

Fractales ?
Cours de DEA SIPT
P.O. Amblard



Vraie, ... ou fausse ?



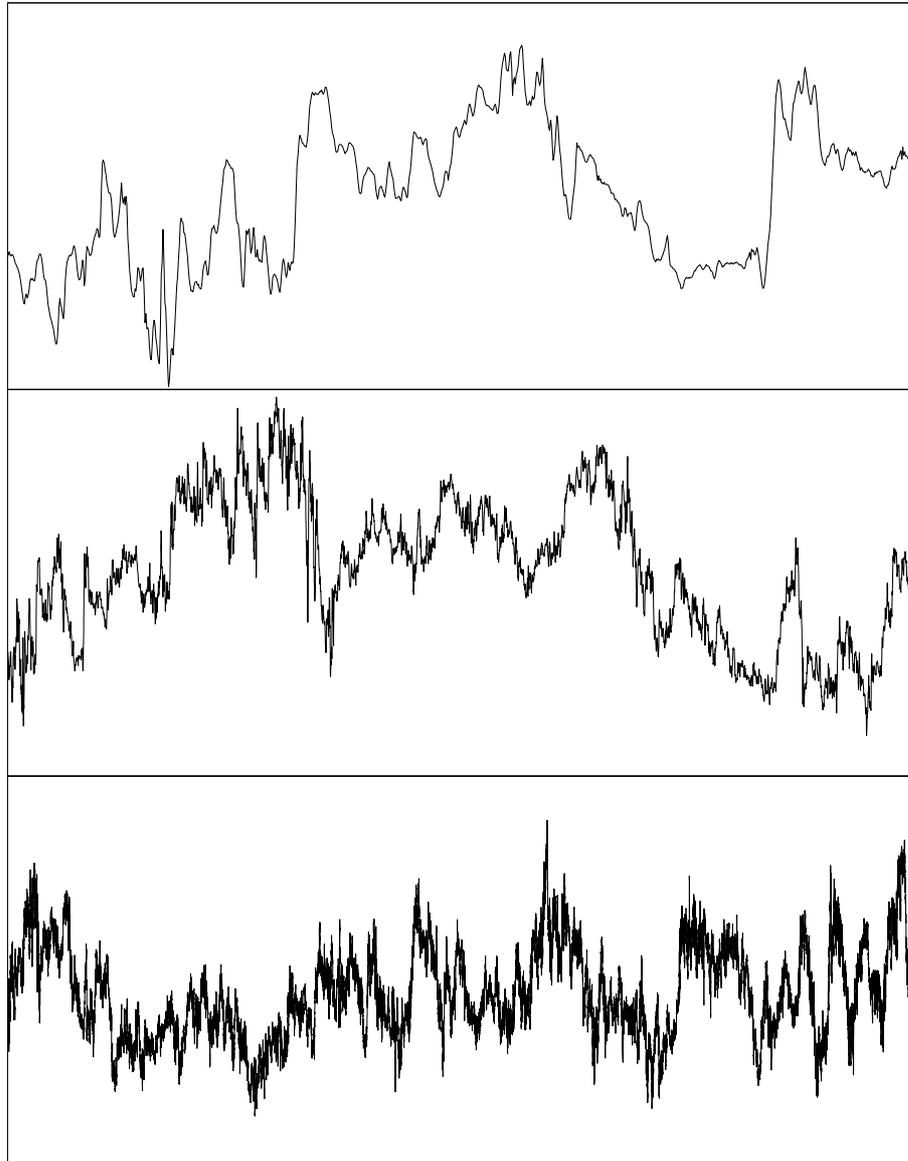
Paysages...





*Les termites se promènent...
et les plantes...*





Un peu de turbulence...

Chapter 1

Introduction aux fractales

1.1 Mesurer des longueurs

Tous les étudiants possèdent une règle graduée qui leur permet de mesurer la longueur de n'importe quel objet. Le principe de base est de considérer la règle comme une unité de longueur l_0 , puis de compter le nombre de fois nécessaire pour couvrir l'objet avec cette unité. Soit $n(l_0)$ ce nombre. La longueur est alors approchée par $L \equiv n(l_0)l_0$. Pour un segment de droite de longueur L , cette formule est exacte si l'on choisit $l_0 = L/n(l_0)$. Pour un cercle de rayon R , divisons en n angles égaux le cercle, et choisissons pour l'unité de longueur la longueur de la corde liant deux points du cercle définissant l'angle $2\pi/n$. Alors $l_0 = 2R \sin(\pi/n)$. Or, il faut $n(l_0) = n$ unité de longueur pour décrire le cercle. Donc la longueur du cercle est approchée par $L \equiv n2R \sin(\pi/n)$. Pour obtenir une mesure fine, n est choisit de plus en plus grand. En développant le sinus on obtient $2R \sin(\pi/n) \approx 2R\pi/n - R\pi^3/(3n^3)$ de sorte que $n(l_0)l_0 \approx 2R\pi - R\pi^3/(3n^2)$, qui tend vers $2\pi R$ quand n tend vers l'infini.

Considérons maintenant l'exemple de la figure (1.1). Cette courbe est appelée courbe de Von Koch, et est obtenue à l'aide d'un processus récursif. L'élément initial est un segment de droite de longueur l_0 . La récursion consiste à remplacer le tiers central du segment par un triangle équilatéral de côté $l_0/3$. La courbe résultante est composée de 4 segments sur lesquels on remplace le tiers central par un triangle équilatéral de côté $l_0/9$. Ce processus est alors répété à l'infini. La limite est appelée courbe de Von Koch. Calculons sa longueur. Initialement, la longueur est $n_0 l_0$ avec $n_0 = 1$. A la première itération, elle est de $n_1 l_0/3$ où $n_1 = 4n_0$. A chaque itération, le nombre de segments est multiplié par 4, et chaque nouveau segment a une longueur trois fois plus petite : donc, $n_k = 4n_{k-1} = 4^k n_0$ et $l_k = l_0/3^k$. La longueur de la courbe à l'itération k est donc $L_k = l_0(4/3)^k$, et L_k tend vers l'infini lorsque k tend vers l'infini. La courbe de Von Koch est donc une courbe continue de longueur infinie!

D'une façon générale, lorsque l'on mesure une longueur, le nombre $n(l_0)$ d'unités de longueur l_0 permet d'obtenir la longueur par

$$l = \lim_{l_0 \rightarrow 0} n(l_0)l_0$$

Pour un segment de droite de longueur L , il faut environ L/l_0 unités de longueur pour couvrir le segment, soit $d = 1$, de sorte que $l = \lim_{l_0 \rightarrow 0} n(l_0)l_0 = L$. D'une façon générale, dès qu'une courbe est rectifiable (dérivable, sauf peut-être en un nombre fini de points), cette limite est finie et appelée longueur de la courbe.

Mais pour la grande majorité des courbes, la rectifiabilité n'est pas vérifiée, et la limite ci-dessus s'avère souvent infinie. Mathématiquement, la limite peut être nulle si $d < 1$, finie si $d = 1$ ou infinie si $d > 1$. Nous avons déjà vu des exemples de courbes rectifiables pour lesquelles la limite est finie. Le cas d'une longueur nulle correspond à un objet qui n'est en fait qu'une poussière de points, et qui porte donc difficilement le nom de courbe (voir l'exemple de l'ensemble de Cantor au

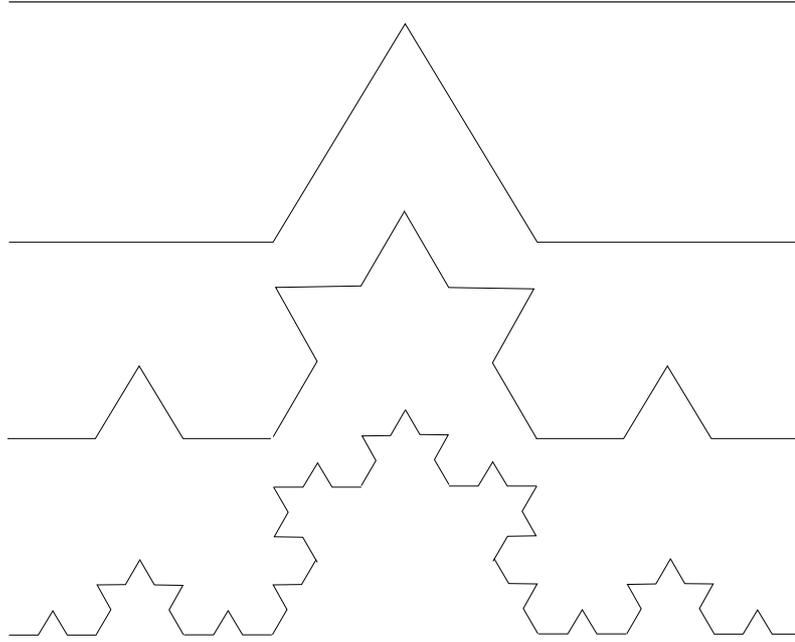


Figure 1.1: Courbe de Von Koch : exemple d'une courbe continue de longueur infinie.

paragraphe 1.1.2, représenté figure 1.2). L'exemple de la courbe de Von Koch pour une longueur infinie est représentatif. Lorsque l'étalon de longueur l_0 diminue, des détails fins apparaissent et peuvent être pris en compte dans le nombre $n(l_0)$. Le premier exemple pris dans les exposés de B. Mandelbrot est celui des côtes maritimes d'un pays (les côtes de Bretagne pour être précis). Lorsque l'étalon de longueur est de l'ordre de la centaine de mètres, les échancrures mesurables sont les grandes baies. A l'échelle de la dizaine de mètres, de nouvelles petites baies deviennent visibles par l'étalon. Par contre, les circonvolutions dues aux rochers ne sont pas prises en compte. Elles le deviennent lorsque l'étalon est de l'ordre du mètre. Chaque échelle fait ainsi apparaître de nouveaux détails, petits mais en nombre suffisant pour faire diverger la mesure de longueur.

En général, le nombre $n(l_0)$ de fois qu'il faut déplacer l'étalon de longueur pour parcourir la courbe se comporte comme l_0^{-d} où d est caractéristique de la courbe. Dans le cas des courbes rectifiables, nous avons noté que $n(l_0) \propto l_0^{-1}$. Si $d < 1$, la longueur tend vers 0 et nous sommes confronté à une poussière alors que si $d > 1$ la longueur de la courbe est infinie.

Ce nombre d est caractéristique de la courbe. Or, en mesurant la longueur de la courbe, nous n'avons accès qu'à la position de d par rapport à 1. Ceci est dû au fait que la mesure de longueur n'est pas adaptée à la mesure de d . Toutefois, si l'on envisage

$$l(D) = \lim_{l_0 \rightarrow 0} n(l_0) l_0^D$$

alors

$$n(l_0) l_0^D \approx l_0^{D-d} \begin{cases} \rightarrow 0 & \text{si } D > d \\ \rightarrow \text{cte} & \text{si } D = d \\ \rightarrow \infty & \text{si } D < d \end{cases}$$

Le nombre d est appelé dimension fractale de la courbe.

Le problème du calcul d'une dimension est très compliqué. Ceci est dû à l'existence d'une multitude de définitions de la dimension d'un ensemble. Nous allons donner la définition qui semble la plus générale, puis voir une autre définition, plus constructive, permettant un calcul pratique.

1.1.1 Dimension de Hausdorff-Besicovitch

Nous nous plaçons dans \mathbb{R}^2 , bien que cette restriction puisse être levée. Soit S un sous-ensemble de \mathbb{R}^2 . Soit maintenant une famille finie ou dénombrable d'ensembles $\{U_i\}$ de diamètres inférieurs ou égaux à δ , *i.e.* $|U_i| = \sup(|x - y|, x, y \in U_i) \leq \delta$ qui couvre S . En d'autres termes, la famille $\{U_i\}$ est telle que

$$\begin{aligned} |U_i| &\leq \delta, \forall i \\ S &\subset \bigcup_{i=0}^{\infty} U_i \end{aligned}$$

On dit alors que la famille $\{U_i\}$ est un δ -recouvrement de S . On pourrait alors mesurer la surface de S en sommant les diamètres des U_i . Mais un problème a déjà été évoqué dans le paragraphe précédent : pour certains ensembles, la mesure peut tendre vers l'infini. Il est alors préférable de sommer les diamètres élevés à une puissance s et trouver le s rendant finie la somme. Un deuxième problème réside dans le fait que cette mesure dépend du recouvrement choisi. Pour cette raison, il faut envisager l'ensemble de tous les δ -recouvrements possibles. On définit alors

$$\mathcal{H}_\delta^s(S) = \inf \left\{ \sum_{i=0}^{+\infty} |U_i|^s, \{U_i\} \text{ } \delta\text{-recouvrement de } S \right\}$$

Tout comme pour mesurer les longueurs dans le paragraphe précédent, on envisage ensuite la limite lorsque δ tend vers zéro de $\mathcal{H}_\delta^s(S)$ et l'on écrit

$$\mathcal{H}^s(S) = \lim_{\delta \rightarrow 0} \mathcal{H}_\delta^s(S)$$

On montre que $\mathcal{H}^s(S)$ est une mesure de l'ensemble S qui peut être éventuellement infinie. Cette mesure est appelée mesure de Hausdorff. Pour comprendre son comportement, supposons que S soit le graphe d'une fonction lisse. Si $s > 1$, on mesure la surface de S qui est nulle. Si $s < 1$ on mesure une grandeur d'ordre plus petit qu'une longueur et le résultat est infini. Pour $s = 1$, on mesure la longueur de la courbe lisse, et l'on obtient un résultat fini. On trouve alors que la mesure de Hausdorff est en général

$$\mathcal{H}^s(S) = \begin{cases} \infty & \text{pour } s \in [0, s_d[\\ \mathcal{H}^{s_d}(S) & \text{pour } s = s_d \\ 0 & \text{pour } s > s_d \end{cases}$$

Le nombre s_d tel que la mesure de Hausdorff soit finie est appelée dimension de Hausdorff de l'ensemble S et est noté $\dim_H(S)$.

On obtient une définition équivalente si l'on remplace les ensembles U_i par des disques par exemple.

Il existe d'autres définitions de la dimension. Le caractère pratique de ces dimensions n'est pas toujours évident. Le paragraphe suivant donne une définition permettant une évaluation simple de la dimension.

1.1.2 Dimension de "comptage de boîtes"

L'idée derrière le calcul de dimension est le recouvrement par des objets simples de l'ensemble à mesurer. La dimension de Hausdorff est fondée sur cette idée, mais est difficilement utilisable en pratique puisque la forme du recouvrement n'est pas explicite.

Dans \mathbb{R}^2 , on considère les carrés de côtés de longueur δ . Soit $N_\delta(S)$ le nombre minimal de ces carrés nécessaire pour recouvrir un ensemble S . La dimension de "comptage de boîtes" de S est alors définie par la limite, lorsqu'elle existe

$$\dim_B(S) = \lim_{\delta \rightarrow 0} -\frac{\log(N_\delta(S))}{\log \delta}$$

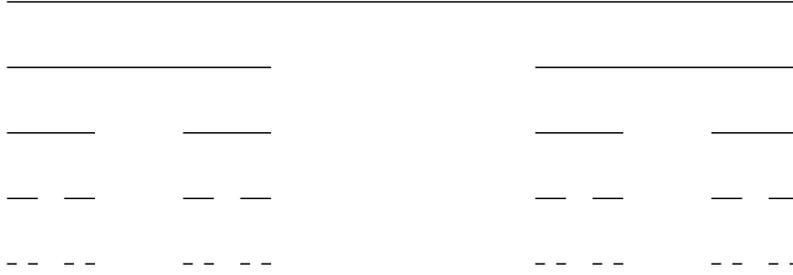


Figure 1.2: Ensemble de Cantor.

Cette dimension porte un certain nombre de noms, parmi lesquels dimension métrique, d'information, de capacité. . .

On peut montrer que la dimension de comptage de boîtes est supérieur ou égale à la dimension de Hausdorff. Toutefois, pour beaucoup d'ensembles, ces dimensions sont égales.

L'avantage d'une telle définition est son utilité pratique, puisqu'il est très facile de l'implanter sur ordinateur. La marche à suivre est la suivante. On quadrille le plan par des carrés de longueur δ_1 et on compte le nombre $N_{\delta_1}(S)$ de ces carrés qui intersectent l'ensemble à mesurer. Puis on diminue δ_1 en δ_2 pour chercher $N_{\delta_2}(S)$, etc. . .

On trace alors $\log(N_{\delta_i}(S))$ en fonction de $\log(\delta_i)$. Pour des valeurs $\log(\delta_i)$ suffisamment petites, ce graphe doit être une droite dont on calcule la pente par régression linéaire.

Exemple : l'ensemble de Cantor L'ensemble de Cantor est obtenu en éliminant le tiers central du segment $[0, 1]$ puis en itérant le processus sur les segments restant. Cette construction est montré figure (1.2).

Pour couvrir l'ensemble de Cantor, on choisit des boîtes de taille $\delta = 1/3^n$. A l'étape n , il faut $N_\delta = 2^n$ boîtes pour recouvrir complètement l'ensemble. Donc

$$-\frac{\log(N_\delta)}{\log \delta} = \frac{\log 2}{\log 3} = \dim_B$$

On remarque également que le recouvrement envisagé est le δ -recouvrement qui donne "l'inf" pour la mesure de Hausdorff. Dans cet exemple, la dimension de Hausdorff et la dimension d'information sont donc égales.

Remarque Nous avons vu deux définitions de la dimension. La notion de dimension fractale est donc difficilement définissable. Toutefois, nous dirons qu'un objet a une dimension fractale si sa dimension de Hausdorff est strictement supérieure à sa dimension topologique. Pour la courbe de Von Koch, la dimension topologique est 1, la dimension de Hausdorff de $\log 4 / \log 3 \approx 1.26$. La courbe de Von Koch est donc fractale. L'ensemble de Cantor est un ensemble discret. Sa dimension topologique est donc $0 < \dim_H \approx 0.64$: l'ensemble de Cantor est donc fractal.

1.1.3 Signaux fractals

Un signal est une fonction en général du temps. Par définition, le signal est dit fractal si le graphe $(t, f(t))$ de la fonction est un ensemble fractal de \mathbb{R}^2 .

La notion de fonction fractale est étroitement liée à la notion de fonctions Höldériennes.

Définition Une fonction $f : [0, 1] \rightarrow \mathbb{R}$ satisfait une condition de Hölder s'il existe $\alpha \in]0, 1[$ tel que

$$|f(t) - f(u)| \leq |t - u|^\alpha, \forall t, u \in [0, 1]$$

Pour ces fonctions, la dimension de Hausdorff du graphe de f est en générale inférieure ou égale à la dimension d'information, elle même en général inférieure ou égale à $2 - \alpha$. Toutefois, pour un bon nombre de fonctions on a le résultat $\dim_B(\text{graphe } f) = 2 - \alpha$. La fonction suivante en est un exemple célèbre.

Fonction de Weierstrass Considérons $\lambda > 1$ et $\alpha \in]0, 1[$. La fonction de Weierstrass est définie sur $[0, 1]$ par

$$f(t) = \sum_{k=1}^{+\infty} \lambda^{-\alpha k} \sin(\lambda^k t)$$

et a un graphe fractal (lorsque λ est suffisamment grand). La figure (1.3) montre quelques unes de ces fonctions pour différentes valeurs de α . La dimension d'information est $\dim_B = 2 - \alpha$. On voit sur la figure que le graphe de f est de plus en plus "chahuté" au fur et à mesure que α diminue. Autrement dit, le graphe de la fonction tend à recouvrir de plus en plus le plan, et sa dimension fractale augmente de plus en plus.

Cet exemple permet également de voir une caractéristique générale des fonctions fractales : leur comportement fréquentiel. En effectuant la transformée de Fourier de la fonction de Weierstrass on obtient pour les fréquences positives

$$\begin{aligned} |F(\nu)| &= \frac{1}{2} \sum_{k=1}^{+\infty} \lambda^{-\alpha k} \delta\left(\nu - \frac{\lambda^k}{2\pi}\right) \\ &= \frac{1}{2} \sum_{k=1}^{+\infty} a_k \delta(\nu - \nu_k) \end{aligned}$$

On s'aperçoit qu'il y a un lien simple entre l'amplitude a_k et la fréquence ν_k . En effet on peut écrire $a_k = (2\pi)^{-\alpha} 1/(\nu_k)^\alpha$. Les amplitudes suivent une loi de puissance en fonction de la fréquence. Cet exemple montre ce fait sur un nombre discret de fréquence, mais la remarque reste en général valable pour un spectre continu de fréquence.

Un signal $x(t)$ présentant une transformation de Fourier en $1/\nu^\alpha$ ($\alpha \in]0, 1[$) est appelé signal en $1/f$ et est en général fractal de dimension $2 - \alpha$. On dit aussi que le signal satisfait à une invariance d'échelle car ces signaux vérifient $x(at) = a^{\alpha-1}x(t)$: changer l'échelle de temps ne change rien au signal (à une normalisation près). Ces signaux ne possèdent donc pas d'échelle de temps caractéristique!

Nous verrons d'autres exemple de fonctions fractales dans les paragraphes 1.2.8 et 1.3.

1.2 Un exemple de construction de courbes fractales : systèmes de fonctions itérés

Les systèmes de fonctions itérés (IFS) constituent une façon de construire des fractales. Leur principe repose sur le théorème du point fixe dans les espaces métriques. Quelques rappels sur les espaces métriques sont maintenant donnés.

1.2.1 Rappels sur les espaces métriques

Soit X un ensemble. Une distance sur cet ensemble est une application $d : X \times X \rightarrow \mathbb{R}$ vérifiant les propriétés suivantes :

1. symétrie : $d(x, y) = d(y, x), \forall x, y \in X$,
2. positivité : $d(x, y) > 0, \forall x, y \in X, x \neq y$,

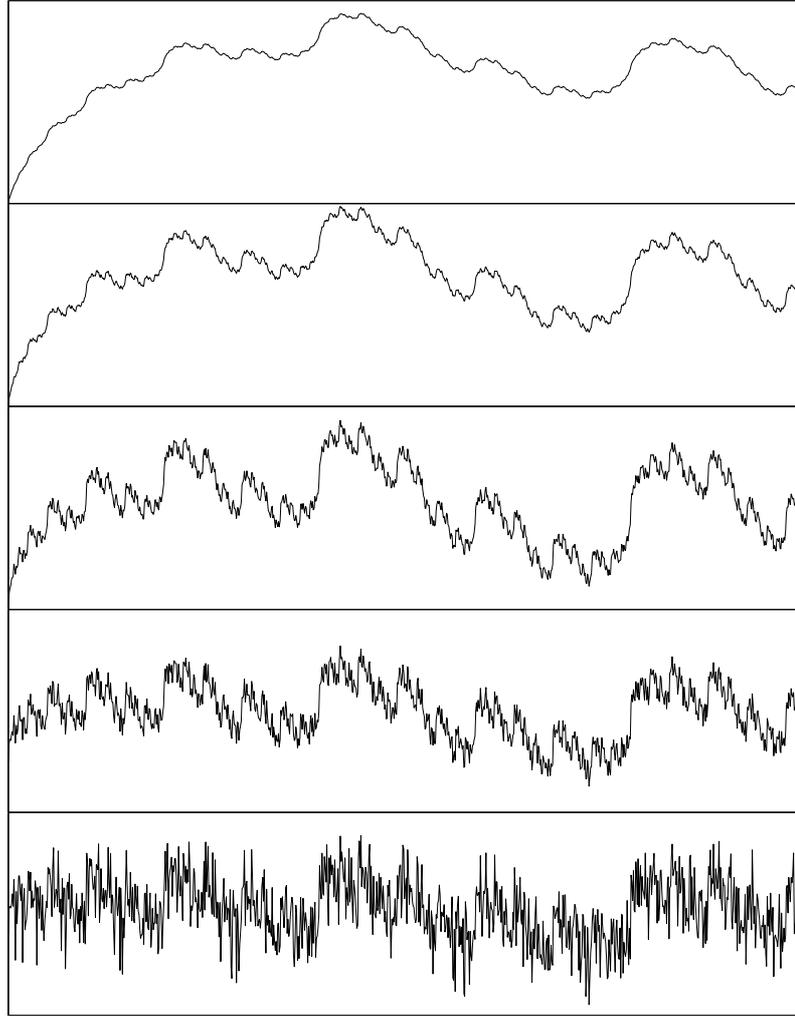


Figure 1.3: Fonction de Weierstrass. De haut en bas, $\alpha = 0.9, 0.7, 0.5, 0.3, 0.1$ de sorte que les dimensions fractales sont de haut en bas de 1.1, 1.3, 1.5, 1.7, 1.9.

3. $d(x, x) = 0, \forall x \in X,$

4. inégalité triangulaire : $d(x, y) \leq d(x, z) + d(z, y), \forall x, y, z \in X.$

Alors (X, d) est appelé *espace métrique*. La distance ou métrique d permet de conférer une structure topologique à l'ensemble X . En particulier, la notion de convergence se déduit de la notion de distance. Rappelons que deux distances d_1 et d_2 sont équivalentes si il existe $0 < c_1 < c_2 < +\infty$ tels que

$$c_1 d_2((x_1, y_1), (x_2, y_2)) \leq d_1((x_1, y_1), (x_2, y_2)) \leq c_2 d_2((x_1, y_1), (x_2, y_2))$$

pour tout point de X .

Une suite $\{x_n\}_{n=0, \dots, \infty}$ de points de l'espace métrique (X, d) est une suite de Cauchy, si $\forall \varepsilon > 0$, il existe un entier $N > 0$ tel que $d(x_n, x_m) < \varepsilon$ pour tout $n, m > N$.

Une suite $\{x_n\}_{n=0, \dots, \infty}$ de points de l'espace métrique (X, d) converge vers un point x de X si, $\forall \varepsilon > 0$, il existe un entier $N > 0$ tel que $d(x_n, x) < \varepsilon$ pour tout $n > N$.

Une telle suite est dite convergente. On montre alors qu'une suite convergente est une suite de Cauchy. La réciproque est fautive en g n ral. Cette remarque conduit   :

Un espace m trique (X, d) est *complet* si toutes les suites de Cauchy de X convergent dans X .

Soit S un sous-ensemble de (X, d) . Un point $x \in X$ est un point limite de S s'il existe une suite de points de $S - \{x\}$ convergente vers x . On appelle fermeture de S , not e \bar{S} , l'union de S et de ses points limites. Un sous-ensemble S est dit ferm  si il est  gal   sa fermeture, $S = \bar{S}$.

Un sous-ensemble S de (X, d) est compact, si de toute suite de points de S on peut extraire une sous suite convergente dans S . Un sous-ensemble S de (X, d) est born  s'il existe un point a de X et $R > 0$ tels que $d(a, x) < R, \forall x \in S$.

Les sous-ensembles compacts de (X, d) sont les ferm s born s.

1.2.2 Espace m trique des ensembles compacts

On se place dans ce qui suit dans $X = \mathbb{R}^2$ bien que des formulations plus g n rale existent. On rend \mathbb{R}^2 m trique en le dotant d'une distance (qui peut  tre la distance euclidienne ou toute autre distance sur \mathbb{R}^2).

Les compacts de \mathbb{R}^2 sont les ensembles ferm s born s. Consid rons deux ensembles compacts A, B . On d finit la distance de Hausdorff entre ces deux ensembles comme suit (voir la figure (1.4)). Soit $d(x, B)$ la distance du point $x \in A$   l'ensemble B d finie par $d(x, B) = \min(d(x, y), y \in B)$. La distance de l'ensemble A   l'ensemble B est alors $d(A, B) = \max(d(x, B), x \in A)$. On ne prend pas le min dans cette d finition car alors la distance entre deux ensembles s'intersectant serait nulle. Le caract re d fini n cessaire   la d finition d'une distance ne serait donc pas obtenu. Mais il reste toutefois un probl me avec cette d finition : la mesure est non sym trique. On d finit alors la distance de Hausdorff entre A et B par

$$d_H(A, B) = \max(d(A, B), d(B, A))$$

Cette d finition satisfait les propri t s d'une distance et d_H constitue alors une m trique sur l'ensemble des ensembles compacts de \mathbb{R}^2 . On peut alors montrer le r sultat suivant :

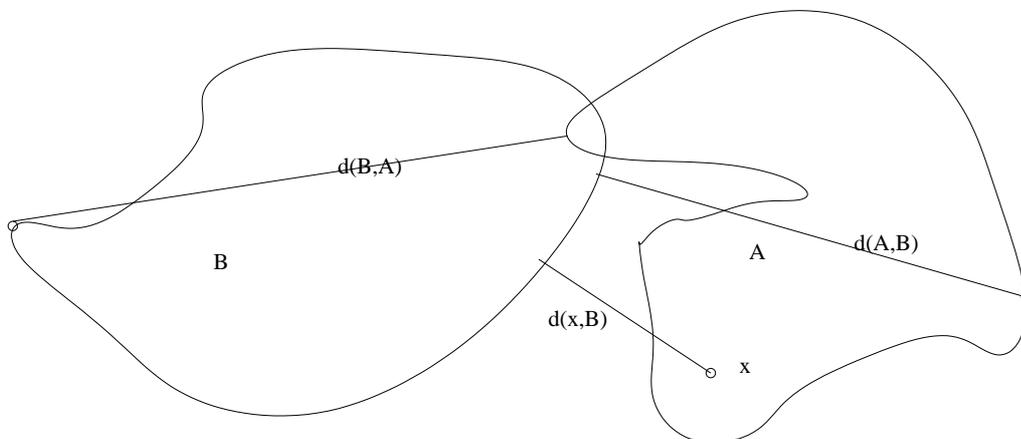


Figure 1.4: Illustration de la d finition de la distance de Hausdorff.

L'ensemble des compacts de \mathbb{R}^2 , not  $\mathcal{H}(\mathbb{R}^2)$, muni de la distance de Hausdorff est un espace m trique complet.

Ainsi, une suite de Cauchy de sous-ensembles compacts de \mathbb{R}^2 converge vers un sous-ensemble compact de \mathbb{R}^2 . Nous verrons que beaucoup de fractales sont des limites de ce type!

1.2.3 Applications contractantes et théorème du point fixe

Une transformation $f : X \rightarrow X$, X étant un espace métrique muni de d , est *contractante* s'il existe un réel $s \in [0, 1[$ tel que

$$d(f(x), f(y)) \leq sd(x, y), \forall x, y \in X$$

On dispose alors du théorème du point fixe :

Théorème Soit $f : X \rightarrow X$ une application contractante de (X, d) , espace métrique complet. Alors, f a un unique point fixe $x_f \in X$, et de plus, $\forall x \in X$, $\lim_{n \rightarrow \infty} f^{(n)}(x) = x_f$.

Démonstration : Soit f contractante de facteur de contraction s . Soit $x \in X$. Soient n, m deux entiers. Alors

$$d(f^{(n)}, f^{(m)}) \leq sd(f^{(n-1)}, f^{(m-1)}) \leq s^{\min(n,m)} d(x, f^{|m-n|})$$

Or, d'après l'inégalité triangulaire,

$$\begin{aligned} d(x, f^{(k)}) &\leq d(x, f(x)) + d(f(x), f^{(2)}(x)) + \dots + d(f^{(k-1)}(x), f^{(k)}(x)) \\ &\leq (1 + s + s^2 + \dots + s^{k-1})d(x, f(x)) \\ &\leq (1 - s)^{-1}d(x, f(x)) \end{aligned}$$

Il vient alors

$$d(f^{(n)}, f^{(m)}) \leq sd(f^{(n-1)}, f^{(m-1)}) \leq s^{\min(n,m)}(1 - s)^{-1}d(x, f(x))$$

Le membre de droite de cette inégalité peut être rendu aussi petit que l'on veut, à condition de choisir n, m suffisamment grands. La suite $\{f^{(n)}(x)\}_{n=0, \dots, \infty}$ est donc de Cauchy. Comme l'espace est complet, cette suite converge dans X vers un point $x_f = \lim_{n \rightarrow \infty} f^{(n)}(x)$. Comme f est contractante, elle est continue, de sorte que

$$f(x_f) = f(\lim_{n \rightarrow \infty} f^{(n)}(x)) = \lim_{n \rightarrow \infty} f^{(n+1)}(x) = x_f$$

De plus, x_f est unique. En effet, si y_f est un autre point fixe, alors $d(x_f, y_f) = d(f(x_f), f(y_f)) \leq sd(x_f, y_f)$, et donc $d(x_f, y_f) \leq 0$ qui implique que $d(x_f, y_f) = 0$ et par suite $x_f = y_f$. ■

1.2.4 Systèmes de fonctions itérés sur \mathbb{R}^2

Nous allons maintenant utiliser le théorème du point fixe sur l'espace des ensembles compacts de \mathbb{R}^2 , $\mathcal{H}(\mathbb{R}^2)$.

Soit w une application contractante de \mathbb{R}^2 dans \mathbb{R}^2 , de facteur de contraction s . Alors, $w : \mathcal{H}(\mathbb{R}^2) \rightarrow \mathcal{H}(\mathbb{R}^2)$ définie par

$$w(B) = \{w(x), x \in B\}, \forall B \in \mathcal{H}(\mathbb{R}^2)$$

est une application contractante sur $(\mathcal{H}(\mathbb{R}^2), d_H)$

Soient alors N applications contractantes w_n , de facteurs de contractions s_n . Alors, $W : \mathcal{H}(\mathbb{R}^2) \rightarrow \mathcal{H}(\mathbb{R}^2)$ définie par

$$W(B) = \bigcup_{n=1}^N w_n(B), \forall B \in \mathcal{H}(\mathbb{R}^2) \quad (1.1)$$

est une application contractante de facteur de contraction $s = \max(s_n)$.

La donnée de l'espace métrique et des N applications, $\{\mathbb{R}^2, \{w_n\}_{n=1, \dots, N}\}$ définit un **système de fonctions itéré hyperbolique (IFS)** de facteur de contraction s . D'après le théorème du point fixe, cette application possède un unique point fixe qui est un compact de \mathbb{R}^2 . Ce point fixe sera appelé attracteur de l'IFS.

Remarque Le terme hyperbolique signifie que l'application définie par (1.1) est contractante avec facteur de contraction $s \in [0, 1[$. Ce terme sera omis par la suite. Notons toutefois que des IFS non hyperboliques existent. De plus, il est possible qu'un IFS ne soit pas hyperbolique par rapport à une distance d mais qu'il le soit pour une autre distance !

Les attracteurs des IFS sont en général des ensembles fractals dont la dimension D est donnée par la solution de

$$\sum_{n=1}^N |s_n|^D = 1 \quad (1.2)$$

où $D \in [0, 2]$.

Exemple : le tapis de Sierpinski Considérons les trois applications suivantes

$$\begin{aligned} w_1 \mathbf{r} &= \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix} \mathbf{r} \\ w_2 \mathbf{r} &= \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix} \mathbf{r} + \begin{pmatrix} 0 \\ 0.5 \end{pmatrix} \\ w_3 \mathbf{r} &= \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix} \mathbf{r} + \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} \end{aligned} \quad (1.3)$$

où \mathbf{r} est un point de \mathbb{R}^2 . Chacune de ces trois applications est une pure contraction par un facteur 0.5 suivie d'une translation. L'attracteur de l'IFS associé est montré sur la figure (1.5). La figure (1.6) montre les premiers termes de la suite $W^{(n)}(A_0)$ pour deux conditions initiales A_0 différentes. Comme prévu par le théorème du point fixe, ces deux suites convergent vers le même point que l'on appelle "tapis de Sierpinski". Le caractère fractal de cet ensemble est clairement vu, puisqu'une petite partie ressemble au tout. On montre à l'aide de l'équation (1.2) que sa dimension fractale est $\log(3)/\log(2) \approx 1.58$.

On comprend facilement sur cet exemple le mécanisme des IFS. En effet, le tapis de Sierpinski peut se décomposer en trois "sous"-triangles, homothétiques au tout (facteur 0.5) et déplacés par une certaine translation. Les trois applications définissant l'IFS réalisent ce montage !

1.2.5 Algorithme aléatoire pour la génération d'IFS

Implanter pratiquement un IFS demande un temps de calcul énorme, puisqu'à chaque étape il faut calculer la transformée d'une image par N applications et réaliser l'union des résultats. Mais il existe un algorithme plus rapide, aléatoire que nous décrivons maintenant.

Associons à l'IFS défini par $\{w_i\}_{i=1,\dots,N}$ un jeu de probabilités $\{p_i\}_{i=1,\dots,N}$. p_i est la probabilité pour que la variable aléatoire discrète k prenne la valeur i

Considérons alors la suite de points de \mathbb{R}^2 définie par

$$\begin{aligned} \mathbf{r}_0 & \\ \mathbf{r}_n &= w_{k_n} \mathbf{r}_{n-1} \end{aligned}$$

où $\{k_n\}_{n=1,\dots,\infty}$ est une suite de variables aléatoires discrètes, indépendantes, identiquement distribuées selon les probabilités $\{p_i\}_{i=1,\dots,N}$

On peut alors montrer que l'ensemble des points $\{\mathbf{r}_n\}$ converge vers l'attracteur de l'IFS !

Cet algorithme est donc très aisément implantable. Partant d'un point de \mathbb{R}^2 initial, on lui applique selon la loi de probabilité p_i l'une des applications de l'IFS. Ce procédé est alors répété récursivement, et la suite de points dessine l'attracteur de l'IFS

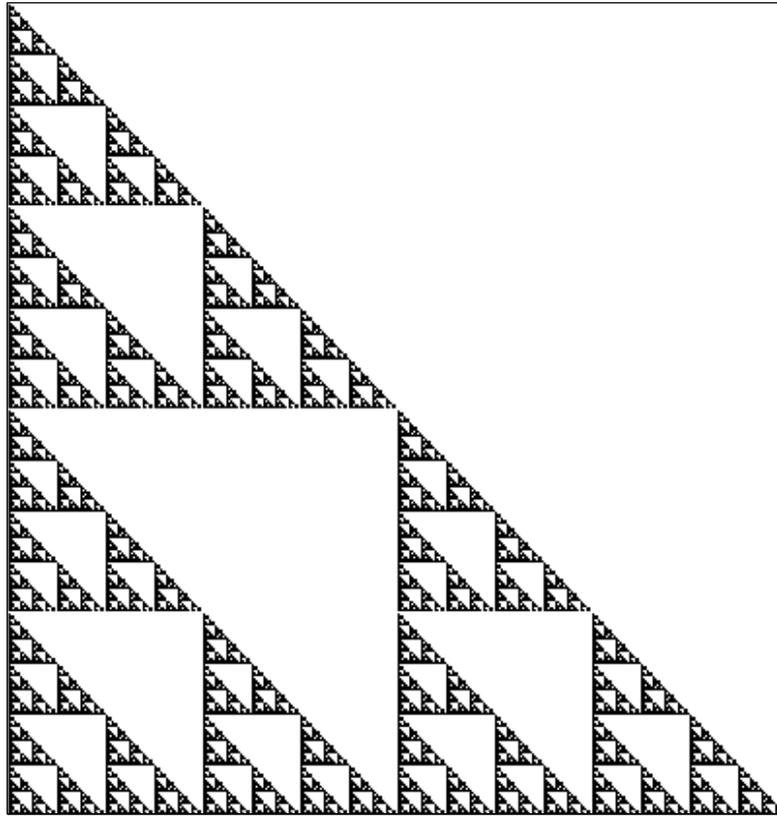


Figure 1.5: Tapis de Sierpinski : attracteur de l'IFS 1.3

Exemple : La feuille de Fougère Un exemple spectaculaire de la théorie des IFS concerne la feuille de Fougère représentée sur la figure (1.7) et calculée à l'aide de l'algorithme aléatoire. La feuille de fougère est l'attracteur de l'IFS défini par les quatres applications suivantes

$$\begin{aligned}
 w_1 \mathbf{r} &= \begin{pmatrix} 0 & 0 \\ 0 & 0.16 \end{pmatrix} \mathbf{r} \\
 w_2 \mathbf{r} &= \begin{pmatrix} 0.85 & 0.04 \\ -0.04 & 0.85 \end{pmatrix} \mathbf{r} + \begin{pmatrix} 0 \\ 1.6 \end{pmatrix} \\
 w_3 \mathbf{r} &= \begin{pmatrix} 0.2 & -0.26 \\ 0.23 & 0.22 \end{pmatrix} \mathbf{r} + \begin{pmatrix} 0 \\ 1.6 \end{pmatrix} \\
 w_4 \mathbf{r} &= \begin{pmatrix} -0.15 & 0.28 \\ 0.26 & 0.24 \end{pmatrix} \mathbf{r} + \begin{pmatrix} 0 \\ 0.44 \end{pmatrix}
 \end{aligned} \tag{1.4}$$

et son jeu de probabilités associé est $[0.01, 0.85, 0.07, 0.07]$. A chaque étape de l'algorithme, les applications w_1, w_2, w_3, w_4 sont choisies respectivement avec les probabilités 0.01, 0.85, 0.07 et 0.07.

Algorithme aléatoire et texture L'utilisation de l'algorithme aléatoire permet de texturer un ensemble fractal. Reprenons l'exemple du tapis de Sierpinski. La figure (1.8) montre cet effet pour

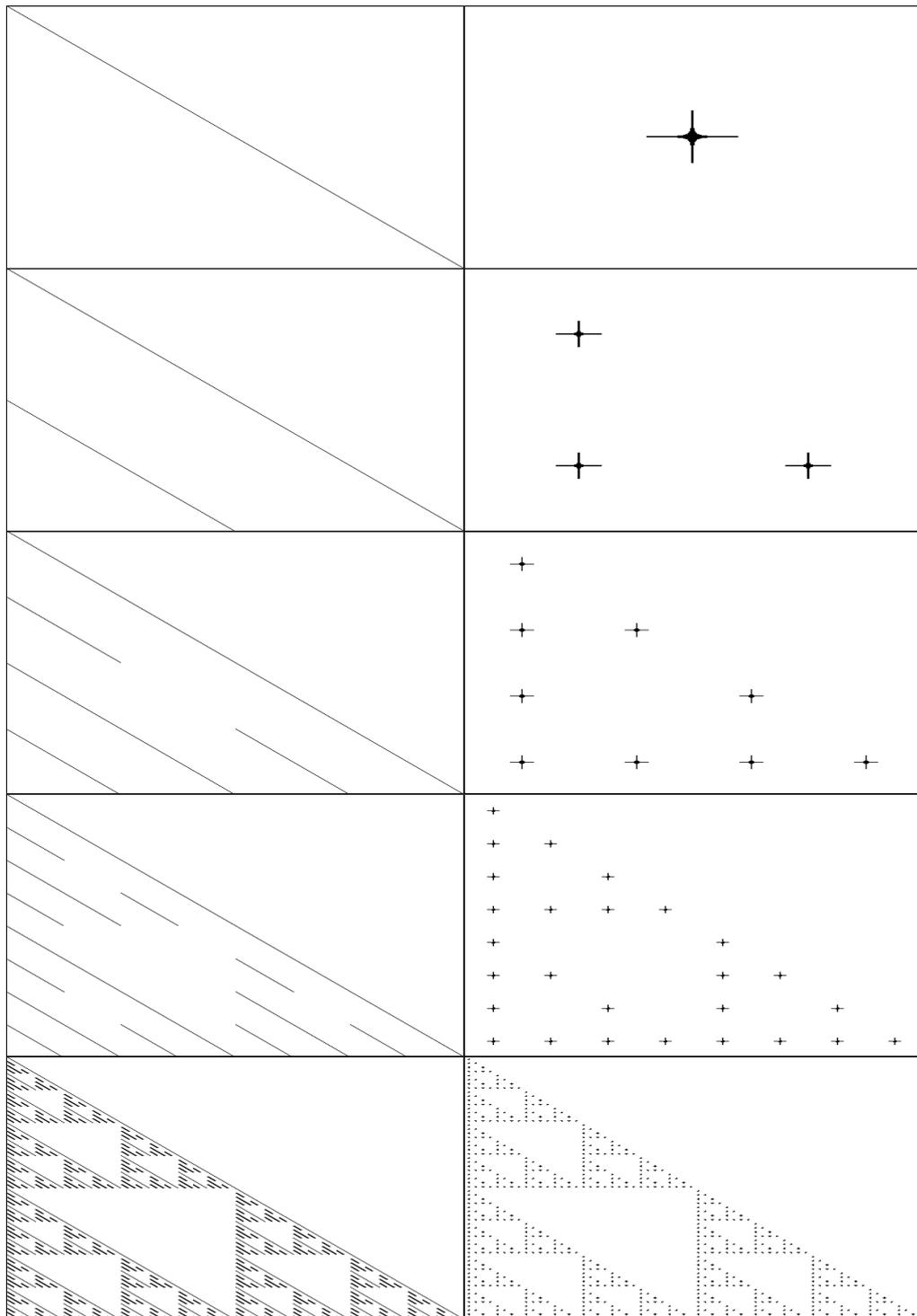


Figure 1.6: Tapis de Sierpinski : premiers termes de la suite définie par l'IFS 1.3 pour deux conditions initiales différentes. La dernière ligne correspond à la 7ème itération.



Figure 1.7: Feuille de fougère : attracteur de l'IFS défini par (1.4), calculé à l'aide de l'algorithme aléatoire. 20000 itérations.

quatre vecteurs de probabilités différents

$$\begin{aligned}
 \mathbf{p}_1 &= [0.5, 0.5, 0.5] \\
 \mathbf{p}_2 &= [0.4, 0.5, 0.1] \\
 \mathbf{p}_3 &= [0.2, 0.2, 0.6] \\
 \mathbf{p}_4 &= [0.01, 0.39, 0.6]
 \end{aligned}$$

En donnant des poids différents aux trois applications, on favorise l'une des trois zones principales du tapis de Sierpinski. Comme toute partie ressemble au tout, une de ces trois zones se découpe en trois plus petites zones parmi lesquelles une est favorisée, *etc.* . . .

On peut également voir cette texture comme la répartition d'une grandeur sur l'attracteur de l'IFS. Cette vision est à la base de la notion de multifractales.

1.2.6 IFS avec “condensation”

Une application de $w_0 : \mathcal{H}(\mathbb{R}^2) \rightarrow \mathcal{H}(\mathbb{R}^2)$ est appelée application avec condensation si $w_0(B) = C, \forall B \in \mathcal{H}(\mathbb{R}^2)$, où C est un élément de $\mathcal{H}(\mathbb{R}^2)$. Une application avec condensation est alors contractante avec un facteur de contraction nul.

On peut alors élargir la définition des IFS en ajoutant éventuellement une application avec condensation. Si l'IFS initial a un facteur de contraction s , l'ajout de w_0 ne le change pas.

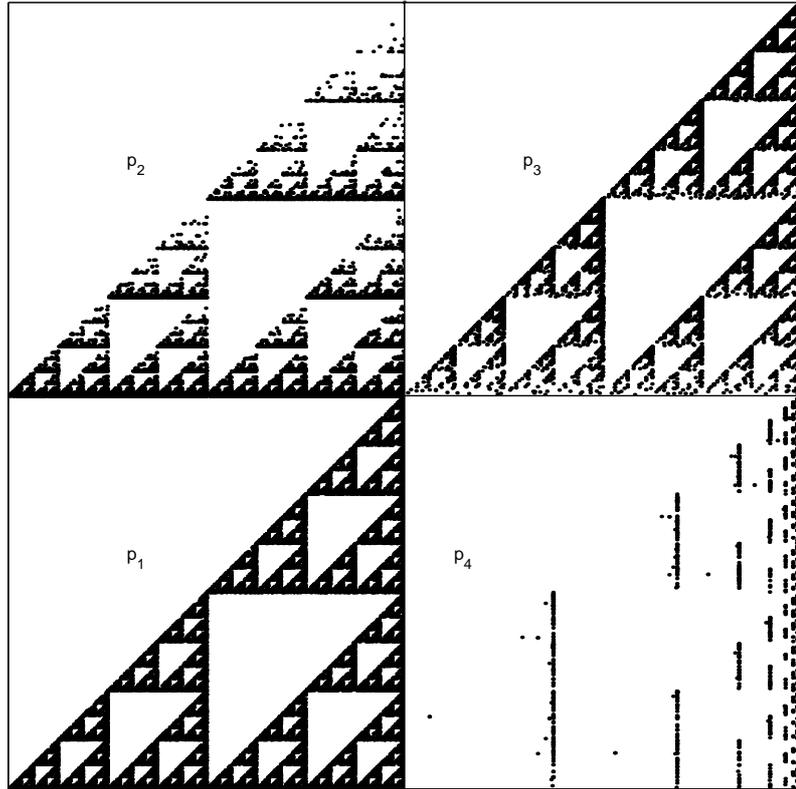


Figure 1.8: Tapis de Sierpinski texturé à l'aide de l'algorithme aléatoire. En variant les probabilités associées aux applications de l'IFS, on modifie la texture de l'attracteur.

A titre d'illustration, on se propose de construire un ifs dont le point fixe est un arbre, tel que celui dessiné sur la figure (1.9). Le tronc de l'arbre est l'ensemble de condensation C , et les deux contractions nécessaires sont composées d'une homothétie de rapport r , suivies de rotations d'angle $\pm\theta$, suivies d'une translation $[0, hauteurdutronc]$.

Notons que le choix de r et θ est soumis à des conditions si l'on souhaite que les branches de l'arbre ne se coupent pas. Un exemple est montré figure (1.10) pour $r = 0.6$ et $\theta = \pi/3$ et 9 itérations.

1.2.7 Le théorème du collage

La théorie des fractales essaie de mieux représenter la nature que ne le fait la géométrie classique. Il est alors intéressant de disposer d'outils permettant d'approcher des objets naturels par des objets mathématiques : peut-on par exemple trouver un algorithme qui dessine une feuille d'arbre? Nous avons vu qu'en ce qui concerne la fougère, la réponse à cette question est oui.

La théorie des IFS donne une réponse générale au problème inverse : étant donné un compact L de \mathbb{R}^2 , peut-on trouver un IFS dont l'attracteur soit le plus proche possible de L ? La solution est contenue dans le théorème du collage :

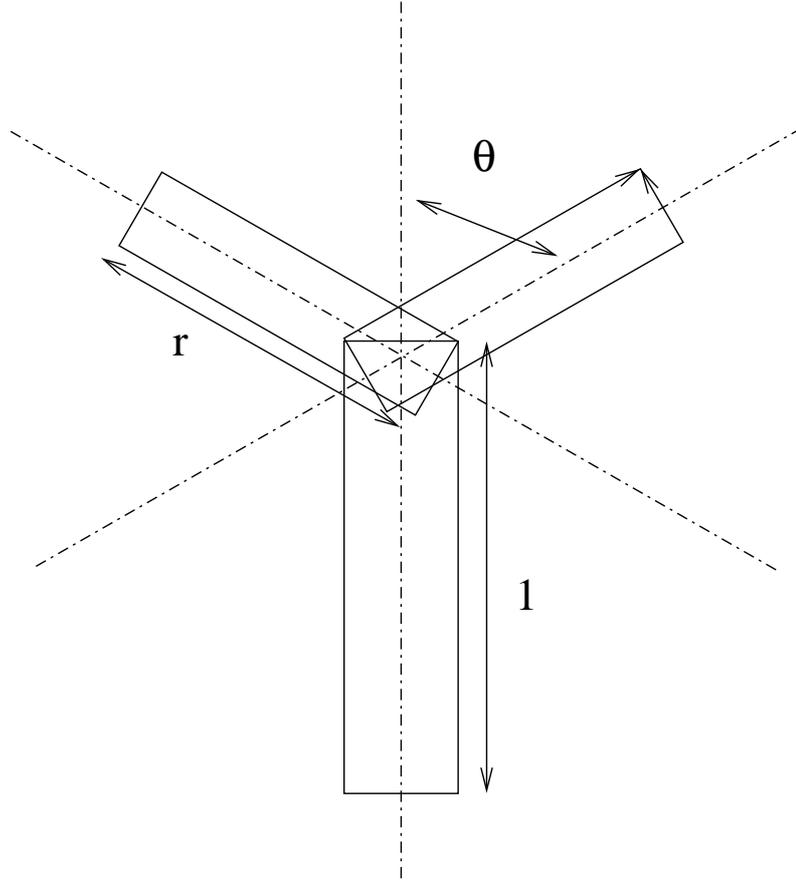


Figure 1.9: Arbre dont on cherche l'IFS avec condensation le g n rant.

Th or me On se place dans l'espace m trique (\mathbb{R}^2, d) . Soit $L \in (\mathcal{H}(\mathbb{R}^2), d_H)$. Soit un IFS $\{w_0, w_1, \dots, w_N\}$ de facteur de contraction $s = \max(s_n)$ tel que

$$d_H(L, \bigcup_{n=0}^N w_n(L)) \leq \varepsilon$$

Alors l'attracteur A de l'IFS v rifie

$$d_H(L, A) \leq \frac{\varepsilon}{1-s}$$

Le th or me du collage porte donc bien son nom : pour approcher un objet par l'attracteur d'un IFS, il suffit que l'union des images de l'objet par les applications de l'IFS soit proche de l'objet!

1.2.8 Cas des fonctions d'interpolation

Consid rions maintenant un ensemble $\{(x_i, y_i)\}_{i=0, \dots, N}$ de points que l'on cherche   interpoler. On peut utiliser la th orie des IFS pour r aliser une interpolation. On utilise pour cela N applications affines (bien que l'affinit  ne soit pas n cessaire). Les N applications r alisent une partition de l'intervalle $[x_0, x_N]$ en $[x_0, x_1] \cup [x_1, x_2] \cup \dots \cup [x_{N-1}, x_N]$. De plus, on veut que l'attracteur de

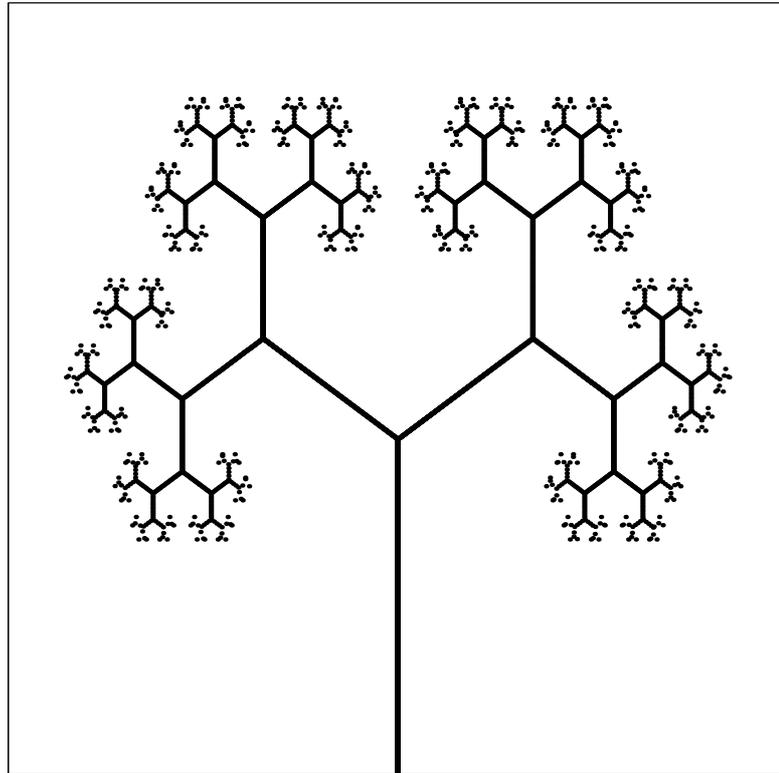


Figure 1.10: Attracteur de l'IFS avec condensation représentant un arbre.

l'IFS soit une fonction. Cette condition impose que les applications transforment les verticales en verticales, de sorte que

$$w_n \mathbf{r} = \begin{pmatrix} a_n & 0 \\ c_n & d_n \end{pmatrix} \mathbf{r} + \begin{pmatrix} e_n \\ f_n \end{pmatrix}$$

On impose alors à w_n d'envoyer l'intervalle $[x_0, x_N]$ sur l'intervalle $[x_{n-1}, x_n]$ et d'assurer la continuité de la fonction interpolante : w_n envoie (x_0, y_0) sur (x_{n-1}, y_{n-1}) et (x_N, y_N) sur (x_n, y_n) (figure 1.11). Ces conditions s'écrivent alors

$$\begin{aligned} x_{n-1} &= a_n x_0 + e_n \\ x_n &= a_n x_N + e_n \\ y_{n-1} &= c_n x_0 + d_n y_0 + f_n \\ y_n &= c_n x_N + d_n y_N + f_n \end{aligned}$$

On est en présence d'un système de quatre équations pour cinq inconnues. Une des inconnues peut être choisie arbitrairement et considérée comme paramètre. Le paramètre d_n règle le facteur de contraction vertical, puisque les verticales le restent par application de w_n . Ce paramètre va permettre de régler la rugosité des fonctions obtenues, et par suite leur dimension fractale.

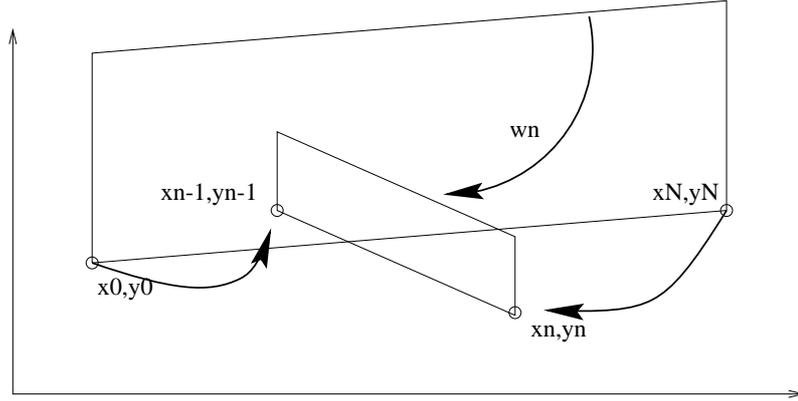


Figure 1.11: Principe de l'interpolation par IFS. L'application w_n envoie sur l'intervalle $[x_{n-1}, x_n]$.

Le système peut alors être résolu et conduit à

$$a_n = \frac{x_n - x_{n-1}}{x_N - x_0} \quad (1.5)$$

$$e_n = \frac{x_N x_{n-1} - x_0 x_n}{x_N - x_0} \quad (1.6)$$

$$c_n = \frac{y_n - y_{n-1} - d_n(y_N - y_0)}{x_N - x_0} \quad (1.7)$$

$$f_n = \frac{x_N y_{n-1} - x_0 y_n - d_n(x_N y_0 - x_0 y_N)}{x_N - x_0} \quad (1.8)$$

Un exemple d'interpolation est montré sur la figure (1.12). Les 11 points à interpoler proviennent de l'échantillonnage d'une sinusoïde. On s'aperçoit que la rugosité de la fonction interpolante diminue lorsque les d_n diminuent.

On peut montrer que la dimension fractale D de la fonction interpolante est la solution réelle de

$$\sum_{n=1}^N |d_n| a_n^{D-1} = 1$$

à condition que la somme des $|d_n|$ soit strictement supérieure à 1.

Cette relation est simple lorsque les x_n sont régulièrement espacés. En effet, on peut alors écrire $x_i = x_0 + i(x_N - x_0)/N$ de sorte que $a_n = 1/N$. Dans ce cas la dimension D est obtenue par

$$D = 1 + \frac{\log\left(\sum_{n=1}^N |d_n|\right)}{\log(N)}$$

Sur la figure (1.12), on obtient alors $D = 1$ pour $d_n = 0.1$, $D \approx 1.48$ pour $d_n = 0.3$ et $D \approx 1.9$ pour $d_n = 0.8$.

Sur la figure (1.13), les points $(0,0)$, $(0.5,1)$ et $(1,0)$ sont interpolés par cette technique, avec $d_1 = 0.6$, $d_2 = -0.7$ pour la courbe du panneau du haut et $d_2 = 0.7$ pour la courbe du panneau du bas. L'effet "d'étirement positif" dans le panneau du bas donne cette concavité moyenne vers le bas. Par contre, sur le panneau du haut, la contraction est alternativement positive et négative et provoque le caractère plus "chahuté" de la fonction. Notons toutefois que ces deux courbes ont même dimension fractale $D \approx 1.38$.

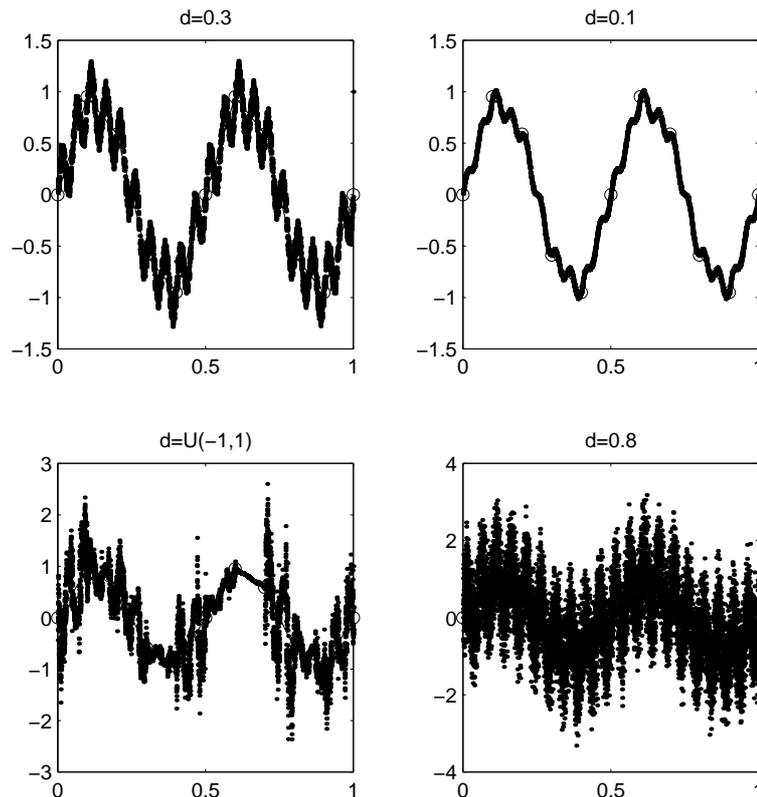


Figure 1.12: Interpolation de 11 points d'une sinusoïde par IFS. d_n est pris ici constant, sauf sur le cadran du bas à gauche où d_n suit une variable aléatoire uniformément répartie sur $[0, 1]$.

1.2.9 Aspects techniques

Dans ce paragraphe nous donnons les démonstrations qui prouvent la convergence du principe d'interpolation décrit avant.

Théorème 1 Soient $\{(x_n, y_n)\}_{n=1, \dots, N}$ $N > 1$ points de \mathbb{R}^2 et $\{\mathbb{R}^2, \{w_n\}_{n=1, \dots, N}\}$ l'IFS d'interpolation associée aux données, dont les paramètres vérifient (1.8). On suppose que les paramètres d_n satisfont $d_n \in [0, 1], \forall n = 1, \dots, N$. Il existe alors une métrique d équivalente à la métrique euclidienne sur \mathbb{R}^2 pour laquelle l'IFS considéré est hyperbolique. L'IFS possède alors un unique point fixe dans $\mathcal{H}(\mathbb{R}^2)$.

Démonstration : Soit $\theta \in \mathbb{R}^{*+}$ et $d_\theta((x_1, y_1), (x_2, y_2)) = |x_1 - x_2| + \theta|y_1 - y_2|$. Il est facile de montrer que d est une métrique sur \mathbb{R}^2 . Montrons qu'elle est équivalente à la distance euclidienne $d_e = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$. Rappelons que deux distances d_1 et d_2 sont équivalentes si il existe $0 < c_1 < c_2 < +\infty$ tels que

$$c_1 d_2((x_1, y_1), (x_2, y_2)) \leq d_1((x_1, y_1), (x_2, y_2)) \leq c_2 d_2((x_1, y_1), (x_2, y_2))$$

pour tout point de \mathbb{R}^2 .

Pour $\theta = 1$, on a toujours $d_1((x_1, y_1), (x_2, y_2))/2 \leq d_e((x_1, y_1), (x_2, y_2)) \leq d_1((x_1, y_1), (x_2, y_2))$ de sorte que d_1 et d_e sont équivalentes. Il suffit alors de montrer que d_θ pour $\theta \neq 1$ est équivalente à d_1 . Supposons $\theta > 1$. Alors $|y_1 - y_2| < \theta|y_1 - y_2|$ et donc $d_1 \leq d_\theta$. De plus $|x_1 - x_2|/\theta < |x_1 - x_2|$

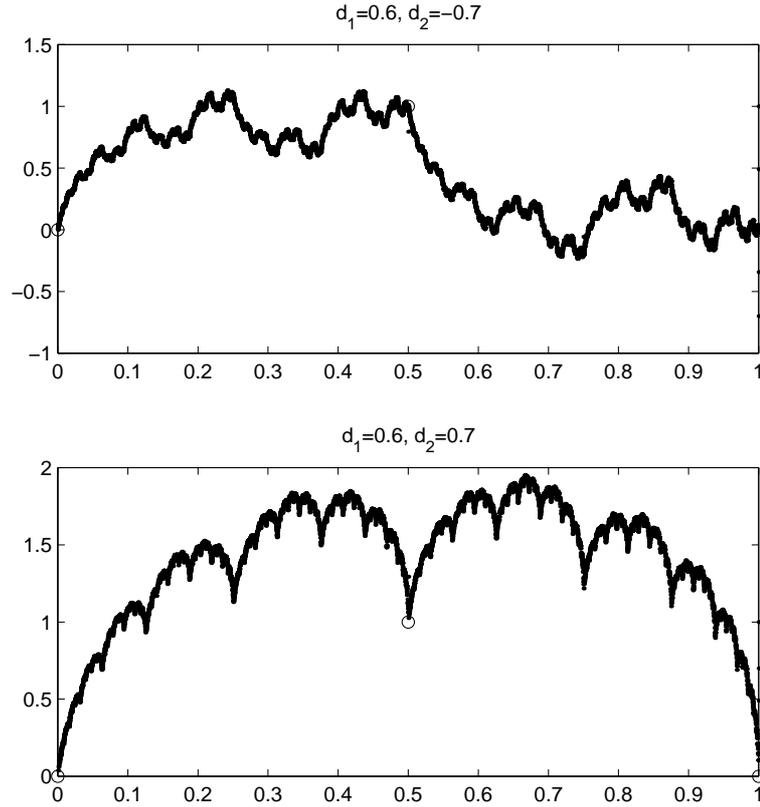


Figure 1.13: Interpolation de $(0,0)$, $(0.5,1)$ et $(1,0)$ par IFS. $d_1 = 0.6$ et $d_2 = \pm 0.7$ de sorte que les deux courbes ont même dimension fractale 1.38.

de sorte que $d_\theta/\theta \leq d_1$. Donc d_θ et d_1 sont équivalentes pour $\theta > 1$. Pour $\theta < 1$, on a le même résultat puisque $\theta d_\theta \leq d_1 \leq d_\theta$.

Il faut maintenant démontrer que l'IFS est hyperbolique avec la distance d_θ . Pour cela il faut montrer que les w_n sont contractantes.

$$\begin{aligned}
 d_\theta(w_n(x_1, y_1), w_n(x_2, y_2)) &= d_\theta((a_n x_1 + e_n, c_n x_1 + d_n y_1 + f_n), (a_n x_2 + e_n, c_n x_2 + d_n y_2 + f_n)) \\
 &= |a_n| |x_1 - x_2| + \theta |c_n (x_1 - x_2) + d_n (y_1 - y_2)| \\
 &\leq (|a_n| + \theta |c_n|) |x_1 - x_2| + \theta |d_n| |y_1 - y_2|
 \end{aligned}$$

qu'il faut majorer par $s d_\theta((x_1, y_1), (x_2, y_2))$. A cet effet, si les c_n sont tous nuls, on choisit $\theta = 1$ et sinon $\theta = \min(2 - |a_n|) / \max(|c_n|)$. Comme $|a_n| < 1$, $\theta > 0$. Alors

$$d_\theta(w_n(x_1, y_1), w_n(x_2, y_2)) \leq a |x_1 - x_2| + \delta \theta |y_1 - y_2|$$

où $a = 1 + |a_n| - \max(|a_n|)/2$ et $\delta = \max(|d_n|)$ qui sont strictement inférieur à 1.

Alors $d_\theta(w_n(x_1, y_1), w_n(x_2, y_2)) \leq \max(a, \delta) (|x_1 - x_2| + \theta |y_1 - y_2|)$ qui montre que w_n est contractante. Par suite l'IFS est contractant. ■

Théorème 2 Sous les hypothèses du théorème 1, l'attracteur G de l'IFS est le graphe d'un fonction continue $f : [x_0, x_N] \rightarrow \mathbb{R}$ qui interpole les données $\{(x_n, y_n)\}_{n=1, \dots, N}$ $N > 1$.

Démonstration : La démonstration de ce thorem introduit un opérateur qui agit sur l'ensemble \mathcal{F} des fonctions continues $f : [x_0, x_N] \rightarrow \mathbb{R}$ telles que $f(x_0) = y_0$ et $f(x_N) = y_N$. Tout d'abord, muni de $d(f, g) = \max(|f(x) - g(x)|, x \in [x_0, x_N])$, \mathcal{F} est un espace métrique complet. On considère les réels a_n, c_n, e_n, f_n définis par l'équation (1.8) et on définit

$$\begin{aligned} T : \mathcal{F} &\rightarrow \mathcal{F} \\ f &\mapsto Tf, (Tf)(x) = c_n l_n^{-1}(x) + d_n f(l_n^{-1}(x)) + f_n, \text{ pour } x \in [x_{n-1}, x_n] \forall n = 1, \dots, N \end{aligned}$$

où

$$\begin{aligned} l_n : [x_0, x_N] &\rightarrow [x_{n-1}, x_n] \\ x &\mapsto l_n(x) = a_n x + e_n \end{aligned}$$

Montrons que T envoie bien \mathcal{F} sur lui-même. Il faut donc montrer que $(Tf)(x_0) = y_0$, que $(Tf)(x_N) = y_N$ et que Tf est continue. On peut aisément montrer que $l_n^{-1}(x_0) = x_0$ et que $l_n^{-1}(x_N) = x_N$. On vérifie alors facilement que $(Tf)(x_0) = y_0$ et $(Tf)(x_N) = y_N$. De plus, comme composée de fonctions continues sur $]x_{n-1}, x_n[$, (Tf) est continue sur chacun de ces intervalles. Il faut alors vérifier la continuité en x_n . Ceci se montre en vérifiant l'égalité $c_n l_n^{-1}(x_n) + d_n f(l_n^{-1}(x_n)) + f_n = c_{n+1} l_{n+1}^{-1}(x_n) + d_{n+1} f(l_{n+1}^{-1}(x_n)) + f_{n+1} = y_n$.

Montrons maintenant que T est un opérateur contractant. Pour $x \in [x_{n-1}, x_n]$ on a

$$|(Tf)(x) - (Tg)(x)| = |d_n| |f(l_n^{-1}(x)) - g(l_n^{-1}(x))| \leq |d_n| d(f, g)$$

de sorte que $d(Tf, Tg) \leq \max(|d_n|) d(f, g)$. Comme les d_n sont plus petits que 1, T est donc contractant. T admet donc un unique point fixe dans \mathcal{F} . Il suffit de vérifier que ce point fixe correspond à l'attracteur de l'IFS.

Pour cela on remarque que la définition de T peut être mise sous la forme $(Tf)(a_n x + e_n) = c_n x + d_n f(x) + f_n, x \in [x_0, x_N]$ pour tout $n = 1, \dots, N$. Le graphe du point fixe de T vérifie donc $G = \bigcup_{n=1}^N w_n(G)$ et est par suite le point fixe de l'IFS. Ceci conclut la démonstration. ■

1.2.10 IFS pour les fonctions

On a vu aux paragraphes précédents comment le formalisme des IFS dans $\mathcal{H}(\mathbb{R}^2)$ permet de travailler sur des fonctions. Toutefois, ce formalisme est restreint au cas où les transformations sont affines. Or il existe un cadre plus général qui fait l'objet de ce paragraphe.

On travaille ici sur des fonctions définies sur un compact de \mathbb{R} . On ne restreint alors pas la généralité en considérant des fonctions $f : [0, 1] \rightarrow \mathbb{R}$. De plus, on considère des espaces métriques de fonctions. En particulier, nous allons travailler avec les espaces $L^p([0, 1])$, définis par

$$L^p([0, 1]) = \left\{ f : [0, 1] \rightarrow \mathbb{R} / \int_0^1 |f(t)|^p dt < +\infty \right\}$$

La distance adoptée est alors la distance classique

$$d_p(f, g) = \left\{ \int_0^1 |f(t) - g(t)|^p dt \right\}^{1/p}$$

Rappelons enfin qu'une fonction est lipschitzienne si il existe une constante C telle que $\forall t, u |f(t) - f(u)| \leq C|t - u|$.

Comme vu dans le cas des fonctions d'interpolation, le principe des IFS pour les fonctions est d'avoir une partition de l'intervalle de travail. Soient alors N applications contractantes w_i de facteurs de contraction c_i défini par

$$c_i = \sup_{(x, y) \in [0, 1]^2} \frac{d(w_i(x), w_i(y))}{d(x, y)}$$

et qui réalisent une partition de $[0, 1]$, c'est-à-dire

$$[0, 1] = \bigcup_{i=1}^N w_i([0, 1])$$

$$\mu(w_i([0, 1]) \cap w_j([0, 1])) = 0, \forall i \neq j$$

μ étant la mesure de travail, Lebesgue en général.

Soient maintenant N fonctions $\varphi_i(t, s)$ uniformément lipschitz en leur première variable, c'est-à-dire $|\varphi_i(t_1, s) - \varphi_i(t_2, s)| \leq K_i |t_1 - t_2|$, où K_i est une constante positive. On définit alors l'opérateur T suivant par

$$T : L^p([0, 1]) \longrightarrow L^p([0, 1])$$

$$f \longmapsto Tf = \sum_{i=1}^N T_i f$$

$$(T_i f)(x) = \varphi_i(f(w_i^{-1}(x)), w_i^{-1}(x)) \mathbf{1}_{w_i(X)}(x)$$

Le fonctionnement de cet opérateur est illustré figure 1.14

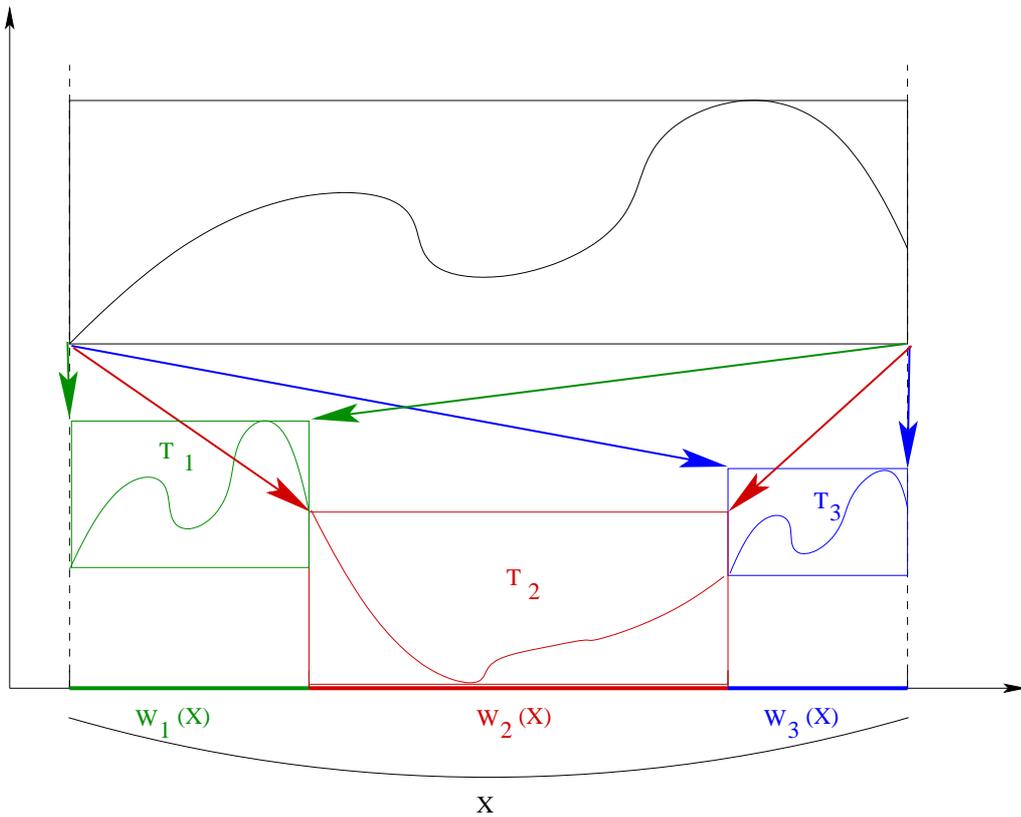


Figure 1.14: Description de l'opérateur définissant un IFS dans un espace de fonctions

On peut alors énoncer et démontrer les deux résultats suivants qui assurent l'existence d'une fonction limite à la suite $T^n f_0$.

Premier résultat : T est contractant dans $(L^p(X), d_p)$.

Démonstration Soient f et g dans $L^p(X)$. Alors

$$\begin{aligned}
\|(Tf)(x) - (Tg)(x)\|_p^p &= \int \left| \sum_{i=1}^N \varphi_i(f(w_i^{-1}(x)), w_i^{-1}(x)) \mathbf{1}_{w_i(X)}(x) - \sum_{i=1}^N \varphi_i(g(w_i^{-1}(x)), w_i^{-1}(x)) \mathbf{1}_{w_i(X)}(x) \right|^p dx \\
&= \int \left| \sum_{i=1}^N [\varphi_i(f(w_i^{-1}(x)), w_i^{-1}(x)) - \varphi_i(g(w_i^{-1}(x)), w_i^{-1}(x))] \mathbf{1}_{w_i(X)}(x) \right|^p dx \\
&= \sum_{i=1}^N \int_{w_i(X)} |\varphi_i(f(w_i^{-1}(x)), w_i^{-1}(x)) - \varphi_i(g(w_i^{-1}(x)), w_i^{-1}(x))|^p dx
\end{aligned}$$

On pose $y = w_i^{-1}(x)$, $dx = |J_i|dy$ où $J_i = dw_i/dy$ est de le Jacobien de la transformation, inférieur ou égal à c_i puisque w_i est contractant. D'où

$$\begin{aligned}
\|(Tf)(x) - (Tg)(x)\|_p^p &= \sum_{i=1}^N \int_X |\varphi_i(f(y), y) - \varphi_i(g(y), y)|^p |J_i| dy \\
&\leq \sum_{i=1}^N c_i \int_X |\varphi_i(f(y), y) - \varphi_i(g(y), y)|^p dy \\
&\leq \sum_{i=1}^N c_i K_i^p \int_X |f(y) - g(y)|^p dy \\
&= \left(\sum_{i=1}^N c_i K_i^p \right) d_p^p(f, g)
\end{aligned}$$

T est contractant à condition que $\sum_{i=1}^N c_i K_i^p < 1$.

On a également le résultat

$$d_p(Tf, Tg) \leq \left(\sum_{i=1}^N c_i^{1/p} K_i \right) d_p(f, g)$$

Comme T est contractant, le théorème du point fixe assure la convergence de la suite $T^n(f_0)$ vers une limite f_* . On a également une vitesse exponentielle :

Deuxième résultat : $T^n f_0$ atteint sa limite f_* avec une vitesse exponentielle.

Démonstration On a pour toute fonction f_0

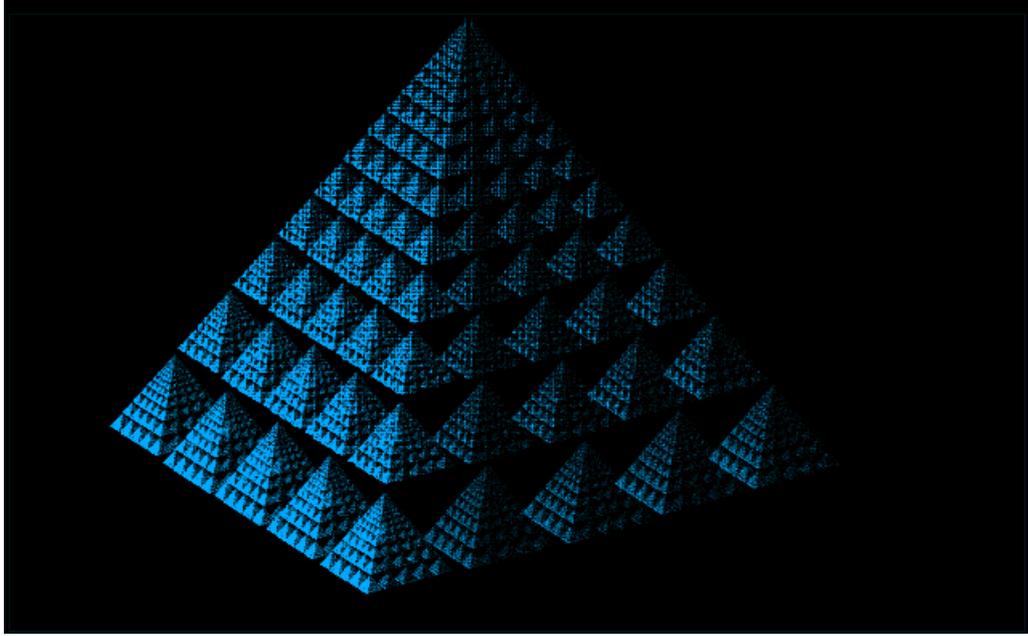
$$\begin{aligned}
d_p(T^n f_0, f_*) &= d_p(T^n f_0, T f_*) \\
&\leq C(p) d_p(T^{n-1} f_0, f_*) \\
&\leq C(p)^n d_p(f_0, f_*)
\end{aligned}$$

Or,

$$\begin{aligned}
d_p(f_0, f_*) &\leq d_p(f_0, T f_0) + d_p(T f_0, f_*) \\
&\leq d_p(f_0, T f_0) + C(p) d_p(f_0, f_*) \\
&\leq \frac{1}{1 - C(p)} d_p(f_0, T f_0)
\end{aligned}$$

Finalement,

$$d_p(T^n f_0, f_*) \leq \frac{C(p)^n}{1 - C(p)} d_p(T f_0, f_0)$$



La constant $C(p)$ intervenant ici est le facteur de contraction de T , soit d'après le premier résultat

$$C(p) = \left(\sum_{i=1}^N c_i K_i^p \right)^{1/p}$$

conditions de continuité

supposons que les w_i soient croissantes et provoquent la partition $a = a_0 \leq a_1 \leq \dots \leq a_{N-1} \leq a_N$, de sorte que $w_i([a, b]) = [a_{i-1}, a_i]$. Soit f_* le point fixe. La continuité est assurée si

$$\begin{aligned} f_*(a_i) &= \varphi_i(f_*(a_N), a_N) \\ &= \varphi_{i+1}(f_*(a_0), a_0) \end{aligned}$$

Or, on a également

$$\begin{aligned} f_*(a_N) &= \varphi_N(f_*(a_N), a_N) \\ f_*(a_0) &= \varphi_1(f_*(a_0), a_0) \end{aligned}$$

qui sont les points fixes de $\varphi_N(x, a_N)$ et $\varphi_1(x, a_0)$. On peut donc énoncer les conditions de continuité :

Soient α et β les points fixes respectifs de $\varphi_1(x, a_0)$ et $\varphi_N(x, a_N)$. Alors f_* est continu si $\varphi_i(\beta, a_N) = \varphi_{i+1}(\alpha, a_0)$

1.3 Une courbe fractale aléatoire : le mouvement Brownien fractionnaire

1.3.1 Rappels sur le mouvement brownien

Le mouvement Brownien est obtenu comme limite d'une marche au hasard définie comme

1. $x(0) = 0$,
2. $x(n \cdot dt) = x((n-1)dt) + a_n$, où a_n prend les valeurs $\pm\eta$ équiprobablement indépendamment des intervalles précédents. Pour les temps négatifs, la marche est définie de même.

Cette définition montre que $x(n \cdot dt) = \sum_{k=1}^n a_k$. Soit $t \in \mathbb{R}$. Alors il existe n tel que $t \in [n \cdot dt, (n+1)dt[$. On obtient alors

$$\begin{aligned} E[x(t)] &= 0 \\ E[x^2(t)] &= n\eta^2 \end{aligned}$$

Evaluons la fonction caractéristique $\varphi_x(u)$ de $x(t)$ afin de trouver sa densité de probabilité $p_x(v)$. Comme les a_n sont indépendantes, il vient

$$\varphi_x(u) = E[e^{iux}] = \prod_{i=1}^n E[e^{iua_n}] = E[e^{iua}]^n = \cos^n(u\eta)$$

Examinons maintenant $\psi_x(u) = \log(\varphi_x(u)) = n \log(\cos(u\eta))$. Supposons η suffisamment petit pour écrire

$$\psi_x(u) \approx n \log(1 - u^2\eta^2/2 + u^4\eta^4/4!) \approx -\frac{nu^2\eta^2}{2}$$

les termes correctifs étant d'ordre 4 en η . On cherche maintenant la limite de $\psi_x(u)$ lorsque $n \rightarrow +\infty$. *A priori*, cette limite semble être infinie. Toutefois, n représente le nombre de pas effectués entre les dates 0 et t , et donc $n = [t/dt]$ (partie entière). Ainsi la limite de $\psi_x(u)$ lorsque $n \rightarrow +\infty$ peut être finie si $n\eta^2$ est finie, c'est-à-dire si on choisit $\eta^2 = cte(n)/n$. Mais si n tend vers l'infini, dt tend vers zéro. Si l'on veut que le processus limite ait une variance finie, il faut que $n\eta^2 = t\eta^2/dt$ soit fini. Cette condition impose que η^2/dt soit constant. On pose alors $\eta^2 = \sigma^2 dt = cte(n)dt/n = cte(n)dt/t$, et il résulte que $cte(n) = \sigma^2 t$. On obtient finalement

$$\lim_{n \rightarrow +\infty} \psi_x(u) = -\frac{u^2\sigma^2 t}{2}$$

Ceci correspond à une densité de probabilité gaussienne, et donc

$$p_x(v) = \frac{1}{\sqrt{2\pi}\sqrt{t}\sigma} \exp\left(-\frac{v^2}{2\sigma^2 t}\right)$$

Ce processus limite est appelée mouvement Brownien ou processus de Wiener-Lévy. On le notera $B(t)$. La variance de ce processus est donc $\sigma^2 t$, ce qui montre sa non-stationnarité.

Par construction, les accroissements du brownien sont indépendants. Ils sont de plus stationnaires. On vérifie également que $B(at) = a^{1/2}B(t)$ en distribution. On dit alors que le processus est

auto-similaire avec comme index d'auto-similarité $1/2$. On peut montrer que le processus de Wiener-Lévy est le seul processus gaussien auto-similaire qui possède des accroissements indépendants et stationnaires. D'autres processus gaussiens auto-similaires existent, mais n'ont donc pas d'accroissements indépendants stationnaires. Ces processus sont maintenant décrits.

1.3.2 fBm : définition, propriétés

Le mouvement Brownien fractionnaire ou fBm, défini par Mandelbrot et Van Ness, est une généralisation du mouvement Brownien. Un fBm d'index $H \in]0, 1[$ est défini par

$$\begin{aligned} B_H(0) &= 0 \\ B_H(t) &= \frac{1}{\Gamma(H + 1/2)} \left[\int_{-\infty}^0 \left(|t - \tau|^{H-1/2} - |\tau|^{H-1/2} \right) dB(\tau) \right. \\ &\quad \left. + \int_0^t |t - \tau|^{H-1/2} dB(\tau) \right] \end{aligned} \quad (1.9)$$

que l'on écrit sous la forme plus compacte

$$B_H(t) = \int K_H(t, \tau) dB(\tau)$$

avec $K_H(t, \tau) = (\Gamma(H + 1/2))^{-1} (|t - \tau|^{H-1/2} \mathbf{1}_{]-\infty, t[}(\tau) - |\tau|^{H-1/2} \mathbf{1}_{]-\infty, 0[}(\tau))$. L'intégrale est une intégrale stochastique au sens de Wiener, $B(t)$ étant un brownien standard.

Dans le cas $H = 1/2$, on obtient $B_{1/2}(t) = B(t)$ et l'on retrouve le mouvement brownien. Il est clair que $B_H(t)$ est un processus gaussien de moyenne nulle, puisque transformée linéaire d'un processus gaussien à moyenne nulle. Nous allons maintenant étudier diverses propriétés de ce processus, et notamment montrer que ses réalisations sont des fonctions fractales.

Auto-similarité En utilisant la définition des fBm et l'auto-similarité du mouvement brownien, on montre facilement

$$B_H(at) = |a|^H B_H(t), \text{ en distribution}$$

Un fBm d'index H est donc auto-similaire d'index H . L'auto-similarité signifie que dilater le temps et normaliser adéquatement l'amplitude ne change pas les propriétés statistiques du processus. Ce fait est illustré sur la figure (1.15).

Structure au second-ordre Les incréments du fBm s'écrivent $B_\delta(t) = B_H(t + \delta) - B_H(t)$, soit

$$B_\delta(t) = \frac{1}{\Gamma(H + 1/2)} \left[\int_{-\infty}^{t+\delta} \left(|t + \delta - \tau|^{H-1/2} \right) dB(\tau) - \int_{-\infty}^t |t - \tau|^{H-1/2} dB(\tau) \right]$$

La variance de ces incréments est donc

$$\begin{aligned} E[B_\delta(t)^2] &= \Gamma(H + 1/2)^{-2} E \left[\left(\int_{-\infty}^{t+\delta} \left(|t + \delta - \tau|^{H-1/2} \right) dB(\tau) - \int_{-\infty}^t |t - \tau|^{H-1/2} dB(\tau) \right)^2 \right] \\ &= |\delta|^{2H} \Gamma(H + 1/2)^{-2} E \left[\left(\int_{-\infty}^1 \left(|1 + u|^{H-1/2} \right) dB(u) - \int_{-\infty}^0 |u|^{H-1/2} dB(u) \right)^2 \right] \\ &= V_H |\delta|^{2H} \end{aligned}$$

où $V_H = \text{Var}[B_H(1)] = \sigma^2 \Gamma(1 - 2H) \cos(\pi H) / (\pi H)$. La deuxième ligne de l'équation précédente est obtenue en faisant le changement $u = (t - \tau) / \delta$ et en utilisant l'auto-similarité du brownien.

Cette variance permet d'obtenir très facilement la fonction de covariance du fBm puisque

$$\begin{aligned} E[B_H(t_1)B_H(t_2)] &= \frac{1}{2} (E[B_H(t_1)^2] + E[B_H(t_2)^2] - E[(B_H(t_1) - B_H(t_2))^2]) \\ &= \frac{V_H}{2} (|t_1|^{2H} + |t_2|^{2H} - |t_1 - t_2|^{2H}) \end{aligned}$$

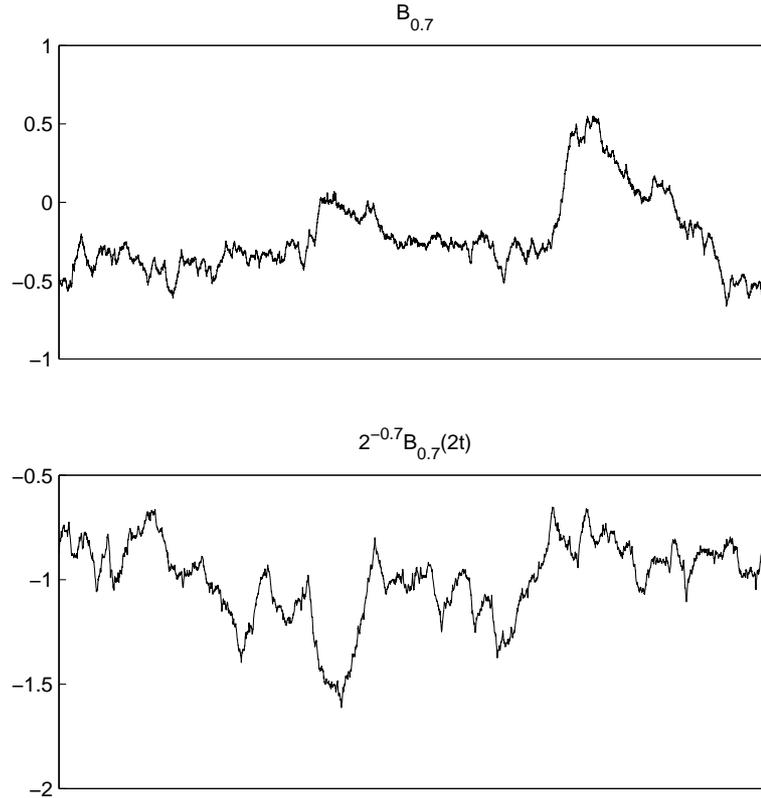


Figure 1.15: Illustration de l'auto-similarité du fBm. Le panneau du haut montre une réalisation de $B_H(t)$ et celui du bas une réalisation de $2^{-H}B_H(2t)$. Ces deux courbes sont statistiquement semblables.

qui montre la non-stationnarité du fBm. En particulier, $\text{Var}[B_H(t)] = V_H|t|^{2H}$, et l'on vérifie que le cas $H = 1/2$ conduit bien au brownien.

Continuité, dérivabilité, signal höldérien On peut montrer que les traces du fBm sont presque-surement continues, mais non dérivables (difficile!). Contentons nous de la moyenne quadratique. On se propose de savoir si le fBm est dérivable en moyenne quadratique, c'est-à-dire la limite suivante existe

$$\lim_{\delta \rightarrow 0} \frac{E[(B_H(t + \delta) - B_H(t))^2]}{\delta^2}$$

Le rapport précédent s'écrit explicitement $|\delta|^{2H-2}V_H$ qui tend vers l'infini lorsque δ tend vers zéro puisque $H \in]0, 1[$. Le fBm n'est donc pas dérivable en moyenne quadratique. De la même façon que dans le cas de mesure des longueurs, on peut chercher un nombre α tel que

$$\lim_{\delta \rightarrow 0} \frac{E[(B_H(t + \delta) - B_H(t))^2]}{\delta^{2\alpha}}$$

exite. Ceci requiert la convergence de $|\delta|^{2H-2\alpha}V_H$. Si $\alpha < H$ ce terme tend vers 0 avec δ alors qu'il diverge si $\alpha > H$. Par contre, ce terme prend une valeur finie si $\alpha = H$.

On dira qu'un signal aléatoire $x(t)$ a un exposant de Hölder $H < 1$ si

$$\sup_{\alpha} \left\{ \lim_{\delta \rightarrow 0} \frac{E[(x(t+\delta) - x(t))^2]}{\delta^{2\alpha}} < +\infty \right\} = H$$

Cet exposant lorsqu'il existe montre que les traces du signal sont continues (en moyenne quadratique), non dérivables, mais possèdent tout de même une certaine régularité : bien que non dérivables, elles sont toutefois "plus que continues".

Les traces des fBm sont donc höldériennes d'exposant H , et par suite non dérivables. Nous avons montré ces faits en moyenne quadratique, et ils sont également avérés presque-surement.

La dérivée du fBm n'existe pas en un sens usuel, mais peut être définie dans le cadre des processus aléatoires généralisés (distributions aléatoires). Derrière ces théorie se cache l'idée simple de régularisation du processus qui conduit au bruit gaussien fractionnaire.

Bruit gaussien fractionnaire (fGn) Pour régulariser le fBm, on l'intègre selon

$$B_H(t, \delta) = \delta^{-1} \int_t^{t+\delta} B_H(s) ds = \int B_H(s) \varphi(t-s) ds$$

où $\varphi(t) = \delta^{-1} \mathbf{1}_{[0, \delta]}(t)$. Si φ est choisie plus douce, on entre dans la théorie des distributions ! Le processus $B_H(t, \delta)$ est dérivable et l'on obtient

$$B'_H(t, \delta) = \delta^{-1} (B_H(t+\delta) - B_H(t))$$

qui représente le taux d'accroissement. Le processus dérivé est appelé bruit gaussien fractionnaire (fGn).

La variance de $B'_H(t, \delta)$ se calcule facilement et $\text{Var}[B'_H(t, \delta)] = V_H \delta^{2H-2}$. La fonction de corrélation du fGn s'obtient par un calcul direct et s'écrit

$$C_H(\tau, \delta) = E[B'_H(t, \delta) B'_H(t+\tau, \delta)] = \frac{V_H \delta^{-2}}{2} (|\tau + \delta|^{2H} + |\tau - \delta|^{2H} - 2|\tau|^{2H})$$

qui démontre la stationnarité du fGn.

Le comportement de la fonction de corrélation du fGn dépend de H et du comportement de τ/δ . En écrivant

$$C_H(\tau, \delta) = \frac{V_H \delta^{2H-2}}{2} \left(\left| \frac{\tau}{\delta} + 1 \right|^{2H} + \left| \frac{\tau}{\delta} - 1 \right|^{2H} - 2 \left| \frac{\tau}{\delta} \right|^{2H} \right)$$

on montre en effectuant un développement limité que

1. si $|\tau|/\delta \gg 1$:

$$C_H(\tau, \delta) \approx V_H H(2H-1) |\tau|^{2H-2}$$

2. si $|\tau|/\delta \ll 1$:

$$C_H(\tau, \delta) \approx C_H(0, \delta) + V_H \delta^{-2} |\tau|^{2H}$$

Le comportement de la corrélation $C_H(\tau, \delta)$ pour $|\tau| \rightarrow +\infty$ est très intéressant. Dans ce cas, $|\tau|/\delta \gg 1$ et le comportement dépend de H :

1. $H > 1/2$: La corrélation est positive, et en $+\infty$, elle décroît vers zéro comme τ^{2H-2} et $2H-2 \in]-1, 0[$. La décroissance de la fonction de corrélation est donc plus lente que $1/\tau$ de sorte que la corrélation n'est pas sommable! On est alors en présence d'un signal présentant le phénomène de dépendance ou mémoire longue. Ce genre de phénomène est observé dans de nombreuses situations (circulation d'information dans les réseaux informatiques, bruits électroniques, cours des rivières, etc).

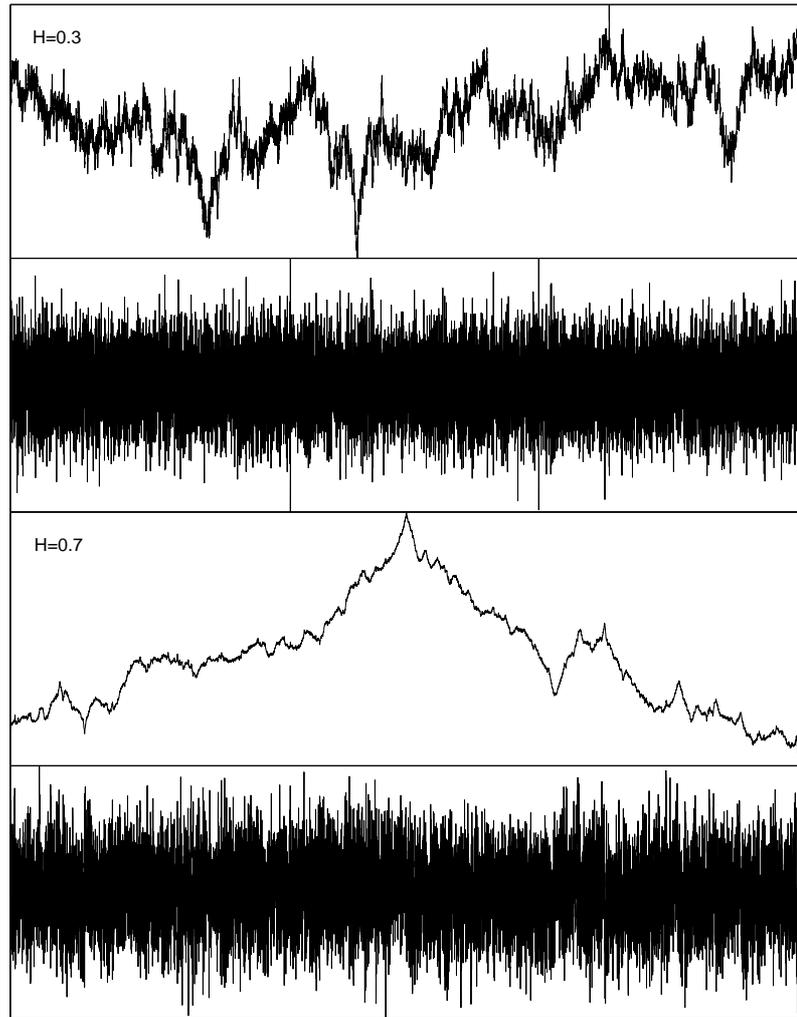


Figure 1.16: Comportement du fBm et du fGn pour $H < 1/2$ et $H > 1/2$. Les deux panneaux du haut montre la trace du fBm pour $H = 0.3$ et la trace du fGn associé. Les deux panneaux du bas montre le fBm et le fGn pour $H = 0.7$.

2. $H > 1/2$: La corrélation est négative pour τ suffisamment grand. On peut alors montrer que la somme de la corrélation est nulle, de sorte que l'on se trouve dans un cas plus classique, bien que la décroissance (en valeur absolue) à l'infini soit lente.

Comme le fGn est directement lié aux incréments du fBm, l'inspection de la corrélation du fGn nous renseigne directement sur le comportement des traces du fBm (figure 1.16). Lorsque $H > 1/2$, les incréments sont à corrélation positive : si la trace du fBm “monte”, elle aura alors tendance à continuer cette montée. Pour $H > 1/2$, le fBm aura des traces relativement douces, présentant des sortes d'oscillations sur de grandes échelles de temps. Dans le cas $H < 1/2$, la corrélation des incréments est négative, de sorte que si la trace du fBm “monte”, elle aura alors tendance à descendre rapidement. Les traces du fBm seront alors très chahutées.

Comportement spectraux Le bruit gaussien fractionnaire étant stationnaire, il possède une densité spectrale $\gamma_H(\nu, \delta)$ définie comme la transformée de Fourier de la corrélation $C_H(\tau, \delta)$. Le

point difficile provient du fait de la non-sommabilité de la corrélation dans le cas $H > 1/2$, et donc de la

non existence de la transformée de Fourier. En fait, ce problème n'est réel que pour la fréquence nulle pour laquelle le spectre est non défini. Plus précisément, lorsque ν tend vers zéro, la densité spectrale diverge et l'on parle de catastrophe "infrarouge". On peut trouver le comportement de cette divergence de la façon heuristique suivante

$$\begin{aligned}\gamma_H(\nu, \delta) &= \int_0^{+\infty} |\tau|^{2H-2} e^{-2i\pi\nu\tau} d\tau \\ &= \int_0^{+\infty} |\nu|^{2-2H} |\nu\tau|^{2H-2} e^{-2i\pi\nu\tau} d\tau \\ &= \frac{1}{|\nu|^{2H-1}} \int_0^{+\infty} |u|^{2H-2} e^{-2i\pi u} du\end{aligned}$$

et donc $\gamma_H(\nu, \delta) \approx cte/|\nu|^{2H-1}$ quand ν tend vers 0.

On est en présence d'un spectre dit en $1/f^\alpha$ ou plus couramment en $1/f$. On s'aperçoit que si $H > 1/2$ on a divergence en zéro, alors que $H < 1/2$ conduit à un spectre tendant vers zéro en zéro, et qui croît : on parle alors de catastrophe "ultraviolette".

Cette forme spectrale donne un moyen pour estimer le paramètre H puisqu'il suffit de tracer ce spectre dans un diagramme log-log et de faire une régression linéaire. En effet, on peut écrire

$$\log(\gamma_H(\nu, \delta)) = -(2H + 1) \log |\nu| + cte$$

Ceci est illustrée sur la figure (1.17) où les spectres du fGn sont tracés pour $H = 0.3$ et $H = 0.7$ dans un diagramme log-log. Les droites de pentes $1 - 2H$ sont superposées.

La notion de spectre n'a pas de sens pour le fBm puisqu'il s'agit d'un processus non stationnaire. Il faut donc en toute rigueur recourir aux outils du non-stationnaire. On peut par exemple utiliser le spectre de Wigner-Ville défini par W. Martin et P. Flandrin

$$\begin{aligned}W_B(t, \nu) &= \int_{-\tau/2}^{\tau/2} r_B(t - \frac{\tau}{2}, t + \frac{\tau}{2}) e^{-2i\pi\nu\tau} d\tau \\ &= (1 - 2^{1-2H} \cos(4\pi\nu t)) \frac{1}{|2\pi\nu|^{2H+1}}\end{aligned}$$

On peut obtenir un spectre moyen en effectuant la moyenne temporelle du spectre de Wigner-Ville et l'on trouve

$$\frac{1}{T} \int_0^T W_B(t, \nu) dt = \frac{1}{|2\pi\nu|^{2H+1}}$$

si $T = k/4\nu$. Pour un autre choix de T on obtient le même résultat à une constante dépendante de T près. Ce spectre moyen est encore une fois une loi de puissance. Il est montré figure (1.18) pour $H = 0.3$ et $H = 0.7$.

Pour terminer : et les fractales? Dans tout ce paragraphe, nous n'avons encore pas parlé de fractales ! En fait, elles sont sous-jacentes dans les termes "höldérien, spectre en $1/f$, corrélation à longue portée, etc".

Nous concluons donc cette partie par le théorème suivant, difficile à démontrer.

Théorème Les réalisations du mouvement Brownien fractionnaire d'index H ont presque-surement

une dimension fractale $D = 2 - H$.

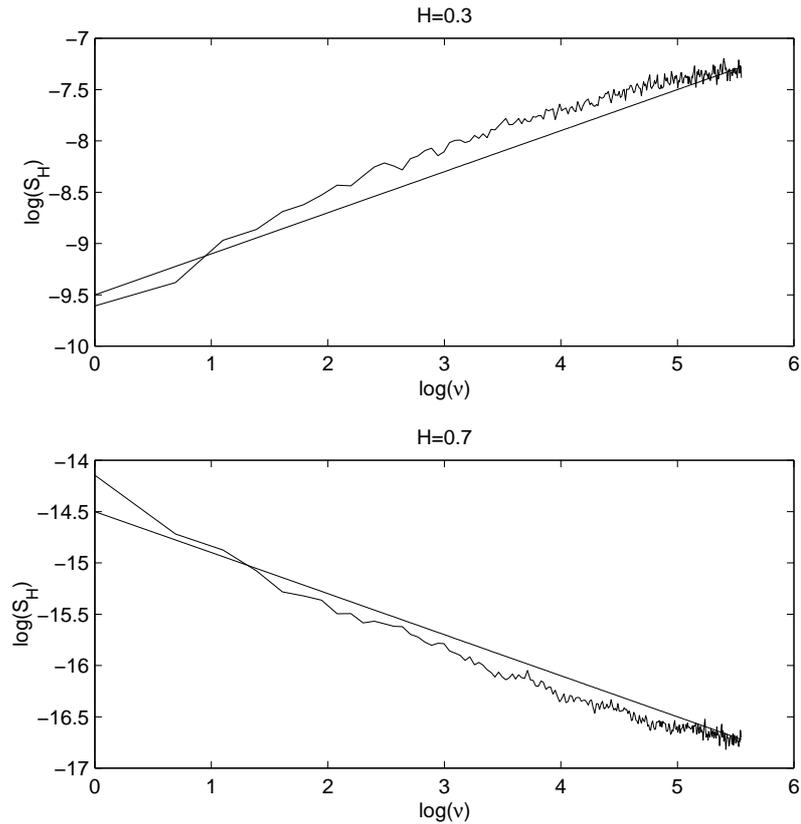


Figure 1.17: Spectre du fGn pour $H = 0.3$ et $H = 0.7$.

Ce résultat est à mettre en regard des différentes remarques effectuées tout au long de ce paragraphe. Par exemple, nous avons vu que pour $H < 1/2$, les traces du fBm sont très chahutées : elles tendent, lorsque H est de plus en plus petit à remplir le plan et conduisent donc à un graphe de dimension de plus en plus proche de 2.

Les fractales sont des objets particuliers, pour lesquels le grossissement d'un microscope ne change pas la vision (réellement ou statistiquement) que l'on a de l'objet. Les signaux fractals ne possèdent alors pas d'échelle de temps caractéristique, contrairement aux signaux usuels.

L'analyse des fractals requiert de prendre en compte ces remarques. En 1982, est apparu le microscope mathématique, outil idéal d'analyse des objets fractals. Ce microscope s'appelle *ondelette*.

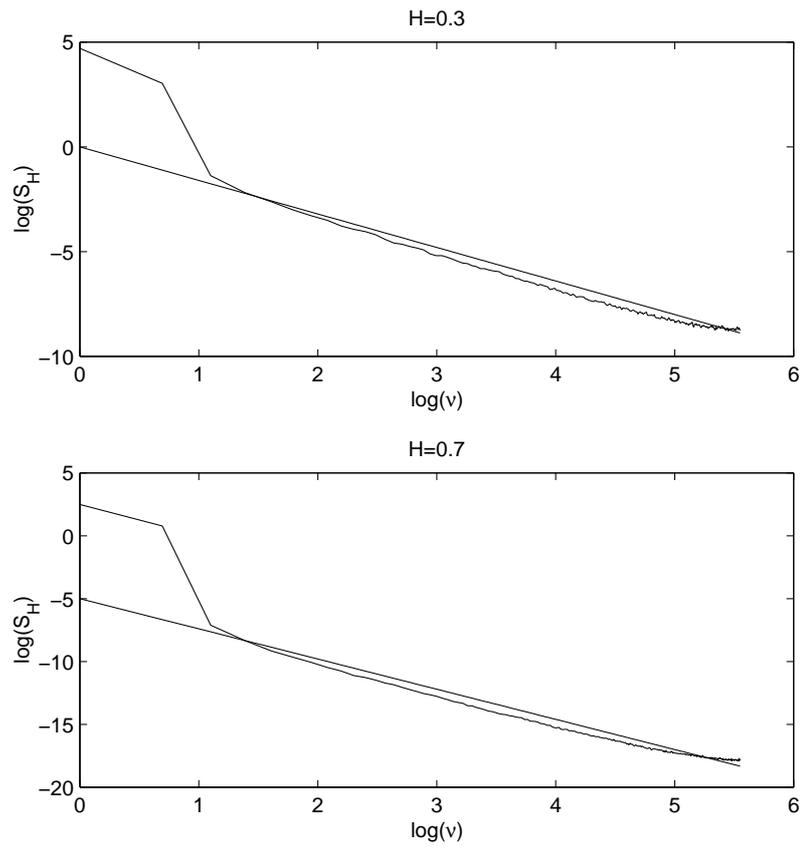


Figure 1.18: Spectre moyen du fBm pour $H = 0.3$ et $H = 0.7$.

Chapter 2

Ondelettes et signaux fractals

2.1 Rappels sur la transformée en ondelettes et sa discrétisation

La transformée en ondelettes continue consiste à regarder un signal $x(t) \in L^2(\mathbb{R})$ à diverses échelles, ceci au cours du temps. Elle se définit comme

$$T_x(a, b) = \frac{1}{\sqrt{a}} \int x(t) \psi^* \left(\frac{t-b}{a} \right) dt \quad (2.1)$$

où ψ est appelée ondelette et appartient à $L^2(\mathbb{R})$, et a est l'échelle. Le paramètre b représente la localisation temporelle de l'analyse. A une date b donnée, on réalise l'analyse en ondelettes en comparant le signal x à des versions dilatées et contractées de l'ondelette ψ . On doit pouvoir reconstruire le signal à partir de sa transformée en ondelettes. Pour cela, l'ondelette doit vérifier la condition dite d'admissibilité

$$C_\psi = \int \frac{|\hat{\psi}(\nu)|^2}{|\nu|} d\nu < +\infty$$

où $\hat{\psi}(\nu)$ est la transformée de Fourier de l'ondelette. Alors, on a le résultat

$$x(t) = \frac{1}{C_\psi} \int \int T_x(a, b) \frac{1}{\sqrt{a}} \psi \left(\frac{t-b}{a} \right) \frac{dadb}{a^2}$$

La condition d'admissibilité impose que la transformée de Fourier de l'ondelette soit nulle pour la fréquence nulle. Autrement dit

$$\int \psi(t) dt = 0$$

qui montre que l'ondelette est oscillante.

Pour implanter la transformée en ondelettes, on doit évidemment posséder une version discrétisée, c'est-à-dire définir une suite a_n d'échelles et une suite b_m de dates pour évaluer $T_x(a_n, b_m) = T_x(n, m)$. Cette discrétisation n'est pas triviale et on procède en général comme suit. On restreint d'abord l'échelle a aux réels positifs et l'on choisit $a_n = a_0^n, n \in \mathbb{Z}$. Pour $n = 0$ l'analyse compare le signal avec l'ondelette $\psi(t - b)$. On discrétise alors b traditionnellement, *i.e.* $b_m = mb_0$, de sorte que la famille des $\psi(t - mb_0)$ recouvre l'ensemble de la droite réelle. Pour $n > 0$, et $a_0 < 1$ par exemple, le support de l'ondelette analysante $\psi(a_0^{-n}t)$ est environ a_0^n fois plus petit que celui de l'ondelette non contractée. Pour couvrir l'ensemble de la droite réelle, il faut alors considérer

plus de points de calcul b_m que pour le cas $n = 0$. On en prend donc a_0^n fois plus, de sorte que $b_m = nb_0a_0^n$. Par simplicité, on choisit en général $a_0 = 1/2$ et $b_0 = 1$ de sorte que la transformée en ondelettes discrète est définie par

$$T_x(n, m) = \int x(t)2^{n/2}\psi^*(2^nt - m)dt$$

Des algorithmes rapides pour implanter cette transformée existent, mais le plus important dans cette formulation est l'existence d'ondelettes ψ pour lesquelles la famille $2^{n/2}\psi(2^nt - m)$, $(n, m) \in \mathbb{Z}$ est une base orthogonale de $L^2(\mathbb{R})$! On entre alors dans la théorie des décompositions en ondelettes orthogonales.

2.2 Transformée en ondelettes orthogonales

On étudie ici la transformée en ondelettes orthogonales. On commence par examiner la notion d'analyse multirésolution.

Analyse multirésolution. La compréhension de la théorie des ondelettes est assez simple lorsque l'on considère la théorie des analyses multirésolutions. L'idée est la suivante. Une fonction peut être vue comme la somme de deux fonctions : la première est une approximation alors que la deuxième corrige l'approximation. Cette deuxième fonction correspond donc aux détails oubliés lors de l'approximation. La notion d'analyse multirésolution formalise mathématiquement cette idée.

Définition : Une analyse multirésolution de $L^2(\mathbb{R})$ est une suite croissante de sous-espaces emboîtés V_j

$$\cdots V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \cdots \quad (2.2)$$

tels que

1. l'intersection des V_j est nulle, soit

$$\bigcap_{i=-\infty}^{+\infty} V_i = \emptyset \quad (2.3)$$

2. L'union des V_j est dense dans $L^2(\mathbb{R})$
3. $f(x)$ est dans V_j si et seulement si sa version contractée par un facteur 2 est dans V_{j+1} , c'est-à-dire

$$f(x) \in V_j \iff f(2x) \in V_{j+1} \iff f(2^{-j}x) \in V_0 \quad (2.4)$$

4. Si $f(x)$ est dans V_j , ses translatées "entières" sont dans V_j ,

$$f(x) \in V_j \iff f(x - k) \in V_j, \forall k \in \mathbb{Z} \quad (2.5)$$

5. Il existe une fonction $\phi(x)$ de V_0 telle que la famille

$$\{\phi(x - k), k \in \mathbb{Z}\} \quad (2.6)$$

est une base orthonormée de V_0 . ■

L'analyse multirésolution est schématisée sur la figure (2.1). Si $f(x)$ est dans V_0 sa version dilatée est dans V_{-1} et sa version contractée dans V_1 . On s'aperçoit qu'en dilatant de plus en plus, l'approximation d'une fonction deviendra de plus en plus grossière. Inversement, pour obtenir des approximations de plus en plus fines, il faut regarder dans les V_j avec j de plus en plus grands.

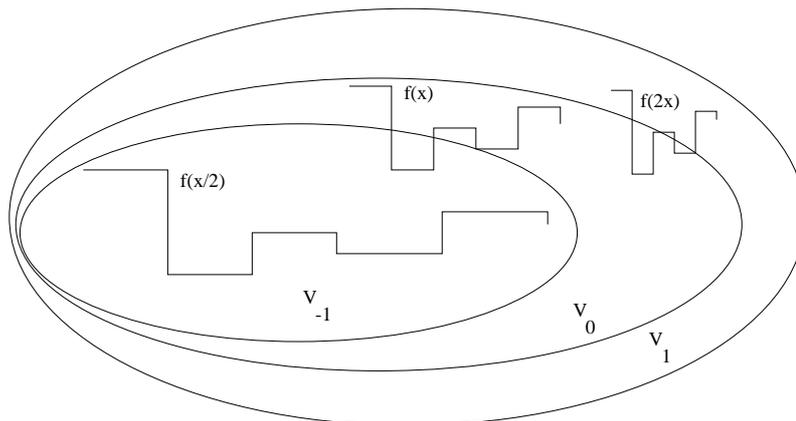


Figure 2.1: Illustration de l'analyse multirésolution.

Les espaces V_j constituent des espaces d'approximation. La projection d'une fonction $f(x)$ de $L^2(\mathbb{R})$ sur le sous-espace V_j constitue une approximation de $f(x)$ à l'échelle 2^j . La condition 2 dans l'énumération précédente assure que toute fonction de $L^2(\mathbb{R})$ peut être approchée dans cette analyse. La dernière condition implique que

$$\left\{ \phi_{j,k}(x) = 2^{j/2} \phi(2^j x - k), k \in \mathbb{Z} \right\} \quad (2.7)$$

est une base orthonormée de V_j . Toute fonction $f(x) \in L^2(\mathbb{R})$ peut être approchée par une fonction de V_j selon

$$P_j f(x) = \sum_{k \in \mathbb{Z}} a_{j,k} \phi_{j,k}(x) \quad (2.8)$$

où les coefficients $a_{j,k}$ sont appelés coefficients d'approximation et sont obtenus par

$$a_{j,k} = \int_{\mathbb{R}} f(x) \phi_{j,k}(x) dx \quad (2.9)$$

On dira que $P_j f$ est une approximation de f à l'échelle j .

Etant donné l'emboîtement des espaces V_j , il existe une très grande redondance d'information dans la représentation précédente. Une représentation plus adéquate prend en compte la différence d'information entre les approximations à deux échelles successives : les détails évoqués plus haut. Cette différence d'information peut être étudiée en considérant le complément orthogonal d'un espace d'approximation dans le suivant. Précisément, soit le sous-espace W_j défini par

$$V_{j+1} = V_j \oplus W_j \quad (2.10)$$

L'approximation à l'échelle $j+1$ peut alors être vue comme le raffinement de l'approximation à l'échelle j en lui ajoutant des détails. La définition de l'analyse multirésolution montre que l'union des W_j est dense dans $L^2(\mathbb{R})$.

Le problème suivant est celui de savoir si l'on dispose d'une base orthonormée dans les W_j de sorte qu'une collection

$$\left\{ \psi_{j,k}(x) = 2^{j/2} \psi(2^j x - k), j \in \mathbb{Z}, k \in \mathbb{Z} \right\} \quad (2.11)$$

constitue une base orthonormée de $L^2(\mathbb{R})$. Pour cela, il suffit d'avoir une fonction $\psi(x)$ de W_0 telle que la famille $\{\psi(x - k), k \in \mathbb{Z}\}$ soit une base orthonormée de W_0 . Cette fonction sera alors appelée ondelette.

Fonction d'échelle La fonction $\phi(x)$ est appelée fonction d'échelle. L'analyse multirésolution lui confère des propriétés intéressantes. Tout d'abord, puisque $\phi(x)$ est dans $V_0 \subset V_1$, elle peut se décomposer dans la base $(\sqrt{2}\phi(2x - k), k \in \mathbb{Z})$ de V_1 . Soit h_k ses coordonnées dans cette base. Alors

$$\phi(x) = \sum_k h_k \sqrt{2} \phi(2x - k) \quad (2.12)$$

La transformée de Fourier de cette équation conduit à

$$\Phi(\nu) = \sum_k h_k \frac{\sqrt{2}}{2} \Phi\left(\frac{\nu}{2}\right) e^{\pi j \nu k} \quad (2.13)$$

On pose

$$H(\nu) = \frac{\sqrt{2}}{2} \sum_k h_k e^{2\pi j \nu k} \quad (2.14)$$

qui est une fonction complexe 1-périodique. Alors

$$\Phi(\nu) = H\left(\frac{\nu}{2}\right) \Phi\left(\frac{\nu}{2}\right) \quad (2.15)$$

Utilisons maintenant le fait que $\phi(x)$ est orthogonale à ses translatées "entières". On a

$$\begin{aligned} \delta_{k,0} &= \int \phi(x) \phi(x - k) dx \\ &= \int |\Phi(\nu)|^2 e^{2\pi j \nu k} d\nu \\ &= \int_0^1 \sum_n |\Phi(\nu - n)|^2 e^{2\pi j \nu k} d\nu \end{aligned}$$

puisque l'exponentielle imaginaire est 1-périodique. Donc, on obtient

$$\sum_n |\Phi(\nu - n)|^2 = 1 \quad (2.16)$$

En séparant la somme en la somme sur les entiers pairs et la somme sur les entiers impairs, on a

$$\begin{aligned} 1 &= \sum_{2n} |\Phi(\nu - n)|^2 + \sum_{2n-1} |\Phi(\nu - n)|^2 \\ &= \sum_n |\Phi(\nu - 2n)|^2 + \sum_n |\Phi(\nu - 2n + 1)|^2 \end{aligned}$$

Utilisant (2.15), on a alors

$$\begin{aligned} 1 &= \sum_n |H(\nu/2 - n) \Phi(\nu/2 - n)|^2 + \sum_n |H(\nu/2 - n + 1/2) \Phi(\nu/2 - n + 1/2)|^2 \\ &= |H(\nu/2)|^2 + |H(\nu/2 + 1/2)|^2 \end{aligned}$$

en vertu de la 1-périodicité de $H(\nu)$ et de (2.16). Cette relation étant vraie pour toute fréquence, on a finalement

$$|H(\nu)|^2 + |H(\nu + 1/2)|^2 = 1$$

Ondelette. Comme $\psi(x)$ est dans $W_0 \subset V_1$ on peut écrire

$$\psi(x) = \sum_k g_k \sqrt{2} \phi(2x - k) \quad (2.17)$$

En procédant de même que pour la fonction échelle, on écrit

$$\Psi(\nu) = G\left(\frac{\nu}{2}\right) \Phi\left(\frac{\nu}{2}\right) \quad (2.18)$$

où

$$G(\nu) = \frac{\sqrt{2}}{2} \sum_k g_k e^{2\pi j \nu k} \quad (2.19)$$

Utilisons maintenant le fait que $W_0 \perp V_0$. Alors,

$$0 = \int \psi(x) \phi(x - k) dx \quad (2.20)$$

$$= \int \Psi(\nu) \Phi^*(\nu) e^{2\pi j \nu k} d\nu \quad (2.21)$$

$$= \int G(\nu/2) H^*(\nu/2) |\Phi(\nu)|^2 e^{2\pi j \nu k} d\nu \quad (2.22)$$

En procédant de même que pour la fonction échelle, on aboutit alors à

$$G(\nu) H^*(\nu) + G\left(\nu + \frac{1}{2}\right) H^*\left(\nu + \frac{1}{2}\right) = 0 \quad (2.23)$$

Supposons H donnée. Une solution de l'équation précédente est alors de la forme

$$G(\nu) = \Lambda(\nu) H^*\left(\nu + \frac{1}{2}\right) \quad (2.24)$$

où $\Lambda(\nu)$ est 1-périodique et vérifie

$$\Lambda(\nu) + \Lambda\left(\nu + \frac{1}{2}\right) = 0 \quad (2.25)$$

Un choix simple est $\Lambda(\nu) = -\exp(2\pi j \nu)$. Alors $G(\nu) = \exp(2\pi j (\nu + 1/2)) H^*((\nu + 1/2))$. On obtient alors $g_k = (-1)^k h_{1-k}$.

Il reste à vérifier que $\{\psi(x - k), k \in \mathbb{Z}\}$ est une base orthonormée de W_0 . Calculons $\langle \psi(x) | \psi(x - k) \rangle$. On a

$$\begin{aligned} \langle \psi(x) | \psi(x - k) \rangle &= \int \psi(x) \psi(x - k) dx \\ &= \int |\psi(\nu)|^2 e^{2\pi j \nu k} d\nu \\ &= \int_0^1 \sum_n |\Psi(\nu - n)|^2 e^{2\pi j \nu k} d\nu \end{aligned}$$

Or,

$$\begin{aligned} \sum_n |\Psi(\nu - n)|^2 &= \sum_n |G(\nu/2 - n/2) \Phi(\nu/2 - n/2)|^2 \\ &= \sum_n |H(\nu/2 - n/2 + 1/2) \Phi(\nu/2 - n/2)|^2 \\ &= |H(\nu/2 + 1/2)|^2 \sum_n \Phi(\nu/2 - n + 1/2) + |H(\nu/2)|^2 \sum_n \Phi(\nu/2 - n) \\ &= 1 \end{aligned}$$

On a donc finalement

$$\langle \psi(x) | \psi(x - k) \rangle = \delta_{k,0}$$

qui montre bien que $\{\psi(x - k), k \in \mathbb{Z}\}$ est une base orthonormée de W_0 . Enfin,

$$\left\{ \psi_{j,k}(x) = 2^{j/2} \psi(2^j x - k), j \in \mathbb{Z}, k \in \mathbb{Z} \right\} \quad (2.26)$$

est une base orthonormée de $L^2(\mathbb{R})$.

Approximation et détails. On a vu que le sous-espace W_j est défini comme le supplément orthogonal de V_j dans V_{j+1} . Ceci suggère une décomposition en cascade. On peut en effet écrire, pour toute fonction $f(x)$ de $L^2(\mathbb{R})$, que l'approximation à l'échelle $j + 1$ résulte de l'addition de détails (contenus dans W_j) à l'approximation à l'échelle j , c'est-à-dire

$$\text{Proj}(f(x)/V_{j+1}) = \text{Proj}(f(x)/V_j) + \sum_k d_{j,k} \psi_{j,k}(x) \quad (2.27)$$

En continuant l'itération, on obtient alors

$$\text{Proj}(f(x)/V_{j+1}) = \text{Proj}(f(x)/V_{j_0}) + \sum_{i=j_0}^j \sum_k d_{i,k} \psi_{i,k}(x) \quad (2.28)$$

On obtient donc une approximation fine de la fonction en ajoutant à une approximation plus grossière de cette fonction les détails successifs nécessaire pour arriver à la bonne résolution.

Algorithmes La structure en cascade décrite précédemment se transpose naturellement à l'algorithme de calcul des coefficients d'approximation et de détails (analyse), et à l'algorithme de reconstruction (synthèse).

On souhaite évaluer les coefficients d'approximation $a_{j,k}$ à l'échelle j . On a alors

$$\begin{aligned} a_{j,k} &= \int f(x) 2^{j/2} \phi(2^j x - k) dx \\ &= \int f(2^{-j}(u + k)) 2^{j/2+1} \phi(u) du \\ &= \sum_j h_l \int f(2^{-j}(u + k)) 2^{j/2+j} 2^{1/2} \phi(2u - l) du \\ &= \sum_l h_l \int f(x) 2^{(j+1)/2} \phi(2^{j+1} x - 2k - l) dx \end{aligned}$$

On voit donc que les coefficients d'approximation se calculent récursivement selon

$$a_{j,k} = \sum_l h_{l-2k} a_{j+1,l} \quad (2.29)$$

Les coefficients d'approximation à l'échelle j s'obtiennent à partir des coefficients d'approximation à l'échelle plus fine $j + 1$ par convolution, puis décimation.

Quant aux coefficients de détails, il vient

$$\begin{aligned}
d_{j,k} &= \int f(x)2^{j/2}\psi(2^j x - k) dx \\
&= \int f(2^{-j}(u+k))2^{j/2+1}\psi(u)du \\
&= \sum_j g_l \int f(2^{-j}(u+k))2^{j/2+j}2^{1/2}\phi(2u-l)du \\
&= \sum_l g_l \int f(x)2^{(j+1)/2}\phi(2^{j+1}x - 2k - l)dx
\end{aligned}$$

et donc

$$d_{j,k} = \sum_l g_{l-2k}a_{j+1,l} \quad (2.30)$$

Les coefficients de détails à l'échelle j s'obtiennent à partir des coefficients d'approximation à l'échelle plus fine $j+1$ par convolution, puis décimation.

Notons que les filtres apparaissant dans les relations d'analyse (2.29) et (2.30) sont les "retournés" dans le temps des filtres h_k et g_k .

Venons en à l'étape de synthèse. On utilise pour cela la formule (2.28) pour écrire

$$\sum_l a_{j,l}\phi_{j,l}(x) = \sum_l a_{j-1,l}\phi_{j-1,l}(x) + \sum_l d_{j-1,l}\psi_{j-1,l}(x)$$

On prend alors le produit scalaire de cette équation avec $\phi_{j,k}(x)$. Comme les $\phi_{j,k}(x)$ constituent une base orthonormée, on a

$$a_{j,k} = \sum_l a_{j-1,l} \langle \phi_{j-1,l}(x) | \phi_{j,k}(x) \rangle + \sum_l a_{j-1,l} \langle \psi_{j-1,l}(x) | \phi_{j,k}(x) \rangle$$

Or,

$$\begin{aligned}
\langle \phi_{j-1,l}(x) | \phi_{j,k}(x) \rangle &= \int 2^{(j-1)/2}\phi(2^{j-1}x - l)2^{j/2}\phi(2^j x - k)dx \\
&= \int 2^{(j-1)/2}2^{j/2}2^{1-j}\phi(u)\phi(2u + 2l - k)du \\
&= \sum_m h_m \int 2\phi(2u - m)\phi(2u + 2l - k)du \\
&= \sum_m h_m \delta_{m+2l-k,0} \\
&= h_{k-2l}
\end{aligned}$$

En procédant de même (on remplace dans ces lignes de calcul $\phi_{j-1,l}$ par $\psi_{j-1,l}$ et h_m par g_m) on obtient

$$\langle \psi_{j-1,l}(x) | \phi_{j,k}(x) \rangle = g_{k-2l}$$

Finalement, l'algorithme de synthèse est

$$a_{j,k} = \sum_l a_{j-1,l}h_{k-2l} + \sum_l d_{j-1,l}g_{k-2l} \quad (2.31)$$

Il procède donc à un sur-échantillonnage par 2 des coefficients d'approximation et de détail à une échelle j , puis au filtrage et addition de ces séries. L'analyse et la synthèse sont résumées sur la figure (2.2).

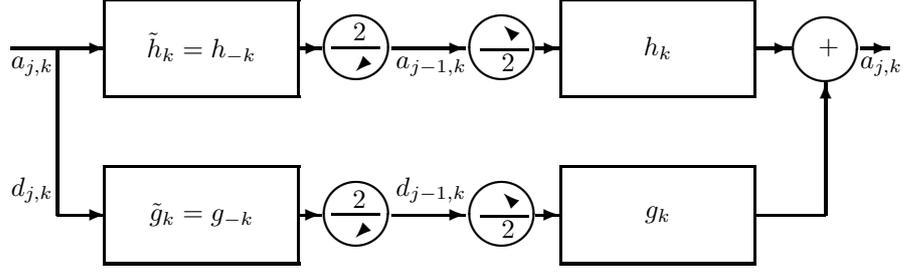


Figure 2.2: Analyse et synthèse dans la transformée orthogonale en ondelettes. L'analyse procède par filtrage et sous-échantillonnage. Les détails sont obtenus par filtrage passe-haut et mis de côté. L'approximation obtenue par filtrage passe-bas puis décimation est à son tour décomposée en approximation et détails. La synthèse est obtenue par l'opération duale.

Filtres H et G . Les filtres H et G ont des propriétés particulières qui leur sont conférées par l'analyse multirésolution.

Le filtre H est un filtre passe-bas. En effet, comme $\Phi(\nu) = H(\nu/2)\Phi(\nu/2)$, on a $H(0) = 1$. De plus, $H(1/2) = 0$. De plus, $G(0) = 0$ et $|G(1/2)| = 1$. G est passe-haut. Le caractère passe-bas de H correspond bien à une approximation, et le caractère passe-haut de G met en lumière la notion de détails.

Ces filtres sont appelées filtres miroirs en quadrature. La figure (2.2) montre qu'il permettent une cascade analyse-synthèse avec exacte reconstruction. Considérons les filtres en jeu indépendamment de la transformée en ondelettes, et cherchons les conditions de parfaite reconstruction. On a les relations

$$\begin{aligned} a_{j-1,k} &= \sum_l \tilde{h}_{2k-l} a_{j,l} \\ d_{j-1,k} &= \sum_l \tilde{g}_{2k-l} a_{j,l} \\ \hat{a}_{j,k} &= \sum_l h_{2k-l} a_{j,l} + \sum_l g_{k-2l} d_{j,l} \end{aligned}$$

En passant en fréquence réduite, on obtient

$$\begin{aligned} a_{j-1}(2\lambda) &= \frac{1}{2} \left(a_j(\lambda) \tilde{H}(\lambda) + a_j(\lambda + \frac{1}{2}) \tilde{H}(\lambda + \frac{1}{2}) \right) \\ d_{j-1}(2\lambda) &= \frac{1}{2} \left(a_j(\lambda) \tilde{G}(\lambda) + a_j(\lambda + \frac{1}{2}) \tilde{G}(\lambda + \frac{1}{2}) \right) \\ \hat{a}_j(\lambda) &= H(\lambda) a_{j-1}(2\lambda) + G(\lambda) d_{j-1}(2\lambda) \end{aligned}$$

Les deux premières équations s'obtiennent comme suit. Soit x_k la sortie du filtre \tilde{h} . Alors $x_k = \sum \tilde{h}_{k-l} a_{j,l}$. La suite sous-échantillonnée s'écrit donc $a_{j-1,k} = x_{2k}$ et sa transformée de Fourier en fréquences réduites est

$$\begin{aligned} a_{j-1}(\lambda) &= \sum_k x_{2k} e^{2\pi j \lambda k} \\ &= \sum_k x_{2k} e^{2\pi j \frac{\lambda}{2} 2k} \end{aligned}$$

Cette somme sur les indices pairs peut s'écrire $1/2 (\sum_k \cdot k + \sum_k \cdot k (-1)^k)$ qui revient à considérer deux fois les indices pairs et à ôter les impairs. Comme $(-1)^k = \exp(\pi j n)$ on a le résultat. Les termes en $\lambda + 1/2$ correspondent au recouvrement.

On écrit maintenant \hat{a}_j en fonction de a_j selon

$$\begin{aligned}\hat{a}_j(\lambda) &= \frac{1}{2} \left(H(\lambda)\tilde{H}(\lambda) + G(\lambda)\tilde{G}(\lambda) \right) a_j(\lambda) \\ &+ \frac{1}{2} \left(H(\lambda)\tilde{H}\left(\lambda + \frac{1}{2}\right) + G(\lambda)\tilde{G}\left(\lambda + \frac{1}{2}\right) \right) a_j\left(\lambda + \frac{1}{2}\right)\end{aligned}$$

Pour éliminer les recouvrements, on impose

$$H(\lambda)\tilde{H}\left(\lambda + \frac{1}{2}\right) + G(\lambda)\tilde{G}\left(\lambda + \frac{1}{2}\right) = 0$$

et la reconstruction parfaite impose

$$H(\lambda)\tilde{H}(\lambda) + G(\lambda)\tilde{G}(\lambda) = 1$$

Le choix adopté dans la transformée en ondelette est

$$\begin{aligned}\tilde{H}(\lambda) &= \frac{1}{\sqrt{2}}h(\lambda) \\ \tilde{G}(\lambda) &= -\frac{1}{\sqrt{2}}\exp(2\pi j\lambda)h(\lambda) = \frac{1}{\sqrt{2}}g(\lambda) \\ G(\lambda) &= \frac{1}{\sqrt{2}}g(-\lambda) \\ H(\lambda) &= \frac{1}{\sqrt{2}}h(-\lambda)\end{aligned}$$

où

$$\begin{aligned}h(\lambda) &= \sqrt{2}\sum h_k e^{2\pi j\lambda k} \\ g(\lambda) &= \sqrt{2}\sum g_k e^{2\pi j\lambda k}\end{aligned}$$

sont les grandeurs calculées lors de l'étude de la fonction échelle et de l'ondelette.

2.2.1 Exemples d'ondelettes

Nous montrons ici quelques exemples d'ondelettes.

Ondelette de Haar

La base de Haar est connue depuis le début du siècle, mais son interprétation dans la théorie des ondelettes est récente. L'ondelette de Haar est définie par

$$\psi^1(t) = \begin{cases} 1 & \text{si } t \in [0, \frac{1}{2}[\\ -1 & \text{si } t \in [\frac{1}{2}, 1[\\ 0 & \text{sinon} \end{cases}$$

et la base de Haar est constituée par la famille $\psi_{j,k}^1(t) = \{2^{j/2}\psi_1(2^j t - k)\}_{j,k \in \mathbb{Z}}$. Il est facile de vérifier que $\langle \psi_{j,k}^1 | \psi_{j',k'}^1 \rangle = \delta_{j,j'} \delta_{k,k'}$. La fonction échelle associée est l'indicatrice de $[0, 1]$, $\phi^1(t) = \mathbf{1}_{[0,1]}(t)$.

Toute fonction de $L^2(\mathbb{R})$ peut se décomposer sur la base de Haar, les coefficients étant donnés par

$$d_{j,k} = \langle x(t) | \psi_{j,k}^1(t) \rangle = \int x(t)\psi_{j,k}^1(t)dt$$

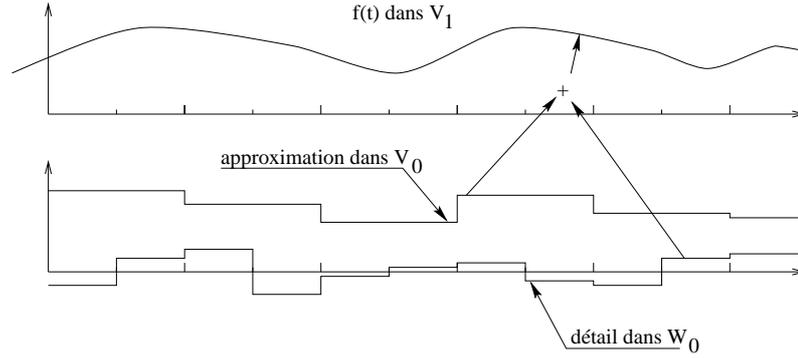


Figure 2.3: Illustration de la décomposition en ondelettes de Haar. Si $f(t)$ est dans V_1 , alors elle s'exprime exactement comme la somme d'une fonction d'approximation de V_0 et d'une fonction de détails dans W_0 .

De plus, on reconstruit x selon

$$x(t) = \sum_{j,k \in \mathbb{Z}} d_{j,k} \psi_{j,k}^1(t)$$

La base de Haar consiste à approcher une fonction par une somme de fonctions constantes par morceaux sur des intervalles de longueur 2^{-j} .

Dans la vision analyse multirésolution, on a

$$\text{Proj}(f(t)/V_{j+1}) = \text{Proj}(f(t)/V_j) + \sum_k d_{j,k} \psi_{j,k}^1(t) \quad (2.32)$$

Supposons que la fonction f soit dans V_1 . Alors l'équation précédente se réécrit

$$f(t) = \sum_k a_{0,k} \phi^1(t-k) + \sum_k d_{0,k} \psi^1(t-k)$$

L'illustration de cette décomposition est présentée sur la figure (2.3). La fonction $f(t)$ est approchée par une fonction constante par intervalle, à laquelle il faut ajouter

la fonction détail pour obtenir f .

Les ondelettes de Haar ont la propriété remarquable d'être à support compact, puisqu'elle sont portées par l'intervalle $[0, 1]$. Cette propriété a d'intéressante conséquence sur les filtres h_k et g_k utilisés pour l'algorithme. Comme h_k et g_k sont les coordonnées de respectivement $\phi^1(t)$ et $\psi^1(t)$ dans la base $\sqrt{2}\phi(2t-k)$ il est facile de voir que

$$\begin{aligned} h_0 &= \frac{\sqrt{2}}{2} \\ h_1 &= \frac{\sqrt{2}}{2} \\ g_0 &= \frac{\sqrt{2}}{2} \\ g_1 &= -\frac{\sqrt{2}}{2} \\ h_k &= g_k = 0, \forall k \neq 0, 1 \end{aligned}$$

Les filtres ainsi définis sont donc des filtres à réponse impulsionnelle finie, très facilement implantable. Toutefois, la régularité de l'ondelette de Haar est bien pauvre puisqu'elle n'est pas

continue! Ceci confère à sa transformée de Fourier des propriétés également pauvre, comme par exemple une très mauvaise localisation fréquentielle. Pour garder le caractère support compact (et donc des filtres associés MA) mais améliorer la régularité des ondelettes, il faut examiner les ondelettes de Ingrid Daubechies, découvertes au milieu des années 80.

Ondelettes de Daubechies

La construction des ondelettes à support compact de Daubechies est délicate. Elle fait appel à des résultats algébriques sur les polynômes (théorème de Bezout) et d'autres théorèmes liés aux polynômes trigonométriques. Nous ne présentons donc pas ici cette construction.

Les ondelettes de Daubechies sont indexées par un paramètre N et s'écrivent $\phi^N(t)$ (fonction échelle) et $\psi^N(t)$ (ondelette). L'ondelette de Haar est donc la première ondelette de Daubechies. Le paramètre N est lié à la régularité de l'ondelette.

En effet, on peut montrer que $\psi^N(t) \in C^{N-1}(\mathbb{R})$ où $C^N(\mathbb{R})$ est l'ensemble des fonctions continues $N - 1$ fois dérivables. Ainsi, ψ^N pour $N > 1$ est continue et peut être dérivée $N - 2$ fois. De plus, pour les ondelettes, cette notion est liée aux nombres de moments nuls de l'ondelette. On montre en effet que $\psi^N(t)$ vérifie

$$\int t^n \psi^N(t) dt = 0, \forall n = 0, 1, \dots, N - 1$$

$\psi^N(t)$ a donc ses N premiers moments nuls. L'ondelette de Haar ψ^1 est centrée (admissibilité de l'ondelette) mais son moment d'ordre 1 est non nul. Le nombre de moments nuls de l'ondelette est lié au comportement de la transformée de Fourier au voisinage de la fréquence nulle, et nous verrons l'importance de ce comportement dans le paragraphe 2.4.

Mentionnons enfin que les filtres associés à ϕ^N et ψ^N ont $2N$ coefficients non nuls, et que le support de la fonction échelle et de l'ondelette est de longueur $2N - 1$. La figure (2.4) montre les fonctions ϕ^N et ψ^N pour $N = 2, 3, 5, 9$. On observe "l'augmentation" de régularité mentionnée précédemment.

2.3 Ondelettes et signaux höldériens

Nous avons déjà rencontré dans le chapitre 1 des signaux höldériens. Il s'agit de signaux continus mais non dérivable, mais pour lesquels un exposant $\alpha \in]0, 1[$ caractérise la régularité. On peut généraliser cette définition en supposant que $\alpha \in]n, n + 1[$, où n est un entier. Précisément, ceci signifie que localement, une fonction peut être approchée par un polynôme d'ordre n et que le reste de l'approximation est puissance de $n + 1 - \alpha$.

Définition : Un signal $f(t)$ est dit höldérien en t_0 d'ordre $n + \alpha$, $\alpha \in]0, 1[$, s'il vérifie

$$|f(t) - P_n(t - t_0)| \leq C|t - t_0|^{n+\alpha}$$

où $P_n(t)$ est un polynôme de degré n . On peut alors énoncer le théorème suivant :

Théorème Soit $f(t)$ un signal höldérien en t_0 d'ordre $n + \alpha$, $\alpha \in]0, 1[$. Si l'ondelette ψ a au moins $n + 1$ moments nuls, si $\int (1 + |t|)^{n+1} |\psi(t)| dt < +\infty$, alors la transformée en ondelette de f en t_0 vérifie.

$$|T_f(a, t_0)| \leq C a^{n+\alpha+1/2}$$

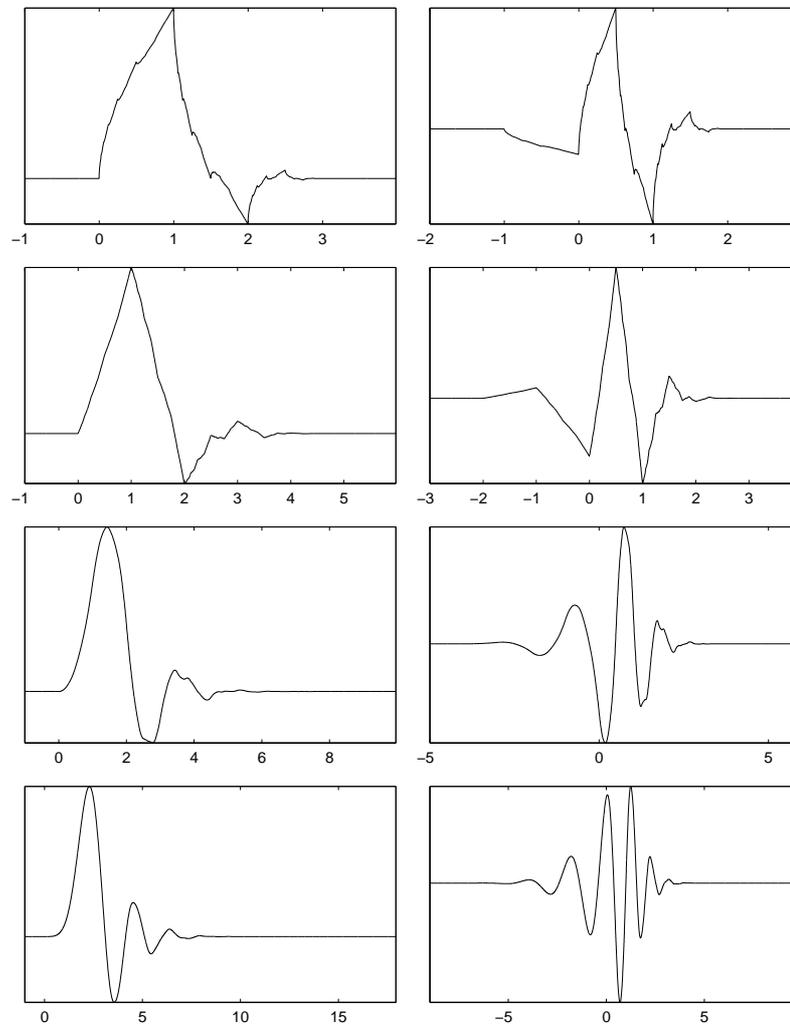


Figure 2.4: Ondelettes de Daubechies pour $N = 2, 3, 5, 9$ de haut en bas, à gauche la fonction échelle et à droite l'ondelette. On observe l'augmentation de la longueur du support et de la régularité avec N . Le nombre d'oscillations augmente également, ce qui assure le nombre de moments nuls correct.

Démonstration : Evaluons la transformée en ondelettes de $f(t)$ lorsque l'ondelette a $N \geq n + 1$ moments nuls

$$\int t^k \psi(t) dt = 0, \forall k = 0, \dots, n + 1, \dots, N$$

Comme l'ondelette a $N > n + 1$ moments nuls, elle est orthogonale aux polynômes d'ordre inférieur ou égal à n . Donc

$$\begin{aligned} T_f(a, b) &= \frac{1}{\sqrt{a}} \int f(t) \psi^* \left(\frac{t-b}{a} \right) dt \\ &= \frac{1}{\sqrt{a}} \int (f(t) - P_n(t - t_0)) \psi^* \left(\frac{t-b}{a} \right) dt \end{aligned}$$

Il vient alors

$$\begin{aligned} |T_f(a, b)| &\leq \frac{1}{\sqrt{a}} \int |f(t) - P_n(t - t_0)| \psi^* (fract - ba) dt \\ &\leq C \int C |t - t_0|^{n+\alpha} \psi^* \left(\frac{t-b}{a} \right) dt \end{aligned}$$

En effectuant le changement de variables $u = (t - b)/a$, et en regardant la transformée en $b = t_0$ on obtient

$$|T_f(a, t_0)| \leq C' a^{n+\alpha+1/2} \blacksquare$$

On voit alors que la transformée en ondelette révèle la structure intime de la fonction. Cette constatation permet alors une estimation du paramètre α caractérisant la régularité de la fonction. Ceci sera illustré dans le paragraphe suivant.

Le même genre de résultat existe aussi pour la transformée en ondelette orthogonal. Dans nos conventions d'écriture, pour un signal höldérien d'ordre $n + \alpha$, les coefficient de détails se comporte en $2^{-j(n+\alpha+1/2)}$.

2.4 Ondelettes et signaux fractals aléatoires

Dans ce paragraphe, nous étudions l'analyse en ondelette du mouvement brownien fractionnaire. On commence par l'analyse par une transformée en ondelette continue, puis nous examinons le cas de la décomposition orthogonale.

2.4.1 Transformée en ondelettes continue

Considérons un fBm d'index H . Sa transformée en ondelette, s'il elle existe, s'écrit

$$T_B(a, b) = \frac{1}{\sqrt{a}} \int B_H(t) \psi^* \left(\frac{t-b}{a} \right) dt$$

Comme le fBm est aléatoire, $T_B(a, b)$ définit un champ aléatoire dont il faut étudier les propriétés statistiques. Le fBm étant centré, le champ est également centré. Sa covariance est donnée par

$$\begin{aligned} \text{Cov}[T_B(a_1, b_1), T_B(a_2, b_2)] &= E[T_B(a_1, b_1) T_B^*(a_2, b_2)] \\ &= \frac{1}{\sqrt{a_1 a_2}} \int E[B_H(t_1) B_H(t_2)] \psi^* \left(\frac{t_1 - b_1}{a_1} \right) \psi \left(\frac{t_2 - b_2}{a_2} \right) dt_1 dt_2 \\ &= \frac{V_H}{2\sqrt{a_1 a_2}} \int (|t_1|^{2H} + |t_2|^{2H} - |t_1 - t_2|^{2H}) \psi^* \left(\frac{t_1 - b_1}{a_1} \right) \psi \left(\frac{t_2 - b_2}{a_2} \right) dt_1 dt_2 \\ &= -\frac{V_H}{2\sqrt{a_1 a_2}} \int |t_1 - t_2|^{2H} \psi^* \left(\frac{t_1 - b_1}{a_1} \right) \psi \left(\frac{t_2 - b_2}{a_2} \right) dt_1 dt_2 \end{aligned}$$

la dernière égalité provenant de la condition d'admissibilité. Divers enseignements peuvent être tirés de ce calcul.

A une échelle fixée a , la covariance s'écrit

$$\text{Cov}[T_B(a, b_1), T_B(a, b_2)] = -\frac{V_H a^{2H+1}}{2} \int \left| u_1 - u_2 + \frac{b_1 - b_2}{a} \right|^{2H} \psi^*(u_1) \psi(u_2) du_1 du_2$$

A une échelle donnée, la transformée en ondelette du fBm est un signal aléatoire stationnaire, puisque la covariance ne dépend que de l'écart entre les deux dates de calcul.

A une date b donnée et une échelle a fixée, on obtient

$$\text{Var}[T_B(a, b)] = a^{2H+1} \frac{-V_H}{2} \int |u_1 - u_2|^{2H} \psi^*(u_1) \psi(u_2) du_1 du_2$$

La variance de la transformée en ondelette varie donc en a^{2H+1} . Cette constatation permet une estimation statistique du paramètre H . Pour ce faire, on calcule le scalogramme $|T_B(a, b)|^2$ et l'on évalue

$$S_B(a) = \frac{1}{T} \int_{-T/2}^{T/2} |T_B(a, b)|^2 db$$

qui varie en a^{2H+1} . Une régression linéaire dans un diagramme log-log permet de trouver H .

Cette technique est illustrée sur les figures (2.5) et (2.6) pour $H = 0.3$ et $H = 0.7$ respectivement. Sur ces figures, nous représentons le signal, le logarithme en base 2 du carré de sa transformée en ondelettes et la somme sur le temps de ce scalogramme. Est superposée à cette somme la régression linéaire. Pour $H = 0.3$, la pente estimée est de -1.54, soit $\hat{H} = 0.27$. Pour $H = 0.7$, on obtient $\hat{H} = 0.62$. La longueur des signaux est de 4096 échantillons, ce qui explique cette estimation moyenne.

L'analyse effectuée sur le fBm est justifiée puisque le fBm est non stationnaire et requiert l'utilisation de techniques dévouées au non stationnaire. Cette analyse peut également être effectuée sur le bruit gaussien fractionnaire, qui est stationnaire. Dans ce cas la justification précédente n'est plus valable puisque ce bruit est stationnaire, et peut donc être analysé avec les techniques standard (analyse spectrale classique par exemple). Toutefois, quelques arguments permettent de montrer la supériorité des ondelettes sur les techniques classiques.

Bruits en $1/f^\alpha$: estimation de α Calculons le scalogramme et le spectrogramme dans un cas général de signal stationnaire dont le spectre s'écrit $S_x(\nu) = C/\nu^\alpha$. En sommant sur le temps ces deux quantités, nous obtiendrons deux estimateurs spectraux à comparer.

Dans le cas du scalogramme, l'estimateur spectral s'écrit

$$\hat{S}c_x(a) = \frac{1}{T} \int_{-T/2}^{T/2} \frac{1}{a} \left| \int x(t) \psi^* \left(\frac{t-b}{a} \right) dt \right|^2 db$$

et dans le cas du spectrogramme

$$\hat{S}p_x(\nu) = \frac{1}{T} \int_{-T/2}^{T/2} \left| \int x(t) h^*(t-b) e^{-2i\pi\nu t} dt \right|^2 db$$

Calculons le biais de ces estimateurs. Il vient dans le cas du scalogramme

$$\begin{aligned} E[\hat{S}c_x(a)] &= \frac{1}{T} \int_{-T/2}^{T/2} \frac{1}{a} \int r_x(u) \psi \left(\frac{v-b}{a} \right) \psi^* \left(\frac{v-u-b}{a} \right) dudvdb \\ &= a \int S_x(\nu) \left| \hat{\psi}(a\nu) \right|^2 d\nu \end{aligned}$$

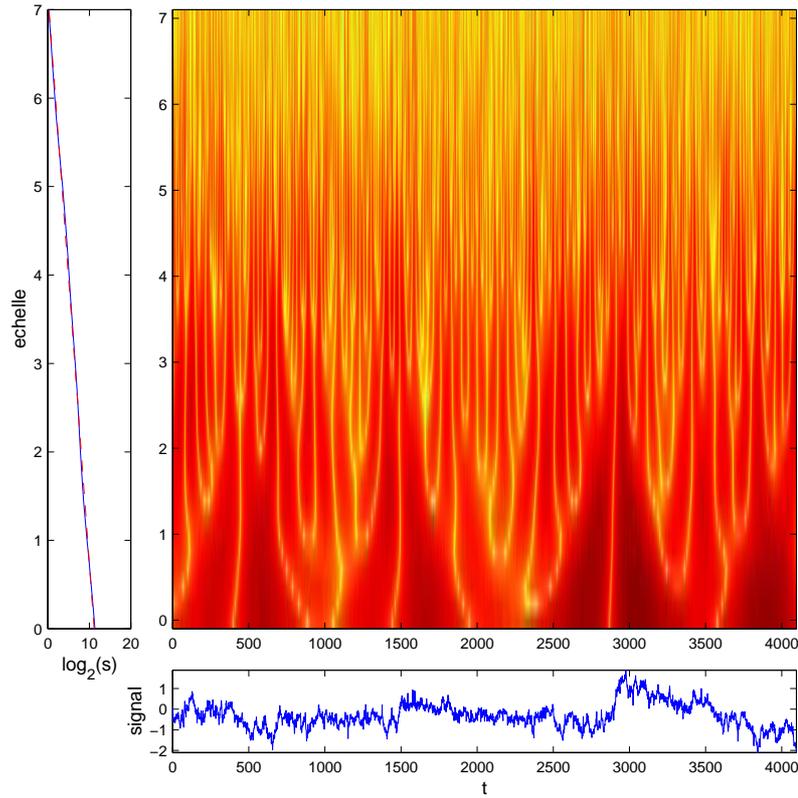


Figure 2.5: Transformée en ondelette d'un fBm de paramètre $H = 0.3$. Le panneau de gauche montre dans un diagramme log-log la somme sur le temps du module carré de la transformée. La régression linéaire superposée permet l'estimation de H . La valeur obtenue est 0.27.

et dans le cas du spectrogramme, on trouve

$$E[\hat{S}p_x(\nu)] = \int S_x(f) |\hat{h}(f - \nu)|^2 df$$

Si $S_x(\nu) = C/\nu^\alpha$, on obtient

$$\begin{cases} E[\hat{S}c_x(a)] &= a^\alpha \int C u^{-\alpha} |\hat{\psi}(u)|^2 du \\ E[\hat{S}p_x(\nu)] &= \nu^{-\alpha} \int C (1 + \frac{u}{\nu})^{-\alpha} |\hat{h}(u)|^2 du \end{cases}$$

En associant à l'échelle a la fréquence ν_0/ν , on s'aperçoit que

$$\begin{cases} E[\hat{S}c_x(a = \nu_0/\nu)] &= C\nu^{-\alpha} b_1 \\ E[\hat{S}p_x(\nu)] &= C\nu^{-\alpha} b_2(\nu) \end{cases}$$

Les deux estimateurs sont donc biaisés multiplicativement. Toutefois, le biais obtenu pour l'estimateur fondé sur le scalogramme est constant (donc facilement réduit), alors que le biais dans le cas du spectrogramme dépend de la fréquence. Ce résultat montre l'adéquation de la transformée en ondelettes pour l'analyse des signaux en $1/f$. Ce fait provient de la nature analyse à facteur de surtension constant de la transformée en ondelette : elle peut être vue comme le filtrage adapté au processus en $1/f$.

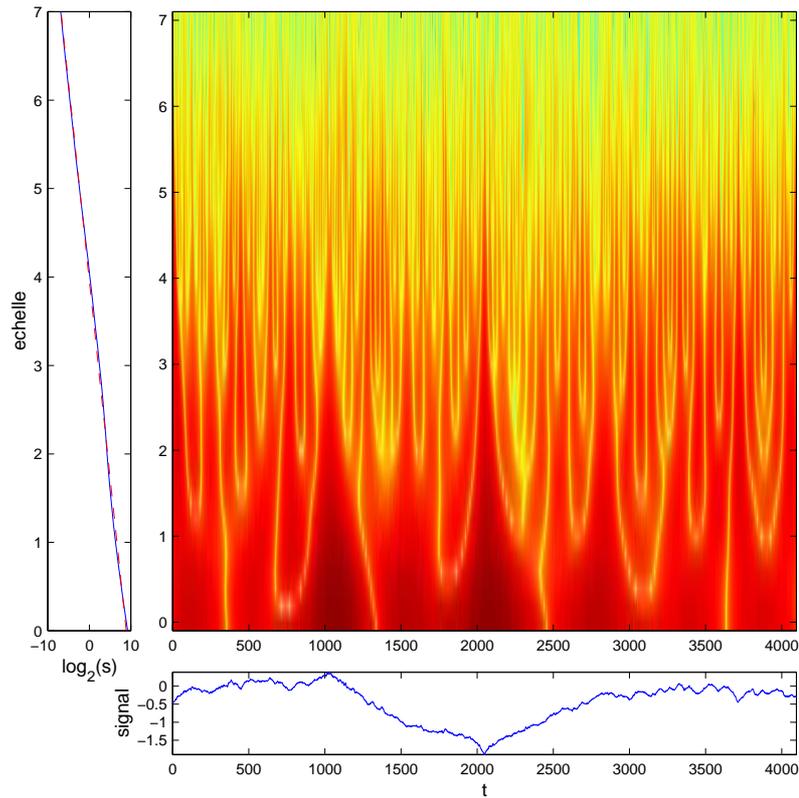


Figure 2.6: Transformée en ondelette d'un fBm de paramètre $H = 0.7$. La régression linéaire donne pour l'estimation de H la valeur 0.62.

Estimation du paramètre α : pourquoi ça marche ? Nous venons de voir que la transformée en ondelettes est un outil plus puissant que l'analyse de Fourier pour analyser des processus en $1/f$. Nous montrons ici les deux arguments fondamentaux du pourquoi de ce succès.

Le premier a déjà été dévoilé précédemment : il s'agit de l'adéquation entre l'analyse à facteur de surtension constant aux lois d'échelle (ou de puissance) en $1/f$.

Le deuxième argument est caché dans l'analyse. Considérons un processus stationnaire en dont le spectre est en $1/f^\alpha$, avec $\alpha \in]0, 1[$. Le comportement à l'infini de la corrélation associé est en $\tau^{\alpha-1}$, de sorte que la corrélation décroît très lentement à l'infini ($\alpha - 1 \in]-1, 0[$). Le processus considéré est à mémoire longue. Tout type d'estimateurs fondé sur la somme des échantillons de ce signal converge alors très lentement. En effet, des échantillons éloignés étant fortement liés, l'apport d'information sur le processus par ajout d'un échantillon dans une somme est très faible. A titre d'exemple, considérons l'estimateur de la moyenne $1/T \int_0^T x(t) dt$. La variance de cet estimateur pour le processus considéré se comporte comme $1/T^{1-\alpha}$. Si le processus est blanc, cette variance se comporte en $1/T$. Pour obtenir une variance de 0.001, il faut donc de l'ordre de 1000 échantillons dans le cas blanc, alors qu'il en faut $10^{3/\alpha-1}$ dans le cas longue dépendance. Si $\alpha = 0.1$ ce nombre vaut environ 4600, si $\alpha = .4$ il vaut 100000! Effectuer une analyse spectrale classique d'un processus à dépendance longue requiert donc de la patience!

Considérons maintenant la transformée en ondelettes. $T_x(a, b)$ à une échelle a fixée peut être vue comme un signal temporel $x_a(b)$. On montre alors facilement que la densité spectrale de ce

signal est

$$S_{x_a}(\nu) = aS_x(\nu) \left| \hat{\psi}(a\nu) \right|^2$$

En effet, la covariance de x_a s'écrit

$$E[x_a(b_1)x_a^*(b_2)] = \int aS_x(\nu) \left| \hat{\psi}(a\nu) \right|^2 e^{2i\pi\nu(b_1-b_2)} d\nu$$

et la forme de la densité spectrale suit par utilisation du théorème de Wiener-Kintchine.

La forme de la densité spectrale nous renseigne sur la fonction de corrélation. On sait en effet que le comportement à l'infini de la corrélation est gouverné par le comportement à l'origine de la densité spectrale. Or $S_{x_a}(\nu) = aC\nu^{-\alpha} \left| \hat{\psi}(a\nu) \right|^2$, de sorte que le comportement de $S_{x_a}(\nu)$ à l'origine est étroitement lié à celui de la transformée de Fourier de l'ondelette. Or le comportement de l'ondelette pour $\nu \rightarrow 0$ est régit par le nombre de moments nuls de l'ondelette, c'est-à-dire le nombre N tel que

$$\int x^n \psi(t) dt = 0, \forall n = 0, 1, \dots, N-1$$

En effet, le développement de Taylor en 0 de la transformée de Fourier de l'ondelette s'écrit

$$\hat{\psi}(\nu) = \sum_{n=0}^k \frac{\nu^n}{n!} \frac{d^n \hat{\psi}}{d\nu^n}(0) + o(\nu^k)$$

où $f(t) = o(t^k) \Leftrightarrow f(t)/t^k \rightarrow 0$ si $t \rightarrow 0$. Or, on sait que $\hat{\psi}^{(n)}(\nu) = (-2i\pi)^n TF(t^n \psi(t))$ de sorte que $\hat{\psi}^{(n)}(0) = (-2i\pi)^n \int t^n \psi(t) dt$. Si l'ondelette a N moments nuls, alors

$$\hat{\psi}(\nu) = Cte\nu^N + o(\nu^N)$$

Dans ce cas, on obtient pour la densité spectrale de $x_a(b)$

$$S_{x_a}(\nu) = aC'\nu^{2N-\alpha} + o(\nu^{2N-\alpha})$$

Le comportement à l'infini de la corrélation est alors

$$r_{x_a}(\tau) \approx \frac{1}{\tau^{2N+1-\alpha}}$$

Ainsi, si $2N+1-\alpha > 1 \Leftrightarrow N > \alpha/2$, la décroissance à l'infini de la corrélation est plus rapide que $1/\tau$ et la longueur de corrélation de x_a est beaucoup plus courte que celle de x . Autrement dit, la transformée en ondelettes effectuée à une échelle donnée un blanchiment imparfait des processus à dépendance longue dès que $N > \alpha/2$, et permet d'obtenir des estimateurs convergeant plus vite.

Pour terminer cette partie listons les avantages de la transformée en ondelettes pour l'analyse des signaux en $1/f$:

- Adéquation du banc de filtres "ondelette" aux processus en $1/f$ grâce à la surtension constante. Cette adéquation joue sur le biais en le rendant constant.
- Stationnarisation des processus non stationnaire en $1/f$. La non-stationnarité des processus en $1/f$ est essentiellement due à la catastrophe infrarouge. La nullité de la transformée de Fourier d'une ondelette admissible pour la fréquence nulle annihile cette catastrophe.
- blanchiment imparfait lorsque le nombre de moments nuls de l'ondelette est strictement plus grand que la moitié du paramètre α . A une échelle donnée, la transformée en ondelette est un signal "mélangeant" : sa corrélation tend vite vers zéro pour les grands retards. Ce point et le précédent joue sur la variance des estimateurs en permettant une convergence plus rapide.

2.4.2 Et la décomposition en ondelettes orthogonales?

Nous examinons maintenant la décomposition en ondelettes orthogonales du fBm. Les coefficients d'approximation et de détails sont donnés par

$$\begin{aligned} a_{j,k} &= \int B_H(t) \phi_{j,k}(t) dt \\ d_{j,k} &= \int B_H(t) \psi_{j,k}(t) dt \end{aligned}$$

Si l'on étudie les statistiques au second ordre des coefficients de détails, on retrouve le même genre de résultats que dans le cas de la transformée en ondelette continues.

Par exemple, Calculons la covariance des $d_{j,k}$ à une échelle 2^j fixée. On obtient alors

$$E[d_{j,k_1} d_{j,k_2}] = -\frac{V_H 2^{-j(2H+1)}}{2} \int |u_1 - u_2 + k_1 - k_2|^{2H} \psi(u_1) \psi(u_2) du_1 du_2$$

qui montre que les coefficients de détails sont stationnaires à une échelle donnée.

La variance de ces coefficients varie en $2^{-j(2H+1)}$, ce qui fournit également une possibilité d'estimation de H .

Toutes les méthodes développées en continu sont transposables au cas ondelettes orthogonales. Toutefois, ces transpositions peuvent être heureuses ou malheureuses, suivant les applications envisagées.

Toutefois, à des fins d'analyse, la transformée continue est en générale plus agréable à utiliser car elle permet une meilleure définition en échelle. De plus elle est invariante par translation temporelle, contrairement à la décomposition orthogonale.

Par contre, lorsqu'une étape de synthèse doit exister dans un traitement, l'utilisation des ondelettes orthogonales semble inévitable.

Cette remarque permet par exemple la définition de simulateur de mouvement Brownien fractionnaire.