Stationarity, Mixing, Distributional Properties and Moments of GARCH(p, q)–Processes

Alexander M. Lindner

Abstract This paper collects some of the well known probabilistic properties of GARCH(p, q) processes. In particular, we address the question of strictly and of weakly stationary solutions. We further investigate moment conditions as well as the strong mixing property of GARCH processes. Some distributional properties such as the tail behaviour and continuity properties of the stationary distribution are also included.

1 Introduction

Since their introduction by Engle (1982), autoregressive conditional heteroskedastic (ARCH) models and their extension by Bollerslev (1986) to generalised ARCH (GARCH) processes, GARCH models have been used widely by practitioners. At a first glance, their structure may seem simple, but their mathematical treatment has turned out to be quite complex. The aim of this article is to collect some probabilistic properties of GARCH processes.

Let $(\varepsilon_t)_{t\in\mathbb{Z}}$ be a sequence of independent and identically distributed (*i.i.d.*) random variables, and let $p \in \mathbb{N} = \{1, 2, \ldots\}$ and $q \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$. Further, let $\alpha_0 > 0$, $\alpha_1, \ldots, \alpha_{p-1} \geq 0$, $\alpha_p > 0$, $\beta_1, \ldots, \beta_{q-1} \geq 0$ and $\beta_q > 0$ be non-negative parameters. A GARCH(p, q) *process* $(X_t)_{t\in\mathbb{Z}}$ with *volatility process* $(\sigma_t)_{t\in\mathbb{Z}}$ is then a solution to the equations

$$X_t = \sigma_t \varepsilon_t, \quad t \in \mathbb{Z}, \tag{1}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2, \quad t \in \mathbb{Z}, \tag{2}$$

Alexander M. Lindner
Technische Universität Braunschweig, Institut für Mathematische Stochastik, Pockelsstrasse 14, D-38106 Braunschweig, e-mail: a.lindner@tu-bs.de

T.G. Anderson et al., *Handbook of Financial Time Series*,

DOI: 10.1007/978-3-540-71297-8_2,

where the process $(\sigma_t)_{t\in\mathbb{Z}}$ is non-negative. The sequence $(\varepsilon_t)_{t\in\mathbb{Z}}$ is referred to as the *driving noise sequence.* GARCH$(p,0)$ processes are called ARCH(p) *processes.* The case of a GARCH$(0,q)$ process is excluded since in that case, the volatility equation (2) decouples from the observed process X_t and the driving noise sequence. Note that in some articles (including the original paper by Bollerslev (1986)) the definition of p and q for GARCH processes is interchanged and the process defined in (1) with volatility given by (2) is referred to as GARCH(q,p) rather than GARCH(p,q).

It is a desirable property that σ_t should depend only on the past innovations $(\varepsilon_{t-h})_{h\in\mathbb{N}}$, i.e. be measurable with respect to the σ-algebra generated by $(\varepsilon_{t-h})_{h\in\mathbb{N}}$. If this condition holds, we shall call the GARCH(p,q) process *causal.* Then X_t is measurable with respect to the σ-algebra $\sigma(\varepsilon_{t-h} : h \in \mathbb{N}_0)$ generated by $(\varepsilon_{t-h})_{h\in\mathbb{N}_0}$. Also, σ_t is independent of $(\varepsilon_{t+h})_{h\in\mathbb{N}_0}$, and X_t is independent of $\sigma(\varepsilon_{t+h} : h \in \mathbb{N})$, for fixed t. Often the requirement of causality is added to the definition of GARCH processes. However, since we shall be mainly interested in strictly stationary solutions which turn out to be automatically causal for GARCH processes, we have dropped the requirement at this point.

The requirement that all the coefficients $\alpha_1,\ldots,\alpha_p$ and $\beta_1,\ldots,\beta_q$ are non-negative ensures that σ_t^2 is non-negative, so that σ_t can indeed be defined as the square root of σ_t^2. The parameter constraints can be slightly relaxed to allow for some negative parameters, but such that σ_t^2 will still be non-negative, see Nelson and Cao (1992). In the present paper, we shall however always assume non-negative coefficients.

The paper is organized as follows: in Section 2 we collect the criteria under which strictly stationary and weakly stationary solutions to the GARCH equations exist. The ARCH(∞) representation for GARCH processes is given in Section 3. In Section 4, we focus on conditions ensuring finiteness of moments, and give the autocorrelation function of the squared observations. Section 5 is concerned with the strong mixing property and an application to the limit behaviour of the sample autocorrelation function when sufficiently high moments exist. In Section 6 we shortly mention the tail behaviour of stationary solutions and their continuity properties. GARCH processes indexed by the integers are addressed in Section 7. Finally, some concluding remarks are made in Section 8.

For many of the results presented in this paper, it was tried to give at least a short sketch of the proof, following often the original articles, or the exposition given by Straumann (2005).

2 Stationary Solutions

Recall that a sequence $(Y_t)_{t\in\mathbb{Z}}$ of random vectors in $\mathbb{R}^d$ is called *strictly stationary*, if for every $t_1,\ldots,t_k \in \mathbb{Z}$, the distribution of $(Y_{t_1+h},\ldots,Y_{t_k+h})$

does not depend on h for $h \in \mathbb{N}_0$. When speaking of a *strictly stationary* GARCH(p,q) *process*, we shall mean that the bivariate process $(X_t, \sigma_t)_{t\in\mathbb{N}_0}$ is strictly stationary.

2.1 Strict stationarity of ARCH(1) and GARCH(1,1)

Now suppose that $(p,q) = (1,1)$ or that $(p,q) = (1,0)$, that $(\varepsilon_t)_{t\in\mathbb{Z}}$ is i.i.d., and that $(X_t, \sigma_t)_{t\in\mathbb{Z}}$ satisfy (1), (2). Hence we have a GARCH$(1,1)$/ARCH(1) process, whose volatility process satisfies

$$\sigma_t^2 = \alpha_0 + \beta_1 \sigma_{t-1}^2 + \alpha_1 \sigma_{t-1}^2 \varepsilon_{t-1}^2 = \alpha_0 + (\beta_1 + \alpha_1 \varepsilon_{t-1}^2)\sigma_{t-1}^2, \tag{3}$$

where $\beta_1 := 0$ if $q = 0$. Denoting

$$A_t = \beta_1 + \alpha_1 \varepsilon_t^2, \quad B_t = \alpha_0, \quad \text{and} \quad Y_t = \sigma_{t+1}^2, \tag{4}$$

it follows that $(Y_t)_{t\in\mathbb{Z}} = (\sigma_{t+1}^2)_{t\in\mathbb{Z}}$ is the solution of the random recurrence equation $Y_t = A_t Y_{t-1} + B_t$, where $(A_t, B_t)_{t\in\mathbb{Z}}$ is i.i.d. As we shall see, every strictly stationary solution $(\sigma_t^2)_{t\in\mathbb{Z}}$ of (3) can be expressed as an appropriate function of the driving noise sequence $(\varepsilon_t)_{t\in\mathbb{Z}}$, so that stationarity of $(\sigma_t^2)_{t\in\mathbb{Z}}$ implies stationarity of $(\sigma_t^2, \varepsilon_t)_{t\in\mathbb{Z}}$ and hence of (X_t, σ_t). Thus, the question of existence of strictly stationary solutions of the GARCH$(1,1)$ process can be reduced to the study of strictly stationary solutions of (3). Since we will need multivariate random reccurence equations for the treatment of higher order GARCH processes, we give their definition already in $\mathbb{R}^d$. So let $d \in \mathbb{N}$, and suppose $(A_t, B_t)_{t\in\mathbb{Z}}$ is an i.i.d. sequence, where A_t is a $(d \times d)$-random matrix and B_t is a d-dimensional random vector. The difference equation

$$Y_t = A_t Y_{t-1} + B_t, \quad t \in \mathbb{Z}, \tag{5}$$

is then called a *random recurrence equation (with i.i.d. coefficients)*, where the solution $(Y_t)_{t\in\mathbb{Z}}$ is a sequence of d-dimensional random vectors. Every such solution then satisfies

$$\begin{aligned} Y_t &= A_t Y_{t-1} + B_t \\ &= A_t A_{t-1} Y_{t-2} + A_t B_{t-1} + B_t = \cdots \\ &= \left(\prod_{i=0}^{k} A_{t-i} \right) Y_{t-k-1} + \sum_{i=0}^{k} \left(\prod_{j=0}^{i-1} A_{t-j} \right) B_{t-i} \end{aligned} \tag{6}$$

for all $k \in \mathbb{N}_0$, with the usual convention that $\prod_{j=0}^{-1} A_{t-j} = 1$ for the product over an empty index set. Letting $k \to \infty$, it is reasonable to hope that for a stationary solution, $\lim_{k\to\infty} \left(\prod_{i=0}^{k} A_{t-i} \right) Y_{t-k-1} = 0$ a.s. and that $\sum_{i=0}^k \left(\prod_{j=0}^{i-1} A_{t-j}\right) B_{t-i}$ converges almost surely as $k \to \infty$. In the GARCH(1,1) and ARCH(1) case, this is indeed the case: let A_t, B_t and Y_t as in (4). By (6), we have

$$\sigma_{t+1}^2 = Y_t = \left(\prod_{i=0}^{k} A_{t-i}\right) \sigma_{t-k}^2 + \alpha_0 \sum_{i=0}^{k} \prod_{j=0}^{i-1} A_{t-j}.$$

Since this is a sum of non-negative components, it follows that $\sum_{i=0}^\infty \prod_{j=0}^{i-1} A_{t-j}$ converges almost surely for each t, and hence that $\prod_{i=0}^k A_{t-i}$ converges almost surely to 0 as $k \to \infty$. Hence if $(\sigma_t^2)_{t\in\mathbb{Z}}$ is strictly stationary, then $\left(\prod_{i=0}^k A_{t-i}\right) \sigma_{t-k}^2$ converges in distribution and hence in probability to 0 as $k \to \infty$. So in the ARCH(1) and GARCH(1,1) case, there is at most one strictly stationary solution $(\sigma_t^2)_{t\in\mathbb{Z}} = (Y_{t-1})_{t\in\mathbb{Z}}$, given by

$$Y_t := \sum_{i=0}^{\infty} \left(\prod_{j=0}^{i-1} A_{t-j}\right) B_{t-i}, \quad t \in \mathbb{Z}. \tag{7}$$

On the other hand, it is clear that if (7) converges a.s. for some and hence all $t \in \mathbb{Z}$, where $(A_t, B_t)_{t\in\mathbb{Z}}$ are the i.i.d. coefficients of the random recurrence equation (5) in $\mathbb{R}^d$, then Y_t, defined by (7), defines a strictly stationary solution of (5).

We have seen that existence of a strictly stationary GARCH(1,1)/ARCH(1) process implies almost sure convergence of $\prod_{i=0}^k A_{-i}$ to 0 as $k \to \infty$. For the converse, we cite the following result:

Proposition 1 (Goldie and Maller (2000), Theorem 2.1)
Let $d = 1$ and $(A_t, B_t)_{t\in\mathbb{Z}}$ be i.i.d. in $\mathbb{R} \times \mathbb{R}$. Suppose that $P(B_0 = 0) < 1$, $P(A_0 = 0) = 0$, that $\prod_{i=0}^n A_{-i}$ converges almost surely to zero as $n \to \infty$, and that

$$\int_{(1,\infty)} \frac{\log q}{T_A(\log q)}\, P_{|B_0|}(dq) < \infty, \tag{8}$$

where $P_{|B_0|}$ denotes the distribution of $|B_0|$ and $T_A(y) := \int_0^y P(|A_0| < e^{-x})\, dx$ for $y \geq 0$. Then $\sum_{i=0}^\infty \left(\prod_{j=0}^{i-1} A_{t-j}\right) B_{t-i}$ converges almost surely absolutely for every $t \in \mathbb{Z}$.

In the GARCH(1,1) / ARCH(1) case, we have $B_0 = \alpha_0 > 0$ and (8) clearly holds. Observe that $\sum_{i=0}^\infty \left(\prod_{j=0}^{i-1} A_{t-j}\right) B_{t-i}$ converges trivially almost surely if $P(A_0 = 0) > 0$, in which case also $\prod_{i=0}^\infty A_{t-i} = 0$ a.s. Hence we see that a strictly stationary solution of GARCH(1,1) / ARCH(1) exists if and only if $\prod_{i=0}^k A_{-i}$ converges almost surely to 0 as $k \to \infty$. If $P(A_0 = 0) > 0$ this is clearly the case, so suppose that $\beta_1 > 0$ or that $P(\varepsilon_0^2 > 0) = 1$. Denoting $W_t := \log A_t$, $\prod_{i=0}^\infty A_{-i} = 0$ a.s. is then equivalent to the almost sure divergence to $-\infty$ of the random walk $S_n := \sum_{i=0}^n W_{-n}$. If $EW_0^+ < \infty$, then it is well known that $S_n \to -\infty$ if and only if $EW_0^+ < EW_0^- \leq \infty$, i.e. either $EW_0^- = \infty$ or $E|W_0| < \infty$ with $EW_0 < 0$. Furthermore, S_n cannot diverge almost surely to $-\infty$ as $n \to \infty$ if $EW_0^- < EW_0^+ = \infty$. Observe that in the GARCH$(1,1)$ case we have $\beta_1 > 0$, so that $W_0 \geq \log \beta_1 > -\infty$, hence $EW_0^- < \infty$, and it follows that there exists a strictly stationary solution of the GARCH$(1,1)$ process if and only if $E\log(\beta_1 + \alpha_1\varepsilon_0^2) < 0$. In the ARCH(1) case, however, $EW_0^- = \infty$ can happen. If $EW_0^- = \infty$, it is known from Kesten and Maller (1996) and Erickson (1973), that $S_n \to -\infty$ a.s. if and only if

$$\int_{(0,\infty)} \frac{x}{E(W_0^- \wedge x)}\, dP(W_0^+ \leq x) < \infty.$$

With $W_0 = \log \alpha_1 + \log \varepsilon_0^2$, the latter condition can be easily seen to be independent of $\alpha_1 > 0$. Summing up, we have the following characterisation of stationary solutions of the GARCH$(1,1)$ and ARCH(1) equations. For the GARCH$(1,1)$ case, and for the ARCH(1) case with $E\log^+(\varepsilon_0^2) < \infty$ this is due to Nelsen (1990). The ARCH(1) case with $E\log^+(\varepsilon_0^2) = \infty$ was added by Klüppelberg et al. (2004). Here, as usual, for a real number x we set $\log^+(x) = \log(\max(1,x))$, so that $\log^+(\varepsilon_0^2) = (\log \varepsilon_0^2)^+$.

Theorem 1 (Nelsen (1990), Theorem 2, Klüppelberg et al. (2004), Theorem 2.1)
(a) The GARCH$(1,1)$ *process with $\alpha_0, \alpha_1, \beta_1 > 0$ has a strictly stationary solution if and only if*

$$-\infty < E\log(\beta_1 + \alpha_1\varepsilon_0^2) < 0. \tag{9}$$

This solution is unique, and its squared volatility is given by

$$\sigma_t^2 = \alpha_0 \sum_{i=0}^{\infty} \prod_{j=0}^{i-1} (\beta_1 + \alpha_1 \varepsilon_{t-1-j}^2). \tag{10}$$

(b) The ARCH(1) *process with $\beta_1 = 0$ and $\alpha_1, \alpha_0 > 0$ has a strictly stationary solution if and only if one of the following cases occurs:*

(i) $P(\varepsilon_0 = 0) > 0$.
(ii) $E|\log \varepsilon_0^2| < \infty$ *and* $E\log \varepsilon_0^2 < -\log \alpha_1$, *i.e.* (9) *holds.*
(iii) $E(\log \varepsilon_0^2)^+ < \infty$ *and* $E(\log \varepsilon_0^2)^- = \infty$.
(iv) $E(\log \varepsilon_0^2)^+ = E(\log \varepsilon_0^2)^- = \infty$ *and*

$$\int_0^\infty x \left(\int_0^x P(\log \varepsilon_0^2 < -y)\, dy \right)^{-1} dP(\log \varepsilon_0^2 \leq x) < \infty. \tag{11}$$

In each case, the strictly stationary solution is unique, and its squared volatility is given by (10).

Observe that condition (9) depends on ε_0^2, α_1 and β_1, while conditions (i), (iii) and (iv) in the ARCH case depend on ε_0^2 only.

Example 1 (*a*) Suppose that $(\varepsilon_t)_{t\in\mathbb{Z}}$ is i.i.d. with $E\varepsilon_0^2 \in (0,\infty)$, and suppose that either $\beta_1 > 0$ (GARCH$(1,1)$) or that $E|\log \varepsilon_0^2| < \infty$. Since

$$E\log(\beta_1 + \alpha_1\varepsilon_0^2) \leq \log E(\beta_1 + \alpha_1\varepsilon_0^2) = \log(\beta_1 + E(\varepsilon_0^2)\,\alpha_1)$$

by Jensen's inequality, a sufficient condition for a strictly stationary solution to exist is that $E(\varepsilon_0^2)\,\alpha_1 + \beta_1 < 0$. Now suppose that ε_0 is standard normally distributed. If $\beta_1 = 0$, then

$$E\log(\alpha_1\varepsilon_0^2) = \log\alpha_1 + \frac{4}{\sqrt{2\pi}}\int_0^\infty \log(x)e^{-x^2/2}\,dx = \log(\alpha_1) - (C_{EM} + \log(2)),$$

where $C_{EM} := \lim_{N\to\infty}\sum_{n=1}^N \frac{1}{n} - \log(N) \approx 0.57721566$ is the Euler-Mascheroni constant. Hence, the ARCH(1) process with standard normal noise has a strictly stationary solution if and only

$$\alpha_1 < 2\exp(C_{EM}) \approx 3.562.$$

Since $\lim_{\beta_1\downarrow 0} E\log(\beta_1 + \alpha_1\varepsilon_0^2) = E\log(\alpha_1\varepsilon_0^2)$, it follows that for every $\alpha_1 < 2\exp(C_{EM})$ there exists some $\overline{\beta}(\alpha_1) > 0$ such that the GARCH$(1,1)$ process with parameters α_0, α_1 and $\beta_1 \in (0, \overline{\beta}(\alpha_1))$ and standard normal innovations has a strictly stationary solution. In particular, strictly stationary solutions of the GARCH$(1,1)$ process with $\alpha_1 + \beta_1 > 1$ do exist. However, observe that while α_1 may be bigger than 1, $\beta_1 < 1$ is a necessary condition for a strictly stationary solution to exist.
For normal noise, $E(\log(\beta_1 + \alpha_1\varepsilon_0^2))$ can be expressed in terms of confluent and generalised hypergeometric functions, which in turn can be calculated numerically. See Nelsen (1990), Theorem 6, for details.
(*b*) Consider the ARCH(1) process with $\alpha_1 > 0$, and let $(\varepsilon_t)_{t\in\mathbb{Z}}$ be i.i.d. such that the distribution of ε_0 has atoms at $\pm\sqrt{2 - E_2(2)}$ with mass $1/4$ each, and an absolutely continuous component with density $f_\varepsilon(x) = (4|x|(\log|x|)^2)^{-1}\mathbf{1}_{(-1/e,1/e)}(x)$. Here, $E_n(x) = \int_1^\infty e^{-xt}/t^n\,dt$ denotes the exponential integral, and it holds $E_2(2) \approx 0.0375$. Since $\int_{-1/e}^{1/e} f_\varepsilon(x)\,dx = \int_{-\infty}^{-1}(2y^2)^{-1}\,dy = 1/2$, f_ε indeed defines a probability distribution. Moreover, since ε_0 is symmetric, we have $E\varepsilon_0 = 0$ and

$$E\varepsilon_0^2 = \frac{1}{2}\int_0^{1/e}\frac{x}{(\log x)^2}\,dx + \frac{1}{2}(2 - E_2(2)) = \frac{1}{2}\int_{-\infty}^{-1}\frac{e^{2y}}{y^2}\,dy + \frac{1}{2}(2 - E_2(2)) = 1.$$

The absolutely continuous component of $\log\varepsilon_0^2$ can be easily seen to have density $x \mapsto (2x^2)^{-1}\mathbf{1}_{(-\infty,-1)}(x)$, so that $E(\log\varepsilon_0^2)^- = \infty$. Since $E(\log\varepsilon_0^2)^+ < \infty$, the ARCH(1) process with $\alpha_1 > 0$ and the given distribution of the $(\varepsilon_t)_{t\in\mathbb{Z}}$ has a unique strictly stationary solution by Case (iii) of the previous

Theorem.
(c) Let $(\varepsilon_t)_{t\in\mathbb{Z}}$ be i.i.d. with marginal density

$$f_\varepsilon(x) = \begin{cases} (2|x|(\log|x|)^{3/2})^{-1}, & |x| > e, \\ (4|x|(\log|x|)^2)^{-1}, & 0 < |x| < 1/e, \\ 0, & \text{else.} \end{cases}$$

Then the density of $\log \varepsilon_0^2$ is given by

$$f_{\log \varepsilon^2}(x) = \begin{cases} x^{-3/2}, & x > 1, \\ (2x^2)^{-1}, & x < -1, \\ 0, & x \in [-1, 1]. \end{cases}$$

We conclude that $E(\log \varepsilon_0^2)^+ = E(\log \varepsilon_0^2)^- = \infty$, and it is easily checked that (11) is satisfied. Hence, a unique strictly stationary solution of the ARCH(1) process with driving noise $(\varepsilon_t)_{t\in\mathbb{Z}}$ exists.

2.2 Strict stationarity of GARCH(p, q)

For higher order GARCH processes, one has to work with multidimensional random recurrence equations. Consider a GARCH(p, q) process $(X_t)_{t\in\mathbb{Z}}$ with volatility $(\sigma_t)_{t\in\mathbb{Z}}$ and driving noise sequence $(\varepsilon_t)_{t\in\mathbb{Z}}$. Let $\widetilde{p} := \max(p, 2)$, $\widetilde{q} := \max(q, 2)$ and define the random $(\widetilde{p} + \widetilde{q} - 1)$–vectors Y_t and B_t by

$$\begin{aligned} & Y_t = (\sigma_{t+1}^2, \ldots, \sigma_{t-\widetilde{p}+2}^2, X_t^2, \ldots, X_{t-\widetilde{q}+2}^2)' \\ \text{and} \quad & B_t = (\alpha_0, 0, \ldots, 0)' \in \mathbb{R}^{\widetilde{p}+\widetilde{q}-1}\ , \end{aligned} \tag{12}$$

respectively. Further, let $\beta_{q+1} = \beta_2 = 0$ if $q \le 1$, and $\alpha_2 = 0$ if $p = 1$, and define the random $(\widetilde{p} + \widetilde{q} - 1) \times (\widetilde{p} + \widetilde{q} - 1)$-matrix A_t by

$$A_t = \begin{pmatrix} \beta_1 + \alpha_1 \varepsilon_t^2 & \beta_2 & \cdots & \beta_{\widetilde{q}-1} & \beta_{\widetilde{q}} & \alpha_2 & \cdots & \alpha_{\widetilde{p}-1} & \alpha_{\widetilde{p}} \\ 1 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 \\ \epsilon_t^2 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix}. \tag{13}$$

These matrices where introduced by Bougerol and Picard (1992a). It is then easy to see that each strictly stationary solution of the GARCH equations

(1), (2) gives rise to a strictly stationary solution of the random recurrence equation (5) with Y_t, B_t and A_t as defined in (12) and (13), and vice versa. Observe that for $p = q = 1$ and for $(p, q) = (1, 0)$, the random recurrence equation with A_t and B_t as in (12) and (13) differs from the one with A_t and B_t as in (4). In fact, the former is a random recurrence equation in $\mathbb{R}^3$, while the latter is one-dimensional.

Strict stationarity of multivariate random recurrence equations is studied in terms of the top Lyapunov exponent. Let $\|\cdot\|$ be any vector norm in $\mathbb{R}^d$. For a matrix $M \in \mathbb{R}^{d\times d}$, the corresponding matrix norm $\|M\|$ is defined by

$$\|M\| := \sup_{x\in\mathbb{R}, x\neq 0} \frac{\|Mx\|}{\|x\|}.$$

Definition 1 Let $(A_n)_{n\in\mathbb{Z}}$ be an i.i.d. sequence of $d \times d$ random matrices, such that $E \log^+ \|A_0\| < \infty$. Then the *top Lyapunov exponent* associated with $(A_n)_{n\in\mathbb{Z}}$ is defined by

$$\gamma := \inf_{n\in\mathbb{N}_0} E\left(\frac{1}{n+1}\|A_0 A_{-1}\cdots A_{-n}\|\right).$$

Furstenberg and Kesten (1960) showed that

$$\gamma = \lim_{n\to\infty} \frac{1}{n+1} \log \|A_0 A_{-1}\cdots A_{-n}\| \tag{14}$$

almost surely, and an inspection of their proof shows that γ is independent of the chosen vector norm (hence matrix norm).

The existence of stationary solutions of random recurrence equations can be described neatly in terms of strict negativity of the associated top Lyapunov exponent. Namely, Bougerol and Picard (1992b) have shown that so called *irreducible* random recurrence equations with i.i.d. coefficients $(A_t, B_t)_{t\in\mathbb{Z}}$, such that $E \log^+ \|A_0\| < \infty$ and $E \log^+ \|B_0\| < \infty$, admit a nonanticipative strictly stationary solution if and only if the top Lyapunov exponent associated with $(A_t)_{t\in\mathbb{Z}}$ is strictly negative. Here, *nonanticipative* means that Y_t is independent of $(A_{t+h}, B_{t+h})_{h\in\mathbb{N}}$ for each t. For GARCH(p, q) cases, it is easier to exploit the positivity of the coefficients in the matrix A_t rather than to check that the model is irreducible. The result is again due to Bougerol and Picard:

Theorem 2 (Bougerol and Picard (1992a), Theorem 1.3) *Let $(\varepsilon_t)_{t\in\mathbb{Z}}$ be an i.i.d. sequence of random variables such that $E(\log \varepsilon_0^2)^+ < \infty$. Let $\alpha_0, \ldots, \alpha_p, \beta_1, \ldots, \beta_q$ be* GARCH(p, q) *parameters, and let the $(\widetilde{p}+\widetilde{q}-1)\times(\widetilde{p}+\widetilde{q}-1)$ random matrices A_t as well as the $(\widetilde{p}+\widetilde{q}-1)$–vectors B_t be defined as in* (13) *and* (12)*, respectively. Then the corresponding* GARCH(p, q) *process admits a strictly stationary solution if and only if the top Lyapunov exponent γ associated with the sequence $(A_t)_{t\in\mathbb{Z}}$ is strictly negative. This solution is unique, and the random vector Y_t defined in* (12) *satisfies* (7).

The fact that every strictly stationary solution must be unique and of the form (7) follows with a refined argument similar to the GARCH(1,1) case, using that every element in the vectors Y_t and in the matrices A_t must be non-negative. In particular this shows that every strictly stationary solution must be causal (the argument here does not require the assumption of finite log-moments). Further, existence of a strictly stationary solution implies $\lim_{k\to\infty} \|A_0 A_{-1} \cdots A_{-k}\| = 0$ a.s. Since $(A_n)_{n\in\mathbb{Z}}$ is i.i.d. and $E\log^+ \|A_0\| < \infty$, this in turn implies strict negativity of the top Lyapunov exponent γ (see Bougerol and Picard (1992b), Lemma 3.4). That $\gamma < 0$ implies convergence of (7) can be seen from the almost sure convergence in (14), which implies

$$\left\| \left(\prod_{j=0}^{k-1} A_{t-j} \right) B_{t-k} \right\| \le C_t e^{\gamma k/2}$$

for some random variable C_t. Hence, the series (7) converges almost surely and must be strictly stationary. That strict negativity of the top Lyapunov exponent implies convergence of (7) and hence the existence of strictly stationary solutions is true for a much wider class of random recurrence equations, see e.g. Kesten (1973), Vervaat (1979), Brandt (1986) or Bougerol and Picard (1992b).

Due to its importance, we state the observation made after Theorem 2 again explicitly:

Remark 1 A strictly stationary solution to the GARCH equations (1) and (2) is necessarily unique and the corresponding vector Y_t defined in (12) satisfies (7). In particular, every strictly stationary GARCH process is causal.

For matrices, it may be intractable to obtain explicit expressions for the top Lyapunov exponent and hence to check whether it is strictly negative or not. Often, one has to use simulations based on (14) to do that. If the noise sequence has finite variance, however, Bollerslev gave a handy sufficient condition for the GARCH process to have a strictly stationary solution, which is easy to check (part (a) of the following theorem). Bougerol and Picard showed that the boundary values in this condition can still be attained under certain conditions, and they have also given a necessary condition for strictly stationary solutions to exist:

Corollary 1 (Bollerslev (1986), Theorem 1, Bougerol and Picard (1992a), Corollaries 2.2, 2.3)
Let $(\varepsilon_t)_{t\in\mathbb{Z}}$ be the driving noise sequence of a GARCH(p,q) *process, and suppose that $0 < E\varepsilon_0^2 < \infty$. Then the following hold:*
(a) If $E(\varepsilon_0^2)\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$, then the GARCH$(p,q)$ *process admits a unique strictly stationary solution.*
(b) If $P(\varepsilon_0 = 0) = 0$, ε_0 has unbounded support, $p, q \ge 2$ and $\alpha_1, \ldots, \alpha_p > 0$, $\beta_1, \ldots, \beta_q > 0$, and $E(\varepsilon_0^2)\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j = 1$, then the GARCH$(p,q)$ *process admits a unique strictly stationary solution.*

(c) If $\sum_{j=1}^q \beta_j \geq 1$*, then no strictly stationary solution of the* GARCH(p,q) *process exists.*

For the proof of Corollary 1, one may assume that $E\varepsilon_0^2 = 1$. The general result then follows by an easy transformation. If $E\varepsilon_0^2 = 1$, Bougerol and Picard (1992a) prove (b) by showing that the spectral radius $\rho(E(A_0))$ of the matrix $E(A_0)$ is equal to 1. Recall that the *spectral radius* $\rho(C)$ of a square matrix C is defined by

$$\rho(C) = \sup\{|\lambda| : \lambda \text{ eigenvalue of } C\}.$$

Since A_0 is almost surely not bounded, neither has zero columns nor zero rows, and has non-negative entries, it follows from Theorem 2 of Kesten and Spitzer (1984) that $\gamma < \log \rho(E(A_0)) = 0$. The proofs of (a) and (c) are achieved by similar reasoning, using estimates between the top Lyapunov exponent and the spectral radius. In particular, in case (a) one has $\gamma \leq \log \rho(E(A_0)) < 0$.

For real data one often estimates parameters α_i and β_j such that $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j$ is close to one, when assuming noise with variance 1. In analogy to the integrated ARMA (ARIMA) process, Engle and Bollerslev (1986) call GARCH processes for which $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j = 1$ *integrated* GARCH(p,q) processes, or IGARCH(p,q) *processes*, for short. Observe that Corollary 1(b) shows that IGARCH processes may have a strictly stationary solution, unlike ARIMA processes where a unit root problem occurs.

Remark 2 Let ε_0, p, q and $\alpha_1, \ldots, \alpha_p$, $\beta_1, \ldots, \beta_q$ be as in Corollary 1(b). Then there exists $\delta > 0$ such that for all $\widetilde{\alpha}_i \geq 0$, $\widetilde{\beta}_j \geq 0$ with $|\widetilde{\alpha}_i - \alpha_i| < \delta$ $(i = 1, \ldots, p)$ and $|\widetilde{\beta}_j - \beta_j| < \delta$ $(j = 1, \ldots, q)$, the GARCH(p,q) process with parameters α_0, $\widetilde{\alpha}_1, \ldots, \widetilde{\alpha}_p$, $\widetilde{\beta}_1, \ldots, \widetilde{\beta}_q$ and noise sequence $(\varepsilon_t)_{t\in\mathbb{Z}}$ admits a unique strictly stationary solution. In particular, there exist strictly stationary GARCH(p,q) processes for which $E(\varepsilon_0^2)\sum_{i=1}^p \widetilde{\alpha}_i + \sum_{j=1}^q \widetilde{\beta}_j > 1$. This follows immediately from Definition 1 and Theorem 2, since for the parameters of Corollary 1(b), the top Lyapunov exponent γ is strictly negative.

2.3 Ergodicity

Let $Y = (Y_t)_{t\in\mathbb{Z}}$ be a strictly stationary time series of random vectors in $\mathbb{R}^k$. Then Y can be seen as a random element in $(\mathbb{R}^k)^{\mathbb{Z}}$, equipped with its Borel-σ-algebra $\mathcal{B}((\mathbb{R}^k)^{\mathbb{Z}})$. Let the backshift operator $\Phi_{BS} : (\mathbb{R}^k)^{\mathbb{Z}} \to (\mathbb{R}^k)^{\mathbb{Z}}$ be given by $\Phi_{BS}((z_i)_{i\in\mathbb{Z}}) = (z_{i-1})_{i\in\mathbb{Z}}$. Then the time series $(Y_t)_{t\in\mathbb{Z}}$ is called *ergodic* if $\Phi_{BS}(\Lambda) = \Lambda$ for $\Lambda \in \mathcal{B}((\mathbb{R}^k)^{\mathbb{Z}})$ implies $P(Y \in \Lambda) \in \{0,1\}$. See e.g. Ash and Gardner (1975) for this and further properties of ergodic time series. In particular, it is known that if $(g_n)_{n\in\mathbb{Z}}$ is a sequence of measurable functions

$g_n : (\mathbb{R}^k)^{\mathbb{Z}} \to \mathbb{R}^d$ such that $g_{n-1} = g_n \circ \Phi_{BS}$ and $Y = (Y_t)_{t\in\mathbb{Z}}$ is strictly stationary and ergodic with values in $\mathbb{R}^k$, then $(g_n(Y))_{n\in\mathbb{Z}}$ is also strictly stationary and ergodic (see e.g. Brandt et al. (1990), Lemma A 1.2.7). Since the sequence $(A_t, B_t)_{t\in\mathbb{Z}}$ is i.i.d. and hence strictly stationary and ergodic for a GARCH process, it follows that every strictly stationary GARCH process is ergodic, since it can be expressed via (7). This is due to Bougerol and Picard (1992a), Theorem 1.3.

2.4 Weak stationarity

Recall that a time series $(Z_t)_{t\in\mathbb{Z}}$ of random vectors in $\mathbb{R}^d$ is called *weakly stationary* or *wide-sense stationary*, if $E\|Z_t\|^2 < \infty$ for all $t \in \mathbb{Z}$, $E(Z_t) \in \mathbb{R}^d$ is independent of $t \in \mathbb{Z}$, and the covariance matrices satisfy

$$\mathrm{Cov}(Z_{t_1+h}, Z_{t_2+h}) = \mathrm{Cov}(Z_{t_1}, Z_{t_2})$$

for all $t_1, t_2, h \in \mathbb{Z}$. Clearly, every strictly stationary sequence which satisfies $E\|Z_0\|^2 < \infty$ is also weakly stationary. For causal GARCH processes, we shall see that the converse is true also, i.e. that every causal weakly stationary GARCH process is also strictly stationary.

Let (X_t, σ_t) be a GARCH process such that σ_t is independent of ε_t, which is in particular satisfied for causal solutions. Then if $P(\varepsilon_0 = 0) < 1$, it follows from (1) and the independence of σ_t and ε_t that for given $r \in (0, \infty)$, $E|X_t|^r < \infty$ if and only if $E|\varepsilon_t|^r < \infty$ and $E\sigma_t^r < \infty$. Suppose $E\varepsilon_0^2 \in (0, \infty)$, and that (X_t, σ_t) is a GARCH(p, q) process such that $E\sigma_t^2 = E\sigma_{t'}^2 < \infty$ for all $t, t' \in \mathbb{Z}$. Then (2) shows that

$$E(\sigma_0^2) = \alpha_0 + \sum_{i=1}^{p} \alpha_i E(\sigma_0^2)\, E(\varepsilon_0^2) + \sum_{j=1}^{q} \beta_j E(\sigma_0^2).$$

Hence we see that a necessary condition for a causal weakly stationary solution to exist is that $E(\varepsilon_0^2)\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$. Now suppose that $(\sigma_t)_{t\in\mathbb{Z}}$ is a causal weakly stationary solution, and for simplicity assume that $E\varepsilon_0^2 = 1$. With Y_t, B_t and A_t as in (12) and (13), Y_t must satisfy (6). Note that then $\sum_{i=0}^{\infty} \left(\prod_{j=0}^{i-1} A_{t-j}\right) B_{t-i}$ converges a.s. to the strictly stationary solution by Corollary 1. By (6), this implies that $\left(\prod_{i=0}^{k} A_{t-i}\right) Y_{t-k-1}$ converges almost surely to some finite random variable as $k \to \infty$. If this limit can be seen to be 0, then it follows that the weakly stationary solution must coincide with the strictly stationary. As remarked after Corollary 1, the spectral radius of $E(A_0)$ is less than 1. Hence there is some $N \in \mathbb{N}$ such that $\|(EA_0)^N\| = \|E(A_0 \cdots A_{-N+1})\| < 1$. By causality and weak stationarity, this implies that $E\left(\left(\prod_{i=0}^{k} A_{t-i}\right) Y_{t-k-1}\right)$ converges to 0 as $k \to \infty$, and since each of the components of $\left(\prod_{i=0}^{k} A_{t-i}\right) Y_{t-k-1}$ is positive, Fatou's lemma shows that its almost sure limit must be 0, so that every causal weakly stationary solution is also strictly stationary. Conversely, if $(Y_t)_{t\in\mathbb{Z}}$ is a strictly stationary solution and $E(\varepsilon_0^2)\sum_{i=1}^{p}\alpha_i + \sum_{j=1}^{q}\beta_q < 1$ with $E\varepsilon_0^2 = 1$ for simplicity, it follows from $\|(EA_0)^N\| < 1$ that $\sum_{i=0}^{\infty} E\left(\left(\prod_{j=0}^{i-1} A_{t-j}\right) B_{t-i}\right)$ is finite, and since each of its components is positive, this implies that $E\|Y_t\| < \infty$ for the strictly stationary solution. Summing up, we have the following characterisation of causal weakly stationary solutions, which was derived by Bollerslev (1986).

Theorem 3 (Bollerslev (1986), Theorem 1)
Let $(\varepsilon_t)_{t\in\mathbb{Z}}$ be such that $E\varepsilon_0^2 < \infty$. Then the GARCH$(p,q)$ *process $(X_t,\sigma_t)_{t\in\mathbb{Z}}$ admits a causal weakly stationary solution if and only $E(\varepsilon_0^2)\sum_{i=1}^{p}\alpha_i + \sum_{j=1}^{q}\beta_j < 1$. In that case, the causal weakly stationary solution is unique and coincides with the unique strictly stationary solution. It holds*

$$E(\sigma_t^2) = \frac{\alpha_0}{1 - E(\varepsilon_0^2)\sum_{i=1}^{p}\alpha_i - \sum_{j=1}^{q}\beta_j}, \quad E(X_t^2) = E(\sigma_t^2)E(\varepsilon_0^2). \tag{15}$$

3 The ARCH(∞) Representation and the Conditional Variance

Often it can be helpful to view a GARCH(p,q) process as an ARCH process of infinite order. In particular, from the ARCH(∞) representation one can easily read off the conditional variance of X_t given its infinite past $(X_s : s < t)$. Originally, Engle (1982) and Bollerslev (1986) defined ARCH and GARCH processes in terms of the conditional variance. Equation (18) below then shows that this property does hold indeed, so that the definition of GARCH processes given here is consistent with the original one of Engle and Bollerslev.

Theorem 4 (Bollerslev (1986), pp. 309–310)
Let $(X_t,\sigma_t)_{t\in\mathbb{Z}}$ be a strictly stationary GARCH(p,q) *process driven by $(\varepsilon_t)_{t\in\mathbb{Z}}$, such that $E\varepsilon_0^2 < \infty$ and $E(\varepsilon_0^2)\sum_{i=1}^{p}\alpha_i + \sum_{j=1}^{q}\beta_j < 1$. Then there is a sequence $(\psi_j)_{j\in\mathbb{N}_0}$ of real constants such that $\psi_0 > 0$, $\psi_j \geq 0$ for all j, $\sum_{j=0}^{\infty}\psi_j < \infty$, and*

$$\sigma_t^2 = \psi_0 + \sum_{i=1}^{\infty}\psi_i X_{t-i}^2. \tag{16}$$

The constants are determined by

$$\psi_0 = \frac{\alpha_0}{1 - \sum_{j=1}^q \beta_j},$$
$$\sum_{j=1}^{\infty} \psi_j z^j = \frac{\sum_{i=1}^p \alpha_i z^i}{1 - \sum_{j=1}^q \beta_j z^j}, \quad z \in \mathbb{C}, \quad |z| \leq 1. \tag{17}$$

In particular, σ_t^2 is measurable with respect to the infinite past $(X_s : s \leq t-1)$, and the conditional expectation and variance of X_t given $(X_s : s < t)$ are given by

$$E(X_t | X_s : s < t) = E(\varepsilon_0)\, \sigma_t \quad \textit{and} \quad V(X_t | X_s : s < t) = V(\varepsilon_0)\, \sigma_t^2, \tag{18}$$

respectively.

For example, if $(\varepsilon_t)_{t\in\mathbb{Z}}$ is i.i.d. standard normal, then conditionally on $(X_s : s < t)$, X_t is $N(0, \sigma_t^2)$ distributed, since σ_t^2 is a Borel function of $(X_s : s < t)$. ARCH(∞) models were introduced in more generality by Robinson (1991). The explicit expression in (16) can be found in Bollerslev (1986) or Nelson and Cao (1992). It can be derived defining

$$S_t := \sigma_t^2 - E(\sigma_t^2), \quad Z_t := X_t^2 - E(X_t^2), \quad t \in \mathbb{Z}. \tag{19}$$

Then (2) is equivalent to

$$S_t - \sum_{j=1}^q \beta_j S_{t-j} = \sum_{i=1}^p \alpha_i Z_{t-i}. \tag{20}$$

This is an ARMA equation for $(S_t)_{t\in\mathbb{Z}}$ such that $\sup_{t\in\mathbb{Z}} E|Z_t| < \infty$ and $E(S_t) = E(Z_t) = 0$. Since $\sum_{j=1}^q \beta_j < 1$, this ARMA equation is causal, and it follows that $S_t = \sum_{j=1}^\infty \psi_j Z_{t-j}$ where $(\psi_j)_{j\in\mathbb{N}}$ are given by (17). An easy calculation prevails that $\psi_j \geq 0$, and resubstituting σ_t^2 and X_t^2 in this ARMA equation shows (16). Hence σ_t is measurable with respect to the σ-algebra generated by $(X_s : s < t)$, while ε_t is independent of this σ-algebra by causality. This then implies (18).

In the literature there exist many other examples of ARCH(∞) models apart from GARCH(p,q). For more information and references regarding ARCH(∞) models, see Giraitis et al. (2006) and (2008).

4 Existence of Moments and the Autocovariance Function of the Squared Process

It is important to know whether the stationary solution has moments of higher order. For example, in Theorem 3, we have seen that the strictly stationary solution has finite second moments if and only if $E(\varepsilon_0^2) \sum_{i=1}^p \alpha_i +$

$\sum_{j=1}^q \beta_j < 1$, and we have given an explicit expression for $E\sigma_t^2$ and EX_t^2. However, one is also interested in conditions ensuring finiteness of moments of higher order, the most important case being finiteness of $E\sigma_t^4$ and EX_t^4. For the GARCH$(1,1)$ process with normal innovations, a necessary and sufficient condition for such moments to exist has been given by Bollerslev (1986), and extended by He and Teräsvirta (1999b) to general noise sequences. Ling (1999) and Ling and McAleer (2002) give a necessary and sufficient condition for moments of higher order to exist. For ARCH(p) processes, a necessary and sufficient condition for higher order moments to exist was already obtained earlier by Milhøj (1985).

Observe that if $P(\varepsilon_0 = 0) < 1$, then by independence of X_t and σ_t for strictly stationary and hence causal solutions, the m'th moment of $X_t = \sigma_t \varepsilon_t$ exists if and only $E\sigma_t^m < \infty$ and $E|\varepsilon_t|^m < \infty$. Hence we shall only be concerned with moment conditions for σ_t^2. In most cases, ε_t will be a symmetric distribution, so that the odd moments of ε_t and hence X_t will be zero. The main concern is hence on even moments of GARCH processes.

4.1 Moments of ARCH(1) and GARCH(1, 1)

The following theorem gives a complete characterisation when the (possible fractional) moment of a GARCH$(1,1)$ or ARCH(1) process exists:

Theorem 5 (Bollerslev (1986), Theorem 2, and He and Teräsvirta (1999b), Theorem 1)
Let (X_t, σ_t) be a strictly stationary GARCH$(1,1)$ *or* ARCH(1) *process as in* (1), (2). *Let $m > 0$. Then the (fractional) m'th moment $E(\sigma_t^{2m})$ of σ_t^2 exists if and only if*

$$E(\beta_1 + \alpha_1 \varepsilon_0^2)^m < 1. \tag{21}$$

If m is a positive integer and this condition is satisfied, and $\mu_j := E(\sigma_t^{2j})$ denotes the j'th moment of σ_t^2, then μ_m can be calculated recursively by

$$\mu_m = (1 - E(\beta_1 + \alpha_1 \varepsilon_0^2)^m)^{-1} \sum_{j=0}^{m-1} \binom{m}{j} \alpha_0^{m-j} E(\beta_1 + \alpha_1 \varepsilon_0^2)^j \mu_j. \tag{22}$$

The $(2m)$'th moment of X_t is given by

$$E(X_t^{2m}) = \mu_m E(\varepsilon_0^{2m}).$$

That condition (21) is necessary and sufficient for finiteness of $E(\sigma_t^{2m})$ $(m \in (0,\infty))$ can be easily seen from representation (10): for if $E(\beta_1 + \alpha_1 \varepsilon_0^2)^m < 1$ and $m \in [1,\infty)$, then Minkowski's inequality shows that

$$(E(\sigma_t^{2m}))^{1/m} \leq \alpha_0 \sum_{i=0}^{\infty} (E(\beta_1 + \alpha_1 \varepsilon_0^2)^m)^{i/m} < \infty,$$

and for $m < 1$ one uses similarly $E(U+V)^m \leq EU^m + EV^m$ for positive random variables U, V. Conversely, if $E(\sigma_t^{2m}) < \infty$, then $E\prod_{j=0}^{i-1}(\beta_1 + \alpha_1 \varepsilon_{t-1-j}^2)^m$ must converge to 0 as $i \to \infty$, which can only happen if (21) holds. Finally, if m is an integer and (21) holds, then (22) follows easily by raising (2) to the m'th power and taking expectations.

Example 2 For an integer m, $E(\sigma_t^{2m})$ is finite if and only if $\sum_{j=0}^{m} \binom{m}{j} \beta_1^{m-j} \alpha_1^j E\varepsilon_t^{2j} < 1$. If ε_t is standard normally distributed, this means that

$$\sum_{j=0}^{\infty} \binom{m}{j} \beta_1^{m-j} \alpha_1^j \prod_{i=1}^{j} (2i-1) < 1.$$

For example, the fourth moment of σ_t exists if and only if $\beta_1^2 + 2\beta_1\alpha_1 + 3\alpha_1^2 < 1$.

As an immediate consequence of Theorem 5, one sees that GARCH processes do not have finite moments of all orders if ε_0 has unbounded support, which is a first indication that GARCH processes will generally have heavy tails:

Corollary 2 *Let $(X_t, \sigma_t : t \in \mathbb{Z})$ be a strictly stationary* GARCH(1,1) *or* ARCH(1) *process and assume that $P(\alpha_1 \varepsilon_0^2 + \beta_1 > 1) > 0$. Then there is $r \geq 1$, such that $E\sigma_0^{2r} = E|X_0|^{2r} = \infty$.*

4.2 Moments of GARCH(p, q)

For GARCH processes of higher order, Ling (1999) and Ling and McAleer (2002) give necessary and sufficient conditions for even moments of σ_t to be finite. In order to state their result, we need the notion of the Kronecker product of two matrices. For an $(m \times n)$-matrix $C = (c_{ij})_{i=1,\ldots,m, j=1,\ldots,n}$ and a $(p \times r)$-matrix D, the *Kronecker product* $C \otimes D$ is the $(mp \times nr)$-matrix given by

$$C \otimes D = \begin{pmatrix} c_{11}D & \cdots & c_{1n}D \\ \vdots & \ddots & \vdots \\ c_{m1}D & \cdots & c_{mn}D \end{pmatrix}.$$

See e.g. Lütkepohl (1996) for elementary properties of the Kronecker product. We then have:

Theorem 6 (Ling and McAleer (2002), Theorem 2.1)
Let $(X_t, \sigma_t)_{t\in\mathbb{Z}}$ be a strictly stationary GARCH(p,q) *process as in* (1), (2), *and assume that $\alpha_1+\beta_1>0$. Let A_t be the $(\widetilde{p}+\widetilde{q}-1)\times(\widetilde{p}+\widetilde{q}-1)$ matrix of* (13). *Let $m\in\mathbb{N}$. Then the m'th moment of σ_t^2 is finite if and only if the spectral radius of the matrix $E(A_t^{\otimes m})$ is strictly less than 1.*

Originally, Ling and McAleer (2002) formulated their result in terms of the spectral radius of a matrix corresponding to another state space representation of GARCH processes than the A_t-matrix defined in (13). The proof, however, is quite similar. We shortly sketch the argument:
Suppose that $\rho(E(A_t^{\otimes m}))=\limsup_{n\to\infty}\|(E(A_t^{\otimes m}))^n\|^{1/n}<1$. Then there is $\lambda\in(0,1)$ such that $\|(E(A_t^{\otimes m}))^n\|\le\lambda^n$ for large enough n, so that the supremum of all elements of $(E(A_t^{\otimes m}))^n$ decreases exponentially as $n\to\infty$. The same is then true for all elements of $(E(A_t^{\otimes m'}))^n$ for every $m'\in\{1,\ldots,n\}$. Now take the m'th Kronecker power of the representation (7) for the vector Y_t defined in (12). For example, for $m=2$, one has (since $B_t=B_{t-i}$ in (12))

$$\begin{aligned}
Y_t^{\otimes 2} &= \sum_{i_1=0}^{\infty}\sum_{i_2=0}^{\infty}\left(\left(\prod_{j_1=0}^{i_1-1}A_{t-j_1}\right)B_t\right)\otimes\left(\left(\prod_{j_2=0}^{i_2-1}A_{t-j_2}\right)B_t\right)\\
&= \sum_{i_1=0}^{\infty}\sum_{i_2=i_1}^{\infty}\left(\prod_{j_1=0}^{i_1-1}A_{t-j_1}^{\otimes 2}\right)\left(\prod_{j_2=i_1}^{i_2-1}(\mathrm{Id}\otimes A_{t-j_2})\right)B_t^{\otimes 2}\\
&\quad+\sum_{i_1=1}^{\infty}\sum_{i_2=0}^{i_1-1}\left(\prod_{j_2=0}^{i_2-1}A_{t-j_2}^{\otimes 2}\right)\left(\prod_{j_1=i_2}^{i_1-1}(A_{t-j_1}\otimes\mathrm{Id})\right)B_t^{\otimes 2},
\end{aligned}$$

where Id denotes the $(\widetilde{p}+\widetilde{q}-1)\times(\widetilde{p}+\widetilde{q}-1)$ identity matrix. Taking expectations and using the exponential decay of the elements, which are all non-negative, this then shows that $E(Y_t^{\otimes m})$ is finite, and hence that $E(\sigma_t^{2m})<\infty$. The converse is established along similar lines: finiteness of $E(\sigma_t^{2m})$ implies finiteness of $E(Y_t^{\otimes m})$. Using the fact that all appearing matrices and vectors have non-negative entries, this then implies finiteness of $\sum_{i=0}^{\infty}(E(A_t^{\otimes m}))^iB_0^{\otimes m}$ as argued by Ling and McAleer (2002), and making use of the assumption $\alpha_1+\beta_1>0$, this can be shown to imply finiteness of $\sum_{i=0}^{\infty}\|(E(A_t^{\otimes m}))^i\|$, showing that $\rho(E(A_t^{\otimes m}))<1$.

To check whether the spectral radius of the matrix $E(A_t^{\otimes m})$ is less than 1 or not may be tedious or only numerically achievable. A simple sufficient condition for the existence of moments can however be obtained by developing the ARCH(∞) representation (16) into a *Volterra series expansion*, as described by Giraitis et al. (2006) and (2008). Accordingly, a sufficient condition for the m'th moment of σ_t^2 in an ARCH(∞) process to exist is that $\sum_{j=1}^{\infty}\psi_j(E(|\varepsilon_0|^{2m}))^{1/m}<1$. This was shown by Giraitis et al. (2000) for $m=2$ and observed to extend to hold for general $m\ge1$ by Giraitis et al. (2006). With (17), this gives for the GARCH(p,q) process:

Proposition 2 (Giraitis et al. (2006), Theorem 2.1)
Let $(X_t, \sigma_t)_{t\in\mathbb{Z}}$ *be a strictly stationary* GARCH(p,q) *process as in* (1), (2), *let* $m \in [1,\infty)$, *and suppose that* $0 < E|\varepsilon_0|^{2m} < \infty$. *Then*

$$\left(\frac{\sum_{i=1}^{p} \alpha_i}{1 - \sum_{j=1}^{q} \beta_j} \right)^m E|\varepsilon_0|^{2m} < 1$$

is a sufficient condition for $E(\sigma_0^{2m}) < \infty$.

Observe that $0 < E|\varepsilon_0|^{2m} < \infty$ implies $\sum_{j=1}^{q} \beta_j < 1$ by Corollary 1(c), so that the expressions in the condition above are well-defined.

In econometrics, the kurtosis is often seen as an indicator for tail heaviness. Recall that the *kurtosis* K_R of a random variable R with $ER^4 < \infty$ is defined by $K_R = \frac{ER^4}{(ER^2)^2}$. If (X_t, σ_t) is a stationary GARCH process which admits finite fourth moment, then if follows from Jensen's inequality that

$$EX_t^4 = E(\varepsilon_t^4)E(\sigma_t^4) \geq E(\varepsilon_t^4)(E(\sigma_t^2))^2 = K_{\varepsilon_0}(E(X_t^2))^2,$$

so that $K_{X_0} \geq K_{\varepsilon_0}$. This shows that the kurtosis of the stationary solution is always greater or equal than the kurtosis of the driving noise sequence, giving another indication that GARCH processes lead to comparatively heavy tails.

While Theorem 6 gives a necessary and sufficient condition for even moments to exist, it does not give any information about the form of the moment. The most important higher order moment is the fourth moment of σ_t, and an elegant method to determine $E\sigma_t^4$ was developed by Karanasos (1999). To illustrate it, suppose that $(X_t, \sigma_t)_{t\in\mathbb{Z}}$ is a strictly stationary GARCH(p,q) process as in (1), (2), such that $E(\sigma_t^4) < \infty$, and denote

$$w := E\varepsilon_0^2, \;\; v := E\varepsilon_0^4, \;\; f := E\sigma_0^4 \;\text{ and }\; g := E\sigma_0^2 = \frac{\alpha_0}{1 - w\sum_{i=1}^{p}\alpha_i - \sum_{j=1}^{q}\beta_j},$$

where we used (15). Then w, v, and g are known and we want to determine f. For $i \in \mathbb{N}$, denote further

$$\lambda_i := E(\sigma_t^2 X_{t-i}^2) \quad \text{and} \quad c_i := E(\sigma_t^2 \sigma_{t-i}^2).$$

Since $E(X_t^2 | \varepsilon_{t-h} : h \in \mathbb{N}) = w\sigma_t^2$, it further holds for $i \in \mathbb{N}$,

$$w\lambda_i = E(X_t^2 X_{t-i}^2), \;\; wc_i = E(X_t^2 \sigma_{t-i}^2), \;\; \text{and} \;\; f = E(X_t^2\sigma_t^2)/w = E(X_t^4)/v.$$

Then, taking expectations in each of the equations

$$\begin{aligned} X_t^2\sigma_t^2 &= X_t^2\left(\alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 + \sum_{i=1}^q \beta_i \sigma_{t-i}^2\right), \\ \sigma_t^2\sigma_{t-j}^2 &= \sigma_{t-j}^2\left(\alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 + \sum_{i=1}^q \beta_i \sigma_{t-i}^2\right), \quad j = 1,\ldots,q, \\ \sigma_t^2 X_{t-j}^2 &= X_{t-j}^2\left(\alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 + \sum_{i=1}^q \beta_i \sigma_{t-i}^2\right), \quad j = 1,\ldots,p, \end{aligned}$$

one obtains

$$wf = \alpha_0 wg + \sum_{i=1}^p w\alpha_i\lambda_i + \sum_{i=1}^q w\beta_i c_i, \tag{23}$$

$$\begin{aligned} c_j &= \alpha_0 g + (w\alpha_j + \beta_j)f + \sum_{i=1}^{j-1}(w\alpha_{j-i} + \beta_{j-i})c_i \\ &\quad + \sum_{i=1}^{p-j}\alpha_{j+i}\lambda_i + \sum_{i=1}^{q-j}\beta_{j+i}c_i, \quad j = 1,\ldots,q, \end{aligned} \tag{24}$$

$$\begin{aligned} \lambda_j &= \alpha_0 wg + (v\alpha_j + w\beta_j)f + \sum_{i=1}^{j-1}(w\alpha_{j-i} + \beta_{j-i})\lambda_i \\ &\quad + \sum_{i=1}^{p-j} w\alpha_{j+i}\lambda_i + \sum_{i=1}^{q-j} w\beta_{j+i}c_i, \quad j = 1,\ldots,p, \end{aligned} \tag{25}$$

where $\alpha_i = 0$ for $i > p$ and $\beta_i = 0$ for $i > q$. Substituting c_q from (24) and λ_p from (25) into (23), one obtains a system of $(p+q-1)$ equations for the unknown variables $(f, c_1, \ldots, c_{q-1}, \lambda_1, \ldots, \lambda_{p-1})$. See Karanasos (1999), Theorem 3.1, for more information. For another approach to obtain necessary conditions for the fourth moment to exist and to obtain its structure, we refer to He and Teräsvirta (1999a), Theorem 1.

4.3 The autocorrelation function of the squares

If the driving noise process of a strictly and weakly stationary GARCH process has expectation $E\varepsilon_0 = 0$, then $EX_t = E(\varepsilon_0)\,E(\sigma_t) = 0$, and for $h \in \mathbb{N}$ it follows from (18) that

$$E(X_t X_{t-h}) = E\,E(X_t X_{t-h} | X_s : s < t) = E(X_{t-h} E(\varepsilon_0)\sigma_t) = 0,$$

so that $(X_t)_{t\in\mathbb{Z}}$ is (weak) White Noise (provided $E\varepsilon_0^2 \neq 0$), i.e. a weakly stationary sequence whose elements are uncorrelated. This uncorrelatedness is however not preserved in the squares of the GARCH process. Rather do the squares $(X_t^2)_{t\in\mathbb{Z}}$ satisfy an ARMA equation. This was already observed by Bollerslev (1986), (1988). More precisely, we have:

Theorem 7 (Bollerslev (1986), Section 4)
Let $(X_t, \sigma_t)_{t\in\mathbb{Z}}$ be a strictly stationary GARCH(p,q) *process such that $E\sigma_0^4 < \infty$, $E\varepsilon_0^4 < \infty$ and $\mathrm{Var}(\varepsilon_0^2) > 0$. Define*

$$u_t := X_t^2 - (E\varepsilon_t^2)\sigma_t^2 = (\varepsilon_t^2 - E(\varepsilon_t^2))\sigma_t^2, \quad t \in \mathbb{Z}. \tag{26}$$

Then $(u_t)_{t\in\mathbb{Z}}$ is a White Noise sequence with mean zero and variance $E(\sigma_0^4)$ $\mathrm{Var}(\varepsilon_0^2)$, and

$$S_t := \sigma_t^2 - \frac{\alpha_0}{1 - (E\varepsilon_0^2)\sum_{i=1}^p \alpha_i - \sum_{j=1}^q \beta_j}, \quad t \in \mathbb{Z},$$

and

$$W_t := X_t^2 - \frac{\alpha_0 E\varepsilon_0^2}{1 - (E\varepsilon_0^2)\sum_{i=1}^p \alpha_i - \sum_{j=1}^q \beta_j}, \quad t \in \mathbb{Z},$$

satisfy the causal ARMA$(\max(p,q), p-1)$ *and causal* ARMA$(\max(p,q), q)$ *equations*

$$S_t - \sum_{i=1}^{\max(p,q)} ((E\varepsilon_0^2)\alpha_i + \beta_i)S_{t-i} = \sum_{i=1}^p \alpha_i u_{t-i}, \quad t \in \mathbb{Z},$$

and

$$W_t - \sum_{i=1}^{\max(p,q)} ((E\varepsilon_0^2)\alpha_i + \beta_i)W_{t-i} = u_t - \sum_{j=1}^q \beta_j u_{t-j}, \quad t \in \mathbb{Z},$$

respectively. Here, $\alpha_i = 0$ for $i > p$ and $\beta_j = 0$ for $j > q$. In particular, the autocovariance and autocorrelation functions of $(\sigma_t^2)_{t\in\mathbb{Z}}$ and that of $(X_t^2)_{t\in\mathbb{Z}}$ are those of the corresponding ARMA *processes.*

The fact that $(u_t)_{t\in\mathbb{Z}}$ is White Noise follows in complete analogy to the White Noise property of $(X_t)_{t\in\mathbb{Z}}$ by using (18). The ARMA representations then follow by inserting (26) into (2), and they are causal by Theorem 3. Observe that the ARMA equation for $(S_t)_{t\in\mathbb{Z}}$ is actually an ARMA$(\max(p,q), p'-1)$-equation driven by $(u_{t-p'})_{t\in\mathbb{Z}}$, where $p' := \min\{j \in \{1,\dots,p\} : \alpha_j \neq 0\}$. For general expressions for the autocoviarance functions of ARMA processes, see Brockwell and Davis (1991), Section 3.3.

5 Strong Mixing

Mixing conditions describe some type of asymptotic independence, which may be helpful in proving limit theorems, e.g. for the sample autocorrelation function or in extreme value theory. There exist many types of mixing conditions, see e.g. Doukhan (1994) for an extensive treatment. For GARCH processes, under weak assumptions one has a very strong notion of mixing, namely β-mixing, which in particular implies strong mixing: let $Y = (Y_t)_{t\in\mathbb{Z}}$ be a strictly stationary time series in $\mathbb{R}^d$, defined on a probability space $(\Omega, \mathcal{F}, P)$. Denote by $\mathcal{F}_{-\infty}^0$ the σ-algebra generated by $(Y_s : s \leq 0)$ and by $\mathcal{F}_t^\infty$ the σ-algebra generated by $(Y_s : s \geq t)$, and for $k \in \mathbb{N}$ let

$$\alpha_k^{(SM)} := \sup_{C\in\mathcal{F}_{-\infty}^0, D\in\mathcal{F}_k^\infty} |P(C\cap D) - P(C)P(D)|,$$

$$\beta_k^{(SM)} := \frac{1}{2}\sup \sum_{i=1}^{I}\sum_{j=1}^{J} |P(C_i \cap D_j) - P(C_i)P(D_j)|,$$

where in the definition of $\beta_k^{(SM)}$ the supremum is taken over all pairs of finite partitions $\{C_1, \ldots, C_I\}$ and $\{D_1, \ldots, D_J\}$ of Ω such that $C_i \in \mathcal{F}_{-\infty}^0$ for each i and $D_j \in \mathcal{F}_k^\infty$ for each j. The constants $\alpha_k^{(SM)}$ and $\beta_k^{(SM)}$ are the *α-mixing coefficients* and *β-mixing coefficients*, respectively, and $(Y_t)_{t\in\mathbb{Z}}$ is called *strongly mixing* (or *α-mixing*) if $\lim_{k\to\infty} \alpha_k^{(SM)} = 0$, and *$\beta$-mixing* (or *absolutely regular*) if $\lim_{k\to\infty} \beta_k^{(SM)} = 0$. It is *strongly mixing with geometric rate* if there are constants $\lambda \in (0,1)$ and c such that $\alpha_k^{(SM)} \leq c\lambda^k$ for every k, i.e. if α_k decays at an exponential rate, and *β-mixing with geometric rate* is defined similarly. Since

$$\alpha_k^{(SM)} \leq \frac{1}{2}\beta_k^{(SM)},$$

β-mixing implies strong mixing.

Based on results of Mokkadem (1990), Boussama (1998) showed that GARCH processes are beta mixing with geometric rate under weak assumptions, see also Boussama (2006). The proof hereby relies on mixing criteria for Markov chains as developed by Feigin and Tweedie (1985), see also Meyn and Tweedie (1996). Observe that the sequence $Y = (Y_t)_{t\in\mathbb{N}_0}$ of random vectors defined by (12) defines a discrete time Markov chain with state space $\mathbb{R}_+^{\widetilde{p}+\widetilde{q}-1}$. Boussama (1998) then shows that under suitable assumptions on the noise sequence this Markov chain is *geometrically ergodic*, i.e. there is a constant $\lambda \in (0,1)$ such that

$$\lim_{n\to\infty} \lambda^{-n} \|p_n(y,\cdot) - \pi(\cdot)\|_{TV} = 0.$$

Here, $p_n(y,E)$ for $y \in \mathbb{R}_+^{\tilde{p}+\tilde{q}-1}$ and $E \in \mathcal{B}(\mathbb{R}_+^{\tilde{p}+\tilde{q}-1})$ denotes the n-step transition probability from y to E, i.e.

$$p_n(y,E) = P(Y_n \in E | Y_0 = y),$$

π denotes the initial distribution of Y_0 which is chosen to be the stationary one, and $\|\cdot\|_{TV}$ denotes the total variation norm of measures. Since geometric ergodicity implies β-mixing of $(Y_t)_{t\in\mathbb{Z}}$ with geometric rate, using the causality it can be shown that this in turn implies β-mixing of $(\sigma_t, \varepsilon_t)_{t\in\mathbb{Z}}$ and hence of $(X_t)_{t\in\mathbb{Z}}$. Originally, the results in Boussama (1998) and (2006) are stated under the additional assumption that the noise sequence has finite second moment, but an inspection of the proof shows that it is sufficient to suppose that $E|\varepsilon_0|^s < \infty$ for some $s > 0$. The next Theorem gives the precise statements. See also Basrak et al. (2002), Corollary 3.5, and Mikosch and Straumann (2006), Theorem 4.5 and Proposition 4.10.

Theorem 8 (Boussama (1998), Théorème 3.4.2)
Let $(X_t, \sigma_t)_{t\in\mathbb{Z}}$ be a strictly stationary GARCH(p,q) *process as in* (1), (2), *and suppose the noise sequence is such that ε_0 is absolutely continuous with Lebesgue density being strictly positive in a neighbourhood of zero, and such that there exists some $s \in (0,\infty)$ such that $E|\varepsilon_0|^s < \infty$. Let Y_t be defined as in* (12). *Then $(Y_t)_{t\in\mathbb{Z}}$ is β-mixing with geometric rate. In particular, $(\sigma_t^2)_{t\in\mathbb{Z}}$, $(X_t^2)_{t\in\mathbb{Z}}$ and $(X_t)_{t\in\mathbb{Z}}$ are β-mixing and hence strongly mixing with geometric rate.*

An important application of strong mixing is the asymptotic normality of the sample autocovariance and autocorrelation function, under suitable moment conditions. Recall that the *sample autocovariance function* of a time series $(Z_t)_{t\in\mathbb{Z}}$ based on observations $Z_1, \ldots, Z_n$ is defined by

$$\gamma_{Z,n}(h) := \frac{1}{n}\sum_{t=1}^{n-h}(Z_t - \overline{Z}_n)(Z_{t+h} - \overline{Z}_n), \quad h \in \mathbb{N}_0,$$

where $\overline{Z}_n := \frac{1}{n}\sum_{t=1}^n Z_t$ denotes the sample mean. Similarly, the *sample autocorrelation function* is given by

$$\rho_{Z,n}(h) := \frac{\gamma_{Z,n}(h)}{\gamma_{Z,n}(0)}, \quad h \in \mathbb{N}_0.$$

If now $(Z_t)_{t\in\mathbb{Z}}$ is a strictly stationary strongly mixing time series with geometric rate such that $E|Z_t|^{4+\delta} < \infty$ for some $\delta > 0$, then for each $h \in \mathbb{N}_0$, also $(Z_t Z_{t+h})_{t\in\mathbb{Z}}$ is strongly mixing with geometric rate and $E|Z_t Z_{t+h}|^{2+\delta/2} < \infty$. Then a central limit theorem applies, showing that $\sqrt{n}\sum_{j=1}^n (Z_t Z_{t+h} - E(Z_t Z_{t+h}))$ converges in distribution to a mean zero normal random variable as $n \to \infty$, see e.g. Ibragimov and Linnik (1971), Theorem 18.5.3. More generally, using the Cramér-Wold device, one can show that the vector $(\sqrt{n}\sum_{j=1}^n (Z_t Z_{t+h} - E(Z_t Z_{t+h})))_{h=0,\ldots,m}$ converges for every $m \in \mathbb{N}$ to a multivariate normal distribution. Standard arguments as presented in Brockwell and Davis (1991), Section 7.3, then give multivariate asymptotic normality of the sample autocovariance function and hence of the autocorrelation function via the delta method. Applying these results to the GARCH process, we have:

Corollary 3 *Suppose that $(X_t, \sigma_t)_{t\in\mathbb{Z}}$ is a strictly stationary* GARCH *process whose noise sequence $(\varepsilon_t)_{t\in\mathbb{Z}}$ is such that ε_0 is absolutely continuous with Lebesgue density being strictly positive in a neighbourhood of zero.*
(a) If there is $\delta > 0$ such that $E|X_t|^{4+\delta} < \infty$, then the sample autocovariance and sample autocorrelation function of $(X_t)_{t\in\mathbb{Z}}$ are asymptotically normal with rate $n^{1/2}$, i.e. for every $m \in \mathbb{N}$ there exists a multivariate normal random vector $(V_0, \ldots, V_m)$ with mean zero such that $(\sqrt{n}(\gamma_{n,X}(h) - \gamma_X(h)))_{h=0,\ldots,m}$ converges in distribution to $(V_0, \ldots, V_m)$ as $n \to \infty$, and $(\sqrt{n}(\rho_{n,X}(h) - \rho_X(h)))_{h=1,\ldots,m}$ converges to $(\gamma_X(0))^{-1}(V_h - \rho_X(h)V_0)_{h=1,\ldots,m}$ as $n \to \infty$. Here, γ_X and ρ_X denote the true autocovariance and autocorrelation function of $(X_t)_{t\in\mathbb{Z}}$, respectively.
(b) If there is $\delta > 0$ such that $E|X_t|^{8+\delta} < \infty$, then the sample autocovariance and sample autocorrelation functions of $(X_t^2)_{t\in\mathbb{Z}}$ are asymptotically normal with rate $n^{1/2}$.

The above statement can for example be found in Basrak et al. (2002), Theorems 2.13 and 3.6. In practice one often estimates GARCH processes with parameters which are close to IGARCH. Hence the assumption on finiteness of $E|X_t|^{4+\delta}$ is questionable. Indeed, in cases when $EX_t^4 = \infty$, one often gets convergence of the sample autocovariance and autocovariance functions to stable distributions, and the rate of convergence is different from $\sqrt{n}$. For the ARCH(1) case, this was proved by Davis and Mikosch (1998), extended by Mikosch and Stărică (2000) to the GARCH$(1,1)$ case, and by Basrak et al. (2002) to general GARCH(p,q). See also Davis and Mikosch (2008).

6 Some Distributional Properties

In this section we shortly comment on two other properties of the strictly stationary solution, namely tail behaviour and continuity properties. We have already seen that the kurtosis of a GARCH process is always greater than or equal to the kurtosis of the driving noise sequence. Furthermore, Corollary 2 shows that under any reasonable assumption, a GARCH$(1,1)$ process will never have moments of all orders. Much more is true. Based on Kesten's (Kesten (1973)) powerful results on the tail behaviour of random recurrence equations (see also Goldie (1991) for a simpler proof in dimension 1), one can deduce that GARCH processes have Pareto tails under weak assumptions. For the ARCH(1) process this was proved by de Haan et al. (1989), for

GARCH$(1,1)$ by Mikosch and Stărică (2000), and for general GARCH(p,q) processes by Basrak et al. (2002). For a precise statement of these results, we refer to Corollary 1 in the article of Davis and Mikosch (2008) in this volume. For example, for a GARCH$(1,1)$ process with standard normal noise, it holds for the stationary solutions $(X_t, \sigma_t)_{t\in\mathbb{Z}}$,

$$\lim_{x\to\infty} x^{2\kappa} P(\sigma_0 > x) = c_\sigma,$$

$$\lim_{x\to\infty} x^{2\kappa} P(|X_0| > x) = c_\sigma E(|\varepsilon_0|^{2\kappa}), \quad \lim_{x\to\infty} x^{2\kappa} P(X_0 > x) = \frac{c_\sigma}{2} E(|\varepsilon_0|^{2\kappa}).$$

Here, κ is the unique solution in $(0,\infty)$ to the equation

$$E(\alpha_1 \varepsilon_0^2 + \beta_1)^\kappa = 1,$$

and c_σ is a strictly positive constant.

Regarding continuity properties of stationary solutions of GARCH(p,q) processes, we shall restrict us to the case of GARCH$(1,1)$ and ARCH(1). Observe that in that case, the strictly stationary solution satisfies the random recurrence equation

$$\sigma_t^2 = \alpha_0 + (\beta_1 + \alpha_1 \varepsilon_{t-1}^2)\sigma_{t-1}^2.$$

Hence if ε_0 is absolutely continuous, so is $\log(\beta_1 + \alpha_1 \varepsilon_{t-1}^2) + \log \sigma_{t-1}^2$ by independence of ε_{t-1} and σ_{t-1}, and we conclude that σ_t^2 must be absolutely continuous. It follows that absolute continuity of ε_0 leads to absolute continuity of the stationary σ_t and hence of the stationary X_t. Excluding the case when ε_0^2 is constant, i.e. when the distribution of σ_t^2 is a Dirac measure, one might wonder whether the stationary distribution σ_t will always be absolutely continuous, regardless whether ε_0 is absolutely continuous or not. For stationary distributions of the related continuous time GARCH processes (COGARCH) introduced by Klüppelberg et al. (2004), this is indeed the case, see Klüppelberg et al. (2006). For the discrete time GARCH$(1,1)$ process, the author is however unaware of a solution to this question. At least there is the following positive result which is an easy consequence of Theorem 1 of Grincevicius (1980):

Theorem 9 (Grincevicius (1980), Theorem 1)
Let $(X_t, \sigma_t)_{t\in\mathbb{Z}}$ be a strictly stationary GARCH$(1,1)$ *or* ARCH(1) *process. Then σ_0 is continuous with respect to Lebesgue measure, i.e. cannot have atoms, unless σ_0 is degenerate to a constant, i.e. unless ε_0^2 is constant. Consequently, X_0 does not have atoms unless ε_0^2 is constant or ε_0 has an atom at zero.*

Actually, Grincevičius' result applies to more general situations, but in the GARCH case says that if $\sigma_0^2 = \alpha_0 \sum_{i=1}^\infty \prod_{j=1}^{i-1} (\beta_1 + \alpha_1 \varepsilon_{-j}^2)$ has an atom, then there must exist a sequence $(S_n)_{n\in\mathbb{N}_0}$ such that $\prod_{n=1}^\infty P(\alpha_0 + (\beta_1 + \alpha_1 \varepsilon_n^2) S_n =$

$S_{n-1}) > 0$. By the i.i.d. assumption on $(\varepsilon_n)_{n\in\mathbb{Z}}$, this can be seen to happen only if ε_0^2 is constant.

7 Models Defined on the Non-Negative Integers

We defined a GARCH process as a time series indexed by the set $\mathbb{Z}$ of integers. This implies that the process has been started in the infinite past. It may seem more natural to work with models which are indexed by the non-negative integers $\mathbb{N}_0$. Let $(\varepsilon_t)_{t\in\mathbb{N}_0}$ be a sequence of i.i.d. random variables, and $p \in \mathbb{N}$, $q \in \mathbb{N}_0$. Further, let $\alpha_0 > 0$, $\alpha_1, \ldots, \alpha_{p-1} \geq 0$, $\alpha_p > 0$, $\beta_1, \ldots, \beta_{q-1} \geq 0$ and $\beta_q > 0$ be non-negative parameters. Then by a GARCH(p,q) *process indexed by* $\mathbb{N}_0$, we shall mean a process $(X_t)_{t\in\mathbb{N}_0}$ with *volatility process* $(\sigma_t)_{t\in\mathbb{N}_0}$ which is a solution to the equations

$$X_t = \sigma_t \varepsilon_t, \quad t \in \mathbb{N}_0, \tag{27}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2, \quad t \geq \max(p,q). \tag{28}$$

The process is called *causal* if additionally σ_t^2 is independent of $(\varepsilon_{t+h})_{h\in\mathbb{N}_0}$ for $t = 0, \ldots, \max(p,q)$. By (28), the latter independence property then easily extends to hold for all $t \in \mathbb{N}_0$.

Recall that every strictly stationary GARCH(p,q) process indexed by $\mathbb{Z}$ is causal by Remark 1. When restricting such a process to $\mathbb{N}_0$, it is clear that we obtain a causal strictly stationary GARCH process indexed by $\mathbb{N}_0$. Conversely, suppose that $(X_t, \sigma_t)_{t\in\mathbb{N}_0}$ is a strictly stationary GARCH process indexed by $\mathbb{N}_0$. Like any strictly stationary process indexed by $\mathbb{N}_0$, it can be extended to a strictly stationary process $(X_t, \sigma_t)_{t\in\mathbb{Z}}$, see Kallenberg (2002), Lemma 10.2. With $\varepsilon_t = X_t/\sigma_t$ for $t < 0$ (observe that $\sigma_t^2 \geq \alpha_0$), one sees that also $(X_t, \sigma_t, \varepsilon_t)_{t\in\mathbb{Z}}$ is strictly stationary. Hence $(\varepsilon_t)_{t\in\mathbb{Z}}$ must be i.i.d., and (27) and (28) continue to hold for $t \in \mathbb{Z}$. Since $(X_t, \sigma_t)_{t\in\mathbb{Z}}$ is strictly stationary, it is causal, and hence so is $(X_t, \sigma_t)_{t\in\mathbb{N}_0}$.

We have seen that there is an easy correspondence between strictly stationary GARCH processes defined on the integers and strictly stationary GARCH processes defined on $\mathbb{N}_0$. This justifies the restriction to GARCH processes indexed by $\mathbb{Z}$, which are mathematically more tractable. Furthermore, strictly stationary GARCH processes indexed by $\mathbb{N}_0$ are automatically causal.

8 Conclusion

In the present paper we have collected some of the mathematical properties of GARCH(p,q) processes $(X_t, \sigma_t)_{t\in\mathbb{Z}}$. The existence of strictly and weakly stationary solutions was characterised, as well as the existence of moments. The GARCH process shares many of the so called *stylised features* observed in financial time series, like a time varying volatility or uncorrelatedness of the observations, while the squared observations are not uncorrelated. The autocorrelation of the squared sequence was in fact seen to be that of an ARMA process. Stationary solutions of GARCH processes have heavy tails, since they are Pareto under weak assumptions. On the other hand, there are some features which are not met by the standard GARCH(p,q) process, such as the leverage effect, to name just one. In order to include these and similar effects, many different GARCH type models have been introduced, such as the EGARCH model by Nelson (1991), or many other models. We refer to the article by Teräsvirta (2008) for further information regarding various extensions of GARCH processes.

Acknowledgement I would like to thank Richard Davis and Thomas Mikosch for careful reading of the paper and many valuable suggestions.

References

Ash, R.B. and Gardner, M.F. (1975): *Topics in Stochastic Processes Vol. 27 of Probability and Mathematical Statistics.* Academic Press, New York.

Basrak, B., Davis, R.A. and Mikosch, T. (2002): Regular variation of GARCH processes. *Stochastic Processes and Their Applications* **99**, 95–115.

Brandt, A. (1986): The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients. *Advances in Applied Probability* **18**, 211–220.

Brandt, A., Franken, P.and Lisek, B. (1990): *Stationary Stochastic Models.* Wiley, Chichester.

Brockwell, P.J. and Davis, R.A. (1991): *Time Series: Theory and Methods.* 2nd edition. Springer, Berlin Heidelberg New York.

Bollerslev, T. (1986): Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **31**, 307–327.

Bollerslev, T. (1988): On the correlation structure for the generalized autoregressive conditional heteroskedastic process. *Journal of Time Series Analysis* **9**, 121–131.

Bougerol, P. and Picard, N. (1992a): Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics* **52**, 115–127.

Bougerol, P. and Picard, N. (1992b): Strict stationarity of generalized autoregressive process. *The Annals of Probability* **20**, 1714–1730.

Boussama, F. (1998): *Ergodicité, mélange et estimation dans le modelés GARCH.* Ph.D. Thesis, Université 7 Paris.

Boussama, F. (2006): Ergodicité des chaînes de Markov à valeurs dans une variété algébraique: application aux modéles GARCH multivariés. *Comptes Rendus. Mathématique. Académie des Sciences* **343**, 275–278.

Davis, R.A. and Mikosch, T. (1998): Limit theory for the sample ACF of stationary processes with heavy tails with applications to ARCH. *The Annals of Statistics* **26**, 2049–2080.

Davis, R.A. and Mikosch, T. (2008): Extreme value theory for GARCH processes. *In: Andersen, T.G., Davis, R.A., Kreiss, J.-P. and Mikosch, T. (Eds.): Handbook of Financial Time Series*, 186–200. Springer, New York.

Doukhan, P. (1994): *Mixing. Properties and Examples. Lecture Notes in Statistics* **85**, Springer, Berlin Heidelberg New York.

Engle, R.F. (1982): Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50**, 987–1008.

Engle, R.F. and Bollerslev, T. (1986): Modelling the persistence of conditional variances. With comments and a reply by the authors. *Econometric Reviews* **5**, 1–87.

Erickson, K.B. (1973): The strong law of large numbers when the mean is undefined. *Transactions of the American Mathematical Society* **185**, 371–381.

Feigin, P.D. and Tweedie, R.L. (1985): Random coefficient autoregressive processes: a Markov chain analysis of stationarity and finiteness of moments. *Journal of Time Series Analysis* **6**, 1–14.

Furstenberg, H. and Kesten, H. (1960): Products of random matrices. *Annals of Mathematical Statistics* **31**, 457–469.

Giraitis, L., Kokoszka, P. and Leipus, R. (2000): Stationary ARCH models: dependence structure and central limit theorem. *Econometric Theory* **16**, 3–22.

Giraitis, L., Leipus, R. and Surgailis, D. (2006): Recent advances in ARCH modelling. *In: Teyssière, G. and Kriman, A. (Eds): Long Memory in Economics*, 3–38. Springer, Berlin Heidelberg New York.

Giraitis, L., Leipus, R. and Surgailis, D (2008): ARCH(∞) models and long memory properties. *In: Andersen, T.G., Davis, R.A., Kreiss, J.-P. and Mikosch, T. (Eds.): Handbook of Financial Time Series*, 70–84. Springer, New York.

Goldie, C.M. (1991): Implicit renwal theory and tails of solutions of random equations. *The Annals of Applied Probobility* **1**, 126–166.

Goldie, C.M. and Maller, R.A. (2000): Stability of perpetuities. *The Annals of Probability* **28**, 1195–1218.

Grincevičius, A.K. (1980): Products of random affine transformations. *Lithuanian Mathematical Journal* **20**, 279–282.

de Haan, L., Resnick, S.I., Rootzén, H. and de Vries, C.G. (1989): Extremal behaviour of solutions to a stochastic difference equation with applications to ARCH processes. *Stochastic Processes and Their Applications* **32**, 213–224.

He, C. and Teräsvirta, T. (1999a): Fourth moment structure of the GARCH(p,q) process. *Econometric Theory* **15**, 824–846.

He, C. and Teräsvirta, T. (1999b): Properties of moments of a family of GARCH processes. *Journal of Econometrics* **92**, 173–192.

Ibragimov, I.A. and Linnik, Yu.V. (1971): *Independent and Stationary Sequences of Random Variables.* Wolters-Noordhoff, Groningen.

Kallenberg, O. (2002): *Foundations of Modern Probability. 2nd edition.* Springer, Berlin Heidelberg New York.

Karanasos, M. (1999): The second moment and the autovariance function of the squared errors of the GARCH model. *Journal of Econometrics* **90**, 63–67.

Kesten, H. (1973): Random difference equations and renewal theory of products of random matrices. *Acta Mathematica* **131**, 207–248.

Kesten, H. and Maller, R. (1996): Two renewal theorems for general random walk tending to infinity. *Probability Theory and Related Fields* **106**, 1–38.

Kesten, H. and Spitzer, F. (1984): Convergence in distribution for products of random matrices. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **67**, 363–386.

Klüppelberg, C., Lindner, A. and Maller, R. (2004): A continuous-time GARCH process driven by a Lévy process: stationary and second order behaviour. *Journal of Applied Probability* **41**, 601–622.

Klüppelberg, C., Lindner, A. and Maller, R. (2006): Continuous time volatility modelling: GARCH versus Ornstein models. *In: Kabanov, Yu., Liptser, R. and Stoyanov, J. (Eds.): From Stochastic Calculus to Mathematical Finance. The Shiryaev Festschrift*, 393–419. Springer, Berlin New York Heidelberg.

Ling, S. (1999): On the probabilistic properties of a double threshold ARMA conditional heteroskedastic model. *Journal of Applied Probability* **36**, 688–705.

Ling, S. and McAleer, M. (2002): Necessary and sufficient moment conditions for the GARCH(r,s) and asymmetric power GARCH(r,s) models. *Econometric Theory* **18**, 722–729.

Lütkepohl, H. (1996): *Handbook of Matrices.* Wiley, Chichester.

Meyn, S.P. and Tweedie, R.L. (1996): *Markov Chains and Stochastic Stability.* 3rd edition. Springer, Berlin Heidelberg New York.

Milhøj, A. (1985): The moment structure of ARCH processes. *Scandinavian Journal of Statistics* **12**, 281–292.

Mikosch, T. and Stărică, C. (2000): Limit theory for the sample autocorrelations and extremes of a GARCH(1,1) process. *The Annals of Statistics* **28**, 1427–1451.

Mikosch, T. and Straumann, D. (2006): Stable limits of martingale transforms with application to the estimation of GARCH parameters. *The Annals of Statistics* **34**, 493–522.

Mokkadem, A. (1990): Propriétés de mélange des processus autorégressifs polynomiaux. *Annales de l'Institut Henri Poincaré. Probabilités et Statistique* **26**, 219–260.

Nelsen, D.B. (1990): Stationarity and persistence in the GARCH(1,1) model. *Econometric Theory* **6**, 318–334.

Nelson, D.B. (1991): Conditional heteroskedasticity in asset returns: a new approach. *Econometrica* **59**, 347–370.

Nelson, D.B. and Cao, C.Q. (1992): Inequality constraints in the univariate GARCH model. *Journal of Business and Economic Statistics* **10**, 229–235.

Robinson, P.M. (1991): Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. *Journal of Econometrics* **47**, 67–84.

Straumann, D. (2005): *Estimation in Conditionally Heterscedastic Time Series Models. Lecture Notes in Statistics* **181**, Springer, Berlin.

Teräsvirta, T. (2008): An introduction to univariate GARCH models. *In: Andersen, T.G., Davis, R.A., Kreiss, J.-P. and Mikosch, T. (Eds.): Handbook of Financial Time Series*, 17–42. Springer, New York.

Vervaat, W. (1979): On a stochastic difference equation and a representation of nonnegative infinitely divisible random variables. *Advances in Applied Probability* **11**, 750–783.

http://www.springer.com/978-3-540-71296-1

Handbook of Financial Time Series
Andersen, T.G.; Davis, R.A.; Kreiss, J.-P.; Mikosch, Th.V. (Eds.)
2009, XXIX, 1050 p., Hardcover
ISBN: 978-3-540-71296-1