

Sur la route des ondelettes

Albert Cohen¹

Depuis près de trois décennies, les ondelettes se sont imposées comme un outil puissant en analyse mathématique, et dans des domaines appliqués tels que le traitement du signal et de l'image, les statistiques ou encore la simulation numérique. La construction de cet outil est en quasi-totalité effectuée durant la décennie 1980-1990, tout comme l'identification de ses principales propriétés. L'attribution du prix Gauss à Yves Meyer nous donne l'occasion de nous repencher sur cette floraison scientifique, à l'interface entre l'analyse harmonique et le calcul numérique, dont il fut à la fois l'artisan, le chef d'orchestre, et le critique avisé. Si certaines des promesses formulées dans l'enthousiasme des premières découvertes n'ont pas toujours été suivies des succès escomptés, d'autres développements actuels inattendus et spectaculaires doivent beaucoup à ce moment fécond.

1. Représenter les fonctions : de Fourier à Gabor

Analyser, reconstruire et représenter des fonctions quelconques à l'aide de fonctions élémentaires, parfois surnommées *atomes*, est une démarche scientifique à l'origine de nombreuses avancées fondamentales et appliquées depuis plusieurs siècles. La possibilité offerte plus récemment par l'ordinateur d'implémenter une telle démarche au moyen d'algorithmes performants lui confère un intérêt supplémentaire pour le calcul numérique.

Dans l'exemple fondateur de la *transformée de Fourier*, on effectue l'analyse en fréquence d'une fonction f , définie sur \mathbb{R} par la formule

$$\hat{f}(\omega) = \int_{-\infty}^{+\infty} f(t)e^{-i\omega t} dt.$$

Sous des hypothèses convenables sur f , la fonction \hat{f} est bien définie et permet la *synthèse* de f par la formule d'inversion de Fourier

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{f}(\omega)e^{i\omega t} d\omega.$$

Dans cet exemple, les atomes sont les oscillations

$$e_{\omega}(t) = e^{i\omega t},$$

aux différentes fréquences $\omega \in \mathbb{R}$, et la combinaison de ces fonctions affectées des poids $\hat{f}(\omega)$ permet de reconstruire f . Les atomes e_{ω} servent aussi à l'analyse, puisque l'on peut formellement écrire $\hat{f}(\omega)$ comme un *produit scalaire*

$$\hat{f}(\omega) = \langle f, e_{\omega} \rangle,$$

avec la notation $\langle f, g \rangle = \int_{-\infty}^{+\infty} f(t)\overline{g(t)} dt$. La fonction \hat{f} est une *représentation* de f au sens où la connaissance de celle-ci est équivalente à celle de f .

¹ Laboratoire Jacques-Louis Lions, Université Pierre et Marie Curie, Paris.

Les mêmes remarques s'appliquent aux *séries de Fourier* pour l'analyse des fonctions définies sur un intervalle borné : si f est définie sur $[a, b]$, on étudie sa décomposition sous la forme de la série de fonctions de période $T = b - a$,

$$f(t) = \sum_{n \in \mathbb{Z}} c_n e^{i \frac{2\pi n t}{T}}.$$

Les coefficients de Fourier c_n peuvent être vus comme les *coordonnées* de f dans la base orthogonale de l'espace de Hilbert $L^2([a, b])$ constituée des fonctions $e_n(t) = e^{i2\pi n t/T}$, et sont donnés par les produits scalaires

$$c_n = c_n(f) = \frac{1}{T} \int_a^b f(t) e^{-i \frac{2\pi n t}{T}} dt.$$

Il est bien connu que le *calcul approché* était l'une des motivations à l'origine des séries de Fourier, l'approximation de f par la somme partielle $\sum_{n=-N}^N c_n e^{i2\pi n t}$ pouvant servir à l'évaluation de solutions d'équations différentielles ou aux dérivées partielles. Les applications des représentations de Fourier se sont intensifiées avec l'apparition des ordinateurs et le développement du calcul numérique, ainsi que l'invention de l'algorithme de transformée de Fourier rapide, permettant de calculer en $\mathcal{O}(N \log N)$ opérations la version discrète de la transformée de Fourier, où N représente la taille du vecteur représentant la fonction discrétisée.

Du point de vue du calcul, la représentation d'une fonction f est d'autant plus efficace qu'elle *concentre* sur un petit nombre de valeurs numériques l'essentiel de l'information permettant de reconstruire cette fonction avec précision. Les représentations de Fourier sont ainsi efficaces pour des combinaisons linéaires d'oscillations de la forme $\sum_{n=1}^N a_n e_{\omega_n}(t)$, puisque leur transformée de Fourier est une combinaison de masses de Dirac aux fréquences ω_n correspondantes, ainsi que pour les fonctions uniformément régulières : l'application itérée de la formule

$$\widehat{f'}(\omega) = (-i\omega)\widehat{f}(\omega)$$

montre que la régularité de f se traduit par la décroissance à l'infini de \widehat{f} .

Ces propriétés de concentration disparaissent lorsque f est régulière sauf en un point isolé de discontinuité, une telle fonction étant perçue comme globalement peu régulière par l'analyse en fréquence. Les transformées de Fourier sont aussi mises en échec lorsque l'on cherche à identifier des fréquences qui apparaissent uniquement dans certains intervalles de temps, telles des notes dans une partition de musique. Ces limitations traduisent le caractère global des transformées de Fourier : toutes les valeurs de f influencent toutes les valeurs de \widehat{f} et réciproquement.

Une première idée pour *localiser* l'analyse en fréquence est de multiplier la fonction f par une fonction g régulière et bien localisée, par exemple la gaussienne $g(t) = e^{-t^2}$ ainsi que le propose Dennis Gabor en 1945, et par ses translatées $g(t - \tau)$, avant d'appliquer la transformée de Fourier. On obtient ainsi la *transformée de Fourier locale* (ou à *fenêtre glissante*)

$$Gf(\omega, \tau) = \int_{-\infty}^{+\infty} f(t) g(t - \tau) e^{-i\omega t} dt,$$

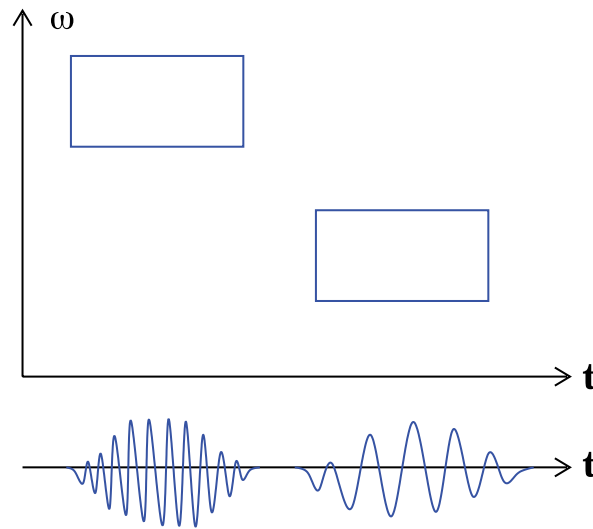


FIG. 1. Atomes temps-fréquence

qui mesure l'intensité de la fréquence ω dans un voisinage de l'instant τ . Elle peut aussi s'écrire $Gf(\omega, \tau) = \langle f, g_{\omega, \tau} \rangle$, où les atomes d'analyse sont donnés par

$$g_{\omega, \tau}(t) = g(t - \tau)e^{-i\omega t}, \quad \omega, \tau \in \mathbb{R}.$$

Remarquons que l'on a $\hat{g}_{\omega, \tau}(\xi) = e^{-i\tau\xi}\hat{g}(\xi - \omega)$, et que \hat{g} est aussi une gaussienne : les atomes sont à la fois bien localisées dans le temps autour de l'instant t et dans le domaine fréquentiel autour de la fréquence ω . On peut les visualiser symboliquement comme des rectangles de formes fixes, localisés à divers emplacements du plan temps-fréquence, repéré par les axes (ω, t) (Figure 1), les dimensions horizontales et verticales du rectangle correspondant aux variances de g et \hat{g} . Le principe d'incertitude limite cette double localisation par une aire minimale imposée à ces rectangles. On peut reconstruire f à partir de sa transformée Gf grâce à la formule

$$f(t) = C \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} Gf(\omega, \tau) e^{i\omega t} d\omega d\tau \quad \text{avec} \quad C = \left(\int_{-\infty}^{+\infty} g(t) dt \right)^{-1},$$

qui est une conséquence immédiate de la formule d'inversion de Fourier.

2. Transformées en ondelettes et frames

La résolution temporelle de la transformée de Fourier locale est limitée par l'échelle du support de g : la détection d'une oscillation localisée sur un intervalle de temps significativement plus petit se heurte inévitablement au même problème que dans le cas de la transformée de Fourier globale.

Au début des années 1980, Jean Morlet propose une solution différente : partant d'une fonction ψ bien localisée dans le plan temps-fréquence et oscillante au sens

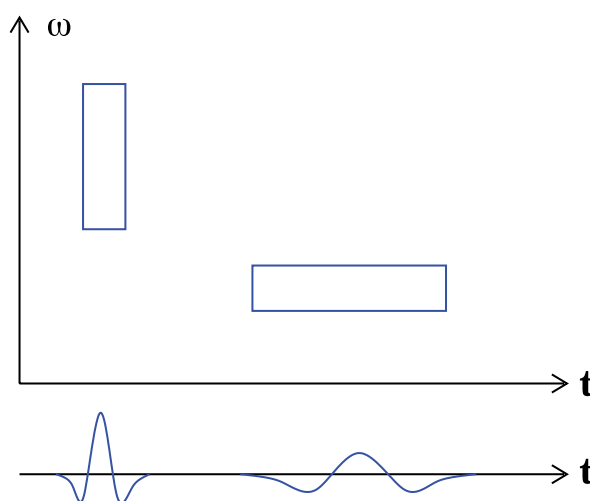


FIG. 2. Ondelettes

où $\int_{-\infty}^{+\infty} \psi(t) dt = 0$, il construit une famille de fonctions analysantes $\psi_{a,b}$, dites *ondelettes*, au moyen de translations et dilatations

$$\psi_{a,b}(t) = a^{-1/2} \psi\left(\frac{t-b}{a}\right), \quad a > 0, \quad b \in \mathbb{R}.$$

Le facteur $a^{-1/2}$ normalise cette famille, au sens où $\|\psi_{a,b}\|_{L^2}$ est indépendante de a et b . La transformée en ondelettes de f est alors définie par

$$Wf(a, b) := \langle f, \psi_{a,b} \rangle = a^{-1/2} \int_{-\infty}^{+\infty} f(t) \psi\left(\frac{t-b}{a}\right) dt.$$

La formule de reconstruction

$$f(t) = C \int_0^{+\infty} \int_{-\infty}^{+\infty} Wf(a, b) \psi_{a,b}(t) \frac{db da}{a^2} \quad \text{avec} \quad C = \left(\int_{-\infty}^{+\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega \right)^{-1},$$

conjecturée par Jean Morlet est démontrée en 1982 par Alexandre Grossman, qui l'identifie comme un cas particulier d'une formule de résolution de l'identité en théorie des représentations des groupes.

Par contraste avec les transformées précédentes, le paramètre d'échelle a ouvre l'accès à l'analyse de phénomènes d'oscillation ou de régularité arbitrairement localisés en temps au prix d'une perte de localisation en fréquence : quand a tend vers 0, les ondelettes $\psi_{a,b}$ sont visualisées par des rectangles fins en temps de l'ordre de l'échelle a et longs en fréquences de l'ordre de $1/a$ (Figure 2).

Dans le cadre du calcul numérique, on est naturellement amené à *échantillonner* les transformées Gf et Wf . Dans le cas de la transformée de Fourier locale, il est assez clair que ceci peut se faire sans perte d'information sur f : si g est à support dans un intervalle $[a, b]$, positive et ne s'annule pas sur un intervalle fermé de longueur t_0 contenu dans $[a, b]$, on peut caractériser f par la donnée de toutes les fonctions $f_m(t) = f(t)g(t - mt_0)$ pour $m \in \mathbb{Z}$, puisque $f(t) =$

$(\sum_{m \in \mathbb{Z}} f_m(t)) / (\sum_{m \in \mathbb{Z}} g(t - mt_0))$, où le dénominateur est strictement positif. Les fonctions f_m sont elles mêmes caractérisées par la donnée de leurs coefficients de Fourier

$$c_n(f_m) = Gf(mt_0, n\omega_0) = \langle f, g_{n\omega_0, mt_0} \rangle, \quad n \in \mathbb{Z},$$

avec $\omega_0 = 2\pi/(b - a)$. L'échantillonnage naturel de Gf est ainsi donné par le réseau $(mt_0, n\omega_0)_{m, n \in \mathbb{Z}}$, ce qui correspond à effectuer un recouvrement du plan temps-fréquence par les rectangles associés aux atomes

$$g_{n, m} := g_{n\omega_0, mt_0}, \quad n, m \in \mathbb{Z}.$$

Dans le cas des ondelettes, un tel recouvrement est plus naturellement associé à un réseau du type $(a_0^n, mb_0 a_0^n)_{n, m \in \mathbb{Z}}$, avec $a_0 > 1$ et $b_0 > 0$ fixés, qui prend en compte le changement des formes des rectangles en fonction de a . Pour ψ , a_0 et b_0 fixés, on définit ainsi la famille

$$\psi_{n, m} := \psi_{a_0^n, mb_0 a_0^n}, \quad n, m \in \mathbb{Z},$$

et l'on cherche à comprendre si l'échantillonnage $Wf(a_0^{-n}, mb_0 a_0^{-n}) = \langle f, \psi_{n, m} \rangle$ permet de caractériser f et de la reconstruire.

La réponse à ce problème est liée à la notion suivante : une suite $(f_n)_{n \geq 0}$ dans un espace de Hilbert H est appelée *frame*, si et seulement si il existe deux constantes $0 < A \leq B$ telles que pour tout $f \in H$,

$$A\|f\|^2 \leq \sum_{n \geq 0} |\langle f, f_n \rangle|^2 \leq B\|f\|^2.$$

Un tel encadrement montre que la suite des produits scalaires $c_n = \langle f, f_n \rangle$ caractérise f de façon stable au sens où la norme de f dans H est équivalente à la norme hilbertienne discrète de la suite $(c_n)_{n \geq 0}$. Il signifie aussi que l'opérateur F qui associe à x la suite (c_n) est continu et inversible à gauche, ce qui permet de construire une deuxième suite $\tilde{f}_n = (F^*F)^{-1}f_n$, dite *frame dual*, qui conduit à la formule de reconstruction

$$f = \sum_{n \geq 0} \langle f, f_n \rangle \tilde{f}_n,$$

que l'on peut implémenter numériquement. Dans le cas des ondelettes, on cherche donc à savoir si la famille $(\psi_{n, m})_{n, m \in \mathbb{Z}}$ est un frame pour l'espace de Hilbert $L^2(\mathbb{R})$. En 1984, Ingrid Daubechies montre que ceci est vrai dès que le réseau d'échantillonnage est suffisamment fin, c'est-à-dire si $a_0 b_0 < K$ où K est une constante qui dépend de ψ .

3. Les bases d'ondelettes d'Yves Meyer

La décomposition d'une fonction f dans un frame est généralement *redondante*, au sens où on n'a pas imposé l'indépendance linéaire des atomes f_n , par contraste avec une base de fonctions. Ceci constitue un défaut lorsque l'on souhaite obtenir une représentation la plus économique possible, ce qui est souvent le cas dans les applications numériques. La construction des bases orthonormées d'ondelettes proposée par Yves Meyer en 1985 est un moment clé qui fait le lien entre l'évolution décrite précédemment et d'autres programmes scientifiques.

Il faut d'abord rappeler que l'existence d'une telle base était connue depuis le début du 20^e siècle : il s'agit du système introduit par Alfred Haar dans sa thèse et qui a la forme

$$\psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k), \quad j, k \in \mathbb{Z},$$

où l'ondelette génératrice

$$\psi = \chi_{[0,1/2[} - \chi_{[1/2,1[},$$

est la fonction constante par morceau qui vaut 1 sur $[0, 1/2[$ et -1 sur $[1/2, 1[$. Il est immédiat de constater que ces fonctions sont orthogonales entre elles. Le fait qu'elles constituent une base de $L^2(\mathbb{R})$ est aussi facile à vérifier, en introduisant pour $f \in L^2(\mathbb{R})$ ses approximations constantes par morceaux

$$P_j f(t) = 2^j \int_{I_{j,k}} f(s) ds, \quad t \in I_{j,k} := [2^{-j}k, 2^{-j}(k+1)[, \quad k \in \mathbb{Z}.$$

On remarque que la projection sur les ondelettes de résolution 2^{-j} coïncide avec la différence de deux approximations successives

$$Q_j f = P_{j+1} f - P_j f = \sum_{k \in \mathbb{Z}} \langle f, \psi_{j,k} \rangle \psi_{j,k},$$

et on conclut par le fait que $\lim_{j \rightarrow +\infty} P_j f = f$ et $\lim_{j \rightarrow -\infty} P_j f = 0$. Le défaut principal du système de Haar est son manque de régularité puisque ψ n'est pas continue, ce qui équivaut à un manque de localisation de $\hat{\psi}$.

Un autre exemple élémentaire de base d'ondelettes $(\psi_{j,k})_{j,k \in \mathbb{Z}}$ est obtenu en définissant la fonction ψ par sa transformée de Fourier

$$\hat{\psi}(\omega) = \chi_{]-2\pi, -\pi]}(\omega) + \chi_{[\pi, 2\pi[}(\omega)$$

qui vaut 1 sur les intervalles $]-2\pi, -\pi]$ et $[\pi, 2\pi[$, et 0 ailleurs. Comme pour le système de Haar, l'orthogonalité des fonctions $\psi_{j,k}$ est facile à établir, et le fait qu'elles constituent une base se montre de façon analogue mais en utilisant des approximations successives définies par troncature en fréquence

$$\widehat{P_j f}(\omega) = \hat{f}(\omega) \chi_{[-2^j \pi, 2^j \pi]}(\omega).$$

La fonction ψ est parfois appelée ondelette de *Shannon* par référence au théorème de Shannon-Nyquist qui caractérise le pas d'échantillonnage permettant de reconstruire les fonctions à support limité en fréquence au moyen d'une base orthonormée. L'ondelette de Shannon est, à l'inverse de celle de Haar, parfaitement localisée en fréquence mais mal localisée en temps.

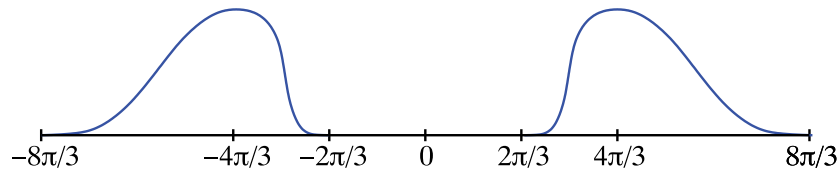


FIG. 3. L'ondelette d'Yves Meyer (graphe de $|\hat{\psi}|$)

L'existence d'une base orthonormée d'ondelettes engendrée par une fonction ψ à la fois bien localisée en temps et en fréquence semble cependant loin d'être acquise, car une obstruction existe pour le même problème dans le cas des atomes temps-fréquence $g_{m,n}$ issus de la discrétisation de la transformée de Gabor. Un théorème dû à Roger Balian et Francis Low affirme en effet que ces atomes ne peuvent constituer une base orthonormée que si l'un des moments quadratiques $\int_{-\infty}^{+\infty} |t|^2 |g(t)|^2 dt$ ou $\int_{-\infty}^{+\infty} |\omega|^2 |\hat{g}(\omega)|^2 d\omega$ est divergent. C'est en se posant la question d'un résultat similaire pour les ondelettes qu'Yves Meyer parviendra finalement à construire une base avec une fonction ψ dont les moments de tous les ordres convergent ainsi que ceux de $\hat{\psi}$. L'ondelette proposée à la forme d'une version régularisée en fréquence de l'ondelette de Shannon (Figure 3) : la fonction $\hat{\psi}$ est infiniment régulière, à support dans $[-8\pi/3, -2\pi/3] \cup [2\pi/3, 8\pi/3]$, son module est une fonction paire vérifiant

$$\sum_{n \in \mathbb{Z}} |\hat{\psi}(\omega + 2n\pi)|^2 = 1 \quad \text{et} \quad \sum_{j \in \mathbb{Z}} |\hat{\psi}(2^j \omega)|^2 = 1.$$

On vérifie aisément par la formule de Poisson que la première relation assure l'orthogonalité des $\psi_{j,k}$ pour un niveau d'échelle j fixé. Un choix astucieux de la phase suivant $e^{i\omega/2}$ pour définir $\hat{\psi}(\omega)$, assure l'orthogonalité entre les échelles.

Une autre base orthonormée d'ondelettes bien localisée en temps et en fréquence avait déjà été construite en 1981 par Jan Olov Stromberg, en suivant une approche tout à fait différente, plus proche de celle qui est discutée dans la section suivante.

Dans sa conception, l'ondelette d'Yves Meyer s'inspire d'une *décomposition dyadique*

$$f = \sum_{j \in \mathbb{Z}} \Delta_j f$$

proposée par John Littlewood et Raymond Paley dans les années 1930 pour l'analyse harmonique des espaces L^p . Les blocs dyadiques de la décomposition de Littlewood-Paley sont définis par filtrage de f aux fréquences de l'ordre de 2^j

$$\widehat{\Delta_j f}(\omega) = \hat{\Psi}(2^{-j}\omega) \hat{f}(\omega),$$

où la fonction $\hat{\Psi}$ est du même type que $|\hat{\psi}|$. Dans la décomposition en ondelettes, les blocs $\Delta_j f$ sont remplacés par les projections

$$Q_j f := \sum_{k \in \mathbb{Z}} \langle f, \psi_{j,k} \rangle \psi_{j,k},$$

l'indice k apportant ainsi un niveau supplémentaire de discrétisation.

L'analyse de Littlewood-Paley permet de caractériser avec précision les propriétés de régularité de fonctions, et Yves Meyer montre qu'il en est de même des décompositions en ondelettes. Par exemple, si α est un nombre positif non-entier et $f \in L^\infty(\mathbb{R})$, on peut montrer que les trois propriétés suivantes sont équivalentes :

(1) La fonction f est uniformément Hölderienne d'exposant α sur \mathbb{R} , c'est-à-dire que pour tout $t \in \mathbb{R}$ il existe un polynôme P_t de degré $m < \alpha$ tel que

$$|f(s) - P_t(s)| \leq A|s - t|^\alpha,$$

où $A > 0$ est indépendant de s et t .

(2) Les blocs de Littlewood-Paley de f vérifient

$$\|\Delta_j f\|_{L^\infty} \leq B 2^{-\alpha j},$$

où $B > 0$ est indépendante de j .

(3) Les coefficients d'ondelettes de f vérifient

$$|\langle f, \psi_{j,k} \rangle| \leq C 2^{-(\alpha + \frac{1}{2})j},$$

où $C > 0$ est indépendante de j et k .

On retrouve ainsi avec plus de précision l'idée, déjà présente dans l'analyse de Fourier, que la régularité se traduit par de la décroissance dans la représentation. La présence de l'indice k dans la décomposition en ondelettes permet en outre d'apporter une information supplémentaire sur la régularité *locale* de f . Nous reviendrons un peu plus loin sur ce point important.

La possibilité de caractériser l'appartenance d'une fonction à la quasi-totalité des espaces fonctionnels classiques (Hölder, Lebesgue, Sobolev, Besov) par des conditions portant uniquement sur les *modules* des coefficients de cette fonction dans une base d'ondelettes traduit le fait que celle-ci a la propriété de *base inconditionnelle* pour ces espaces. Par définition, une base $(f_n)_{n \geq 0}$ d'un espace de Banach X est inconditionnelle si et seulement si il existe une constante $C > 0$ telle que pour tout $N > 0$, on ait la propriété

$$|c_n| \leq |d_n|, \quad n = 0, \dots, N \Rightarrow \left\| \sum_{n=0}^N c_n f_n \right\|_X \leq C \left\| \sum_{n=0}^N d_n f_n \right\|_X.$$

Cette propriété est fondamentale pour les applications numériques car elle assure la stabilité dans la norme $\|\cdot\|_X$ d'opérations telles que la mise à zéro de certains coefficients ou leur approximation à une certaine précision par l'ordinateur.

4. Analyse multirésolution et filtres numériques

Dans le domaine du traitement numérique du signal et de l'image, les ingénieurs utilisent depuis les années 1970 des décompositions hiérarchiques présentant des similarités avec l'analyse de Littlewood-Paley. Ainsi, partant d'une image discrétisée sur une grille de pixels, on peut considérer ses approximations obtenues en moyennant successivement l'intensité lumineuse sur des carrés de 2×2 , 4×4 , 8×8 pixels etc. De telles approximations sont *hiérarchisées* au sens où l'image approchée à une certaine résolution se déduit de celle approchée à la résolution deux fois plus fine, et on cherche alors à caractériser l'information manquante entre chaque niveau.

En 1986, Stéphane Mallat et Yves Meyer remarquent que ce point de vue conduit naturellement aux bases d'ondelettes, à travers la notion fondamentale d'*analyse multirésolution*. Cette notion désigne une suite emboîtée d'espaces

$$\{0\} \subset \dots \subset V_{j-1} \subset V_j \subset V_{j+1} \subset \dots \subset L^2(\mathbb{R}),$$

l'entier j étant associé à la résolution 2^{-j} au sens où V_j est engendré par une base du type

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

La fonction φ est appelée *fonction d'échelle*. Dans le cas des espaces de fonctions constantes par morceaux qui mène au système de Haar, on a tout simplement $\varphi = \chi_{[0,1]}$, c'est-à-dire $\varphi(t) = 1$ sur $[0, 1]$ et 0 ailleurs. On désigne par P_j la projection orthogonale sur V_j et on exige que pour tout $f \in L^2(\mathbb{R})$,

$$\lim_{j \rightarrow -\infty} \|P_j f\|_{L^2} = 0 \text{ et } \lim_{j \rightarrow +\infty} \|f - P_j f\|_{L^2} = 0.$$

La deuxième propriété signifie que les espaces V_j permettent d'approcher n'importe quelle fonction lorsque l'on fait tendre la résolution 2^{-j} vers 0.

En pratique, on s'intéresse souvent à d'autres types de convergence de $P_j f$ vers f (par exemple la convergence uniforme), et on souhaite de plus avoir des informations sur la vitesse de cette convergence. Dans le cas des fonctions constantes par morceaux, il est facile de vérifier que si f est suffisamment régulière au sens où $f' \in L^p(\mathbb{R})$, on a

$$\|f - P_j f\|_{L^p} \leq 2^{-j} \|f'\|_{L^p},$$

et qu'une telle estimation ne peut être améliorée même si f est beaucoup plus régulière. Ceci reflète le fait que la précision de l'approximation par des fonctions constantes est d'ordre 1 en fonction de la résolution. Afin d'augmenter l'ordre d'approximation, une idée naturelle est de remplacer les constantes par des polynômes de degré plus élevé, ce qui permet aussi d'augmenter la régularité des fonctions de V_j . On peut ainsi définir V_j comme le sous-espace de $L^2(\mathbb{R})$ constitué des fonctions polynômiales de degré m sur les intervalles $I_{j,k}$ définis plus haut, et globalement $m - 1$ fois continûment dérivables : ce sont des *fonctions splines* de degré m , bien connues des ingénieurs pour la modélisation numérique des courbes et surfaces. Dans ce cas, on peut prendre pour fonction d'échelle φ la *B-spline* obtenue en effectuant $m + 1$ produits de convolution de $\chi_{[0,1]}$ par elle-même

$$\varphi = \chi_{[0,1]} * \dots * \chi_{[0,1]} = (*)^{m+1} \chi_{[0,1]}.$$

On note qu'en dehors du cas $m = 0$, la base $(\varphi_{j,k})_{k \in \mathbb{Z}}$ n'est pas orthonormée.

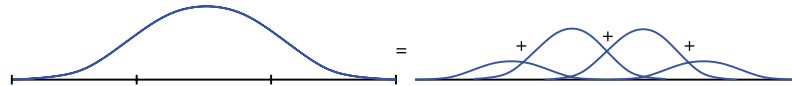


FIG. 4. B-spline quadratique
 $\varphi(t) = \frac{3}{4}(\varphi(2t - 1) + \varphi(2t - 2)) + \frac{1}{4}(\varphi(2t) + \varphi(2t - 3))$

Partant d'une analyse multirésolution, on cherche à présent à construire une ondelette ψ permettant de caractériser l'information manquante $P_{j+1}f - P_j f$ entre deux niveaux d'approximation, au moyen des fonctions $(\psi_{j,k})_{k \in \mathbb{Z}}$. Une remarque essentielle pour cette construction est que la fonction d'échelle vérifie une équation d'autosimilarité

$$\varphi(t) = \sum_{n \in \mathbb{Z}} h_n \varphi(2t - n),$$

qui exprime l'emboîtement des espaces V_j . On peut vérifier, sous des hypothèses minimales, que le support de la suite h_n est de même taille que celui de la fonction φ . Par exemple, dans le cas de la B-spline de degré m , on a $\text{Supp}(\varphi) = [0, m + 1]$ et

$h_n = 2^{-m} \binom{m+1}{n}$ pour $n = 0, \dots, m+1$ et $h_n = 0$ sinon. La Figure 4 illustre ainsi cette équation dans le cas de la B-spline quadratique ($m = 2$).

Yves Meyer démontre que si la base $(\varphi_{j,k})_{k \in \mathbb{Z}}$ de V_j est orthonormée, on peut définir l'ondelette ψ par la relation

$$\psi(t) = \sum_{n \in \mathbb{Z}} g_n \varphi(2t - n),$$

avec $g_n = (-1)^n h_{1-n}$, les fonctions $(\psi_{j,k})_{k \in \mathbb{Z}}$ obtenues par changement d'échelle constituant alors une base du supplémentaire orthogonal de V_j dans V_{j+1} . Par le même raisonnement que pour le système de Haar et l'ondelette de Shannon on aboutit ainsi à une base orthonormée $(\psi_{j,k})_{j,k \in \mathbb{Z}}$ de $L^2(\mathbb{R})$.

Le problème qui subsiste est la construction de fonctions d'échelles φ telles que $(\varphi_{j,k})_{k \in \mathbb{Z}}$ constitue une base orthonormée, et ayant des propriétés intéressantes pour les applications : support compact, régularité, ordre élevé d'approximation pour les espaces V_j correspondants. On peut chercher à utiliser les splines en construisant une nouvelle fonction d'échelle dont les translatés sont orthogonaux, mais le défaut de cette construction est que la fonction φ obtenue n'est plus à support compact, et par conséquent plus difficile à manipuler numériquement que la B-spline de départ.

Ce problème est finalement entièrement résolu par Ingrid Daubechies en 1988 dans un travail qui fait l'unanimité pour sa pertinence et sa profondeur. L'approche suivie consiste à définir directement la fonction φ comme une solution de l'équation d'échelle, et de mettre au point les coefficients h_n dans cette équation de façon à assurer les propriétés requises sur la solution φ . Le tour de force accompli par Ingrid Daubechies est à la fois d'identifier la nature de ces propriétés sous forme d'équations algébriques portant sur les h_n et de résoudre ces équations. Elle construit pour tout m une suite de coefficients $(h_n)_{n=0, \dots, 2m-1}$, telle que la solution φ de l'équation d'échelle, à support compact dans $[0, 2m-1]$, engendre une base orthonormée $(\varphi_{j,k})_{k \in \mathbb{Z}}$ d'espaces d'approximation d'ordre m , et montre en outre que la régularité de la fonction φ tend vers $+\infty$ avec m . Le cas $m = 1$ correspond simplement au système de Haar. Pour $m > 1$, les fonctions φ et ψ n'ont pas d'expression mathématique explicite (contrairement aux splines), et sont déterminées implicitement par la donnée des coefficients h_n . Cette construction est ensuite étendue par Albert Cohen, Ingrid Daubechies et Jean-Christophe Feauveau en 1992, au cadre plus général des bases *biorthogonales*, qui est le plus utilisé en pratique en raison de sa flexibilité (on peut ainsi en particulier obtenir des ondelettes splines à support compact).

L'introduction de l'analyse multirésolution donne non seulement un éclairage nouveau sur la construction des bases d'ondelettes, mais elle permet aussi par sa flexibilité d'adapter cette construction à de nombreuses situations géométriques : on sait ainsi construire des analyses multirésolutions et des ondelettes sur des domaines multidimensionnels (par exemple en partant d'espaces d'éléments finis emboîtés), ou plus généralement sur certains types de variétés.

Le cadre des approximations hiérarchiques joue en outre un rôle essentiel pour la mise en œuvre d'algorithmes permettant le calcul des coefficients d'ondelettes en partant de données discrétisées à un certain niveau de résolution 2^{-j} . Ces données sont assimilées à une fonction d'un espace V_j et on calcule de proche en proche les

coefficients aux échelles $j' \leq j$ en allant vers la résolution plus grossière. Ces algorithmes utilisent les suites h_n et g_n intervenant dans les équations d'échelles puisque ce sont celles-ci qui relient les ondelettes et fonctions d'échelles aux différentes résolutions. Ces suites jouent le rôle de filtres numériques, et les algorithmes de décomposition obtenus sont du même type que ceux développés par les ingénieurs dans le domaine du traitement du signal et de l'image. De même, les algorithmes de reconstruction, qui remontent vers la résolution plus fine, se révèlent être du même type que les *schémas de subdivision* proposés en modélisation géométrique pour générer des courbes et surfaces à partir d'un petit nombre de points de contrôle par des raffinement itératifs.

5. Du traitement du signal à la simulation numérique

En à peine dix ans, la théorie des ondelettes est ainsi forgée dans toute sa variété, grâce à des apports provenant de nombreux domaines de recherche : analyse harmonique, traitement du signal et de l'image, théorie de l'approximation, modélisation des courbes et surfaces. Le premier grand congrès dédié aux ondelettes (Luminy 1989) témoigne de cette diversité thématique et de l'enthousiasme suscité par l'émergence d'un domaine permettant de telles rencontres, ainsi que par la multiplicité des applications envisagées.

En premier lieu, c'est la capacité des ondelettes à analyser les phénomènes localisés, à la manière d'un microscope mathématique, qui intéresse pour l'étude de signaux provenant de phénomènes acoustiques, électromagnétiques, économiques ou biologiques. Un effort particulier est porté sur l'analyse de la turbulence développée, autour de spécialistes tels que Uriel Frisch et Marie Farge, avec l'espoir de comprendre plus finement les lois de transfert d'énergie à travers les échelles.

Revenons en particulier sur la caractérisation de la régularité Hölderienne d'une fonction par la décroissance des coefficients d'ondelettes en fonction du niveau d'échelle suivant l'estimation

$$|\langle f, \psi_{j,k} \rangle| \leq C 2^{-(\alpha + \frac{1}{2})j}.$$

Cette estimation est vérifiée uniformément sur j et k si la fonction f est uniformément Hölderienne d'exposant α sur tout \mathbb{R} , et en supposant que l'ondelette vérifie les relations de moments nuls

$$\int_{-\infty}^{+\infty} t^n \psi(t) dt = 0,$$

pour tout entier $n < \alpha$. Si l'ondelette ψ est à support compact, il est facile de montrer que si f est Hölderienne d'exposant α seulement en un point t , alors l'estimation ci-dessus sera assurée pour les indices j et k tels que $|2^j t - k| \leq M$ où M est une constante arbitraire (mais dont la taille influence la constante C), c'est-à-dire pour les ondelettes dont les supports se concentrent vers le point t lorsque la résolution j augmente. En 1991, Stéphane Jaffard montre une réciproque à ce résultat sous une hypothèse de régularité uniforme minimale. Il est ainsi possible de caractériser la *régularité locale* d'une fonction par la taille de ses coefficients d'ondelettes.

Cette propriété ouvre des perspectives nouvelles en *analyse multifractale*, domaine dans lequel on cherche à quantifier avec précision la taille des ensembles de

points où une fonction f donnée présente une certaine régularité. On s'intéresse en particulier aux propriétés du *spectre de singularité* de f qui est la fonction $\alpha \mapsto d(\alpha)$ mesurant la dimension de Hausdorff de l'ensemble des points t où l'exposant de Hölder α est atteint. L'observation des signaux turbulents amène Uriel Frisch et Giorgio Parisi à proposer une conjecture appelée *formalisme multifractal*. Cette conjecture relie le spectre de singularité avec la transformée de Legendre de la fonction $s \mapsto s(p)$ qui mesure l'ordre de dérivabilité de f dans l'espace L^p (c'est-à-dire la valeur maximale de s telle que f admet une dérivée fractionnaire d'ordre s dans L^p). L'utilisation des bases d'ondelettes permet à Stéphane Jaffard de montrer en toute rigueur que le formalisme multifractal est mis en défaut par certaines fonctions, le spectre de singularité pouvant avoir une forme quelconque, mais que sa validité est néanmoins générique au sens de Baire sur les espaces fonctionnels classiques.

Si les ondelettes permettent d'identifier un point régulier dans un environnement moins régulier, c'est surtout la propriété inverse qui intéresse les numériciens : si f est une fonction régulière par morceaux présentant des singularités en des points isolés, alors ses coefficients d'ondelettes décroissent rapidement en fonction du niveau de résolution j , à l'exception d'une minorité d'entre eux pour lesquels le support de $\psi_{j,k}$ rencontre les points de singularité. La Figure 6, illustre cette propriété dans le cas où la fonction décomposée est une image : les coefficients d'ondelettes, représentés par l'image du centre, sont presque tous nuls ou très proche de 0, à l'exception de ceux qui sont localisés au voisinage des contours de l'image. Ces contours correspondent à des discontinuités de l'intensité lumineuse localisées sur des courbes isolées.

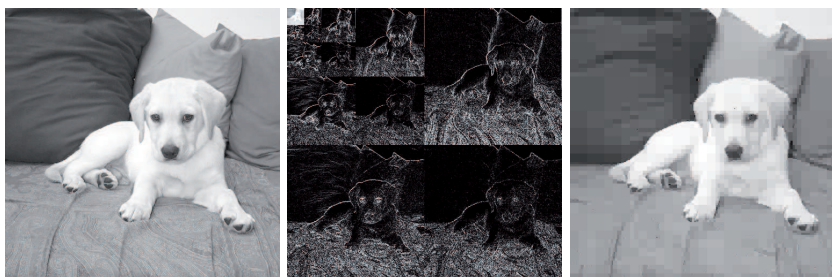


FIG. 5. Image 512×512 pixels, coefficients d'ondelettes, compression Haar 1/100

Cette propriété de concentration de l'information sur un petit nombre de coefficients – souvent appelée *parcimonie* ou *sparsity* – signifie qu'il est possible de reconstruire avec précision l'image en ne conservant qu'un petit nombre de coefficients obtenus par une procédure de *seuillage*. L'image de droite sur la Figure 5 montre ainsi la reconstruction obtenue à partir des 2500 plus grands coefficients dans la base de Haar (adaptée aux images bidimensionnelles), soit environ 1/100 du nombre total de coefficients d'ondelettes qui est égal à la dimension 512×512 de l'image de départ. On constate des artefacts visuels, liés à l'utilisation du système de Haar et qui disparaissent avec des ondelettes plus régulières, mais qui permettent

ici d'illustrer le fait que la résolution est *localement adaptée* et se raffine au voisinage des contours puisque les coefficients aux résolutions fines y ont été conservés par la procédure de seuillage.

On voit ainsi apparaître un lien entre les procédures de seuillage des coefficients d'ondelettes et le raffinement de maillage. Ces procédures sont par essence *adaptatives* et *non-linéaires*, l'ensemble des coefficients ou le maillage retenu variant selon la fonction que l'on approche. La *théorie de l'approximation non-linéaire*, développée autour de Ronald DeVore à partir des années 1980, offre un cadre général pour analyser leurs performances. Un résultat élémentaire de cette théorie propose une mesure mathématique du degré de parcimonie d'une représentation dans une base hilbertienne : si $\sigma_N(f)$ désigne l'erreur en norme L^2 obtenue si l'on ne conserve que les N plus grands coefficients de f , et si $(d_n)_{n>0}$ est la suite des modules des coefficients réarrangés par ordre décroissant, on a l'équivalence

$$\sup_{N>0} N^s \sigma_N(f) < \infty \Leftrightarrow \sup_{n>0} n d_n^p < \infty \quad \text{avec} \quad \frac{1}{p} = \frac{1}{2} + s.$$

La première estimation signifie que $\sigma_N(f)$ décroît comme N^{-s} , et la deuxième affirme que la suite des coefficients est *faiblement* ℓ^p sommable. L'espace $w\ell^p$ des suites faiblement ℓ^p sommables est proche mais plus grand que l'espace ℓ^p (par exemple, la suite $1/n$ appartient à $w\ell^1(\mathbb{N})$ mais sa série diverge logarithmiquement). Le degré de parcimonie est ainsi mesuré de manière équivalente par vitesse de convergence de l'approximation non-linéaire à N termes et par l'appartenance de la suite des coefficients aux espace $w\ell^p$, et ce degré augmente lorsque $p \rightarrow 0$ ou $s \rightarrow +\infty$. Dans le cas des bases d'ondelettes, il est également possible de relier ces propriétés à l'appartenance de la fonction à certains espaces fonctionnels mesurant la régularité (en particulier les espaces de Besov). Nous évoquons dans la section suivante un exemple important d'un tel résultat.

Une première application de la parcimonie est la compression des données. La représentation adaptative de l'image par ses N plus grands coefficients d'ondelettes n'est pas à proprement parler un algorithme de compression d'image puisque le codage de ces coefficients sans erreur d'arrondi exige a priori un nombre infini de bits. Cette représentation suggère néanmoins une stratégie de compression naturelle, consistant à ne coder avec précision que les plus grands coefficients d'ondelettes, et à seuiller les plus petits. Une telle stratégie sous-entend que l'on code aussi les indices correspondant aux coefficients numériquement significatifs. Ces idées simples sont à la base des algorithmes de compression les plus efficaces à l'heure actuelle, et qui ont été intégrés à la fin des années 1990 dans le standard JPEG 2000. Puisque les ingénieurs disposaient déjà des outils de décompositions multiéchelle des images au moyen de filtres numériques, il est légitime de se demander si le détour par les bases d'ondelettes proposé par les mathématiciens a joué un rôle fondamental dans le développement de nouvelles stratégies de compression. Il semble que l'on peut affirmer sans exagération que les travaux sur l'approximation non-linéaire par les bases d'ondelettes ont fait évoluer l'intuition dans ce domaine d'application : la propriété essentielle recherchée dans une représentation du signal ou de l'image en vue de la compression n'est plus la décorrélation (qui menait à la recherche de composantes principales) mais la parcimonie.

Les ondelettes reçoivent aussi un fort écho en statistiques à partir de 1990, à travers les travaux de David Donoho, Iain Johnstone, Dominique Picard et Gérard Kerkycharian sur les méthodes d'estimation par seuillage des coefficients. Ces méthodes consistent à ajuster le seuil au niveau du bruit de mesure, afin de capturer uniquement les coefficients numériquement significatifs qui émergent au-dessus de ce niveau. Leur succès est par conséquent conditionné aux propriétés de parcimonie des décompositions en ondelettes. Il s'agit là aussi de procédures non-linéaires, par contraste avec les techniques de filtrage plus anciennes qui présentent le défaut de lisser les discontinuités. Dans le domaine du traitement d'image, ces méthodes sont à comparer avec celles utilisant les EDP de diffusion non-linéaires ou avec celles plus récentes reposant sur des moyennages non-locaux, proposées par Jean-Michel Morel et qui constituent l'état de l'art actuel.

Depuis le début des années 1990, les bases d'ondelettes suscitent enfin un intérêt croissant dans le domaine de la simulation numérique des équations aux dérivées partielles. Si elles ne diagonalisent pas les opérateurs différentiels intervenant dans les EDP, les coefficients des matrices obtenues se concentrent néanmoins autour de la diagonale, et, comme le remarque Yves Meyer, c'est aussi le cas pour les opérateurs intégraux de type Calderon-Zygmund. Ces remarques sont à l'origine d'un algorithme proposé en 1990 par Gregory Beylkin, Ronald Coifman et Vladimir Rokhlin, pour le calcul rapide de produits matrices-vecteurs.

On peut d'autre part chercher à utiliser les représentations dans les bases d'ondelettes pour calculer sous une forme approchée la solution d'une EDP de la façon la plus économique possible. Nous avons observé que le seuillage des coefficients revient à effectuer un raffinement local de l'approximation aux voisinages des singularités d'un signal ou d'une image. En simulation numérique de tels raffinements sont nécessaires aux voisinages des singularités des solutions : ondes de chocs, ruptures, singularités. Contrairement à un signal que l'on comprime, la solution de l'EDP n'est pas une donnée explicite (elle est donnée sous forme implicite par l'équation), et on ne connaît pas par avance l'emplacement de ses coefficients d'ondelettes les plus significatifs. Les *stratégies adaptatives* visent à acquérir ces informations au fil du calcul. En 1998, Albert Cohen, Wolfgang Dahmen et Ronald DeVore montrent, dans le cadre des EDP elliptiques, qu'il existe des stratégies *optimales* en un sens suggéré par la théorie de l'approximation non-linéaire : la précision $\varepsilon(N)$ obtenue en fonction du nombre de coefficients d'ondelettes N calculés pour décrire la solution de l'EDP est comparable à la précision optimale que l'on obtiendrait si cette solution était une donnée connue que l'on approche en conservant ses N plus grands coefficients d'ondelettes. Il serait aujourd'hui inexact d'affirmer que les méthodes ondelettes ont surpassé les méthodes d'éléments finis adaptatifs qui sont leur concurrents naturels, mais il est intéressant de constater que le point de vue offert par l'approximation non-linéaire a souvent été repris par les numériciens pour mieux comprendre les performances des méthodes de raffinement de maillage.

6. À la recherche de la parcimonie

Afin d'évaluer plus précisément les performances des méthodes fondées sur les ondelettes pour le traitement d'image, il est nécessaire de disposer d'un modèle pour décrire l'ensemble des images plausibles. L'espace BV des *fonctions à variation bornée* est fréquemment utilisé, car il semble bien adapté pour décrire la structure

régulière par morceaux des images peu texturées. Rappelons qu'une fonction f est dans l'espace $BV(\Omega)$ pour Ω domaine de \mathbb{R}^d si et seulement si son gradient ∇f est une mesure de Radon finie. La variation totale de f est la masse de cette mesure, que l'on peut définir par dualité suivant

$$TV(f) := \sup \left\{ \int_{\Omega} f \operatorname{div}(g) : g \in \mathcal{D}(\Omega), \|g\|_{L^\infty} \leq 1 \right\},$$

où $\mathcal{D}(\Omega)$ est l'ensemble des fonctions C^∞ à support compact, ou par l'intégrale $\int_{\Omega} |\nabla f|$ lorsque $\nabla f \in L^1(\Omega)$. En particulier, une fonction de deux variables f régulière par morceaux présentant une discontinuité sur un ensemble Γ de courbes rectifiables de longueur finie, a une variation totale bornée et donnée par

$$TV(f) = \int_{\Omega \setminus \Gamma} |\nabla f| + \int_{\Gamma} |[f]|,$$

où $[f]$ désigne le saut de f . L'espace BV est aussi le cadre utilisé pour décrire la régularité des solutions faibles d'EDP hyperboliques non linéaires qui développent des profils de chocs en temps fini.

Cet espace n'est pas séparable, et n'admet pas de base inconditionnelle, mais il est néanmoins presque possible de le caractériser au moyen des bases d'ondelettes : en 1998, Albert Cohen, Ronald DeVore, Pencho Petrushev et Hong Xu démontrent que si $(d_n)_{n>0}$ est la suite des modules des coefficients d'ondelettes d'une image f réarrangés par ordre décroissant, on a l'encadrement

$$A \sup_{n>0} n d_n \leq TV(f) \leq B \sum_{n>0} d_n,$$

où $0 < A < B$ sont indépendantes de f , ce qui signifie que l'espace des coefficients d'ondelettes des fonctions BV en dimension 2 contient ℓ^1 et est contenu dans $w\ell^1$.

Il est d'autre part possible de montrer que le comportement en $1/n$ de la suite d_n est optimal et ne peut pas être amélioré, même pour des images très simples telles que l'indicatrice d'un polygone ou d'un cercle. D'après les remarques de la section précédente, ceci signifie aussi que la vitesse d'approximation par seuillage des N plus grands coefficients se comporte en $N^{-1/2}$ et pas mieux pour de telles fonctions. Ce résultat cerne ainsi les limites des ondelettes et du modèle BV en traitement d'image.

D'un point de vue intuitif, la parcimonie des représentations en ondelettes d'objets bidimensionnels présentant des contours est limitée par l'*isotropie* du changement d'échelle : le nombre d'ondelettes interceptant un contour à la résolution 2^{-j} est de l'ordre de 2^j . Afin de briser cette limite, il est nécessaire de mettre au point d'autres représentations dans lesquelles les changements d'échelles sont *anisotropes*, et qui peuvent en particulier tirer parti de la *régularité géométrique* des contours. Ainsi Emmanuel Candes et David Donoho proposent en 2000 la famille des *curvelets* qui ont la forme générale

$$\psi_{j,k,l}(x) = 2^{3j/2} \psi(D_j R_l x - k), \quad j \in \mathbb{Z}, \quad k \in \mathbb{Z}^2, \quad l = 0, \dots, 2^j - 1,$$

où ψ est une fonction de deux variables localisée en espace et en fréquence, et d'intégrale nulle dans la direction verticale, R_l est la matrice de rotation d'angle $2^{-j}l\pi$, et D_j est une matrice de dilatation diagonale anisotrope de facteur 2^j et 2^{2j}

dans les directions horizontale et verticale. Ils montrent que de telles familles constituent des frames de $L^2(\mathbb{R}^2)$ et que les coefficients d'une image régulière par morceaux avec des contours géométriquement réguliers appartiennent à l'espace $w\ell^p$ pour tout $p > 2/3$, améliorant ainsi le degré de parcimonie obtenu avec les ondelettes.

Dans la foulée, de nombreuses autres familles apparaissent avec les mêmes objectifs, à quelques variantes près qui peuvent se révéler importantes dans la pratique du calcul numérique. Citons en particulier la construction en 2005 des *bandelettes* par Erwan Le Pennec et Stéphane Mallat qui permettent d'améliorer ces résultats lorsque la régularité des contours augmente. Un concurrent naturel à ces approches est le raffinement de maillage anisotrope dans le cadre des méthodes d'éléments finis.

L'anisotropie joue aussi un rôle important lorsque l'on cherche à approcher numériquement des fonctions d'un nombre de variables $d \gg 1$. Les méthodes de discrétisation classiques, telles que l'approximation à partir des valeurs échantillonnées sur un réseau carré, se heurtent à une difficulté de complexité : par exemple, deux points de discrétisation par variable en dimension $d = 40$, donnent un nombre total de 2^{40} points, ce qui n'est pas gérable par l'ordinateur. L'idée des *sparse grids* souvent utilisée par les numériciens confrontés aux problèmes de grande dimension, est directement héritée des bases d'ondelettes : on considère la base anisotrope obtenue par tensorisation

$$\Psi_{j,k}(x) = \psi_{j_1,k_1}(x_1) \cdots \psi_{j_d,k_d}(x_d), \quad j = (j_1, \dots, j_d), \quad k = (k_1, \dots, k_d)$$

et on retient uniquement les indices tels que $j_1 + \dots + j_d \leq J$, où J pilote la précision recherchée, ce qui signifie qu'une résolution fine en une variable est compensée par une résolution plus grossière pour les autres variables.

La recherche de représentations parcimonieuses des signaux et des images conduit aussi à remplacer une base orthonormée fixée, par une famille de plusieurs bases, ou plus généralement par un *dictionnaire* redondant. En augmentant la taille du dictionnaire, l'objectif est de mieux capturer les fonctions d'intérêt avec très peu de termes. Les exemples caricaturaux tels qu'un dictionnaire contenant toutes les fonctions, et dans lequel n'importe quelle fonction est capturée par un seul terme, sont rapidement disqualifiés par des considérations de théorie de l'information et de calcul : la taille du dictionnaire entre en jeu lorsque l'on souhaite pouvoir coder ses éléments par un nombre raisonnable de bits, ou effectuer la recherche d'un élément en un temps limité. En pratique, on considère par exemple des dictionnaires \mathcal{G} contenant des atomes temps-fréquences et temps-échelles, ou d'autres fonctions élémentaires. Puisqu'il ne s'agit plus de bases, la représentation d'une fonction f suivant

$$f = \sum_{g \in \mathcal{G}} c_g g,$$

n'est pas unique. Le problème qui se pose est alors le suivant : si f admet une représentation exacte ou approchée utilisant un petit nombre d'éléments de \mathcal{G} , comment identifier ceux-ci et leurs coefficients ? L'approche classique des moindres carrés, qui consiste à minimiser $\sum_{g \in \mathcal{G}} |c_g|^2$ parmi toutes les suites (c_g) telles que $\sum_{g \in \mathcal{G}} c_g g = f$, conduit à une représentation qui n'est pas parcimonieuse. Ce

problème se relie à une question souvent considérée en statistique : identifier un petit nombre de variables *explicatives* du phénomène observé.

En 1998, David Donoho propose l'approche du *basis pursuit*, qui consiste à remplacer le critère des moindres carrés par la somme des modules $\sum_{g \in \mathcal{G}} |c_g|$, c'est-à-dire minimiser la norme ℓ^1 sous la contrainte $\sum_{g \in \mathcal{G}} c_g g = f$. Il constate, et démontre dans certains cas, que la solution de norme ℓ^1 minimale est aussi la plus parcimonieuse, c'est-à-dire celle qui contient le plus petit nombre de termes. Ces idées aboutissent à partir de 2005 aux travaux d'Emmanuel Candes et Terence Tao qui ouvrent la voie au domaine actuellement très actif du *compressed sensing*. L'un des résultats marquant du *compressed sensing* affirme que si x est un vecteur de taille $N \gg 1$ qui ne contient que k composantes non-nulles en des emplacements inconnus, il est possible de reconstruire exactement ce vecteur à partir d'un vecteur y de m mesures linéaires

$$y = \Phi x,$$

où Φ est une matrice $m \times n$ et m est de l'ordre de $k \log(N/k)$. L'algorithme utilisé est précisément le *basis pursuit*, c'est-à-dire la minimisation $\|z\|_{\ell^1}$ sous la contrainte $\Phi z = y$. Le point clé réside dans le choix de la matrice de mesure Φ , les seules constructions optimales étant actuellement obtenues par des techniques probabilistes. Ces résultats sont fortement liés à ceux obtenus depuis les années 1970 en géométrie des espaces de Banach sur les épaisseurs des convexes, tout en partant d'une motivation concrète en traitement du signal et de l'image issue de la recherche de représentations parcimonieuses. Signalons aussi le lien récemment établi par Yves Meyer et Basarab Matei entre le *compressed sensing* et la *théorie des modèles*, introduite à la fin des années 1960 par Yves Meyer et qui fournit la description mathématique des quasi-cristaux.

Tant par leurs fascinantes propriétés que par leurs limitations intrinsèques, les ondelettes ont ainsi stimulé l'activité scientifique dans de nombreuses directions.

7. Un peu de lecture

Deux références fondamentales sur la théorie des ondelettes :

- [1] Y. MEYER, *Wavelets and operators*, vol. 1 and 2, revised version of 3, Cambridge university press, 1992-1997.
- [2] I. DAUBECHIES, *Ten lectures on wavelets*, SIAM, Philadelphia, 1992.

Un ouvrage orienté vers les applications des ondelettes :

- [3] S. JAFFARD, Y. MEYER and R. RYAN, *Wavelets, tools for science and technology*, SIAM, Philadelphia, 2001.

Autour des applications au traitement du signal et de l'image, et de la notion de parcimonie :

[4] S. MALLAT, *A wavelet tour of image processing. Third edition : the sparse way*, Academic Press, 2008.

Autour des applications à la simulation numérique :

[5] W. DAHMEN, *Wavelets and operator equations*, Acta Numerica, 1997.

[6] A. COHEN, *Numerical analysis of wavelet methods*, Elsevier, 2003.

Une référence sur l'approximation non-linéaire :

[7] R. DEVORE, *Nonlinear approximation*, Acta Numerica, 1998.

Autour de l'analyse de l'espace BV et des phénomènes oscillants :

[8] Y. MEYER, *Oscillating patterns in image processing and in some nonlinear evolution equations*, Jacqueline Lewis Lectures, AMS, 2001.