Modélisation de séries financières par un modèle multifractal

Céline Azizieh

Mémoire présenté à l'Université Libre de Bruxelles en vue d'obtenir le diplôme d'actuaire

Supervision: W. Breymann (RiskLab, Suisse) P. Devolder (Université Libre de Bruxelles).

Année académique 2001-2002

ii

Table des matières

	Introduction						
Ι	Préliminaires mathématiques						
1	Pro	cessus	auto-similaires	7			
	1.1	Prélin	inaires: variables, vecteurs et processus stables	7			
		1.1.1	Variables aléatoires stables	7			
		1.1.2	Vecteurs aléatoires stables	10			
		1.1.3	Processus stochastiques stables	10			
	1.2	Proces	ssus auto-similaires	11			
		1.2.1	Processus auto-similaires à accroissements stationnaires	14			
		1.2.2	Mouvement brownien fractionnaire	16			
		1.2.3	Représentation intégrale d'un mouvement brownien fractionnaire .	19			
		1.2.4	Bruit gaussien fractionnaire	21			
		1.2.5	Propriété des trajectoires et simulations	23			
		1.2.6	Intégrale stochastique par rapport à un mouvement brownien frac-				
			tionnaire	28			
2	Processus multifractals						
	2.1	2.1 Exposants de singularité					
	2.2 Analyse multifractale		se multifractale	34			
		2.2.1	Spectres multifractals	35			
		2.2.2	Lien entre les spectres	36			
		2.2.3	Fonction de structure et spectre de Legendre	38			
		2.2.4	Exposants déterministes	40			
		2.2.5	Lois d'échelle	41			
		2.2.6	Intégrale stochastique par rapport à un processus multifractal \therefore	43			
	2.3	Casca	des multiplicatives	43			
		2.3.1	Mesures multinomiales	43			
		2.3.2	Analyse multifractale de la mesure binomiale déterministe	46			

II Modélisation pour le marché des changes par un modèle multifractal 51

3	Motivation de ce type de modèle							
	3.1	Le marché des changes et les données à haute fréquence Les données B Désaisonnalisation des données						
	3.2							
	3.3							
	3.4	Les faits stylisés						
		3.4.1	Queue lourde et non normalité de la distribution des returns	62				
		3.4.2	Comportement des volatilités réalisées des returns	71				
		3.4.3	Lois d'échelle	80				
		3.4.4	Hétérogénéité du marché	87				
4	Modèle multifractal en cascade							
	4.1	Rappe	l: Modèles ARCH	89				
	4.2	Modèl	es de volatilité stochastique	91				
	4.3	3 Le modèle en cascade						
		4.3.1	Analogie avec la turbulence	91				
		4.3.2	Le modèle	93				
		4.3.3	Le modèle MMAR de Mandelbrot–Fisher–Calvet	97				
	4.4	4.4 Résultats sur simulations du modèle						
		4.4.1	Une première impression	98				
		4.4.2	Fonction de vraisemblance	113				
		4.4.3	Approche plus quantitative	114				
		4.4.4	Lois d'échelle	117				
		4.4.5	Conclusion et perspectives	121				
	Con	onclusion 12						
Α	Dimensions fractales							
	A.1	1 Dimension de boîte ou capacité						
	A.2	Dimen	ision de Hausdorff	128				
		A.2.1	Mesure de Hausdorff d-dimensionnelle	128				
		A 2 2	Dimension de Hausdorff	128				
	A.3	Lien ei	ntre les deux dimensions	129				
	11.0							

Introduction

Le premier modèle d'évolution des cours d'un actif S(t) en fonction du temps est donné par l'équation différentielle stochastique

$$dS(t) = S(t)(\mu dt + \sigma dB_t) \tag{1}$$

où B_t est un mouvement brownien standard et μ, σ sont des constantes (modèle de Black-Scholes). Dans ce cas, il est très facile de voir que le cours S(t) est de loi log-normale:

$$S(t) = S(0)e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma B_t} \quad \forall t \ge 0.$$

Diverses généralisations de (1) ont été considérées par la suite (par exemples, modèles où la volatilité σ n'est pas constante mais variable et même stochastique). Puis on est passé au cas où B_t est remplacé par une martingale M(t), et on a constaté que l'on pouvait même aller jusqu'à une semi-martingale Y(t) pour conserver des propriétés intéressantes au niveau des marchés (notamment l'absence d'opportunité d'arbitrage).

Dans le modèle de Black-Scholes, si on définit le return logarithmique r(t) par

$$r(t) := \ln S(t) - \ln S(t-1)$$

pour tout $t \in \mathbb{N}$, alors

$$r(t) = \mu - \frac{1}{2}\sigma^2 + \sigma(B_t - B_{t-1}),$$

ce qui implique que les returns r(t) sont i.i.d. de loi $\mathcal{N}(\mu - \frac{1}{2}\sigma^2, \sigma^2)$.

En pratique cependant, on observe des distributions de returns ayant des queues plus lourdes que des normales. Ce fait a été remarqué pour la première fois par Mandelbrot dans [26] pour les prix du coton et a été depuis lors observé pour différents marchés. On a également observé le fait que l'indice de queue de la distribution des returns semble augmenter lorsque l'intervalle de temps considéré (pour le calcul des returns) augmente, jusqu'à se rapprocher d'une loi normale pour de grands intervalles de temps.

Une autre observation importante est celle de longue mémoire des marchés financiers: si le processus des returns semble effectivement non autocorrélé, il n'en est pas de même pour celui des returns absolus (i.e. en valeur absolue) ou carrés dont la fonction d'autocorrélation semble faiblement décroître dans le temps avec une décroissance de l'ordre d'une puissance (décroissance hyperbolique). De façon générale, on observe de la longue mémoire dans le processus des volatilités réalisées.

Un autre fait assez universellement observé dans les marchés financiers est les comportements d'échelle, c'est-à-dire le fait que, si l'on définit

$$S_q(\Delta t) := \frac{\Delta t}{N} \sum_{j=1}^{N/\Delta t} |r(j\Delta t, \Delta t)|^q$$

où $r(t, \Delta t)$ désigne le return logarithmique pour l'intervalle Δt à l'instant t et N la longueur de la série, alors $S_q(\Delta t)$ semble proportionnelle à $|\Delta t|^{\xi(q)}$ pour un certain exposant $\xi(q)$ dépendant de q de façon concave. Ces comportements sont typiques des processus auto-similaires (avec dans ce cas $\xi(q)$ linéaire) et des processus multifractals (si $\xi(q)$ est non linéaire).

Toutes ces observations ont conduit à des modèles d'évolution des cours plus généraux que ceux comportant un mouvement brownien ou une semi-martingale, par exemple contenant des mouvements browniens fractionnaires ou des processus multifractals. Ce genre de modèle semble alors mieux tenir compte des propriétés statistiques précitées, mais le problème est alors au niveau probabiliste: on ne sait jusqu'à présent pas définir une intégrale stochastique satisfaisante par rapport à ce genre de processus. Il y a donc de ce côté énormément de questions ouvertes en probabilité, mais nous ne les aborderons pas dans ce mémoire.

Dans le domaine de l'étude des propriétés statistiques des marchés, les choses ont fort évolué grâce aux données à hautes fréquences, c'est-à-dire ayant des fréquences supérieures à (1/) 1 jour. Ce type de données a fait son apparition grâce à l'informatisation des marchés. Il est actuellement possible de disposer de prix espacés par des intervalles de seulement parfois quelques secondes (pour certaines grandes devises du marché des changes).

Ce mémoire s'intéresse à la modélisation des cours par des processus qui ne sont pas des semi-martingales, et plus précisément, par des processus multifractals.

Dans une première partie plus mathématique, nous donnons quelques notions relatives aux processus auto-similaires, comprenant en particulier les mouvements browniens fractionnaires (chapitre 1), puis relatives aux processus dits multifractals (chapitre 2).

Les processus fractals seront essentiellement des processus dont les trajectoires présenteront certaines irrégularités à toutes les échelles. Un processus sera dit multifractal si le long d'une trajectoire quelconque, il présente des comportements d'échelle locaux non uniformes, ce qui se traduira par un type d'irrégularité (par exemple au sens d'Hölder) qui variera le long de la trajectoire considérées. Plus formellement, au lieu d'avoir un seul coefficient de Hausdorff, on trouvera tout un ensemble de valeurs distinctes, tout un spectre. On peut essayer de quantifier les irrégularités des trajectoires à l'aide de divers exposants de singularité dont certains sont introduits au chapitre 2. Ces exposants permettent alors d'introduire différents spectres (exemple: spectre de Hausdorff), liés également aux comportements d'échelle, et base d'une analyse multifractale.

Dans une deuxième partie, nous commençons par rappeler dans le chapitre 3 les propriétés statistiques observées dans les marchés financiers et rappelées ci-dessus, et nous en illustrons certaines sur des données issues du marché des changes. Ces données, constituées de plus de 65 000 points, contiennent le cours horaire du franc suisse par rapport au dollar américain sur une période de 10 ans. Il s'agit de données à haute fréquence provenant de la base de données d'Olsen¹. Ces données² ont été préparées par Wolfgang Breymann (actuellement chercheur à l'ETH Zürich) ayant travaillé chez Olsen durant plusieurs années.

La désaisonnalisation des données consiste à espacer régulièrement celles-ci dans une nouvelle échelle de temps liée au niveau d'activité du marché. Essentiellement, on augmentera la fréquence d'observations aux périodes de grande activité et on la diminuera aux périodes plus creuses. Le but est d'effacer la structure saisonnière observée dans le processus des volatilités réalisées (mesure de l'agitation du marché au voisinage d'un instant donné) et dans la fonction d'autocorrélation des returns absolus. Cette saisonnalité est un effet assez important dû au fait que l'activité du marché tourne autour de la planète avec un cycle de 24 heures, et cet effet est susceptible d'en voiler d'autres potentiellement observables mais plus subtils. Dans l'idée d'étudier plus finement les données (à forte saisonnalité), celles-ci ont été au préalable traitées en vue d'éliminer cette saisonnalité.

Les principales propriétés observées dans les données sont

- la non-normalité et la queue lourde de la distribution des returns,
- la lente décroissance de la fonction d'autocorrélation des returns absolus et le phénomène d'accumulation des volatilités (volatility clustering),
- les comportements d'échelle.

Une dernière observation importante est celle d'hétérogénéité du marché, c'est-à-dire le fait que les agents aient des perceptions du marché, des degrés d'information et des règles de comportement différents. Dans [29], les horizons de temps sur lesquels agissent les traders sont considérés comme une différence essentielle entre eux. Ils observent une asymétrie de la fonction d'autocorrélation des volatilités réalisées (calculées sur différentes échelles de temps) et interprètent cela comme un flux d'information net entre agents à long terme vers ceux à court terme.

¹Olsen Group, Seefeldstrasse 233, Zürich, Suisse

²données rendues accessibles par Olsen dans le cadre du projet RiskLab Volatility Estimation and Risk Measurement: From Short to Long-Time Horizons

Dans le chapitre 4, nous décrivons le modèle multifractal de Breymann–Ghashghaie– Talkner [7]. Ce modèle est à volatilité stochastique avec un processus de volatilité du même type que les cascades multiplicatives associées aux mesures multinomiales de Mandelbrot. Ce modèle est basé sur une analogie entre les phénomènes de turbulence et les marchés financiers développée dans [19]. Le point est qu'en turbulence (en théorie de l'intermittence), on observe le même genre de comportements statistiques que ceux mentionnés plus hauts pour des données financières, et le flux d'information à travers les horizons de temps est vu comme l'analogue de la cascade de Richardson (cascade de tourbillons à travers les échelles d'espaces).

Dans la dernière partie du chapitre 4, nous donnons un aperçu d'une comparaison entre les données et des simulations du modèle. Nous comparons les deux en ce qui concerne la fonction d'autocorrélation des returns absolus, la forme des distributions et les comportements d'échelle. Nous considérons également un ajustement des paramètres sur simulations du modèle, basé sur ces trois aspects. Cette dernière partie n'est en fait pas terminée, est en devenir et suscite encore beaucoup de questions ouvertes.

Ce travail est issu d'un séjour que j'ai effectué comme visiteur au RiskLab[®], ETH Zürich (Ecole Polytechnique Fédérale de Zürich), d'octobre à décembre 2001. J'y ai travaillé sur le projet intitulé *Volatility Estimation and Risk Measurement: From Short* to Long-Time Horizons, projet proposé au RiskLab[®] par W. Breymann.

La présente recherche a été supportée par Credit Suisse Group, Swiss Re et UBS AG par l'intermédiaire du RiskLab, Suisse.

Remerciements

Je voudrais tout d'abord remercier Wolfgang Breymann pour avoir accepté ma participation à son projet, et m'avoir aidée et suivie tout au long de mon séjour à Zürich. Je voudrais ensuite remercier le Professeur Uwe Schmock qui, lorsqu'il était directeur de la recherche au RiskLab, a permis que j'y effectue ce séjour comme visiteur. Je remercie également le Professeur Paul Embrechts pour les différents entretiens qu'il m'a accordés et ses conseils qui m'ont éclairée à bien des points de vue. Je voudrais enfin remercier le Professeur Pierre Devolder pour avoir accepté d'être mon promoteur de mémoire à l'Université Libre de Bruxelles et pour les discussions que nous avons eues.

Partie I

Préliminaires mathématiques

Chapitre 1

Processus auto-similaires

1.1 Préliminaires: variables, vecteurs et processus stables

Nous commençons par donner un très bref aperçu sur les variables aléatoires et processus stables. Cette théorie a été développée dans les années 1920 et 1930 par P. Lévy et A. Y. Khinchine. Les processus stables seront des généralisations des processus gaussiens, avec des seconds moments n'existant pas toujours. On peut trouver les définitions qui suivent dans la plupart des livres de probabilité. Nous suivons essentiellement le livre de Samorodnitsky et Taqqu [37].

On supposera donné un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ sur lequel seront définies toutes les variables aléatoires considérées dans ce chapitre. Elles seront toutes à valeurs réelles.

1.1.1 Variables aléatoires stables

Nous commençons par donner une définition des distributions stables ainsi que certaines propriétés (sans démonstration).

Définition 1.1.1 Une variable aléatoire X est dite de distribution stable si pour tout A, B > 0, il existe C > 0 et $D \in \mathbb{R}$ tels que

$$AX_1 + BX_2 \stackrel{\mathrm{d}}{=} CX + D \tag{1.1}$$

où X_1, X_2 sont des copies indépendantes de X. La variable X est dite strictement stable si D = 0 dans (1.1).

Remarquons que dans la classe des distributions stables, on trouve toutes les distributions dégénérées.

Le résultat suivant est prouvé dans [16]

Théorème 1.1.2 Pour toute variable aléatoire X stable, il existe $\alpha \in (0,2]$ tel que A, B, C dans (1.1) satisfont

$$C^{\alpha} = A^{\alpha} + B^{\alpha}$$

On dit que X est α -stable, et on qualifie α d'indice de stabilité (ou d'exposant caractéristique).

Exemple 1.1.3 (Distribution normale) Si $X \sim \mathcal{N}(\mu, \sigma^2)$, alors X est stable d'indice de stabilité $\alpha = 2$ car si X_1, X_2 sont deux copies indépendantes de X, on a

$$AX_1 + BX_2 \sim \mathcal{N}((A+B)\mu, (A^2+B^2)\sigma^2)$$

et donc on retrouve bien (1.1) avec $C = (A^2 + B^2)^{1/2}$ et $D = (A + B - C)\mu$.

On peut voir que la définition qui suit est équivalente à la première (cf. [37] pour une idée de la preuve de cette équivalence et [20]).

Définition 1.1.4 Une variable aléatoire est dite de distribution α -stable s'il existe $\alpha \in (0,2], \sigma \geq 0, \beta \in [-1,1]$ et $\mu \in \mathbb{R}$ tels que la fonction caractéristique de X soit de la forme:

$$\mathbb{E}\left[e^{i\theta X}\right] = \begin{cases} \exp(-\sigma^{\alpha}|\theta|^{\alpha}(1-i\beta(\operatorname{sign}\theta)\tan\frac{\pi\alpha}{2}) + i\mu\theta) & \text{si } \alpha \neq 1, \\ \exp(-\sigma|\theta|(1+i\beta\frac{2}{\pi}(\operatorname{sign}\theta)\ln|\theta|) + i\mu\theta) & \text{si } \alpha = 1, \end{cases}$$
(1.2)

avec

$$\operatorname{sign} \theta := \left\{ \begin{array}{ll} 1 & \operatorname{si} \theta > 0, \\ 0 & \operatorname{si} \theta = 0, \\ -1 & \operatorname{si} \theta < 0. \end{array} \right.$$

Les paramètres σ, β, μ sont uniques.

Remarquons que β peut-être choisi arbitrairement si $\alpha = 2$. On retrouve dans ce cas une loi normale de moyenne μ et de variance $2\sigma^2$ (et non σ^2 !).

Comme dans la définition ci-dessus il n'y a que 4 paramètres qui caractérisent la distribution, on notera la distribution stable de paramètres $(\alpha, \beta, \sigma, \mu)$ par $S_{\alpha}(\sigma, \beta, \mu)$. Lorsque $\beta = \mu = 0$, la fonction caractéristique de X est réelle, entraînant la symétrie de X. On notera alors simplement $X \sim S\alpha S$ ("symétrique α -stable"). L'expression de la fonction caractéristique devient alors simplement:

$$\mathbb{E}[e^{i\theta X}] = \exp(-\sigma^{\alpha}|\theta|^{\alpha}).$$

On peut voir assez facilement que si $\alpha > 1$, μ est en fait la moyenne de X.

Nous donnons ci-dessous quelques propriétés des distributions stables. La preuve de ces propriétés se trouve dans [37].

Proposition 1.1.5 (Somme de deux variables α -stables) Soient X_1, X_2 indépendantes avec $X_i \sim S_\alpha(\sigma_i, \beta_i, \mu_i), i = 1, 2$. Alors $X_1 + X_2 \sim S_\alpha(\sigma, \beta, \mu)$ avec

$$\sigma = (\sigma_1^{\alpha} + \sigma_2^{\alpha})^{1/\alpha}, \quad \beta = \frac{\beta_1 \sigma_1^{\alpha} + \beta_2 \sigma_2^{\alpha}}{\sigma_1^{\alpha} + \sigma_2^{\alpha}}, \quad \mu = \mu_1 + \mu_2.$$

Proposition 1.1.6 Soit $X \sim S_{\alpha}(\sigma, \beta, \mu)$ et $a \in \mathbb{R}$. Alors $X + a \sim S_{\alpha}(\sigma, \beta, \mu + a)$.

Proposition 1.1.7 Soit $X \sim S_{\alpha}(\sigma, \beta, \mu)$ et soit $a \in \mathbb{R}_0$. Alors

$$aX \sim S_{\alpha}(|a|\sigma, \operatorname{sign}(a)\beta, a\mu) \qquad si \ \alpha \neq 1$$
$$aX \sim S_{1}(|a|\sigma, \operatorname{sign}(a)\beta, a\mu - \frac{2}{\pi}a\ln|a|\sigma\beta) \qquad si \ \alpha = 1$$

Proposition 1.1.8 *Pour tout* $\alpha \in (0, 2)$ *,*

$$X \sim S_{\alpha}(\sigma, \beta, 0) \Leftrightarrow -X \sim S_{\alpha}(\sigma, -\beta, 0).$$

Proposition 1.1.9 Une variable aléatoire $X \sim S_{\alpha}(\sigma, \beta, \mu)$ est symétrique si et seulement si $\beta = \mu = 0$. Elle est symétrique autour de μ si et seulement si $\beta = 0$.

La preuve des propositions 1.1.5, 1.1.6, 1.1.7, 1.1.8, 1.1.9 utilise simplement l'expression de la fonction caractéristique d'une variable aléatoire stable.

On qualifie μ de paramètre de shift (ou de localisation) à cause de la Proposition 1.1.6. Le paramètre σ est appelé paramètre d'échelle. Remarquons néanmoins que si $\alpha = 1$, une multiplication par une constante affecte le paramètre de position de manière non linéaire. On qualifie β de paramètre de symétrie à cause de la Proposition 1.1.9.

Proposition 1.1.10 Soit $X \sim S_{\alpha}(\sigma, \beta, \mu)$ avec $\alpha \neq 1$. Alors X est strictement stable si et seulement si $\mu = 0$. Si cette fois $\alpha = 1$, alors X est strictement stable si et seulement si $\beta = 0$.

La preuve de la Proposition 1.1.10 utilise les propositions 1.1.5, 1.1.6, 1.1.7 ainsi que la première définition d'une variable aléatoire stable (on traite séparément les cas $\alpha \neq 1$ et $\alpha = 1$).

La proposition suivante est prouvée dans [37].

Proposition 1.1.11 (Existence des moments) Soit $X \sim S_{\alpha}(\sigma, \beta, \mu)$ avec $\alpha \in (0, 2)$. Alors

$$\mathbb{E}[|X|^p] < +\infty \quad \forall p \in (0, \alpha), \\ \mathbb{E}[|X|^p] = +\infty \quad \forall p \in [\alpha, +\infty).$$

On voit donc qu'excepté dans le cas gaussien $\alpha = 2$, les moments d'ordre supérieurs ou égaux à 2 des distributions stables n'existent pas.

On peut voir (cf. [43]) que les densités de probabilité des variables aléatoires stables existent et sont continues, mais à part quelques exceptions, on ne peut pas les exprimer analytiquement de façon explicite. Voici quelques exceptions¹ (on peut justifier ce qui suit en considérant les fonctions caractéristiques):

¹il y a encore une exception, le cas où $\alpha = 2^{-n}$ où $n \in \mathbb{N}$

- $S_2(\sigma, 0, \mu) = \mathcal{N}(\mu, 2\sigma^2)$
- $S_1(\sigma, 0, \mu) = \text{loi de Cauchy, fonction de densité } f(x) = \sigma/(\pi((x \mu)^2 + \sigma^2)).$
- $S_{1/2}(\sigma, 1, \mu)$ = distribution de Lévy, de densité

$$f(x) = \left(\frac{\sigma}{2\pi}\right)^{\frac{1}{2}} \frac{1}{(x-\mu)^{\frac{3}{2}}} \exp\left(-\frac{\sigma}{2(x-\mu)}\right).$$

En fait, $F(x) = 2(1 - \Phi\left(\sqrt{\frac{\sigma}{x}}\right))$ où Φ est la fonction de répartition d'une normale standard.

On peut voir que l'on peut exprimer la densité d'une loi stable par une intégrale d'une certaine fonction élémentaire faisant évidemment intervenir les paramètres α , β , $\mu\sigma$ (voir [41, 42]), mais vu le caractère peu maniable de l'expression, on préfère utiliser la fonction caractéristique.

1.1.2 Vecteurs aléatoires stables

Définition 1.1.12 Un vecteur aléatoire $\underline{X} = (X_1, \ldots, X_n)$ est dit stable (resp. strictement stable) si pour tout A, B > 0, il existe C > 0 et $D \in \mathbb{R}^n$ tels que

$$A\underline{X}_{(1)} + B\underline{X}_{(2)} \stackrel{\mathrm{d}}{=} C\underline{X} + D \tag{1.3}$$

où $\underline{X}_{(1)}, \underline{X}_{(2)}$ sont deux copies indépendantes de \underline{X} (resp. avec D = 0). On dit que \underline{X} est de distribution stable multivariée.

Le résultat suivant est prouvé dans [37], et sa preuve utilise la définition 1.1.12 et le Théorème 1.1.2.

Théorème 1.1.13 Soit \underline{X} un vecteur aléatoire stable. Alors il existe $\alpha \in (0,2]$ tel que $C = (A^{\alpha} + B^{\alpha})^{\frac{1}{\alpha}}$ dans (1.3). De plus, toute combinaison linéaire des composantes de \underline{X}

$$Y := \sum_{k=1}^{n} b_k X_k \tag{1.4}$$

est nécessairement α -stable. En outre, \underline{X} est strictement α -stable si et seulement si toute combinaison linéaire (1.4) est α -stable strictement. Enfin, \underline{X} est α -stable avec $\alpha \geq 1$ si et seulement si toute combinaison linéaire (1.4) est α -stable.

1.1.3 Processus stochastiques stables

Les processus stables ont joué un rôle important en modélisation en finance. Mandelbrot ([26]) a proposé un modèle pour les prix du coton selon un processus stable à variance infinie. Ce modèle a été fort critiqué à cause de la non existence des seconds moments du processus (cf. Proposition 1.1.11; il y a eu toute une controverse dans la communauté des chercheurs au sujet de l'existence des moments d'ordre 2 dans les séries financières). Ces processus sont une généralisation des processus gaussiens. Voir également [41, 42].

Définition 1.1.14 Un processus $\{X(t), t \in T\}$ $(T \subset \mathbb{R})$ est dit α -stable (resp. strictement stable) si pour tout $t_1, \ldots, t_n \in T$ et $n \in \mathbb{N}$, le vecteur aléatoire $(X(t_1), \ldots, X(t_n))$ est α -stable (resp. strictement stable).

Le résultat suivant est prouvé dans [37]. C'est une conséquence du Théorème 1.1.13.

Théorème 1.1.15 Soit $\{X(t), t \in T\}$ un processus stochastique. Alors:

1. $\{X(t), t \in T\}$ est strictement stable si et seulement si toute combinaison linéaire

$$\sum_{k=1}^{n} b_k X(t_k), \quad t_1, \dots, t_n \in T, \quad n \in \mathbb{N}, \quad b_1, \dots, b_n \in \mathbb{R}$$

est strictement stable.

2. Si $\alpha \geq 1$, alors $\{X(t), t \in T\}$ est α -stable si et seulement si toute combinaison linéaire du processus est α - stable.

Exemple 1.1.16 (Mouvement de Lévy α -stable) Un processus stochastique $\{X(t), t \in \mathbb{R}^+\}$ est un mouvement de Lévy standard α -stable si

- (1) X(0) = 0 p.s.
- (2) X a des accroissements indépendants
- (3) $X(t) X(s) \sim S_{\alpha}((t-s)^{1/\alpha}, \beta, 0)$ pour tout $0 \leq s \leq t < \infty$, où $\alpha \in (0, 2]$ et $\beta \in [-1, 1]$ (en particulier, accroissements stationnaires).

On retrouve un mouvement brownien si $\alpha = 2$. Ces processus jouent un rôle par rapport aux processus de Lévy similaire au rôle joué par les mouvements browniens par rapport aux processus gaussiens.

1.2 Processus auto-similaires

Les processus auto-similaires seront des processus de distribution invariante par certains changements d'échelle de temps et d'espace. Ils ont été introduits et étudiés par Kolmogorov à l'occasion de ses études sur les fluides turbulents. Le terme "auto-similaire" est dû à Mandelbrot, et est devenu standard. Cependant, Mandelbrot utilise quant à lui le terme auto-affin pour qualifier les processus auto-similaires (auto-similaire étant réservé pour lui à des objets géométriques).

Nous suivrons essentiellement [37] ainsi que le nouveau livre de Embrechts-Maejima [13]. On supposera toujours que les processus sont définis pour t appartenant à $T = \mathbb{R}$ ou \mathbb{R}^+ .

Définition 1.2.1 Un processus $\{X(t); t \in T\}$ est dit auto-similaire (self-similar) d'indice (de Hurst) H > 0 si pour tout $a > 0, n \in \mathbb{N}, t_1, \ldots, t_n \in T$, on a²

$$(X(at_1), X(at_2), \dots, X(at_n)) \stackrel{d}{=} (a^H X(t_1), a^H X(t_2), \dots, a^H X(t_n)).$$
(1.5)

On notera

$$\{X(at), t \in T\} \stackrel{\mathrm{d}}{=} \{a^H X(t), t \in T\}$$

à la place de (1.5) (lorsqu'on a cette relation pour tout $t_1, \ldots, t_n \in T$). On dira également qu'un processus est *H*-ss (*H* self-similar) lorsque celui-ci est auto-similaire d'indice *H*.

Exemple 1.2.2 (Mouvement brownien standard) Il est bien connu (cf. par exemple [40]) que l'on peut définir un mouvement brownien standard $\{X(t), t \in \mathbb{R}\}$ comme un processus gaussien de moyenne nulle et tel que

$$\mathbb{E}[X(t_1)X(t_2)] = \min(t_1, t_2) \quad \forall t_1, t_2 \in \mathbb{R}.$$

De ce fait, comme $\{X(at)\}$ (a > 0) est encore un processus gaussien de moyenne nulle et que

$$\mathbb{E}[X(at_1)X(at_2)] = \min(at_1, at_2) = a\min(t_1, t_2) = \mathbb{E}[a^{1/2}X(t_1)a^{1/2}X(t_2)],$$

on en déduit que

$$\{X(at), t \in \mathbb{R}\} \stackrel{\mathrm{d}}{=} \{a^{1/2}X(t), t \in \mathbb{R}\},\$$

et que le mouvement brownien standard est auto-similaire d'indice H = 1/2.

Exemple 1.2.3 (Mouvement de Lévy α -stable) On vérifie facilement que le processus défini dans l'exemple 1.1.16 est auto-similaire de paramètre $H = 1/\alpha$ (du moins si $\alpha \neq 1$). En effet, si a > 0, le processus $\{Y(t) := a^{-1/\alpha}X(at), t \in \mathbb{R}^+\}$ est un processus tel que Y(0) = 0, à accroissements indépendants et tel que $Y(t) - Y(s) \stackrel{d}{=}$ $a^{-1/\alpha}(X(at) - X(as)) \stackrel{d}{=} a^{-1/\alpha}Z$ où $Z \sim S_{\alpha}(a^{1/\alpha}(t-s)^{1/\alpha},\beta,0)$. Par la Proposition 1.1.7, $a^{-1/\alpha}Z$ est de loi $S_{\alpha}((t-s)^{1/\alpha},\beta,0)$ si $\alpha \neq 1$. Le processus $\{Y(t)\}$ est donc luimême un processus de Lévy α -stable de mêmes paramètres que le processus de départ $\{X(t)\}$. Cela prouve l'auto-similarité.

Remarque 1.2.4

1. Dans la définition de processus auto-similaire, on demande que les vecteurs aléatoires (de dimension finie) $(X(at_1), X(at_2), \ldots, X(at_n))$ soient invariants en loi par changement d'échelle. Contrairement à l'autosimilarité déterministe, on ne demande pas que les réalisations, les trajectoires du processus se répètent lorsqu'on agrandit certaines parties. C'est l'allure générale des trajectoires qui sera invariante par changement d'échelle (leurs propriétés statistiques).

 $^{^2 {\}rm par}$ un théorème de Kolmogorov, la distribution du processus est entièrement caractérisée par les distributions de dimension finie

1.2. Processus auto-similaires

2. Si $T = \mathbb{R}$, alors pour tout $t \neq 0$ fixé,

$$X(t) \stackrel{\mathrm{d}}{=} |t|^H X(\operatorname{sign} t) = \begin{cases} t^H X(1) & \operatorname{si} t > 0\\ |t|^H X(-1) & \operatorname{si} t < 0 \end{cases}$$

3. Si un processus $\{X(t), t \in T\}$ est *H*-ss, et si $0 \in T$, alors pour tout a > 0, on a $X(0) = X(0a) \stackrel{\text{d}}{=} a^H X(0)$. Cela implique que l'on a nécessairement

$$X(0) = 0 \ p.s.$$

4. Un processus H-ss (avec $T = \mathbb{R}$ ou \mathbb{R}^+) non dégénéré ne peut pas être stationnaire. En effet, si le processus était stationnaire, pour tout $t_1, \ldots, t_n \in T$ et pour tout $h \in \mathbb{R}$,

$$(X(t_1+h),\ldots,X(t_n+h)) \stackrel{\mathrm{d}}{=} (X(t_1),\ldots,X(t_n))$$

et on aurait en particulier que pour tout a>0 et t>0 fixé, que

$$X(t) \stackrel{\mathrm{d}}{=} X(at + (1-a)t) \stackrel{\mathrm{d}}{=} X(at) \stackrel{\mathrm{d}}{=} a^{H}X(t)$$

On arrive alors à une contradiction car $a^H X(t) \to \infty$ p.s. sur $[X(t) \neq 0]$ si $a \to +\infty$.

Il y a cependant un lien entre les processus auto-similaires et stationnaires:

Proposition 1.2.5 Si $\{X(t), t \in (0, +\infty)\}$ est H-ss, alors le nouveau processus

$$Y(t) := e^{-tH} X(e^t), \quad t \in \mathbb{R}$$

est stationnaire. Réciproquement, si le processus $\{Y(t), t \in \mathbb{R}\}$ est stationnaire, alors le processus

$$X(t) = t^H Y(\ln t), \quad t \in (0, +\infty)$$

est H-ss.

Preuve: La preuve est assez immédiate. On utilise la définition d'un processus *H*-ss et stationnaire, et on prouve que pour tout $\theta_1, \ldots, \theta_d \in \mathbb{R}$ et pour tout t_1, \ldots, t_d ,

$$\sum_{j=1}^{d} \theta_j Y(t_j + h) \stackrel{\mathrm{d}}{=} \sum_{j=1}^{d} \theta_j Y(t_j),$$

ce qui prouve la stationnarité. L'idée est la même pour prouver la condition suffisante. \blacksquare

Exemple 1.2.6 (Application de la Proposition 1.2.5) On part d'un mouvement brownien standard $\{X(t), t \ge 0\}$. C'est un processus auto-similaire d'indice H = 1/2, donc par la Proposition 1.2.5, le processus

$$Y(t) = e^{-t/2} X(e^t), \quad t \in \mathbb{R}$$

est un processus stationnaire. Il est également gaussien de moyenne nulle, et de fonction d'autocovariance:

$$\mathbb{E}[Y(t_1)Y(t_2)] = e^{-\frac{t_1+t_2}{2}} \mathbb{E}[X(e^{t_1})X(e^{t_2})]$$
$$= e^{-\frac{t_1+t_2}{2}} \min(e^{t_1}, e^{t_2})$$
$$= e^{-\frac{|t_1+t_2|}{2}}$$

On appelle un tel processus $\{Y(t)\}$ processus d'Ornstein-Uhlenbeck.

La proposition précédente permet de construire de nouveaux processus auto-similaires.

On va considérer dans la suite des processus H-ss à accroissements stationnaires.

1.2.1 Processus auto-similaires à accroissements stationnaires

Définition 1.2.7 On dira qu'un processus $\{X(t); t \in T\}$ est à accroissements stationnaires si

$${X(t+h) - X(h); t \in T} \stackrel{d}{=} {X(t) - X(0); t \in T} \quad \forall h \in T.$$

On qualifiera les processus auto-similaires à accroissements stationnaires de processus H-sssi (self-similar with stationary increments).

Exemple 1.2.8 (Mouvements browniens) On retrouve bien entendu dans la classe des processus *H*-sssi les mouvements browniens. En effet, on a vu dans l'exemple 1.2.2 que de tels processus étaient *H*-ss avec H = 1/2. Par définition, un mouvement brownien est à accroissements stationnaires.

On remarquera que si un processus $\{X(t); t \in \mathbb{R}\}$ est *H*-sssi, comme X(0) = 0 (cf. Remarque 1.2.4), on a pour tout *t* fixé:

$$X(-t) = X(-t) - X(0) \stackrel{d}{=} X(0) - X(t) = -X(t)$$

On a donc

$$X(-t) \stackrel{\mathrm{d}}{=} -X(t)$$

pour tout $t \in \mathbb{R}$ fixé.

On va s'intéresser à l'existence des moments de ce genre de processus. Le résultat suivant montre que si l'on désire l'existence de certains moments, cela limite les valeurs admissibles pour H.

Proposition 1.2.9 Si $\{X(t); t \in \mathbb{R}\}$ est H-sssi et si $\mathbb{P}[X(1) \neq 0] > 0$, alors la relation

$$\mathbb{E}[|X(1)|^{\gamma}] < \infty$$

implique

$$\left\{ \begin{array}{ll} H \in (0,1/\gamma) & si \ 0 < \gamma < 1 \\ H \in (0,1] & si \ \gamma \geq 1. \end{array} \right.$$

La preuve de la Proposition 1.2.9 utilise le lemme suivant, dont la démonstration est essentiellement technique (cf. [37]).

Lemme 1.2.10 Si $\{X(t); t \in \mathbb{R}\}$ est H-sssi, alors pour tout $s \neq 0$, pour tout $t \neq 0$,

$$\mathbb{P}[X(s) = 0, X(t) = 0] = \mathbb{P}[X(1) = 0]$$
$$\mathbb{P}[X(s) \neq 0, X(t) \neq 0] = \mathbb{P}[X(1) \neq 0]$$

Preuve de la Proposition 1.2.9:

PREMIER CAS: $\gamma \in (0, 1)$.

Comme $\gamma < 1$, $(a+b)^{\gamma} \le a^{\gamma} + b^{\gamma}$ pour tout $a, b \ge 0$, avec inégalité stricte si a > 0 et b > 0. On a donc en particulier

$$|X(2)|^{\gamma} \le |X(2) - X(1)|^{\gamma} + |X(1)|^{\gamma},$$

et cette inégalité est stricte sur l'événement $A := [X(1) \neq 0, X(2) - X(1) \neq 0]$. Par le lemme 1.2.10 et par l'hypothèse $\mathbb{P}[X(1) \neq 0] > 0$, on a $\mathbb{P}[A] > 0$. De ce fait,

$$\mathbb{E}[|X(2)|^{\gamma}] < \mathbb{E}[|X(2) - X(1)|^{\gamma}] + \mathbb{E}[|X(1)|^{\gamma}].$$
(1.6)

Par la stationnarité des accroissements et le fait que X(0) = 0 p.s., $X(2) - X(1) \stackrel{d}{=} X(1)$. Cela implique que (1.6) se ramène à $\mathbb{E}[|X(2)|^{\gamma}] < 2\mathbb{E}[|X(1)|^{\gamma}]$. Or, par l'autosimilarité du processus, on a $\mathbb{E}[|X(2)|^{\gamma}] = 2^{H\gamma}\mathbb{E}[|X(1)|^{\gamma}]$. On arrive donc finalement à $2^{H\gamma}\mathbb{E}[|X(1)|^{\gamma}] < 2\mathbb{E}[|X(1)|^{\gamma}]$, ce qui implique, du fait que $\mathbb{E}[|X(1)|^{\gamma}] < \infty$ par hypothèse, $2^{H\gamma} < 2$ ou encore $H < 1/\gamma$.

Deuxième cas: $\gamma \geq 1$

Si $\gamma \geq 1$, alors nécessairement $\mathbb{E}[|X(1)|^p] < \infty$ pour tout p < 1, et donc par le premier cas, H < 1/p pour tout p < 1, ce qui implique finalement $H \leq 1$.

La condition $\mathbb{P}[X(1) \neq 0] > 0$ dans la Proposition 1.2.9 est relativement faible. Elle est satisfaite dès que le processus n'est pas identiquement nul sur \mathbb{R}^+ . On a en particulier le corollaire suivant:

Corollaire 1.2.11 Soit $\{X(t); t \in \mathbb{R}\}$ un processus α -stable non dégénéré, $\alpha \in (0, 2]$ et *H*-sssi. Alors

$$\begin{array}{rcl} \alpha < 1 & \Rightarrow & 0 < H \leq 1/\alpha, \\ \alpha \geq 1 & \Rightarrow & 0 < H \leq 1. \end{array}$$

Preuve: Il suffit d'utiliser les Propositions 1.2.9 et 1.1.11.

On peut voir (cf. [37]) que pour tout couple (α, H) , il existe au moins un processus α stable *H*-sssi. On va voir cela dans le cas $\alpha = 2$, où il y aura même unicité du processus *H*-sssi. Ce sera le mouvement brownien fractionnaire d'indice *H*. On peut voir par contre que si $\alpha < 2$, il peut correspondre plusieurs processus *H*-sssi pour un même couple (α, H) . On identifiera par convention des processus différant seulement par une constante multiplicative.

1.2.2 Mouvement brownien fractionnaire

Ce seront des processus auto-similaires gaussiens à accroissements stationnaires, mais non nécessairement indépendants (comme c'est la cas pour le mouvement brownien). Il s'agit d'un type de processus très important au niveau applications, surtout à cause de son caractère à longue mémoire. Ces processus ont été étudiés la première fois par Kolmogorov. Ils ont été utilisés en sciences naturelles, particulièrement en hydrologie, géophysique, turbulence mais aussi en économie dans différents modèles.

Il est bien connu qu'un processus gaussien $\{X(t), t \in \mathbb{R}\}$ est entièrement déterminé par sa fonction de moyenne $(\mu(t) = \mathbb{E}[X(t)])$ et sa fonction de covariance (A(t,s) = Cov(X(t), X(s))), cf. par exemple [40]). On peut voir cela en considérant la fonction caractéristique des distributions fini-dimensionnelles du processus, fonction caractéristique ayant la forme suivante:

$$\mathbb{E}[\exp(i\sum_{j=1}^{n}\theta_j X(t_j))] = \exp\left\{-\frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n}A(t_j,t_k)\theta_k\theta_k + i\sum_{j=1}^{n}\mu(t_j)\theta_j\right\}$$

pour tout t_1, \ldots, t_n , pour tout $\theta_1, \ldots, \theta_n$, pour tout $n \in \mathbb{N}_0$. La fonction $A(t, s), s, t \in \mathbb{R}$ est semi-définie positive, c'est-à-dire pour tout $n \in \mathbb{N}_0, t_1, \ldots, t_n \in \mathbb{R}, x_1, \ldots, x_n \in \mathbb{R}$,

$$\sum_{j=1}^n \sum_{k=1}^n A(t_j, t_k) x_j x_k \ge 0.$$

Donc à toute couple de fonctions $(\mu(t), A(t, s))$ où A est semi-définie positive, correspond un et un seul processus gaussien.

Fixons maintenant $H \in (0, 1)$. On peut voir que la fonction $A(t, s) := |s|^{\alpha} + |t|^{\alpha} - |s - t|^{\alpha}, s, t \in \mathbb{R}$ est semi-définie positive, pour tout $\alpha \in (0, 2]$. La preuve de ce résultat est un peu technique et se trouve dans [37] (Lemme 2.10.8, p. 106).

Par ce résultat, il existe donc un et un seul processus gaussien $\{X(t), t \in \mathbb{R}\}$ de moyenne nulle et de fonction de covariance

$$R_H(t,s) = \frac{1}{2} (|s|^{2H} + |t|^{2H} - |s - t|^{2H}) \operatorname{Var}(X(1)).$$
(1.7)

1.2. Processus auto-similaires

On peut voir que le processus ainsi défini est *H*-sssi. En effet, il est à accroissements stationnaires car si $s, t \in \mathbb{R}$, X(s) - X(t) est de loi normale de moyenne nulle et de variance

$$\operatorname{Var}[X(s) - X(t)] = \operatorname{Var}[X(s)] + \operatorname{Var}[X(t)] - 2R_H(t,s)$$
$$= |s - t|^{2H} \operatorname{Var}[X(1)]$$

par (1.7). Pour voir le caractère *H*-ss, il suffit de voir que pour tout a > 0, pour tout $t_1, \ldots, t_n \in \mathbb{R}$, pour tout $n \in \mathbb{N}_0$,

$$(X(at_1),\ldots,X(at_n)) \stackrel{\mathrm{d}}{=} (a^H X(t_1),\ldots,a^H X(t_n)).$$

Comme il s'agit de deux vecteurs gaussiens, il suffit de voir que les vecteurs moyennes et matrices de covariances coïncident. C'est déjà le cas pour les moyennes (nulles par hypothèse), et pour ce qui est des covariances, on a:

$$Cov(X(at_i), X(at_j)) = \frac{1}{2} (|at_i|^{2H} + |at_j|^{2H} - |at_i - at_j|^{2H}) Var(X(1))$$

= $a^{2H} Cov(X(t_i), X(t_j))$
= $Cov(a^H X(t_i), a^H X(t_j)).$

Si l'on suppose maintenant que H = 1, que la fonction de covariance est donnée par $R_H(t)$ mais que la moyenne $\mu(t)$ vaut t, il est facile de voir que le processus gaussien engendré est encore H-sssi (preuve essentiellement la même que lorsque $H \neq 1$). On appelle ce type de processus mouvement brownien fractionnaire:

Définition 1.2.12 (Mouvement brownien fractionnaire) Soit $H \in (0, 1]$. Un processus gaussien de moyenne $\mu(t)$ nulle si H < 1 et égale à t si H = 1 et de fonction de covariance égale à $R_H(t, s)$ est appelé mouvement brownien fractionnaire.

Le résultat suivant (cf. [37]) implique qu'il n'y a pas d'autre processus gaussien H-sssi pour H < 1 (grâce à la caractérisation des processus gaussiens par leurs fonctions de moyenne et de covariance).

Lemme 1.2.13 Supposons que $\{X(t), t \in \mathbb{R}\}$ soit un processus H-sssi non dégénéré et tel que $\operatorname{Var}(X(1)) < \infty$. Alors $H \in (0, 1]$, X(0) = 0 p.s. et $\operatorname{Cov}(X(s), X(t)) = R_H(s, t)$. De plus, pour tout $t \in \mathbb{R}$,

$$\mathbb{E}[X(t)] = 0 \quad si \ H < 1, \tag{1.8}$$

$$X(t) = tX(1)$$
 p.s. $si H = 1.$ (1.9)

Preuve: On a toujours, si $s, t \in \mathbb{R}$,

$$\mathbb{E}[X(s)X(t)] = \frac{1}{2} \left\{ \mathbb{E}[X(s)^2] + \mathbb{E}[X(t)^2] - \mathbb{E}[(X(s) - X(t))^2] \right\}.$$
 (1.10)

Par la stationnarité des accroissements et l'auto-similarité du processus (et le fait que X(0) = 0 p.s. par la remarque 1.2.4),

$$(1.10) = \frac{1}{2} \left\{ \mathbb{E}[X(s)^2] + \mathbb{E}[X(t)^2] - \mathbb{E}[(X(t-s) - X(0))^2] \right\} \\ = \frac{1}{2} \left\{ |s|^{2H} + |t|^{2H} - |s-t|^{2H} \right\} \mathbb{E}[X(1)^2].$$
(1.11)

En effet, $\mathbb{E}[X(t)^2] = \mathbb{E}[|t|^{2H}X(\operatorname{sign}(t))^2]$, et par la remarque page 14, $X(1) \stackrel{\mathrm{d}}{=} X(-1)$, ce qui implique $\mathbb{E}[X^2(1)] = \mathbb{E}[X^2(-1)]$. Par la Proposition 1.2.9 et l'hypothèse sur l'existence des variances du processus, on a $0 < H \leq 1$.

Supposons que H < 1. Par la stationnarité des accroissements, $\mathbb{E}[X(1)] = \mathbb{E}[X(2) - X(1)] = (2^H - 1)\mathbb{E}[X(1)]$, et donc nécessairement $\mathbb{E}[X(1)] = 0$ (puisque $H \neq 0$). Comme $\mathbb{E}[X(t)] = |t|^H \mathbb{E}[X(\operatorname{sign}(t))]$ et $\mathbb{E}[X(-1)] = -\mathbb{E}[X(1)]$, on en conclut que $\mathbb{E}[X(t)] = 0$ pour tout $t \in \mathbb{R}$. De ce fait, (1.11) nous donne bien l'expression annoncée de la fonction de covariance et on a prouvé (1.8).

Supposons maintenant que H = 1. Dans ce cas, (1.11) devient:

$$\mathbb{E}[X(s)X(t)] = \frac{1}{2} \{s^2 + t^2 - (s-t)^2\} \mathbb{E}[X^2(1)] \\ = st \mathbb{E}[X^2(1)]$$

et donc

$$\mathbb{E}[(X(t) - tX(1))^2] = \mathbb{E}[X^2(t)] - 2t\mathbb{E}[X(t)X(1)] + t^2\mathbb{E}[X^2(1)]$$

= $t^2\mathbb{E}[X^2(1)] - 2t^2\mathbb{E}[X^2(1)] + t^2\mathbb{E}[X^2(1)]$
= 0.

Cela implique directement (1.9). On a alors également

$$Cov(X(s), X(t)) = st Var[X(1)],$$

ce qui nous donne bien la forme annoncée pour la fonction de covariance.

Par ce résultat, tous les processus H-sssi gaussiens ont nécessairement une fonction d'autocovariance d'une certaine forme. Ils sont donc déterminés à une constante multiplicative près, et peuvent différer par leur moyenne si H = 1. On peut grâce à ce lemme donner une définition d'un mouvement brownien fractionnaire équivalente à la précédente.

Définition 1.2.14 (Mouvement brownien fractionnaire) Un processus *H*-sssi gaussien tel que $0 < H \leq 1$ est appelé mouvement brownien fractionnaire (FBM, fractional Brownian motion) et est noté $\{B_H(t), t \in \mathbb{R}\}$. Il est dit standard si $\operatorname{Var}[X(1)] = 1$.

Remarque 1.2.15 Il existe également des processus H-sssi de variance finie et non gaussiens (cf. des exemples donnés dans [37] et les références qui y sont citées).

Exemple 1.2.16 (Mouvement brownien) Lorsque H = 1/2, on retrouve un mouvement brownien puisque l'on a alors

$$\mathbb{E}[X(t)X(s)] = \frac{1}{2}(|s|+|t|-|s-t|)\operatorname{Var}[X(1)] = \begin{cases} \operatorname{Var}[X(1)]\min(s,t) & \text{si sign}(s) = \operatorname{sign}(t), \\ 0 & \text{sinon.} \end{cases}$$

On a vu que cela caractérisait le mouvement brownien.

Exemple 1.2.17 Lorsque H = 1, on a par le Lemme 1.2.13 que $B_1(t) = tB_1(1)$ p.s., et comme $B_1(1)$ est de loi normale $\mathcal{N}(\mu, \operatorname{Var}[B_1(1)])$, on a finalement $B_1(t) = t(Z + \mu)$ où $Z \sim \mathcal{N}(0, 1)$ et $\mu \in \mathbb{R}$. Ici, il reste encore la moyenne à déterminer. La valeur du processus en un seul instant détermine entièrement le processus. A cause de cette particularité, on supposera toujours dans la suite que $H \neq 1$.

1.2.3 Représentation intégrale d'un mouvement brownien fractionnaire

On veut obtenir une représentation intégrale du mouvement brownien fractionnaire de la forme:

$$\int_{\mathbb{R}} f_t(x) dB_x$$

où B_x est un mouvement brownien standard, et où $f_t(x)$ est une fonction déterministe satisfaisant $\int_{\mathbb{R}} f_t^2(x) dx < \infty$.

Par les propriétés de l'intégrale stochastique,

$$\mathbb{E}[I(f_s)I(f_t)] = \int_{\mathbb{R}} f_s(x)f_t(x)dx.$$
(1.12)

Dans ce qui suit, on adoptera la convention que $0^0 = 0$.

Proposition 1.2.18 (Représentation intégrale) Soit $H \in (0, 1)$. Alors tout mouvement brownien fractionnaire standard $\{B_H(t), t \in \mathbb{R}\}$ possède la représentation intégrale:

$$\frac{1}{C_1(H)} \int_{\mathbb{R}} (((t-x)_+)^{H-1/2} - ((-x)_+)^{H-1/2}) dB_x, \quad \forall t \in \mathbb{R}$$
(1.13)

оù

$$C_1(H) = \left(\int_0^{+\infty} \left((1+x)^{H-1/2} - x^{H-1/2}\right)^2 dx + \frac{1}{2H}\right)^{\frac{1}{2}}.$$

Lorsque H = 1/2, $C_1(1/2) = 1$ et (1.13) doit s'interpréter comme $\int_0^t dB_x$ si $t \ge 0$ et $-\int_t^0 dB_x$ si t < 0.

Preuve: Soit X(t) := (1.13) et $f_t(x)$ l'intégrante. Le cas H = 1/2 est trivial. Supposons donc que 0 < H < 1 avec $H \neq 1/2$. Pour que (1.13) ait un sens, il suffit que $f_t(x)$ soit

localement de carré sommable, ce qui est bien le cas ici. Pour que $\mathbb{E}[X^2(t)]$ existe, il faut que

$$\int_{\mathbb{R}} f_t^2(x) \, dx < \infty. \tag{1.14}$$

On a que $f_t(x)$ se comporte asymptotiquement pour $x \to -\infty$ comme $(H-1/2)(-x)^{H-3/2}$. En effet,

$$\lim_{x \to -\infty} \frac{(t-x)^{H-1/2} - (-x)^{H-1/2}}{(-x)^{H-3/2}} = \lim_{y \to +\infty} \frac{(t+y)^{H-1/2} - (y)^{H-1/2}}{y^{H-3/2}}$$
$$= \lim_{y \to +\infty} \frac{\frac{\mathrm{d}}{\mathrm{dt}} (t+y)^{H-1/2}_{|_{t=0}} t + g(t,\bar{t})}{y^{H-3/2}} = (H-1/2) + \lim_{y \to +\infty} \frac{g(t,\bar{t})}{y^{H-3/2}} = H - 1/2$$

où $g(t,\bar{t}) = \frac{1}{2} \frac{d}{dt} (t+y)_{|t=\bar{t}}^{H-1/2} t^2$ pour un certain $\bar{t} \in (0,t)$. La limite vaut bien H - 1/2 puisque t est fixé dans le raisonnement. Comme $(-x)^{H-3/2}$ est de carré intégrable près de $-\infty$, f_t^2 est intégrable près de $-\infty$. Si maintenant x tend vers t, on voit facilement que $f_t(x)$ se comporte asymptotiquement comme $(t-x)_+^{H-1/2}$, qui est bien de carré sommable près de x = t. Un raisonnement analogue peut être tenu pour $x \to +\infty$ ou $x \to 0$. On en conclut finalement que l'on a bien (1.14).

Du fait que l'intégrale stochastique (1.13) est de loi normale de moyenne nulle pour tout $t \in \mathbb{R}$, il suffit donc maintenant de vérifier que le processus $\{X(t)\}$ a une fonction de covariance donnée par (1.7) avec Var [X(1)] = 1, ce qui prouvera alors que $\{X(t)\}$ est bien un mouvement brownien fractionnaire standard d'indice H.

On voit facilement que X(0) = 0 p.s., et par (1.12), que si t > 0,

$$\begin{split} \mathbb{E}[X^{2}(t)] &= \frac{1}{C_{1}(H)^{2}} \int_{\mathbb{R}} (((t-x)_{+})^{H-1/2} - ((-x)_{+})^{H-1/2})^{2} dx \\ &= \frac{1}{C_{1}(H)^{2}} t^{2H-1} \int_{\mathbb{R}} \left(\left(1 - \frac{x}{t}\right)_{+}^{H-1/2} - \left(-\frac{x}{t}\right)_{+}^{H-1/2} \right)^{2} dx \\ &= \frac{1}{C_{1}(H)^{2}} t^{2H} \int_{\mathbb{R}} \left((1-u)_{+}^{H-1/2} - (-u)_{+}^{H-1/2} \right)^{2} du \\ &= \frac{1}{C_{1}(H)^{2}} t^{2H} \left[\int_{-\infty}^{0} \left((1-u)^{H-1/2} - (-u)^{H-1/2} \right)^{2} du + \int_{0}^{1} (1-u)^{2H-1} du \right] \\ &= \frac{1}{C_{1}(H)^{2}} t^{2H} \left[\int_{-\infty}^{0} \left((1+x)^{H-1/2} - x^{H-1/2} \right)^{2} dx + \frac{1}{2H} \right] \\ &= t^{2H} \end{split}$$

où on a utilisé successivement le changement de variable u = x/t, le fait que $\int_0^1 (1 - u)^{2H-1} du = 1/2H$, et la définition de $C_1(H)$. Un raisonnement analogue montre que $\mathbb{E}[X^2(t)] = |t|^{2H}$ si t < 0.

1.2. Processus auto-similaires

Maintenant, $\mathbb{E}[(X(t)-X(s))^2]=\mathbb{E}[Z^2]$ où

$$Z := \frac{1}{C_1(H)} \int_{\mathbb{R}} (((t-x)_+)^{H-1/2} - ((s-x)_+)^{H-1/2}) dB(x).$$

 $\mathbb{E}[Z^2]$ vaut donc:

$$\mathbb{E}[Z^2] = \frac{1}{C_1(H)^2} \int_{\mathbb{R}} (((t-x)_+)^{H-1/2} - ((s-x)_+)^{H-1/2})^2 dx$$

= $\frac{1}{C_1(H)} \int_{\mathbb{R}} (((t-s-u)_+)^{H-1/2} - ((-u)_+)^{H-1/2})^2 du$
= $|t-s|^{2H}$

en faisant le changement de variable u = x - s et par ce qui précède pour le calcul de $\mathbb{E}[X^2(t)]$. De ce fait,

$$Cov(X(t), X(s)) = \mathbb{E}[X(t)X(s)] = \frac{1}{2} \{ \mathbb{E}[X^2(t)] + \mathbb{E}[X^2(s)] - \mathbb{E}[(X(t) - X(s))^2] \}$$
$$= \frac{1}{2} \{ |t|^{2H} + |s|^{2H} - |s - t|^{2H} \},$$

ce qui est bien l'expression désirée (1.7) avec $\operatorname{Var}[X(1)] = 1$.

1.2.4 Bruit gaussien fractionnaire

Puisque tout mouvement brownien fractionnaire $\{B_H(t), t \in \mathbb{R}\}$ a des accroissements stationnaires, ses accroissements

$$Y_n := B_H(n+1) - B_H(n), \quad n \in \mathbb{Z}$$

constituent donc un processus (à temps discret³) stationnaire. Ce processus $\{Y_n, n \in \mathbb{Z}\}$ est appelé bruit gaussien fractionnaire (fractional Gaussian noise, FGN). On le qualifie de standard lorsque Var $[Y_n] \equiv 1$ (c'est-à-dire lorsque le mouvement brownien sous-jacent est standard).

La représentation intégrale que l'on vient de voir dans la section précédente peut se réécrire pour $\{Y_n\}$:

$$Y_n \stackrel{\mathrm{d}}{=} \frac{\operatorname{Var}[Y_n]}{C_1(H)} \int_{-\infty}^{n+1} ((n+1-x)_+^{H-1/2} - (n-x)_x^{H-1/2}) dB_x \quad \forall n \in \mathbb{Z}.$$

Un bruit gaussien fractionnaire est donc un processus gaussien stationnaire de moyenne nulle et de variance égale à Var $[B_H(1)]$, où $B_H(.)$ est le processus sous-jacent.

On définit les fonctions suivantes:

³On pourrait également définir un processus à temps continu en remplaçant simplement $n \in \mathbb{Z}$ par $t \in \mathbb{R}$ dans la définition précédente

Définition 1.2.19 La fonction d'autocovariance du processus, $r : \mathbb{Z} \to \mathbb{R}$, est définie par $r(n) := \mathbb{E}[Y_0Y_n]$ pour tout $n \in \mathbb{Z}$. La fonction d'autocorrélation peut se définir simplement comme $\frac{r(n)}{\operatorname{Var}[Y_n]} = \frac{r(n)}{\operatorname{Var}[Y_0]}$ (puisque le processus est stationnaire).

On a l'expression analytique suivante pour la fonction d'autocovariance.

Proposition 1.2.20 Tout bruit gaussien fractionnaire est de fonction d'autocovariance

$$r(n) = \frac{\operatorname{Var}\left[Y_0\right]}{2}(|n+1|^{2H} + |n-1|^{2H} - 2|n|^{2H}), \quad n \in \mathbb{Z}.$$

Preuve: Supposons que Var $[Y_0] = 1$. On a par définition de r:

$$\begin{aligned} r(n) &= \mathbb{E}[Y_0 Y_n] \\ &= \mathbb{E}[B_H(1)(B_H(n+1) - B_H(n))] \\ &= R_H(1, n+1) - R_H(1, n) \\ &= \frac{1}{2}(1 + |n+1|^{2H} - |n|^{2H} - 1 - |n|^{2H} + |n+1|^{2H}) \\ &= \frac{1}{2}(|n+1|^{2H} - 2|n|^{2H} + |n-1|^{2H}), \end{aligned}$$

où R_H , donné par (1.7), est la fonction de covariance de $\{B_H\}$.

Proposition 1.2.21 (Comportement asymptotique de la fonction d'autocovariance) Si $H \neq 1/2$, on a le comportement asymptotique suivant pour $n \rightarrow +\infty$:

$$r(n) \sim \operatorname{Var}[Y_0] H(2H-1)n^{2H-2}.$$

Rappelons que dans le cas où H = 1/2, le mouvement brownien étant un processus à accroissements indépendants, $r(n) \equiv 0$.

Preuve: Il suffit de reprendre d'expression de r(n) obtenue dans la proposition précédente, de mettre n^{2H-2} en évidence dans l'expression pour obtenir:

$$r(n) = \frac{\operatorname{Var}\left[Y_0\right]}{2} n^{2H-2} n^2 \left[\left(1 + \frac{1}{n}\right)^{2H} - 2 + \left(1 - \frac{1}{n}\right)^{2H} \right].$$

On développe alors $(1+x)^{2H} + (1-x)^{2H}$ en série de Taylor autour de x = 0, on remplace x par $\frac{1}{n}$ et on passe à la limite pour $n \to +\infty$.

On voit donc que r(n) décroît à l'infini comme une puissance et tend vers 0. Pour $H \in (1/2, 1), r(n)$ est toujours positive, et comme l'exposant 2H - 2 > -1, la série $\sum_{n=0}^{+\infty} r(n)$ diverge. On dit dans ce cas que le processus est "à longue mémoire" (long range dependence, long memory process). Les processus à longue mémoire constituent un sujet important en statistique (cf. [3]). Si par contre $H \in (0, 1/2)$, alors pour tout n suffisamment grand, r(n) < 0. On parle alors de processus avec dépendance négative. Dans ce cas, on peut voir que la série $\sum_{n=0}^{+\infty} r(n)$ converge vers $-\operatorname{Var}[Y_0]/2$, et converge même absolument. La figure 1.1 donne la fonction d'autocovariance d'un bruit gaussien fractionnaire standard pour différentes valeurs de H.



Figure 1.1: Fonction d'autocovariance r(n).

1.2.5 Propriété des trajectoires et simulations

Lorsque deux processus $\{X(t), t \in T\}$ et $\{Y(t), t \in T\}$ satisfont pour tout $t \in T$, $\mathbb{P}[X(t) = Y(t)] = 1$, on dit que l'un est une modification de l'autre. Il est bien connu (voir par exemple [40]), qu'un mouvement brownien a une modification dont les trajectoires sont continues, mais toute modification possède des trajectoires différentiables nulle part.

On dira qu'un processus $\{X(t), t \in T\}$ est höldérien d'ordre $\gamma \in (0, 1)$ s'il existe une variable aléatoire h positive p.s. et une constante $\delta > 0$ telles que

$$\mathbb{P}\left[\sup_{\substack{0 < t - s < h \\ s, t \in T}} \frac{|X(t) - X(s)|}{|t - s|^{\gamma}} \leq \delta\right] = 1$$

Dans [13] se trouve le résultat suivant:

Lemme 1.2.22 Si un processus $\{X(t)\}$ satisfait:

$$\exists \delta > 0, \varepsilon > 0, C > 0 \text{ tels que } \mathbb{E}[|X(t) - X(s)|^{\delta}] \le C|t - s|^{1+\varepsilon} \quad \forall s, t$$

alors le processus $\{X(t)\}$ possède une modification dont les trajectoires sont höldérienne d'ordre $\gamma \in [0, \varepsilon/\delta]$.

Grâce à ce résultat, on peut prouver (cf. [13]) que tout mouvement brownien fractionnaire $\{B_H(t)\}$ possède une modification dont les trajectoires sont höldériennes d'ordre $\gamma \in [0, H)$. En effet, on fixe $\gamma \in (0, H)$ et on utilise l'autosimilarité et la stationnarité des accroissements pour obtenir:

$$\mathbb{E}[|B_H(t) - B_H(s)|^{1/\gamma}] = \mathbb{E}[|B_H(t-s)|^{1/\gamma}] = |t-s|^{H/\gamma} \mathbb{E}[|B_H(1)|^{1/\gamma}].$$

Par le Lemme 1.2.22, le processus est höldérien d'ordre $\beta < H - \gamma$. Puisque γ peut être pris arbitrairement petit, on a le résultat annoncé.

Simulation d'un mouvement brownien fractionnaire

Plusieurs méthodes permettent de simuler un mouvement brownien fractionnaire. Par ce qu'on a vu précédemment, il est facile de voir que la matrice de covariance Σ d'un bruit gaussien fractionnaire sur [0, n] est de la forme

$$\Sigma_{ij} = \gamma(i-j)$$

pour une certaine fonction γ . Pour générer n observations d'un bruit gaussien fractionnaire (ou plus généralement d'un processus gaussien stationnaire), on peut d'abord générer n variables aléatoires i.i.d. de loi $\mathcal{N}(0,1)$, disons un vecteur Z. Ensuite, on décompose Σ en un produit $\Sigma = LL^T$ où L est une matrice triangulaire inférieure (décomposition de Choleski⁴). On multiplie ensuite Z à gauche par L, et le vecteur LZest alors bien un vecteur gaussien de moyenne nulle et de matrice de covariance Σ . Le problème est que pour un bruit gaussien fractionnaire (et un processus à longue mémoire en général), on aime bien générer de longues séries, ce qui donne alors à Σ une très grande dimension, nécessitant alors beaucoup de place mémoire. On peut également avoir des problèmes numériques pour la décomposition de Choleski (Σ n'est pas nécessairement une matrice très creuse). Il existe donc d'autres méthodes évitant de telles grandes matrices. Pour ce qui est du bruit gaussien fractionnaire, on peut utiliser une méthode basée sur une représentation intégrale (dans \mathbb{C}) du processus (obtenue en gros par transformation de Fourier à partir de la représentation vue auparavant). Cette représentation peut faire intervenir la densité spectrale, i.e. la transformée de Fourier de la fonction d'autocovariance. On peut alors utiliser la technique de la Fast Fourier Transform. Simuler un bruit gaussien ou un mouvement brownien est en fait un problème en soit. Dans [37, 13, 3] sont discutées des méthodes de simulation.

Les figures 1.2, 1.3 représentent des trajectoires de mouvements browniens fractionnaires simulés à partir d'algorithmes de simulations provenant de [3] (basés sur la dernière méthode citée ci-dessus). On "voit" bien que la trajectoire du mouvement fractionnaire avec H = 0.8 est beaucoup plus "lisse" (a un exposant d'Hölder local plus grand) que celle du processus avec H = 0.5 et encore plus que lorsque H = 0.2.

La figure 1.4 donne à gauche un mouvement brownien fractionnaire simulé suivant cette fois 40000 itérations, et à droite ce même mouvement auquel on a fait subir un changement d'échelle d'un facteur a = 1/2. Plus précisément, on dilaté d'un facteur 2 suivant l'axe des x le processus (en ne prenant que les 20000 premiers points), ce qui correspond à tracer X(t/2), et on l'a multiplié par 2^{0.2}. Puisque $X(t) \stackrel{d}{=} X(t/2)2^{0.2}$, on doit avoir l'impression que l'allure générale des deux processus est la même. C'est bien une auto-similarité statistique et non déterministe comme pour certains fractals géométriques, les deux trajectoires ne sont donc pas identiques, elles ont juste la même allure générale, traduisant les propriétés statistiques identiques des processus sous-jacents. Idem avec un mouvement brownien fractionnaire avec H = 0.8 (figure 1.5).

⁴en analyse numérique matricielle, cas particulier d'une décomposition LU dans le cas de matrices définies positives, où il y a moyen d'avoir $L = U^T$



Figure 1.2: Simulations d'un mouvement brownien fractionnaire d'indice H = 0.2, H = 0.5 et H = 0.8.



Figure 1.3: Bruits gaussiens fractionnaires correspondants (1000 premières itérations)



Figure 1.4: Auto-similarité d'un mouvement brownien fractionnaire pour H = 0.2.



Figure 1.5: Auto-similarité d'un mouvement brownien fractionnaire pour H = 0.8.

1.2.6 Intégrale stochastique par rapport à un mouvement brownien fractionnaire

Les processus stochastique auto-similaires sont devenus populaires dans beaucoup de domaines d'applications (physique, télécommunications, finance), et il y a de ce fait une demande naturelle de calcul stochastique basé sur ce type de processus. En particulier, des équations différentielles stochastiques contenant des mouvements browniens fractionnaires seraient les bienvenues. Le problème est que le mouvement brownien fractionnaire n'est pas une semi-martingale dès que l'indice de Hurst H est différent de 1/2 (cf. [13] pour une preuve et des références de preuves). Cela empêche donc d'avoir recours à la construction standard d'intégrale stochastique (voir [35]).

Diverses approches pour définir une intégrale stochastique satisfaisante ont été proposées jusqu'à présent, mais il n'y a pas encore d'unanimité concernant une méthode pleinement satisfaisante au point de vue théorique. Certaines personnes se sont également penchées sur le problème d'un modèle de Black-Scholes fractionnaire pour différentes approches de l'intégrale stochastique par rapport à un FBM.

Tout ceci est évidemment crucial en finance puisque la notion d'intégrale stochastique est liée aux concepts importants d'opportunités d'arbitrage, de marchés complets,... Voir [13] pour un tour de la question.

Chapitre 2 Processus multifractals

En deux mots, on dira que l'on a affaire à une entité fractale en présence d'"irrégularités dans sa forme et sa complexité"¹. On parle par exemple d'objets fractals géométriques lorsque ces objets ont une dimension (à définir, voir appendice) non entière, contrairement aux objets classiques (ceux que l'on a plus l'habitude de rencontrer en géométrie) qui ont une dimension entière. Il y a également la notion d'auto-similarité qui caractérise certains fractals géométriques (comme le célèbre ensemble de Mandelbrot par exemple, ou encore l'ensemble triadique de Cantor). Si l'on s'intéresse maintenant aux processus stochastiques, on dira que l'on a affaire à un processus fractal si les trajectoires de celuici possèdent également une certaine complexité. Par exemple, un mouvement brownien fractionnaire $\{B_H(t)\}$ sera considéré comme processus fractal. Ses trajectoires sont continues partout mais différentiables nulle part (elles sont en fait Höldériennes d'exposant H), et leurs graphes auront une dimension fractale non entière. Comme on l'a vu, le paramètre de Hurst H (paramètre d'auto-similarité) est intimement lié à la régularité du processus. La variation du processus sur un petit intervalle de temps Δt sera de l'ordre de $(\Delta t)^H$ (comportement d'échelle). Il s'agit ici d'une auto-similarité statistique et non géométrique, c'est-à-dire qu'un changement d'échelle d'une trajectoire du processus ne sera pas exactement identique à la trajectoire de départ, mais aura la même allure, car le processus "rescalé" aura les mêmes propriétés statistiques.

Parmi les processus à longue mémoire, les mouvements browniens fractionnaires ont beaucoup de succès à cause de leur auto-similarité qui en facilite l'analyse. Cette autosimilarité donne lieu à des comportement d'échelle avec exposants linéaires. Or dans la pratique, les séries observées dans les applications présentent parfois des comportements d'échelle (locaux) avec exposants non linéaires. Cela limite l'utilisation du FBM en modélisation, et permet d'introduire une classe plus générale de processus: les processus multifractals.

Les processus multifractals ont eu des applications dans divers domaines: turbulence (en théorie de l'intermittence, voir section 4.3.1), traitement d'image, données médicales, géophysique, et en modélisation des marchés financiers.

 $^{^{1}}$ cf.[36]

La théorie mathématique de ces processus a surtout été étudiée durant ces dix dernières années, même si leur utilisation en turbulence est plus ancienne. Il y a donc jusqu'à présent relativement peu de références bibliographiques traitant du sujet dans son ensemble. Les sections qui suivent reprennent des concepts et résultats provenant essentiellement de [36, 27, 8, 17]. On peut également consulter [38, 39] ainsi que les références présentes dans [36].

2.1 Exposants de singularité

On considère un processus $\{X(t), t \in T\}$ sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P}_{\Omega})$ où T est supposé être un intervalle fermé borné de \mathbb{R} . On supposera sans perte de généralité que T = [0, 1]. Ci-dessous, nous donnons quelques exposants de singularités (cf. [36]). Ces exposants sont reliés aux comportement d'échelle locaux.

Définition 2.1.1 L'exposant de Hölder local en le point t d'une trajectoire fixée du processus $\{X(t), t \in T\}$ est défini par

$$H(t) := \sup\{h \mid X \in C_t^h\}$$

où l'on dira que $X \in C_t^h$ s'il existe un polynôme P_t (= $P_t(h)$) et un voisinage U_t de t tel que

$$\frac{|X(u) - P_t(u)|}{|u - t|^h}$$

reste borné si $u \in U_t$.

On peut imaginer le cas où pour h donné, le polynôme P_t qui convient dans la définition est constant (par rapport à u). Dans ce cas, si la trajectoire en question de X est de plus continue, cela implique que $P_t = X(t)$. Cela suggère la définition suivante (qui est d'ailleurs parfois prise comme définition de l'exposant d'Hölder local).

Définition 2.1.2 On définit l'exposant h(t) par

$$h(t) := \liminf_{\varepsilon \to 0} \frac{\log_2 \left(\sup_{|u-t| < \varepsilon} |X(u) - X(t)| \right)}{\log_2(2\varepsilon)},$$

avec la convention que $\log_2(0) := -\infty$.

On constate que pour tout h < h(t), on a nécessairement $|X(u) - X(t)| \le C|u - t|^h$ pour une certaine constante C et pour tout u dans un voisinage de t. Cela implique donc déjà $h(t) \le H(t)$.

En fait, si pour tout h tel que $X \in C_t^h$, $P_t = P_t(h) = X(t)$, alors h(t) = H(t). En effet, il suffit de montrer que $h(t) \ge H(t)$. Comme

$$|X(u) - X(t)| \le C_h |u - t|^h \quad \forall h < H(t)$$

2.1. Exposants de singularité

pour une certaine constante C_h et pour $u \in U_{t,h}$ (un voisinage de t dépendant a priori de h), si on suppose par l'absurde que h(t) < H(t), alors il existe h > h(t) tel que pour tout $\varepsilon > 0$ suffisamment petit,

$$\sup_{|u-t|<2\varepsilon} |X(u) - X(t)| \le C_h |u-t|^h \le C_h (2\varepsilon)^h$$

ce qui implique

$$\log_2(\sup_{|u-t|<\varepsilon} |X(u) - X(t)|) \le h \log_2(2\varepsilon) + D_h$$

pour une certaine constante D_h , ou encore

$$\frac{\log_2(\sup_{|u-t|<\varepsilon}|X(u)-X(t)|)}{\log_2(2\varepsilon)} \ge h + \frac{D_h}{\log_2(2\varepsilon)},\tag{2.1}$$

et ce pour tout ε suffisamment petit, ce qui implique que la limite inférieure pour $\varepsilon \to 0$ du membre de gauche de (2.1) est supérieure ou égale à h, entraînant $h(t) \ge h$, une contradiction.

Le résultat suivant (Lemme 2.3 de [36]) vient préciser le lien entre h(t) et H(t).

Lemme 2.1.3 Si $h(t) \notin \mathbb{N}_0$, alors P_t est constant $(P_t = X(t))$ et h(t) = H(t).

Preuve: La preuve donnée ici est un peu différente de celle contenue dans [36]. Montrons que si H(t) > h(t), alors $h(t) \in \mathbb{N}_0$. Par hypothèse, il existe h > h(t), un polynôme P_t (associé à h) et une constante C_h tels que

$$|X(u) - P_t(u)| \le C_h |u - t|^h$$
(2.2)

pour tout u dans un certain voisinage de t. Remarquons que P_t ne peut dès lors pas être constant, puisque si cela avait été le cas, cela aurait impliqué $h \leq h(t)$ comme on vient de le voir dans le raisonnement précédant l'énoncé du Lemme. De ce fait, il existe un naturel $m \geq 1$ et un polynôme Q ne s'annulant pas en u = t tel que

$$Q(u)(u-t)^m = P_t(u) - X(t)$$

(puisque $P_t(u) - X(t)$ est un polynôme en u s'annulant en u = t). Comme

$$X(u) - X(t) = (X(u) - P_t(u)) + (P_t(u) - X(t)),$$

on a que pour u dans un voisinage de t,

$$|X(u) - X(t)| \le C_h |u - t|^h + D|u - t|^m.$$

Si $h \leq m$, alors $|X(u) - X(t)| \leq D''|u - t|^h$ pour une constante D'' et u dans un voisinage de t, impliquant $h \leq h(t)$, une contradiction. On en déduit donc que m < h, et ce raisonnement peut être tenu quel que soit h > h(t) pour lequel il existe un polynôme P_t tel que (2.2) ait lieu. On en conclut finalement que $m \leq h(t)$. Montrons que m < h(t)

est impossible. Par définition de m, si m < h(t), il existe une suite $u_n \to t$ et $\alpha_n \to +\infty$ telles que

$$|X(t) - P_t(u_n)| \ge \alpha_n |u_n - t|^{h(t)}$$

Or, pour tout n suffisamment grand,

$$|X(u_n) - X(t)| \geq |X(t) - P_t(u_n)| - |X(u_n) - P_t(u_n)|$$

$$\geq \alpha_n |u_n - t|^{h(t)} - C_h |u_n - t|^h.$$

Comme par hypothèse, h > h(t), $-C_h |u_n - t|^h > -C_h |u_n - t|^{h(t)}$ pour tout n suffisamment grand, ce qui implique finalement que

$$|X(u_n) - X(t)| \ge (\alpha_n - C_h)|u_n - t|^{h(t)}$$

pour tout n suffisamment grand, contredisant ainsi la définition de h(t). On en conclut donc que h(t) = m, ce qui termine la preuve.

Soit $t \in [0, 1]$ et $k_n(t) := \lfloor t2^n \rfloor$. Alors il est assez clair que $k_n(t)$ est l'unique entier tel que $t \in I_{k_n(t)}^{(n)} := [k_n(t)2^{-n}, (k_n(t)+1)2^{-n})$. Si *n* augmente, ces intervalles deviennent de plus en plus petits (puisque de longueur 2^{-n}). Le but est d'utiliser ces entiers $k_n(t)$ pour trouver une approximation de h(t).

Définition 2.1.4 On définit, pour t < 1, $h_{k_n(t)}^{(n)}$ par

$$h_{k_n(t)}^{(n)} := -\frac{1}{n} \log_2 \left(\sup\{|X(u) - X(t)| : u \in [(k_n(t) - 1)2^{-n}, (k_n(t) + 2)2^{-n})\} \right)$$

On a alors le résultat suivant d'approximation (cf. [36]):

Lemme 2.1.5 $h(t) = \liminf_{n \to \infty} h_{k_n(t)}^{(n)}$.

Preuve: Choisissons, à $\varepsilon > 0$ fixé, un entier *n* tel que $2^{-n+1} \leq \varepsilon < 2^{-n+2}$. Notons, pour *t* fixé, k_n à la place de $k_n(t)$. Dès lors,

$$[(k_n - 1)2^{-n}, (k_n + 2)2^{-n}) \subset [t - \varepsilon, t + \varepsilon) \subset [(k_{n-2} - 1)2^{-n+2}, (k_{n-2} + 2)2^{-n+2}), \quad (2.3)$$

où les deux inclusions sont vérifiées par définition de I_{k_n} et par le choix de n par rapport à ε . On a donc

$$h_{k_{n-2}} \le \frac{\log_2 \sup_{|u-t| < \varepsilon} |X(u) - X(t)|}{\log_2(2\varepsilon)} \le h_{k_n}$$

pour ce couple (ε, n) . Cela prouve le Lemme.

Remarque 2.1.6 Dans ce qui précède, on a considéré une trajectoire fixée du processus pour définir les différents exposants. Ces exposants sont en réalité des variables aléatoires si on considère l'ensemble des trajectoires possibles du processus.

On peut voir que l'exposant d'Hölder local d'un mouvement brownien fractionnaire est (p.s.) identiquement égal à son indice de Hurst H.
Décomposition en ondelettes et exposants de singularité

On peut encore considérer la décomposition en ondelettes de chaque trajectoire X(t), et établir un lien entre le comportement asymptotique des coefficients de la décomposition et l'exposant d'Hölder introduit dans la définition 2.1.2. Ce résultat est classique (voir [36, 40]). On peut alors introduire un autre exposant de régularité lié aux coefficients de la décomposition en ondelettes (voir [36]). Voici l'idée générale de tout ceci.

On introduit le système de fonctions:

$$\psi_{j,k}(t) := 2^{\frac{j}{2}} \psi(2^{j}t - k), \ \phi_{j,k}(t) := 2^{\frac{j}{2}} \phi(2^{j}t - k), \quad k \in \{0, \dots, 2^{j} - 1\}, \ j \in \mathbb{N}$$
(2.4)

où

$$\phi(t) := \begin{cases} 1 & \text{si } 0 \le t \le 1 \\ 0 & \text{sinon} \end{cases} \quad \psi(t) := \begin{cases} 1 & \text{si } 0 \le t < 1/2 \\ -1 & \text{si } 1/2 \le t \le 1 \\ 0 & \text{sinon} \end{cases}$$

On peut voir (cf. [31]) que ces fonctions forment un système orthonormal complet dans $L^2(0, 1)$ et que la convergence des séries de Fourier correspondantes est meilleure que dans la base trigonométrique classique (convergence uniforme vers toute fonction continue sans autre hypothèse que la continuité). On peut voir que l'on peut même se restreindre à utiliser les fonctions $\psi_{k,n}$ (base de Haar). On peut en fait généraliser cette procédure et considérer une fonction ψ intégrable telle que $\int_{\mathbb{R}} \psi(t) dt = 0$. On peut voir que le système résultant des $\psi_{n,k}$ est encore complet. On supposera que le processus Xest défini sur \mathbb{R} et que $k, j \in \mathbb{Z}$ dans (2.4). On a alors le résultat suivant (prouvé dans [36]):

Lemme 2.1.7 Soit $t \in \mathbb{R}$ fixé et $k_n = k_n(t)$. Si $|X(s) - X(t)| = O(|s-t|^h)$ pour $s \to t$ et si la fonction ψ est intégrable et à support compact avec $\int_{\mathbb{R}} \psi(t) dt = 0$, si $\{C_{n,k}\}$ sont les coefficients de Fourier de la trajectoire X(s) dans la base correspondante $\{\psi_{n,k}\}$, alors

$$2^{n/2}|C_{n,k_n}| = 2^n \left| \int_{\mathbb{R}} X(s)\psi(2^n s - k_n) ds \right| = O(2^{-nh}) \quad pour \ n \to \infty.$$

Ceci suggère l'introduction des exposants $(w_{k_n}^{(n)})$ définis par

$$w_{k_n(t)}^{(n)} := -\frac{1}{n} \log_2 |2^{n/2} C_{n,k_n(t)}|.$$

On définit ensuite l'exposant de singularité $w(t) := \liminf_{n \to \infty} w_{k_n(t)}^{(n)}$. Cet exposant permet d'étudier les processus multifractals par des techniques d'ondelettes.

Exposants de singularité pour des mesures

Considérons une mesure μ sur [0, 1] et introduisons sa fonction de répartition

$$\mathcal{M}(t) := \mu([0,t)).$$

Nous nous intéresserons au cas où μ est singulière, c-à-d que la mesure ne possédera pas de fonction de densité car sa fonction de répartition ne sera dérivable nulle part. La mesure μ sera éventuellement une mesure aléatoire, c'est-à-dire une application μ : $(\Omega, \mathcal{F}, \mathbb{P}) \to \mathfrak{M}_{[0,1]}$ où $\mathfrak{M}_{[0,1]}$ désigne l'ensemble des mesures positives sur $([0,1], \mathcal{B})$, où \mathcal{B} est la σ -algèbre de Borel. Dans ce cas, pour tout $\omega \in \Omega$, $\mu(\omega)$ est une mesure sur [0,1], et on demandera que pour tout borélien B, $\mu(B)$ soit une variable aléatoire. $\mathcal{M}(t)$ nous définit ainsi un processus sur [0,1] croissant presque partout. On peut dès lors appliquer les définitions précédentes pour ce processus $\mathcal{M}(t)$, et le supremum intervenant dans la définition de h_{k_n} devient ainsi

$$\max\{\mathcal{M}((k_n(t)+2)2^{-n}) - \mathcal{M}(t), \mathcal{M}(t) - \mathcal{M}((k_n(t)-1)2^{-n})\}.$$

En fait, vu la croissance du processus, on introduira une suite de coefficients $(\alpha_{k_n(t)}^{(n)})$ définis simplement par

$$\alpha_{k_n(t)}^{(n)} := -\frac{1}{n} \log_2\left(\left|\mathcal{M}((k_n(t)+1)2^{-n}) - \mathcal{M}(k_n(t)2^{-n})\right|\right) = -\frac{1}{n} \log_2\left(\mu\left(I_{k_n(t)}^{(n)}\right)\right), \quad (2.5)$$

et l'exposant

$$\alpha(t) := \liminf_{n \to \infty} \alpha_{k_n(t)}^{(n)}.$$
(2.6)

D'autres exposants de singularité peuvent encore être introduits, mesurant comme précédemment les oscillations d'un processus (ou d'une fonction déterministe), et reliés éventuellement à la dimension des trajectoires (ou du graphe de la fonction) (cf. les références citées dans [36]).

Remarque 2.1.8 Le lien entre les différents exposants de singularité n'est pas toujours un problème si évident. Ces exposants tentent tous essentiellement de capturer, de décrire le type de singularité, d'oscillation de la fonction au voisinage d'un point fixé t, mais à chaque fois un peu différemment. On peut trouver des processus (découlant éventuellement de mesures singulières) pour lesquels les exposants de singularités introduits plus haut ne coïncideront pas parfaitement (voir [36]).

2.2 Analyse multifractale

Ce type d'analyse a été introduit dans le contexte de la turbulence, ensuite en physique et en mathématique. On a d'abord découvert que souvent, les exposants de singularité $h_{k_n}^{(n)}$, $\alpha_{k_n}^{(n)}$ et $w_{k_n}^{(n)}$ (qui sont liés aux comportements d'échelle locaux du processus) ne sont pas uniformes en t, c'est-à-dire que les exposants limites h(t), $\alpha(t)$ ou w(t) ne sont pas constants par rapport à t mais prennent tout un ensemble de valeurs distinctes. On est alors en présence d'objets dont l'irrégularité va elle-même avoir une certaine structure. Lorsque l'on a affaire à un processus fractal pour lequel l'un des exposants de régularité varie en fonction du point t considéré de la trajectoire, on parlera de processus multifractal.

2.2. Analyse multifractale

La structure de tels processus est donc très riche, et on va montrer comment on peut la caractériser soit géométriquement en considérant la dimension fractale des t de même exposant de régularité, ou statistiquement, en terme de moments et propriétés d'échelle pour ces moments. Il sera donc possible d'étudier la structure d'un processus multifractal (ou de mettre en évidence des propriétés multifractales d'un processus) à l'aide de propriétés statistiques, et donc pas seulement en regardant le comportement des trajectoires au voisinage d'un point, hautement sujettes à du bruit dans la pratique (erreurs d'observation en physique, erreurs dans les données pas suffisamment "bien" filtrées en finance, et de toutes façons, discrétisation dans le temps des observations). Ce dernier point est capital, il donne un sens à la modélisation de phénomènes par de tels processus.

On a vu que l'on pouvait décrire les irrégularités d'un processus ou d'une fonction grâce à différents exposants. Il n'y a, a priori, pas de raison d'en choisir un plutôt qu'un autre pour une analyse, il faut juste être consistant en utilisant toujours le même pour les propriétés géométriques et statistiques que l'on étudie. On considérera donc ici une suite d'exposants de singularité $s_k^{(n)}$ (où $k = 0, ..., 2^n - 1$ pour couvrir tout l'intervalle [0, 1], et $n \in \mathbb{N}$), qui sont des variables aléatoires si X est un processus. On notera alors, pour t fixé, s(t) pour la limite inférieure des $s_{k_n(t)}^{(n)}$ pour n tendant vers l'infini.

2.2.1 Spectres multifractals

On s'intéresse à la régularité locale des trajectoires d'un processus. Le but va être de quantifier géométriquement et statistiquement quelles valeurs de l'exposant de singularité s apparaissent sur une trajectoire et quelles sont leurs fréquences d'apparition.

On introduit, pour tout $a \in \mathbb{R}$ les ensembles

$$E^{(a)} := \{ t \, | \, s(t) = \liminf_{n \to \infty} s^{(n)}_{k_n(t)} = a \}$$

et

$$K^{(a)} := \{ t \mid \lim_{n \to \infty} s^{(n)}_{k_n(t)} \text{ existe et vaut } a \}.$$

Le support de X peut donc se décomposer en une réunion disjointe de ces ensembles $E^{(a)}$. Typiquement, les processus multifractals auront des sous-ensembles $E^{(a)}$ s'entremêlant de façon complexe dans le support de X, et seront des ensembles fractals, dans le sens qu'ils auront une dimension (fractale) non entière. On parlera de décomposition multifractale du support de X. Il est donc intéressant en soit de considérer la dimension de ces ensembles. Typiquement également, il existera une valeur \bar{a} pour laquelle $E^{(\bar{a})}$ prend toute la mesure de Lebesgue du support de X, avec les autres $E^{(a)}$ de mesure nulle. Le problème est que la plus grande partie de la variation d'une trajectoire du processus (la partie "intéressante" de ce processus) se concentrera en général sur des t caractérisés par un exposant différent de \bar{a} , ou dans le cas de mesures multifractales, la plus grande partie de la masse de la mesure. C'est pour ces raisons que l'on considérera la dimension fractale de ces ensembles, et l'on prendra la dimension de Hausdorff. La fonction $a \mapsto \dim(E^{(a)})$ donnera finalement une représentation compacte de la structure complexe de X. Remarquons que par définition, $K^{(a)}$ est un sous-ensemble de $E^{(a)}$.

Définition 2.2.1 Le spectre de Hausdorff est la fonction

$$\mathbb{R} \to \mathbb{R} : a \mapsto \dim(E^{(a)})$$

où dim désigne la dimension de Hausdorff d'un sous-ensemble de $\mathbb R.$

Le spectre suivant sera lié à la fréquence d'apparition d'une valeur fixée a comme exposant de singularité. On considérera, à n fixé, les exposants $s_k^{(n)}$ $(k = 0, ..., 2^n - 1)$ et on comptera le nombre de k pour lesquels le $s_k^{(n)}$ correspondant est dans un voisinage de la valeur a que l'on s'est donnée. Formellement, on définit pour $a \in \mathbb{R}$ et $\varepsilon > 0$ fixés

$$N^{(n)}(a,\varepsilon) := \# \left\{ k \in \{0, \dots, 2^n - 1\} \, | \, a - \varepsilon \le s_k^{(n)} < a + \varepsilon \right\}.$$

L'idée est que pour une réalisation donnée du processus, $N^{(n)}(a,\varepsilon)/2^n$ peut être vu comme la probabilité pour qu'un naturel k pris au hasard uniformément dans $\{0,\ldots,2^n-1\}$ soit tel que $a - \varepsilon \leq s_k^{(n)} < a + \varepsilon$.

Définition 2.2.2 On définit le spectre f(a) $(a \in \mathbb{R})$ par

$$f(a) := \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log_2(N^{(n)}(a, \varepsilon))$$

et on définit de même

$$\underline{f}(a) := \liminf_{\varepsilon \to 0} \liminf_{n \to \infty} \frac{1}{n} \log_2(N^{(n)}(a,\varepsilon)).$$

On peut déjà constater que si $s(t) \neq a$ pour tout t dans [0, 1], alors $f(a) = -\infty$.

On peut définir un processus multifractal comme un processus pour lequel le spectre f(a) prend tout un ensemble de valeurs distinctes (parmi celles différentes de $-\infty$).

On a introduit ces nouvelles notions à cause de la difficulté que l'on a parfois à calculer les dimensions de Hausdorff. La différence entre les deux approches est que avec la mesure de Hausdorff, on fait un recouvrement de $E^{(a)}$ par des ensembles arbitraires (voir Appendice A.2), alors qu'ici, on considère un recouvrement donné par les $I_{k_n}^{(n)}$. C'est plus facile à calculer en pratique.

2.2.2 Lien entre les spectres

Le résultat suivant est prouvé dans [36]. Il améliore le résultat prouvé par Riedi et Mandelbrot selon lequel dim $(K^{(a)}) \leq f(a)$ (la preuve utilisant le fait que $N^{(n)}(a,\varepsilon)$ peut être utilisé pour estimer des dimensions de boîte, voir Appendice A.1).

Lemme 2.2.3

$$\dim(E^{(a)}) \le f(a),\tag{2.7}$$

$$\dim(K^{(a)}) \le f(a), \tag{2.8}$$

où dim désigne la dimension de Hausdorff d'un sous-ensemble de \mathbb{R} .

Preuve: Fixons $a \in \mathbb{R}$ et prouvons (2.7). Soit $\gamma > f(a)$ arbitraire et $\eta > 0$ tel que $\gamma > f(a) + 2\gamma$. Alors, par définition de f(a), il existe $\varepsilon > 0$ et $m_0 \in \mathbb{N}$ tel que

$$N^{(n)}(a,\varepsilon) \le 2^{n(f(a)+\eta)} \quad \forall n > m_0.$$

Soit

$$J(m) := \bigcup_{n \ge m} \left\{ k \in \{0, \dots, 2^n - 1\} : a - \varepsilon \le s_k^{(n)} \le a + \varepsilon \right\}.$$

Alors, si $t \in E^{(a)}$, pour tout $\varepsilon > 0$, il existe $N \in \mathbb{N}$ tel que $n \ge N$ implique $a - \varepsilon < s_{k_n(t)}^{(n)} \le a + \varepsilon$, et donc que $k_n(t) \in J(m)$ pour *n* suffisamment grand. Comme par définition, $t \in I_{k_n(t)}^{(n)}$ pour tout *n*, on en conclut que

$$\bigcup_{n\geq m} \bigcup_{\substack{k\in J(m)\\k\in\{0,\dots,2^n-1\}}} I_k^{(n)}$$

forme un recouvrement de $E^{(a)}$. Or, la mesure de Lebesgue $|I_{k_n(t)}|$ de $I_{k_n(t)}$ est telle que

$$|I_{k_n(t)}|^{\gamma} = 2^{-n\gamma} \le 2^{-m\gamma} \quad \forall n \ge m.$$

Comme si $k \in J(m)$ avec $k \in \{0, ..., 2^n - 1\}$, $a - \varepsilon \leq s_k^{(n)} \leq a + \varepsilon$ par définition de J(m), on a

$$\#\{k \in J(m) \cap \{0, \dots, 2^n - 1\}\} = \#N^{(n)}(a, \varepsilon)$$

ce qui implique que pour tout $m > m_0$,

$$\sum_{n \ge m} \sum_{\substack{k \in J(m) \\ k \in \{0, \dots, 2^n - 1\}}} |I_k^{(n)}|^{\gamma} = \sum_{n \ge m} 2^{n(\gamma - f(a) - \eta)} \le \sum_{n \ge m} 2^{-n\gamma}.$$

Comme le dernier terme tend vers $0 ext{ si } m \to \infty$, on en déduit que la mesure γ -dimensionnelle de Hausdorff de $E^{(a)}$ est nulle (par définition de la mesure de Hausdorff, et du fait du recouvrement de $E^{(a)}$ par les $I_k^{(n)}$ obtenu plus haut). Donc dim $(E^{(a)}) \leq \gamma$ (par définition de la dimension de Hausdorff), et comme ceci est vrai pour tout $\gamma > f(a)$ arbitraire, (2.7) est prouvé.

Prouvons maintenant (2.8). On considère à nouveau $\gamma > \underline{f}(a)$ et $\eta > 0$ tels que $\gamma > \underline{f}(a) + 2\eta$. Dans ce cas, on a qu'il existe une infinité de naturels n pour lesquels

$$N^{(n)}(a,\varepsilon) \le 2^{n(f(a)+\eta)}$$

Soient les ensembles auxiliaires

$$K_{l} := \left\{ t \in \mathbb{R} : a - \varepsilon < s_{k}^{(n)} < a + \varepsilon \text{ pour } k = k_{n}(t) \text{ et } n > l \right\}$$

définis pour tout $l \in \mathbb{N}$. Par définition, on peut recouvrir K_l par l'union

a

$$\bigcup_{\substack{0 \le k \le 2^n - 1\\ -\varepsilon < s_k^{(n)} < a + \varepsilon}} I_k^{(n)}$$

pour tout n > l fixé. Par définition de $N^{(n)}(a, \varepsilon)$, on a, pour un ensemble infini de $n \in \mathbb{N}$,

$$\sum_{\substack{0 \le k \le 2^n - 1\\ a - \varepsilon < s_k^{(n)} < a + \varepsilon}} |I^{(n)}|^{\gamma} = N^{(n)}(a,\varepsilon) \cdot 2^{-n\gamma} \le 2^{-n(\gamma - \underline{f}(a) - \eta)}.$$

Par définition de la mesure de Hausdorff et par le résultat d'inclusion ci-dessus, on en déduit que la mesure de Hausdorff γ -dimensionnelle de K_l est nulle, et donc que $\dim(K_l) \leq \gamma$. Puisque $K^{(a)} = \bigcup_l K_l$, on en déduit également que $\dim(K^{(a)}) \leq \gamma$ par la σ -continuité de la dimension de Hausdorff (voir appendice). On fait ensuite tendre γ vers $\underline{f}(a)$ pour obtenir le résultat annoncé.

On peut voir que pour une classe de "bons" processus multifractals, $\dim(E^{(a)}) = f(a)$.

2.2.3 Fonction de structure et spectre de Legendre

Par définition, on pourrait voir $\frac{N^{(n)}(a,\varepsilon)}{2^n}$ comme la probabilité, pour une trajectoire fixée du processus, pour qu'un naturel k pris au hasard (uniformément) dans $\{0, \ldots, 2^n - 1\}$ soit tel que $s_k^{(n)} \in [a - \varepsilon, a + \varepsilon]$. Typiquement, il y aura une valeur \hat{a} qui sera la valeur la plus fréquente de $\lim_{n\to\infty} s_k^{(n)}$. En fait, ceci est suggéré par la relation $\frac{N^{(n)}(a,\varepsilon)}{2^n} \sim 2^{f(a)-1}$, et si a n'est pas la valeur la plus probable, f(a) mesurera donc la vitesse de décroissance d'observer a (ou des valeurs proches de a, parmi les s_k^n).

Définition 2.2.4 La fonction de structure d'une trajectoire d'un processus X est définie pour tout $q \in \mathbb{R}$ par

$$\tau(q) := \liminf_{n \to \infty} -\frac{1}{n} \log_2 S^{(n)}(q) \tag{2.9}$$

où

$$S^{n}(q) := \sum_{k=0}^{2^{n}-1} \exp\left(-qns_{k}^{(n)}\ln(2)\right) = \sum_{k=0}^{2^{n}-1} 2^{-nqs_{k}^{(n)}},$$

où l'on a posé par convention ci-dessus que $2^{-qn\infty} = 0$ pour tout $q \in \mathbb{R}$.

Remarquons que si l'on considère la variable aléatoire $A_n := -ns_K^{(n)} \ln(2)$ où K est une variable aléatoire uniforme sur $\{0, \ldots, 2^n - 1\}$, $S^n(q)$ n'est rien d'autre que $2^n \mathbb{E} [\exp(qA_n)]$. $\tau(q)$ apparaît donc comme une "fonction génératrice des moments logarithmique".

Le résultat suivant est prouvé dans [36].

2.2. Analyse multifractale

Théorème 2.2.5 Si dans (2.9), la limite existe dans \mathbb{R} (et pas seulement la limite inférieure) pour tout $q \in \mathbb{R}$, et si $\tau(.)$ est différentiable, alors la limite

$$f(a) = \lim_{\epsilon \downarrow 0} \lim_{n \to \infty} \frac{1}{n} \log_2 N^{(n)}(a, \epsilon)$$

existe, et

$$f(a) = \inf_{q \in \mathbb{R}} (qa - \tau(q)) := \tau^*(a)$$

pour tout $a \in \mathbb{R}$. En particulier, f(a) = f(a).

On appelle $\tau^*(q)$ transformée de Legendre de la fonction $\tau(q)$. La preuve de ce résultat fait appel à la théorie des grandes déviations. Dans la pratique cependant, les hypothèses de ce résultat sont parfois trop restrictives. On a néanmoins le lemme suivant (prouvé dans [36]).

Lemme 2.2.6 Pour tout $a \in \mathbb{R}$, on $a f(a) \leq \tau^*(a)$.

Preuve: Soit $q \in \mathbb{R}$ quelconque et $a \in \mathbb{R}$ tel que $f(a) > -\infty$. Soit $\gamma < f(a)$ et $\varepsilon > 0$ fixé. Alors, par définition de $N^{(n)}(a,\varepsilon)$, il existe $N \in \mathbb{N}$ tel que $n \ge N$ implique $N^{(n)}(a,\varepsilon) \ge 2^{n\gamma}$. Pour de tels n, on a

$$S^{(n)}(q) = \sum_{k=0}^{2^{n}-1} 2^{-nqs_{k}^{(n)}} \ge \sum_{\substack{k=0,\dots,2^{n}-1\\|s_{k}^{(n)}-a|<\varepsilon}} 2^{-nqs_{k}^{(n)}} \ge N^{(n)}(a,\varepsilon)2^{-n(qa+|q|\varepsilon)}$$

puisque le nombre de termes de la seconde somme est égal à $N^{(n)}(a,\varepsilon)$ et que chaque terme de cette somme est plus grand ou égal à $2^{-n(qa+|q|\varepsilon)}$. Puisque $n \ge N$, cette dernière somme est supérieure ou égale à $2^{-n(qa-\gamma+|q|\varepsilon)}$. De ce fait,

$$\tau(q) = \liminf_{n \to \infty} -\frac{1}{n} \log_2 S^{(n)}(q)$$

$$\leq \liminf_{n \to \infty} -\frac{1}{n} \log_2 (2^{-n(qa-\gamma+|q|\varepsilon)})$$

$$= qa - \gamma + |q|\varepsilon.$$

On fait ensuite tendre ε vers 0 et γ vers f(a). On obtient alors $\tau(a) \leq qa - f(a)$ ce qui équivaut à $f(a) \leq qa - \tau(q)$. Ceci est évidemment le cas si $f(a) = -\infty$. On en déduit que cette relation est vraie pour tout $(a, q) \in \mathbb{R}^2$, ce qui prouve le lemme.

Remarquons que si en particulier $\tau(q)$ est linéaire, c-à-d $\tau(q) = cq - 1$ pour une constante c, alors la transformée de Legendre de τ vaut $-\infty$ pour tout $a \neq c$, et vaudra 1 en a = c. Donc la fonction f(a) ne prendra qu'une seule valeur distincte de $-\infty$. C'est pourquoi on caractérise parfois les multifractals comme des processus pour lesquels $\tau(q)$ n'est pas linéaire. La même remarque sera encore valable pour les fonctions de structures déterministes (voir ci-dessous). **Remarque 2.2.7** Lorsque $X = \mathcal{M}$ est le processus associé à une mesure aléatoire μ , si on choisit comme exposant $s_k^{(n)} = \alpha_k^{(n)}$ défini par (2.5), alors, en reprenant les définitions, on voit facilement que

$$S^{(n)}(q) = \sum_{k=0}^{2^n - 1} \left| \mathcal{M}((k+1)2^{-n}) - \mathcal{M}(k2^{-n}) \right|^q = \sum_{k=0}^{2^n - 1} (\mu(I_k^{(n)}))^q.$$

C'est sous cette forme que les fonctions $S^{(n)}(q)$ et $\tau(q)$ ont été introduites historiquement en analyse multifractale (dans le cadre de la turbulence, voir les références citées dans [36]). C'est également essentiellement cette fonction $S^{(n)}(q)$ qui sera considérée plus loin dans la modélisation de taux de change par le modèle en cascade de Breymann *et al.* de même que celle calculée sur les données, à condition de voir le processus des volatilités comme $\mu[t, t + \Delta t]$ pour une certaine mesure aléatoire.

2.2.4 Exposants déterministes

On considère ici les exposants de singularité $s_k^{(n)}(\omega)$ comme des variables aléatoires sur $\Omega \times \{0, 1, \ldots, 2^n - 1\}$ où k est tiré au hasard uniformément dans $\{0, 1, \ldots, 2^n - 1\}$ indépendamment de $\omega \in \Omega$. On supposer toujours ici que $t \in [0, 1]$.

Définition 2.2.8 On définit la fonction de structure déterministe du processus X par

$$T(q) := \liminf_{n \to \infty} -\frac{1}{n} \log_2 \mathbb{E}_{\Omega}[S^{(n)}(q)].$$

Le passage de $\tau(q)$ à T(q) consiste à remplacer des moyennes sur des exposants pour une trajectoire donnée à des moyennes sur des exposants *et* sur les trajectoires. Formellement, pour un ω fixé (pour un état du monde), on calcule un exposant $\tau(q)$ que l'on notera désormais $\tau(q, \omega)$. On remarquera que $\mathbb{E}[\tau(q)] \ge T(q)$. Le résultat suivant (prouvé dans [36]) précise le lien entre $\tau(q, \omega)$ et T(q).

Lemme 2.2.9 $\tau(q,\omega) \geq T(q)$ pour tout $\omega \in A \subset \Omega$ où $\mathbb{P}[A] = 1$ et pour tout q tel que $T(q) < \infty$.

Remarque 2.2.10 On peut introduire ces notions également dans le cas de processus définis sur \mathbb{R}^+ . On définira dans ce cas

$$S^{(n)}(q) := \lim_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N2^n - 1} 2^{-nqs_k^{(n)}},$$

et on fait le même genre d'adaptation pour les autres exposants.

On peut encore introduire un spectre déterministe basé sur f(a) défini comme suit:

Définition 2.2.11

$$F(a) := \lim_{\varepsilon \downarrow 0} \limsup_{n \to \infty} \frac{1}{n} \log_2 \mathbb{E}[N^{(n)}(a,\varepsilon)],$$

$$\underline{F}(a) := \lim_{\varepsilon \downarrow 0} \liminf_{n \to \infty} \frac{1}{n} \log_2 \mathbb{E}[N^{(n)}(a,\varepsilon)].$$

2.2. Analyse multifractale

On peut démontrer le résultat suivant (voir [36]):

Théorème 2.2.12 Si tous les spectres sont calculés sur base du même exposant de singularité, alors

$$\dim(K^{(a)}) \le \underline{f}(a) \le f(a) \le \tau^*(a) \stackrel{p.s.}{\le} T^*(a),$$

où * désigne la transformée de Legendre et où les trois premières relations ont lieu le long d'une trajectoire. De même,

$$\dim(E^{(a)}) \le f(a) \stackrel{p.s.}{\le} F(a) \le T^*(a).$$

On a également le résultat fondamental suivant:

Théorème 2.2.13 Considérons une réalisation de X. Alors

• Si $\{s_k^{(n)} : n \in \mathbb{N}, k = 0, ..., 2^n - 1 \text{ et } s_k^{(n)} < \infty\}$ est borné, alors

$$\tau(q) = f^*(q) \quad \forall q \in \mathbb{R}.$$

Si {s_k⁽ⁿ⁾ : n ∈ N, k = 0,..., 2ⁿ − 1 et s_k⁽ⁿ⁾ < ∞} est non borné supérieurement mais bien inférieurement, alors

$$\tau(q) = \begin{cases} f^*(q) & \forall q > 0\\ -\infty & \forall q < 0 \end{cases}$$

Si {s_k⁽ⁿ⁾ : n ∈ N, k = 0,..., 2ⁿ − 1 et s_k⁽ⁿ⁾ < ∞} est borné supérieurement mais non borné inférieurement, alors

$$\tau(q) = \begin{cases} -\infty & \forall q > 0\\ f^*(q) & \forall q < 0 \end{cases}$$

• Si $\{s_k^{(n)} : n \in \mathbb{N}, k = 0, \dots, 2^n - 1 \text{ et } s_k^{(n)} < \infty\}$ est non borné supérieurement ni inférieurement, alors

$$\tau(q) = -\infty \quad \forall q \neq 0.$$

2.2.5 Lois d'échelle

Dans [27], on définit la notion de processus multifractal en partant des lois d'échelle. Leur définition est la suivante:

Un processus $X(t), t \in T$ est multifractal s'il est à accroissements stationnaires et si

$$\mathbb{E}[|X(t)|^{q}] = c(q)t^{\tau_{q}+1} \tag{2.10}$$

pout tout $t \in T$ et $q \in Q$, où T est intervalle de \mathbb{R}^+ contenant 0 et Q un intervalle de \mathbb{R} contenant [0, 1], et où τ_q et c(q) sont deux fonctions déterministes de $q \in Q$. On peut voir facilement (par l'inégalité d'Hölder) que cette fonction τ_q ainsi définie est concave. En effet, si $q_1, q_2 \in Q$ et $\lambda \in [0, 1]$,

$$\mathbb{E}[|X(t)|^{(1-\lambda)q_1+\lambda q_2}] = c(q_1(1-\lambda)+\lambda q_2)t^{\tau_{(1-\lambda)q_1+\lambda q_2}}, \qquad (2.11)$$

pour tout $t \in T,$ mais par l'inégalité d'Hölder, cette même espérance est inférieure ou égale à

$$\mathbb{E}[|X(t)|^{q_1}]^{(1-\lambda)}\mathbb{E}[|X(t)|^{q_2}]^{\lambda} = c(q_1)^{(1-\lambda)}c(q_2)^{\lambda}t^{(1-\lambda)\tau_{q_1}+\lambda\tau_{q_2}}.$$
(2.12)

On prend ensuite le logarithme des deux membres de (2.11), (2.12), on les divise par $\ln t$ et on passe à la limite pour $t \to 0$ pour obtenir finalement

$$\tau_{(1-\lambda)q_1+\lambda q_2} \le (1-\lambda)\tau_{q_1} + \lambda \tau_{q_2}.$$

Notons qu'une classe particulière de processus vérifiant (2.10) est donnée par les processus auto-similaires. Dans ce cas, on a clairement $\tau_q = Hq - 1$.

Si X(t) est la fonction de répartition associée à une mesure aléatoire (c-à-d $X(t) = \mu([0,t))$, ou $\mu([0,t])$ selon la définition que l'on choisira ...), alors (2.10) peut se réécrire en terme de la mesure μ par

$$\mathbb{E}[|\mu([t,t+\Delta t))|^q] = c(q)(\Delta t)^{\tau_q+1}$$
(2.13)

grâce à la stationnarité des accroissements supposée. En particulier, si l'on choisit l'intervalle $[t, t + \Delta t)$ de la forme $I_k^{(n)}$ pour un certain $n \in \mathbb{N}$ et $k \in \{0, \ldots, 2^n - 1\}$, on aura $\mathbb{E}[|\mu(I_k^{(n)})|^q] = c(q)2^{-n(\tau_q+1)}$. On en déduit qu'alors, si on s'intéresse à l'exposant de singularité $\alpha(t)$, on obtient pour la fonction $S^{(n)}(q)$ correspondante:

$$\mathbb{E}[S^{(n)}(q)] = \sum_{k=0}^{2^n - 1} \mathbb{E}[|\mu(I_k^{(n)})|^q] = \sum_{k=0}^{2^n - 1} c(q) 2^{-n(\tau_q + 1)} = c(q) 2^{-n\tau_q},$$

donc

$$-\frac{1}{n}\log_2 \mathbb{E}[S^{(n)}(q)] = -\frac{1}{n}\log_2(c(q)) + \tau_q,$$

ce qui implique que la limite $\lim_{n\to\infty} -\frac{1}{n}\log_2 \mathbb{E}[S^{(n)}(q)]$ existe et vaut simplement τ_q . Donc la fonction de structure déterministe $T_{\alpha}(q)$ introduite à la section 2.2.4 vaut simplement dans ce cas-ci l'exposant d'échelle τ_q .

Remarque 2.2.14 De façon générale, même si le processus ne satisfait pas une loi d'échelle au sens strict défini plus haut, mais est la fonction de répartition \mathcal{M} d'une mesure aléatoire μ , et est à accroissements stationnaires, on peut réécrire $T_{\alpha}(q)$:

$$T_{\alpha}(q) = -1 + \liminf_{n \to \infty} -\frac{1}{n} \log_2 \mathbb{E}[|\mu(I_0^{(n)})|^q].$$

Cela signifie que même si le processus aléatoire ne satisfait pas de lois d'échelles au sens strict comme dans (2.10), il semblera en satisfaire pour des intervalles de temps très petits (par rapport à l'intervalle de temps sur lequel le processus est défini) avec des exposants τ_q égaux à T(q). Il en satisfera pour des intervalles de temps infinitésimaux. On peut parler de comportements d'échelle locaux. Mais on pouvait déjà parler de lois d'échelle locales (mais avec des puissances seulement égales à q = 1) dès l'introduction des exposants de singularité.

2.2.6 Intégrale stochastique par rapport à un processus multifractal

Au vu des difficultés qui existent déjà pour définir une intégrale stochastique satisfaisante par rapport à un mouvement brownien fractionnaire, il va sans dire que l'on est encore loin d'en définir une par rapport à des processus multifractals.

2.3 Cascades multiplicatives

Dans cette section, nous rappelons d'abord les notions de mesures binomiales et multinomiales, puis illustrons sommairement les concepts de l'analyse multifractale sur une mesure binomiale.

2.3.1 Mesures multinomiales

Un exemple simple: la mesure binomiale déterministe

Commençons par un exemple simple de mesure multifractale: la mesure binomiale déterministe. Cette mesure, de même que les généralisations considérées plus loin ont été introduites par Mandelbrot.

Sur l'espace de mesure ([0, 1], \mathcal{B}), où \mathcal{B} est la σ -algèbre de Borel sur [0, 1], on va considérer une suite de mesures (μ_k) définies suivant le schéma itératif suivant. On commence par se donner deux réels $m_0, m_1 > 0$ avec $m_0 + m_1 = 1$.

A l'étape k = 0, on met une masse unitaire uniforme sur [0, 1]. Ceci nous définit une mesure μ_0 .

A l'étape k = 1, on scinde l'intervalle [0, 1] en deux sous-intervalles de même longueur [0, 1/2] et [1/2, 1], et on met sur [0, 1/2] une masse uniforme m_0 et sur [1/2, 1] une masse uniforme m_1 . Ceci nous définit une mesure μ_1 .

A l'étape k = 2, on divise chaque intervalle de l'étape k = 1 en deux sous-intervalles de même longueur, on répartit uniformément sur le sous-intervalle de gauche une fraction m_0 de la masse de l'intervalle et sur celui de droite une fraction m_1 . En clair, on répartit uniformément sur [0, 1/4] une fraction m_0 de $\mu_1[0, 1/2]$ et sur [1/4, 1/2], une fraction m_1 de cette même masse. On fait de même avec l'intervalle [1/2, 1]. Cela nous définit une mesure μ_2 uniforme par morceaux telle que

$$\mu_2[0, 1/4] = m_0 m_0, \qquad \mu_2[1/4, 1/2] = m_0 m_1, \mu_2[1/2, 3/4] = m_0 m_1, \qquad \mu_2[3/4, 1] = m_1 m_1.$$

On itère ainsi cette procédure. Cela nous donne une suite de mesures μ_k uniformes par morceaux, telle que, si $t \in [0, 1]$ avec

$$t = \sum_{i=1}^{k} \xi_i 2^{-i}, \qquad \xi_i \in \{0, 1\} \ \forall i = 1, \dots, k,$$
(2.14)

 μ_{k+1} est définie à partir de μ_k en mettant une fraction m_0 de la masse $\mu_k([t, t+2^{-k}])$ uniformément sur $[t, t+2^{-k-1}]$ et une fraction m_1 de cette même masse sur $[t+2^{-k-1}, t+2^{-k}]$, et cela pour tout t de la forme (2.14).

La mesure binomiale μ sera par définition la limite de cette suite de mesures au sens suivant. Si t de la forme (2.14), alors

$$\mu_{k+j}([t,t+2^{-k}]) = \mu_k([t,t+2^{-k}]) \quad \forall j \in \mathbb{N}$$
(2.15)

car $m_0 + m_1 = 1$. Donc la limite $\mu([t, t + 2^{-k}]) = \lim_{l\to\infty} \mu_l([t, t + 2^{-k}])$ existe. μ est donc définie pour tout intervalle dyadique, et donc possède trivialement une extension unique à toutes les réunions finies disjointes de tels intervalles. Maintenant, si $t \in [0, 1]$ est quelconque, on peut définir $\mu([0, t)) := \lim \mu([0, k_n(t)2^{-n}])$. Du fait que $\mu([0, 1]) = 1$ et que μ est positive, la suite ci-dessus est croissante bornée supérieurement et donc convergente. De ce fait, on peut étendre de façon unique μ à tout intervalle semi-ouvert du type [a, b), l'ensemble de ces intervalles formant un semi-anneau. On peut voir que μ est positive et σ -additive sur ce semi-anneau. Par le théorème d'extension de Lebesgue, il existe donc une unique extension de μ aux boréliens de [0, 1].

Du fait que $\mu([0, k2^{-n}]) = \mu([0, k2^{-n}])$ pour tout k, n, on peut voir que $\mu([0, t]) = \mu([0, t))$ pour tout t, ce qui implique la continuité de la fonction \mathcal{M} associée. On peut voir cependant que \mathcal{M} n'est dérivable nulle part. On parle dans ce cas de mesure singulière.

Remarquons que nous avons finalement

$$\mu([t, t+2^{-k}]) = m_0^{k-n_1} m_1^{n_1}$$

avec

$$n_1 = \sum_{i=1}^k \xi_i.$$

On parle également de mesure conservative car la masse de l'intervalle $[t, t + 2^{-k}]$ est conservée aux étapes suivantes (voir (2.15)) puisque $m_1 + m_0 = 1$.

La mesure μ est naturellement déterministe.

Les figures 2.1, 2.2 illustrent la construction de μ . A la limite, la densité de mesure n'est continue nulle part, donnant lieu à un \mathcal{M} dérivable nulle part.

Mesure multinomiale déterministe

On peut étendre la construction précédente en considérant un naturel b > 2 et des nombres $m_0, \ldots, m_{b-1} > 0$ tels que $\sum_{\beta=0}^{b-1} m_{\beta} = 1$, et en suivant le même genre de procédure itérative. A la (k + 1)è étape, on divise chaque intervalle b-adique de la forme $[t, t + b^{-k}]$ où $t = \sum_{i=1}^{k} \xi_i b^{-i}$ en b sous-intervalles de longueur égale et on répartit uniformément sur chaque sous-intervalle la masse $\mu_k([t, t + b^{-k}])$ suivant des fractions m_0, \ldots, m_{b-1} .



Figure 2.1: Mesure binomiale: Trois premières itérations μ_1, μ_2, μ_3 , avec $m_0 = 1/3, m_1 = 2/3$



Figure 2.2: Mesure binomiale: μ_7, μ_{10} , avec $m_0 = 1/3, m_1 = 2/3$

Mesures binomiales aléatoires

On peut rendre aléatoire une mesure binomiale en remplaçant dans le schéma itératif m_0, m_1 par deux variables aléatoires positives à chaque étape. L'idée est la même qu'auparavant: si $t \in [0, 1]$, on a déjà vu qu'il existait une unique suite d'entiers $(k_n) = (k_n(t))$ tels que les intervalles correspondants $I_{k_n}^{(n)}$ contiennent t pour tout n. Ces intervalles sont de longueur $2^{-n} \to 0$. Pour une telle suite de naturels (k_n) , la mesure binomiale μ sera définie par

$$\mu(I_{k_n}^{(n)}) = M_{k_n}^{(n)} M_{k_{n-1}}^{(n-1)} \dots M_{k_1}^{(1)} M_0^{(0)}$$
(2.16)

où l'on supposera cette fois:

- (H1) $M_0^{(0)}$ est une variable aléatoire positive (représentant la masse totale de [0, 1]),
- (H2) les différents facteurs $M_{k_n}^{(n)}, M_{k_{n-1}}^{(n-1)}, \ldots, M_{k_1}^{(1)}, M_0^{(0)}$ sont indépendants et positifs,
- (H3) $M_{2k_n}^{(n+1)} + M_{2k_n+1}^{(n+1)} = 1$ p.s. pour tout *n* (conservation de la masse, mesure conservative)

$$M_k^{(n)} \stackrel{\mathrm{d}}{=} \begin{cases} M_0 & \text{si } k \text{ est pair} \\ M_1 & \text{sinon} \end{cases}$$

où M_0, M_1 sont deux variables aléatoires positives.

Donc essentiellement, on a remplacé m_0 et m_1 par deux variables aléatoires positives, avec indépendance entre les étapes de la procédure itérative.

A nouveau, la mesure μ est bien définie puisque définie sur chaque intervalle dyadique. La fonction de répartition correspondante \mathcal{M} sera définie par $\mu[0, t]$. De façon générale, \mathcal{M} sera continue à gauche comme toute fonction de répartition, mais sera continue en tout point $t \in [0, 1]$ sauf si $M_{k_n(t)}^{(n)} = 1$ pour tout *n* suffisamment grand. On appelle \mathcal{M} une cascade multiplicative (ou cascade binomiale).

On peut également ne plus supposer la conservation de la masse à chaque étape, et ne la supposer plus qu'en moyenne: imposer juste que $\mathbb{E}[M_0 + M_1] = 1$ à la place de (H3). Dans ce cas,

$$\mu(I_{k_n}^{(n)}) = M_{k_n}^{(n)} M_{k_{n-1}}^{(n-1)} \dots M_{k_1}^{(1)} M_0^{(0)} \lim_{m \to \infty} \sum_{(i_1, \dots, i_m)} M_{2k_n + i_1}^{(n+1)} M_{4k_n + i_2}^{(n+2)} \dots M_{2^m k_m + i_m}^{(n+m)}$$
(2.17)

où la somme a lieu sur toutes les m-uples (i_1, \ldots, i_m) tels que

$$I_{2^{l}k_{n}+i_{l}}^{(n+l)} \subset I_{2^{l-1}k_{n}+i_{l-1}}^{(n+l-1)} \quad \forall l = 1, \dots, m$$

Remarquons que dans le cas où l'on suppose (H3), cette dernière somme vaut simplement 1. Mais si l'on n'a pas (H3) et que l'on désire faire des simulations numériques de cette mesure, le produit apparaissant dans (2.16) ne donnera plus $\mu(I_{k_n}^{(n)})$.

On pourrait également considérer des constructions de mesures multinomiales aléatoires en considérant par exemple que les réels positifs m_{β} sont remplacés par des variables aléatoires M_{β} prenant les valeurs $m_0, m_1, \ldots, m_{b-1}$ avec des probabilités respectives p_0, \ldots, p_{b-1} , avec indépendance entre les étapes et $\sum_{\beta=0}^{b-1} M_{\beta} = 1$ (mesure conservative), ou seulement en moyenne. Dans ce cas-ci donc, les variables M_{β} sont identiquement distribuées. On peut supposer également que M_{β} sont de distribution non discrète, ou encore que les M_{β} ne sont pas identiquement distribuées mais satisfont une relation du type (H4).

2.3.2 Analyse multifractale de la mesure binomiale déterministe

Voici, en guise d'illustration des concepts définis dans la Section 2.2, une analyse multifractale sommaire de la mesure binomiale déterministe introduite plus haut. Commençons par voir à quoi ressemblent les exposants $\alpha(t)$, pour $t \in [0, 1]$. Si t est quelconque dans [0, 1], on peut considérer son développement en binaire:

$$t = \sum_{k=1}^{\infty} \xi_k(t) 2^{-k}, \quad \xi_k(t) \in \{0, 1\} \quad \forall k \in \mathbb{N}.$$

46

2.3. Cascades multiplicatives

Ce développement sera défini par induction par:

$$\xi_1 = \lfloor t2 \rfloor, \quad \xi_k = \left[(t - \sum_{l=1}^{k-1} \xi_l 2^{-l}) 2^k \right] \quad k = 2, 3, \dots$$

Dans ce cas, à $n \in \mathbb{N}$ fixé,

$$k_n(t) = \lfloor t2^n \rfloor = \left\lfloor \sum_{k=1}^n \xi_k(t)2^{n-k} + \sum_{k=n+1}^\infty \xi_k(t)2^{n-k} \right\rfloor = \sum_{k=1}^n \xi_k(t)2^{n-k},$$

du moins à condition que l'on n'ai pas $\xi_k = 1$ pour tout k > n (un tel cas est en fait toujours évité à condition de définir le développement en binaire comme plus haut, sauf si t = 1, auquel cas on ne définira pas $I_{k_n(t)}^{(n)}$) et donc les extrémités de $I_{k_n(t)}^{(n)}$ sont

$$k_n(t)2^{-n} = \sum_{k=1}^n \xi_k(t)2^{n-k}, \quad (k_n(t)+1)2^{-n} = \sum_{k=1}^n \xi_k(t)2^{n-k} + 2^{-n},$$

et donc évidemment, $\mu(I_{k_n}^{(n)}) = m_0^{n-n_1} m_1^{n_1}$ où $n_1 = n_1(n,t) = \sum_{k=1}^n \xi_k(t)$. Par définition,

$$\alpha_{k_n(t)}^{(n)} = -\frac{1}{n} \log_2(\mu(I_{k_n}^{(n)})) = -\frac{1}{n} ((n-n_1) \log_2(m_0) + n_1 \log_2(m_1)),$$

et donc

$$\alpha(t) = -\log_2(m_0) + \liminf_{n \to \infty} \frac{1}{n} \sum_{k=1}^n \xi_k(t) (\log_2(m_0) - \log_2(m_1))$$

On voit que $\alpha(t)$ va varier entre les valeurs extrêmes $\alpha_0 := -\log_2(m_0)$ et $\alpha_1 := -\log_2(m_1)$. En fait, on a toujours que $n_1(n,t) \leq n$ pour tout n,t. Dès que $n_1(n,t)$ est $O(n^\beta)$ pour un exposant $\beta < 1$, la limite inférieure dans $\alpha(t)$ sera nulle et $\alpha(t) = \alpha_0$. Si par exemple t a un développement fini en binaire, c'est à dire $t = \sum_{k=1}^{N} \xi_k(t) 2^{-k}$ pour un certain naturel N, alors clairement $\alpha(t) = \alpha_0$. Donc il existe un ensemble dense dans [0, 1] tel que $\alpha(t)$ soit égal à α_0 sur cet ensemble.

Si par contre t = 1/3, on voit que $\xi_k(t) = 1$ si et seulement si k est pair, et donc $n_1(n,t) = n/2$ si n est pair et (n-1)/2 si n est impair, ce qui donne finalement $\alpha(1/3) = -1/2 \log_2(m_0) - 1/2 \log_2(m_1)$.

En fait, on peut voir que $\alpha(t)$ peut prendre comme valeur n'importe quel réel du type

$$\left(1-\frac{p}{q}\right)\alpha_0 + \frac{p}{q}\alpha_1$$

où $p, q \in \mathbb{N}_0$, p < q. Il suffit de prendre pour cela t tel que $\xi_k(t) = 1$ si et seulement si $k \mod q \leq p$. En clair, dans le développement en binaire de t, on aura des cycles de longueur q avec des 1 pour les p premières positions d'un cycle, et des 0 pour les q - ppositions suivantes du cycle. Dans ce cas, on peut voir que

$$n_1(n,t) = \begin{cases} (n - n \mod q)\frac{p}{q} + n \mod q & \text{si } n \mod q \le p \\ (n - n \mod q + q)\frac{p}{q} & \text{si } n \mod q > p \end{cases}$$

et on en conclut donc le résultat pour $\alpha(t)$. Ici, on peut voir que $t = \frac{1-2^{-p}}{1-2^{-q}}$.

Calculons maintenant $S^{(n)}(q)$ et $\tau(q)$ correspondant à l'exposant de singularité α . Par définition,

$$S^{(n)}(q) = \sum_{k=0}^{2^n - 1} 2^{-nq\alpha_k^{(n)}} = \sum_{n_1=0}^n m_0^{q(n-n_1)} m_1^{qn_1} \begin{pmatrix} n \\ n_1 \end{pmatrix} = (m_0^q + m_1^q)^n$$

puisqu'il y a exactement $\binom{n}{n_1}$ naturels k dans $\{0, \dots, 2^n - 1\}$ tels que $n_1(n, k2^{-n}) = n_1$. On en déduit donc que

$$\log_2(S^{(n)}(q)) = n \log_2(m_0^q + m_1^q),$$

et finalement

$$\tau(q) = -\log_2(m_0^q + m_1^q),$$

ce qui est bien une fonction différentiable de q. Par le Théorème 2.2.5, $f(\alpha)$ satisfait

$$f(\alpha) = \tau^*(\alpha) = \inf_{q \in \mathbb{R}} (q\alpha - \tau(q))$$

pour tout $\alpha \in \mathbb{R}$, cet infimum valant éventuellement parfois $-\infty$. On peut voir que si α se trouve strictement entre α_0 et α_1 , la valeur de q minimisant l'expression ci-dessus est donnée par

$$q(\alpha) = \frac{\ln\left[\ln(m_1/a)/\ln(a/m_0)\right]}{\ln(m_0/m_1)}, \quad a := 2^{-\alpha}.$$

On en déduit alors l'expression de $f(\alpha)$. Si par contre $\alpha = \alpha_0$ (ou α_1), $q\alpha - \tau(q) = \log_2(1 + (m_1/m_0)^q)$ (ou $\log_2(1 + (m_0/m_1)^q)$) ce qui prend comme valeur minimum la valeur 0. On en déduit que $f(\alpha_0) = f(\alpha_1) = 0$. Si α n'est pas entre α_0 et α_1 , on peut voir que la fonction $q\alpha - \tau(q)$ n'est pas bornée inférieurement, donc l'infimum vaut $-\infty$. En fait, cela découle simplement de la définition de $f(\alpha)$, sans devoir passer par $\tau(q)$.

On peut déduire de tout ceci que l'exposant $\alpha(t)$ peut prendre n'importe quelle valeur entre α_0 et α_1 . Vu la continuité de $q(\alpha)$ entre α_0 et α_1 et celle de $\tau(q)$, on en déduit celle de $f(\alpha)$ entre α_0 et α_1 .

Voir la figure 2.3 pour les graphes de $f(\alpha)$ et $\tau(q)$ dans le cas où l'on a choisi $m_0 = 0.3$ et $m_1 = 0.7$. On voit que $f(\alpha)$ a une valeur maximum égale à 1.

Tout ceci montre que la mesure binomiale déterministe est bien un cas particulier de multifractal.

Lois d'échelle pour la mesure binomiale aléatoire

Supposons que μ soit égale à la mesure binomiale aléatoire introduite précédemment (avec les hypothèses (H1) jusque (H4)). Dans ce qui suit on supposera que $\mathbb{E}[|M_0|^q]$ et $\mathbb{E}[|M_1|^q]$ existent pour tout $q \in \mathbb{R}$.



Figure 2.3: Mesure binomiale: spectres $f(\alpha)$ et $\tau(q)$ si $m_0 = 0.3$ et $m_1 = 0.7$.

Si on s'intéresse à $\mathbb{E}[|\mu(I_k^{(n)})|^q]$ où $k \in \{0, \dots, 2^n - 1\}$, et que l'on définit $\xi_i \in \{0, 1\}$ par

$$k \, 2^{-n} = \sum_{i=1}^{n} \xi_i 2^{-i},$$

alors par définition de la mesure binomiale aléatoire et l'hypothèse d'indépendance (H2),

$$\mathbb{E}[|\mu(I_k^{(n)})|^q] = \mathbb{E}[|M_0^{(0)}|^q]\mathbb{E}[|M_0|^q]^{n-n_1}\mathbb{E}[|M_1|^q]^{n_1}$$

où l'on a posé

$$n_1 = \sum_{i=1}^n \xi_i.$$

Du fait que le logarithme de $\mathbb{E}[|M_0|^q]^{n-n_1}\mathbb{E}[|M_1|^q]^{n_1}$ va dépendre de n_1 , et donc de l'intervalle $I_k^{(n)}$ considéré, on n'aura pas de loi d'échelle (au sens strict) qui sera vérifiée pour une telle mesure. Si par contre on remplace l'hypothèse (H4) par (H4'):

(H4')

$$M_k^{(n)} \stackrel{\mathrm{d}}{=} M \quad \forall n \in \mathbb{N}_0, \forall k \in \{0, \dots, 2^n - 1\},$$

pour une certaine variable aléatoire M positive

(on impose juste que $M_0 \stackrel{d}{=} M_1$) avec E[M] = 1/2 à la place de (H3), alors

$$\mu(I_k^{(n)}) = M_{k_n}^{(n)} M_{k_{n-1}}^{(n-1)} \dots M_{k_1}^{(1)} M_0^{(0)} N_{k_1,\dots,k_n}$$

où N_{k_1,\dots,k_n} est la limite dans (2.17). On peut voir que ces variables N_{k_1,\dots,k_n} sont i.i.d. $\stackrel{\text{d}}{=} N$. On a alors

$$\mathbb{E}[|\mu(I_k^{(n)})|^q] = c(q)(2^{-n})^{\tau_q+1}$$

où l'on a posé $\tau_q := -\log_2(\mathbb{E}[|M|^q]) - 1$ et $c(q) = \mathbb{E}[|M_0^{(0)}N|^q]$, ce qui signifie que l'on a effectivement bien une loi d'échelle du type (2.13) pour tout $(t, \Delta t)$ de la forme

$$\begin{cases} t = \sum_{i=1}^{n} \xi_i 2^{-i}, & \xi_i \in \{0, 1\} \forall i \\ \Delta t = 2^{-n} \end{cases}$$

Replaçons-nous à nouveau sous (H3-H4) et calculons $T_{\alpha}(q)$, la fonction de structure déterministe correspondant à l'exposant $\alpha(t)$. Nous obtenons

$$\mathbb{E}[S^{(n)}(q)] = \sum_{k=0}^{2^n - 1} \mathbb{E}[|\mu(I_k^{(n)})|^q] = \mathbb{E}[|M_0^{(0)}|^q] \sum_{n_1=0}^n \mathbb{E}[|M_0|^q]^{n-n_1} \mathbb{E}[|M_1|^q]^{n_1} \begin{pmatrix} n\\ n_1 \end{pmatrix}$$
$$= \mathbb{E}[|M_0^{(0)}|^q] \left(\mathbb{E}[|M_0|^q] + \mathbb{E}[|M_1|^q]\right)^n,$$

et on en déduit que

$$T(q) = -\log_2 \left(\mathbb{E}[|M_0|^q] + \mathbb{E}[|M_1|^q] \right).$$

En particulier, dans le cas où l'on suppose (H4') à la place de (H4), on retrouve que $T(q) = \tau_q$, l'exposant d'échelle.

En fait, dans [36] est prouvé le fait que dès que \mathcal{M} est un processus croissant presque sûrement, alors pour tout q,

$$T_{\alpha}(q) = T_h(q)$$

et pour presque toute trajectoire,

$$\tau_{\alpha}(q,\omega) = \tau_h(q,\omega).$$

On peut voir également que dans le cas d'une mesure binomiale, $f(\alpha)$ est égal à la dimension de Hausdorff de $K^{(\alpha)}$ et à $T^*(\alpha)$ pour tout α tel que $T^*(\alpha) > 0$.

On en déduit que si l'on définit des exposants $\alpha_k^{(n)}$ en utilisant une autre base que 2 pour les logarithmes ainsi que pour les subdivisions successives de l'intervalle [0, 1] des sous-intervalles, on va retomber sur les mêmes fonctions de structure $\tau(q)$ et T(q). Ceci est intéressant si l'on désire étudier des mesures multinomiales.

Partie II

Modélisation pour le marché des changes par un modèle multifractal

Chapitre 3

Motivation de ce type de modèle

3.1 Le marché des changes et les données à haute fréquence

Il s'agit du plus grand marché financier au point de vue du montant des transactions. Pour avril 92, la Bank for International Settlements (BIS) avait estimé le montant total journalier des transactions (pour le marché au comptant et à terme) à USD 832 milliards, c'est-à-dire plus que les réserves non-or de tous les pays industrialisés en 92. Pour avril 95, l'estimation était de USD 1190 milliards et pour avril 98 de USD 1500 milliards.

Le marché des changes produit des données à haute fréquence (i.e. intra-journalières) ayant joué un rôle central en finance à haute fréquence. Contrairement aux autres marchés, ces données peuvent être disponibles sur de longues périodes et à de hautes fréquences, et 24 heures par jour ouvrable. Comme pour la plupart des marchés financiers, ce marché est informatisé et les prix cotés arrivent (pour certaines devises) à des intervalles de temps de seulement parfois quelques secondes. Ce marché est également très liquide.

Depuis le début des années 1990, on s'intéresse aux données intra-journalières, ce qui a permis de mettre en évidence des nouveaux comportements qui n'apparaissaient pas dans l'analyse des données journalières. Par exemple, l'homogénéité des agents disparaît selon certains. Le marché des devises est caractérisé par une répartition géographique des différents marchés et différents types d'agents, avec des profils de risques différents et des contraintes institutionnelles différentes. Avant l'étude des données à haute fréquence, ces caractéristiques structurelles n'étaient pas apparentes dans les données, et peu considérées en modélisation.

On peut diviser le marché en essentiellement deux parties: le marché au comptant (spot market) et le marché à terme (forward market)¹. On ne parlera que du marché au comptant.

¹Il y a encore le marché des produits dérivés sur devises, mais c'est une plus petite partie du marché (en développement cependant)

Actuellement, la plupart des transactions du marché au comptant se font via des systèmes de cotation automatisés (par exemple, l'Electronic Broking Services ou Reuters Dealing ...). Cela donne lieu à une source de données à très haute fréquence, contenant les prix et volumes des transactions. Ces données ne sont cependant pas toujours facilement accessibles aux chercheurs.

A coté des transactions via des systèmes de cotations électroniques centralisés, une grande part des transactions se fait sans intermédiaire, directement entre banques. On appelle cette partie du marché le marché "over-the-counter" (OTC). En pratique, les cours acheteurs et vendeurs des grandes institutions financières sont communiqués (électroniquement) à leurs clients potentiels par des fournisseurs de données comme Reuters, Bloomberg ou Bridge avec un délai aussi court que possible. Les transactions se négocient par téléphone, les prix de transactions ne sont donc pas disponibles. Les prix du marché OTC sont qualifiés de prix cotés, par opposition aux prix de transaction. Typiquement, pour ce marché, un tick complet contient l'instant, le cours acheteur et vendeur (bid and ask price), et en général l'origine du tick. On ne dispose ni des volumes ni des prix de transaction. Le prix réel est en général contenu dans l'intervalle cours acheteur–cours vendeur (bid-ask spread). Ce marché est une source très importante de données pour les chercheurs.

Le marché des changes n'a pas de limitation horaire: n'importe quel intermédiaire peut émettre un nouveau prix, et les grandes institutions financières ont des branches partout dans le monde. Cependant, les prix émis ont quand-même tendance à se concentrer géographiquement, et on peut distinguer trois grandes zones d'activité, correspondant à l'Europe (au sens large), l'Est asiatique et l'Amérique, avec comme grands centres respectivement Londres, Tokyo et New-York.

Olsen & Associates (actuellement Olsen Group) posséde une très importante banque de données pour le marché au comptant. La fréquence de ces ticks a augmenté assez fort ces dix dernières années. Par exemple, sur le plus grand marché (EUR-USD), on comptait plus de 13 000 ticks par jour en moyenne de janvier 1999 à mai 2000. Pour le cours USD-CHF, environ 2400 ticks par jour en moyenne.

La collecte des données intra-journalières présente toute une série de problèmes pratiques comme les délais de transmission, les pannes, les données fausses dues à des erreur humaines ou de machine ... Il y a donc tout un travail préliminaire important de filtrage des données à réaliser avant toute analyse. Dans le marché au comptant, les filtres seront utilisés pour lutter contre les phénomènes suivants:

- Certains traders publient parfois des prix cotés simplement copiés à partir de (ou obtenus à partir de moyennes sur des) prix publiés par d'autres traders, pour signaler simplement la présence dans le marché de leur institution. On doit en tenir compte lors du filtrage des données.
- Des erreurs décimales peuvent bien-sûr apparaître.
- Certaines institutions abusent parfois du canal de transmission simplement pour

3.2. Les données

faire certains tests (d'affichage par exemple), et on essaiera de se débarrasser de ces ticks par filtrage.

D'autres phénomènes sont également observés sur le marché des changes (mais dont le filtrage ne tiendra pas compte):

- les agents ont parfois une réelle préférence pour vendre ou acheter. Ils publient alors de nouveaux prix pour n'attirer que les autres traders désirant réaliser une transaction dans l'une ou l'autre direction. Pour cela, le cours acheteur (resp. vendeur) sera compétitif alors que le cours vendeur (resp. acheteur) ne le sera pas du tout. On aboutit alors à un bid-ask spread non réaliste.
- Il y a des problèmes de délais pour les prix cotés chez certains traders, par rapport au marché réel.
- Certains traders (de mauvaise réputation) tentent également de manipuler le marché dans une certaine direction

De façon générale, le traitement des données à haute fréquence nécessite la mise en oeuvre de toute une série de méthodes (désaisonnalisation, filtrage, interpolation,...). Le récent livre [10] donne une bonne idée des méthodes utilisées.

3.2 Les données

Les données dont nous avons disposé proviennent de la base de données de Olsen & Associates², et ont été préparées par W. Breymann (ayant travaillé chez Olsen plusieurs années). Elles contiennent le cours du franc suisse par rapport au dollar américain pour chaque heure vingt-quatre (en temps de marché, voir plus bas) du 02/01/1991 au 30/05/2001. Plus précisément, le cours s'exprimera comme le nombre de francs suisses correspondant à 1 USD.

Pour les taux de change, les prix cotés se présentent toujours sous la forme de paires cours acheteur - cours vendeur (bid-ask pairs). Dans notre cas, les données contiennent des prix milieux logarithmiques (middle logarithmic prices), c'est-à-dire se présentent sous forme d'une suite (x_j) correspondant à

$$x_j = x(t_j) := \frac{\log p_{bid}(t_j) + \log p_{ask}(t_j)}{2} = \log \sqrt{p_{bid}(t_j)p_{ask}(t_j)},$$

où $p_{ask}(t_j)$ est le cours vendeur à l'instant t_j et $p_{bid}(t_j)$ le cours acheteur.

On définit le return logarithmique à l'instant t pour un certain intervalle Δt comme

$$r(t, \Delta t) = x(t) - x(t - \Delta t).$$

²Actuellement Olsen Group, Seefeldstrasse 233, Zürich, Suisse

L'avantage de considérer ces prix moyens logarithmiques est leur antisymétrie: si x est le cours USD-CHF, alors -x est le cours CHF-USD. Donc des résultats statistiques basés sur les prix absolus seront les mêmes pour les cours USD-CHF et CHF-USD. Un autre avantage de considérer des prix logarithmiques est que les returns correspondants sont sans dimension (indépendants de l'unité monétaire utilisée, important lorsque l'on prend des logarithmes d'une certaine quantité).

On préfère analyser le processus des returns plutôt que directement celui des prix pour plusieurs raisons. C'est d'abord la variable d'intérêt des traders, pour qui c'est une mesure directe du succès de leur investissements. De plus, la distribution des returns sera plus symétrique et plus stationnaire que la distribution des prix.

La longueur de notre série est de 65171. La figure 3.1 illustre la série temporelle des données (prix logarithmiques) et la figure 3.2 les returns pour $\Delta t = 1$ heure 24 en "temps de marché" (ce qui correspond à une moyenne d'environ 1 heure en temps physique pendant les jours ouvrables, voir plus loin).



Figure 3.1: Données: cours USD-CHF du 02/01/1991 au 30/05/2001.

Lorsque des professionnels (traders) ont accès à des données financières à haute fréquence, ils en connaissent généralement très bien le contexte (l'état du marché à tel moment, le niveau de vraisemblance des cours à un certain instant,...). Ils sont donc capables de détecter les erreurs dans les données et les nettoient implicitement avant de s'en servir. A l'inverse, des chercheurs ne sont pas capables de discerner aussi facilement de telles erreurs. En données à haute fréquence, le filtrage est donc une nécessité.



Figure 3.2: Returns logarithmiques pour $\Delta t = 1$ heure.

Un filtre de données se présentant sous forme d'un algorithme devient nécessaire lorsque l'on se trouve face à des milliers de données à traiter (comme c'est le cas en finance à haute fréquence). Une alternative à ces filtres est de faire de l'inférence robuste, mais ces méthodes ne sont pas universellement applicables.

Les erreurs peuvent être de plusieurs type: erreurs humaines tout d'abord (prix mal introduit, mal tapé, ou prix intentionnellement faux produit juste pour raisons techniques, comme par exemple un agent qui envoie un mauvais prix tôt le matin pour voir si la connection au distributeur de données est opérationnelle), ou erreurs machine. Le filtre devra tenir compte de la nature possible de ces erreurs, et certaines ne seront pas faciles à détecter.

Le type de filtre qui a été utilisé ici est très bien décrit dans [10] au chapitre 4. Ce sujet est assez technique dans son ensemble. L'idée de base sera de mesurer au fur et à mesure la crédibilité (à quantifier) des nouveaux ticks en tenant compte des ticks précédents (et de leur crédibilité) se trouvant dans une certaine fenêtre de temps.

3.3 Désaisonnalisation des données

Un fait important dans les données à haute fréquence: on a remarqué (cf. [10]) que les données financières à haute fréquence espacées régulièrement en temps physique présentent une structure saisonnière dans le processus des volatilités réalisées (mesure de l'agitation du marché autour d'un certain instant, voir section 3.4.2 pour les définitions): les returns absolus moyens par exemple seront moins élevés durant le week-end, les vacances ou encore l'heure de midi,... donc pendant les périodes de moindre activité. La plus faible période d'activité a lieu durant le temps de midi au Japon (nuit en Amérique et en Europe, et pause de midi au Japon). C'est à ce moment également que l'on trouve la plus faible valeur des returns moyens absolus. La fonction d'autocorrélation des returns absolus présente également le même genre de saisonalité: on peut observer des pics pour des délais multiples entiers de 24h. Cette saisonalité est un effet assez important dû au fait que les informations politiques et économiques de même que l'activité du marché tournent autour de la terre avec un cycle de 24 heures, et cet effet est susceptible d'en voiler d'autres éventuellement observables, mais plus subtils. Dans l'idée d'étudier empiriquement (et plus finement) des données financières à forte saisonalité comme c'est le cas pour des données issues du marché des changes, on va traiter au préalable ces données en vue d'éliminer cet effet de saisonalité. Comme cette saisonalité est directement liée à l'activité du marché, une possibilité sera d'introduire une nouvelle échelle de temps faisant intervenir cette activité (à formaliser), ce qui devrait faire disparaître une grande partie de la saisonalité. Les données seront ensuite espacées régulièrement dans cette nouvelle échelle de temps. On appellera ce nouveau temps "temps de marché".

Voici une description de la nouvelle échelle de temps utilisée dans le cas de nos données. Elle a été introduite par l'équipe de recherche d'Olsen³, et est décrite dans [10] (d'où provient directement ce qui suit).

Le temps de marché $\theta(t)$ sera un changement d'échelle de temps tel que dans celle-ci, les données ne présenteront (presque) plus de saisonalité intra-journalière. Ce sera une fonction $\theta : \mathbb{R} \to \mathbb{R}$ strictement croissante, et le processus des prix dans cette nouvelle échelle sera un processus $\{x^*(\theta), \theta \in \mathbb{R}\}$ défini par $x^*(\theta(t)) = x(t)$ pour tout $t \in \mathbb{R}$. L'idée est donc de se débarrasser de la saisonalité due uniquement à l'activité du marché et le principe du temps $\theta(t)$ sera d'être proportionnel à l'activité du marché. Cette activité sera modélisée sur des données réelles à partir des lois d'échelles.

Le modèle est le suivant. On modélise l'activité du marché au cours du temps par une certaine fonction a(t). Comme le marché des devises se subdivise en trois grands marchés (Europe, Amérique, Asie), on décompose l'activité totale a(t) en les activités de chaque zone:

$$a(t) := \sum_{k=1}^{3} a_k(t).$$

Chacun des trois marchés sera considéré dans le modèle comme soit ouvert, soit fermé.

³le groupe de recherche d'Olsen la qualifie de " θ -time"

3.3. Désaisonnalisation des données

Cependant, on supposera que l'activité pendant les heures de fermeture reste à un niveau constant très faible $a_{0,k}$ (car l'activité sera modélisée sur base des returns). On définit alors $a_{1,k}$ par

$$a_{1,k}(t) := a_k(t) - a_{0,k}$$

de sorte que

$$a(t) = \sum_{k=1}^{3} (a_{0,k} + a_{1,k}(t)) = a_0 + \sum_{k=1}^{3} a_{1,k}(t)$$

où l'on a posé $a_0 := \sum_{k=1}^3 a_{0,k}$ (a_0 sera l'un des paramètres du modèle). L'activité $a_{1,k}(t)$ sera ensuite modélisée par un polynôme (facilement intégrable et dérivable analytiquement).

Le temps t sera mesuré en heures et variera de 0 (le lundi à 0h00 GMT) à t = 168 (le dimanche soir à 24h00 GMT).

Avant de donner explicitement le polynôme de $a_{1,k}(t)$, introduisons le temps auxiliaire $T_k(t)$ (k = 1, 2, 3) par:

$$T_k(t) := (t + \Delta T_k) \mod 24 - \Delta T_k$$

où

$$\Delta T_k := \begin{cases} 9 & \text{pour l'Asie} \\ 0 & \text{pour l'Europe} \\ -5 & \text{pour l'Amérique} \end{cases}$$

Essentiellement, T_k ramène le temps t sur 24 heures dans chaque marché k, en tenant compte du changement de date déjà survenu ou non sur le marché en question. Exemple: si on s'intéresse au marché européen, si t = 26 (mardi 2h00 à Londres), alors $T_k = 2$. Par contre, sur le marché américain, cela donne $T_k = 26$, car on n'a pas encore changé de date en Amérique. Si t = 20 (lundi 20h00 à Londres mais déjà mardi en Asie), alors pour le marché asiatique, $T_k = -4$ car on a déjà changé de date.

La condition d'être déjà arrivé au week-end (week-end condition) s'exprime pour chaque marché k par:

$$(t + \Delta T_k) \mod 168 \ge 120. \tag{WEC}$$

On écrit alors

$$a_{1,k}(t) = \begin{cases} 0 & \text{si } (T_k < o_k) \text{ ou } (T_k > c_k) \text{ ou } (\text{WEC}) \\ a_{open,k}(t) & \text{si } (o_k < T_k < c_k) \text{ et non(WEC)} \end{cases}$$

où o_k et c_k sont les heures d'ouverture et de fermeture de chaque marché. On modélise ensuite $a_{open,k}(t)$ par un polynôme comme annoncé plus haut:

$$a_{open,k}(t) = \frac{\omega_k}{\frac{o_k + c_k}{2} - s_k} (T_k - o_k)^2 (T_k - c_k)^2 (T_k - s_k) [(T_k - m_k)^2 + d_k^2]$$

où m_k est l'instant du minimum d'activité de la pause de midi pour le marché en question, d_k est lié à l'intensité d'activité lors de ce minimum, ω_k est la pondération du kè marché au niveau activité dans le marché total, et s_k est lié à l'asymétrie de la courbe d'activité. Les facteurs $(T_k - o_k)^2$ et $(T_k - c_k)^2$ sont là pour annuler le polynôme en $T_k = o_k$ et $T_k = c_k$ de façon plus lisse (zéro d'ordre 2).

Pour chaque marché, le résultat est une fonction à deux maxima locaux, et un minimum local (si existence d'une pause de midi) entre les deux maxima. La fonction est asymétrique si $s_k \neq 0$. Pour l'Amérique (où traditionnellement, il n'y a pas de baisse d'activité durant la pause de midi), on n'indique simplement pas le troisième facteur.

On demande que $a_0 > 0, \omega_k > 0$ et $(s_k \leq o_k \text{ ou } s_k \geq c_k)$ pour que l'on ait bien $a_{open,k}(t) \geq 0$. De même, on demande de façon évidente que $o_k < m_k < c_k$ (l'heure de la pause de midi est entre les heures d'ouverture et de fermeture du marché).

Ajustement des paramètres

Il faut encore ajuster en fonction des observations les paramètres intervenant dans $a_k(t)$. La méthode d'ajustement consistera à minimiser la somme des carrés des écarts entre a(t) définie plus haut et une mesure empirique de l'activité du marché définie comme suit. On part de la loi d'échelle (cf. section 3.4.3) en temps physique

$$\mathbb{E}[|r(\Delta t)|] \sim \Delta t^D$$

avec $\Delta t = 1$ heure. On subdivise la semaine en 168 heures comme plus haut, on ne considère que la *i*è heure de chaque semaine dans les données, et en notant r_i le return sur cette *i*è heure, on calcule une estimation (sur base d'un échantillon de plusieurs années) de $\mathbb{E}[|r_i|]$. L'idée du temps de marché θ sera d'être proportionnel à l'activité du marché: si on note $a_{1,2}$ l'activité "moyenne" du marché entre les instants (en temps physique) t_1, t_2 , alors l'intervalle correspondant en temps de marché $\Delta \theta_{1,2}$ sera défini par

$$\Delta \theta_{1,2} := (t_2 - t_1) a_{1,2}.$$

Cela revient en clair à contracter le temps en périodes d'activité faible et à le dilater en période d'activité intense. Si on suppose vraie la loi d'échelle pour tout intervalle de temps Δt avec les mêmes paramètres c, D, en prenant $\Delta t = \Delta t_i$ correspondant à la *i*è heure, on arrive à

$$\Delta \theta_i \sim \left(\mathbb{E}[|r_i|] \right)^{1/D}$$

et on définira finalement l'activité moyenne pour le *i*è intervalle d'une heure Δt_i par:

$$\hat{a}_i := \frac{c'}{\Delta t_i} \left(\mathbb{E}[|r_i|] \right)^{1/D}, \qquad \Delta t_i = 1 \text{ heure.}$$

La constante de renormalisation c' sera définie par

$$\frac{1}{168} \sum_{i=1}^{168} \hat{a}_i = 1.$$

La fonction $i \mapsto \hat{a}_i$ sera considérée comme une mesure (empirique) de l'activité et sera utilisée pour estimer les paramètres du modèle. Essentiellement, on minimisera la somme des carrés des écarts entre $a(t_i)$ et \hat{a}_i .

Nouvelle échelle de temps

Suivant la démarche annoncée plus haut, la nouvelle échelle de temps $\theta(t)$ sera définie par

$$\theta(t) := \int_{t_0}^t a(s) \, ds = a_0(t - t_0) + \sum_{k=1}^3 \int_{t_0}^t a_{1,k}(s) \, ds.$$

On peut voir que grâce à la normalisation imposée aux \hat{a}_i et donc à a(t), cela aura un sens d'exprimer θ dans les mêmes unités que t (en heures, semaines, mois,...). Vu la saisonalité des données et la définition de $\theta(t)$, si $t_2 - t_1 = 1$ semaine (càd 168 heures), alors $\theta(t_2) - \theta(t_1) = 168$ également⁴.

Retour aux données

Les données contiennent le cours USD-CHF pour chaque heure vingt-quatre en temps θ . Cela veut dire que les instants sont espacés par des intervalles $\Delta \theta = 1h24$. En fait, 1h24 = 1.4 heure et 1.4 = 168/120, où 168 est le nombre d'heures dans une semaine week-end inclus et 120 est le nombre d'heures dans une semaine week-end exclu. Comme l'activité totale est quasiment nulle durant les week-ends mais toujours positive à n'importe quel moment de la semaine, l'intervalle moyen $\overline{\Delta t}$ correspondant à $\Delta \theta = 1.4$ heure est d'environ 1 heure en temps physique.

Notons que toutes les manipulations de données citées dans la suite auront toujours été faites sur nos données désaisonnalisées. Quand nous dirons $\Delta t = 1$ heure, cela voudra dire en fait $\Delta \theta = 1h24$ et $\Delta t = 1$ heure en moyenne pendant les jours ouvrables.

3.4 Les faits stylisés

Les données à haute fréquence ont ouvert un nouveau champ d'exploration et ont mis en lumière certains comportements qui ne pouvaient pas être observés à des fréquences plus faibles. Nous donnons ci-dessous les principaux faits stylisés que l'on peut observer sur des données financières à haute fréquence. Ils sont exposés notamment dans [10], mais seront illustrés sur nos données. Ce sont ces différents faits qui motiveront l'introduction du modèle en cascade. Ces faits sont en réalité observés également dans d'autres types de données financières, pas uniquement pour le marché des changes, même si on s'intéressera dans cette section uniquement aux taux de change. Comme nous l'avons vu précédemment, l'avantage de ce marché est son activité 24 heures sur 24 et sa très grande liquidité.

⁴ce sera en fait un peu plus que 168 heures à cause de corrections à apporter du fait des jours fériés; en réalité, 4 ans en temps physique correspondront à 4 ans en temps de marché

3.4.1 Queue lourde et non normalité de la distribution des returns

Ce fait a été observé pour la première fois par Mandelbrot en 63 dans [26] dans les prix du coton, et a été depuis lors observé pour d'autres types de marchés. Ce sera encore plus flagrant si l'on considère les returns pour des petits intervalles, ce qui n'est possible que grâce aux données à haute fréquence.

De façon générale, il existe assez bien de modèles différents pour les distributions des returns. Certains pensent que les taux de change ont des returns assez proches de distributions stables, d'autres, des distributions de Student (qui ne font pas partie de la classe des distributions stables). Il y a également la classe des modèles conditionnels hétéroscédastiques (ARCH-GARCH, cf. [14, 4, 5]), et beaucoup d'auteurs pensent que ces modèles décrivent mieux les données que des modèles non conditionnels.

Les chercheurs sont néanmoins assez d'accord sur le caractère à queue lourde des distributions des returns. On est donc loin d'une modélisation par un processus gaussien.

On supposera dans la suite que les returns constituent un processus stationnaire. On verra plus loin qu'ils sont quasiment non autocorrélés. Voici quelques caractéristiques de la distribution des returns des données 5:

Δt	Moyenne	Médiane	Variance	Coefficient	Kurtosis
				d'asymétrie	
1 heure	5.163204 e-006	0.000000 e+000	1.729830 e-006	2.850384 e-001	15.42766
6 heures	3.073667 e-005	2.800000 e-005	1.336694 e-005	7.624225 e-003	4.7543024
1 jour	1.232066 e-004	4.300000 e-004	5.508311 e-005	2.043161 e-001	2.750464
1 semaine	6.160331 e-004	1.218000 e-003	2.812772 e-004	2.007003 e-002	1.185754

Ces résultats sont assez similaires à ceux donnés dans [10] (Chap. 5). La moyenne est très proche de 0, et les valeurs du coefficient d'asymétrie suggèrent que la distribution des returns est quasiment symétrique. Le kurtosis⁶ est d'autant plus élevé que l'intervalle Δt est petit. Les valeurs élevées du kurtosis pour $\Delta t = 1$ heure ou 6 heures suggèrent que la distribution des returns est beaucoup plus aplatie qu'une normale (et donc a une queue plus lourde) lorsque Δt est de l'ordre de quelques heures. Mais on retrouve une distribution plus proche d'une normale pour les returns sur 1 semaine (mais le kurtosis est néanmoins encore > 0). Les valeurs observées dans dans [10] sont assez proches de celles-ci, et du même ordre pour d'autres devises. Le phénomène est également observé dans [9]. Tous ces résultats suggèrent de regarder le comportement de queue de la distribution des returns. L'analyse de queue donnée dans [10, 34] semble indiquer que les 2è et 3è moments existent, mais pas le 4è (voir ci-dessous) de même que certains autres travaux

 $^{{}^{5}}$ Ces estimations sont faites sur des returns sans "overlapping", ce qui diminue la dépendance entre les returns (dépendance forte pour des returns avec overlapping). La fonction d'autocorrélation des returns est par ailleurs très faible, voir plus loin

 $^{{}^{6}=(}r(\Delta t)-\overline{r(\Delta t)})^{4}/s(r(\Delta t))^{4}-3$, si > 0, plus aplati qu'une normale, et d'autant plus élevé que l'indice de queue est petit

sur la question, et pas uniquement pour des séries issues du marché des devises (voir [9]).

La figure 3.3 donne un quantile-quantile plot (par rapport à une normale standard) des returns pour $\Delta t = 1$ heure, 6 heures, 1 jour et 1 semaine. La forme en S des graphiques pour $\Delta t = 1$, 6 ou même 24 heures est caractéristique des distributions à queue lourde. Voir également la figure 3.4. La figure 3.5 donne la densité logarithmique des données comparée avec celle d'une Student à 4 degrés de liberté (à gauche et à droite) et 3 degrés de liberté (à droite). Sur la figure de droite sont représentées toutes les valeurs extrêmes. La figure 3.6 montre un qqplot par rapport à une Student à 3, 4 et 5 degrés de liberté. Ces graphes semblent confirmer les résultats d'indice de queue (pour des returns sur une heure environ) obtenus pour ce type de données (indice entre 3 et 4), et la loi de Student semble assez bien ajuster la distribution des returns pas uniquement dans la queue.



Figure 3.3: Qqplot des returns (centrés réduits) pour $\Delta t = 1$ heure, 6 heures, 1 jour et 1 semaine par rapport à une normale standard.



Figure 3.4: Fonction de densité (à gauche) et logarithme de la fonction de densité (à droite) des returns centrés réduits pour $\Delta t = 1$ h et comparaison avec une normale standard.



Figure 3.5: Fonction de densité logarithmique des returns centrés réduits pour $\Delta t = 1$ h et comparaison avec une Student à 3 ou 4 degrés de liberté.



Figure 3.6: QQ plot des returns (centrés réduits) pour $\Delta t=1{\rm h}$ et comparaison avec une Student à 3, 4 et 5 degrés de liberté.

Indice de queue de la distribution des returns

Le type de queue d'une distribution peut être résumé grâce à son indice de queue. On peut classer les distributions en trois types:

- les distributions à queue légère, où tous les moments de la distribution existent, et (1 -) la fonction de répartition décroît exponentiellement dans la queue.
- les distributions à queue lourde, caractérisées par (1) une fonction de répartition qui décroît comme une certaine puissance,
- les distributions bornées, prenant des valeurs contenues dans un borné, et ne possédant donc pas de queue.

En cas de que ue lourde, on peut définir l'indice de que ue par le réel α tel que

$$\lim_{x \to \infty} (1 - F(x)) |x|^{\alpha}$$

existe dans $(0, +\infty)$. Si l'indice de queue de la distribution d'une variable aléatoire vaut α , alors seuls les k premiers moments $\mathbb{E}[X^k]$ pour $k < \alpha$ sont finis. En cas de queue légère (comme pour la distribution normale), on peut également définir l'indice de queue par $\alpha = \infty$.

Connaître l'indice de queue des returns est intéressant en pratique pour évaluer le risque des marchés financiers, mais aussi pour la modélisation en finance en général, car beaucoup de modèles reposent sur l'hypothèse de l'existence de la variance des returns (il a subsisté toute une controverse entre chercheurs quand à l'existence de la variance de la distribution des returns). Évaluer les indices de queue est très important en théorie des valeurs extrêmes pour les marchés financiers. La BIS a par exemple établi des règles à suivre par les banques pour contrôler leurs risques, mais la plupart des modèles utilisés dans ces règles sont encore basés sur l'hypothèse que les actifs financiers sont distribués suivant une distribution normale. Avoir plus d'information sur l'indice de queue des séries financières est donc une question d'intérêt en soi.

Estimer les indices de queue (sans distribution connue a priori) est une tache difficile et est un sujet actuel de recherche en statistique. Pour déterminer l'indice de queue d'une distribution, il faut énormément d'observations (puisque lié aux valeurs extrêmes de la distribution). En ce sens, les données à haute fréquence constituent un atout, puisqu'elles fourniront de grands échantillons. Diverses méthodes issues de la théorie des valeurs extrêmes existent, et ce sujet est trop vaste pour que l'on puisse s'y étendre dans un mémoire. Dans [34] est proposée une méthode d'estimation des indices de queue par une méthode de bootstrap basée sur l'estimateur de Hill. Dans ce qui suit est exposée brièvement l'idée de [34]. Mentionnons également la méthode du "peak over threshold" (POT, cf. [25]).

On suppose que l'on est en présence d'un échantillon de n observations X_1, \ldots, X_n d'un processus stationnaire i.i.d. de fonction de répartition F inconnue. On suppose

3.4. Les faits stylisés

que la distribution est à queue lourde. On notera les statistiques d'ordre décroissantes par $X_{(1)} \ge \ldots \ge X_{(n)}$. Sans entrer dans les détails, voici ci-dessous l'idée générale de leur estimation. Les auteurs considèrent d'abord différents estimateurs de l'inverse de l'indice de queue α de la distribution, $\gamma = 1/\alpha$:

• L'estimateur de Pickands:

$$\hat{\gamma}_{n,m}^P := \frac{1}{\ln 2} \ln \frac{X_{(m)} - X_{(2m)}}{X_{(2m)} - X_{(4m)}},$$

• L'estimateur de Hill:

$$\hat{\gamma}_{n,m}^{H} := \frac{1}{m-1} \sum_{i=1}^{m-1} \ln X_{(i)} - \ln X_{(m)}$$

où m > 1. On peut montrer que cet estimateur est tel que $(\hat{\gamma}_{n,m}^H - \gamma)m^{1/2}$ est asymptotiquement normal de moyenne nulle et de variance γ^2 , mais pour des échantillons de taille finie, cet estimateur est biaisé. Le problème est que ce biais est inconnu a priori et dépend de m (problème non trivial si le type de distribution est inconnu a priori).

• L'estimateur de De Haan et Resnick:

$$\hat{\gamma}_{n,m}^R := \frac{\ln X_{(1)} - \ln X_{(m)}}{\ln m},$$

• Une extension de l'estimateur de Hill proposée par Dekkers, Einmahl et De Haan:

$$\hat{\gamma}_{n,m}^{D} := \hat{\gamma}_{n,m}^{H} + 1 - \frac{1}{2} \left[1 - \frac{(\hat{\gamma}_{n,m}^{H})^2}{\hat{\gamma}_{n,m}^{H(2)}} \right]^{-1}$$

où

$$\hat{\gamma}_{n,m}^{H(2)} := \frac{1}{m-1} \sum_{i=1}^{m-1} (\ln X_{(i)} - \ln X_{(m)})^2.$$

On peut prouver également la normalité de cet estimateur.

Concernant les estimateurs donnés ci-dessus et leurs propriétés, on peut aller voir les références citées dans [34] (dont [11, 12, 22, 33]). Ils requièrent tous que $m(n) \to \infty$ pour $n \to \infty$, mais on ne sait pas comment choisir m de façon optimale pour un échantillon de taille finie. En effet, le problème est que ces estimateurs ont un biais dépendant de m et n, et le problème sera d'être le plus proche possible de la vraie valeur de γ avec la variance la plus petite possible.

Dans [34] sont simulées en grand nombre des données suivant des lois pour lesquelles l'indice de queue est connu a priori; ils calculent ensuite sur ces simulations les différents estimateurs définis ci-dessus (voir [34] pour les détails au sujet de ces simulations). Les lois considérées sont des distributions de Student, de Cauchy, stables (non gaussiennes) et normale (pour laquelle on a $\alpha = \infty$). Ils calculent ces estimateurs pour un grand nombre de simulations et effectuent des moyennes sur ces différentes simulations (car certains estimateurs varient très fort d'une simulation à l'autre). Leur conclusion est que les estimateurs donnant les meilleurs résultats pour les cas examinés sont les estimateurs de Hill et de Dekkers-Einmahl-De Haan. L'estimateur de Hill a fait l'objet de beaucoup de recherches et il y a déjà assez bien de résultats théoriques le concernant, c'est pourquoi les auteurs ont décidé de concentrer la suite de leur travail sur cet estimateur.

Ils calculent ensuite les moyennes et variances de $\hat{\gamma}_{n,m}^{H}$ sur 100 simulations de chaque loi considérée. Ils constatent que la variance croît vers l'infini lorsque *m* s'approche de 0, mais que le biais augmente avec *m*. De plus, le biais dépend de la distribution considérée. Il est d'autant plus petit que la distribution a un indice α grand.

Le principal problème qui se pose avec cet estimateur est de trouver comment choisir m de manière optimale, c'est-à-dire de façon à minimiser le biais existant pour m grand, mais aussi minimiser la variance de l'estimateur. Prendre m petit ne semble absolument pas raisonnable au vu de la variance très élevée dans ce cas. Il faut donc faire un compromis entre la variance et le biais.

Une possibilité est de minimiser la somme des carrés des écarts entre $\hat{\gamma}_{n,m}^{H}$ et la valeur théorique γ . Mais le problème est que dans la plupart des études empiriques sur des données, la valeur théorique de γ est inconnue.

Pour justifier leur méthode d'estimation, ils commencent par calculer la moyenne et la variance de $\hat{\gamma}_{n,m}^{H} \stackrel{not}{=} \gamma_{n,m}$. Ils supposent pour cela que les *n* observations sont i.i.d. de loi *F*. Ils commencent par calculer des expressions de $\mathbb{E}[\gamma_{n,m}]$ et $\operatorname{Var}[\gamma_{n,m}]$ en utilisant la formule de Bayes, en conditionnant par rapport à la *m*è statistique d'ordre $X_{(m)}$. Ils utilisent alors la définition de l'indice de queue α en écrivant (après avoir fait un changement de variable sur l'axe des *x* pour "recentrer" la distribution, de sorte que F(0) = 1/2 après changement de variable):

$$F(x) = 1 - a(x - c)^{-\alpha} [1 + b(x - c)^{-\beta} + o((x - c)^{-\beta})]$$

où $a > 0, \alpha > 0$ est l'indice de queue, $\beta > 0$ et $b, c \in \mathbb{R}$. Les valeurs de x correspondant à la partie positive de la queue sont telles que $x \gg |c|$. On développe alors F(x) pour |c|/x dans un voisinage de zéro (i.e. $x \gg |c|$) en puissances négatives de x, pour obtenir une expression un peu plus maniable que la précédente:

$$F(x) = 1 - ax^{-\alpha} [1 + b^* x^{-\beta^*} + o(x - \beta^*)]$$

où $\beta^* := \min(\beta, 1)$ et

$$b^* := \begin{cases} b & \text{if } \beta < 1\\ b + c\alpha & \text{if } \beta = 1\\ c\alpha & \text{if } \beta > 1 \end{cases}$$

Notons pour la suite que β^* est maintenant inférieur ou égal à 1 sauf si c = 0.
3.4. Les faits stylisés

On reporte alors l'expression de F(x) dans les formules bayésiennes obtenues pour $\mathbb{E}[\gamma_{n,m}]$ et on obtient finalement (après quelques développements, et après avoir remplacé β^* et b^* par β, b pour simplifier les notations):

$$\mathbb{E}[\gamma_{n,m}] = \frac{1}{\alpha} + B$$

où le biais B vaut:

$$B = -\frac{1}{\alpha} \frac{\beta b'}{\alpha + \beta} \frac{\Gamma(m + \frac{\beta}{\alpha})}{\Gamma(m)} (a(n - m))^{-\beta/\alpha}, \quad b' = b(1 + o(1))$$

pour $|c|/x \to 0$, où γ est la fonction gamma d'Euler. En utilisant un développement du quotient $\frac{\Gamma(m+\frac{\beta}{\alpha})}{\Gamma(m)}$ par la formule de Stirling, ils obtiennent finalement

$$B = -\frac{1}{\alpha} \frac{\beta b}{\alpha + \beta} a^{-\beta/\alpha} (m/n)^{\beta/\alpha} \left\{ 1 + O(1/m) + O(m/n) + o(1) \right\}$$

pour $1/m \to 0$ et $m/n \to 0$. Un calcul similaire est fait pour la variance de $\gamma_{n,m}$, pour obtenir finalement:

$$\operatorname{Var}[\gamma_{n,m}] = \frac{1}{\alpha^2 m} \left(1 + O[1/m] + O[(m/n)^{\beta/\alpha}] \right).$$

Des résultats précédents, on peut calculer l'erreur moyenne quadratique

$$\mathbb{E}\left[\left(\gamma_{n,m} - \frac{1}{\alpha}\right)^2\right] = B^2 + \mathbb{E}\left[\left(\gamma_{n,m} - \mathbb{E}[\gamma_{n,m}]\right)^2\right]$$
$$= \frac{1}{\alpha^2} \frac{\beta^2 b^2}{(\alpha + \beta)^2} a^{-2\beta/\alpha} (m/n)^{2\beta/\alpha} + \frac{1}{\alpha^2 m}$$

Cette erreur est grande lorsque m est grand par rapport à n ou si m est proche de 0.

Dans ce qui précède, les expressions comportent encore, outre α que l'on cherche à estimer, des paramètres inconnus β , a et b. On peut maintenant essayer de minimiser l'erreur quadratique, ou estimer le biais B, mais il faut au préalable estimer les paramètres inconnus.

On peut voir que la valeur de m minimisant l'erreur quadratique moyenne est donné par

$$\bar{m} = \left(\frac{\alpha(\alpha+\beta)^2}{2\beta^3 b^2}\right)^{\alpha/(\alpha+2\beta)} (a\,n)^{2\beta/(\alpha+2\beta)} \tag{3.1}$$

Si on reporte cette valeur de m dans l'expression de l'erreur quadratique, on obtient le comportement asymptotique suivant par rapport à n:

$$\mathbb{E}\left[(\gamma_{n,m}-\frac{1}{\alpha})^2\right] \propto n^{-\frac{2\beta}{\alpha+2\beta}},$$

ce qui montre que l'erreur diminue lorsque la taille de l'échantillon augmente, et d'autant plus si α est petit et β est grand. Dans ce qui suit, nous donnons l'idée de la méthode de bootstrap utilisée pour trouver le meilleur m à prendre dans l'estimateur de Hill.

Comme (3.1) contient des paramètres inconnus, on ne peut pas l'utiliser directement pour obtenir le meilleur m. Hall a proposé dans [21] une méthode de bootstrap pour résoudre ce genre de problème.

On considère un sous-échantillon de taille réduite $n_1 < n$, et on cherche le m optimal correspondant à n_1 en calculant une estimation de:

$$\min_{m_1} \mathbb{E}\left[(\gamma_{n_1,m_1} - \gamma_0)^2 | F_n \right]$$
(3.2)

où $\gamma_0 = \gamma_{n,m_0}$ est une estimation réalisée sur tout l'échantillon avec un m initial (non optimal) m_0 , et où F_n est la fonction de répartition empirique de tout l'échantillon. Cette dernière expression peut être calculée explicitement sur base des données. La quantité γ_0 est une bonne estimation de $1/\alpha$ pour le sous-échantillon, puisqu'elle est basée sur un échantillon de longueur plus grande. Etant donné $\overline{m_1}$ atteignant le minimum dans (3.2), la valeur optimale \overline{m} peut être obtenue en prenant le rapport $\overline{m}/\overline{m_1}$ et en utilisant (3.1):

$$\bar{m} = \bar{m}_1 \left(\frac{n}{n_1}\right)^{\frac{2\beta}{2\beta+\alpha}}$$

C'est cette statistique \bar{m} sur laquelle sera fait le bootstrap. On notera \hat{m} est l'estimation par bootstrap de \bar{m} . On choisit ensuite comme estimation de α : $\hat{\alpha} = 1/\gamma_{n,\hat{m}}$.

Problème: il faut se débarrasser des paramètres α (que l'on cherche à estimer) et β qui apparaissent dans l'expression de \hat{m} (pour bien avoir une statistique). Ils proposent d'utiliser comme valeur de α simplement $1/\gamma_0$. L'estimation de β est plus compliquée. Ils évaluent l'exposant $\frac{2\beta}{2\beta+\alpha}$ pour différentes valeurs de β et comparent les résultats, qui semblent ne pas trop dépendre de la valeur de β .

Pour le cours USD-CHF, ils trouvent dans [10] des indices de queue autour de 3.5 pour les returns avec $\Delta t = 30$ min, et jusqu'à 5.6 pour $\Delta t = 1$ jour. Les valeurs pour les autres devises sont assez proches. Ils remarquent également que les indices pour les taux croisés (c'est-à-dire lorsque les deux devises sont différentes du USD) des monnaies faisant partie (anciennement) du Système Monétaire Européen sont plus faibles. D'après eux, les distributions des returns sont toujours à queue lourde et non stable.

Question qui se pose: quelle est la qualité de leur estimation de l'indice de queue? L'estimation par estimateur de Hill (dans le cas non paramétrique) n'est toujours pas un problème résolu et est assez ardu en soit. Quel est l'indice de queue de nos données et de notre modèle (introduit plus loin)? Déjà tout un programme en soit.

3.4.2 Comportement des volatilités réalisées des returns

Notion de volatilité réalisée

Une mesure de l'agitation du marché au cours du temps peut-être obtenue grâce à la volatilité réalisée définie comme suit.

Définition 3.4.1 Supposons donnée une série chronologique homogène (i.e. espacée régulièrement dans le temps, pour une certaine échelle de temps) de returns $\{r(t_i, \Delta t)\}$. La volatilité réalisée (ou volatilité historique) au temps t_i est définie par

$$v(t_i) = v(\Delta t, n, p; t_i) := \left(\frac{1}{n} \sum_{j=1}^n |r(t_{i-n+j}, \Delta t)|^p\right)^{\frac{1}{p}}.$$

Pour calculer cette volatilité, on a donc besoin de fixer:

- 1. l'intervalle de temps Δt sur lequel on considère les returns (avec $t_i t_{i-1} = \Delta t$ dans le cas de returns sans recouvrement (non overlapped)),
- 2. la fenêtre de temps sur laquelle on se place $(n\Delta t)$, ou encore le nombre *n* d'observations que l'on fait intervenir,
- 3. l'exposant p (souvent, on choisit p = 1 ou 2)

On peut encore faire un changement d'échelle de temps dans les volatilités (et considérer alors les volatilités annuelles, mensuelles,...). En général, les praticiens utilisent toujours la volatilité annualisée. Remarquons qu'il s'agit d'un p-norme des returns.

Pour chaque instant t_i , on peut donc calculer la volatilité réalisée $v(t_i)$ pour ainsi obtenir une nouvelle série chronologique.

On pourrait également centrer la volatilité réalisée en la définissant par

$$v'(t_i) := \left\{ \frac{1}{n-1} \sum_{j=1}^n \left| r(t_{i-n+j}, \Delta t) - \frac{1}{n} \sum_{k=1}^n r(t_{i-n+k}, \Delta t) \right|^p \right\}^{\frac{1}{p}}.$$

Si p = 2, on retrouve l'écart-type (échantillon) des returns. Les deux définitions données ci-dessus sont réellement différentes en pratique si les returns ont une moyenne fort différente de 0, c'est-à-dire en présence d'un trend.

Choisir p grand donne évidemment plus de poids aux valeurs extrêmes. En fait, si la distribution des returns est à queue lourde, prendre p trop grand (plus grand que l'indice de queue) peut aboutir à une volatilité réalisée de moyenne asymptotiquement infinie.

Un cas particulier de volatilités réalisées peut être obtenu par la séries des returns absolus ou des returns carrés. C'est le cas où n = 1 et p = 1 ou 2. Lorsque p = 1et n est quelconque, on parle de returns moyens absolus (mean absolute returns). On mesurera en général la volatilité du marché par les returns moyens absolus (cas p = 1), ce qui est mieux que les racines carrées des carrés des returns (cas p = 2) car on a besoin de l'existence du 4è moment des returns pour pouvoir considérer la fonction d'autocorrélation des returns carrés (nous indiquant dans quelle mesure les volatilités ont de la mémoire), et cette existence n'est pas évidente pour les données empiriques comme on l'a déjà vu précédemment. Une autre raison de préférer les returns absolus est encore leur apparente meilleure stabilité par rapport à la taille de l'échantillon. Quand on calcule ces volatilités réalisées, leur résolution est très importante: certains événements ne sont observables qu'à de hautes résolutions (par exemple une agitation des cours pendant 30 minutes suivie d'une accalmie et d'un retour au cours initial ne sera pas nécessairement perçue à une résolution d'une heure).

Volatility clustering – Autocorrélation des returns absolus

Les fonctions d'autocovariance c_t et d'autocorrélation r_t des returns $r(t, \Delta t)$ pour l'intervalle Δt est définie comme suit. Pour chaque délai t, on définit

$$c_t := \frac{1}{n} \sum_{s=1}^{n-t} (r(t+s, \Delta t) - \overline{r(., \Delta t)}) (r(s, \Delta t) - \overline{r(., \Delta t)}), \quad r_t := \frac{c_t}{c_0}, \qquad t \in \{1, \dots, Lag_{max}\}$$

où n est la longueur de la série, Lag_{max} est le délai maximum considéré et

$$\overline{r(.,\Delta t)} := \frac{1}{n} \sum_{s=1}^{n} r(s,\Delta t).$$

Une propriété importante observée universellement sur des séries financières est le phénomène de "volatility clustering": le fait qu'il semble y avoir des périodes de volatilité importante et d'autres de volatilité plus faible, et cela, même en temps de marché. C'est le fait qu'il y ait une accumulation des volatilités, ce qui n'est pas le cas dans un mouvement brownien par exemple. Ce phénomène a été observé pour la première fois par Mandelbrot [26].

Ce phénomène est directement relié à (et peut être mis en évidence par) l'observation selon laquelle la fonction d'autocorrélation des returns absolus décroît lentement et est positive, bien en dehors de l'intervalle de confiance à 95% pour des observations i.i.d. normales. On l'observe également en turbulence (cf.[18]).

Dans [10] sont étudiées la fonction d'autocorrélation des returns, des returns absolus et de leurs carrés, mais pour des données non désaisonnalisées. Ils constatent que la fonction d'autocorrélation des returns est très faible, mais qu'au contraire celle des returns absolus et carrés est positive et nettement en dehors de l'intervalle de confiance à 95 % sous l'hypothèse d'observations indépendantes gaussiennes.

Dans [10] est également étudiée la fonction d'autocorrélation de puissances des returns absolus $(|r|^p)$ en fonction de l'exposant de la puissance. Ils constatent que cette fonction d'autocorrélation décroît si p augmente. Comme dans ce cas le poids relatif des événements extrêmes augmente dans la fonction d'autocorrélation, ils en déduisent que les événements extrêmes sont moins corrélés entre eux que les autres.

Nous calculons également ces fonctions d'autocorrélation pour nos données. La figure 3.9 représentent la fonction d'autocorrélation des returns absolus et returns carrés pour nos données désaisonnalisées. On remarque que malgré la désaisonnalisation, il subsiste encore dans la fonction d'autocorrélation des pics environ toutes les 24 heures en temps physique. Cela suggère que la désaisonnalisation n'est pas encore "parfaite". Les valeurs de la fonction d'autocorrélation pour des délais du même ordre que ceux considérés dans [10] sont légèrement inférieures à celles qu'ils ont observées. La fonction d'autocorrélation des returns eux-mêmes est négligeable puisqu'elle reste dans l'intervalle de confiance à 95 %.

Ce qui est donc le plus flagrant, c'est la relativement lente décroissance de la fonction d'autocorrélation des returns absolus (et carrés).

Les figures suivantes illustrent le phénomène de volatility clustering sur nos données. La figure 3.7 représente les volatilités réalisées pour $\Delta t = 1$ heure, n = 10 et p = 1(le premier graphique donne ces volatilités en seulement 1000 points, le deuxième en 3500 points et le troisième en 6500 points). La figure 3.8 donne les returns absolus pour les données et des simulations de mouvement brownien et brownien fractionnaire (les returns suivent alors le bruit gaussien ou gaussien fractionnaire correspondant). On constate clairement que le nombre de grands mouvements est plus important dans le cas de nos données, et le phénomène de "volatility clustering" peut être perçu. Mais cela ne devient convaincant qu'en calculant la fonction d'autocorrélation de returns absolus (qui permet de mettre en évidence le phénomène).

La figure 3.10 représente la fonction d'autocorrélation des returns absolus et des returns carrés en échelle bilogarithmique. Ces graphes suggèrent que ces fonctions décroissent à l'infini comme une puissance (décroissance hyperbolique). Les volatilités réalisées semblent donc constituer ce que l'on appelle un processus à longue mémoire.

La figure 3.11 représente la fonction d'autocorrélation des returns absolus pour différents intervalles de temps (d'une heure à 1 semaine, plus précisément pour des intervalles en temps de marché $\Delta\theta$ tels que le Δt correspondant soit en moyenne égal à 1 heure, 6 heures, ...) (délais maxima: de 5 à 25 semaines).

La figure 3.13 donne la fonction d'autocorrélation des returns absolus de nos données et de simulations d'un mouvement brownien et brownien fractionnaire.

La figure 3.12 montre la fonction d'autocorrélation des returns pour différents intervalles de temps (délais maxima: de 5 à 25 semaines). On constate que celle-ci est contenue essentiellement dans l'intervalle de confiance à 95%.



Figure 3.7: Données: Volatilités réalisées pour $\Delta t = 1$ heure, n = 10 et p = 1. Les différentes figures représentent la même fonction regardée (en haut) sur les 1500 premiers points, (au milieu) sur les 3000 premiers et (en bas) sur les 7000 premiers.



Figure 3.8: Returns absolus pour les données (en haut), pour un mouvement brownien standard (au milieu) et un FBM d'indice de Hurst H = 0.8 (en bas).



Figure 3.9: Données: fonction d'autocorrélation des returns absolus (en haut) et des returns carrés (en bas) (returns pour $\Delta t = 1$ heure). La bande en pointillé représente l'intervalle de confiance à 95 % pour la fonction d'autocorrélation sous l'hypothèse d'observations indépendantes gaussiennes. L'unité de temps est l'heure.



Figure 3.10: Données: fonction d'autocorrélation des returns absolus (à gauche) et carrés (à droite) ($\Delta t = 1$ heure) en échelle bilogarithmique.



Figure 3.11: Données: fonction d'autocorrélation des returns absolus pour différents intervalles (en haut à gauche: $\Delta t = 1$ heure, en haut à droite: 6 heures, en bas à gauche: 24 heures et en bas à droite: une semaine). Bande en pointillé: intervalle de confiance à 95 % pour la fonction d'autocorrélation d'observations indépendantes gaussiennes. Unité de temps des graphiques: une semaine.



Figure 3.12: Données: fonction d'autocorrélation des returns pour différents intervalles (en haut à gauche: $\Delta t = 1$ heure, en haut à droite: 6 heures, en bas à gauche: 24 heures et en bas à droite: une semaine). Bande en pointillé: intervalle de confiance à 95 % pour la fonction d'autocorrélation d'observations indépendantes gaussiennes. Unité de temps des graphiques: une semaine.



Figure 3.13: Fonction d'autocorrélation des returns absolus pour les données (en haut, avec la semaine comme unité de temps et l'intervalle de calcul des returns Δt égal à 1 heure), un mouvement brownien standard (au milieu) et un FBM d'indice H = 0.8 (en bas).

3.4.3 Lois d'échelle

Ces idées de propriétés d'échelle remontent à [26] pour le prix du coton. Des lois d'échelle ont ensuite été observées empiriquement dans [28] pour des taux de change, ainsi que dans d'autres types de séries financières par d'autres personnes. Depuis, on en parle beaucoup, et certains auteurs disent observer des comportements de loi d'échelle dans beaucoup de données financières.

On dira que l'on est en présence d'un tel comportement pour un processus de prix lorsque, si l'on définit la fonction de structure des returns par:

$$S_q(\Delta t) := \frac{\Delta t}{N} \sum_{j=1}^{N/\Delta t} |r(j\Delta t, \Delta t)|^q,$$

avec N la longueur de la série (returns toujours pris ici sans overlapping), on observe le comportement

$$S_q(\Delta t) \propto (\Delta t)^{\tau_q+1}$$

pour une certaine fonction τ_q (la même donc pour tout Δt , de même que le coefficient de proportionnalité).

Dans le cas où le processus des prix (logarithmiques) est un mouvement brownien, on sait que $\mathbb{E}[|r(\Delta t)|^q] = (\Delta t)^{\frac{1}{2}q}$, et dans le cas plus général d'un mouvement brownien fractionnaire d'indice $H \in (0, 1)$, la puissance dans la relation précédente est remplacée par Hq. On devrait donc trouver dans les deux cas précédents des comportements d'échelle avec des exposants $\tau_q + 1$ égaux à 1/2q ou Hq.

On peut essayer d'observer de tels comportement dans nos données, et tenter de calculer les exposants $\xi(q) := \tau_q + 1$ pour différentes valeurs de q. Pour réaliser les graphiques suivants, on a d'abord calculé $S_q(\Delta t)$ pour différentes valeurs de Δt et de q. On a ensuite tracé dans une échelle bilogarithmique $S_q(\Delta t)$ en fonction des valeurs de Δt . On a choisi de faire varier q de 1 (ou de 0.5) à 4 par pas de 0.5 et Δt de 2 à 685 par pas multiplicatif de 1.7. Sur la figure 3.14, on voit qu'en échelle bilogarithmique, les valeurs de log $S_q(\Delta t)$ ont l'air de se répartir le long de droites, ce qui semble indiquer un comportement d'échelle. On a effectué ensuite une simple régression linéaire (non pondérée) des valeurs log $S_q(\Delta t)$ en fonction des log Δt et tracé les droites de régression. Les coefficients angulaires de ces droites ont alors été pris comme $\xi(q)$. Les résultats sont les suivants:

$\xi(0.5)$	$\xi(1)$	$\xi(1.5)$	$\xi(2)$	$\xi(2.5)$	$\xi(3)$	$\xi(3.5)$	$\xi(4)$
0.27	0.53	0.77	1.01	1.23	1.44	1.64	1.82



Figure 3.14: $\log S_q(\Delta t)$ en fonction de $\log \Delta t$ et régression linéaire



Figure 3.15: Données: exposants d'échelle $\xi(q)$, et comparaison notamment avec la droite y = 0.5x (graphique de gauche)

Dans notre calcul des $\xi(q)$ sur les données, les coefficients de corrélation carrés R^2 des régressions linéaires étaient tous supérieurs à 0.95. On constate que la valeur de $\xi(1)$ est légèrement supérieure à 0.5, comme ça aurait été le cas pour un mouvement brownien. Dans [10], ils estiment $\xi(1)$ à une valeur proche de 0.6. Le graphe de la fonction $\xi(q)$ (figure 3.15) est très légèrement concave (et donc non linéaire), mais quelle est la pertinence de l'observation des lois d'échelle pour q grand, là où les moments n'existent pas nécessairement et où il faut énormément d'observations pour l'estimation (alors que l'on est limité par la taille de l'échantillon)? Quelle est la stabilité du $\xi(q)$ ainsi calculé? Est-ce une bonne estimation de $\log(\mathbb{E}[|r(\Delta t)|^q|])/\log(\Delta t)$? Est-ce que cela prouve l'existence de lois d'échelle ou la multifractalité?

A titre d'exemple, on peut suivre la même démarche sur des simulations d'un mouvement brownien et d'un mouvement brownien fractionnaire (suivant le même algorithme de simulation qu'à la section 1.2.5, seulement 40000 points par simulation) pour lesquels on connaît a priori la valeur de $\log(\mathbb{E}[|r(\Delta t)|^q|])/\log(\Delta t)$. L'estimation est satisfaisante pour le mouvement brownien, mais pour le FBM avec H = 0.8, selon la simulation considérée on obtient des valeurs différentes, mais qui semblent néanmoins se placer quand-même le long d'une droite et non d'une courbe concave (voir figure 3.16). Ceci indique une relativement lente convergence de $S_q(\Delta t)$ vers $\mathbb{E}[|r(.,\Delta t)|^q]$, expliquée par le caractère à longue mémoire du mouvement brownien fractionnaire d'indice H = 0.8.

Dans le cas de nos données, on peut également scinder notre série en plusieurs sous-échantillons (11 sous-échantillons de 15000 points chacun, se recouvrant en partie évidemment). On voit par exemple que $\tau_1 + 1$ a une légère tendance à diminuer sur la fin de la période considérée, ce qui est également mentionné dans [10].

$\xi(0.5)$	$\xi(1)$	$\xi(1.5)$	$\xi(2)$	$\xi(2.5)$	$\xi(3)$	$\xi(3.5)$	$\xi(4)$
0.28	0.54	0.80	1.04	1.27	1.49	1.70	1.89
0.28	0.54	0.78	1.01	1.22	1.42	1.59	1.74
0.29	0.55	0.81	1.05	1.28	1.49	1.69	1.87
0.27	0.53	0.78	1.00	1.21	1.39	1.55	1.69
0.28	0.54	0.79	1.02	1.23	1.42	1.59	1.75
0.27	0.54	0.79	1.03	1.24	1.44	1.61	1.77
0.27	0.53	0.78	1.01	1.24	1.46	1.66	1.85
0.27	0.54	0.79	1.03	1.26	1.46	1.65	1.83
0.27	0.53	0.77	1.01	1.23	1.44	1.64	1.82
0.26	0.51	0.75	0.97	1.19	1.39	1.58	1.75
0.27	0.52	0.77	1.00	1.22	1.43	1.63	1.81
0.27	0.53	0.78	1.02	1.24	1.44	1.63	1.80
0.01	0.01	0.02	0.02	0.03	0.04	0.05	0.06

Les deux dernières lignes de ce tableau donnent la moyenne et l'écart-type des différentes observations sur les différents sous-échantillons. Les valeurs moyennes sont assez proches de celles calculées précédemment sur tout l'échantillon. La figure 3.17 montre graphiquement les différentes valeurs calculées sur les différents sous-échantillons



Figure 3.16: Exposants d'échelle $\xi(q)$ pour un mouvement brownien (à gauche) et un mouvement brownien fractionnaire avec H = 0.8 (à droite), pour différentes séries simulées. En ligne continue, la droite y = Hx, comparée avec les valeurs calculées des $\xi(q)$.



Figure 3.17: Données: $\log S_q(\Delta t)$ en fonction de $\log \Delta t$ pour différents sous-échantillons



Figure 3.18: Données: comparaison entre la moyenne sur les différents sous-échantillons des exposants $\xi(q)$ et leurs valeurs calculés sur tout l'échantillon.

(voir également la figure 3.18). On voit bien que les fluctuations sont relativement plus importantes lorsque q est grand.

Lien avec la fonction de structure de l'analyse multifractale

Si le facteur multiplicatif que l'on prend entre les différents intervalles Δt pour lesquels on calcule $S_q(\Delta t)$ est égale à 2 et si $N/\Delta t$ est une puissance de 2, alors, les fonctions $S_q(\Delta t)$ calculées sont en fait (essentiellement) égales aux fonctions $S_{\alpha}^{(n)}(q)/2^n$ introduites au chapitre 2 (voir la définition 2.9). Le coefficient angulaire des droites le long desquelles $\ln(S_q(\Delta t))$ ont l'air de très bien se disposer est calculé essentiellement de la façon suivante:

$$\frac{\ln\left(\frac{S^{(n)}(q)}{N/\Delta t}\right) - \ln\left(\frac{S^{(n-1)}(q)}{2N/\Delta t}\right)}{\ln(2^{-n}) - \ln(2^{-n+1})} = \log_2(S^{(n-1)}(q)) - \log_2(S^{(n)}(q)) + 1$$
$$= -(1-n)\frac{\log_2(S^{(n-1)}(q))}{n-1} - n\frac{\log_2(S^{(n)}(q))}{n} + 1.$$

Or, si n est suffisamment grand (autrement dit, si $N/\Delta t$ est suffisamment grand, c'està-dire si la longueur de la série est suffisamment grande par rapport aux intervalles Δt), ce que l'on calcule comme fonction de structure est une estimation de $\mathbb{E}[|r(\Delta t)|^q]$ (convergence presque sûre vers cette valeur si N tend vers l'infini). En calculant le coefficient angulaire de ces droites, si $N/\Delta t$ est suffisamment grand, cela peut être vu comme une estimation (au sens que l'on a convergence presque sûre de $\ln(S_q(\Delta t)$ vers $\mathbb{E}[|r(\Delta t)|^q])$ de

$$\frac{\ln(\mathbb{E}[|r(I_0^{(n-1)})|^q]) - \ln(\mathbb{E}[|r(I_0^{(n)})|^q])}{\ln(2^{n-1}) - \ln(2^n)} = \log_2(\mathbb{E}[|r(I_0^{(n-1)})|^q]) - \log_2(\mathbb{E}[|r(I_0^{(n)})|^q])$$

pour un certain n lié à $N/\Delta t$, ce qui tend vers 1 + T(q) pour $n \to \infty$ (ou du moins au pire en limite inférieure). On pourrait donc voir ces coefficients angulaires comme une estimations des T(q) + 1.

Comme nous l'avons vu, même si le processus ne possède pas de loi d'échelle au sens strict (global), si les exposants T(q) existent, il en possédera à la limite pour des intervalles de temps infinitésimaux(localement), et aura donc un comportement proche d'un comportement d'échelle pour des intervalles suffisamment petits. Le raisonnement ci-dessus montre que l'on obtient par ce procédé, à condition que $N/\Delta t$ soit suffisamment grand, une estimation (peut-être très mauvaise) des T(q). On a vu que pour des vrais multifractals, T(q) n'était pas une fonction linéaire mais concave strictement. Une indication de multifractalité du processus sera donc donnée par une fonction $\xi(q)$ non linéaire. Ces comportement d'échelle peuvent être donc observés en théorie pour des processus fractals ne satisfaisant pas nécessairement une loi d'échelle au sens strict comme définie dans la section 2.2.5 (globale), mais possédant des exposants T(q) (lois d'échelle pour Δt tendant vers 0). C'est donc une indication de fractalité, et de multifractalité si

 $\xi(q)$ est non linéaire.

On peut voir que si l'on prend des intervalles de temps trop grands par rapport à N, les $\ln(S_q(\Delta t))$ s'éloigneront des droites. Cela peut être dû au fait que l'estimation de $\mathbb{E}[|r(\Delta t)|^q]$ n'est plus satisfaisante, ou bien que le n correspondant dans le raisonnement ci-dessus à $N/\Delta t$ est trop petit pour que l'on soit suffisamment proche de T(q).

Le problème pour pouvoir tirer des conclusions à partir d'observations de comportements d'échelle est que la vitesse de convergence des $S_q(\Delta t)$ vers $\mathbb{E}[|r(\Delta t)|^q]$ va dépendre du caractère à longue mémoire ou non des données (réelles ou simulées d'un modèle que l'on désire comparer aux données réelles). Le caractère à longue mémoire des returns absolus peut très fortement ralentir cette vitesse de convergence, de sorte que ce que l'on calcule n'a peut-être plus grand chose à voir avec T(q). En plus, il y a la limite dans la définition de T(q).

Question: tout ceci constitue-t-il une indication pertinente du caractère multifractal des données, ou de la présence ou non de lois d'échelle du type (2.10)? Cette question n'est pas du tout évidente, et porte pour le moment réellement à polémique dans la communauté scientifique. Le problème est réellement la qualité de l'estimation des T(q)par ces $\xi(q)$.

Mentionnons brièvement les deux articles [1, 6].

Dans [1], on montre qu'il est possible de calculer des estimations des moments à partir de simulations de processus pour lesquels on sait a priori qu'il n'y a pas de loi d'échelle du type (2.10), mais tels qu'en échelle bilogarithmique, la relation entre $S_q(\Delta t)$ et Δt est malgré tout très proche d'une droite. Ils montrent cela a partir d'un processus de Lévy gaussien inverse, et arrivent même à ajuster les paramètres du modèle pour reproduire le comportement d'échelle des données du taux USD-DEM dont ils disposaient.

Dans [6] est construit un modèle obéissant cette fois bien à une loi d'échelle avec des exposants linéaires (monofractal) mais sur des simulations du modèle, les exposants d'échelle calculés semblent former une fonction concave. Plus précisément, le processus des prix logarithmiques x(t) est donné par $x(t) = \sum_{k=1}^{N} r_k$, $N = t/\tau$ où τ est une échelle de temps très petite, et $r_k = \varepsilon_k |\sigma_k|$ avec ε_k i.i.d. de loi $\mathcal{N}(0,1)$ et les σ_k sont des variables aléatoires normales de moyenne nulle mais de fonction de corrélation non nulle, se comportant à l'infini comme une puissance comparable à ce que l'on observe pour les volatilités réalisées de séries financières (voir section 3.4.2). Par construction, tous les moments des returns sont finis, et on peut montrer analytiquement que le processus satisfait une loi d'échelle de type (2.10) avec un exposant d'échelle linéaire (processus monofractal). Cependant, en calculant les fonctions de structure, celles-ci se rangent bien le long de droites en échelle bilogarithmique, mais les exposants d'échelle $\xi(q)$ correspondant forment une fonction concave non linéaire, et ce malgré le relativement grand nombre de simulations du modèle (500 000 itérations). D'après eux, cette anomalie serait due au caractère à longue mémoire du processus des volatilités, caractère que l'on retrouve dans des données financières au niveau de la volatilité réalisée comme nous l'avons vu. Il faudrait probablement un nombre astronomique de simulations pour le faire disparaître, la même chose au niveau des données financières étant évidemment impossible en pratique.

La conclusion de tout ceci est qu'il faut manipuler ces concepts de lois et d'exposants d'échelle avec la plus grande prudence, et certainement ne pas baser une modélisation uniquement là-dessus, mais également sur d'autres critères.

3.4.4 Hétérogénéité du marché

Une hypothèse classique en économie est celle d'un marché homogène, où tous les investisseurs se comportent de la même façon, en investisseurs rationnels maximisant leur fonction d'utilité. Les investisseurs réagissent donc tous de la même façon aux événements.

L'hypothèse de marché hétérogène est une nouvelle idée en modélisation. Cette hypothèse suppose que les agents ont des perceptions du marché, des degrés d'information et des règles de comportement différents. Pour le marché des changes, il apparaît que les agents ont des profils de risques, des localisations géographiques et des contraintes institutionnelles différents. Il y a eu un certain nombre d'articles dans les années 1990 rejetant l'hypothèse traditionnelle de traders rationnels interprétant toutes les informations de la même manière et n'ayant aucune raison d'avoir des horizons de temps différents, et ce pas uniquement pour le marché des devises. Dans [30, 29] est soutenue la thèse que c'est l'horizon de temps sur lequel les investisseurs influencent le marché qui est l'aspect le plus important des différences de comportement et de perception du marché. Pour eux, les différences entre traders vont donc essentiellement se traduire par des horizons de temps différents. Ils tentent d'observer cela pour le marché des devises. L'idée est que les traders agissant à court terme évaluent le marché à de plus hautes fréquences et ont une plus courte mémoire que ceux travaillant à plus long terme. Une augmentation rapide de 0.5% suivie par une baisse rapide de 0.5% est un événement important pour un trader à court terme, mais pas pour une banque centrale ou un autre investisseur à long terme. Les traders à long terme s'intéressent quant à eux plus à la tendance générale sur de longs intervalles. Ils regardent donc le marché (et sa volatilité) sur un maillage plus grossier. Les traders à court terme et à long terme n'auront donc pas les mêmes ensembles d'opportunités, ils réagiront donc différemment à une même information.

Dans [29], les volatilités réalisées pour différents horizons de temps seront considérées comme la traduction des perceptions et des actions des différentes composantes du marché, des différents groupes de traders.

Dans [29], on s'intéresse à la corrélation croisée entre les volatilités réalisées pour des intervalles de temps différents (maillages plus ou moins fins). Le point est qu'ils remarquent une certaine asymétrie entre les deux fonctions d'autocorrélation, asymétrie d'après eux non due aux seuls effets statistiques. Ils observent plus précisément une asymétrie autour de $\rho_1 - \rho_{-1} = -0.1$ et $\rho_2 - \rho_{-2} = -0.1$ à -0.05 selon l'échantillon considéré (ils considèrent une dizaine d'échantillons, avec différentes devises), dans le sens que la volatilité à long terme semble plus influencer celle à court terme que le contraire. Pour eux, cela va dans le sens de l'hypothèse des marchés hétérogènes. Ils vont plus loin: ils interprètent cela comme un flux d'information des traders à long terme vers ceux à court terme (flux d'information à travers les échelles de temps). L'idée qu'il y a derrière est la suivante: les traders à très court terme sont les "faiseurs de marché", ce sont eux qui réagissent en premier lieu aux nouvelles informations. Dès qu'ils ont réagi, si ces informations sont suffisamment importantes, les changements de prix vont atteindre également les traders à court terme vont réagir à la réaction des traders à plus long terme, et cela se traduira au niveau des volatilités. D'après [10], cet effet de feedback pourrait expliquer la longue mémoire des volatilités des marchés financiers. D'après [10, 29], il faut en tenir compte pour de futures modélisations. Ils proposent d'ailleurs dans [29] un nouveau modèle de type ARCH: les modèles HARCH qui incorporent dans leur formulation mathématique l'hypothèse de marché hétérogène (voir section 4.1).

Chapitre 4 Modèle multifractal en cascade

Nous commençons par rappeler quelques notions élémentaires relatives aux modèles ARCH, qui constituent une classe importante de modèles de régression en finance. Le but de ce chapitre ne sera pas de discuter d'une éventuelle modélisation par un tel modèle. Nous les donnons juste à titre d'information complémentaire.

4.1 Rappel: Modèles ARCH

Dans les années 70 s'est développée la classe des modèles ARMA, (Auto-Regressive Moving Average). Ces modèles sont fondés sur une écriture de la valeur présente de la série temporelle comme fonction linéaire de ses valeurs passées et de la valeur présente d'un certain bruit. Ce type de modèle est assez restreint au niveau applications, notamment pour les problèmes financiers. Les modèles ARCH (AutoRegressive Conditional Heteroskedasticity) ont été introduits par Engle en 1982 ([14]). Dans ces modèles, ce seront les seconds moments qui varieront en fonction du temps suivant un certain modèle de régression. Cela correspond au fait que l'on s'est rendu compte que les moments d'ordre 2 de séries temporelles issues des marchés spéculatifs dépendent du temps, et les chercheurs en finance ont donc commencé à modéliser cette dépendance. Ces modèles peuvent capturer le phénomène d'accumulation des volatilités (volatility clustering) des données financières. Les modèles ARCH ont joué (et jouent encore) un rôle important dans la modélisation de la variation de ces seconds moments en fonction du temps. Cependant, ils ne rendent pas compte des comportements d'échelle.

Dans la définition qui suit, on suppose donné un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ muni d'une filtration $\{\mathcal{F}_t\}$. On considère un processus discret $\{r_t\}$ de la forme

$$r_t = \sigma_t \xi_t$$

où

$$\xi_t \text{ i.i.d. }, \mathbb{E}[\xi_t] = 0, \text{ Var}[\xi_t] = 1,$$

où de plus $\sigma_t \in \mathcal{F}_t$ et est indépendante de ξ_t . Le processus des r_t est univarié ou multivarié, mais nous ne présentons ici que le cas univarié. Dans la plupart des applications financières, r_t sera simplement le processus des returns.

Voici un aperçu de quelques modèles de type ARCH utilisés en finance.

1. Le modèle linéaire ARCH(q)

Il s'agit du premier modèle de type ARCH¹, dû à Engle ([14]). On suppose la paramétrisation suivante pour σ_t^2 :

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i r_{t-i}^2$$

où les paramètres satisfont $\omega > 0$ et $\alpha_i \ge 0$ pour tout *i*. On peut réduire le nombre de paramètres et imposer une influence décroissante linéairement en fonction du temps du passé (imposer que les événements plus lointains aient moins d'influence) en rajoutant la condition $\alpha_i = \alpha(q+1-i)/(q(q+1))$ (cf. [14]).

2. Le modèle linéaire GARCH(p,q) (=Generalized AutoRegressive Conditional Heteroskedasticity)

Cette généralisation du modèle ARCH a été introduite par Bollerslev en 1986 (cf. [4]) et fait intervenir directement dans la régression les variances passées, ce qui rend ce modèle toujours plus flexible que le précédent (si q est grand). Le modèle est défini par

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i \sigma_{t-i}^2.$$

C'est le modèle de type ARCH qui est utilisé de façon standard.

Ces deux modèles sont linéaires dans les seconds moments et univariés. Il y a à coté de cela un nombre impressionnant d'autres modèles de ce type (des versions non linéaires et/ou multivariées), cf. [5].

3. Modèle HARCH(n) (= Heterogeneous interval AutoRegressive Conditional Heteroskedasticity)

Ce modèle est introduit dans [29] pour rendre compte de l'hypothèse d'hétérogénéité des marchés. On suppose ici que

$$\sigma_t^2 = c_0 + \sum_{j=1}^n c_j \left(\sum_{i=1}^j r_{t-i}\right)^2$$

où $c_0 > 0, c_n > 0$ et $c_j \ge 0$ pour $j = 1, \ldots, n-1$. La somme $\sum_{i=1}^{j} r_{t-i}$ n'est rien d'autre que le return sur un intervalle de longueur j. On fait donc intervenir le

¹Quand on parle du modèle ARCH, on fait en fait référence à celui-ci.

carré des returns sur des intervalles de longueurs différentes. Ce qui différencie ce modèle des modèles ARCH ou GARCH, c'est donc de considérer des volatilités de changements de prix sur des horizons de temps différents. Exemple: HARCH(2):

$$\sigma_t^2 = c_0 + (c_1 + c_2)r_{t-1}^2 + c_2r_{t-2}^2 + 2c_2r_{t-1}r_{t-2},$$

donc on est comme dans un modèle ARCH(2) avec un terme en plus: le terme croisé. Ici, non seulement les returns absolus vont influencer σ_t^2 , mais également leur signe, contrairement aux deux modèles précédents. On peut encore citer le modèle EMA-HARCH (cf. [10]).

4.2 Modèles de volatilité stochastique

Dans le modèle de Black-Scholes, les returns logarithmiques (pour $\Delta t = 1$) vérifient $r_t = \mu + \sigma U_t$, avec μ et σ constants et U_t i.d.d. de loi normale standard. On peut généraliser ceci en supposant que σ n'est plus constante mais est lui-même un processus stochastique. On parle alors de modèle à volatilité stochastique. On peut supposer également la non normalité des U_t .

4.3 Le modèle en cascade

4.3.1 Analogie avec la turbulence

En couverture de son volume 381, la revue *Nature* utilise les termes: *Financial turbulence* et publie l'article de Ghashghaie–Breymann–Peinke–Talkner–Dodge ([19]), article proposant une analogie entre les phénomènes de turbulence et la dynamique des marchés spéculatifs, et plus précisément le marché des changes.

La turbulence hydrodynamique

L'évolution du champ des vitesses dans le temps et l'espace d'un fluide incompressible est décrite par les équations de Navier-Stokes (équation aux dérivées partielles d'évolution en la vitesse).

Une caractéristique importante d'un fluide en mouvement est son nombre de Reynolds R défini par UL/ν où ν est la viscosité du fluide, U sa vitesse moyenne et L lié à la géométrie du domaine où se fait l'écoulement.

Lorsque ce nombre est inférieur à une certaine valeur critique, le fluide va se mouvoir suivant un écoulement laminaire, que l'on peut bien étudier avec les équations de Navier-Stokes. Lorsque ce nombre augmente, il commence à apparaître des fluctuations importantes (tourbillons) dans le mouvement du fluide à des grandes échelles d'espace, et si Raugmente suffisamment, ces fluctuations deviennent instables, et se brisent pour donner naissance à de nouvelles à des échelles d'espace plus petites. Il s'agit du phénomène de la cascade de Richardson, un mécanisme qui permet de dissiper dans un fluide visqueux de grandes quantités d'énergie. C'est le fait d'avoir un flux d'énergie (cinétique) se transmettant à travers les échelles d'espace, des grandes vers les plus petites. En clair, l'énergie est introduite dans le système à une (grande) échelles l_0 (exemple: dans l'atmosphère par un avion), ensuite, elle descend dans la hiérarchie des tourbillons: les gros tourbillons donnent naissance à de plus petits, et ainsi de suite. A la fin de la cascade, l'énergie est chassée du système par dissipation (sous forme de chaleur), en bas de l'échelle d'espace. Les générations successives de tourbillons ont des échelles $l_n = l_0 r^n$, $n = 0, 1, 2, \ldots$, pour un certain $r \in (0, 1)$. C'est cette cascade qui donne lieu aux propriétés d'échelle pour les moments des différences de vitesse (entre deux points): la moyenne de la différence de vitesse entre deux points est proportionnelle à la distance entre ces deux points à une certaine puissance.

Finalement, le mouvement du fluide présentera des tourbillons à toutes les échelles d'espace. C'est ce que l'on appelle la turbulence développée. Pour étudier ce genre de phénomène, les méthodes statistiques semblent beaucoup plus adaptées que les déterministes (instabilité croissante avec R de Navier-Stokes).

La turbulence intervient dans beaucoup de situations: mouvements de l'atmosphère, prévisions météorologiques, formation des galaxies, écoulements autour des automobiles et avions, circulation sanguine, ...

Kolmogorov a beaucoup étudié ces phénomènes, et c'est à cette occasion qu'il a notamment développé la théorie des processus auto-similaires. Dans son modèle de 41 (voir [23]), une hypothèse importante est celle d'auto-similarité du champ des vitesses. On postule en fait dans ce modèle que les solutions des équations de Navier-Stokes ont la même invariance par changement d'échelle, mais dans le sens probabiliste, que les équations elles-mêmes. Dans ce modèle on trouve également la cascade de Richardson.

Dans les années 60 (cf. [24, 32]), Kolmogorov raffine sa théorie et intègre le phénomène d'intermittence de la turbulence, c'est-à-dire le fait que le fluide contienne des régions de très haute activité et d'autres de basse activité. On retrouve la cascade de Richardson, en supposant cette fois que la fraction de volume occupée par les tourbillons à chaque étape de la cascade décroît d'un facteur $\beta < 1$. On peut voir que dans ce cas, les exposants d'échelle sont non linéaires. C'est ici que les modèles multifractals apparaissent, ainsi que les cascades multiplicatives.

Pour un bon aperçu du sujet, on peut consulter le livre de Frisch [18].

Ressemblance entre turbulence et finance

Dans [19], Ghashghaie–Breymann–Peinke–Talkner–Dodgeest proposent une analogie entre l'évolution d'actifs financiers (taux de change) et la turbulence hydrodynamique. Cet article montre les similitudes dans le comportement statistique des données financières et des données de turbulence. Voir également [2]. En finance, on observe les prix d'actifs, en mécanique des fluides, on observe le champ des vitesses. Les analogies entre les deux phénomènes sont essentiellement les suivantes:

- queue lourde de la distribution des différences de vitesse et des returns, avec une tendance à tendre vers une normale si l'intervalle de temps (pour les returns) ou d'espace (pour les différences de vitesse) est de plus en plus grand;
- l'intermittence en turbulence développée, et la variation des volatilités réalisées dans le temps avec accumulation (volatility clustering);
- l'apparente longue mémoire des deux processus;
- la présence de lois d'échelle, avec des exposants d'échelle non linéaires (ou du moins qui semblent l'être pour les données financières);
- cascade d'énergie (cascade de tourbillons) dans le cas d'un fluide turbulent, flux d'information (ou de volatilités) à travers les échelles de temps, des grands vers les petits horizons de temps (cf. section 3.4.4 et plus bas).

L'observation de [29] selon laquelle la fonction d'autocorrélation croisée des volatilités réalisées est une fonction asymétrique est ici interprétée comme l'existence d'un flux d'information entre les traders actant sur différents horizons de temps, des horizons longs vers les courts.

On est donc tenté de modéliser les séries financières par le même genre de modèle, à condition de faire la correspondance:

 $\begin{array}{rcl} & \mbox{énergie} & \leftrightarrow & \mbox{volatilité réalisée} \\ & \mbox{distance dans l'espace} & \leftrightarrow & \mbox{intervalle de temps} \\ & \mbox{la vitesse en le point } x & \leftrightarrow & \mbox{le prix à l'instant } t \end{array}$

Tout ceci motive le modèle en cascade.

4.3.2 Le modèle

Le modèle en cascade a été introduit explicitement dans [7] par Breymann-Ghasghaie-Talkner (2000), mais basé sur l'analogie avec la turbulence hydrodynamique décrite dans la section précédente. Dans [7] est proposé le modèle en cascade suivant: on considère une cascade d'horizons de temps $\tau^{(1)} > \ldots > \tau^{(m)}$, sensés représenter les horizons des groupes de traders, où $\tau^{(1)}$ est typiquement de l'ordre d'une année et $\tau^{(m)}$ de l'ordre de quelques minutes. On note le return logarithmique à l'instant t par r_t . Plus précisément, on considérera $\tau^{(m)}$ comme un intervalle de temps élémentaire, et on s'intéressera à $r_t = r(t, \tau^{(m)})$, où en pratique $t = n\tau^{(m)}$ pour $n \in \mathbb{N}$.

On supposera que

où ξ_t sont des variables aléatoires i.i.d. ~ t_N , Student avec N degrés de liberté. Au vu des indices de queue observés sur ce genre de données financières (voir section 3.4.1), on s'intéressera typiquement au cas N = 3 ou 4. On assigne ensuite à chaque horizon de temps une volatilité $\sigma_t^{(k)}$ qui, pour rendre compte de l'observation de [29], sera déterminée par la volatilité au niveau k - 1 et par un facteur aléatoire dépendant du temps $a_t^{(k)}$:

$$\sigma_t^{(k)} = a_t^{(k)} \sigma_t^{(k-1)}$$

On arrive alors à une volatilité de la forme:

$$\sigma_t^{(m)} = \sigma_t = \sigma_0 \prod_{k=1}^m a_t^{(k)},$$

où l'on supposera que σ_0 est une constante et $\{a_t^{(k)}, k = 1, \ldots, m\}$ sont des variables aléatoires suivant un certain processus de renouvellement sensé traduire le flux d'information des grands vers les petits horizons de temps.

On supposera par simplicité que les horizons de temps satisfont:

$$\frac{\tau^{(k)}}{\tau^{(k-1)}} = p$$

pour une certaine constante p < 1.

On suppose qu'à l'instant initial t_0 , toutes les variables $a_t^{(k)}$ sont générées suivant des log-normales indépendantes $LN(a_k, \lambda_k^2)$ (où $a_k = \mathbb{E}[\log(a_{t_0}^{(k)})]$ et $\lambda_k^2 = \operatorname{Var}\left[\log(a_{t_0}^{(k)})\right]$.) Ensuite, pour passer de l'instant t_n à $t_{n+1} = t_n + \tau^{(m)}$, on suppose que

$$a_{t_{n+1}}^{(1)} = \begin{cases} a_{t_n}^{(1)} & \text{avec probabilité } 1 - \omega^{(1)} \\ \text{renouvelée} & \sim LN(a_1, \lambda_1^2) & \text{avec probabilité } \omega^{(1)} \end{cases}$$

pour une certaine probabilité de renouvellement $\omega^{(1)}$. En cas de renouvellement de $a_{t_{n+1}}^{(1)}$, on suppose que $a_{t_{n+1}}^{(k)}$ est aussi renouvelée pour tout k > 1 suivant des lois log-normales $LN(a_k, \lambda_k^2)$ indépendantes. Dans l'autre cas (pas de renouvellement), on suppose que

$$a_{t_{n+1}}^{(2)} = \begin{cases} a_{t_n}^{(2)} & \text{avec probabilité } 1 - \omega^{(2)} \\ \text{renouvelée} & \sim LN(a_2, \lambda_2^2) & \text{avec probabilité } \omega^{(2)} \end{cases}$$

et ainsi de suite. On peut voir que le processus ainsi défini est stationnaire. Les probabilités de renouvellement $\omega^{(k)}$ sont choisies de sorte que l'intervalle de temps moyen séparant deux renouvellements de $a_t^{(k)}$ est égal à l'horizon de temps $\tau^{(k)}$.

Par construction, le nombre de renouvellements de $a_t^{(1)}$ est un processus de Bernouilli (avec les unités de temps choisies égales à $\tau^{(m)}$) de probabilité $\omega^{(1)}$. Il est bien connu que dans un tel processus, les instants de succès (ici de renouvellement) T_k sont de loi binomiale négative et les intervalles entre deux succès consécutifs sont i.i.d. de loi

4.3. Le modèle en cascade

géométrique de paramètre $\omega^{(1)}$. On en déduit que l'intervalle de temps moyen entre deux renouvellements ($\mathbb{E}[T_k - T_{k-1}]$) est égal à $1/\omega^{(1)}$. Si l'on désire que cette espérance soit égale à $\tau^{(1)}/\tau^{(m)}$ unités de temps $\tau^{(m)}$, comme $\tau^{(1)}/\tau^{(m)} = p^{1-m}$, il faut imposer

$$\omega^{(1)} = p^{m-1}.$$

Considérons maintenant les renouvellements des $a_t^{(2)}$. Si l'on note $S_n^{(2)}$ le nombre de renouvellements après n étapes (après n intervalles de temps de longueur $\tau^{(m)}$), et si l'on note $X_n^{(1)}$ la variable aléatoire indicatrice d'un renouvellement à l'étape n pour les $a_t^{(1)}$, on a

$$\mathbb{P}[S_n^{(2)} - S_{n-1}^{(2)} = 1] = \mathbb{E}[\mathbb{P}[S_n^{(2)} - S_{n-1}^{(2)} = 1 \mid X_n^{(1)}]]$$

= $\omega^{(1)} + \omega^{(2)}(1 - \omega^{(1)})$
 $\stackrel{\text{not}}{=} \tilde{\omega}^{(2)}.$

Du fait que les renouvellements des $a_t^{(1)}$ se font de façon indépendante dans le temps, il en sera de même pour ceux des $a_t^{(2)}$. Les variables aléatoires $S_n^{(2)} - S_{n-1}^{(2)}$ sont donc des variables de Bernouilli i.i.d. de paramètre $\tilde{\omega}^{(2)}$, et le nombre de renouvellements en fonction du temps est à nouveau un processus de Bernouilli, de paramètre $\tilde{\omega}^{(2)}$. L'intervalle de temps moyen entre deux renouvellements est donc donné par $1/\tilde{\omega}^{(2)}$, et pour que ce soit égal à $\tau^{(2)}/\tau^{(m)} = p^{2-m}$, il faut donc que

$$\omega^{(1)} + \omega^{(2)}(1 - \omega^{(1)}) = p^{m-2},$$

c'est-à-dire que

$$\omega^{(2)} = \frac{p^{m-2} - p^{m-1}}{1 - p^{m-1}}.$$

On peut continuer à faire ce raisonnement pour les étapes suivantes de la cascade. De manière générale, la probabilité de renouvellement des $a_t^{(k)}$ à une étape sera donnée par

$$\tilde{\omega}^{(k)} = \tilde{\omega}^{(k-1)} + \omega^{(k)} (1 - \tilde{\omega}^{(k-1)}),$$

et l'on choisira de prendre $\tilde{\omega}^{(k)}=p^{m-k}=\tau^{(m)}/\tau^{(k)},$ ce qui nous donnera

$$\omega^{(k)} = \frac{p^{m-k} - p^{m-k+1}}{1 - p^{m-k+1}}, \quad k = 2, \dots, m.$$

On pourrait voir le modèle (les returns absolus ainsi que les volatilités) comme relié à une mesure multiplicative, mais plus générale que les mesures multinomiales introduites au Chapitre 2. Plus générale puisque $a_{t_n}^{(1)}$ n'est pas égal p.s. à $a_{t_{n+1}}^{(1)}$ mais seulement en loi. Comme l'intervalle moyen entre 2 renouvellements est $\tau^{(1)}$, on pourrait voir cela comme une mesure multinomiale "en moyenne", mais pas au sens strict. Elle sera également non conservative.

Évidemment, du fait que l'on s'arrête aux m premiers niveaux, il ne s'agit pas non plus d'une "vraie" mesure multinomiale (même en moyenne) puisque l'on s'arrête aux m premières étapes. En nous arrêtant à la mè étape, on ne fait en réalité que donner une approximation de la cascade "théorique". Mais si l'on veut simuler numériquement le modèle, on est contraint de s'arrêter à un certain moment. Définir la cascade de cette façon-ci (en négligeant l'équivalent du second terme de (2.17)) revient au même dans la pratique, au niveau du moins des simulations du modèle.

On pourrait essayer de calculer les exposants T(q) (pour le processus des prix logarithmiques). On a évidemment le problème que l'on a arrêté la cascade après les mpremières subdivisions. En fait, si l'on suppose que l'espérance des facteurs qui auraient dû se rajouter (l'espérance de la somme dans (2.17)) est bornée uniformément en m, on peut se limiter aux m premiers facteurs dans le calcul de T(q) pour l'expression de $\mathbb{E}[|r(0, \tau^{(m)})|^q].$

On a vu que l'on avait toujours, pour un processus croissant presque sûrement, que $T_{\alpha}(q) = T_h(q)$ pour tout q, ce qui impliquait que l'on pouvait éventuellement choisir d'autres bases que la base 2 dans la définition de $T_{\alpha}(q)$. On a alors

$$T_{\alpha}(q) = -1 + \lim_{m \to \infty} -\frac{1}{m} \log_{1/p} \mathbb{E}[|r(0, \tau^{(m)})|^q].$$

Calculons $\mathbb{E}[|r(0, \tau^{(m)})|^q]$. Du fait de l'indépendance à t fixé des facteurs $a_t^{(k)}$, $\prod_{k=1}^m a_t^{(k)}$ est de loi $LN(a, \Lambda^2)$ où $a := \sum_{k=1}^m a_k$ et

$$\Lambda^2 := \sum_{k=1}^m \lambda_k^2. \tag{4.1}$$

Or, si une variable aléatoire X est de loi log-normale $LN(\mu, \sigma^2)$, alors $\mathbb{E}[|X|^q] = \mathbb{E}[e^{qY}]$ où $Y \sim \mathcal{N}(\mu, \sigma^2)$, ce qui vaut $e^{\frac{\sigma^2 q^2}{2} + \mu q}$. Donc

$$\mathbb{E}\left[\left(\prod_{k=1}^{m} a_t^{(k)}\right)^q\right] = e^{\frac{\Lambda^2 q^2}{2} + aq}.$$

Si l'on note maintenant $\mathbb{E}[|\xi_t|^q]$ par f(q) (fini ssi q < N, le nombre de degrés de liberté de la Student), on arrive finalement à

$$\ln \mathbb{E}[|r(0,\tau^{(m)})|^q] = \frac{\Lambda^2 q^2}{2} + aq + \ln(f(q)) + c_0$$

pour une certaine constante c_0 dépendant de σ_0 , et donc

$$T(q) = -1 + \lim_{m \to \infty} -\frac{1}{m} \log_{1/p}(e) \left(\frac{\Lambda^2 q^2}{2} + aq + \ln(f(q)) + c_0\right)$$
$$= -1 + \lim_{m \to \infty} -\frac{1}{m} \log_{1/p}(e) \left(\frac{\Lambda^2 q^2}{2} + aq\right)$$

où l'on a en fait $\Lambda^2 = \Lambda^2(m)$ et a = a(m). Au cas où $\lambda_k = \lambda$ pour tout k et $a_k = \alpha$, on arrive alors à

$$T(q) = -1 + \log_{1/p}(e) \left(\frac{\lambda^2 q^2}{2} + \alpha q\right),$$

une fonction du second degré en q. Cependant, il faut être prudent, car ici, je suppose que les ξ_t sont toujours de loi Student à un nombre fixé de degrés de liberté, quel que soit l'intervalle infinitésimal τ^m considéré.

4.3.3 Le modèle MMAR de Mandelbrot–Fisher–Calvet

Le modèle MMAR (Multifractal Model of Asset Returns) est décrit dans [27, 8, 17]. Dans ce modèle, on suppose que le processus de prix logarithmiques $X(t) = \ln P(t) - \ln P(0)$ est donné par

$$X(t) = B_H(\theta(t))$$

où $B_H(t)$ est un mouvement brownien fractionnaire d'indice H et $\theta(t)$ un processus stochastique croissant positif (représentant un temps de marché). On suppose que $\theta(t)$ est la fonction de répartition $\mu[0, t]$ d'une mesure μ multifractale (au sens de Mandelbrot) définie sur un intervalle [0, T]. On suppose également l'indépendance entre $B_H(t)$ et $\theta(t)$.

On peut voir (cf. [27]) que sous ces hypothèses, le processus X(t) est multifractal au sens de Mandelbrot (voir section 2.2.5) avec exposant d'échelle $\tau_X(q) = \tau_{\theta}(Hq)$, et est à accroissements stationnaires.

On choisit alors μ de façon à avoir ou non une queue lourde, et on choisit H pour ajuster le type de mémoire (longue ou non) du processus. Si par exemple, on veut des accroissements indépendants et stationnaires, on prendra H = 1/2. On peut voir qu'alors, le processus X devient une martingale (voir [27]).

La différence essentielle entre cette approche-ci et le modèle en cascade est que dans le MMAR, on a un temps de marché multifractal, que l'on compose avec un processus monofractal, alors que dans l'approche décrite dans les sections précédentes, on fait d'abord une désaisonnalisation des données, puis on se place en temps de marché pour modéliser. Et on utilise alors non un modèle monofractal mais multifractal. On traite donc séparément les problèmes de désaisonnalisation et de modélisation en temps de marché. Fondamentalement, cela devrait revenir au même, en désaisonnalisant. On relie d'abord le temps de marché à la volatilité réalisée, celle-ci étant quelque part liée à σ_t . Ensuite, on modélise la volatilité en temps de marché par un multifractal. Mais ce caractère multifractal doit encore être présent dans la volatilité en temps physique, suggérant que notre temps de marché, s'il devait être modélisé, le serait aussi.

Ce modèle a été introduit en parallèle et indépendamment du modèle multifractal en cascade.

4.4 Résultats sur simulations du modèle

Nous donnons dans les pages qui suivent un aperçu d'une partie expérimentale tentant de confronter le modèle aux données. Cette partie n'est pas entièrement développée (par rapport à ce que j'ai fait) pour la simple raison que beaucoup de travail doit encore être accompli, et qu'une estimation rigoureuse des paramètres n'est pas pour bientôt, vu l'apparente difficulté du problème. Nous donnerons donc une idée de ce qui a été accompli.

Wolfgang Breymann a écrit un programme de simulation du modèle en C++. Ce programme a été appelé depuis le logiciel statistique S-plus pour pouvoir ensuite traiter les données simulées et les comparer aux données réelles. La valeur du paramètre pchoisie était $p = 1/\sqrt{2}$. Il n'y a a priori pas de raison de choisir cette valeur plutôt qu'une autre, mais il fallait bien supposer quelque chose pour le premier essai. Il est vrai que l'on aurait pu prendre 1/p entier, pour se rattacher totalement aux mesures multinomiales décrites dans la première partie. On a ensuite choisi le nombre de degrés de liberté de la Student ξ_t égal à 3 ou 4 (au vu des résultats concernant les indices de queue de la section 3.4.1).

Pour ce qui est des autres valeurs des paramètres, on ne fera varier que Λ^2 et σ_0 .

Dans les simulations, on a choisi de prendre $\tau^{(m)}$ de l'ordre de 1/2 heure, et $\tau^{(1)}$ de l'ordre d'une année. On décide alors de prendre seulement 1 point sur 2 de la série simulée (pour avoir des données horaires). 200 000 itérations correspondent donc à une série de prix horaires de 100 000 points. On peut voir que dans ce cas, m est aux alentours de 30 (28 exactement).

Les paramètres λ_k ont été générés de la manière suivante. Si on note $\gamma := -\log(p)$, on suppose que

$$\lambda_k^2 = \gamma c_1 + \gamma^2 k c_2.$$

Cela donne comme variance totale pour le logarithme de σ_t :

$$\Lambda^{2} = \sum_{i=0}^{m} \lambda_{i}^{2} = \gamma m c_{1} + \frac{\gamma^{2}}{2} (m^{2} - m) c_{2}.$$

Si en outre, on suppose (au contraire de ce qui est donné dans l'introduction du modèle) que σ_0 est aléatoire, il faut rajouter également un terme c_0 dans l'expression qui précède. Ici, on a choisi $c_0 = 0$.

Voici une illustration d'une série de returns simulés suivant le modèle (figure 4.1).

4.4.1 Une première impression

Pour donner une première idée de l'adéquation du modèle avec les données, nous donnons ci-dessous quelques figures illustrant



Figure 4.1: Returns simulés suivant le modèle.

- la fonction d'autocorrélation des returns absolus, comparaison modèle données
- les fonctions de densité (logarithmiques), comparaison modèle données
- les exposants $\xi(q)$ calculés sur les données comparés à ceux calculés à partir de séries simulées du modèle

Ces figures fournissent des comparaisons visuelles entre le modèle et les données.

De façon générale, beaucoup de problèmes sont survenus par rapport aux exposants d'échelle $\xi(q)$. Ils variaient très fort d'une simulation à l'autre, même si la longueur de la série simulée était grande (de l'ordre de 300 000 ou 400 000 itérations). Une idée a alors été de simuler un certain nombre de fois le modèle, de faire des moyennes sur les différents $\xi(q)$ obtenus pour chaque simulation, suivant un ensemble prédéfini de random seeds, puis de simuler à nouveau plusieurs fois le modèle avec d'autres jeux de paramètres mais toujours suivant le même ensemble de random seeds. On a fait cela également pour la fonction d'autocorrélation des returns absolus et pour les fonctions de densité (qui cependant étaient plus stables que les exposants $\xi(q)$).

Nous donnons ci-dessous quelques illustrations comparant les deux modèles.

La figure 4.2 donne la fonction d'autocorrélation pour les données et des séries simulées du modèle (200 000 itérations par simulation, suivant le même jeu de paramètres, dont 3 degrés de liberté pour la Student). On voit que l'allure générale fluctue d'une simulation à l'autre mais raisonnablement. On a surtout une fonction d'autocorrélation pas très "lisse". Ce n'est d'ailleurs pas le cas non plus pour les données évidemment, mais dans l'idée de comparer plus quantitativement les deux, il semblait bon de lisser un peu celle du modèle simulé. C'est pourquoi, dans l'essai d'ajustement des paramètres (voir plus loin), on a considéré des moyennes de cette fonction sur différentes simulations. Il est facile de discerner sur les graphiques la fonction d'autocorrélation des données simulées et celle des données réelles car cette dernière présente encore des faibles pics toutes les 24 h (du moins pour des petits délais).



Figure 4.2: Fonction d'autocorrélation des returns absolus: Comparaison modèle données sur différentes simulations (random seeds non contrôlés, même jeu de paramètres dont 3 degrés de liberté pour la loi de Student, $c_1 = 0.05, c_2 = -0.002$). L'unité de temps choisie est 1 heure. La fonction d'autocorrélation des données réelles est celle qui présente encore des pics toutes les 24 h.

La figure 4.3 donne cette même fonction d'autocorrélation pour les données et le modèle en faisant varier le nombre de degrés de la Student de 2 à 5 (autres paramètres inchangés) et avec le même random seed pour chaque simulation.

La figure 4.4 est identique à la figure 4.3, mais avec un autre random seed.

Le modèle semble donc assez bien reproduire la lente décroissance de la fonction d'autocorrélation des returns absolus. Un nombre de 3 degrés de liberté semblent le mieux convenir pour ces valeurs de Λ^2 .



Figure 4.3: Fonction d'autocorrélation des returns absolus: Comparaison modèle - données, random seed fixé, nombre de degrés de la Student =2 à 5 (de gauche à droite et de haut en bas), autres paramètres fixés



Figure 4.4: Fonction d'autocorrélation des returns absolus: Comparaison modèle données, random seed fixé (mais différent de celui de la figure 4.3), nombre de degrés de la Student = 2 à 5 (de gauche à droite et de haut en bas), autres paramètres fixés

On peut voir que les returns eux-mêmes restent bien quasiment non autocorrélés comme cela découle du modèle(voir figure 4.5).



Figure 4.5: Fonction d'autocorrélation des returns pour le modèle simulé: non autocorrélation des returns, à comparer à la figure 3.12.

La figure 4.6 compare les fonctions de densité logarithmiques des returns (centrés réduits) des données et de simulations du modèle suivant le même jeu de paramètres à chaque fois dont 3 degrés de liberté pour la Student. D'après ces graphes, le modèle a l'air de déjà assez bien reproduire les données. Dans ce qui suit, on va faire varier certains paramètres pour avoir une idée plus fine de ce qui conviendrait le mieux comme jeu de paramètres. En réalité, les valeurs de Λ^2 que nous utilisons ici ont déjà fait l'objet d'un premier ajustement, expliquant les assez bons résultats de la figure 4.6. On peut voir que si on s'écarte significativement de cette valeur (voir figure 4.12), les résultats sont moins bons.

La figure 4.7 compare encore les fonctions de densité logarithmique des returns (centrés réduits) des données et de simulations du modèle suivant le même jeu de paramètres à chaque fois mais avec 4 degrés de liberté pour la Student et $c_1 = 0.05, c_2 = -0.002$.

La figure 4.8 est l'équivalent des deux précédentes, avec cette fois 5 degrés de liberté pour la Student.

La figure 4.9 montre les fonctions de densité logarithmiques pour les données et des simulations du modèle, en faisant varier uniquement le degré de liberté de la Student (de 2 à 6), en fixant le random seed d'une simulation à l'autre.

Les figures 4.10, 4.11, 4.12 comparent à nouveau les fonctions de densité logarithmiques, mais en faisant varier cette fois le paramètre Λ^2 (ou plutôt c_1 lié à Λ^2 , de sorte que si c_1 augmente, Λ^2 augmente). Les random seeds ont été fixés dans ces différentes figures, et égaux à celui de la figure 4.9. On voit que l'indice de queue semble augmenter



Figure 4.6: Returns: Fonction de densité logarithmique: Comparaison modèle - données (en pointillé), même paramètres (dont nombre de degrés de la Student = 3, $c_1 = 0.05, c_2 = -0.002$), random seeds différents

lorsque l'on augmente Λ^2 .

La figure 4.13 compare des qqplots des deux jeux de données (simulées et réelles) (avec le même random seed). En abcisse figure le modèle et en ordonnée les données réelles. Dans le cas de 3 degrés de libertés, la forme légèrement en S suggère qu'un nombre de 3 degrés est peut-être trop faible.

A première vue, au niveau de la densité logarithmique et de l'autocorrélation des returns absolus, 3 ou 4 degrés de liberté pour la Student et le paramètre Λ^2 fixé tel quel donne un ajustement plus ou moins satisfaisant. Mais au niveau de la forme de la densité, 4 degrés ont l'air de mieux convenir. Remarquons que le paramètre d'échelle σ_0 n'influence ni la fonction d'autocorrélation des returns absolus, ni la densité des returns centrés réduits.

Ce modèle permet donc de reproduire les queues lourdes de la distribution des returns grâce à la présence de la loi de Student, et semble assez bon également loin des queues.


Figure 4.7: Returns: Fonction de densité logarithmique: Comparaison modèle - données (en pointillé), même paramètres (deg Student = 4, $c_1 = 0.05, c_2 = -0.002$), random seeds différents



Figure 4.8: Returns: Fonction de densité logarithmique: Comparaison modèle - données (en pointillé), même paramètres (deg Student = 5), $c_1 = 0.05, c_2 = -0.002$, random seeds différents



Figure 4.9: Returns: Fonction de densité logarithmique: Comparaison modèle - données (en pointillé), $c_1 = 0.05, c_2 = -0.002$, nbre degrés Student = 2 à 6 (de gauche à droite et de haut en bas), random seed fixé



Figure 4.10: Returns: Fonction de densité logarithmique: Comparaison modèle - données (en pointillé), $c_1 = 0.035$ à 0.065 (de gauche à droite et de haut en bas), 3 degrés de liberté, $c_2 = -0.002$



Figure 4.11: Returns: Fonction de densité logarithmique: Comparaison modèle - données (en pointillé), $c_1 = 0.035$ à 0.065, nombre de degrés de la Student = 4, $c_2 = -0.002$



Figure 4.12: Returns: Fonction de densité logarithmique: Comparaison modèle - données (en pointillé), $c_1 = 0.1$ à 0.22, nombre de degrés de la Student = 4, $c_2 = -0.002$



Figure 4.13: Returns: qqplot des returns centrés réduits: Comparaison modèle - données, nombre de degrés de la Student = 2 à 6 (de gauche à droite et de haut en bas), $c_1 = 0.05, c_2 = -0.002$.

Considérons maintenant les comportements d'échelle. Les figures 4.14 et 4.15 montrent les exposants d'échelle calculés sur des données simulées selon le modèle et les données réelles. Pour le modèle, les valeurs de $\ln(S_q(\Delta t))$ se placent bien le long de droites en fonction de $\ln(\Delta t)$ (voir en illustration les figures 4.16, 4.17, 4.18).



Figure 4.14: Exposants d'échelle: $\xi(q)$ en fonction de q, comparaison modèle – données (courbe du dessus), mêmes paramètres (dont degré Student = 3, $c_1 = 0.05, c_2 = -0.002$), random seeds différents

On peut constater que les exposants $\xi(q)$ fluctuent assez fort d'une simulation à l'autre. C'est la raison pour laquelle on a décidé de fixer les random seeds, et de faire des moyennes pour le calcul des $\xi(q)$ (voir plus loin). En fait, dans le cas de 3 degrés de liberté, cela n'a pas de sens de considérer les $\xi(q)$ pour $q \geq 3$, puisque les moments correspondants n'existent pas $(\ln(S_q(\Delta t)) \to \infty \sin N/\Delta t \text{ tend vers } \infty, \text{ mais probablement}$ assez lentement puisqu'il s'agit d'un logarithme). C'est précisément pour ces valeurs-là de q que $\xi(q)$ varie très fort d'une simulation à l'autre. Le cas de 4 degrés de liberté semble à première vue assez bon par contre par rapport aux données, mais il faut malgré tout rester prudent si l'on veut en tirer des conclusions au vu des problèmes rencontrés par rapport à ces comportements d'échelle (voir section 3.4.3).



Figure 4.15: Exposants d'échelle: Comparaison modèle – données, mêmes paramètres (dont nombre de degrés de liberté de la Student = 4, $c_1 = 0.05, c_2 = -0.002$), random seeds différents

dof= 3, nr iter= 200002 , q = 1 --> 4 (haut --> bas)



Figure 4.16: Exposants d'échelle: log $S_q(\Delta t)$ en fonction de log Δt , nombre de degrés de la Student = 3, $c_1 = 0.05, c_2 = -0.002$



Figure 4.17: Exposants d'échelle: $\log S_q(\Delta t)$ en fonction de $\log \Delta t$, nombre de degrés de la Student = 4, $c_1 = 0.05, c_2 = -0.002$



Figure 4.18: Exposants d'échelle: $\log S_q(\Delta t)$ en fonction de $\log \Delta t$, nombre de degrés de la Student = 5, $c_1 = 0.05, c_2 = -0.002$

4.4.2 Fonction de vraisemblance

L'idée maintenant est d'essayer de quantifier un peu plus les comparaisons entre le modèle et les données. Une idée naturelle serait de faire une estimation des (ou de certains) paramètres par maximum de vraisemblance. Dans ce qui suit, nous cherchons à quoi ressemble la fonction de vraisemblance pour ce modèle.

Si la taille de la série simulée est N, la fonction de vraisemblance s'écrit

$$f(r_1, \sigma_1, r_2, \sigma_2, \dots, r_t, \sigma_t, \dots, r_N, \sigma_N) = f(r_1, \sigma_1) \cdot f(r_2, \sigma_2 | r_1, \sigma_1)$$

$$\cdots f(r_t, \sigma_t | \{r_{t'}, \sigma_{t'}; t' < t\}) \cdots f(r_N, \sigma_N | \{r_{t'}, \sigma_{t'}; t' < N\})$$

Il vient d'abord

$$f(r_t | \{r_{t'}, \sigma_{t'}; t' < t\}, \sigma_t) = g_{\sigma_t}(r_t)$$
(4.2)

où $g_{\sigma}(x)$ est la fonction de densité d'une variable aléatoire X telle que $X = \sigma Y$ où Y est de loi Student (avec le nombre de degrés de liberté qui convient, un des paramètres du modèle). Cherchons ensuite $f(\sigma_t | \{r_{t'}, \sigma_{t'}; t' < t\})$. On voit facilement que la probabilité pour qu'à l'instant t, il n'y ait pas de renouvellement à l'échelon 2 de la cascade est donnée par

$$1 - \tilde{\omega}^{(2)} = (1 - \omega^{(2)})(1 - \omega^{(1)}).$$

Ceci est en fait la probabilité pour qu'il n'y ait pas de renouvellement jusqu'à l'échelon 2. De manière générale, on peut voir que la probabilité pour qu'il n'y ait pas de renouvellement jusqu'à l'échelon k vaut

$$\prod_{i=1}^{k} (1 - \omega^{(i)}).$$

On peut donc écrire

$$f(\sigma_t | \{r_{t'}, \sigma_{t'}; t' < t\}) = \delta_{\sigma_{t-1}} \prod_{i=1}^m (1 - \omega^{(i)}) + \prod_{i=1}^{m-1} (1 - \omega^{(i)}) \omega^{(m)} d_m(a_t^{(m)}) f(\prod_{i=1}^{m-1} a_t^{(i)}) | \{r_{t'}, \sigma_{t'}; t' < t\}) + \prod_{i=1}^{m-2} (1 - \omega^{(i)}) \omega^{(m-1)} d_{m-1}(a_t^{(m-1)} a_t^{(m)}) f(\prod_{i=1}^{m-2} a_t^{(i)}) | \{r_{t'}, \sigma_{t'}; t' < t\}) + \ldots + \omega^{(1)} d_1(\prod_{i=1}^m a_t^{(i)})$$

$$(4.3)$$

où $d_k(x)$ $(1 \le k \le m)$ est la fonction de densité évaluée en x de σ_0 fois une log-normale $LN(a_m + \ldots + a_k, \lambda_m^2 + \ldots + \lambda_k^2)$, et $f(\prod_{i=1}^{m-1} a_t^{(i)} | \{r_{t'}, \sigma_{t'}; t' < t\})$ désigne la fonction de densité de $\prod_{i=1}^{m-1} a_t^{(i)}$ conditionnellement aux valeurs de $\{r_{t'}, \sigma_{t'}; t' < t\}$ en $\prod_{i=1}^{m-1} a_t^{(i)}$ (abus de notation). Il faut ensuite faire le produit de ces deux expressions (4.2), (4.3), la deuxième comportant en pratique une trentaine de termes. On fait ensuite le produit des N facteurs ainsi obtenus. Ceci semble a première vue assez compliqué à faire analytiquement, et le calcul numérique de la fonction de vraisemblance sur un échantillon pour un jeu de paramètre n'a pas non plus l'air évident à implémenter dans des temps raisonnables.

Pour cette raison, il semblait à première vue plus simple de faire appel à des simulations du modèle et de calculer certaines statistiques sur base de ces simulations et des données réelles. Tout le problème est de savoir quelle statistique considérer. Or, tout ceci va au delà des modèles considérés jusqu'à présents, des méthodes statistiques pour estimer les paramètres n'ont pas encore été développées. On est donc pour le moment dans une phase d'exploration du problème, et on a choisi de construire des fonctions de "distance" entre les données réelles et simulées. Ceci est l'objet des sections suivantes. Remarquons que tout ceci n'a jamais été réalisé pour ce modèle, et est encore à un stade très empirique.

4.4.3 Approche plus quantitative

J'ai écrit une fonction S-plus comparant la fonction d'autocorrélation des returns absolus pour les données ainsi que pour des données simulées. On simule N séries temporelles avec chacune 200 000 itérations (100 000 points), on calcule la fonction d'autocorrélation de chaque série, sa moyenne sur les N simulations ainsi que son écart-type (cela revient à lisser la fonction d'autocorrélation). On prend ensuite la différence absolue entre la fonction d'autocorrélation des données et la moyenne obtenue pour les simulations, puis on intègre le tout (on somme les valeurs de la fonction). On calcule également son intégrale pondérée par les inverses des écarts-types, ou par ces inverses multipliés par la fonction 1/x ou une autre fonction décroissante (car il semblait que la queue de la fonction d'autocorrélation était plus difficile à lisser). Il s'agit en fait de distances dans L^1 entre les fonctions d'autocorrélation.

Pour donner une idée, voici ce que l'on obtient lorsque l'on fait uniquement varier le degré de liberté de la Student et que N = 10 (random seeds non contrôlés, Λ^2 fixé à travers c_1, c_2 , égaux à 0.05 et -0.002 respectivement):

	Intégrale de
	$ ACF_d - ACF_s / \sigma(ACF_s)$
2 degrés	
0.021518898	3.005267147
0.022781824	4.427369818
0.019223282	2.743260278
0.021640184	3.932198661
0.022653440	3.856411183
3 degrés	
0.006627878	0.406286930
0.011639741	0.641786662
0.010581950	0.844828828
0.005901922	0.317886545
4 degrés	
0.015072756	1.283765537
0.02012434	1.10990500
0.027430249	1.199908343
0.018572655	1.313744595
5 degrés	
0.018286709	1.252879476
0.020403325	1.299791302
0.029405089	2.499487940
0.028463259	2.413913984
6 degrés	
0.024985785	1.443276670
0.045497077	1.928952715
0.023932331	1.641971497
0.025285574	1.760893244

où:

 ACF_d : fonction d'autocorrélation des données, ACF_s : fonction d'autocorrélation moyenne sur les N simulations, $\sigma(ACF_s)$: (vecteur des) écarts-types sur les N simulations de la fonction d'autocorrélation.

A première vue, selon ce critère-ci, et pour ces valeurs de Λ^2 , 3 degrés de liberté ont l'air de mieux convenir. On peut voir que le paramètre σ_0 n'influence pas cette fonctionci (puisque c'est un paramètre d'échelle).

Voici maintenant la même fonction où le random seed est fixé entre les simulations correspondant à des paramètres différents (20 simulations de 200000 itérations à chaque fois).

Intégrale de	Intégrale de	Intégrale de
$ ACF_d - \overline{ACF_s} $	$\frac{ ACF_d - \overline{ACF_s} }{\sigma(ACF_s)}$	$\frac{ ACF_d - \overline{ACF_s} }{\sigma(ACF_s)} * \frac{1}{x}$
3 degrés de liberté, $v_0 = 0.075$,	$c_0 = 0, c_1 = 0.05, c_2 = -0.002$	
0.006461732	0.647529493	0.008614758
4 degrés de liberté		
0.01693444	1.54198099	0.02132949
5 degrés de liberté		
0.021260117	1.690645536	0.023140281
6 degrés de liberté		
0.02399745	1.86285459	0.02649899
$\sigma_0 = 0.1$		
0.006461732	0.647529557	
$c_1 = 0.1$		
0.023227944	1.510953712	
$c_1 = 0.001$		
0.034467713	11.545447543	
$c_1 = 0.02$		
0.027461458	7.942332594	0.076329128
$c_1 = 0.03$		
0.01325220	2.33025044	0.02011893
$c_1 = 0.04$		
0.0048227105	0.6116508475	0.0045178597
$c_1 = 0.06$	1 01 00 1000	0.010011005
0.011681382	1.016040097	0.012811285
$c_2 = -0.001$		
0.012375942	1.080478993	0.012428082
$c_2 = -0.0015$		
0.009273773	0.863384715	
$c_2 = -0.0025$		
0.0046880591	0.5151766678	0.0067202726
$c_2 = -0.003$		
0.0048364897	0.6007170753	

Allure de la fonction $\int (|ACF_d - \overline{ACF_s}| / \sigma(ACF_s)) * 1/x$

Calcul pour $(c_1,c_2) \in [0.01,0.06] \times [-0.0035,-0.001],$ 3 degrés de liberté pour la Student:

	$c_1 = 0.01$	$c_1 = 0.015$	$c_1 = 0.02$	$c_1 = 0.025$	$c_1 = 0.03$	$c_1 = 0.035$
$c_2 = -0.0035$	0.1477411	0.14384629	0.13005424	0.10585496	0.077584586	0.046855240
$c_2 = -0.0030$	0.1472998	0.13993079	0.11886500	0.08885968	0.054971775	0.029064552
$c_2 = -0.0025$	0.1459711	0.13203653	0.10202604	0.06467080	0.034328167	0.016690253
$c_2 = -0.0020$	0.1427648	0.11699425	0.07632913	0.04055422	0.020118930	0.008581741
$c_2 = -0.0015$	0.1326874	0.09034741	0.04797976	0.02409746	0.010657561	0.004762115
$c_2 = -0.0010$	0.1070733	0.05691208	0.02873154	0.01319591	0.005498531	$\underline{0.004512118}$
	$c_1 = 0.04$	$c_1 = 0.045$	$c_1 = 0.05$	$c_1 = 0.0$	$c_1 = 0.$	06
$c_2 = -0.0035$	0.024584321	0.0113432	61 0.005612	116 0.00611	4393 0.00787	75120
$c_2 = -0.0030$	0.013770687	0.0060219	98 0.005434	774 0.00733	64272 0.00950	08164
$c_2 = -0.0025$	0.007011649	0.00472738	87 0.006720	273 0.00908	0.01124	41127
$c_2 = -0.0020$	0.004517860	0.0060473	65 0.008614	758 0.01096	0.01281	1285
$c_2 = -0.0015$	0.005288213	0.0080758	03 0.010630	062 0.01264	3187 0.01416	61504
$c_2 = -0.0010$	0.007454605	0.0102331	94 0.012428	082 0.01408	0.01531	15955



Figure 4.19: Valeurs de $\int (|ACF_d - \overline{ACF_s}| / \sigma(ACF_s)) * 1/x$, $(c_1, c_2) \in [0.01, 0.06] \times [-0.0035, -0.001]$. Les différents graphes en 3D sont des graphes de la même fonction vus sous des angles différents

On peut encore pondérer de différentes manières la différence absolue entre les fonctions d'autocorrélation. Cela donne des résultats similaires à ceci.

4.4.4 Lois d'échelle

Comme déjà annoncé, comparer les comportements d'échelle pour le modèle et les données n'est pas trivial. J'ai d'abord écrit une fonction calculant la distance quadratique entre les exposants d'échelle du modèle simulé et ceux des données. Plus précisément, cette fonction calcule:

$$d := \left(\sum_{i=2}^{8} (\overline{\tau_s}(i/2) - \tau_d(i/2))^2\right)^{1/2}$$

<u>où</u> $\overline{\tau_s}(q)$ est l'exposant d'échelle obtenu en faisant une régression linéaire pondérée de $\log S_q(\Delta t)$ sur $\log(\Delta t)$, avec

$$\overline{\log S_q(\Delta t)} := \frac{1}{N} \sum_{i=1}^N \log S_q^{(i)}(\Delta t)$$

avec $S_q^{(i)}(\Delta t)$ la fonction de structure calculée à la *i*ème simulation. La régression linéaire est pondérée par les inverses des variances des $S_q^{(i)}(\Delta t)$ sur les différentes simulations.

Pour donner une idée de la variabilité de cette distance d en fonction des simulations, voici quelques résultats où l'on a fait tourner 10 fois la fonction, avec pour chaque tours 15 simulations de 200000 itérations chacune (3 degrés de liberté pour la Student et les autres paramètres fixés comme plus haut). On s'intéresse également à la variation de chaque distance individuelle

$$d_q := |\overline{\tau_s}(q) - \tau_d(q)|.$$

Essai 1: 10 tours de scaling.3, N =15 simulations, niter =200000 itérations / simulation.

	d	d_1	$d_{1.5}$	d_2	$d_{2.5}$	d_3	$d_{3.5}$	d_4
Moyenne	0.3548	0.0112	0.0063	0.0134	0.0541	0.1150	0.1897	0.2708
Ecart-type	0.0750	0.0024	0.0032	0.0057	0.0140	0.0262	0.0407	0.0558

Essai 2: 10 tours de scaling.3, N = 15 simulations, 200000 itérations / simulation.

	d	d_1	$d_{1.5}$	d_2	$d_{2.5}$	d_3	$d_{3.5}$	d_4
Moyenne	0.3436	0.0117	0.0065	0.0123	0.0522	0.1118	0.1840	0.2618
Ecart-type	0.0492	0.0023	0.0037	0.0050	0.0086	0.0159	0.0261	0.0383

Comme on pouvait s'y attendre, les résultats sont très mauvais pour d_4 , ce qui est normal puisque le nombre de degrés de la Student est égal à 3 et que les moments d'ordre 3 de la Student n'existent pas. En fait, d_4 est assez importante, mais son écart-type est relativement faible par rapport à sa moyenne. On constate d'autre part que $\xi(1.5)$ est déjà assez bien ajusté. On peut choisir de se limiter à $q \leq 3$. Juste à titre d'illustration supplémentaire:

Essai 1: 10 tours de scaling.3, N = 15 simulations, 200000 itérations / simulation.

	$\mid d$	d_1	$d_{1.5}$	d_2	$d_{2.5}$	d_3
Moyenne	0.1254	0.0110	0.0057	0.0138	0.0531	0.1119
Ecart-type	0.0158	0.0024	0.0028	0.0046	0.0071	0.0144

Essai 2: 10 tours de scaling.3, N = 20 simulations, 300000 itérations / simulation.

	d	d_1	$d_{1.5}$	d_2	$d_{2.5}$	d_3
Moyenne	0.1183	0.0119	0.0071	0.0106	0.0484	0.1064
Ecart-type	0.0078	0.0022	0.0031	0.0040	0.0046	0.0066

Essai 3: 10 tours de scaling.3, N = 25 simulations, 300000 itérations / simulation.

	d	d_1	$d_{1.5}$	d_2	$d_{2.5}$	d_3
Moyenne	0.1132	0.0127	0.0090	0.0079	0.0449	0.1024
Ecart- type	0.0100	0.0013	0.0025	0.0034	0.0049	0.0091

Essai 4: 10 tours de scaling.3, N = 30 simulations, 200 000 itérations / simulation.

	d	d_1	$d_{1.5}$	d_2	$d_{2.5}$	d_3
Moyenne	0.1174	0.0117	0.0068	0.0112	0.0488	0.1052
Ecart-type	0.0114	0.0012	0.0019	0.0033	0.0059	0.0101

Essai 5: 10 tours de scaling.3, N = 40 simulations, 120 000 itérations / simulation.

	d	d_1	$d_{1.5}$	d_2	$d_{2.5}$	d_3
Moyenne	0.1240	0.0109	0.0053	0.0143	0.0530	0.1103
Ecart-type	0.0107	0.0021	0.0036	0.0047	0.0062	0.0088

Essai 6: 10 tours de scaling.3, N =45 simulations, 100 000 itérations / simulation.

	d	d_1	$d_{1.5}$	d_2	$d_{2.5}$	d_3
Moyenne	0.1148	0.0109	0.0055	0.0130	0.0495	0.1019
Ecart-type	0.0097	0.0019	0.0024	0.0036	0.0051	0.0082

Essai 7: 10 tours de scaling.3, avec $q \neq 1.5$, N = 20 simulations, 200 000 itérations / simulation.

	d	d_1	d_2	$d_{2.5}$	d_3
Moyenne	0.1245	0.0111	0.0126	0.0517	0.1119
Ecart-type	0.0160	0.0016	0.0039	0.0076	0.0141

On peut encore continuer à faire de tels essais en faisant varier les paramètres (ce que j'ai fait). On peut augmenter le nombre de simulations dans des limites acceptables en pratique, mais sans réelle amélioration au niveau stabilité des exposants d'échelle. Si on moyennise au niveau des $\tau_s(q)$ (et non sur les fonctions $S_q(\Delta t)$), on peut voir que cela ne change pas grand chose tant au niveau stabilité de la distance correspondante que pour la valeur de cette distance.

On peut introduire la distance relative suivante:

$$d^{(2)} := \left(\sum_{2q=1}^{6} \frac{\left(\overline{|\tau_d(q) - \bar{\tau}_s(q)|}\right)^2}{Var(|\tau_d(q) - \bar{\tau}_s(q)|)}\right)^{1/2}$$

En clair: on calcule sur N simulations des exposants $\bar{\tau}_s(q)$ (en prenant des moyennes sur les $S_q(\Delta t)$ comme dans la fonction précédente), et on refait cela un certain nombre M de fois, puis on prend la moyenne, sur les M jeux de simulations, des distances $\tau_d(q) - \bar{\tau}_s(q)$, ce que nous notons $\overline{|\tau_d(q) - \bar{\tau}_s(q)|}$. On calcule de même la variance sur ces M jeux de ces distances. On fait ensuite la somme sur q des carrés des moyennes divisées par les variances, et on prend la racine carrée de la somme.

Cette distance a l'air de se montrer plus stable que la précédente, ce qui est normal puisque l'on divise chaque distance individuelle par sa variance sur les M jeux de N simulations. L'intérêt de cette distance relative est bien sûr que l'on pourra également la comparer avec d'autres distances relatives liées à la fonction d'autocorrélation ou la forme de la fonction de densité.

On peut voir que plus on augmente le degré de liberté de la Student, plus la distance $d^{(2)}$ semble diminuer. A la limite, 8 degrés de liberté semblent encore donner encore de bons résultats:

3 x 10 tours de 20 simulations, avec 200 000 itérations / simulation:

 $d^{(2)}$ 11.646890 16.00930 14.71883

3 x 15 tours de 20 simulations, avec 200 000 itérations / simulation:

 $d^{(2)}$ 9.403307 15.49933 11.623658

Dans les essais qui suivent, on calcule sur M = 10 tours de N = 25 simulations, avec 200 000 itérations / simulation, la distance $d^{(2)}$ en faisant varier le nombre de degrés de liberté de la Student (différentes simulations du modèle avec random seeds non contrôlés):

2 degrés	33.73556	30.18153	25.86971	37.72401
3 degrés	13.781164	14.367043	13.388145	22.928287
4 degrés	5.120204	6.789409	7.607776	7.191657
5 degrés	9.386347	5.222071	6.058095	
6 degrés	5.062179	5.653010	5.563561	
8 degrés	4.384150	6.524564		

Tout ceci illustre le problème de la stabilité de ces exposants d'échelle calculés sur base de simulation et la nécessité de fixer le jeu de random seeds à utiliser pour les $N \times M$ simulations lorsque l'on veut faire varier les paramètres.

Distance relative avec contrôle du générateur de nombres aléatoires: un aperçu

Dans ce qui suit, nous avons calculé la fonction $d^{(2)}$ introduite plus haut en faisant varier certains paramètres du modèle en contrôlant les random seeds entre les différentes valeurs des paramètres. En clair: pour un jeu de paramètres, on fait $M \times N$ simulations pour calculer $d^{(2)}$ suivant un ensemble de $M \times N$ random seeds, puis on répète l'opération avec un autre jeu de paramètres mais les mêmes $M \times N$ random seeds.

Dans ce qui suit, on fait varier le paramètre Λ^2 (à travers les paramètres c_1 et c_2) pour un premier jeu de random seeds, puis pour un second (3 degrés de liberté pour la Student). Voir les figures 4.20, 4.21 pour les graphes correspondants de $d^{(2)}$.

Premier essai	$c_1 = 0.02$	$c_1 = 0.025$	$c_1 = 0.03$	$c_1 = 0.035$	$c_1 = 0.04$
$c_2 = -0.0035$	12.990	13.826	14.294	11.903	11.003
$c_2 = -0.003$	13.319	14.371	11.905	10.761	10.585
$c_2 = -0.0025$	13.777	11.949	10.595	10.288	10.373
$c_2 = -0.002$	11.995	10.475	10.064	10.042	10.268
$c_2 = -0.0015$	10.400	9.889	9.770	9.913	10.225
	1				
	$c_1 = 0.045$	$c_1 = 0.05$	$c_1 = 0.055$	$c_1 = 0.06$	$c_1 = 0.065$
$c_2 = -0.0035$	$c_1 = 0.045$ 10.955	$c_1 = 0.05$ 11.221	$c_1 = 0.055$ 11.652	$c_1 = 0.06$ 12.152	$c_1 = 0.065$ 12.683
$c_2 = -0.0035 c_2 = -0.003$	$ \begin{array}{r} c_1 = 0.045 \\ 10.955 \\ 10.774 \end{array} $	$c_1 = 0.05$ 11.221 11.154	$c_1 = 0.055 \\ 11.652 \\ 11.636$	$c_1 = 0.06$ 12.152 12.159	$ \begin{array}{r} c_1 = 0.065 \\ 12.683 \\ 12.688 \end{array} $
$c_2 = -0.0035$ $c_2 = -0.003$ $c_2 = -0.0025$	$ \begin{array}{r} c_1 = 0.045 \\ 10.955 \\ 10.774 \\ 10.690 \end{array} $	$c_1 = 0.05$ 11.221 11.154 11.134	$c_1 = 0.055$ 11.652 11.636 11.643	$c_1 = 0.06$ 12.152 12.159 12.174	$ \begin{array}{c} c_1 = 0.065 \\ 12.683 \\ 12.688 \\ 12.702 \end{array} $
$c_2 = -0.0035$ $c_2 = -0.003$ $c_2 = -0.0025$ $c_2 = -0.002$	$c_1 = 0.045$ 10.955 10.774 10.690 10.663	$c_1 = 0.05$ 11.221 11.154 11.134 11.145	$c_1 = 0.055$ 11.652 11.636 11.643 11.665	$c_1 = 0.06$ 12.152 12.159 12.174 12.192	$c_1 = 0.065$ 12.683 12.688 12.702 12.720

Second essai	$c_1 = 0.01$	$c_1 = 0.015$	$c_1 = 0.02$	$c_1 = 0.025$	$c_1 = 0.03$
$c_2 = -0.0035$	<u>6.39</u>	7.02	8.27	8.32	9.225
$c_2 = -0.0030$	<u>6.43</u>	7.30	7.81	8.60	8.040
$c_2 = -0.0025$	<u>6.63</u>	7.62	7.95	7.51	7.780
$c_2 = -0.0020$	6.91	7.33	7.13	7.29	7.920
$c_2 = -0.0015$	6.89	6.88	6.93	7.42	8.185
$c_2 = -0.0010$	6.72	6.61	6.99	7.64	8.470
- 4					
	$c_1 = 0.04$	$c_1 = 0.045$	$c_1 = 0.05$	$c_1 = 0.055$	$c_1 = 0.06$
$c_2 = -0.0035$	$c_1 = 0.04$ 8.62	$c_1 = 0.045$ 9.15	$c_1 = 0.05$ 10.010	$c_1 = 0.055$ 11.09	$c_1 = 0.06$ 12.135
$c_2 = -0.0035 c_2 = -0.0030$	$c_1 = 0.04$ 8.62 8.39	$c_1 = 0.045$ 9.15 9.26	$c_1 = 0.05$ 10.010 10.315	$c_1 = 0.055$ 11.09 11.39	$ \begin{array}{c} c_1 = 0.06 \\ 12.135 \\ 12.350 \end{array} $
$c_2 = -0.0035$ $c_2 = -0.0030$ $c_2 = -0.0025$	$ \begin{array}{c} c_1 = 0.04 \\ 8.62 \\ 8.39 \\ 8.53 \end{array} $	$ \begin{array}{r} c_1 = 0.045 \\ 9.15 \\ 9.26 \\ 9.52 \end{array} $	$c_1 = 0.05$ 10.010 10.315 10.620	$c_1 = 0.055$ 11.09 11.39 11.66	$ \begin{array}{c} c_1 = 0.06 \\ 12.135 \\ 12.350 \\ 12.510 \end{array} $
$c_2 = -0.0035$ $c_2 = -0.0030$ $c_2 = -0.0025$ $c_2 = -0.0020$	$c_1 = 0.04 \\ 8.62 \\ 8.39 \\ 8.53 \\ 8.80$	$c_1 = 0.045$ 9.15 9.26 9.52 9.85	$c_1 = 0.05$ 10.010 10.315 10.620 10.910	$c_1 = 0.055$ 11.09 11.39 11.66 11.85	$c_1 = 0.06$ 12.135 12.350 12.510 12.580
$c_2 = -0.0035$ $c_2 = -0.0030$ $c_2 = -0.0025$ $c_2 = -0.0020$ $c_2 = -0.0015$	$c_1 = 0.04 \\ 8.62 \\ 8.39 \\ 8.53 \\ 8.80 \\ 9.10$	$c_1 = 0.045$ 9.15 9.26 9.52 9.85 10.13	$c_1 = 0.05$ 10.010 10.315 10.620 10.910 11.120	$c_1 = 0.055$ 11.09 11.39 11.66 11.85 11.98	$c_1 = 0.06$ 12.135 12.350 12.510 12.580 12.600

On peut recommencer cela avec 4 degrés de liberté pour la Student (même random seeds que dans le premier essai ci-dessus).

	$c_1 = 0.01$	$c_1 = 0.015$	$c_1 = 0.02$	$c_1 = 0.025$	$c_1 = 0.03$	$c_1 = 0.035$
$c_2 = -0.0035$	14.62782	14.308039	11.429880	9.117836	8.273006	7.387496
$c_2 = -0.003$	15.12912	13.144978	10.715348	8.545241	7.794348	6.600938
$c_2 = -0.0025$	15.22857	12.109424	9.343433	8.410832	6.908823	6.222977
$c_2 = -0.002$	14.20069	10.056798	9.344068	7.399204	6.487432	5.945138
$c_2 = -0.0015$	12.84649	10.378956	8.079492	6.897464	6.161649	5.722104
$c_2 = -0.001$	11.54281	8.760789	7.505674	6.507246	5.890076	5.541744
	$c_1 = 0.04$	$c_1 = 0.045$	$c_1 = 0.05$	$c_1 = 0.055$	$c_1 = 0.06$	
$c_2 = -0.0035$	6.445698	6.115933	6.083947	6.796860	6.724253	
$c_2 = -0.003$	6.101086	5.907163	6.008367	6.904624	6.651849	
$c_2 = -0.0025$	5.862597	5.760510	6.031087	6.515537	6.634867	
$c_2 = -0.002$	5.682086	5.674454	6.307031	6.386983	6.706706	
$c_2 = -0.0015$	<u>5.544477</u>	5.650281	6.651786	6.349086	6.815612	
$c_2 = -0.001$	<u>5.443872</u>	5.709791	6.292636	6.367602	6.945288	

4.4.5 Conclusion et perspectives

D'après ce qui précède, le modèle semble assez bien reproduire les données au point de vue de la fonction d'autocorrelation des returns absolus ainsi que de la fonction de



Figure 4.20: Premier essai: allure de $d^{(2)}$ pour $(c_1, c_2) \in [0.02, 0.065] \times [-0.0035, -0.0015]$, 3 degrés de liberté.

densité. L'adéquation semble assez bonne lorsque le nombre de degrés de liberté de la Student est choisi égal à 3 ou 4 et Λ^2 choisi aux alentours de $c_1 = 0.04$ et $c_2 = -0.002$. On a vu cela d'abord par simple comparaison visuelle puis de façon plus quantitative avec différentes fonctions de distance.

Une procédure d'optimisation

Dans la suite, j'ai encore regroupé les deux fonctions de distances précédentes (la somme (éventuellement pondérée) de $d^{(2)}$ ainsi que de l'une des distances considérées pour les fonctions d'autocorrélation) et regardé l'allure générale de la nouvelle fonction ainsi obtenue. On pourrait encore y joindre une troisième distance quantifiant les différences de forme dans les fonctions de densité (ou de fonction de répartition). On peut alors essayer de minimiser (à l'aide d'une procedure de minimisation) la ou les fonctions ainsi obtenue(s). On peut utiliser diverses procédures de minimisation, la plus simple étant la fonction nlminb de S-plus. La minimisation n'a pour le moment pas encore été concluante, et va probablement encore nécessiter du travail, mais est en cours.

Toute cette partie expérimentale n'est donc pas terminée. Tout ce qui précède mériterait d'être poussé beaucoup plus loin, mais ceci sort du cadre d'un simple mémoire, pour des raisons évidentes de temps. Je compte continuer à regarder ces fonctions de distance, et notamment à accélérer la vitesse de calcul qui a été un réel problème jusqu'à présent et une des causes du nombre nombre limité d'observations que j'ai données, ainsi



Figure 4.21: Deuxième essai: allure de $d^{(2)}$ pour $(c_1, c_2) \in [0.01, 0.055] \times [-0.0035, -0.001]$, 3 degrés de liberté pour la Student.

que du non succès de l'optimisation. Le but de cette section n'était donc que de donner un aperçu de la situation.

Cet aspect expérimental est évidemment pour le moment empirique et n'a pas de justification statistique. Le but de tout ceci n'est pas pour l'instant de donner une méthode d'estimation rigoureuse des paramètres, mais juste d'essayer de comprendre mieux ce qu'il se passe et d'essayer de fournir des pistes pour une future estimation. C'est pour le moment la seule estimation que l'on peut faire. Dans ce domaine, pour ce genre de modèle et ce genre de propriétés statistiques, nous n'en sommes encore qu'à l'étape empirique, des estimateurs statistiques n'existent pas encore, notamment pour les exposants d'échelle. Et on a vu que dans ce domaine, beaucoup de contradictions, de contre-exemples peuvent apparaître, ce qui porte actuellement à polémique dans le monde scientifique (voir section 3.4.3). Des études plus détaillées doivent être faites pour répondre à ces questions.



Figure 4.22: Allure de $d^{(2)}$ pour $(c_1,c_2) \in [0.01,0.055] \times [-0.0035,-0.001],$ 4 degrés de liberté.

Conclusion

Dans le modèle de Black-Scholes, les returns logarithmiques sont i.i.d. gaussiens. Dans la pratique, on observe des propriétés statistiques des returns en contradiction avec ce modèle: queue lourde et non normalité de la distribution des returns lorsque l'intervalle de temps est suffisamment petit, la longue mémoire du processus des volatilités réalisées et le phénomène de "volatility clustering", ainsi que des comportements d'échelle. Pour les taux de change, on a également observé une dissymétrie de la fonction de corrélation croisée des volatilités réalises, phénomène interprété comme un flux net d'information des agents agissant à long terme vers ceux à court terme.

Différents modèles d'évolution des cours ont alors été proposés en vue de tenir compte de ces observations, et les modèles multifractals en sont un exemple.

Le modèle présenté dans ce mémoire se base sur une analogie entre les marchés financiers et les phénomènes de turbulence en mécanique des fluides. L'analogie se situe au niveau des propriétés statistiques des deux phénomènes, et le flux d'information à travers les échelles de temps est mis en parallèle avec la cascade de Richardson en turbulence.

Le modèle de Breymann-Ghashghaie-Talkner est un modèle où les returns suivent une loi de mélange, avec une volatilité qui s'exprime comme une cascade multiplicative, processus de nature multifractale. Il tente de traduire directement le flux d'information à travers les horizons de temps. Ce modèle semble reproduire assez bien la lente décroissance de la fonction d'autocorrélation des returns absolus, permet grâce à la présence de la Student, de reproduire la queue lourde et la non normalité de la distribution des returns, et les propriétés d'échelle (avec exposants non linéaires) semblent également pouvoir être reproduits assez bien. Il y a même moyen d'ajuster plus finement les paramètres pour obtenir des résultats assez satisfaisants selon plusieurs critères.

Enormément de questions ouvertes sont liées à ce type de modélisation.

Estimer les indices de queue de la distribution des returns par estimateur de Hill n'est pas évident en soit si la distribution des returns est supposée inconnue a priori. On pourrait également tenter d'utiliser la méthode du peak over thresholds.

La mise en évidence de lois d'échelle n'est pas non plus un problème simple. On a cité des travaux où des comportements d'échelle apparaissent dans des séries simulées à partir de modèles n'en présentant pas en théorie. Des comportements d'échelle avec exposants non linéaires (typiques des multifractals) peuvent aussi apparaître dans des simulations de modèles monofractals, probablement à cause du caractère à longue mémoire du processus des volatilités. Or les données financières présentent une telle longue mémoire. Que signifient donc les comportements d'échelle observé universellement sur les données financières? Les données possèdent-elles réellement des propriétés multifractales? Doit-on y accorder de l'importance en modélisation? Y a-t-il un avantage à utiliser des modèles multifractals par rapport aux modèles actuels? Ces questions portent actuellement à polémique dans la communauté scientifique.

Dans le modèle multifractal en cascade, le problème du fit des paramètres n'est pas non plus trivial, et n'a jusqu'à présent pas été réalisé jusqu'au bout. Une estimation par maximum de vraisemblance semble assez compliquée, et un essai de fit est en cours, tentant d'optimiser (minimiser) une ou plusieurs fonctions de "distance" entre les données réelles et simulées selon le modèle, distance s'appuyant sur la fonction d'autocorrélation des returns absolus, la forme des fonctions de densité et les lois d'échelle. Une telle optimisation est évidemment empirique, mais donnerait une méthode d'estimation des paramètres alternative, même si de qualité discutable à première vue. C'est pour le moment la seule chose que l'on puisse faire. Cette méthode pourrait également donner des pistes pour une estimation plus rigoureuse. Beaucoup de travail statistique mérite donc d'être accompli autour de ce modèle pour obtenir de bons estimateurs. En particulier, beaucoup de travail doit être réalisé autour des propriétés d'échelle.

Appendice A Dimensions fractales

A.1 Dimension de boîte ou capacité

On se place dans \mathbb{R}^N muni de la norme usuelle $||x|| = \left(\sum_{i=1}^N x_i^2\right)^{1/2}$ et on considère un sous-ensemble borné E. La dimension de boîte de E est définie de la façon suivante. On recouvre E par un maillage de cubes N-dimensionnels d'arête de longueur ε , et on note $\tilde{N}(\varepsilon)$ le nombre minimum de cubes nécessaires pour recouvrir E. On fait cela en prenant des ε de plus en plus petits. La dimension de boîte de l'ensemble E (ou capacité de E) est définie par

$$\dim_0(E) := \lim_{\varepsilon \downarrow 0} \frac{\ln(N(\varepsilon))}{\ln(1/\varepsilon)}$$

à condition que cette limite existe.

Exemples:

- $E = \{a_1, \ldots, a_k\}, k$ points isolés. Dans ce cas, $\tilde{N}(\varepsilon) = k$ pour tout ε suffisamment petit, donc dim₀(E) = 0.
- E = l'ensemble triadique de Cantor (ou poussière de Cantor). Rappelons très brièvement sa construction: on considère le segment [0, 1] ⊂ ℝ auquel on va enlever successivement une suite d'intervalles ouverts. A l'étape 1, on divise [0, 1] en trois intervalles de même longueur, et on enlève celui du milieu sans ses extrémités, à savoir (¹/₃, ²/₃). A l'étape 2, on subdivise chacun des deux intervalles restants, [0, ¹/₃] et [²/₃, 1], en trois sous-intervalles de même longueur, et on enlève ceux du milieux, c'est-à-dire (¹/₉, ²/₉) et (⁷/₉, ⁸/₉). Et ainsi de suite. On peut voir que la longueur des sous-intervalles restants après chaque étape n est égale à (1/3)ⁿ, et que le nombre de sous-intervalles est égal à 2ⁿ.

Pour mesurer sa dimension, on prend une suite $(\varepsilon_n) \to 0$, et par définition, sa dimension de boîte, en admettant qu'elle existe, sera égale à $\lim \frac{\ln(\tilde{N}(\varepsilon_n))}{\ln(1/\varepsilon_n)}$. Par construction de l'ensemble de Cantor, on a intérêt à choisir $\varepsilon_n = (1/3)^n$. Dans ce cas, $\tilde{N}(\varepsilon_n) = 2^n$, ce qui nous donne finalement $\dim_0(E) = \ln 2/\ln 3$. On peut voir que cet ensemble est auto-similaire, c'est à dire que certains de ses sous-ensembles seront, après une transformation affine, égaux à l'ensemble de départ (exemple: $E \cap [0, 1/3]$, après la transformation $x \mapsto 3x$). On peut voir également qu'il est infini non dénombrable, mesurable mais de mesure de Lebesgue nulle.

A.2 Dimension de Hausdorff

La définition de la dimension de Hausdorff est basée sur celle de mesure de Hausdorff.

A.2.1 Mesure de Hausdorff d-dimensionnelle

On se place à nouveau dans \mathbb{R}^N muni de la norme euclidienne. Si $E \subset \mathbb{R}^N$, on définit son diamètre par

$$|E| = \sup_{x,y \in E} ||x - y||,$$

c'est-à-dire la plus grande distance entre deux éléments quelconques de E. Soit $\delta > 0$ fixé, et soit un recouvrement de E par un nombre au plus dénombrable d'ensembles S_i , $i \in I$, tels que $|S_i| \leq \delta$ pour tout i. On peut voir qu'un tel recouvrement existe bien (considérer des boules de rayon et de centre rationnels). On définit alors, pour $d \in \mathbb{R}^+$,

$$\Gamma^d_H(\delta) := \inf_{S_i} \sum_{i \in I} |S_i|^d,$$

où l'infimum est pris sur tous les recouvrements de ce type. On fait ensuite tendre δ vers 0 pour obtenir la mesure d-dimensionnelle de Hausdorff:

$$\Gamma^d_H(E) := \lim_{\delta \downarrow 0} \Gamma^d_H(\delta).$$

L'intérêt de cette mesure par rapport à la mesure de Lebesgue est que l'on peut la définir pour n'importe quel sous-ensemble de \mathbb{R}^N .

A.2.2 Dimension de Hausdorff

On peut montrer que $\Gamma^d_H(E)$ sera toujours du type

$$\Gamma_H^d(E) = \begin{cases} +\infty & \text{si } d < d_{\text{crit}} \\ 0 & \text{si } d > d_{\text{crit}}. \end{cases}$$

Par définition, la dimension de Hausdorff de E, $\dim_H(E)$, sera égale à cette valeur critique d_{crit} .

A.3 Lien entre les deux dimensions

Dans la dimension de boîte, on recouvre E non pas par des sous-ensembles quelconques dont on contrôle le diamètre, mais par des cubes dont on contrôle la longueur d'arête. Pour $\delta > 0$ fixé, si S_i est un cube d'arête $\varepsilon = \varepsilon(\delta) = \delta/\sqrt{N}$, alors le diamètre de S_i vaut δ . Si l'on recouvre un ensemble borné E par de tels cubes,

$$\Gamma^d_H(\delta) \le \sum_{i \in I} \delta^d = \tilde{N}(\varepsilon(\delta))(\varepsilon(\delta))^q N^{\frac{d}{2}}.$$

Comme il existe une constante K > 0 telle que $\tilde{N}(\varepsilon) \leq K\varepsilon^{-\dim_0(E)}$ pour tout ε suffisamment petit, on en conclut que

$$\lim_{\delta \to 0} \Gamma_H^d(\delta) \le C \lim_{\delta \to 0} (\varepsilon(\delta))^{d - \dim_0(\mathsf{E})}$$

pour une certaine constante C > 0. Comme cette dernière limite est nulle ssi $d > \dim_0(\mathbf{E})$, on en déduit que

$$\dim_H(E) \le \dim_0(E).$$

On peut voir que pour certains sous-ensembles, cette inégalité est stricte (exemple: $E = \{1/n \mid n \in \mathbb{N}_0\}$).

Le résultat suivant est prouvé dans [15] (page 29).

Lemme A.3.1 (σ -continuité de la dimension de Hausdorff) Si $(F_n)_{n \in \mathbb{N}}$ est une suite d'ensembles alors

$$\dim_H(\bigcup_{i=1}^{\infty} F_i) = \sup\{\dim_H F_i : i \in \mathbb{N}\}.$$

Preuve: Comme $E \subset F$ implique $\dim_H E \leq \dim_H F$, on sait que pour tout $i \in \mathbb{N}$, $\dim_H(\bigcup_{i=1}^{\infty}F_i) \geq \dim_H F_i$. On en déduit que $\dim_H(\bigcup_{i=1}^{\infty}F_i) \geq \sup\{\dim_H F_i : i \in \mathbb{N}\}$. D'autre part, si $s > \dim F_i$ pour tout i, alors la mesure de Hausdorff $\mathcal{H}^s(F_i) = 0$, ce qui implique que $\mathcal{H}^s(\bigcup_{i\in\mathbb{N}}F_i) = 0$.

Table des Figures

1.1	Fonction d'autocovariance $r(n)$	23
1.2	Simulations d'un mouvement brownien fractionnaire d'indice $H = 0.2$, H = 0.5 et $H = 0.8$.	25
1.3	Bruits gaussiens fractionnaires correspondants (1000 premières itérations)	26
1.4	Auto-similarité d'un mouvement brownien fractionnaire pour $H = 0.2$.	27
1.5	Auto-similarité d'un mouvement brownien fractionnaire pour $H = 0.8$.	27
2.1	Mesure binomiale: Trois premières itérations μ_1, μ_2, μ_3 , avec $m_0 = 1/3, m_1 = 2/3$	45
2.2	Mesure binomiale: μ_7, μ_{10} , avec $m_0 = 1/3, m_1 = 2/3$	45
2.3	Mesure binomiale: spectres $f(\alpha)$ et $\tau(q)$ si $m_0 = 0.3$ et $m_1 = 0.7$	49
3.1	Données: cours USD-CHF du 02/01/1991 au 30/05/2001	56
3.2	Returns logarithmiques pour $\Delta t = 1$ heure	57
3.3	Qqplot des returns (centrés réduits) pour $\Delta t = 1$ heure, 6 heures, 1 jour et 1 semaine par rapport à une normale standard	63
3.4	Fonction de densité (à gauche) et logarithme de la fonction de densité (à droite) des returns centrés réduits pour $\Delta t = 1$ h et comparaison avec une	
	normale standard	64
3.5	Fonction de densité logarithmique des returns centrés réduits pour $\Delta t =$	
	1h et comparaison avec une Student à 3 ou 4 degrés de liberté	64
3.6	QQplot des returns (centrés réduits) pour $\Delta t = 1$ h et comparaison avec	~
3.7	une Student à 3, 4 et 5 degrés de liberté	65
	les 7000 premiers	74
3.8	Returns absolus pour les données (returns pour $\Delta t = 1$ heure, unités de	
	temps du graphique: 1 heure), pour un mouvement brownien standard et un FBM	75
3.9	Données: fonction d'autocorrélation des returns absolus et carrés pour	
	$\Delta t = 1h \dots \dots \dots \dots \dots \dots \dots \dots \dots $	76
3.10	Données: fonction d'autocorrélation des returns absolus en échelle bilog-	
	arithmique	76

3.11	Données: fonction d'autocorrélation des returns absolus pour différents	
2 19	Intervalles de temps	((
0.12	de temps	78
3.13	Fonction d'autocorrélation des returns absolus pour les données (en haut, avec la semaine comme unité de temps et l'intervalle de calcul des returns Δt égal à 1 heure), un mouvement brownien standard (au milieu) et un FBM d'indice $H = 0.8$ (en bas).	79
3.14	$\log S_q(\Delta t)$ en fonction de $\log \Delta t$ et régression linéaire	81
3.15	Données: exposants d'échelle $\xi(q)$, et comparaison notamment avec la droite $y = 0.5x$	81
3.16	Exposants d'échelle $\xi(q)$ pour un mouvement brownien et un mouvement brownien fractionnaire avec $H = 0.8$.	83
3.17 3.18	Données: $\log S_q(\Delta t)$ en fonction de $\log \Delta t$ pour différents sous-échantillons Données: comparaison entre la movenne sur les différents sous-échantillons	84
	des exposants $\xi(q)$ et leurs valeurs calculés sur tout l'échantillon	84
4.1	Returns simulés suivant le modèle.	99
4.2	Fonction d'autocorrélation des returns absolus: Comparaison modèle - données sur différentes simulations	100
4.34.4	Fonction d'autocorrélation des returns absolus: Comparaison modèle - données, random seed fixé, nombre de degrés de la Student $=2$ à 5 (de gauche à droite et de haut en bas), autres paramètres fixés Fonction d'autocorrélation des returns absolus: Comparaison modèle - données, random seed fixé (mais différent de celui de la figure 4.3), nombre	101
4 5	de degres de la Student = 2 a 5 (de gauche a droite et de naut en bas), autres paramètres fixés	102
4.5	Fonction d'autocorrelation des returns pour le modele simule: non auto- corrélation des returns	103
4.6	Returns: Fonction de densité logarithmique: Comparaison modèle - données (en pointillé), même paramètres (dont nombre de degrés de la Student =	
4.7	3, $c_1 = 0.05, c_2 = -0.002$), random seeds différents	104
4.8), random seeds différents	105
	(en pointillé), même paramètres (deg Student = 5), $c_1 = 0.05, c_2 = -0.002$, random seeds différents	106
4.9	Returns: Fonction de densité logarithmique: Comparaison modèle - données (en pointillé), $c_1 = 0.05, c_2 = -0.002$, nbre degrés Student = 2 à 6 (de	105
/ 10	gauche à droite et de haut en bas), random seed fixé	107
4.10	(en pointillé), $c_1 = 0.035$ à 0.065	107

TABLE DES FIGURES

4.11	Returns: Fonction de densité logarithmique: Comparaison modèle - données	
	(en pointillé), $c_1 = 0.035$ à 0.065, nombre de degrés de la Student = 4,	
	$c_2 = -0.002 \dots \dots \dots \dots \dots \dots \dots \dots \dots $	108
4.12	Returns: Fonction de densité logarithmique: Comparaison modèle - données	
	(en pointillé), $c_1 = 0.1$ à 0.22, nombre de degrés de la Student = 4,	
	$c_2 = -0.002 \dots \dots \dots \dots \dots \dots \dots \dots \dots $	108
4.13	Returns: qqplot des returns centrés réduits: comparaison modèle - données	109
4.14	Exposants d'échelle: $\xi(q)$ en fonction de q , comparaison modèle – données	
	(courbe du dessus), mêmes paramètres (dont degré Student = 3, c_1 =	
	$0.05, c_2 = -0.002$, random seeds différents	110
4.15	Exposants d'échelle: Comparaison modèle – données, mêmes paramètres	
	(dont nombre de degrés de liberté de la Student = 4, $c_1 = 0.05, c_2 =$	
	-0.002), random seeds différents	111
4.16	Exposants d'échelle: $\log S_q(\Delta t)$ en fonction de $\log \Delta t$, nombre de degrés	
	de la Student = 3, $c_1 = 0.05, c_2 = -0.002$	111
4.17	Exposants d'échelle: $\log S_q(\Delta t)$ en fonction de $\log \Delta t$, nombre de degrés	
	de la Student = 4, $c_1 = 0.05, c_2 = -0.002$	112
4.18	Exposants d'échelle: $\log S_q(\Delta t)$ en fonction de $\log \Delta t$, nombre de degrés	
	de la Student = 5, $c_1 = 0.05, c_2 = -0.002$	112
4.19	Valeurs d'une fonction distance pour la fonction d'autocorrélation pour	
	différentes valeurs des paramètres	117
4.20	Premier essai: allure de $d^{(2)}$ pour $(c_1, c_2) \in [0.02, 0.065] \times [-0.0035, -0.0015]$,	
	3 degrés de liberté	122
4.21	Deuxième essai: allure de $d^{(2)}$ pour $(c_1, c_2) \in [0.01, 0.055] \times [-0.0035, -0.001]$,	
	3 degrés de liberté pour la Student	123
4.22	Allure de $d^{(2)}$ pour $(c_1, c_2) \in [0.01, 0.055] \times [-0.0035, -0.001], 4$ degrés de	
	liberté	124

TABLE DES FIGURES

Bibliographie

- O.E. Barndorff-Nielsen, K. Prause, Apparent scaling, *Finance and Stochastics* (1999).
- [2] O.E. Barndorff-Nielsen, Probability and statistics: self decomposability, finance and turbulence, Probability towards 2000 (New York, 1995), 47-57, *Lecture Notes* in Statist., **128**, Springer, New York (1998).
- [3] J. Beran, Statistics for long-memory processes, Chapman & Hall, New-York (1994).
- [4] T. Bollerslev, Generalized autoregressive conditional heteroscedasticity, J. Econ., 31 (1986), 307-327.
- [5] T. Bollerslev, R.Y. Chou, K.F. Kroner, ARCH modelling in finance. A review of the theory and empirical evidence, J. Econom, 52, n. 1/2, 5-59 (1992).
- [6] J.-P. Bouchaud, M. Potters, M. Meyer, Apparent multifractality in financial time series, preprint 1999.
- [7] W. Breymann, S. Ghashghaie et P. Talkner, A stochastic cascade model for FX dynamics, *Journal of Theoretical and Applied Finance*, 3, (2000) 357-360.
- [8] L. Calvet, A. Fisher, B.B. Mandelbrot, Large deviations and the distribution of price changes, Cowles Fondation Discussion Paper # 1165 (1997).
- [9] Rama Cont, Statistical Finance: Empirical and Theoretical Approaches to the Statistical Modelling of Price Variation in Speculative Markets, Ph. D. thesis, CMAP-Ecole Polytechnique (1998).
- [10] M. Dacorogna, R. Gençay, U. A. Müller, R. B. Olsen, O. V. Pictet, An introduction to high frequency finance, Academic Press, 2001.
- [11] L. De Haan et S.I. Resnick, Fighting the arch-ennemy with mathematics, *Statistica Nerlandica*, 44, (1980) 45-68.
- [12] A.L.M. Dekkers, J.H.J. Einmahl et L. de Haan, A moment estimator for the index of an extreme value distribution, *Annals of Statistics*, 17, (1990) 1833-1855.
- [13] P. Embrechts et M. Maejima, Selfsimilar processes, preprint to be edited.

- [14] R. Engle, Autoregressive conditional heteroskedasticity with estimates of the varianceof U.K. inflation, *Ecomometrica* **50**, (1982) 987-1008.
- [15] K. Falkoner, Fractal geometry: Mathematical Foundations and Applications, John Wiley and Sons, New York (1990).
- [16] W. Feller, An introduction to probability theory and its applications, Vol. 2, 2nd edn, Wiley, New-York.
- [17] A. Fisher, L. Calvet, B.B. Mandelbrot, Multifractality of Deutschemark/US Dollar exchange rates, Cowles Fondation Discussion Paper 1165 (1997).
- [18] U. Frisch, Turbulence, Cambridge University Press, Cambridge (1995).
- [19] S. Ghashghaie, W. Breymann, J. Peinke, P. Talkner, Y. Dodge, Turbulent cascades in foreign exchange markets, *Nature* 381, 776-770 (1996).
- [20] B.V. Gnedenko et A.N. Kolmogorov, Limit distributions for sums of independent random variables, Addison-Wesley, Readings, MA (1954).
- [21] P. Hall, Using the bootstrap to estimate mean square error and select smoothing parameter in nonparametric problem, *Journal of Multivariate Analysis*, **32** (1990), 117-203.
- [22] B.M. Hill, A simple general approach to inference about the tail of a distribution, Annals of Statistics, 13, (1975), 331-341.
- [23] A.N. Kolmogorov, The local structure of turbulence in incompressible viscous fluid for every large Reynolds numbers, *Dokl. Acad. Nauk. SSSR* 30, (1941), 301-305.
- [24] A.N. Kolmogorov, A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds numbers, J. Fluid. Mech. 13, (1962) 82-85.
- [25] A. McNeil, T. Saladin, The peak over thresholds method for estimating high quantiles of loss distributions, *Proceedings of the 28th ASTIN Colloquium* (1997).
- B.B. Mandelbrot, The variation of certain speculative prices, *Journal of Business*, 36, (1963), 394-419.
- [27] B.B. Mandelbrot, A. Fisher, L. Calvet, A multifractal Model of Asset Returns, Cowles Fondation Discussion Paper # 1164 (1997).
- [28] U. A. Müller, M.M. Dacorogna, R.B. Olsen, O.V. Pictet, M. Schwartz, C. Morgenegg, Statistical study of foreign exchange rates, empirical evidence of a price change scaling law, and intrady analysis, *Journal of Banking and Finance*, 14, (1990) 1189-1208.

- [29] U. A. Müller, M.M. Dacorogna, R.D. Davé, R.B. Olsen, O.V. Pictet, J.E. von Weizsächer, Volatilities of different time resolutions – Analysing the dynamics of market components, *Journal of Empirical Finance* 4 (1997) 213-239.
- [30] U.A. Müller, M.M. Dacorogna, R.D. Davé, O.V. Pictet, R.B. Olsen, J.R. Ward, Fractals and intrinsic time – A challenge to econometricians, Invited presentation at the XXXIXth Int. AEA Conf. on Real Time Econometrics, 14-15 Oct 1993 in Luxembourg, and the 4th Int. PASE Workshop, 22-26 Nov 1993 in Ascona, Switzerland.
- [31] B. Sz.-Nagy, Introduction to Real Functions and Orthogonal Expansions, Oxford University Press, New York (1965).
- [32] A. M. Obukhov, Some specific features of atmospheric turbulence, J. Fluid. Mech. 13, (1962) 77-81.
- [33] Pickands J., Statistical inference using extreme order statistics, The Annals of Statistics, 3(1), (1975) 119-131.
- [34] O.V. Pictet, M.M. Dacorogna, U.A. Müller, Hill, bootstrap and jackknife estimators for heavy tails, dans A practical guide to heavy tails: statistical techniques for analysing heavy tailed distributions, R.J. Adler, R.E. Feldman and M.S. Taqqu, Eds., pp. 283-310.
- [35] P. Protter, Stochastic Integration and Differential Equations, A New Approach, Springer-Verlag (1990).
- [36] R.H. Riedi, Multifractal processes, preprint to be published.
- [37] G. Samorodnitsky et M. S. Taqqu, Stable non-gaussian random processes, Stochastic models with infinite variance, Chapman & Hall, 1994.
- [38] T. Tél, Fractals, Multifractals, and Thermodynamics, An Introductory Review, Z. Nturforsch. 43a, 1154-1174 (1988).
- [39] T. Tél, A. Fülop and T. Vicsek, Determination of fractal dimensions for geometrical multifractals, *Physica A* 159, 155-166 (1989).
- [40] J. M. Steele, Stochastic calculus and financial applications, Springer, 2000.
- [41] C. Walter, Levy-stables distributions and fractal structure on the Paris market: an empirical examination, *Proceedings of the 1st AFIR international colloquium*, vol. 3, Paris, 241-259 (1990).
- [42] C. Walter, Levy-stability-under-addition and fractal structure of markets: imlications for the actuaries and emphasized axamination of MATIF national contract, *Proceedings of the 5st AFIR international colloquium*, 1285-1330.

BIBLIOGRAPHIE

[43] A. Zolotarev, One-dimensional stable distributions, Vol. 65 de Translations of mathematical monographs, AMS (1986). Traduction de l'édition originale en russe de 1983.