OPÉRATEURS DIFFÉRENTIELS FRACTIONNAIRES EN ÉCONOMÉTRIE ET EN FINANCE

F. COMTE

Abstract. Cet article expose comment les opérateurs pseudo-différentiels sont progressivement apparus dans les modèles économétriques et notamment financiers et quelles sont les propriétés occasionnées par ces opérateurs qui intéressent plus particulièrement les modélisateurs du domaine économique.

1. Introduction

Les opérateurs pseudo-différentiels ne sont pas vraiment connus sous ce nom-là en Econométrie. Pourtant, ils ont joué depuis 20 ans un rôle prépondérant aussi bien du côté économique (modélisation) que du côté statistique (méthodes d'estimation et comportements limites des estimateurs des paramètres de ces modèles). La présentation qui suit va tenter bien davantage de distinguer les étapes marquantes de cette évolution, plutôt que de présenter une liste exhaustive de ses acteurs (et de ses auteurs). Des revues de la littérature existent, concernant surtout l'étape centrale de notre description : Lardic et Mignon (1999), Beran (1994), mais tel n'est pas l'objectif du présent exposé. Il s'intéressera surtout aux jalons de l'histoire, et les citations retenues seront seulement partielles et donc nécessairement partiales.

On peut considérer que le modèle aléatoire le plus "simple" impliquant un opérateur pseudo-différentiel est le Brownien fractionnaire, introduit par Mandelbrot et Van Ness (1968). C'était un processus à temps continu, associé à un opérateur de dérivation fractionnaire continu. Il est resté pendant longtemps d'utilité marginale en économétrie, sauf peut-être pour la statistique du paramètre de Hurst (1951) d'autosimilarité qui lui a été associée.

Un grand nombre de modèles économiques étant plutôt définis à temps discret, c'est lorsque des opérateurs fractionnaires ont été définis en temps discret (Granger et Joyeux (1980), Hosking (1981)), que le déclic s'est produit. L'opérateur de différence fractionnaire avait en effet le mérite d'être relativement simple à définir, d'étendre de façon mathématique rigoureuse et claire une classe de modèles d'usage courant et de présenter malgré tout des propriétés à la fois très distinctes de la classe usuelle et très faciles à interpréter. Cette interprétation apparaît dans le terme de "mémoire longue" qui a depuis lors été accolé à ce type de dynamique : il illustre bien l'idée que les variables observées dépendent très longtemps les unes des

Date: December 2001.
Université Paris V, Laboratoire MAP5, email: comte@biomedicale.univ-paris5.fr.

autres et que le passé s'oublie lentement; cette dépendance est simplement mesurée par le biais de la fonction d'autocorrélation dans le domaine temporel ou par la non bornitude inférieure ou supérieure de la densité spectrale, dans le domaine des fréquences.

Un obstacle a pourtant surgi sur le plan statistique : ces modèles ne vérifiaient plus les hypothèses des théories de l'estimation standard, et il a fallu s'adapter à ce cadre nouveau (Fox et Taqqu (1986), Dahlhaus (1989), Robinson (1994)).

Les opérateurs continus étaient sous-jacents, par le biais du Brownien fractionnaire, via des tentatives d'insérer dans la nouvelle classe de processus le bruit fractionnaire (accroissement du Brownien fractionnaire sur une unité de temps, Mandelbrot et Wallis (1969)) dans la nouvelle classe de modèles (Geweke et Porter-Hudak (1983)), et via des théorèmes limites faisant apparaître le Brownien fractionnaire comme loi limite (Akonom et Gouriéroux (1988), Sowell (1990)).

Cependant, lorsqu'on a voulu appliquer ces idées en Finance, on s'est rapidement convaincu, puisque beaucoup de modèles existaient en temps continu, qu'il fallait représenter en temps continu également la mémoire longue soupçonnée dans la volatilité des rendements des actifs (Comte et Renault (1998), Comte, Coutin et Renault (2001)). Le Brownien fractionnaire a donc retrouvé sa place prépondérante, et les opérateurs pseudo-différentiels leur rôle de départ.

Ce sont les grands traits de l'histoire qui va être détaillée dans ce qui suit.

2. Un processus continu peu usité

Au départ, en statistique plutôt que spécifiquement en économétrie, il y a sans doute eu le coefficient de Hurst H, défini en hydrologie par Hurst (1951) qui souhaitait évaluer statistiquement la quantité d'eau à mettre dans un réservoir pour qu'il ne déborde jamais ni ne soit jamais vide. Ainsi fut définie la statistique $Q = R/s$ (étendue (Range) sur écart type (standard deviation)), associée à une série temporelle X_t, $t = 1, \ldots, n$ de moyenne empirique $\bar{X}_n = (1/n)\sum_{i=1}^n X_i$,

$$Q_n = R_n/s_n, \begin{cases} R_n = \max_{1\le k\le n} \sum_{i=1}^{k}(X_j - \bar{X}_n) - \min_{1\le k\le n} \sum_{i=1}^{k}(X_j - \bar{X}_n) \\ s_n = \left[\dfrac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2\right]^{1/2} \end{cases}$$

L'exposant de Hurst H est alors défini comme le nombre, s'il existe, tel que

$$\frac{1}{n^H}\frac{R_n}{s_n} \text{ converge en loi vers une loi non dégénérée.}$$

Dès lors, si l'on définit un processus $(X(t), t \in \mathbb{N})$ H-self similaire (Hss) comme un processus tel que $X(at)$ et $a^H X(t)$ ont même lois finies-dimensionnelles, $\forall a > 0$, il apparaît que R_n a même loi que $n^H R(1)$ et que si de plus X est stationnaire ergodique de variance finie, s_n converge en probabilité vers $\mathrm{Var}(X(1))$. Ainsi, l'exposant de Hurst correspond à l'indice d'autosimilarité.

C'est ce coefficient qui a donné son nom au Brownien fractionnaire dans ses définitions originelles (Mandelbrot et Van Ness (1968)) :
(1)
$$B_H(t)-B_H(0)=\frac{1}{\Gamma(H+\frac{1}{2})}\left\{\int_{-\infty}^{0}[(t-s)^{H-\frac{1}{2}}-(-s)^{H-\frac{1}{2}}]dB(s)+\int_{0}^{t}(t-s)^{H-\frac{1}{2}}dB(s)\right\}.$$
Disons-le tout de suite, nous appellerons indifféremment le Brownien fractionnaire B_H ou B_α où $\alpha = H - 1/2$. Le paramètre α varie entre $-1/2$ et $1/2$ quand H varie entre 0 et 1; le centre du domaine (qui ramène du Brownien fractionnaire au Brownien standard) est en 0 pour α quand il se trouve en $1/2$ pour H.
Les applications se sont d'ailleurs très vite intéressées aux accroissements de Brownien fractionnaire autant qu'au Brownien fractionnaire lui-même. Appelons ces accroissements "bruit fractionnaire" (Mandelbrot et Wallis (1969)) :

$$\varepsilon_t^{(\alpha)} = B_\alpha(t) - B_\alpha(t-1). \tag{2}$$

Ce bruit ajoute aux propriétés du Brownien fractionnaire des propriétés de stationnarité souvent très prisées en modélisation.
Le coefficient H (ou α) se retrouve dans plusieurs propriétés du bruit fractionnaire qui sont sans rapport avec l'auto-similarité. En effet, il joue un rôle sur quatre plans :

(1) indice d'autosimilarité,
(2) ordre de régularité trajectorielle,
(3) vitesse de décroissance des autocorrélations
(4) ordre de l'équivalent de la densité spectrale en 0

Et cela a donné lieu à de nombreuses confusions. Les deux dernières propriétés sont liées, mais elles sont très distinctes des deux premières. Les propriétés (3) et (4) sont connues comme caractéristiques de la mémoire longue, et c'est à elles que nous allons nous intéresser plus particulièrement. La propriété (1) correspond, on l'a dit, à la $\alpha + 1/2$-self similarité. La propriété (2) correspond au caractère β-Hölder (à une modification près) des trajectoires, pour tout $\beta < \alpha + \frac{1}{2}$. Comme on le verra, les processus utilisés pour généraliser le Brownien fractionnaire peuvent conserver des propriétés de mémoire à l'exclusion des autres, ou bien conserver la mémoire et la régularité. L'auto-similarité est une caractéristique qui a trouvé peu d'échos dans les applications économiques. Tout au plus quelques développements récents proposent-ils en Finance d'intéressants tests d'auto-similarité de trajectoires de rendements d'actifs (Bardet (2000a,b)); les estimateurs sont mis en place par analogie avec le Brownien Fractionnaire mais la modélisation sous-jacente est sommaire (Brownien fractionnaire seul), ce qui fait quelque peu défaut pour la prévision ou l'interprétation. Ces travaux sont le prolongement de résultats statistiques plus purement théoriques énoncés par Guyon et Leon (1989) puis Istas et Lang (1997) pour l'estimation du coefficient de Hölder des trajectoires d'un processus. On trouve souvent, cependant, les modèles à longue mémoire mis en regard de l'exposant de Hurst et de la statistique R/s, avec lesquels ils n'ont finalement que peu de rapport.

Seuls quelques travaux de Mandelbrot (voir Mandelbrot (1997)) sont à relier avec cet aspect des choses.
Pour finir, précisons les propriétés (3) et (4) du bruit fractionnaire qui vont le faire resurgir à quelques années de là dans un contexte plus économique :

$$\gamma_k = \mathrm{cov}(\varepsilon_t^{(\alpha)}, \varepsilon_{t+k}^{(\alpha)}) = \frac{1}{2}\left[|k+1|^{2\alpha+1} + |k-1|^{2\alpha+1} - 2|k|^{2\alpha+1}\right]$$

ce qui entraîne

$$\gamma_k \sim_{k\to+\infty} \alpha(2\alpha+1)k^{2\alpha-1}.$$

De plus, la transformée de Fourier continue de cette fonction γ, notée $f(\lambda)$ (pour peu que l'on parvienne à la définir, voir Samko et al. (1993)), est proportionnelle à $\lambda^{-2\alpha}$, ce qui, en fonction du signe de α, la rend nulle ou explosive lorsque λ tend vers 0. Ce sont ces deux caractérisations qui ont ensuite le plus fréquemment hanté la littérature économétrique. Cependant, ce fut dans un premier temps, en escamotant totalement les processus à temps continu d'origine et en se plaçant exclusivement en temps discret. Il faut reconnaître que pour étudier des processus à relevés hebdomadaires ou mensuels, les modèles discrets se sont imposés naturellement.

3. L'ÉVOLUTION FRACTIONNAIRE DES MODÈLES À TEMPS DISCRET

3.1. **Les ARMA, un outil simple et bien défini.** Parmi ces modèles, certains parviennent à réconcilier rigueur mathématique et signification économique (contrairement aux modèles ARCH un peu plus tard). En effet, les modèles ARMA représentent une variable ou un vecteur de variables comme linéairement expliqué par son passé immédiat et plus lointain, à des termes d'erreurs près, eux-mêmes plus ou moins indépendants. Le cas multidimensionnel est très porteur et riche de sens en particulier puisqu'il permet aussi de modéliser des influences croisées entre diverses variables, en plus des interactions temporelles. Nous considèrerons néanmoins pour simplifier des modèles unidimensionnels. Ainsi, si L est l'opérateur retard (L pour lag) et $LX_t = X_{t-1}$, $L^kX_t = X_{t-k}$, on peut définir des polynômes en ces opérateurs visant à généraliser le tout premier de la série : $X_t = \rho X_{t-1} + \varepsilon_t$ pour des ε_t i.i.d. centrés, à des modèles prenant en compte des dépendances plus complexes :

$$X_t = \sum_{i=1}^{p} a_i X_{t-i} + \varepsilon_t + \sum_{i=1}^{q} b_i \varepsilon_{t-i}$$

soit encore

$$A(L)X_t = B(L)\varepsilon_t, \quad A(x) = 1 - \sum_{i=1}^{p} a_i x^i \text{ et } B(x) = 1 + \sum_{i=1}^{q} b_i x^i. \tag{3}$$

Les ε_t sont toujours i.i.d. centrés de variance finie σ^2. Les modèles définis par l'équation (3) sont appelés ARMA(p,q). Ils sont définis dans $\mathbb{L}_2$, et connus comme étant stationnaires (au sens $\mathbb{L}_2$ et au sens strict) et inversibles (à condition de faire remonter l'origine des temps jusqu'à $-\infty$) dès que les racines des polynômes A et B sont de module strictement supérieur à 1. Dans ce cas, on sait que la fraction

$R(x) = B(x)/A(x)$ admet un développement en série infinie $R(x) = 1 + \sum_{i=1}^{+\infty} r_i x^i$, de même que la fraction $S(x) = A(x)/B(x) = 1 + \sum_{i=1}^{\infty} b_i x^i$. Ainsi, on peut donner un sens dans $\mathbb{L}_2$ aux représentations dites MA(∞) et AR(∞) respectivement :

$$X_t = \varepsilon_t + \sum_{i=1}^{\infty} r_i \varepsilon_{t-i}, \qquad X_t + \sum_{i=1}^{\infty} s_i X_{t-i} = \varepsilon_t. \tag{4}$$

Cela permet au passage de vérifier que les équations de récurrence sous-jacentes admettent effectivement une solution, qu'elle est unique et strictement stationnaire pour des ε_t i.i.d. Ces modèles sont présentés dans de nombreux ouvrages désormais, comme par exemple Brockwell et Davis (1991), aussi bien du point de vue des propriétés probabilistes que du point de vue des méthodes statistiques utilisées pour estimer les paramètres intervenant dans la représentation des processus. La partie probabiliste est relativement simple et la méthodologie statistique, connue sous le nom de méthode de Box et Jenkins (1970), est parfaitement au point, en théorie comme en pratique.
Du point de vue de la théorie statistique, on a utilisé des méthodes temporelles basées sur les covariances $\gamma_k = \mathrm{cov}(X_t, X_{t+k})$ connues pour être à décroissance rapide, lorsque les racines des polynômes A et B sont de module strictement supérieur à 1 :

$$\gamma_k \sim_{k\to+\infty} C_1 e^{-C_2 k}, \ \ C_1, C_2 \text{ constantes.} \tag{5}$$

On a utilisé également des méthodes basées sur la représentation spectrale des processus stationnaires et en particulier sur la densité spectrale,

$$f(\lambda) = \frac{1}{2\pi} \sum_{k\in\mathbb{Z}} \gamma_k e^{i\lambda k},$$

qui dans le cas de processus modélisés par (3) vaut

$$f(\lambda) = \frac{\sigma^2}{2\pi} \frac{|B(e^{i\lambda})|^2}{|A(e^{i\lambda})|^2}. \tag{6}$$

Notons que sous les conditions imposées à A et B (racines de module > 1), $f(\lambda)$ est bien défini et $f(\lambda) > 0$, pour tout $\lambda \in [0, 2\pi]$. Cette condition est à la base de toute la statistique paramétrique fondée sur la densité spectrale, en plus de quelques autres conditions de régularité standard.

De nombreux développements ont immédiatement suivis pour étudier ce que l'on peut voir mathématiquement comme un cas limite, le cas où 1 est racine du polyôme A, mais qui économiquement était porteur d'un message simple : la représentation cherchée pour la variable était en fait à chercher pour ses accroissements. En effet, si l'on peut factoriser l'opérateur $1 - L$, on voit apparaître un terme $(1 - L)X_t = X_t - X_{t-1}$ qui empêche le modèle d'être stationnaire. En d'autres termes, ce ne sont pas les niveaux mais les flux qui sont significatifs. Ainsi les économètres ont-ils été amenés à accorder une attention toute particulière aux modèles dits ARIMA(p, d, q) (où le I est pour intégré) définis par

$$A(L)(1-L)^d X_t = B(L)\varepsilon_t, \quad d \in \mathbb{N}$$

avec A et B comme dans (3). La présence d'un opérateur $(1-L)^d$ est connue dans la littérature économétrique sous le nom de "racine unité". Si l'on persiste à travailler sur le processus X_t, on est confronté à de nombreux problèmes statistiques dûs à la non-stationnarité du modèle. Néanmoins, avant de différencier le processus (ce qui peut poser des difficultés d'initialisation), encore faut-il statistiquement repérer la présence de la racine problématique : d'où de nombreux développements sur les tests de racine unité, visant à discriminer entre $d = 0$ et $d \neq 0$ pour des d entiers naturels.

3.2. **Généralisations fractionnaires.** Au cours d'hésitations plus prolongées peut-être entre $d = 0$ et $d = 1$, une généralisation des ARIMA est alors naturellement apparue. Concomitamment, Granger et Joyeux (1980) et Hosking (1981) ont proposé de considérer des modèles appelés tantôt ARFIMA tantôt FARIMA, où l'intégration envisagée c'est-à-dire le degré de l'opérateur $(1-L)$ dans la représentation peut être un réel positif et non plus seulement un entier. Ainsi apparaît en économétrie un opérateur pseudo-différentiel discret. La définition probabiliste de ces modèles reste elle encore simple et donne, de façon naturelle un sens dans $\mathbb{L}_2$ à l'opérateur $(1-L)^\alpha$ pour $-1/2 < \alpha < 1/2$ par le biais de son développement en série :

$$(1-L)^\alpha = \sum_{k=0}^{\infty} \frac{\Gamma(k-\alpha)}{\Gamma(k+1)\Gamma(-\alpha)} L^k.$$

Dans ce cas, les racines unités sont qualifiées de fractionnaires. Une des propriétés les plus remarquables des modèles

$$A(L)(1-L)^\alpha X_t = B(L)\varepsilon_t \tag{7}$$

pour des ε_t i.i.d. centrés de variance finie σ^2 et pour $-1/2 < \alpha < 1/2$, est qu'ils restent stationnaires. De plus, ils anéantissent la dichotomie $\alpha = 0$ contre $\alpha = 1$ pour autoriser un continuum de valeurs, ce qui est intellectuellement satisfaisant.
Les modèles ARFIMA disposent d'autres propriétés remarquables, qui ne sont pas étrangères à leur immense succès en économétrie au cours des vingt dernières années. Ces propriétés, de surcroît, sont aussi simples que remarquables, alors même qu'elles sont contradictoires avec le continuum apparent que je viens de mentionner. Lorsqu'on évalue le comportement de leur fonction d'autocovariance, on s'aperçoit que l'équivalent à décroissance exponentielle donné par (5) n'est plus valide; il faut lui substituer une décroissance hyperbolique (arithmétique), dès lors que $\alpha \neq 0$, $-1/2 < \alpha < 1/2$:

$$\gamma_k \sim_{k \to +\infty} C k^{2\alpha - 1}. \tag{8}$$

Cette décroissance lente de la fonction d'autocovariance arrivait à point pour modéliser des phénomènes observés et permet d'illustrer très simplement (en apparence beaucoup plus simplement qu'avec des coefficients de mélange, par exemple, dont la

définition n'est pas très intuitive) l'idée d'une "forte dépendance" ou d'une "mémoire longue". On constate d'ailleurs au vu de (8) que le statut de $\alpha > 0$ et $\alpha < 0$ ne sont pas tout à fait les mêmes car pour $-1/2 < \alpha < 0$, les covariances restent sommables, et donc la densité spectrale est aisée à définir. Le problème est moins aisé pour $0 < \alpha < 1/2$. Ainsi le terme de mémoire longue désigne-t-il en général exclusivement la plage $0 < \alpha < 1/2$, le cas $-1/2 < \alpha < 0$ ayant suscité moins d'intérêt et étant soit ignoré soit qualifié de "mémoire intermédiaire".
Pour ce qui est du comportement spectral, on peut prolonger (6), définir et calculer la densité spectrale pour $\lambda \neq 0$:

$$f(\lambda) = \frac{\sigma^2}{2\pi} \frac{|1 - e^{i\lambda}|^{-2\alpha} |B(e^{i\lambda})|^2}{|A(e^{i\lambda})|^2}.$$

On obtient alors

$$f(\lambda) \sim_{\lambda \to 0} C\lambda^{-2\alpha}, \tag{9}$$

où C est une constante, équivalent qui permet encore de clairement séparer les singularités liées à $\alpha < 0$ (annulation) ou à $\alpha > 0$ (explosion). Ce comportement de la densité spectrale au voisinage de zéro est également associé à la mémoire longue et fournit un autre moyen de la détecter.

3.3. **Développements économiques et statistiques.** Deux types de développement ont immédiatement suivi la définition de cette nouvelle classe de processus : des développements empiriques de chercheurs plutôt économistes qui ont mené des investigations pour déterminer si les séries qu'ils manipulaient couramment pouvaient prétendre à des propriétés de mémoire; des développements théoriques de chercheurs en statistique ayant pour objectif de mettre en place des méthodes d'estimation performantes pour ces modèles. On comprendra que les études des premiers ont fatalement progressé en fonction des propositions des seconds. Ainsi les applications de ces modèles, par exemple en Finance, ne sont-elles apparues qu'une dizaine d'année plus tard (Lo (1991), Backus et Zin (1993), Cheung (1993)) afin d'étudier la présence de mémoire dans des séries de taux d'intérêt ou de taux de change. On verra que ces études se sont intensifiées lorsqu'on s'est attaché à étudier, plutôt que ces variables, leurs variance conditionnelle dans le cadre de modèles dits ARCH. Dans l'intervalle, les années 80 (puis 90) ont vu surtout une multiplication de travaux statistiques : on a signalé que la bornitude inférieure et supérieure avait été une condition de base de l'estimation spectrale, par exemple; on a donc été confronté à un cadre d'hypothèse dans lequel on ne disposait plus d'aucun résultat de convergence ni de normalité asymptotique des estimateurs des paramètres. Geweke et Porter-Hudak (1983) ont proposé un estimateur de α dont la validité théorique n'a été prouvée que bien plus tard par Robinson (1995), qui a par ailleurs réalisé de nombreux travaux fondateurs sur l'estimation semi-paramétrique du paramètre de mémoire (Robinson (1994)). Dans le cadre de modèles paramétriques généraux (i.e. des modèles ARFIMA pour lesquels on souhaite, à partir d'une série d'observation, proposer des estimateurs des coefficients a_i, b_i, α par exemple), on citera essentiellement les résultats de Fox et Taqqu (1986) puis Dahlhaus (1989) qui obtenaient des résultats de consistence et de normalité asymptotiques des estimateurs des paramètres pour des modèles gaussiens de ce type. De nouveaux développements statistiques ont encore lieu actuellement : on le voit, le sujet s'est révélé suffisamment porteur du point de vue statistique pour alimenter déjà vingt ans de recherche. La statistique actuelle se préoccuppant beaucoup de méthodes adaptatives, Iouditsky, Moulines et Soulier (2001) viennent par exemple de proposer une méthode adaptative d'estimation du paramètre de mémoire α. Autrement dit, les opérateurs pseudo-différentiels discrets ont généré des modèles probabilistement simples mais d'une nouveauté statistique et modélisatrice très profonde.

Notons pour finir que les propriétés de mémoire se sont révélées, comme cela semble naturel, fort gênantes pour faire de la prévision : en effet, la décroissance lente de l'autocovariance se retrouve sur la vitesse de décroissance des coefficients de la représentation AR(∞) de type (4), s_i, lorsque i augmente. On voit aisément que cette représentation permet d'évaluer la valeur future de X_{t+1} en fonction des observations des X_t, X_{t-1}, ... Dans le cadre standard de coefficients à décroissance exponentielle, les formules sont tronquées à un nombre de retards faibles et nécessitent l'estimation d'un petit nombre de coefficients, tout en fournissant des prévisions de bonne qualité. Dans le cas longue mémoire, le nombre de coefficients à estimer est considérablement plus élevé, ce qui occasionne davantage d'erreurs d'estimation et génère donc des prévisions de moindre qualité (avec davantage de travail!). Ces problèmes additionnels expliquent néanmoins que les modèles à mémoire longue aient fait couler beaucoup d'encre à de nombreux titres. Ils illustrent aussi que leur simplicité formelle n'est qu'apparente.

3.4. **Le Brownien fractionnaire comme processus sous-jacent.** Il est tout fait naturel que le processus $\varepsilon_t^{(\alpha)}$ défini par (2) soit immédiatement apparu comme lié à ces modèles. En effet, les propriété caractéristiques (8) et (9) avaient déjà été obtenues, on l'a noté, dans le cadre du bruit fractionnaire. Dès 1983, Geweke et Porter-Hudak ont tenté d'établir un lien entre les processus fractionnaires et des processus ARMA sur un bruit fractionnaire, c'est-à-dire des processus ARMA gouvernés par des équation de type (3) mais en substituant $\varepsilon_t^{(\alpha)}$ à ε_t. Il s'agissait finalement d'identifier des relations entre des opérateurs pseudo-différentiels discrets et des opérateurs pseudo-différentiels continus, mais ces considérations ont mis du temps à se décanter. Les économètres ont certainement espéré pendant un temps que l'on pourrait trouver un lien permettant des passer des coefficients $(a_i)_{1\leq i\leq p}, (b_i)_{1\leq i\leq q}$ associés à un modèle ARFIMA(p,d,q) de représentation (7) donnée, aux coefficients $(\tilde{a}_i)_{1\leq i\leq p}, (\tilde{b}_i)_{1\leq i\leq q}$ d'un modèle $\tilde{A}(L)X_t = \tilde{B}(L)\varepsilon_t^{(\alpha)}$ qui lui correspondrait. Précisément, le mélange des opérateurs différentiels discrets $(1-L)^{-\alpha}\varepsilon_t$ et continus $\varepsilon_t^{(\alpha)} = (1-L)I^{(\alpha)}B(t)$ va empêcher les représentations de s'harmoniser. C'est pourquoi Geweke et Porter-Hudak (1983), pour montrer l'analogie entre les deux classes de processus doivent passer par une "équivalence spectrale". Ils définissent un *bruit gaussien fractionnaire général* comme un processus dont la densité spectrale est de la forme $f_1(\lambda,\alpha)f_u(\lambda)$ où

$$f_1(\lambda,\alpha)=\sigma^2(2\pi)^{-2\alpha-3}\Gamma(2\alpha+2)\sin(\pi(\alpha+\frac{1}{2}))\sin^2(\lambda/2)\sum_{n=-\infty}^{+\infty}|n+(\lambda/2\pi)|^{-2\alpha-2}$$

est la densité spectrale du bruit fractionnaire $\varepsilon_t^{(\alpha)}$ et $f_u(\lambda)$ est une fonction positive continue, $0<m<f_u(\lambda)<M$ pour tout $\lambda\in[-\pi,\pi]$, où m,M désignent des constantes positives. Ils définissent par ailleurs un *processus intégré général* comme un processus dont la densité spectrale est de la forme $f_2(\lambda,\alpha)f_u(\lambda)$ où f_u est comme précédemment et

$$f_2(\lambda,\alpha)=(\sigma^2/2\pi)(4\sin^2(\lambda/2))^{-\alpha}$$

est la densité spectrale du processus X défini par $(1-L)^\alpha X_t=\varepsilon_t$. Leur résultat s'énonce alors ainsi : $\{X_t\}$ est un processus général intégré de paramètre α si et seulement si c'est aussi un bruit fractionnaire général de paramètre α. Mais bien sûr, aucun lien explicite entre les coefficients des représentations ne peut être exhibé.

Notons que l'artefact de la différence discrète opérée sur le bruit fractionnaire n'est pas anodine. Pour des classes de processus en temps continu dont la densité spectrale continue ($f_c(\lambda)=\int\gamma_s e^{i\lambda s}ds$) admet au voisinage de 0 un équivalent $f(\lambda)\sim_0 C\lambda^{-2\alpha}$, cette propriété de mémoire n'est stable par discrétisation que pour $\alpha>0$. En effet, la densité spectrale discrète f_d associée à un échantillon discret au pas 1 du processus vaut

$$f_d(\lambda)=\sum_{n\in\mathbb{Z}}f_c(\lambda+2n\pi) \tag{10}$$

ce qui fait disparaître l'annulation pour $\alpha<0$. Cela accentue la dichotomie entre les cas $\alpha>0$ et $\alpha<0$. D'un coté la "vraie" mémoire longue est donc plus stable, de l'autre le cas d'annulation de la densité spectrale associée au processus continue serait bien complexe à traiter. L'idée que suggère le cas du bruit fractionnaire (nécessitée dans ce cas par la non stationnarité du Brownien), est que l'on peut "stabiliser" la propriété de mémoire intermédiaire par la combinaison des opérateurs discrets et continus, c'est-à-dire en appliquant l'opérateur $1-L$ à une intégrale d'ordre $\alpha+1$, $\alpha+1$ étant positif même pour $-1/2<\alpha<0$. Dans ce cas, on discrétise en fait de la mémoire longue, que l'on rend intermédiaire par différentiation.

Le second point sur lequel le Brownien fractionnaire persiste à apparaître concerne les théorèmes limites, qui constituent la seule façon de le faire re-surgir dans le domaine temporel. Le premier résultat à ce propos a été obtenu par Akonom et Gouriéroux (1987) dans un travail resté non publié pour le moment. Ils prouvaient que pour un processus X ARFIMA défini par (7), $X_T^*(r)=(1/T^{\alpha-1/2})X_{[Tr]}, r\in[0,1]$ converge faiblement, quand T tend vers l'infini, vers le processus

$$\sigma\frac{B(1)}{A(1)}\frac{1}{\Gamma(\alpha)}\int_0^r(r-s)^{\alpha-1}dB(s), r\in[0,1].$$

Un peu plus tard, Sowell (1990) a montré la convergence faible, quand N tend vers l'infini, de $\sigma_N^{-1} S_{[Nt]}$ vers

$$W_\alpha(t) = \frac{1}{\Gamma(1+\alpha)} \int_0^t (t-s)^\alpha dB(s) \tag{11}$$

où les S_N sont les sommes partielles d'un processus ARFIMA$(0,\alpha,0)$. L'écart-type $\sigma_N = \sqrt{\mathrm{Var}(S_N)}$ est explicitement calculée et d'ordre $\sqrt{N^{1+2\alpha}}$ pour N grand. De nombreux travaux ont ensuite fait apparaître diverses fonctions ou transformées du Brownien fractionnaire ou de sa version tronquée dans diverses lois limites issus de processus discrets à mémoire longue (Marinucci et Robinson (2000)).

3.5. **L'extension aux modèles ARCH.** [1] Nous utilisons la présentation des modèles ARCH exposée dans Renault (1999). Les études menées sur des séries financières ont assez rapidement amené à la conclusion que les rendements des actifs ne présentaient pas véritablement de propriété de mémoire en niveau, mais plutôt en variance (processus de variance ou bien variance conditionnelle, appelés volatilité). Cette remarque a sans doute été faite dans un contexte discret, où, dans le cadre de modèles dits modèles ARCH dont nous allons dire un mot maintenant, des comportements presque intégrés et néanmoins stationnaires ont été observés, en même temps que des effets de persistence des variations de la volatilité. Supposons donc que ε_t au lieu d'être une suite de variable i.i.d. de variance constante σ^2 s'écrive $\varepsilon_t = \sigma_t z_t$ où les z_t sont i.i.d. centrées de variance 1 (souvent même suposées $\mathcal{N}(0,1)$) et où la variance conditionnelle de ε_t s'écrit :

$$\sigma_t^2 = \psi(1)\sigma^2 + [1-\psi(L)]\varepsilon_t^2, \qquad \psi(L) = 1 - \sum_{j=1}^{\infty} \psi_j L^j.$$

On appelle alors ARCH(p) le modèle obtenu quand les ψ_j sont nuls pour $j > p$, GARCH lorsque les ψ_j sont à décroissance exponentielle, longue mémoire lorsqu'ils sont à décroissance hyperbolique. En particulier si $\psi(L) = A(L)(1-L)^\alpha / B(L)$, on obtient :

$$A(L)(1-L)^\alpha \varepsilon_t^2 = \Theta(L) u_t, \;\; u_t = \varepsilon_t^2 - \sigma_t^2$$

où les u_t sont centrés non corrélés. Ainsi peut-on penser que l'on s'est ramené à des modèles exactement identiques aux modèles ARMA. C'est d'ailleurs ce que les modélisateurs ont considéré après l'introduction des ARCH par Engle (1982) puis leur généralisation GARCH par Bollerslev (1986) et enfin leur version multidimensionnelle par Engle et Kroner (1995). Pourtant, même dans le cas de modèles GARCH(p,q) sans mémoire, l'existence et l'unicité de solutions strictement stationnaires aux équations de récurrence ainsi générées ne sont rien moins que simples (le cas général est un résultat non publié encore de Boussama (1998)), alors que c'est une condition quasiment nécessaire de consistance des estimateurs. Le problème

[1] ARCH = AutoRégressif Conditionnellement Hétéroscédastique et GARCH = ARCH Généralisé.

est que la représentation de type ARMA des ARCH n'est pas utilisable pour un tel objectif car les u_t ne sont pas i.i.d. D'ailleurs, la consistance et la normalité asymptotique des estimateurs de (pseudo) maximum de vraisemblance couramment utilisés par les économètres n'ont été démontrées que très récemment (Boussama (2000) en dimension 1, Comte et Lieberman (2001) en dimension supérieure). On est loin de pouvoir traiter le cas longue mémoire. Après l'article fondateur de Baillie, Bollerslev et Mikkelsen (1996) introduisant les GARCH intégrés d'ordre fractionnaire, l'économétrie de la Finance s'est néanmoins emparée de ces modèles pour en faire un usage intensif.
Pourtant, les modèles de la Finance sont très souvent en temps continu, fondés sur des bruits browniens. Il est donc naturel que la boucle se soit bouclée là et le brownien fractionnaire, de sous-jacent, soit redevenu prépondérant.

4. L'unification du continu : opérateurs pseudo-différentiels appliqués à la finance

4.1. **Premiers modèles à temps continus.** Comte et Renault (1996) ont introduits les premiers modèles à temps continu et à mémoire longue en considérant des généralisation de processus de type moyenne mobile $X(t) = \int_0^t A(t-s)dB(s)$ à des intégrales par rapport au Brownien fractionnaire tronqué donné par (11):

$$X(t) = \int_0^t A(t-s)dW_\alpha(s). \tag{12}$$

Les seules fonctions A envisagées étaient déterministes et les définitions posées au sens $\mathbb{L}_2$ car la théorie de l'intégration stochastique à ce moment-là n'autorisait que des intégrales de processus par rapport à des semi-martingales et le Brownien fractionnaire était connu pour ne pas en être une. De nombreux développements théoriques ultérieurs ont rendu très primaires ces précautions (Alos, Mazet, Nualart (1999), Carmona et Coutin (1999), Coutin et Decreusefond (2000)), mais la rigueur nous imposait alors la prudence. Nous avons alors fait usage des opérateurs pseudo différentiels d'intégration

$$I_0^{(\alpha)}(X)(t) = \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} X(s)ds \tag{13}$$

pour $\alpha > 0$ et de dérivation pour $\alpha < 0$:

$$D_0^{(|\alpha|)}(X)(t) = I_0^{(\alpha)}(X)(t) = \frac{d}{dt}\left[\int_0^t \frac{(t-s)^{\alpha}}{\Gamma(1+\alpha)} X(s)ds\right], \tag{14}$$

avec des notations conformes au Samko et al. (1993). On peut alors calculer que $D_0^{(\alpha)}(X)(t) = \int_0^t A(t-s)dB(s)$ pour $X(t)$ défini par (12). Cette remarque permet d'unifier les représentations et de résoudre élégamment dans ce cadre le problème de Geweke et Porter-Hudak (1983) : la même fonction A intervient dans la représentation Moyenne Mobile de X sur un Brownien fractionnaire ou dans la représentaion Moyenne Mobile de la *dérivée* d'ordre α de X sur un Brownien

standard. Ainsi peut-on vérifier que sous des hypothèses simple sur la fonction A, le processus défini par (12) est bien à mémoire longue (lui ou son équivalent stationnaire) du point de vue du comportement de sa fonction de covariance ou de sa densité spectrale continue. En revanche, sa dérivée fractionnaire est bien courte mémoire. De plus, l'exemple simple de l'"Ornstein-Uhlenbeck fractionnaire" associé à $A(x) = e^{-kx}\sigma$, c'est-à-dire un processus moralement solution de $dX(t) = -kX(t)dt+\sigma dW_\alpha(t)$ et dont la dérivée fractionnaire est solution de $dX(t) = -kX(t)dt+ \sigma dB(s)$ permet de vérifier le rapport entre les opérateurs discrets et continus : $D_0^{(\alpha)}(X)(t)$ est à temps discret un AR(1) (sans approximation), $(1-L)^\alpha X(t)$ n'est plus à mémoire longue, donc les deux opérateurs éliminent la mémoire longue. Mais les coefficients court terme (l'écriture ARMA) des deux processus sont sans rapport, ceux du premier sont très simples (un seul coefficient non-nul), ceux du seconds présentent par exemple des corrélogrammes à décroissance beaucoup moins vertigineuse que le premier, ce qui est révélateur d'une représentation ARMA complexe. L'homogénéité de la représentarion en continu est donc importante pour simplifier la manipulations des processus fractionnaires, en particulier lorsqu'intervient une étape de modélisation.
Pour finir, signalons que cette fois, le discret est sous-jacent, car les observations ont toujours été considérées comme émanant de relevés à temps discret, de sorte que la partie statistique de ces modèles est traitée en utilisant des résultats existant, une majeure partie du travail consistant à vérifier que les discrétisés de ces modèles vérifiaient les hypothèses requises. Il est heureux que le cas considéré comme intéressant en général soit bien le cas $\alpha > 0$, car c'est le seul qui soit stable par discrétisation, comme on l'a remarqué avec la formule (10) pour un pas unitaire. Sinon, il faudrait considérer les différences discrètes d'un processus fractionnaire d'ordre $\alpha + 1$.

4.2. **Applications à la Finance.** Un champ d'application naturel de ces modèles, qui constituait une partie de la motivation du passage au temps continu, était la modélisation financière. Les balbutiements de la Finance stochastique ont débuté avec le modèle de Black et Scholes (1973) qui modélisait le prix S_t d'un actif par la dynamique

$$dS_t = S_t(\mu dt + \sigma dB_t), \tag{15}$$

où μ et σ sont des constantes, σ étant connue sous le nom de "volatilité" constante. Le succès mondial de ce modèle lui permet de fixer universellement les prix des actifs et des options sur ces actifs. Pourtant, la tentation de faire varier la volatilité, supposée constante dans le modèle, s'est très vite imposée; ainsi a-t-on utilisé les prix des options pour re-calculer des volatilités dites implicites, que l'on a ensuite éventuellement ré-injectées dans l'équation (15) affectée de coefficients réactualisés. Les théoriciens ont donc cherché à rendre compte de cet aspect variable, soit en considérant une fonction de volatilité (une volatilité fonction déterministe du temps), soit en considérant des volatilités aléatoires, elles-mêmes gouvernées par une équation propre. On a alors etudié le modèle de Hull et White (1987) :

$$\left\{ \begin{array}{l} dS(t)/S(t) = \mu(t, S(t))dt + \sigma(t)dB^{(1)}(t) \\ \sigma^2(t) = f(\eta(t)) \\ d\eta(t) = m(\eta(t))dt + g(\eta(t))dB^{(2)}(t) \end{array} \right. \tag{16}$$

où $(B^{(1)}, B^{(2)})$ est un mouvement brownien bi-dimensionnel standard, $S(t)$ le prix de l'actif sous-jacent, $\sigma(t)$ la volatilité instantanée et f une fonction positive. En particulier, Renault et Touzi (1998) (après les travaux précurseurs de Scott (1987) et Melino et Turnbull (1990)) ont étudié ces modèles et ont mis en évidence qu'il permettait d'apporter la justification de nombreux faits observés.
Il n'y avait pas de raison d'introduire des opérateurs pseudo-différentiels ni des browniens fractionnaires dans la première équation car aucun phénomène de mémoire n'avait été observé sur les prix d'actifs. Il y avait même des raisons de ne pas le faire puisque cela aurait créé des opportunités d'arbitrage (possibilité de gagner de l'argent sans investir) ainsi que le démontra peu après Rogers (1997) : or, on cherche à éviter que les modèles autorisent ce genre de phénomène. En revanche, des comportements de persistence avaient été détectés dans les volatilités, ceux-là mêmes qui ont inspiré simultanément l'émergence des modèles FIGARCH présentés précédemment. Nous avons donc très naturellement introduit dans Comte et Renault (1998) le modèle :

$$\left\{ \begin{array}{l} dS(t)/S(t) = \mu(t, S(t))dt + \sigma(t)dB^{(1)}(t) \\ d\ln(\sigma(t)) = k(\theta - \ln(\sigma(t))dt + \gamma dW_\alpha^{(2)}(t) \end{array} \right. \tag{17}$$

avec $W_\alpha^{(2)}(t) = \int_0^t \frac{(t-s)^\alpha}{\Gamma(1+\alpha)} dB^{(2)}(t)$.
L'équation sur $\ln(\sigma)$ définit encore un processus d'Ornstein-Uhlenbeck, mais gouverné par un Brownien fractionnaire au lieu d'un Brownien standard, et le passage au ln garantit la positivité de σ. Notons que la solution de l'équation pouvait être calculée "à la main" sans prendre le risque d'effleurer une théorie encore inexistante concernant les équations différentielles stochastiques fractionnaires, développée ultérieurement par Carmona et Coutin (1999), Alos, Mazet et Nualart (1999). Mieux, la linéarité de l'équation évitait d'avoir à choisir entre processus solution d'un équation fractionnaire et processus intégrale fractionnaire de la solution d'une Equation Différentielle Stochastique (EDS). Ainsi, à condition de s'autoriser quelques allers-retours entre une version non-stationnaire et une version stationnaire du processus de volatilité dont la légitimité paraît rétrospectivement quelque peu périlleuse, avons-nous mis en place et justifié un modèle d'actifs à volatilité stochastique et fractionnaire, qui laissait en apparence inchangée la formule de prix d'option et qui permettait d'illustrer les phénomènes de mémoire longue dans les volatilités. Le nombre de paramètres requis était plutôt moindre que pour les ARCH fractionnaires et la rigueur de la définition plus assurée. Du point de vue statistique, l'utilisation des résultats à temps discret obtenus par Dahlhaus (1989) et Robinson (1994) pouvaient s'appliquer très rigoureusement. Soulignons qu'une difficulté

statistique liée à cette gamme de modèles réside dans le fait que seul le processus S est observé, alors qu'il y a deux sources de bruit. Les problèmes mal résolus résidaient dans le choix entre opérateurs tronqués ou complets (i.e. entre processus non stationnaires ou stationnaires), les méthodes de discrétisation qui résultaient de schéma d'Euler naïfs pour lesquels on ne disposait d'aucune vitesse de convergence, et l'aspect "pratique" des formules de prix d'option qui ne pouvaient plus ici s'appuyer sur des propriétés markoviennes pour simplifier les calculs et les rendre opérationnels. Le modèle suivant résout les deux premières questions.

4.3. **Un modèle rigoureux s'appuyant sur des opérateurs de dérivation fractionnaire.** Dernièrement, les modélisations de la volatilité se sont orientées vers des équations qui ont été introduites à l'origine par Cox, Ingersoll et Ross (1985) (CIR) pour modéliser des dynamiques de taux d'intérêt. L'idée est toujours de trouver une équation simple (c'est-à-dire nécessitant un faible nombre de paramètres) admettant une solution positive et fournissant une dynamique raisonnable pour le processus de volatilité. Cela conduit au modèle particulier de type Hull et White (1987) suivant :

$$\left\{ \begin{array}{l} dS(t)/S(t) = \mu(t, S(t))dt + \sigma(t)dB^{(1)}(t) \\ d\sigma^2(t) = k(\tilde{\theta} - \sigma^2(t))dt + \gamma\sigma(t)dB^{(2)}(t) \end{array} \right. \tag{18}$$

L'équation de la volatilité n'est simple qu'en apparence puisque ses coefficients ne vérifient par les conditions standard d'existence et d'unicité de solution d'EDS (simplement car $f(x) = \gamma\sqrt{x}$ n'est pas lispchitzienne). Néanmoins, il a été démontré (voir Revuz et Yor (1999)) qu'à condition initiale positive dnnée, le processus solution existait, était unique et presque sûrement positif. L'intérêt qui a été trouvé à ce modèle est qu'il permet d'exprimer simplement les moments de tous ordres de la volatilité intégrée:

$$\frac{1}{T-t}\int_t^T \sigma^2(u)du.$$

Cette quantité intéresse tout particulièrement les financiers parce que, en vertu de la première équation du modèle, elle est le "paramètre" qui remplace la volatilité constante de Black et Scholes (1973) dans la formule des prix d'options. De plus, les praticiens considèrent souvent que s'ils disposent de données hautes fréquences (relevés toutes les dix minutes sur un an par exemple), ils atteignent l'asymptotique et disposent de la valeur de $\int_t^T \sigma^2(u)du$ qui représente la variation quadratique de $R_t = \ln(S(t))$, vérifiant

$$\int_t^T \sigma^2(u)du = P\lim_{|t_{i+1}-t_i|\to 0} \sum (R_{t_{i+1}} - R_{t_i})^2, \tag{19}$$

où Plim désigne la limite en probabilité et où les t_i constituent une partition de l'intervalle $[t, T]$ considéré. On dispose donc ainsi d'une pseudo-observation de la variable non observée. Enfin, la variance conditionnelle des rendements des actifs $V_t(\ln(S_{t+h}/S_t))$ mesure l'hétéroscedasticité du processus auquel on s'intéresse et, pour des h petits, est du même ordre que $\mathbb{E}_t\left(\int_t^{t+h}\sigma(u)^2du\right)$. On a donc cherché à introduire de la mémoire dans ce modèle en conservant les propriétés particulières dues à l'équation CIR. On a donc étudié dans Comte, Coutin, Renault (2001) un modèle faisant cette fois directement et explicitement appel à un opérateur pseudo différentiel :

$$\left\{\begin{array}{l} dS(t)/S(t)=\mu(t,S(t))dt+\sigma(t)dB^{(1)}(t)\\ \sigma(t)=\sqrt{|Y(t)|},\\ Y(t)=\theta+I^{(\alpha)}(\widetilde{\sigma}^2(t)-\tilde{\theta}),\\ d\widetilde{\sigma}^2(t)=k(\tilde{\theta}-\widetilde{\sigma}^2(t))dt+\gamma\widetilde{\sigma}(t)dB^{(2)}(t)\end{array}\right. \tag{20}$$

où l'on s'intéresse à $\alpha>0$ et où l'on a même envisagé de la corrélation entre les Browniens (corrélation connue sous le nom d'effet “de levier”). L'opérateur $I^{(\alpha)}$ d'intégration fractionnaire d'ordre α, pour $0<\alpha<1$ est cette fois défini par :

$$I^{(\alpha)}(X)(t)=\int_{-\infty}^t\frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)}X(s)ds \tag{21}$$

Pour $-1<\alpha<0$, la dérivation se définit de façon analogue par

$$I^{(\alpha)}(X)(t)=\frac{d}{dt}\left[\int_{-\infty}^t\frac{(t-s)^{\alpha}}{\Gamma(1-\alpha)}X(s)ds\right]. \tag{22}$$

Cet opérateur est défini pour les processus au sens $\mathbb{L}^2$ (c'est-à-dire que l'on vérifie que $\mathbb{E}[(I^{(\alpha)}(X)(t))^2]<+\infty$) et doit par conséquent être appliqué à un processus centré. Cet opérateur a le mérite de conserver la stationarité du processus sous-jacent, contrairement à l'opérateur $I_0^{(\alpha)}(X)(t)=\int_0^t\frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)}X(s)ds$ qui, lui, sacrifierait la stationarité en conservant la positivité (c'est-à-dire si l'on avait défini $\sigma^2(t)=I_0^{(\alpha)}(\widetilde{\sigma}^2(.))(t)$). Pour des valeurs raisonnables des paramètres, on vérifie par le biais de simulations que la probabilité qu'un tel processus passe par des valeurs négatives est extrèmement faible (non observée pour des lots de 100 trajectoires de très grande longueur). C'est la raison pour laquelle on note abusivement $\sigma^2(t)=Y(t)$.

Notons que ce modèle aurait eu une alternative consistant à substituer au Brownien $B^{(2)}$ de l'équation (18) un Brownien fractionnaire; le modèle n'étant plus linéaire, cette solution ne coïncide a priori plus avec celle retenue. Cela aurait eu plusieurs inconvénients. Le premier est que, comme dans le cas Brownien standard, l'équation ne vérifie pas les conditions d'existence et d'unicité des solutions données par Coutin et Decreusefond (2000) et que l'on n'est pas en mesure aujourd'hui de démontrer qu'une telle équation admet une unique solution, ni d'étudier son signe. Le second problème est que cette extension aurait détruit les bonnes propriétés de l'équation CIR. Enfin, cela aurait rendu complexe le moindre calcul concernant ce processus de sorte que nous aurions eu beaucoup de mal à justifier les propriétés du nouveau modèle.

Ici, on constate simplement que l'on obtient très rigoureusement des propriétés de longue mémoire pour σ^2. Ainsi, de façon très générale, si A est un processus stationnaire courte mémoire (c'est-à-dire admettant une densité spectrale f_A telle que $0 < m < f_A(\lambda) < M$ pour tout λ réel et deux constantes m, M), et $B = I^{(\alpha)}(A)$, alors B est un processus à mémoire longue d'ordre α au sens où

$$f_B(\lambda) = \lambda^{-2\alpha} f_A(\lambda).$$

Plus spécifiquement, si l'on calcule $c_{\sigma^2}(h)$, la fonction d'autocovariance de $\sigma^2 = Y$, en exploitant les propriétés de $\tilde{\sigma}^2$ en plus de celles de l'opérateur $I^{(\alpha)}$, on obtient :

(1) $c_{\sigma^2}(h) = c_{\sigma^2}(0) + Ch^{2\alpha+1} + o(h^2)$, $h \to 0$,
(2) $c_{\sigma^2}(h) \sim C' h^{2\alpha-1}$ quand $h \to +\infty$,

où C et C' sont explicitement connues en fonction des paramètres. Notons que l'ordre $\alpha+1/2$ qui apparaît dans *1.* est caractéristique de la régularité des trajectoires du processus, alors que l'ordre $2\alpha - 1$ de *2.* mesure l'ordre de la mémoire longue. Comme on l'a déjà remarqué, on a perdu les propriétés d'autosimilarité du Brownien fractionnaire. Signalons de plus que les propriétés de la volatilité se répercutent sur les volatilités intégrées et les volatilités implicites approximées au premier ordre (voir Bandi et Perron (2001)) par

$$\sigma^2_{imp}(t) = \int_0^1 \mathbb{E}(\sigma^2(t+u)|\mathcal{F}_t)du$$

où $\mathcal{F}_t$ est la tribu engendrée par le passé du Brownien $(B^{(1)}(s), B^{(2)}(s))$ pour $s \le t$. Enfin, il est utile de remarquer que pour $Z_t = \int_t^{t+1} \sigma^2(s)ds$,

$$I^{(-\alpha)}(Z-\theta)(t) = \frac{d}{dt}\int_{-\infty}^t \frac{(t-s)^{-\alpha}}{\Gamma(1-\alpha)}(Z(s)-\theta)ds = \int_t^{t+1} (\tilde{\sigma}^2(s) - \tilde{\theta})ds,$$

ce qui permet de se ramener au processus standard sans mémoire dès lors que l'on sait mettre en oeuvre ou approximer l'opérateur de dérivation fractionnaire. Il faut insister sur le fait que l'équation CIR permet d'exprimer simplement tous les moments de $\tilde{\sigma}^2$, ce qui est très utile d'un point de vue statistique (utilisation des moments empiriques pour estimer les paramètres).

4.4. **Utilisation de représentations diffusives pour la discrétisation.** Pour étudier le modèle, on s'intéresse alors à la discrétisation d'intégrales fractionnaires, que ce soit la discrétisation de l'opérateur d'intégration fractionnaire en vue de simuler le modèle proposé (et tester les méthodes statistiques), ou bien la discrétisation l'opérateur de dérivation fractionnaire lorsqu'on va souhaiter mettre en place une méthode d'estimation. Ainsi est-on amené à exploiter des résultats de Carmona, Coutin et Montseny (2000) et de Coutin et Pontier (2001), qui ont étudié ce problème dans divers cadres. Leurs travaux amènent à considérer que l'intégrale :

$$I^{(\alpha)}(f)(t) = \int_{-\infty}^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} f(s)ds$$

pour une fonction f donnée est convenablement approximée par

$$I_0^{(\alpha)}(f)(t) = \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} f(s)ds$$

pour laquelle, par Laplace inverse:

$$(t-s)^{\alpha-1} = \frac{1}{\Gamma(1-\alpha)} \int_0^\infty x^{-\alpha} e^{-x(t-s)} dx.$$

En appliquant alors un théorème de Fubini à l'intégrale temps-fréquence, on obtient :

$$I_0^{(\alpha)}(f)(t) = \frac{1}{\Gamma(\alpha)\Gamma(1-\alpha)} \int_0^\infty x^{-\alpha} \left(\int_0^t e^{-x(t-s)} (\tilde{\sigma}^2(s) - \tilde{\theta}) ds \right) dx.$$

Cette représentation, est une représentation diffusive de l'intégrale fractionnaire et a été utilisée par Carmona, Coutin et Montseny (2000) pour écrire, par analogie avec ce qui est fait en Physique :

$$I_0^{(\alpha)}(f)(t) = \frac{1}{\Gamma(\alpha)\Gamma(1-\alpha)} \int_0^\infty x^{-\alpha} \Psi(x,t,f) dx, \tag{23}$$

où $\Psi(x,t,f)$ est un opérateur linéaire aussi étudié par Coutin et Pontier (2001), défini par :

$$\Psi(x,t,f) = \int_0^t e^{-x(t-s)} f(s) ds \tag{24}$$

et vérifie $\partial\Psi(x,t,f)/\partial t = -x\Psi(x,t,f) + f(t)$. On utilise alors une subdivision géométrique en fréquence (variable x) et un schéma d'Euler standard en temps (variable t) pour justifier le schéma global suivant. Supposons choisis $r \in [1,2]$, un entier n et un pas Δ,

$$(\sigma^2)^{r,n,\Delta}(t_j) = \theta + \frac{1}{\Gamma(\alpha)\Gamma(1-\alpha)} \sum_{i=-n}^{n-1} c_i \Psi^\Delta(\eta_i, t_j, \tilde{\sigma}^2 - \tilde{\theta})$$

où c_i, η_i sont donnés par

$$\begin{cases} c_i &= \displaystyle\int_{x_i}^{x_{i+1}} x^{-\alpha} dx = \frac{(r^{1-\alpha}-1)}{1-\alpha} r^{(1-\alpha)i}, \\ \eta_i &= \displaystyle\frac{1}{c_i} \int_{x_i}^{x_{i+1}} x^{1-\alpha} dx = \frac{1-\alpha}{2-\alpha} \frac{r^{2-\alpha}-1}{r^{1-\alpha}-1} r^i, \end{cases}$$

et Ψ^Δ est donné par

$$\begin{aligned} \Psi^\Delta(x, t_{j+1}, f) &= \Psi^\Delta(x, t_j, f) e^{-x\Delta} + f(t_j) \frac{1-e^{-x\Delta}}{x} \\ \Psi^\Delta(x, t_0, f) &= 0. \end{aligned}$$

Alors on dispose d'un schéma récursif efficace, dont on peut aussi évaluer la vitesse de convergence vers le vrai processus en fonction de r, n, Δ (précisément $(r-1)^2(1+\Delta^\beta) + r^{-\alpha n}$).

Ainsi peut-on mettre en place stratégie pour l'inférence statistique en plusieurs temps, et la tester par le biais de simulations.

(1) Les données hautes fréquences permettent d'estimer des intégrales $\int \sigma^2(u)du$ sur de petits intervalles. Donc on extrait $Z_t = \int_t^{t+1} \sigma^2(u)du$ des observations par variation quadratique (voir formule (19)).
(2) On estime $\theta = \mathbb{E}(Z_t)$ par $\hat{\theta} = (1/n)\sum_{i=1}^n Z_i$ la moyenne empirique.
(3) Z est stationnaire et à mémoire longue. Sa densité spectrale vérifie les hypothèses des résultats de Robinson (1994) ou Fox et Taqqu (1986), Dahlhaus (1989), sauf la gaussianité. On propose donc d'estimer α via $\hat{\alpha}$ issu d'une méthode (avec éventuellement d'autres paramètres) de minimum de contraste de Whittle ou régression log-périodogramme. On mène une étude par simulation pour vérifier que ces techniques sont robustes à la non-gaussianité. Des résultats théoriques sont en cours d'élaboration concernant ce problème mais en dehors du cadre longue mémoire (Velasco (2000)).
(4) Comme,
$$I^{(-\alpha)}(Z-\theta)(t) = \int_t^{t+1} (\tilde{\sigma}^2(u) - \tilde{\theta})du$$
on utilise $\hat{\theta}$ et $\hat{\alpha}$ et une approximation de l'opérateur $I^{(-\alpha)}$ pour avoir des valeurs estimées de
$$F(t) = \int_t^{t+1} (\tilde{\sigma}^2(s) - \tilde{\theta})ds.$$
(5) On utilise les expressions connues des moments de $\tilde{\sigma}^2 - \tilde{\theta}$ pour estimer k, γ, $\tilde{\theta}$ par des moments empiriques.

On obtient ainsi une procédure d'estimation complète que l'on peut tester par le biais de données simulées, avant l'épreuve (encore non atteinte) des données réelles.

5. Conclusion

En conclusion, on voit que l'histoire s'achève (pour l'instant) comme elle a commencé, avec des opérateurs pseudo-différentiels continus; entre temps, les économètres se sont convaincus de leur utilité par le biais de leur version discrète. La classe de modèles discrets ainsi générés a nécessité la mise en place de techniques statistiques spécifiques qui, en retour, ont pu s'appliquer aussi bien à ces modèles directement que plus tard à des échantillons discrets issus des modèles continus.

On peut souligner par ailleurs que les modèles de Finance considérés ne sauvegardent que deux des trois grandes caractéristiques du bruit fractionnaire originel, et que les modèles discrets n'en sauvegardaient qu'une seule, appauvrissant ainsi considérablement l'exemple de départ : on conserve dans les deux cas la mémoire longue (covariance d'ordre k d'ordre $k^{2\alpha-1}$ pour k grand et densité spectrale en $\lambda^{-2\alpha}$ notamment au voisinage de 0) et dans le cas des modèles continus seulement la régularité $\beta < \alpha + 1/2$-Hölder des trajectoires, mais on perd dans tous les cas la

$\alpha - 1/2$-autosimilarité du processus. Cette propriété paraît difficile à maintenir dans des modèles de ce type, mais il n'est pas facile de savoir s'il est souhaitable ou non de chercher à la ré-introduire : voir Mandelbrot (1997).
D'autres développements du côté de la finance fractionnaire, sont à signaler. Des travaux de Hu et Oksendal (2000) et Hu, Oksendal et Sulem (2000) s'attachent à définir une nouvelle forme d'intégration fractionnaire permettant de donner un sens à une équation de Black et Scholes généralisée qui s'écrirait : $dS_t/S_t = \mu dt + \sigma dB^{(\alpha)}(t)$, malgré les objections qui ont pu être formulée sur ce modèle. Ils montrent que la définition nouvelle qu'ils proposent permet d'éviter les opportunités d'arbitrage (convenablement redéfinies). Il serait intéressant de savoir s'il est possible d'estimer le paramètre qu'ils introduisent de la sorte. On peut citer certaines directions multi-fractionnaires (Ayache et Levy-Vehel (1999), Benassi et al. (2000), N'Doye (2001)) très liées à des méthodes de traitement d'image, peut-être davantage qu'à de la modélisation aléatoire. Le mouvement Brownien multifractionnaire est défini comme un mouvement Brownien fractionnaire dont le paramètre α varierait en fonction du temps et serait remplacé par une fonction $\alpha(t)$. L'idée est intéressante et donnera peut-être lieu à des applications porteuses, mais plusieurs difficultés surgissent : le processus de base devient tellement complexe qu'il sert pour l'instant de "modèle" à lui seul, le paramètre α lorsqu'il était constant posait déjà de tels problèmes statistiques qu'il est probable qu'une estimation de type non-paramétrique pour évaluer la fonction $\alpha(t)$ sera délicate. De nombreuses études restent donc à mener dans ces directions.

References

[1] AKONOM, J. and GOURIÉROUX, C. (1987). A functional limit theorem for fractional processes. *Working paper* 8801.

[2] ALOS, E., O. MAZET and D. NUALART (1999) Stochastic Calculus with respect to fractional brownian motion with Hurst parameter less than 1/2. *Working paper.*

[3] AYACHE, A., J. LEVY VEHEL (1999) Generalized multifractional Brownian motion: Definition and preliminary results. Dekking, Michel (ed.) et al., Fractals: Theory and applications in engineering, London: Springer, 17–32.

[4] BACKUS, D.K. and S.E. ZIN (1993): "Long-memory inflation uncertainty: evidence from the term structure of interest rates," *Journal of Money, Credit, and Banking*, 3, 681–700.

[5] BAILLIE, R.T., T. BOLLERSLEV and H.O. MIKKELSEN (1996): "Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity," J. Economet., **74**, 3–30.

[6] BANDI, F.M. and B. PERRON (2001): "Long memory and the relation between implied and realized volatility". *working paper.*

[7] BARDET, J-M. (2000a) Testing for the presence of self-similarity of Gaussian time series having stationary increments. *J. of Time Ser. Anal.* **21**, 497–515.

[8] BARDET, J-M. (2000b) Les cours d'actifs financiers sont-ils autosimilaires? *Journal de la SFDS* **141**, 137–148.

[9] BERAN, J. (1994) *Statistics for long memory processes.* Monographs on Statistics and applied Probability, 61. Chapman and Hall, New York.

[10] BENASSI, A., P. BERTRAND, S. COHEN and J. ISTAS (2000) Identification of the Hurst index of a step fractional Brownian motion. *Stat. Inference Stoch. Process.* **3**, 101–111.

[11] BLACK, F. et M. SCHOLES (1973) The pricing of options and corporate liabilities. *Journal of Political Economy* **3**, 637–654.
[12] BOLLERSLEV, T. (1986) Generalized autoregressive conditional heteroskedasticity, *J. Economet.* **31**, 307–327.
[13] BOLLERSLEV, T. et H.O. MIKKELSEN (1996) Modelling and pricing long memory in stock market volatility. *J. Econometrics* **73**, 151–184.
[14] BOUSSAMA, F. (1998) Ergodicity, Mixing and Estimation in GARCH Models, Ph.D. dissertation, University Paris 7.
[15] BOUSSAMA, F. (2000) Asymptotic normality for the quasi-maximum likelihood estimator of a GARCH model, *C. R. Acad. Sci. Paris* Série I 331, 81–84.
[16] BOX, G. et G. JENKINS (1970) *Time series analysis: forecasting and control.* Holden Day.
[17] BROCKWELL, P.J. et R.A. DAVIS (1991) *Time series: theory and methods.* Second edition. Springer.
[18] CARMONA, P. and COUTIN, L. (1999) Stochastic integration with respect to fractional Brownian motion. *Working paper.*
[19] CARMONA, P., L. COUTIN and G. MONTSENY (2000) Approximation of some gaussian processes. *Statistical Inference for Stochastic Processes* **3**, 161–171.
[20] CHEUNG, Y-W. (1993) Long memory in foreign exchange rates. *Journal of Business and Economic Statistics* **11**, 93–101.
[21] COMTE, F. et E. RENAULT (1998) Long memory in continuous time stochastic volatility models. *Mathematical Finance* **8**, 291–323.
[22] COMTE, F., L. COUTIN et E. RENAULT (2001) Affine Fractional Stochastic Volatility Models. *Working paper.*
[23] COMTE, F. et O. LIEBERMAN (2001) Asymptotic Theory for Multivariate GARCH Processes. *Journal of Multivariate Analysis*, to appear.
[24] COUTIN, L. et L. DECREUSEFOND (2000) Volterra differential equations with singular kernels. *Proceedings on the Workshop on Mathematical and Stochastic analysis*, Cruzeiro, A.B. and Zambrini, J.C. Birkhauser.
[25] COUTIN, L. et M. PONTIER (2001) Approximation d'un drap brownien fractionnaire. *Working paper.*
[26] COX, J., J. Jr INGERSOLL et S. ROSS (1985) An intertemporal general equilibrium model of asset prices. *Econometrica* **53**, 363–384.
[27] DAHLHAUS, R. (1989) Efficient parameter estimation for self-similar processes. *Annals of Statistics* **17**, 1749–1766.
[28] ENGLE, R. F. (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation, *Econometrica* **50**, 987–1007.
[29] ENGLE, R.F. et K.F. KRONER (1995) Multivariate simultaneous generalized ARCH. *Economet. Theory* **11**, 122–150.
[30] FOX, R. et M.S. TAQQU (1986) Large sample properties of parameter estimates for strongly dependent time series. *Annals of Statistics* **14**, 517–532.
[31] GEWEKE, J. et S. PORTER-HUDAK (1983) The estimation and application of long memory time series models. *J. of Time Ser. Anal.* **4**, 221–238.
[32] GRANGER, C.W.J. et R. JOYEUX (1980) An introduction to long memory time series models and fractional differencing. *J. of Time Ser. Anal.* **1**, 15–29.
[33] GUYON, X. et J. LEON (1989) Convergence en loi des H-variations d'un processus gaussien stationnaire sur $\mathbb{R}$. *Ann. Inst. Henri Poincaré* **3**, 265–282.
[34] HARVEY, A.C. (1993) Long Memory in Stochastic Volatility. *Working paper of the London School of Economics.*
[35] HOSKING, J.M.R. (1981) Fractional differencing. *Biometrika* **68**, 165–176.

[36] HU, Y. et B. OKSENDAL (2000) Fractional white noise calculus and applications to Finance. *Working paper.*

[37] HU, Y., B. OKSENDAL et A. SULEM (2000) Optimal consumption and portfolio in a Black-Scholes market driven by fractional Brownian motion. *Working paper 28*, Department of Mathematics, University of Oslo.

[38] HULL, J. et A. WHITE (1987) The pricing of options on assets with stochastic volatilities. *The Journal of Finance* **3**, 281–300.

[39] HURST, H.E. (1951) Long term storage capacity of reservoirs. *Transactions of the American Society of Civil Engeneers* **116**, 770–799.

[40] IOUDITSKY, A., E. MOULINES et P. SOULIER (2001). Adaptive estimation of the fractional differencing coefficient. *Bernoulli* **7**, 699–731.

[41] ISTAS, J. et G. LANG (1997) Quadratic variations and estimation of the local Hölder index of a Gaussian process. *Ann. Inst. Henri Poincaré* **4**, 407–436.

[42] LARDIC, S. et V. MIGNON (1999) La mémoire longue en économie : une revue de la littérature. *Journal de la SFdS*, 140-2.

[43] LO, A.W. (1991) Long term memory in stock market prices. *Econometrica* **59**, 1279–1313.

[44] MANDELBROT, B.B. (1997) *Fractals and scaling in finance. Discontinuity, concentration, risk.* Springer, New York.

[45] MANDELBROT, B.B. et J.W. VAN NESS (1968) Fractional Brownian motions, fractional noises and applications. *SIAM Review* **10**, 422–437.

[46] MANDELBROT, B.B. et R. WALLIS (1969) Computer experiments with fractional gaussian noises. *Water Resources Research* **5**, 228–267.

[47] MARINUCCI, D. et P.M. ROBINSON (2000). Weak convergence of multivariate fractional processes. *Stochastic Processes and their applications* **81**, 103–120.

[48] MELINO, A. et S.M. TURNBULL (1990) Pricing foreign currency options with stochastic volatility. *J. Econometrics* **45**, 239–265.

[49] NDOYE, M. (2001) Intégrale stochastique par rapport à un Brownien multifractionnaire et applications en Finance. Thèse de l'Université Paris IX.

[50] RENAULT, E.(1999) Discussion de : La mémoire longue en économie : une revue de la littérature. *Journal de la SFdS*, 140-2, 79–89.

[51] RENAULT, E. et N. TOUZI (1996) Option hedging and implicit volatilities in a stochastic volatility model. *Mathematical Finance* **6**, 279–302.

[52] REVUZ, D. et M. YOR (1999) *Continuous Martingales and Brownian Motion.* Springer-Verlag, Third edition.

[53] ROBINSON, P.M. (1994) Semiparametric analysis of long-memory time series. *Annals of Statistics* **22**, 515–539.

[54] ROBINSON, P.M. (1995) Log periodogram Regression of Time Series with Long Range Dependence. *Annals of Statistics* **23**, 1630–1660.

[55] ROGERS, L.C.G. (1997) Arbitrage with Fractional Brownian Motion. *Mathematical Finance* **7**, 95–105.

[56] SAMKO, S.G., KILBAS, A.A. and MARICHEV, O.I. (1993) *Fractional Integrals and Derivatives, Theory and Applications.* Gordon and Breach Science Publishers, Amsterdam.

[57] SCOTT, L. (1987) Option pricing when the variance changes randomly: estimation and an application. *Journal of Financial and Quantitative Analysis* **22**, 419–438.

[58] SOWELL, F. (1990). The fractional unit root distribution. *Econometrica* **58**, 495–505.

[59] VELASCO, C. (2000) Non-Gaussian log-periodogram regression. *Econometric Theory* **16**, 44–79.