné comme le signal $\widehat{\varphi}$ d'énergie finie défini dans l'espace des fréquences par

$$\widehat{\varphi}(\omega) := \lim_{\substack{L^2(\mathbb{R}, dt) \\ N \rightarrow \infty}} \int_{-N}^N \varphi(t) e^{-i\omega t} dt.$$

C'est d'ailleurs pour satisfaire aux exigences qu'implique la convergence de ce produit infini que nous avons dû imposer que la contrainte additionnelle (2.1) soit satisfaite pour un certain $\epsilon > 0$. Le signal φ_0 est tel que la collection $(\varphi_0(\cdot - n))_n$ forme un système orthonormé de $L^2(\mathbb{R})$; en effet, pareille condition implique que la fonction

$$\Psi_0 : \omega \mapsto \sum_{n \in \mathbb{Z}} |\widehat{\varphi}_0(\omega + 2n\pi)|^2$$

vérifie

$$\Psi_0(2\omega) = (|m_0(\omega)|^2 + |m_0(\omega + \pi)|^2) \Psi_0(\omega) = \Psi_0(\omega), \quad \omega \in \mathbb{R},$$

et, par itération

$$\Psi_0(\omega) \equiv \Psi_0(0) = \widehat{\varphi}_0(0) = 1.$$

Ce n'est cependant pas nécessairement cet atome φ_0 que nous aurons en tête dans l'énoncé du lemme de décomposition que voici :

Lemme 2.1 *Soit φ un signal d'énergie finie sur \mathbb{R} tel que le système $(\varphi(\cdot - n))_{n \in \mathbb{Z}}$ soit un système orthonormé (par exemple $\varphi = \varphi_0$ associé à m_0 comme proposé ci-dessus pourvu que le produit infini converge). Soient ϕ et ψ les deux signaux déduits de φ par les relations*

$$\begin{aligned} \phi &:= \sum_{n \in \mathbb{Z}} h(n) \varphi(\cdot - n) \\ \theta &:= \sum_{n \in \mathbb{Z}} g(n) \varphi(\cdot - n), \end{aligned}$$

et V_0, V_1, W_1 les trois sous-espaces fermés de $L^2(\mathbb{R})$ engendrés respectivement par les systèmes $(\varphi(\cdot - n))_{n \in \mathbb{Z}}$, $(\phi(\cdot - 2n))_{n \in \mathbb{Z}}$, et $(\theta(\cdot - 2n))_{n \in \mathbb{Z}}$. Alors, on a la décomposition orthogonale de V_0

$$V_0 = V_1 \overset{\perp}{\oplus} W_1$$

et de plus, les deux systèmes $(\phi(\cdot - 2n))_{n \in \mathbb{Z}}$ et $(\theta(\cdot - 2n))_{n \in \mathbb{Z}}$ sont des bases hilbertiennes respectivement de V_1 et de W_1 .

Commentaire. La suite des coordonnées dans la base orthonormée $(\varphi(t - n))_{n \in \mathbb{Z}}$ de V_0 du signal $\text{pr}_{V_1}[s^{(0)}]$, où

$$s^{(0)} := \sum_{n \in \mathbb{Z}} s_n \varphi(t - n)$$

correspond à la “version brouillée” du signal dans l'algorithme pyramidal ; la suite des coordonnées dans cette même base de la projection sur W_1 du même signal $s^{(0)}$ correspond au signal “détails à l'échelle 1” du signal digital $(s_n)_n$; l'idée nouvelle (par rapport à l'algorithme pyramidal) est que l'on peut continuer à traiter avec ce lemme de décomposition les deux signaux digitaux de $l^2(2\mathbb{Z})$ que sont

- d’une part, la suite des coefficients de $\text{pr}_{V_1}[s^{(0)}]$, élément de V_1 , dans la base orthonormée (de V_1 cette fois) que constitue le système $(\phi(\cdot - 2n))_{2n \in 2\mathbb{Z}}$;
- d’autre part, la suite des coefficients de $\text{pr}_{W_1}[s^{(0)}]$, élément de W_1 , dans la base orthonormée (de W_1 cette fois) que constitue le système $(\theta(\cdot - 2n))_{2n \in 2\mathbb{Z}}$.

Dans le cas de l’algorithme pyramidal, les détails n’étaient plus traités ; seul le “résumé” (projection sur V_1) se trouvait retravaillé éventuellement.

Preuve. La preuve de ce lemme est un calcul, où le caractère unitaire de la matrice

$$\begin{pmatrix} m_0(\omega) & m_1(\omega) \\ m_0(\omega + \pi) & m_1(\omega + \pi) \end{pmatrix},$$

ainsi que le fait que les translatés de φ forment un système orthonormé, jouent un rôle fondamental. Ce calcul fait apparaître les résultats suivants :

- la projection orthogonale sur V_1 du signal $s^{(0)}$ associé au signal digital $(s_n)_n$ s’écrit

$$\sum_{n \in \mathbb{Z}} \left(\sum_{\nu \in \mathbb{Z}} s_\nu \overline{h(\nu - 2n)} \right) \phi(t - 2n);$$

- la projection orthogonale sur W_1 du signal $s^{(0)}$ associé au signal digital $(s_n)_n$ s’écrit

$$\sum_{n \in \mathbb{Z}} \left(\sum_{\nu \in \mathbb{Z}} s_\nu \overline{g(\nu - 2n)} \right) \theta(t - 2n);$$

Si l’on note R_0 et D_0 les opérateurs (considérés de $l^2(\mathbb{Z})$ dans lui-même)

$$\begin{aligned} R_0 & : (s_n)_n \mapsto \left(\sum_{\nu \in \mathbb{Z}} s_\nu \overline{h(\nu - 2n)} \right)_{n \in \mathbb{Z}} \\ D_0 & : (s_n)_n \mapsto \left(\sum_{\nu \in \mathbb{Z}} s_\nu \overline{g(\nu - 2n)} \right)_{n \in \mathbb{Z}} \end{aligned}$$

(il serait plus judicieux de les penser de $l^2(\mathbb{Z})$ dans $l^2(2\mathbb{Z})$), alors les coordonnées dans la base $(\varphi(t - n))_n$ cette fois de la projection orthogonale de $s^{(0)}$ sur V_1 sont données par la suite $R_0^* R_0[(s_n)_n]$ tandis que celles de la projection orthogonale de $s^{(0)}$ sur W_1 dans la même base sont données par la suite $D_0^* D_0[(s_n)_n]$; on a donc la pseudo-identité de Bézout

$$\text{Id}_{l^2(\mathbb{Z})} = R_0^* R_0 + D_0^* D_0$$

que nous retrouverons plus loin dans l’analyse temps-échelles dyadique.

2.3 L’algorithme “Split and Merge” de V. Wickerhauser

2.3.1 Le cadre des signaux digitaux

Soit s un signal digital de longueur 2^N (que l’on suppose, pour tous les calculs qui suivent, prolongé par 0 hors de l’ensemble $\{0, 1, \dots, 2^N - 1\}$).

s_0	s_1	s_2	s_3	s_4	s_5	s_6	s_7
r_0	r_1	r_2	r_3	d_0	d_1	d_2	d_3
rr_0	rr_1	dr_0	dr_1	rd_0	rd_1	dd_0	dd_1
rrr_0	drr_0	$rdrr_0$	ddr_0	rrd_0	$drrd_0$	rdd_0	ddd_0

FIG. 2.1 – Décomposition d'une information 1D

On se donne en “filigramme” une fonction φ d'énergie finie sur \mathbb{R} et dont les translatées forment un système orthonormé de $L^2(\mathbb{R}, dt)$. Le signal digital s est alors pensé comme le signal analogique (et non plus digital) S

$$S : t \rightarrow \sum_{n \in \mathbb{Z}} s_n \varphi(t - n),$$

signal que l'on décompose suivant ses projections orthogonales $\text{pr}_{V_1}[S]$ et $\text{pr}_{W_1}[S]$ sur les deux sous-espaces (orthogonaux) V_1 et W_1 respectivement engendrés par les $(\phi(t-2n))_{n \in \mathbb{Z}}$ et $(\theta(t-2n))_{n \in \mathbb{Z}}$ du lemme de décomposition (les deux suites $(h(n))_{n \in \mathbb{Z}}$ et $(g(n))_{n \in \mathbb{Z}}$ ayant été choisies de manière à ce que les polynômes trigonométriques m_0 et m_1 correspondants répondent aux exigences du lemme).

Le signal digital s (de longueur 2^N et prolongé par 0) donne ainsi naissance à deux signaux digitaux de longueur 2^{N-1} , les signaux $\{r_0, \dots, r_{2^{N-1}-1}\}$ et $\{d_0, \dots, d_{2^{N-1}-1}\}$, correspondant aux coefficients dans les bases orthonormées respectives $(\phi(t-2n))_{n \in \mathbb{Z}}$ et $(\theta(t-2n))_{n \in \mathbb{Z}}$ des projections $\text{pr}_{V_1}[S]$ et $\text{pr}_{W_1}[S]$; la concaténation de ces deux listes fournit la liste des coordonnées de S dans un autre système orthonormé (celui obtenu en concaténant les deux systèmes $(\phi(t-2n))_{n \in \mathbb{Z}}$ et $(\theta(t-2n))_{n \in \mathbb{Z}}$, au lieu de considérer le système $(\varphi(t-n))_{n \in \mathbb{Z}}$). Ces deux listes représentent, l'une une version “résumé” de l'information (c'est la suite $\{r_0, \dots, r_{2^{N-1}-1}\}$ des coefficients de la projection de S sur V_1), l'autre une version décimée des “détails” de l'information (c'est la suite $\{d_0, \dots, d_{2^{N-1}-1}\}$ des coefficients de la projection de S sur W_1). Chacune de ces deux suites (cette fois de longueur 2^{N-1}) peut être retraitée suivant le même mécanisme et l'on peut visualiser sur le tableau ci-dessous ($N = 3$) figurant sur la figure 2.1 les résultats des calculs.

À chaque niveau de la décomposition, on peut comparer l'entropie de l'information figurant dans une case donnée C (on prend l'entropie de Shannon, mais on pourrait, on l'a vu, envisager d'autres critères utilisant une notion d'entropie différente) à la somme des entropies des informations contenues dans les deux cases situées sous la case C (ce nouveau calcul correspond au calcul de l'entropie de l'information figurant dans la case C , mais calculée cette fois dans une autre base orthonormée qui privilégie le scindage résumé/détails). S'il s'avère que cette somme d'entropies est plus petite que l'entropie calculée directement dans la case C , alors le choix de cette nouvelle base (et donc de cette nouvelle décomposition) se justifie pour une lecture plus adaptée de l'information. Le fait que l'entropie se plie au critère de Shannon soutend la construction (proposée par Victor Wickerhauser il y a une dizaine d'années) d'un algorithme (dit “*split and merge*”) permettant d'accéder à une base orthonormée

s_0	s_1	s_2	s_3	s_4	s_5	s_6	s_7
r_0	r_1	r_2	r_3	$d_0 \quad d_1 \quad d_2 \quad d_3$			
rr_0	rr_1	$dr_0 \quad dr_1$		rd_0	rd_1	dd_0	dd_1
rrr_0	drr_0	rdr_0	ddr_0	rrd_0	$drrd_0$	rdd_0	ddd_0

FIG. 2.2 – Un premier exemple de base optimale

s_0	s_1	s_2	s_3	s_4	s_5	s_6	s_7
$r_0 \quad r_1 \quad r_2 \quad r_3$				d_0	d_1	d_2	d_3
rr_0	rr_1	dr_0	dr_1	rd_0	rd_1	$dd_0 \quad dd_1$	
rrr_0	drr_0	rdr_0	ddr_0	rrd_0	$drrd_0$	rdd_0	ddd_0

FIG. 2.3 – Un second exemple de base optimale

dans laquelle la décomposition de l'information digitale s soit optimale au niveau du critère de minimisation de l'entropie. Nous avons sur les figures 2.2 et 2.3 proposé des modèles de décomposition dans de telles bases (les cases colorées indiquent les cases choisies par l'arbre de sélection de l'algorithme). L'information contenue dans la décomposition optimale est mieux organisée que l'information s ; la compression, voire l'ajout d'une signature (comme dans les techniques de *watermarking*) se trouve mieux “guidée” face à une telle décomposition que face à la décomposition originelle.

L'action des filtres correspondant aux suites $(h_n)_{n \in \mathbb{Z}}$ et $(g_n)_{n \in \mathbb{Z}}$ correspond, l'une à l'action d'un filtre “moyennisateur”, d'autre d'un filtre “dérivateur” ; si l'information continue (correspondant à l'information s) est le signal analogique

$$S : t \rightarrow \sum_{n \in \mathbb{Z}} s_n \varphi(t - n),$$

l'information analogique correspondant à la suite $([aa \cdots aa]_n)_n$ figurant dans la case d'étiquette $aa \cdots aa$ (où la lettre **a** signifie **r** ou **d** suivant que l'on est descendu à gauche ou à droite dans l'arbre de sélection conduisant à cette case) est en fait le signal analogique

$$S_{aa \cdots aa} := \sum_{n \in \mathbb{Z}} [aa \cdots aa]_n \times ([aa \cdots aa](\varphi))(t - kn),$$

où k désigne la longueur du mot correspondant à la case et $[aa \cdots aa]$ désigne le signal analogique obtenu à partir de φ à partir des actions successives des filtres intégrateur et dérivateur (suivant le message codé dans le mot correspondant à l'étiquette de la case), combinées chaque fois avec une décimation de l'information.

L'atome $[aa \cdots aa](\varphi)$ apparaît alors comme un train d'ondes (on dit aussi un "paquet d'ondes") et l'étiquette de la case permet de repérer la gamme de fréquences impliquée dans l'analyse de Fourier du signal analogique $S_{aa \cdots aa}$. En ce sens, la recherche de la base optimale (au niveau du critère de minimisation de l'entropie) correspond à une exploration simultanée du signal analogique correspondant à s à la fois en temps, en échelles, et en fréquences. L'indicateur du paramètre d'échelle d'une case est le niveau dyadique où elle se trouve, l'indicateur du paramètre de fréquence étant son étiquette. Des lemmes de décomposition non plus dyadiques, mais p -adiques (comme on en verra dans la section 2.4) joueront aussi le rôle de décomposition de l'information suivant un banc de filtres (en même temps que dans une gamme p -adique d'échelles).

2.3.2 Le cas des images digitales

Le principe de la décomposition des images (et de la recherche d'une base optimale) est le même, excepté que le scindage de l'image n'est plus un scindage en deux, mais un scindage en quatre ("résumé" dans les deux directions, résumé "en horizontal" et détails "en vertical", résumé "en vertical" et détails "en horizontal", "détails" dans les deux directions, comme sur la figure 1.13). Le processus de décision se réduit à décider si oui ou non, on scinde une image en quatre ou si on la garde telle quelle. Au bout du compte, on détermine une base optimale, dont voici un modèle sur la figure 2.4 (pour une image 8×8).

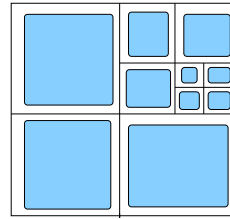


FIG. 2.4 – Un exemple de base optimale pour la décomposition d'une image

2.3.3 Implémentation algorithmique en 1D

Nous proposons de la faire dans le cas dyadique à partir soit de la fonction $m_0^{(N)}$, $N = 2$ (c'est dans ce cas la routine **wpack4** que nous utilisons), ou $N = 4$ (avec la routine **wpack8**). Voici par exemple le synopsis (sous **MATLAB**) de la routine **wpack8** :

```
function [U,etiq]=wpack8(s,nlev);
%      [U,etiq]=wpack8(s,nlev);
%decomposition en paquets d'ondelettes
%(analyse multi-resolution de Daubechies
% a 8 parametres); nlev est le nombre de
%niveaux d'echelle; les paquets sont les
%lignes de U; les etiquettes correspondantes
%%sont les lignes de etiq (R pour resume, D pour
%detail,0 pour l'operateur identite). L'entropie
%est l'entropie de Shannon non normalisee.
```

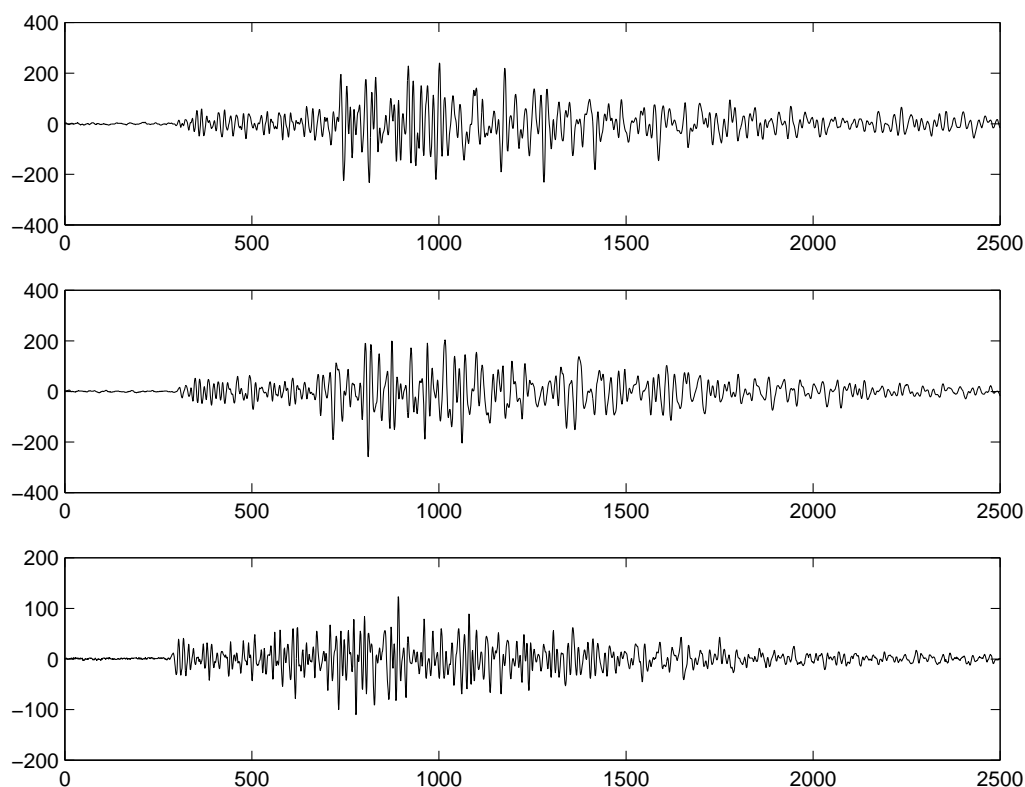


FIG. 2.5 – Un signal de secousse sismique (est-ouest, nord-sud, vertical)

La routine **viewpack** permet dans un second temps de visualiser les divers composants obtenus.

```
function viewpack(U,etiq,m);
% viewpack(U,etiq,m)
% permet la visualisation de la decomposition en
% paquets d'ondelettes (analyse multiresolution
% de Daubechies a 4 ou 8 parametres); on prend
% [U,etiq]=wpackM(s) avec M=4 ou 8; m est le nombre
% de composants par figure.
```

Les signaux obtenus après enregistrement d'une secousse sismique sont de bons modèles de signaux au sens où précisément interviennent les trois paramètres que sont le temps, l'échelle, et le nombre d'onde. Ce nombre d'onde joue un rôle dans les processus de réflexion ou de transmission des ondes sismiques en milieu non homogène et il est important, si l'on désire avoir des informations sur le milieu de propagation à partir de l'enregistrement de la secousse, de tenter de décomposer le phénomène en tenant compte de ces trois paramètres (c'est le principe du SONAR ou, en optique médicale, de l'échographie).

Voici dans un premier temps un signal correspondant à l'enregistrement d'une secousse sismique, de fait à trois enregistrements, dans les trois directions est-ouest, nord-sud, et suivant l'axe vertical. Ce signal a été décomposé en composants orthogonaux suivant un algorithme consistant (dans le sens descendant) à privilégier la minoration de l'entropie (non normalisée). Nous avons indiqué la décomposition

des signaux correspondant aux enregistrements est-ouest, nord-sud et vertical, selon quatre niveaux d'échelle, avec ici $m_0 = m_0^{(4)}$.

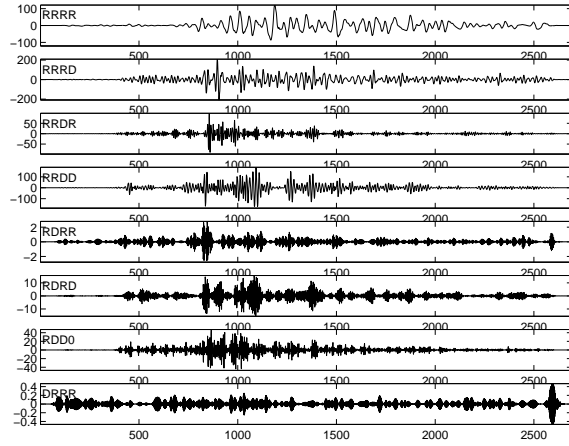


FIG. 2.6 – Analyse est-ouest

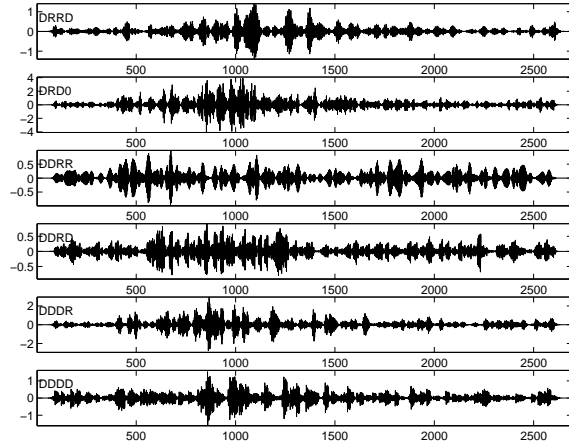


FIG. 2.7 – Analyse est-ouest (suite)

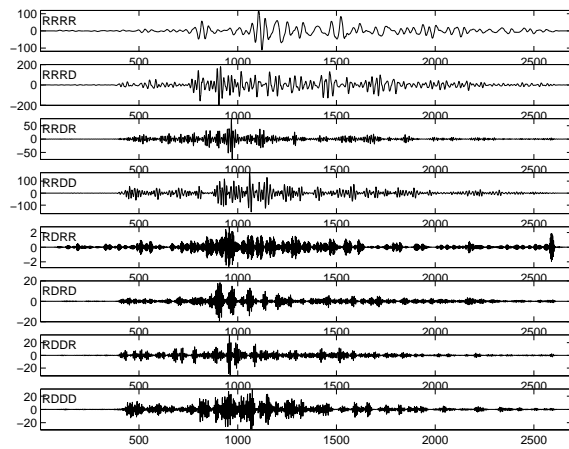


FIG. 2.8 – Analyse nord-sud

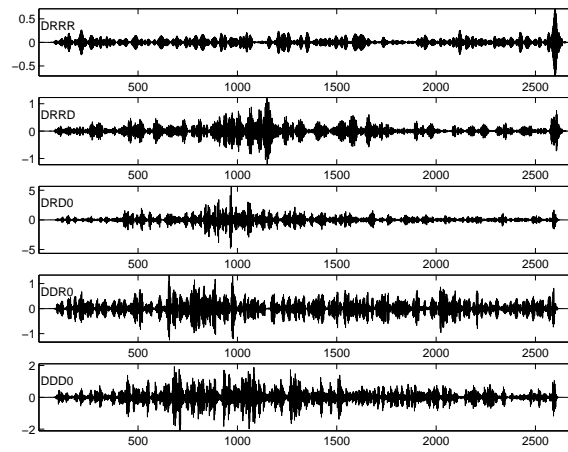


FIG. 2.9 – Analyse nord-sud (suite)

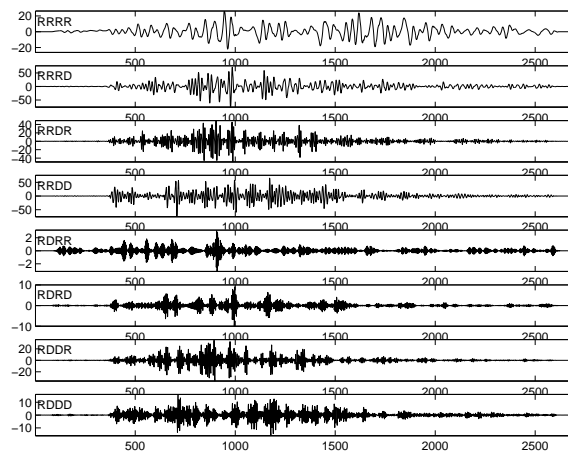


FIG. 2.10 – Analyse enregistrement vertical

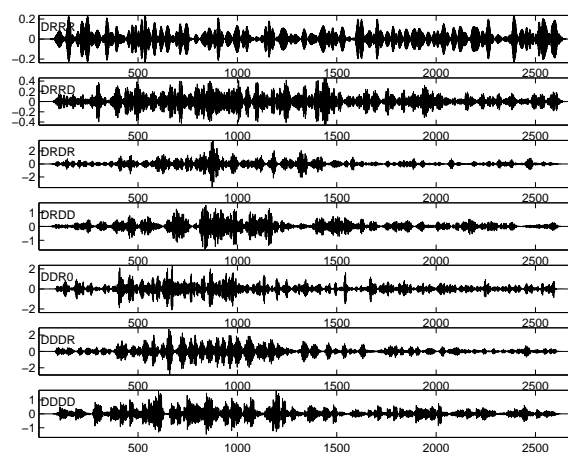


FIG. 2.11 – Analyse enregistrement vertical (suite)

Pour donner un exemple de la manière dont on peut exploiter ces diverses analyses, on s'appuiera sur le fait suivant : il est intéressant de savoir discerner dans une onde

sismique les instants d'arrivée des ondes à propagation horizontale (ondes P), ceux d'arrivée des ondes à propagation verticale (ondes S), ainsi qu'éventuellement les arrivées successives retardées par les effets de diffusion ou de réflexion dans le milieu inhomogène où s'effectue le transit. Pour cela, on peut par exemple construire la fonction de polarisation de l'onde, consistant à calculer, pour chaque niveau de décomposition (disons j) commun aux décompositions des trois composantes envisagées (les signaux de la décomposition correspondants étant ew_j, ns_j, z_j) la matrice de corrélation glissante C_j (pour la notion de corrélation glissante, on se référera au chapitre 1) des variables (ew_j, ns_j, z_j) , puis, pour chaque position temporelle t , la quantité

$$1 - \frac{\lambda_{j2}(t)}{\lambda_{j1}(t)},$$

où $\lambda_{j1}(t) \geq \lambda_{j2}(t)$ sont les deux plus grandes valeurs propres de cette matrice de corrélation $C_j(t)$; la fonction de polarisation

$$F(t) := \prod_j \left(1 - \frac{\lambda_{j2}(t)}{\lambda_{j1}(t)}\right)$$

rend compte de la localisation des instants d'arrivée des P -ondes. Voici le résultat concernant par exemple le cas de notre signal. Sur la figure suivante, nous avons indiqué le graphe de cette fonction de polarisation (en haut) et reproduit le graphe de l'enregistrement est-ouest de la secousse.

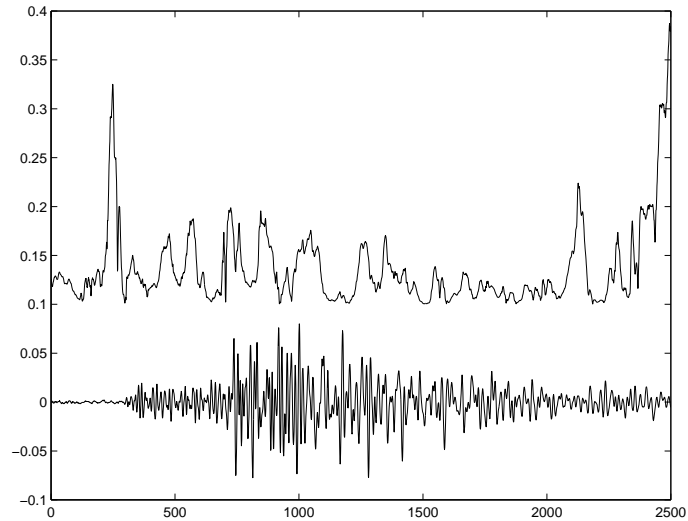


FIG. 2.12 – Polarisation

Une fois l'azimut déterminé (via le calcul du vecteur propre associé à la plus grande valeur propre aux instants supposés d'arrivée de la P -onde), on choisit le bon repère pour analyser à nouveau le problème afin de détecter les instants d'arrivée des ondes de type S . On voit en tout cas sur cet exemple combien ce procédé simple construit sur le modèle de l'analyse de la vision peut s'avérer utile pour envisager l'analyse de signaux où se superposent des informations relatives au temps, à l'échelle où l'on se place, ou au nombre d'onde (ou encore fréquence). C'est volontairement que nous avons choisi ici d'illustrer notre présentation de l'algorithme “*Split and Merge*”

grâce à des signaux empruntés à la sismique. Signalons toutefois que le choix de la fonction m_0 n'est pas anodin, si l'on pense que lui est couplé le choix d'une fonction φ particulière, en l'occurrence φ_0 ; la fonction φ_0 et surtout la fonction ψ_0 donnée par

$$\widehat{\psi}(\omega) = m_1(\omega/2) \prod_{j=2}^{+\infty} m_0(2^{-j}\omega)$$

peut s'avérer plus ou moins bien adapté à la "forme" des ondes présentes dans un signal correspondant à l'enregistrement d'une secousse sismique; le choix de $m_0^{(4)}$ utilisé dans les calculs ci-dessus n'est pas du tout le choix optimum de fonction m_0 pour de tels signaux; nous l'avons plus utilisé dans un souci didactique.

2.3.4 La généralisation au cadre p -adique

Si l'on a en vue les applications de ce type de lemme de décomposition à un cadre de nature plus arithmétique (ou si l'on souhaite plus simplement enrichir la finesse de la décomposition), il est intéressant de savoir que les idées introduites ci-dessus peuvent s'étendre au cadre non plus dyadique, mais p -adique, avec p nombre premier cette fois strictement supérieur à 2.

Supposons que p soit un tel nombre premier (d'ailleurs le fait que p soit premier ne joue pas de rôle ici) et que nous disposions de p fonctions m_0, \dots, m_{p-1} de $L^2(\mathbb{R}/2\pi\mathbb{Z})$ (c'est-à-dire des fonctions définies presque partout, 2π -périodiques, d'énergie finie sur $[0, 2\pi]$)

$$m_j = \frac{1}{\sqrt{p}} \sum_{n \in \mathbb{Z}} h_j(n) e^{-in(\cdot)}, \quad j = 0, \dots, p-1,$$

telles que, pour presque tout ω dans \mathbb{R} , la matrice

$$M = \left[m_j\left(\omega + \frac{2k\pi}{p}\right) \right]_{0 \leq j, k \leq p-1}$$

soit une matrice unitaire, qu'il existe $\epsilon > 0$ tel que

$$\sum |h_0(n)| |n|^\epsilon < \infty$$

et que $m_0(0) = 1$.

On peut comme précédemment associer à m_0 un signal φ_0 défini via l'équation fonctionnelle

$$\widehat{\varphi}_0(p\omega) = m_0(\omega) \widehat{\varphi}_0(\omega)$$

sous la contrainte $\widehat{\varphi}_0(0) \neq 0$, c'est-à-dire le signal défini via son spectre par

$$\widehat{\varphi}_0(\omega) := \prod_{j=1}^{+\infty} m_0(p^{-j}\omega)$$

(la convergence est assurée par les conditions sur la suite $(h_0(n))_n$). Ce signal φ_0 est tel que la collection $(\varphi(\cdot - n))_n$ forme encore un système orthonormé dans $L^2(\mathbb{R})$. En effet, pareille condition implique cette fois que la fonction

$$\Psi_0 : \omega \mapsto \sum_{n \in \mathbb{Z}} |\widehat{\varphi}_0(\omega + 2n\pi)|^2$$

vérifie

$$\Psi_0(p\omega) = \left(|m_0(\omega)|^2 + \cdots + \left| m_0\left(\omega + \frac{2(p-1)\pi}{p}\right) \right|^2 \right) \Psi_0(\omega) = \Psi_0(\omega), \quad \omega \in \mathbb{R},$$

et, par itération

$$\Psi_0(\omega) \equiv \Psi_0(0) = \widehat{\varphi}_0(0) = 1.$$

Nous avons alors la version suivante du *splitting lemma*, extension naturelle au cadre p -adique du lemme 2.1 :

Lemme 2.2 *Soit φ un signal d'énergie finie sur \mathbb{R} tel que le système $(\varphi(\cdot - n))_{n \in \mathbb{Z}}$ soit un système orthonormé (par exemple $\varphi = \varphi_0$ associé à m_0 comme proposé ci-dessus pourvu que le produit infini converge, mais ce n'est pas le seul choix possible). Soient ϕ_j , $j = 0, \dots, p-1$, les p signaux déduits de φ par les relations*

$$\begin{aligned} \phi : &= \sum_{n \in \mathbb{Z}} h_0(n) \varphi(\cdot - n) \\ \theta_j : &= \sum_{n \in \mathbb{Z}} h_j(n) \varphi(\cdot - n), \quad j = 1, \dots, p-1. \end{aligned}$$

et $V_0, V_1, W_{1,1}, \dots, W_{1,p-1}$ les $p+1$ sous-espaces fermés de $L^2(\mathbb{R})$ engendrés respectivement par les systèmes $(\varphi(\cdot - n))_{n \in \mathbb{Z}}$, $(\phi(\cdot - pn))_{n \in \mathbb{Z}}$, et $(\theta_j(\cdot - pn))_{n \in \mathbb{Z}}$, $j = 1, \dots, p-1$. Alors, on a la décomposition orthogonale de V_0

$$V_0 = V_1 \oplus \bigoplus_{1 \leq j \leq p-1}^{\perp} W_{1,j}$$

et de plus, les systèmes $(\phi(\cdot - pn))_{n \in \mathbb{Z}}$, $(\theta_j(\cdot - pn))_{n \in \mathbb{Z}}$, $j = 1, \dots, p-1$ sont des bases hilbertiennes respectivement de $V_0, W_{1,1}, \dots, W_{1,p-1}$.

Commentaire. La suite des coordonnées dans la base orthonormée $(\varphi(t - n))_{n \in \mathbb{Z}}$ de V_0 du signal $\text{pr}_{V_1}[s^{(0)}]$, où

$$s^{(0)} := \sum_{n \in \mathbb{Z}} s_n \varphi(t - n)$$

correspond à la “version brouillée” du signal dans une version de l'algorithme pyramidal où la décimation aurait été remplacée par une “ p -cimation” ; la suite des coordonnées dans cette même base de la projection sur un sous-espace $W_{1,j}$, $j = 0, \dots, p-1$, du même signal $s^{(0)}$ correspond à une certaine gamme de détails “détails à l'échelle 1 et dans la gamme j ” du signal digital $(s_n)_n$; comme dans le cas dyadique, on peut toujours continuer à traiter avec ce lemme de décomposition les p signaux digitaux de $l^2(p\mathbb{Z})$ que sont

- d'une part, la suite des coefficients de $\text{pr}_{V_1}[s^{(0)}]$, élément de V_1 , dans la base orthonormée (de V_1 cette fois) que constitue le système $(\phi(\cdot - pn))_{pn \in p\mathbb{Z}}$;
- d'autre part, pour chaque $j = 1, \dots, p-1$, la suite des coefficients de $\text{pr}_{W_{1,j}}[s^{(0)}]$, élément de $W_{1,j}$, dans la base orthonormée (de $W_{1,j}$ cette fois) que constitue le système $(\theta_j(\cdot - pn))_{pn \in p\mathbb{Z}}$.

Preuve. La preuve de ce lemme est analogue à celle du lemme de décomposition dyadique classique ; le fait que la matrice M et le fait que les translatés de φ forment un système orthonormé jouent un rôle fondamental. Ce calcul fait apparaître les résultats suivants :

- la projection orthogonale sur V_1 du signal $s^{(0)}$ associé au signal digital $(s_n)_n$ s'écrit

$$\sum_{n \in \mathbb{Z}} \left(\sum_{\nu \in \mathbb{Z}} s_\nu \overline{h_0(\nu - pn)} \right) \phi(t - pn);$$

- la projection orthogonale sur $W_{1,j}$, $j = 1, \dots, p-1$, du signal $s^{(0)}$ associé au signal digital $(s_n)_n$ s'écrit

$$\sum_{n \in \mathbb{Z}} \left(\sum_{\nu \in \mathbb{Z}} s_\nu \overline{h_j(\nu - pn)} \right) \theta_j(t - pn);$$

Si l'on note R_0 et $D_{0,j}$, $j = 1, \dots, p-1$, les opérateurs (considérés de $l^2(\mathbb{Z})$ dans lui-même)

$$\begin{aligned} R_0 & : (s_n)_n \mapsto \left(\sum_{\nu \in \mathbb{Z}} s_\nu \overline{h_0(\nu - pn)} \right)_{n \in \mathbb{Z}} \\ D_{0,j} & : (s_n)_n \mapsto \left(\sum_{\nu \in \mathbb{Z}} s_\nu \overline{h_j(\nu - pn)} \right)_{n \in \mathbb{Z}}, \quad j = 1, \dots, p-1 \end{aligned}$$

(il serait plus judicieux de les penser de $l^2(\mathbb{Z})$ dans $l^2(p\mathbb{Z})$), alors les coordonnées dans la base $(\varphi(t-n))_n$ cette fois de la projection orthogonale de $s^{(0)}$ sur V_1 sont données par la suite $R_0^* R_0[(s_n)_n]$ tandis que celles de la projection orthogonale de $s^{(0)}$ sur $W_{1,j}$, $j = 1, \dots, p-1$, dans la même base sont données par la suite $D_{0,j}^* D_{0,j}[(s_n)_n]$; on a donc la pseudo-identité de Bézout

$$\text{Id}_{l^2(\mathbb{Z})} = R_0^* R_0 + \sum_{j=1}^{p-1} D_{0,j}^* D_{0,j},$$

décomposition de l'identité plus affinée qu'elle ne l'était dans le cadre dyadique. De fait, l'indice j peut être (tout au moins dans une première approche) perçu comme l'indice d'un canal fréquentiel. Pareille décomposition est à mettre en parallèle avec la décomposition que matérialise le codage en sous-bande (on pourrait d'ailleurs retrouver exactement ce codage en sous bande modulo le choix d'une judicieuse famille m_0, \dots, m_{p-1}).

Un tel lemme de décomposition génère (comme dans le cas dyadique) un algorithme de sélection de base optimale au niveau de la minimization de l'entropie. L'étiquette d'une case (dans l'alphabet $R_0, D_{0,1}, \dots, D_{0,p-1}$) est en relation avec l'indicateur fréquentiel; en effet, les "paquets d'onde" déduits de l'atome φ et engendrant le sous-espace correspondant à la case correspondent à une fréquence correspondant (si rapportée à l'intervalle utile $[0, \pi]$) au nombre dont le développement en base p correspond à ce mot (R_0 correspondant à 0, $D_{0,j}$ à j , $j = 0, \dots, p-1$).

Un exemple de famille (m_0, \dots, m_{p-1}) (codage en sous-bandes) :

On peut prendre pour m_0 la fonction (2π -périodique) obtenue en périodisant la la fonction caractéristique de $[-\pi/p, \pi/p]$; on prend alors pour m_j ($j = 1, \dots, p-1$) la fonction 2π -périodique obtenue en périodisant la fonction

$$\chi_{[-(j+1)/p, -j/p]} + \chi_{[j/p, (j+1)/p]};$$

on construit ainsi un p -uplet de fonctions (m_0, \dots, m_p) utilisable dans le *splitting lemma* p -adique.

Un autre exemple de fonction m_0 :

Soit V_0 l'espace des fonctions affines par morceaux, avec noeuds toutes les unités, d'énergie finie, c'est-à-dire les fonctions du type :

$$\sum_{n \in \mathbb{Z}} s_n \Delta(t - n), \quad (s_n)_n \in l^2(\mathbb{Z}),$$

où Δ désigne la *fonction triangle*

$$\Delta(t) := \max(0, 1 - |t|), \quad t \in \mathbb{R};$$

notons que le signal continu

$$\sum_{n \in \mathbb{Z}} s_n \Delta(\cdot - n)$$

représente la “visualisation” naturelle du signal digital $(s_n)_n$ affichée sur l'écran de l'ordinateur. Cependant, Δ ne saurait être considéré comme un atome φ auquel les lemmes de décomposition s'appliquent, puisque Δ et $\Delta(t - 1)$ (ou Δ et $\Delta(\cdot + 1)$) ne sont clairement pas orthogonaux ; la fonction φ appartenant à V_0 et telle que ses translatés $(\varphi(\cdot - n))_n$ forment une base orthonormée de V_0 est la fonction dont le spectre est donné par

$$\hat{\varphi} = m\hat{\Delta},$$

où m est la fonction continue 2π -périodique

$$m : \quad \omega \in \mathbb{R} \mapsto \frac{1}{\sqrt{\sum_{n \in \mathbb{Z}} |\hat{\Delta}(\omega + 2n\pi)|^2}} ;$$

le calcul de m se fait grâce à la formule de Plancherel, puisque

$$\sum_{n \in \mathbb{Z}} |\hat{\Delta}(\omega + 2n\pi)|^2 = \int_0^1 |\Delta(t) + e^{i\omega} \Delta(t - 1)|^2 dt = \frac{2}{3} \left(1 + \frac{\cos \omega}{2}\right),$$

d'où

$$\hat{\varphi}(\omega) = \sqrt{\frac{3}{2 + \cos \omega}} \left(\frac{\sin(\omega/2)}{(\omega/2)} \right)^2.$$

Soit, pour un nombre premier p , $m_0^{[p]}$ la fonction 2π -périodique sur \mathbb{R} telle que

$$\forall \omega \in \mathbb{R}, \quad \hat{\varphi}(p\omega) = m_0^{[p]}(\omega) \hat{\varphi}(\omega);$$

il s'agit d'une fonction 2π -périodique réelle, facile à calculer, au moins pour les petits nombres premiers ; par exemple :

$$\begin{aligned} m_0^{[2]}(\omega) : &= \sqrt{\frac{2 + \cos \omega}{2 + \cos 2\omega}} (\cos(\omega/2))^2 \\ m_0^{[3]}(\omega) : &= \sqrt{\frac{2 + \cos \omega}{2 + \cos 3\omega}} \left(\frac{2 \cos(\omega) + 1}{3} \right)^2 \\ m_0^{[5]}(\omega) : &= \sqrt{\frac{2 + \cos \omega}{2 + \cos 5\omega}} \left(\frac{4(\cos(\omega))^2 \cos \omega + \cos 2\omega}{5} \right)^2 \end{aligned}$$

L'opérateur de projection orthogonale de l'espace V_0 sur le sous-espace $V_1^{[p]}$ des fonctions affines par morceaux avec noeuds toutes les p unités s'exprime (si l'on ramène les signaux de V_0 aux suites de leurs coefficients dans la base $(\varphi(t-n))_{n \in \mathbb{Z}}$) sous la forme :

$$\text{pr}_{V_1^{[p]}} \left[\sum_{n \in \mathbb{Z}} s_n \varphi(\cdot - n) \right] = \sum_{n \in \mathbb{Z}} \tilde{s}_n^{[p]} \varphi(t - n),$$

où

$$(\tilde{s}_n^{[p]})_n = R_0^{[p]*} R_0^{[p]} [(s_n)_n],$$

où

$$R_0^{[p]} [(s_n)_n] := \left(\sum_{n \in \mathbb{Z}} s_n h_0^{[p]}(\nu - pn) \right)_n \in l^2(p\mathbb{Z}) \simeq l^2(\mathbb{Z})$$

si

$$m_0^{[p]}(\omega) = \frac{1}{\sqrt{p}} \sum_{n \in \mathbb{Z}} h_0^{[p]}(n) e^{-in\omega}.$$

Un calcul simple montre que l'opérateur de $l^2(\mathbb{Z})$ dans lui-même donné comme $R_0^{[p]*} R_0^{[p]}$ se lit en continu comme l'opérateur $\mathcal{L}_0^{[p]}$ de V_0 dans lui-même

$$\mathcal{L}_0^{[p]}[s] = \sum_{j=0}^{p-1} L_{p,j} \left[e^{-\frac{2ij(\cdot)}{p}} s \right],$$

où $L_{p,j}$, $j = 0, \dots, p-1$, est le filtre digital de transformée de Fourier la fonction 2π -périodique

$$\omega \mapsto m_0^{[p]}(\omega) \overline{m_0^{[p]} \left(\omega + \frac{2j\pi}{p} \right)} = m_0^{[p]}(\omega) m_0^{[p]} \left(\omega + \frac{2j\pi}{p} \right).$$

Il s'agit donc de la somme des actions d'une cascade de filtres, chacun d'eux agissant sur une version déphasée du signal d'entrée s . On parle pour un tel opérateur de *pseudo-filtre*. Nous retrouverons ces pseudo-filtres $\mathcal{L}_0^{[p]}$ dans à propos de la décomposition dite de Franklin.

Remarque 2.1 Les mécanismes de filtrage (convolution discrète) + décimation (ou p -cimation) introduits dans ce chapitre pour guider la construction d'une base où l'entropie d'une information soit minimale pourraient être transposés du cadre analytique où ils sont introduits à un cadre arithmétique. Le principe du critère d'entropie est bien sûr alors à revoir, mais on peut imaginer un substitut arithmétique dans le cadre où les signaux digitaux sont maintenant pensés comme des signaux définis sur un corps fini et à valeurs dans ce corps. Pareille transposition ne semble pas avoir été faite; elle fournirait pourtant un éventail d'idées plus en relation avec le thème de ce cours, à savoir la cryptologie.

Chapitre 3

La lecture temps-échelles d'une information

3.1 L'analyse temps-échelles ; pourquoi faire ?

Dans cette section, nous allons reprendre avec plus de soin l'idée de traitement de l'information (inspirée du mécanisme de la vision) que nous avons exposé en exposant le principe de l'algorithme dit "pyramidal".

Traiter un signal digital ou une image, c'est savoir en même temps en dégager les structures cohérentes (dans nombre de situations, en particulier en régime stationnaire, ces structures sont les composantes moyennes ou basses-fréquences du signal ou de l'image, mais ce peut être aussi certaines structures motiviques apparaissant de manière apériodique) et les composantes hautes-fréquences, les lignes de contraste ou les singularités.

L'outil Fourier (utilisé comme outil d'analyse ou de traitement comme dans le principe des algorithmes de compression type **jpeg** basés sur l'utilisation bloc par bloc de la transformée en cosinus) exploite de manière fondamentale l'idée de base suivant laquelle tout signal ou toute image se présente comme un empilement de composants dits "harmoniques". D'un autre côté, les algorithmes de traitement de signaux ou d'images basés sur la recherche d'une meilleure base (au sens d'une base dans laquelle l'entropie de l'information soit minimale) permettent la recherche, à partir d'un signal ou d'une image digitale donnée, d'un "dictionnaire" d'atomes (qui ne soit pas nécessairement le dictionnaire des fonctions trigonométriques inhérent à l'analyse de Fourier, mais plutôt un dictionnaire de motifs extraits de l'information elle-même. On peut aussi se fixer comme objectif la mise en évidence des "accidents" du signal ou de l'image, ce qui relève d'une analyse (permettant également le traitement) différente, l'analyse dans la gamme des échelles.

Notons qu'un phénomène "haute fréquence", s'il relève de la catégorie des signaux à lire "à petite échelle" (du fait que les ondes qu'il renferme ont une petite amplitude), relève aussi, du fait de sa stationnarité (l'onde est répétée de manière périodique) de la classe des "grandes échelles" (la "périodicité" se voit seulement si l'on examine le signal de loin). On ne peut donc pas dire qu'une analyse temps-échelles corresponde à une analyse temps-fréquences. Ce sont deux points de vue *a priori* différents (et souvent complémentaires). D'ailleurs, le principe d'incertitude de Heisenberg, qui prive l'analyse de Fourier d'être une analyse locale, empêche la compatibilité de nos

deux exigences (un “accident” se lit bien en temps-échelles, beaucoup moins bien en “temps-fréquences”). Les modèles inspirés du *mécanisme de la vision* ainsi que du concept *d'être fractal* (sur lequel nous reviendrons) nous incitent à proposer pour l'analyse des signaux ou des images l'idée simple suivant laquelle l'analyse se ferait par le biais d'une loupe (que l'on s'octroierait la possibilité de retourner), de manière à pouvoir disposer de versions “grossies” du signal (de plus en plus riches du point de vue de la quantité d'information) ainsi que d'images “prises de plus en plus loin” (donc de moins en moins riches du point de vue quantité d'information).

3.2 La décomposition de Franklin d'un signal digital

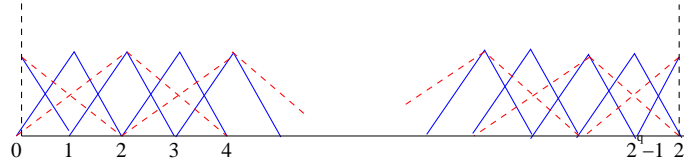


FIG. 3.1 – La base de Franklin

Soit N un entier de la forme 2^q , $q \in \mathbb{N}$. Si l'on considère un signal digital sur $\{0, \dots, N\}$ comme la version discrétisée (aux points $0, \dots, N$) d'un signal affine par morceaux (avec noeuds aux points entiers), on peut penser ce signal digital comme un élément de l'espace $V_{0,N}$ des signaux sur $[0, N]$ qu'engendrent les $N + 1$ fonctions spline sur $[0, N]$

$$t \mapsto \Delta_{0,n}(t) := \max(0, 1 - |t - n|), \quad n = 0, \dots, N.$$

On vérifiera que ce que l'on conviendra d'appeler la *matrice de Gram* de ce système, à savoir la matrice

$$G_{0,N} := [\langle \Delta_{0,n_1}, \Delta_{0,n_2} \rangle]_{0 \leq n_1, n_2 \leq N},$$

est ici la matrice bande

$$\begin{pmatrix} 1/3 & 1/6 & 0 & \dots & 0 & 0 \\ 1/6 & 2/3 & 1/6 & 0 & \dots & 0 \\ 0 & 1/6 & \cdot & \cdot & \dots & 0 \\ \vdots & \cdot & \cdot & \dots & \dots & \vdots \\ 0 & \dots & 0 & 1/6 & 2/3 & 1/6 \\ 0 & \dots & 0 & 0 & 1/6 & 1/3 \end{pmatrix}$$

Considérons le sous espace $V_{1,N}$ de $V_{0,N}$ engendré par les signaux affines par morceaux sur $[0, N]$, à noeuds tous les entiers pairs ; cet espace est de dimension $2^{q-1} + 1$ et une base de cet espace est donnée par les fonctions spline

$$t \mapsto \Delta_{1,n}(t) := \max(0, 1 - |t/2 - n|), \quad n = 0, \dots, 2^{q-1}.$$

Si $G_{0,N}$ désigne la matrice de Gram de ce nouveau système, la projection $R_0[s]$ du signal digital

$$s = \sum_{n=0}^N s_n \Delta_{0,n}$$

sur $V_{1,N}$ est donnée par

$$R_0[s] := \sum_{n=0}^{2^q-1} c_{1,n}(s) \Delta_{1,n},$$

où

$$G_{0,N} \begin{bmatrix} c_{1,0}(s) \\ \vdots \\ c_{1,2^q-1}(s) \end{bmatrix} = \begin{bmatrix} \sum_{k=0}^N s(k) \langle \Delta_{0,k}, \Delta_{1,0} \rangle \\ \vdots \\ \sum_{n=0}^N s_n \langle \Delta_{0,n}, \Delta_{1,2^q-1} \rangle \end{bmatrix}.$$

Nous pouvons retrouver une expression de $R_1[s]$ dans la base originelle de V_N , c'est à dire visualiser $R_1[s]$ comme un signal digital sur $\{0, \dots, N\}$, en utilisant les relations immédiates

$$\begin{aligned} \Delta_{1,0} &= \Delta_{0,0} + \frac{\Delta_{0,1}}{2} \\ \Delta_{1,n} &= \frac{\Delta_{0,n-1}}{2} + \Delta_{0,n} + \frac{\Delta_{0,n+1}}{2}, n = 1, \dots, 2^{q-1} - 1 \\ \Delta_{1,2^q-1} &= \Delta_{0,N} + \frac{\Delta_{0,N-1}}{2} \end{aligned}$$

On obtient ainsi une décomposition orthogonale

$$s = R_0[s] + (s - R_0[s]) = R_0[s] + D_0[s],$$

(“ R ” pour “résumé”, “ D ” pour “détails”). Le signal $R_0[s]$ peut être imaginé comme une version “floue” (ou disons “dégrossie”) du signal digital originel s , tandis que $D_0[s] := s - R_0[s]$ représente le signal correspondant aux détails que le dégrossissage a en quelque sorte gommé. L’algorithme que nous proposons requiert les deux sous-programmes **gram** et **frankaux** et s’exploite sous la forme

$$R_0[s] = \mathbf{franklin}(0, q, s),$$

lorsque s désigne un signal digital sur $\{1, \dots, 2^q + 1\}$.

```
>function f=franklin(1,q,s);
% Ce programme calcule la projection d'un signal digital s
% de taille (1,N=2^q+1), affine par morceaux avec noeuds
% aux points 1+k2^1, k=0,1,..., (1 entre 0 et q-1) sur le sous espace
% des signaux de taille (1,N), affines par morceaux avec
% noeuds aux points 1+k2^(1+1), k=0,1,...
% L'algorithme peut etre itere comme suit pour
% realiser une decomposition orthogonale de s en resume-details:
% On part de s, size(s)=(1,2^q+1)
% R1=franklin(0,q,s);
% D1=s-R1;
% R2=franklin(1,q,R1);
% D2=R1-R2;
% R3=franklin(2,q,R2);
%...
% Rq=franklin(q-1,q,R(q-1));
% les sous-programmes utilises sont gram et frankaux
```


L'algorithme peut s'itérer comme décrit dans le synopsis ci dessus et l'on peut obtenir une décomposition orthogonale

$$s = D_0[s] + D_1[s] + D_2[s] + \cdots D_k[s] + R_k[s]$$

lorsque k désigne un entier entre 1 et q . Le programme **plfr** nous permet de visualiser cette décomposition.

Nous avons par exemple affiché sur la figure 3.13 les divers niveaux de détails (et le résumé) du signal digital s de 2049 points affiché sur la figure 3.2.

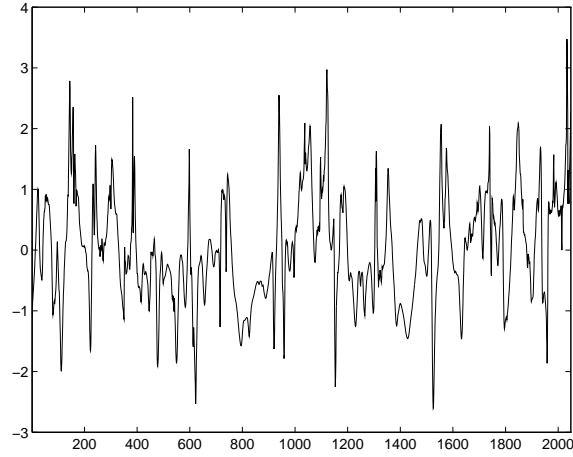


FIG. 3.2 – $s = s(k)$, $k = 0 : 1 : 2048$

La décomposition est affichée sur la figure 3.3 ; la somme des lignes redonne le signal digital s .

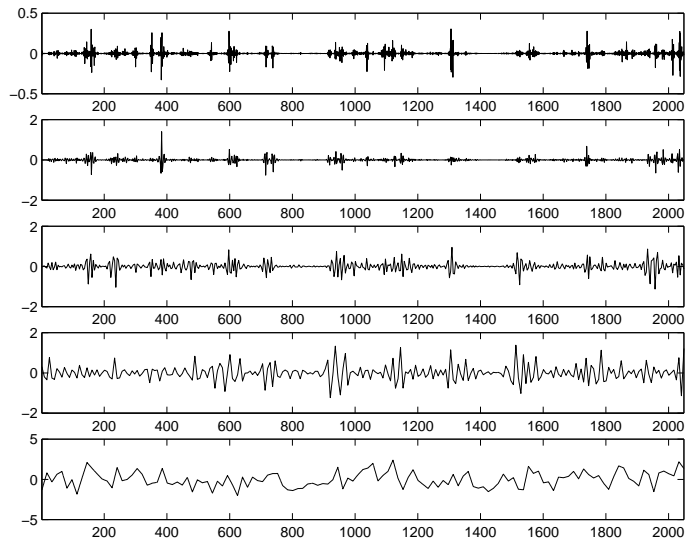


FIG. 3.3 – détails d_0, d_1, d_2, d_3 et résumé r_3 de la décomposition de s

3.3 L'instrument d'analyse : l'analyse multi-résolution dyadique

Cette promenade dans les échelles aux fins de l'analyse se trouve modélisée dans le cadre discret par la notion d' *analyse multi-résolution dyadique*, dont nous allons fixer le cadre.

Une analyse multi-résolution pour $L^2(\mathbb{R})$ est une suite $(V_j)_{j \in \mathbb{Z}}$ de sous-espaces fermés

$$\cdots \subset V_k \subset V_{k-1} \subset \cdots \subset V_1 \subset V_0 \subset V_{-1} \subset \cdots$$

emboîtés, d'intersection réduite à $\{0\}$ et d'union dense, dont la hiérarchie se trouve régie par la règle de promenade dans les échelles

$$s \in V_j \iff s(2(\cdot)) \in V_{j-1}, \quad j \in \mathbb{Z}.$$

On exige de plus qu'il existe une fonction père φ dans V_0 dont tous les translatés

$$t \mapsto \varphi_n(t) := \tau_n[\varphi](t) = \varphi(t - n), \quad n \in \mathbb{Z},$$

constituent une base orthonormée de V_0 (pour le produit scalaire usuel de $L^2(\mathbb{R})$) ; c'est là l'exigence cruciale, la fonction φ conditionnant toute la construction (même si φ , comme la fonction ψ qui lui est associée, on verra plus loin comment, n'apparaît du point de vue pratique qu'en filigramme du traitement correspondant de l'information digitale.

La promenade dans les échelles implique qu'aller vers la droite (vers les V_j lorsque j tend vers $-\infty$) correspond à une excursion vers le domaine des petites échelles (l'exploration “à la loupe” de l'information) tandis qu'aller vers la gauche (vers les V_j lorsque j tend vers $+\infty$) correspond à une excursion vers le domaine des grandes échelles (l'exploration “à la loupe retournée” de l'information).

Lorsque l'on traitera une information digitale suivant une analyse multi-résolution, celle-ci sera implicite, au sens où elle sera invisible dans la démarche de décomposition (mais de fait constamment présente). Changer d'analyse nous fera changer de décomposition.

L'information digitale $(s_n)_{n \in \mathbb{Z}}$ est en fait traitée comme l'élément de V_0 défini (presque partout puisque les objets sont dans $L^2(\mathbb{R})$) par

$$s(t) := \sum_{n \in \mathbb{Z}} s_n \varphi(t - n).$$

Pour que ceci soit cohérent du point de vue pratique, on est amené à exiger de la fonction φ qu'elle soit bien localisée autour du segment “de base” $[0, 1]$.

Voici trois exemples importants ; le dernier (analyse multi-résolution de Franklin) est proche de l'outil de décomposition orthogonale de l'information que nous avons introduit au chapitre 2 (décomposition de Franklin).

- L'analyse multi-résolution de Haar. C'est historiquement la plus ancienne (et aussi la plus simple) ; le sous-espace V_0 est le sous espace des signaux d'énergie finie, en escalier et à pas entiers, c'est à dire

$$V_0^{\text{haar}} := \left\{ \sum_{n \in \mathbb{Z}} a_n \tau_n[\chi_{[0,1]}]; \sum_{n \in \mathbb{Z}} |a_n|^2 < \infty \right\}.$$

Le père en est le signal $\chi_{[0,1]}$.

- L'analyse multi-résolution de Shannon. C'est le pendant de celle de Haar, mais transposée de l'espace des temps à l'espace des fréquences. On a en effet

$$V_0^{\text{shan}} := \{s \in L^2(\mathbf{R}); \hat{s} = 0 \text{ p.p. hors de } [-\pi, \pi]\}.$$

Le père en est le signal dont le spectre est la fonction caractéristique de $[-\pi, \pi]$, soit le signal

$$\varphi : t \mapsto \text{sinc}(\pi t) := \frac{\sin(\pi t)}{\pi t},$$

puisque

$$\lim_{L^2} \left[\int_{-N}^N \frac{\sin(\pi t)}{\pi t} e^{-i(\cdot)t} dt \right] = \chi_{[-\pi, +\pi]}.$$

- L'analyse multi-résolution de Franklin. Le sous-espace V_0 qui génère cette dernière analyse est constitué des signaux d'énergie finie, affines par morceaux, avec noeuds au points de \mathbf{Z} , soit encore

$$V_0^{\text{frank}} = \left\{ \sum_{n \in \mathbf{Z}} a_n \Delta(t - n); \sum_{l \in \mathbf{Z}} |a_n|^2 < \infty \right\},$$

où Δ est le signal défini par

$$\Delta(t) := \max(0, 1 - |t|).$$

Notons que si $(s_n)_{n \in \mathbf{Z}}$ est un signal digital, le signal continu

$$\sum_{n \in \mathbf{Z}} s_n \Delta(t - n)$$

est celui que l'on "visualise" sur l'écran de l'ordinateur (la machine interpole linéairement au niveau du rendu visuel). Malheureusement ici Δ n'est pas une fonction père, car il y a redondance entre Δ , $\Delta(\cdot - 1)$, $\Delta(\cdot + 1)$. Il faut travailler un peu plus pour trouver un père pour l'analyse multi-résolution générée par le sous-espace V_0^{frank} . Le spectre du signal Δ est donné par

$$\hat{\Delta}(\omega) = \left(\frac{\sin \frac{\omega}{2}}{\frac{\omega}{2}} \right)^2, \quad \omega \in \mathbf{R},$$

comme on le vérifie par un calcul immédiat. Comme la fonction continue périodique

$$\omega \mapsto \Psi(\omega) := \sum_{n \in \mathbf{Z}} |\hat{\Delta}(\omega + 2n\pi)|^2 = \sum_{n \in \mathbf{Z}} \left(\frac{\sin \frac{\omega}{2}}{\frac{\omega}{2} + n\pi} \right)^4$$

ne s'annule pas sur \mathbf{R} , elle est encadrée par deux constantes strictement positives c et C . Il est d'ailleurs possible de calculer aisément le père de l'analyse multi-résolution de Franklin; en effet, la formule de Plancherel nous donne

$$\Psi(\omega) = \int_0^1 |\Delta(t) + e^{i\omega} \Delta(t - 1)|^2 dt = \frac{2}{3} \left(1 + \frac{\cos \omega}{2} \right).$$

Par conséquent, le père φ de l'analyse multi-résolution de Franklin est donné via son spectre par

$$\hat{\varphi}(\omega) = \sqrt{\frac{3}{2 + \cos \omega}} \hat{\Delta}(\omega) = \sqrt{\frac{3}{2 + \cos \omega}} \left(\frac{\sin \frac{\omega}{2}}{\frac{\omega}{2}} \right)^2, \quad \omega \in \mathbf{R}.$$

On pourrait d'ailleurs construire sur ce modèle d'analyse-multirésolution un modèle plus général où V_0 serait l'espace des signaux d'énergie finie qui sont des m -splines ($m = 0, 1, \dots$). Ce sont les analyses multi-résolution de fonctions spline.

3.4 Analyse temps-échelles I. De l'information à son "résumé".

La donnée d'une analyse-multirésolution génère à chaque niveau d'échelle $j \in \mathbb{Z}$, une décomposition orthogonale de l'espace V_j sous la forme

$$V_j = V_{j+1} \overset{\perp}{\oplus} W_{j+1},$$

où W_{j+1} est par définition l'orthogonal du sous-espace fermé V_{j+1} (considéré comme sous-espace de V_j) dans V_j , et avec elle une paire d'opérateurs Rés_j , Dét_j (Rés pour "Résumé", Dét pour "Détails", l'indice nous rappelant le niveau d'échelle auquel on se place), où

$$\text{Rés}_j : V_j \mapsto V_{j+1}$$

est la projection orthogonale de V_j sur V_{j+1} et

$$\text{Dét}_j : V_j \mapsto W_{j+1}$$

est la projection orthogonale de V_j sur W_{j+1} .

Ce que l'on voit immédiatement est que, du fait que l'intersection des V_j se réduit au signal nul presque partout alors que l'union est, elle, dense, ce découpage en " tiroirs " de chaque V_j nous permet d'affirmer, suite à un argument basé sur l'aspect "télescopique" du découpage, que

$$L^2(\mathbb{R}) = \overline{\bigcup_{j=-\infty}^{\infty} (V_{j+1} \overset{\perp}{\oplus} W_{j+1})} = \bigoplus_{j \in \mathbb{Z}} W_j$$

Ainsi, tout signal de $L^2(\mathbb{R})$ se décompose sous la forme

$$s = \sum_{j \in \mathbb{Z}} \text{pr}_{W_j}[s]$$

("pr" pour "projection orthogonale"); en particulier notre signal continu

$$s = \sum_{n \in \mathbb{Z}} s_n \varphi(\cdot - n)$$

que nous avons de force décidé d'associer à notre signal digital $(s_n)_{n \in \mathbb{Z}}$ se décompose lui, en tant qu'élément de V_0 , sous la forme

$$s = \sum_{j=1}^N \text{pr}_{W_j}[s] + \text{pr}_{V_N}[s],$$

lorsque N est un entier supérieur ou égal à 1.

Étant donné un signal dans l'espace V_j , décrivons plus soigneusement ce qu'est l'action de cet opérateur de "prise de Résumé" qu'est Rés_j .

Remarquons d'abord que l'on dispose d'une base orthonormée pour "disséquer" l'information contenue dans un signal à l'échelle j ; en effet, le fait que $(\varphi((\cdot) - n))_{n \in \mathbb{Z}}$ constitue une base orthonormée de V_0 et que les V_j se déduisent les uns des autres par

jeu de la promenade dans les échelles implique que la famille des fonctions $(\varphi_{j,n})_{n \in \mathbb{Z}}$, où

$$\varphi_{j,n}(t) = 2^{-j/2} \varphi(2^{-j}t - n), \quad n \in \mathbb{Z},$$

est une base orthonormée de V_j ; notons que, si le père de l'analyse φ est à peu près localisé autour de $[0, 1]$, le signal $\varphi_{j,n}$ est, lui, localisé autour de $[n, n + 2^j]$, version dilatée (ou contractée) suivant que $j > 0$ ou $j < 0$, puis translatée, de $[0, 1]$.

Comme le père φ de l'analyse est dans V_0 , sa version dilatée $\varphi_{1,0}$ est dans V_1 , donc aussi dans V_0 ; on a donc

$$\varphi_{1,0} = \sum_{n \in \mathbb{Z}} h_n \varphi(t - n)$$

pour une certaine suite $(h_n)_n$ dans $l^2(\mathbb{Z})$; en prenant les transformées de Fourier, on trouve

$$\frac{1}{\sqrt{2}} \int_{\mathbb{R}} \varphi(t/2) e^{-i\omega t} dt = \sqrt{2} \int_{\mathbb{R}} \varphi(t) e^{-2i\omega t} dt = \sum_{n \in \mathbb{Z}} h_n e^{-in\omega} \int_{\mathbb{R}} \varphi(t) e^{-it\omega} dt,$$

soit

$$\widehat{\varphi}(2\omega) = m_0(\omega) \varphi(\omega),$$

où $m_0 \in L^2(\mathbb{R}/2\pi\mathbb{Z})$ (ce qui signifie qu'il s'agit d'une fonction définie presque partout sur \mathbb{R} , 2π -périodique, et d'énergie finie sur $[0, 2\pi]$) est définie par

$$m_0 = \frac{1}{\sqrt{2}} \sum_{n \in \mathbb{Z}} h(n) e^{-in(\cdot)};$$

le fait que le système $(\varphi(\cdot) - n)_{n \in \mathbb{Z}}$ soit orthonormé rejaillit en l'identité

$$\sum_{n \in \mathbb{Z}} |\widehat{\varphi}(\omega + 2n\pi)|^2 \equiv 1, \quad \omega \in \mathbb{R},$$

qui elle-même implique

$$|m_0(\omega)|^2 + |m_0(\omega + \pi)|^2 \equiv 1, \quad \omega \in \mathbb{R}$$

(on perçoit la rigidité des contraintes; pareille identité est impossible à réaliser avec une fonction 2π -périodique m_0 arbitraire).

Tout ceci étant posé, un petit calcul simple nous permet de vérifier la proposition suivante, permettant le calcul des coordonnées du résumé de s (dans la base $(\varphi_{j+1,n})_{n \in \mathbb{Z}}$) en fonction des coordonnées de s (dans la base $(\varphi_{j,n})_{n \in \mathbb{Z}}$) lorsque $s \in V_j$. Voici le résultat :

Proposition 3.1 *L'image du signal*

$$s = \sum_{n \in \mathbb{Z}} s_n^{(j)} \varphi_{j,n}$$

de V_j par l'opérateur de projection orthogonale Rés_j est le signal

$$\text{Rés}_j[s] = \sum_{n \in \mathbb{Z}} s_n^{(j+1)} \varphi_{j+1,n},$$

où

$$s_n^{(j+1)} = \sum_{\nu \in \mathbb{Z}} s_\nu^{(j)} \overline{h(\nu - 2n)}, \quad n \in \mathbb{Z}. \quad (\dagger)$$

Remarque 3.1 Avant de faire le petit calcul, une remarque s'impose; si s est le signal de V_0 correspondant à notre signal digital $(s_n)_{n \in \mathbb{Z}}$ de référence, la suite $s_n^{(0)} = s_n$ se lit à l'échelle initiale (ici $j = 0$) où se trouve donnée l'information; en revanche, si l'on envisage au niveau du signal $\text{Rés}[s] \in V_1$ le retour au "digital", il ne faut pas perdre de vue que ce nouveau signal est à lire à l'échelle 1 (il y a eu décimation de l'information); ceci se reflète d'ailleurs dans la présence du $2n$ au membre de droite des formules donnant les coefficients

$$s_n^{(1)} = \sum_{\nu \in \mathbb{Z}} s_n \overline{h(\nu - 2n)}, \quad n \in \mathbb{Z}.$$

On retrouve la le même phénomène que dans l'algorithme pyramidal : si l'on part d'un signal de 2048 points, le signal digital induit par la suite des coefficients de la projection sur V_1 est un signal digital de 1024 points; si l'on veut le représenter sur la même figure que le signal digital $(s_n)_n$ il faut en dilater par deux (dans le sens des abscisses) son graphe.

Preuve de la proposition. Elle se résume à la linéarité de l'opérateur Rés_j et au fait que

$$\begin{aligned} \text{Rés}_j[\varphi_{j,n}] &= \sum_{\nu \in \mathbb{Z}} \langle \varphi_{j,n}, \varphi_{j+1,\nu} \rangle \varphi_{j+1,\nu} \\ &= \frac{1}{2\pi} \sum_{\nu \in \mathbb{Z}} \left(\int_{\mathbb{R}} \widehat{\varphi_{j,n}}(\omega) \overline{\widehat{\varphi_{j+1,\nu}}(\omega)} d\omega \right) \varphi_{j+1,\nu} \\ &= \frac{1}{\sqrt{2\pi}} \sum_{\nu \in \mathbb{Z}} \left(\int_{\mathbb{R}} \widehat{\varphi}(\omega) \overline{\widehat{\varphi}(2\omega)} e^{-i(l-2\nu)\omega} d\omega \right) \varphi_{j+1,\nu} \\ &= \frac{1}{\sqrt{2\pi}} \sum_{\nu \in \mathbb{Z}} \left(\int_{\mathbb{R}} |\widehat{\varphi}(\omega)|^2 m_0(\omega) e^{-i(n-2\nu)\omega} d\omega \right) \varphi_{j+1,\nu} \\ &= \frac{1}{\sqrt{2\pi}} \sum_{\nu \in \mathbb{Z}} \left(\int_0^{2\pi} \overline{m_0(\omega)} e^{-i(n-2\nu)\omega} d\omega \right) \varphi_{j+1,\nu} \\ &= \sum_{\nu \in \mathbb{Z}} \overline{h(n-2\nu)} \varphi_{j+1,\nu}. \end{aligned}$$

Les formules (†) en résultent. \diamond

Comme il est impossible que $\widehat{\varphi}(0) = 0$, on a $m_0(0) = 1$ et par conséquent

$$\sum_{l \in \mathbb{Z}} h(l) = \sqrt{2},$$

ce qui nous autorise à considérer le filtre de réponse impulsionnelle $(h(-n))_{n \in \mathbb{Z}}$ comme un *filtre passe-bas*; l'action de l'opérateur Rés_j se traduit par la composition de l'action de ce filtre passe-bas avec une opération de décimation, et nous retrouvons donc ici le mécanisme qui soutend la phase "calcul des résumés" dans l'*algorithme pyramidal*. La phase "redistribution du résumé $\text{Rés}_j[s]$ dans l'échelle originelle j " correspond, quant à elle, à l'expression de

$$\text{Rés}_j[s] = \sum_{n \in \mathbb{Z}} s_n^{(j+1)} \varphi_{j+1,n}$$

(élément de $V_{j+1} \subset V_j$) dans la base orthonormée $(\varphi_{j,n})_{n \in \mathbb{Z}}$, soit encore

$$\text{Rés}_j[s] = \sum_{n \in \mathbb{Z}} \widetilde{s}_n^{(j)} \varphi_{j,n},$$

où

$$\tilde{s}_n^{(j)} = \sum_{\nu_1 \in \mathbb{Z}} \sum_{\nu_2 \in \mathbb{Z}} s_{\nu_2}^{(j)} \overline{h(\nu_2 - 2\nu_1)} h(n - 2\nu_1), \quad n \in \mathbb{Z}. \quad (\dagger\dagger)$$

Si les formules (\dagger) s'écrivent de manière condensée

$$(s_n^{(j+1)})_n = R_j[(s_n^{(j)})_n]$$

(R_j correspondant au niveau discret à une opération filtrage+décimation), les formules $(\dagger\dagger)$ s'écrivent, quant à elles,

$$(\tilde{s}_n^{(j)})_l = R_j^* R_j[(s_n^{(j)})_n].$$

L'opérateur $R_j^* R_j$, considéré comme un opérateur sur l'espace des signaux discrets d'énergie finie, n'est pas un filtre (car il n'a pas la propriété d'invariance par le shift). Néanmoins, de part son rôle consistant à isoler de la suite discrète correspondant à un signal de V_j les structures cohérentes organisées selon l'échelle "double" $j + 1$, on qualifie un tel opérateur de *pseudo-filtre passe-bas*.

3.5 La mère de l'analyse multi-résolution

Si l'on définit la fonction m_1 à partir de m_0 comme

$$m_1(\omega) := e^{-i(\omega+\pi)} \overline{m_0(\omega + \pi)},$$

alors la matrice

$$\begin{pmatrix} m_0(\omega) & m_1(\omega) \\ m_0(\omega + \pi) & m_1(\omega + \pi) \end{pmatrix}$$

est, pour presque tout $\omega \in \mathbb{R}$, une matrice orthogonale; il en résulte que, si ψ est le signal de V_{-1} lié au père φ par la formule

$$\widehat{\psi}(2\omega) = m_1(\omega) \widehat{\varphi}(\omega), \quad \omega \in \mathbb{R},$$

alors non seulement ψ est en fait dans W_0 , mais encore la collection $(\psi((\cdot) - n))_{n \in \mathbb{Z}}$ constitue une base orthonormée de W_0 .

Cette fonction ψ , intimement liée à φ comme on vient de le voir, est appelée "mère" de l'analyse, ou encore "ondelette mère".

On peut par exemple la calculer dans les trois exemples (Haar, Shannon, Franklin) que nous avons proposé comme modèle d'analyse multi-résolution. On trouve :

$$\begin{aligned} \psi^{\text{haar}}(t) &= \chi_{[0,1/2]}(t) - \chi_{[1/2,1]}(t) \\ \psi^{\text{shan}}(t) &= -\cos\left(\frac{3\pi}{2}\left(t - \frac{1}{2}\right)\right) \operatorname{sinc} \frac{\pi(t - \frac{1}{2})}{2} \\ \widehat{\psi^{\text{frank}}} &= -\sqrt{3} e^{-i\omega/2} \sqrt{\frac{2 - \cos(\omega/2)}{(2 + \cos \omega)(2 + \cos(\omega/2))}} (\operatorname{sinc}(\omega/4))^2. \end{aligned}$$

Notons d'ailleurs que la connaissance de la fonction m_0 suffit à générer tout l'outillage de l'analyse (elle induit la construction du père et de la mère, donc celle de

V_0) via les formules :

$$\begin{aligned}\widehat{\varphi}(\omega) &= m_0(\omega/2)\widehat{\varphi}(\omega/2) = \cdots = \prod_{j=1}^{\infty} m_0(2^{-j}\omega) \\ \widehat{\psi}(\omega) &= e^{-i(\omega+\pi)/2} m_0((\omega+\pi)/2) \prod_{j=2}^{+\infty} m_0(2^{-j}\omega).\end{aligned}$$

La contrainte sur m_0 fait que N étant un entier fixé, il existe un et un seul polynôme trigonométrique $m_0^{(N)}$ satisfaisant à la fois $m_0^{(N)}$ tel que

$$m_0^{(N)}(\omega) = \left(\frac{1 + e^{-i\omega}}{2}\right)^N (1 + \gamma_{N,1}e^{-i\omega} + \cdots + \gamma_{N,N-1}e^{-i(N-1)\omega}).$$

Les programmes **daub4**, **daub8** auxquels nous ferons appel par la suite pour illustrer la décomposition temps-échelles impliquent les analyses multi-résolution générées respectivement par les polynômes trigonométriques $m_{0,2}$ et $m_{0,4}$.

Le rôle de la mère est capital puisque, pour chaque $j \in \mathbb{Z}$, le système constitué des fonctions $(\psi_{j,n})_{n \in \mathbb{Z}}$, où

$$\psi_{j,n}(t) = 2^{-j/2} \psi(2^{-j}t - n),$$

forme une base orthonormée de W_j ; ceci implique, puisque la somme orthogonale des W_j a pour adhérence l'espace $L^2(\mathbb{R})$ tout entier, que la collection des vagues obtenues à partir du motif ψ en le dilatant ou le contractant de manière dyadique, puis le translatant via les entiers, c'est-à-dire la collection

$$(\psi_{j,n})_{j \in \mathbb{Z}, n \in \mathbb{Z}},$$

constitue une base hilbertienne de $L^2(\mathbb{R})$; tout signal de cet espace se décompose comme suit :

$$s = \sum_{j,n \in \mathbb{Z}} \langle s, \psi_{j,n} \rangle \psi_{j,n},$$

où

$$\langle s, \psi_{j,n} \rangle := \int_{\mathbb{R}} s(t) \overline{\psi_{j,n}(t)} dt, \quad j, n \in \mathbb{Z}.$$

Ainsi, notre signal

$$s = \sum_{n \in \mathbb{Z}} s_n \varphi(t - n)$$

attaché à notre information digitale $(s_n)_{n \in \mathbb{Z}}$ se décompose sous la forme

$$s = \sum_{j=1}^{+\infty} \left(\sum_{n \in \mathbb{Z}} \langle s, \psi_{j,n} \rangle \psi_{j,n} \right).$$

Ce sont les diverses suites

$$(\langle s, \psi_{j,n} \rangle)_n,$$

pour $j = 1, \dots, +\infty$ (notons que si l'on revient au point de vue digital, ces suites représentent une information à lire respectivement à l'échelle $0, 1, \dots$, c'est-à-dire correspondent à des suites indexées respectivement par $\mathbb{Z}, 2\mathbb{Z}, 4\mathbb{Z}, \dots$) que nous allons calculer à partir de la suite $(s_n)_{n \in \mathbb{Z}}$ dans la section suivante, ce qui induira ce que l'on appellera la décomposition en ondelettes discrètes du signal digital $(s_n)_n$ suivant l'analyse multi-résolution (c'est-à-dire en fait la fonction m_0) que l'on s'est fixé.

3.6 Analyse temps-échelles II : de l'information à ses détails

La construction même de m_1 à partir de m_0 fait que cette fois $m_1(0) = 0$ et que

$$m_1(\omega) = \frac{1}{\sqrt{2}} \sum_{n \in \mathbb{Z}} g(n) e^{-in\omega},$$

avec

$$g(n) = (-1)^n \overline{h(1-n)}, \quad n \in \mathbb{Z}.$$

Le calcul de la projection sur W_{j+1} d'un signal appartenant à V_j se fait alors suivant les règles suivantes :

Proposition 3.2 *L'image du signal*

$$s = \sum_{n \in \mathbb{Z}} s_n^{(j)} \varphi_{j,n}$$

de V_j par l'opérateur de projection orthogonale Dét_j est le signal

$$\text{Dét}_j[s] = \sum_{n \in \mathbb{Z}} d_n^{(j+1)} \psi_{j+1,n},$$

où

$$d_n^{(j+1)} = \sum_{\nu \in \mathbb{Z}} s_\nu^{(j)} \overline{g(\nu - 2n)} = \sum_{\nu \in \mathbb{Z}} (-1)^\nu s_\nu^{(j)} h(2n + 1 - \nu), \quad n \in \mathbb{Z}. \quad (*)$$

Preuve. Elle se résume à la linéarité de l'opérateur Dét_j et au fait que

$$\begin{aligned} \text{Dét}_j[\varphi_{j,n}] &= \sum_{\nu \in \mathbb{Z}} \langle \varphi_{j,n}, \psi_{j+1,\nu} \rangle \psi_{j+1,\nu} \\ &= \frac{1}{2\pi} \sum_{\nu \in \mathbb{Z}} \left(\int_{\mathbb{R}} \widehat{\varphi_{j,n}}(\omega) \overline{\widehat{\psi_{j+1,\nu}}(\omega)} d\omega \right) \psi_{j+1,\nu} \\ &= \frac{1}{\sqrt{2\pi}} \sum_{\nu \in \mathbb{Z}} \left(\int_{\mathbb{R}} \widehat{\varphi}(\omega) \overline{\widehat{\psi}(2\omega)} e^{-i(l-2\nu)\omega} d\omega \right) \psi_{j+1,\nu} \\ &= \frac{1}{\sqrt{2\pi}} \sum_{\nu \in \mathbb{Z}} \left(\int_{\mathbb{R}} |\widehat{\varphi}(\omega)|^2 \overline{m_1(\omega)} e^{-i(l-2\nu)\omega} d\omega \right) \psi_{j+1,\nu} \\ &= \frac{1}{\sqrt{2\pi}} \sum_{\nu \in \mathbb{Z}} \left(\int_0^{2\pi} \overline{m_1(\omega)} e^{-i(l-2\nu)\omega} d\omega \right) \psi_{j+1,\nu} \\ &= \sum_{\nu \in \mathbb{Z}} \overline{g(n - 2\nu)} \varphi_{j+1,\nu}. \end{aligned}$$

Les formules (*) en résultent. \diamond

L'action de l'opérateur Dét_j , comme celle de l'action de l'opérateur Rés_j , se traduit, au niveau des suites de coefficients qui représentent les coordonnées de l'entrée (exprimée dans la base $(\varphi_{j,n})_n$ de V_j et de la sortie (exprimée dans la base $(\psi_{j+1,n})_n$), par un filtrage (cette fois *passé-haut*), suivi d'une décimation. Si

$$s = \sum_{l \in \mathbb{Z}} s_n^{(j)} \varphi_{j,n},$$

on voit que

$$\text{Dét}_j[s] = \sum \tilde{d}_n^{(j)} \varphi_{j,n},$$

où

$$\tilde{d}_n^{(j)} = \sum_{\nu_1 \in \mathbb{Z}} \sum_{\nu_2 \in \mathbb{Z}} c_{\nu_2}^{(j)} \overline{g(\nu_2 - 2\nu_1)} g(n - 2\nu_1), \quad n \in \mathbb{Z}. \quad (**)$$

Si les formules (*) s'écrivent de manière condensée

$$(d_n^{(j+1)})_l = D_j[(s_n^{(j)})_n],$$

les formules (**) s'écrivent, quant à elles

$$(\tilde{d}_n^{(j)})_l = D_j^* D_j[(c_n^{(j)})_n],$$

et les deux opérateurs $R_j^* R_j$ et $D_j^* D_j$ de V_j dans V_j (correspondant, si l'on rapporte V_j à sa base hilbertienne $(\varphi_{j,n})_n$, aux opérateurs de projection orthogonale Rés_j et Dét_j) sont appelés pseudo-filtres miroirs en quadrature au seuil d'échelle j de l'analyse multi-résolution. On a de plus l'identité

$$\text{Id}_{V_j} = R_j^* R_j + D_j^* D_j.$$

Étant donné un signal de référence s supposé dans V_0 , on peut ainsi le décomposer sous la forme résumés+détails, et ce de manière orthogonale (c'est là un fait nouveau par rapport à l'algorithme pyramidal) et itérative, suivant le mécanisme

$$s = R_0^* R_0[s] + D_0^* D_0[s] = R_1^* R_1 R_0^* R_0[s] + D_1^* D_1 R_0^* R_0[s] + D_0^* D_0[s] = \dots$$

Le signal s était ici identifié au signal digital fourni par la liste $(s_n)_n$ de ses coordonnées dans la base $(\varphi_{0,n})_n$.

Si l'on sort du cadre dyadique (2 étant remplacé par un nombre premier $p > 2$, ce qui peut s'avérer intéressant si l'on a en vue un cadre plus arithmétique qu'analytique), on verra plus loin comment exploiter une "pseudo-identité de Bézout" quadratique

$$\text{Id}_{V_j} = \sum_{k=0}^{p-1} R_{j,k}^* R_{j,k};$$

la dichotomie résumé-détails se trouve affinée dans ce nouveau cadre ; l'information est scindée (toujours de manière orthogonale) en non plus deux, mais en p composants, ceci si l'on se place sur le plan discret (alors il y a scindage de l'information digitale $(s_n)_n$, chaque composant $R_{0,k}[(s_n)_n]$, $k = 0, \dots, p-1$, du scindage correspondant à un signal digital à lire toutes les p -unités si le signal digital est lui à lire toutes les unités) que sur le plan continu (les p composants $R_{0,k}^* R_{0,k}[s]$ sont des signaux continus orthogonaux dont l'empilement permet de restituer le signal s originel de V_0 dont on est parti).

3.7 La décomposition traitée sur un exemple

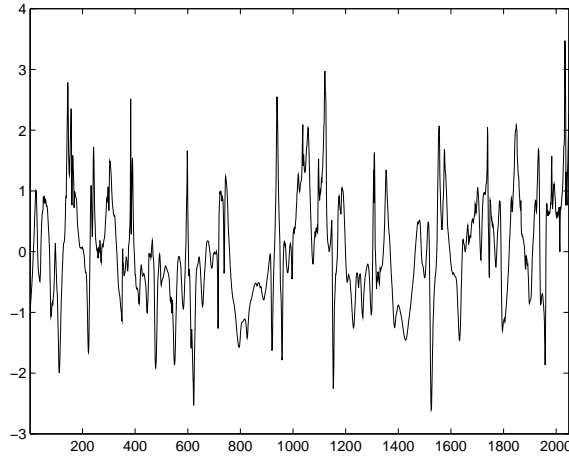


FIG. 3.4 – $s = v_x(t)$, $t = 0 : 1 : 2048$

On va traiter le signal digital de $2048 = 2^N$, $N = 11$, points représenté affiché ci-dessus comme un signal de l'espace V_0 correspondant aux fonctions $m_0^{(2)}$ et $m_0^{(4)}$; on affichera, pour $j = 1, 2, \dots$, les suites $(d_n^{(j)})_n$ des coordonnées (dans la base $(\psi_{j,n})_n$) des projections sur W_j , $j = 1, 2, \dots$, du signal

$$\sum_{n \in \mathbb{Z}} s_n \varphi(t - n)$$

extrapolant l'information digitale $(s_n)_n$; pour j donné, la suite numérique $(d_n^{(j)})_n$ obtenue est en fait une suite de longueur 2^{N-j} , d'où la nécessité de la représenter à l'aide de "fonctions créneau" de pas 2^j si l'on prétend visualiser ces diverses suites sur une même figure, ce que nous avons fait pour notre exemple.

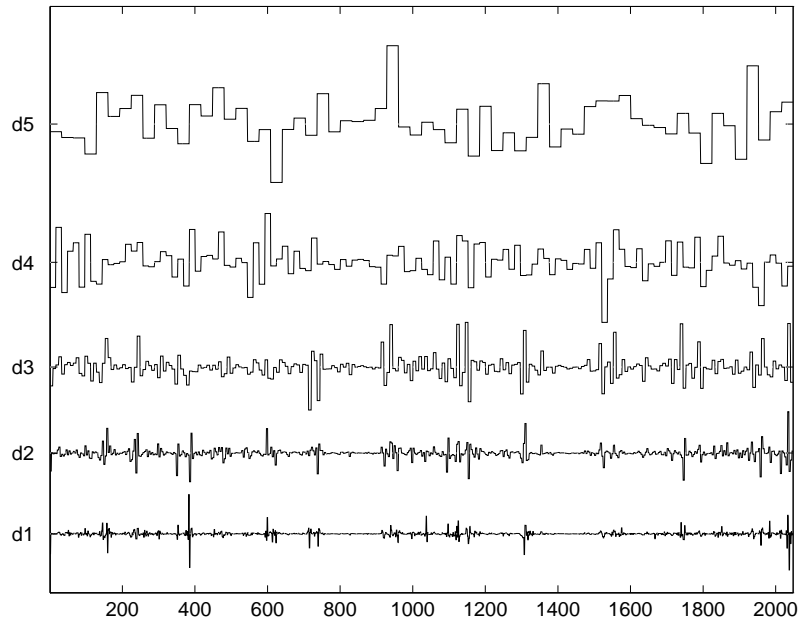


FIG. 3.5 – analyse de s en ondelettes discrète, $\text{AMR-}m_0^{(2)}$

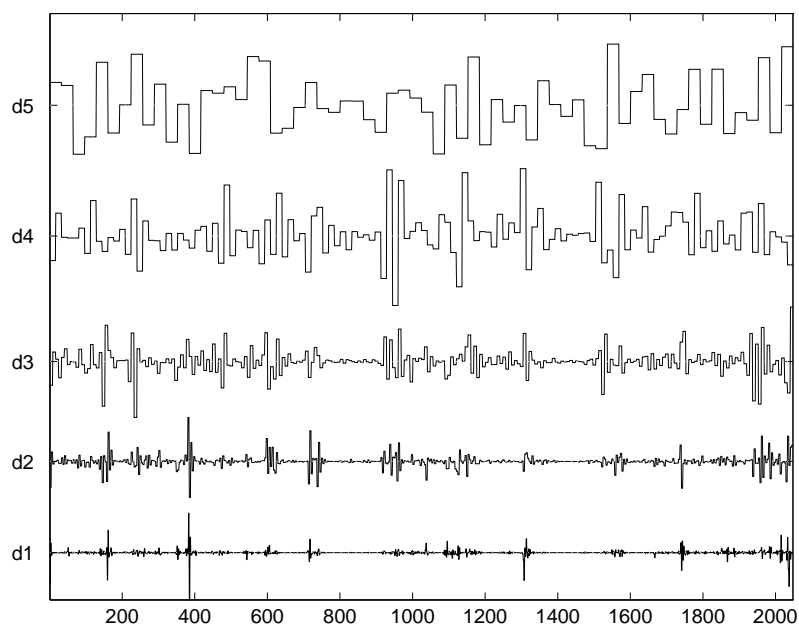


FIG. 3.6 – analyse de s en ondelettes discrète, $\text{AMR-}m_0^{(4)}$

Même si certaines informations relatives au contenu temps-échelles du signal se recoupent au niveau des résultats obtenus dans les deux analyses, on voit que ceux-ci sont sensiblement différents. La localisation d'un "accident" du signal à une certaine échelle dans la gamme des détails étudié peut s'avérer différente d'une analyse à l'autre (on peut observer la migration de certains coefficients d'ondelette significatifs d'un niveau d'échelle dyadique vers un autre par exemple).