Penn Institute for Economic Research
Department of Economics
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19103-6297

pier@ssc.upenn.edu
http://www.econ.upenn.edu/pier

PIER Working Paper 97-033

"Degenerate Continuous Time Limits of GARCH and GARCH-type Processes"

by

Valentina Corradi

Degenerate Continuous Time Limits of GARCH and GARCH-type Processes

by Valentina Corradi*

University of Pennsylvania
Department of Economics
3718 Locust Walk
Philadelphia, PA 19104-6297
e-mail: corradi@econ.sas.upenn.edu
tel. (215) 898-1505

First Draft: November, 1996
This version: July 1997

Keywords: degenerate diffusion, diffusion approximation, GARCH, modified semi-strong GARCH, support of the transition function

*I owe special thanks to Torben Andersen for very helpful comments and suggestions and for having pointed out a mistake in an earlier version. I also wish to thank Rob Engle, Dean Foster, Christian Gourieroux, Grant Hillier as well as the seminar participants at University of Southampton, CORE and 1997 Winter Meeting of the Econometric Society for very useful comments and suggestions.

Abstract

The purpose of this paper is twofold. On one hand we reconsider the continuous time limit of the GARCH(1,1) process and show that, by choosing different reparameterizations, as a function of the discrete interval h, we obtain, as $h \to 0$, either a non-degenerate or a degenerate diffusion limit. On the other hand, we introduce a class of GARCH-type processes, that we call "modified semi-strong GARCH" and derive their continuous time limit. Special attention is paid to the analysis of the support of the transition function of the different diffusion limits we obtain.

1 Introduction

It is a well known fact that asset returns exhibit conditional heteroskedasticity; that is big shocks tend to cluster together. ARCH models are discrete time, parsimonious parametric models able to reproduce such a stylized fact. For this reason, since the seminal paper of Engle (1982), ARCH models have been extensively used in empirical analysis of financial data. Also several extensions of the basic ARCH models allow to reproduce other stylized facts, such as heavy tails and long memory properties (see Journal of Econometrics 1996, issue on ARCH and long memory). However in modern financial theory, in general asset prices are modeled according to continuous time processes. One reason for modeling asset prices in continuous time is that exact option replication is infeasible in discrete time. Among the class of continuous time processes, typically asset prices and returns are modeled according to diffusion processes; the nondifferentiability of the paths well captures the absence of arbitrage opportunities in frictionless markets (e.g. Harrison, Pitbladdo and Schaefer 1984). In order to fill the gap between the use of diffusion processes in the theoretical financial literature and the use of discrete time ARCH models in the empirical financial literature, Nelson (1990) studied the continuous time limit of ARCH models, as the discrete interval, say h, approaches zero. More recently Fornari and Mele (1996) analyzed the continuous time limit for a class of nonlinear ARCH models proposed by Ding, Granger and Engle (1993).

In his seminal paper, Nelson (1990) showed that different classes of GARCH processes. e.g. GARCH(1,1) and exponential ARCH, EARCH, after a proper reparameterization, as the time interval shrinks, converge in distribution to a two-dimensional non-degenerate diffusion; i.e. to a diffusion in $R \times R^{+}$ driven by two Brownian motions, whose covariance matrix is non-singular. Furthermore, Nelson obtains the limiting distribution of the volatility process as the time span goes to infinity.

The aim of this paper is twofold. On one hand we reconsider the continuous time limit of the GARCH(1,1) process, on another hand we introduce a new class of GARCH-type processes and derive their continuous time limit. Special attention will be paid to the analysis of the support of the transition functions of the different diffusion limits we obtain.

We begin by showing that Nelson's (1990) result, although technically correct, is the outcome of a very particular reparameterization; more precisely it hinges on the fact that, when we redefine the parameters as a function of the time interval h, the coefficient on the autoregressive term is linked in a very special manner to the coefficient on the moving average term. This particular reparameterization leads to a non-degenerate limit, i.e. a two-dimensional diffusion driven by two independent Brownian motions. Then we show that, by choosing a somewhat different parameterization, we obtain a degenerate continuous time limit, i.e. a two-dimensional diffusion driven by only one Brownian motion. It should be stressed that by redefining the parameters as a particular function of the discrete interval h, we are indeed choosing a particular approximation scheme. Thus in our context to choose a particular parameterization, as a function of h, is equivalent to choose a particular approximation scheme.

While in principle there is no reason why a particular parameterization, or approximation scheme, should be preferable to another, the parameterization we propose has a nice feature. Broadly speaking, we start from a GARCH(1,1) process, we redefine the parameters in terms of the discrete interval h, we take the limit as $h \to 0$ and obtain a degenerate diffusion limit; then if we take an Euler approximation (Euler approximations are not unique) of such degenerate diffusion we obtain the same GARCH(1,1) process we started from. In this sense GARCH(1,1) models can be seen as Euler approximations to degenerate diffusion. On the other hand, any Euler approximation of a non-degenerate diffusion limit leads to a processes characterized by stochastic (non predictable) volatility.

Both the Nelson's non degenerate limit and the degenerate limit we obtain show a drawback: they are the continuous time limit of the couple (Y_k, σ_k^2), where Y_k are cumulative returns, and σ_k^2 is a F_k-measurable process which represents the volatility of Y_{k+1}. This is due to the particular way we write down the discrete time GARCH(1,1) process we start from; in fact following Bollerslev, Engle and Nelson (1994), we use the following specification $Y_k - Y_{k-1} = \sigma_{k-1}\epsilon_k$ and $\sigma_k^2 = \omega_0 + \omega_1\sigma_{k-1}^2 + \omega_2\sigma_{k-1}^2\epsilon_k^2$. Then we consider the continuous approximation of (Y_k, σ_k^2), but σ_k^2 is clearly the volatility at time $k+1$. Also the support properties of the diffusion limit we obtain are not completely satisfactory: in Nelson's case the support is the entire state space, say $R \times R^+$, in our case the support is just a line in $R \times R^+$, in that, for given initial conditions,

σ_t^2 is a deterministic processes. Thus we introduce a new class of GARCH- type processes that we shall call "modified semi-strong GARCH" for their similarity with the semi-strong GARCH processes defined by Drost and Nijman (1993). For such models we derive the continuous time limit for the couple (Y_k, σ_k^2), where now σ_k^2 is indeed the conditional variance of Y_k and not of Y_{k+1}. The continuos time limit is still a degenerate diffusion, but different from the one obtained from the GARCH(1,1) model. In particular it is a process with differentiable paths, but no longer F_0-measurable. The support of the transition function is a two-dimensional subset of the state space, but with possibly empty interior.

Degenerate and non-degenerate diffusions have also different interpretations from an economic/financial point of view. In the degenerate diffusion case, the cumulative return process has paths of unbounded variation, and so nowhere differentiable, however the volatility process has paths of bounded variation and so differentiable. Now continuous time processes with smooth paths are characterized by significative positive correlation between successive increments of time (e.g. Harrison, Pitbladdo and Schaefer 1984), thus changes in volatility can be predicted even over very short intervals. This property seems to be the natural continuous time counterpart of the property of predictability of volatility in discrete time. Also if we do not consider volatility as a tradeable asset, then the degeneracy of the diffusion preserves the completeness of the market, thus allowing for unique preference independent prices for contingent claims (see Hobson and Rogers, 1994); on the other hand non-degenerate diffusions lead to market incompleteness. Of course if instead volatility were a tradeable asset, then non-degenerate diffusions, i.e. two state variables driven by two non perfectly correlated Brownian motions, would preserve market completeness.

A related area of research (Drost and Werker (1996), Meddahi and Renault (1996)) have analyzed the link between continuous time and discrete time volatility models from a different perspective; that is starting from a non-degenerate diffusion they derive and analyze the implied discrete time volatility process. Drost and Werker (1996) show that the implied discrete time volatility process falls into the class of weak GRACH procesesses (as defined in Drost and Nijman, 1993); while Meddahi and Renault (1996) show that the implied discrete time volatility process falls into the class of $SR - SARV(1)$ processes (square root autoregressive stochastic volatility of order 1). Meddahi and Renault also show that (semi) strong GARCH processes (as defined in Drost and Nijman 1993) are $SR - SARV(1)$ processes subject to few restrictions; however while these restrictions hold for a given frequency, they are not robust to temporal aggregation. Furthermore the $SR - SARV(1)$ processes, implied by non-degenerate diffusions, do not satisfy these restrictions. Thus the discrete processes implied by non-degenerate diffusions are characterized by stochastic volatility; in fact, as pointed out by Meddahi and Renault, also weak GARCH models are characterized by stochastic (non predictable) volatility. As we are mainly concerned about the continuous time counterparts of volatility models characterized by predictable volatility, we won't follow the approach suggested by Drost and Werker (1996) and Meddahi and Renault (1996).

The paper is organized as follows. Section 2 analyzes the continuous time limit of the GARCH(1,1) process via two different reparameterizations, as a function of the time interval h. Section 3 analyzes GARCH models as diffusion approximations. Section 4 introduces a new class GARCH-type models and derive their continuos time limit. Section 5 gives some definitions and facts concerning the support, the accessibility and the controllability of nonlinear degenerate diffusions. Finally in Section 6, using the tools described in Section 5, we analyze the support of the transition function of the degenerate diffusion limits we obtain.

2 Continuous time limits of the GARCH(1,1) process

This section considers two different continuous time approximations of the GARCH(1,1) process, one leading to a degenerate diffusion and another leading to a non-degenerate diffusion (Nelson's result).

Let $Y_k - Y_{k-1}$, $k = 1, 2, \ldots$ be returns on a generic financial asset and so let Y_k denote the cumulative returns. We begin by considering the following discrete time GARCH(1,1) process, written as in Bollerslev et al. (1994),

$$Y_k - Y_{k-1} = \sigma_{k-1}\epsilon_k \tag{1}$$

$$\sigma_k^2 = \omega_0 + \omega_1\sigma_{k-1}^2 + \omega_2\sigma_{k-1}^2\epsilon_k^2 \tag{2}$$

with $\omega_0, \omega_1, \omega_2 > 0$ and $\omega_1 + \omega_2 < 1$, $\epsilon_k \sim iidN(0,1)$.

Let $F_k = \sigma(Y_0, \sigma_0^2, Y_1, \ldots Y_k)$, now σ_k^2 is F_k-measurable; however $E((Y_k - Y_{k-1})^2 | F_{k-1}) = \sigma_{k-1}^2$, so that the conditional variance at time k is F_{k-1}-measurable, thus preserving the predictability of volatility; simply σ_k^2 is the conditional variance at time $k+1$. Let h be the discrete time interval, and consider the following approximation scheme

$$Y_{kh} - Y_{(k-1)h} = \sigma_{(k-1)h}\epsilon_{kh} \tag{3}$$

$$\sigma_{kh}^2 - \sigma_{(k-1)h}^2 = \omega_{0h} + (\omega_{1h} - 1)\sigma_{(k-1)h}^2 + \omega_{2h}\sigma_{(k-1)h}^2\epsilon_{kh}^2 \tag{4}$$

where $\epsilon_{kh} \sim iidN(0,h)$.

We shall compute the first two conditional moments, and then, after few mild technical conditions, appeal to the theorems for weak convergence of Markov chains to diffusion processes by Strook and Varadhan (1979, Ch.11) or by Ethier and Kurtz (1986, Ch.8). Now let $F_{(k-1)h} = \sigma(Y_{(k-1)h}, \sigma_{(k-1)h}^2)$, note that $\forall k$ and for any given h, $((Y_k - Y_{(k-1)h}), (\sigma_{(k-1)h}^2, \sigma_{kh}^2))$ are first order Markov chains. Now

$$h^{-1}E((Y_{kh} - Y_{(k-1)h})|F_{(k-1)h}) = 0 \tag{5}$$

$$h^{-1}E((Y_{kh} - Y_{(k-1)h})^2|F_{(k-1)h}) = \sigma_{(k-1)h}^2 \tag{6}$$

$$h^{-1}E((\sigma_{kh}^2 - \sigma_{(k-1)h}^2)|F_{(k-1)h}) = h^{-1}\omega_{0h} + h^{-1}(\omega_{1h} - 1)\sigma_{(k-1)h}^2 + \omega_{2h}^2\sigma_{(k-1)h}^2 \tag{7}$$

$$(8) h^{-1}E((\sigma_{kh}^2 - \sigma_{(k-1)h}^2)^2|F_{(k-1)h}) = h^{-1}\omega_{0h}^2 + h^{-1}(\omega_{1h} - 1)^2\sigma_{(k-1)h}^4 + 3h^{-1}h^2\omega_{2h}\sigma_{(k-1)h}^2$$
$$+2h^{-1}\omega_{0h}(\omega_{1h} - 1)\sigma_{(k-1)h}^2 + 2\omega_{0h}\omega_{2h}\sigma_{(k-1)h}^2 + 2(\omega_{1h} - 1)\omega_{2h}\sigma_{(k-1)h}^4$$

We shall consider the following parameterization, as a function of h,

$$\omega_{0h} = h\omega_0 \tag{9}$$

$$\lim_{h \to 0} h^{-1}(\omega_{1h} - 1) = \theta < 0 \tag{10}$$

$$\omega_{2h} = \omega_2, \forall h \tag{11}$$

Note that given the parameterization in (9)-(11), the right hand side of (8) is of probability order h and so converges to zero as $h \to 0$.

Now let consider the following right continuous with left limit (CADLAG) processes, $Y_t^h =$

$Y_{kh}, \sigma_t^{2,h} = \sigma_{kh}^2, kh \leq t < (k+1)h$, where Y_{kh}, σ_{kh}^2 are defined as in (3)(4). Hereafter $\Rightarrow$ denotes weak convergence, P_0^h, P_0 denotes the probability measure of $(Y_0^h, \sigma_0^{2,h})$ and of (Y_0, σ_0^2) respectively.

PROPOSITION 2.1

If as $h \to 0$, $(Y_0^h, \sigma_0^{2,h}) \to (Y_0, \sigma_0^2)$ or $P_0^h \to P_0$, then, under the parameterization in (9)-(11), as $h \to 0$, $(Y_t^h, \sigma_t^{2,h}) \Rightarrow (Y_t, \sigma_t^2)$, where (Y_t, σ_t^2) is a diffusion process solution to

$$dY_t = \sigma_t dW_t \tag{12}$$

$$d\sigma_t^2 = (\omega_0 + (\theta + \omega_2)\sigma_t^2)dt \tag{13}$$

□

PROOF: see Appendix

Thus the parameterization chosen in (9)-(11) leads to weak convergence to a degenerate diffusion, that is to a diffusion with two state variables driven by only one Brownian motion.

We now show that there exists a very special parameterization (Nelson's) that leads to a non-degenerate diffusion, that is to a diffusion with two state variables driven by two non perfectly correlated Brownian motions. Starting from (1)-(2), we consider the following (Nelson's) approximation,

$$Y_{kh} - Y_{(k-1)h} = \sigma_{(k-1)h}\epsilon_{kh} \tag{14}$$

$$\sigma_{kh}^2 - \sigma_{(k-1)h}^2 = \omega_{0h} + (\omega_{1h} - 1)\sigma_{(k-1)h}^2 + h^{-1}\omega_{2h}\sigma_{(k-1)h}^2\epsilon_{kh}^2 \tag{15}$$

with $\epsilon_{kh} \sim iidN(0,h)$. The first two conditional moments of $Y_{kh} - Y_{(k-1)h}$ are as in (5)-(6), thus we concentrate on the first two conditional moments of $\sigma_{kh}^2 - \sigma_{(k-1)h}^2$, we have

$$h^{-1}E((\sigma_{kh}^2 - \sigma_{(k-1)h}^2)|F_{(k-1)h}) = h^{-1}\omega_{0h} + h^{-1}(\omega_{1h} - 1 + \omega_{2h})\sigma_{(k-1)h}^2 \tag{16}$$

$$h^{-1}E((\sigma_{kh}^2 - \sigma_{(k-1)h}^2)^2|F_{(k-1)h}) = h^{-1}\omega_{0h}^2 + h^{-1}(\omega_{1h} - 1 + \omega_{2h})^2\sigma_{(k-1)h}^4 \tag{17}$$

$$+2h^{-1}\omega_{2h}^2\sigma_{(k-1)h}^4 + 2h^{-1}\omega_{0h}(\omega_{1h} - 1)\sigma_{(k-1)h}^2 + 2h^{-1}\omega_{0h}\omega_{2h}\sigma_{(k-1)h}^2$$

Now ω_{0h} is as in (9), instead Nelson requires that

$$\lim_{h \to 0} 2h^{-1}\omega_{2h}^2 = \alpha^2 \tag{18}$$

$$\lim_{h \to 0} h^{-1}(\omega_{1h} - 1 + \omega_{2h}) = \theta \tag{19}$$

with $\theta < 0$.

From (18), we see that ω_{2h} is of order $h^{1/2}$; thus the last term on the right side of (15) is of probability order $h^{1/2}$, recalling that ϵ_{kh}^2 is of probability order h; on the other hand the last term on the right hand side of (4) is of probability order h. This is the key difference that leads to a non-degenerate and to a degenerate diffusion limit respectively. In fact the continuous approximation chosen in (15) and the parameterization chosen in (18)(19) assure that the second conditional moment does not vanish as h goes to zero. As ω_{2h} is of order $h^{1/2}$, when scaled by h^{-1} diverges at rate $h^{-1/2}$. Now (19) basically implies that

$$\omega_{1h} - 1 = \theta h - \omega_{2h} \tag{20}$$

and so as $h \to 0$,

$$h^{-1}(\omega_{1h} - 1 + \omega_{2h}) \to \theta$$

In fact without the restriction imposed by (20), (19) cannot hold and the first conditional moment, scaled by h^{-1}, will tend to diverge at rate $h^{-1/2}$. Needless to say if the scaled first conditional moment diverges, we cannot have any weak convergence to a diffusion process. As in discrete time, except for the restriction $\omega_1 + \omega_2 < 1$, there is no cross restrictions between the values taken by ω_1 and ω_2, condition (20) does not seem very "natural". However Nelson's weak convergence result hinges on such a condition.

Summarizing, given the parameterization in (18)-(19), it is easy to see that the right hand side of (17) does not vanish as $h \to 0$, in fact it is equal to $2h^{-1}\omega_{2h}^2\sigma_{(k-1)h}^2$ plus a term vanishing with h. It is also immediate to see that

$$h^{-1}E((\sigma_{kh}^2 - \sigma_{(k-1)h})(Y_{kh} - Y_{(k-1)h})|Y_{(k-1)h} = y, \sigma_{(k-1)h}^2 = \sigma^2) = O(\sqrt{h}) \tag{21}$$

Let's define the right continuous, with left limit, approximations $Y_t^h = Y_{kh}$ and $\sigma_t^{2,h} = \sigma_{kh}^2$, for $kh \leq t < (k+1)h$, where Y_{kh} and σ_{kh}^2 are as (14)-(15). We have

PROPOSITION 2.2

If as $h \to 0$, $(Y_0^h, \sigma_0^{2,h}) \to (Y_0, \sigma_0^2)$, or $P_0^h \to P_0$, then given (20) and the parameterization in (9),(18)-(19), as $h \to 0$, $(Y_t^h, \sigma_t^{2,h}) \Rightarrow (Y_t, \sigma_t^2)$, where (Y_t, σ_t^2) is a diffusion process solution to

$$dY_t = \sigma_t dW_{1t} \tag{22}$$

$$d\sigma_t^2 = (\omega_0 + \theta\sigma_t^2)dt + \alpha\sigma_t^2 dW_{2t} \tag{23}$$

where $\theta < 0$, and (W_{1t}, W_{2t}) are two standard independent Brownian motions. □

PROOF: see Appendix

Note that, because of the specific restriction implied by (20), Proposition 2.1 is not a special case of Proposition 2.2, obtained by setting $\lim_{h\to 0} 2h^{-1}\omega_{2h}^2 = o(\alpha^2)$.

3 GARCH as diffusion approximations

We now want to analyze GARCH(1,1) processes as approximation of diffusion processes. We shall see that there exists an Euler approximation of the diffusion in (12)-(13) that is indeed a GARCH(1,1) processes. Let's consider again

$$dY_t = \sigma_t dW_t$$

$$d\sigma_t^2 = (\omega_0 + \theta\sigma_t^2)dt + \omega_2\sigma_t^2 dt$$

Now because of the conventional multiplication tables of stochastic calculus (see e.g. Karatzas and Shreve 1991, p. 153),

$$(dY_t)^2 = \sigma_t^2 (dW_t)^2 = \sigma_t^2 dt$$

So we can rewrite the degenerate diffusion above as,

$$dY_t = \sigma_t dW_t \tag{24}$$

$$d\sigma_t^2 = (\omega_0 + \theta\sigma_t^2)dt + \omega_2 (dY_t)^2 \tag{25}$$

We now consider an Euler discrete approximation to (24)(25),

$$Y_{kh} - Y_{(k-1)h} = \sigma_{(k-1)h}\epsilon_{kh}$$

$$\sigma_{kh}^2 - \sigma_{(k-1)h}^2 = h\omega_0 + \theta h \sigma_{(k-1)h}^2 + \omega_2 \sigma_{(k-1)h}^2 \epsilon_{kh}^2$$

with $\epsilon_{kh} \sim iidN(0,h)$. Let $\omega_{1h} - 1 = \theta h$, and $\omega_{2h} = \omega_2, \forall h$, we have

$$Y_{kh} - Y_{(k-1)h} = \sigma_{(k-1)h}\epsilon_{kh} \tag{26}$$

$$\sigma_{kh}^2 = \omega_{0h} + (\omega_{1h} - 1)\sigma_{(k-1)h}^2 + \omega_{2h}\sigma_{(k-1)h}^2\epsilon_{kh}^2 \tag{27}$$

We now observe that (26)-(27) is the same as (3)-(4), the continuous approximation scheme of the discrete GARCH(1,1) processes. The same outcome holds even if we allow for a drift component in the return equation. Suppose

$$dY_t = (c + \gamma\sigma_t^2)dt + \sigma_t dW_t$$

Now

$$(dY_t)^2 = (c + \gamma\sigma_t^2)^2(dt)^2 + \sigma_t^2\,(dW_t)^2 + 2\sigma_t(c + \gamma\sigma_t^2)dtdW_t = \sigma_t^2 dt,$$

in fact, again from the conventional multiplication table of stochastic calculus (Karatzas and Shreve, p.153), $dW_t dt = (dt)^2 = 0$. The same outcome will follow.

It should be note that an Euler approximation to (12)-(13) would also give

$$Y_{kh} - Y_{(k-1)h} = \sigma_{(k-1)h}^2\epsilon_{kh}$$

$$\sigma_{kh}^2 - \sigma_{(k-1)h}^2 = \omega_{0h} + (\omega_2 + \theta)h\sigma_{(k-1)h}^2$$

Thus both the GARCH(1,1) process and a process characterized by F_0-measurable (deterministic given σ_0^2) volatility can be seen as possible Euler approximations to the degenerate diffusion limit obtained in Proposition 2.1

We now consider the non-degenerate diffusion limit in (22)-(23). Drost and Werker (1996, p.34-35) have shown that (22)(23) is a GARCH-diffusion, in the sense that the implied discrete differences $(Y_{kh} - Y_{(k-1)h})$, for any fixed h, are weak GARCH with parameter $\xi_h = (\psi_h, \alpha_h, \beta_h, c_h)$. For weak GARCH $\forall h$, they mean that (see Definition 2.1 in Drost and Werker 1996), there exists a covariance stationary process,

$$\sigma_{kh}^2 = \psi_h + \alpha_h(Y_{kh} - Y_{(k-1)h})^2 + \beta_h\sigma_{(k-1)h}^2,$$

such that, $\forall k$ and for any given time interval h, $\sigma^2_{(k-1)h}$ is the best linear predictor of $(Y_{kh} - Y_{(k-1)h})^2$. The fourth parameter c_h denotes the kurtosis of the implied discrete differences and is relevant only for the case of jump-diffusions. Meddahi and Renault show that the discrete time volatility process implied by a, slightly different, non-degenerate diffusion belongs to the class of $SR-SARV(1)$ processes. Indeed both weak GARCH and $SR-SARV(1)$ processes display stochastic (non predictable) volatility. In fact it is immediate to see that Euler approximations to the diffusion in (22)-(23) are not GARCH processes, but stochastic volatility processes, such as

$$Y_{kh} - Y_{(k-1)h} = \sigma_{(k-1)h}\epsilon_{1,kh}$$

$$\sigma^2_{kh} - \sigma^2_{(k-1)h} = \omega_{0h} + \theta h \sigma^2_{(k-1)h} + \omega_2 \sigma^2_{(k-1)h}\epsilon_{2,kh}$$

where $\epsilon_{kh} = (\epsilon_{1,kh}, \epsilon_{2,kh})'$ and $\epsilon_{kh} \sim N(0, Ih)$.

4 Modified semi-strong GARCH processes and their continuous approximation

Regardless the different parameterizations we chose in Section 2, we have considered the continuous time approximation of the couple (Y_k, σ^2_k), where σ^2_k is the volatility of Y_{k+1}, (see eqs. (1)-(2)). In fact if instead of starting from the discrete model

$$Y_k - Y_{k-1} = \sigma_{k-1}\epsilon_k$$

$$\sigma^2_k = \omega_0 + \omega_1\sigma^2_{k-1} + \omega_2\sigma^2_{k-1}\epsilon^2_k$$

we would have started from the following discrete model

$$Y_k - Y_{k-1} = \sigma_k\epsilon_k$$

$$\sigma^2_k = \omega_0 + \omega_1\sigma^2_{k-1} + \omega_2\sigma^2_{k-1}\epsilon^2_{k-1} \tag{28}$$

we would have not reached the same diffusion limits, regardless of the chosen approximation scheme, that we obtained in Section 2. As an example we can take a continuous approximation of (28)

$$\sigma^2_{kh} - \sigma^2_{(k-1)h} = \omega_{0h} + (\omega_{1h} - 1)\sigma^2_{(k-1)h} + \omega_{2h}\sigma^2_{(k-1)h}\epsilon^2_{(k-1)h}$$

We now consider the first conditional moment

$$h^{-1}E((\sigma^2_{kh}-\sigma^2_{(k-1)h})|F_{(k-1)h}) = h^{-1}\omega_{0h} + h^{-1}(\omega_{1h}-1)\sigma^2_{(k-1)h} + h^{-1}\omega_{2h}\sigma^2_{(k-1)h}\epsilon^2_{(k-1)h},$$

as $\epsilon^2_{(k-1)h}$ is $F_{(k-1)h}$-measurable. Thus the presence of $\epsilon^2_{(k-1)h}$ in the first two conditional moments, would not have led to the same continuous time limit. Basically the main difference is that $h^{-1}E(\sigma^2_{(k-1)h}\epsilon^2_{kh}|F_{(k-1)h}) = h^{-1}\sigma^2_{(k-1)h}E(\epsilon^2_{kh}) = \sigma^2_{(k-1)h}$, while instead $h^{-1}E(\sigma^2_{(k-1)h}\epsilon^2_{(k-1)h}|F_{(k-1)h}) = h^{-1}\sigma^2_{(k-1)h}\epsilon^2_{(k-1)h}$.

As the main peculiarity of GARCH type models in discrete time is the predictability or F_{k-1}-measurability of the volatility process, both the continuous time limits we obtained in Section 2 are not completely satisfactory. The non-degenerate diffusion limit is unsatisfactory because in the continuous time limit volatility is a processes of unbounded variation, nowhere differentiable, and so unpredictable. On the other hand, in the degenerate diffusion limit, volatility is a F_0-measurable process and so, for given initial conditions, deterministic.

We now introduce a class of models that we shall call *modified semi-strong GARCH* , for their analogy to the semi-strong GARCH of Drost and Nijman (1993), and we shall derive their continuous time limit as well as their diffusion approximation "capability". In particular we shall derive the continuous time limit for the couple (Y_k, σ^2_k), where σ^2_k is the conditional variance of Y_k, and so we avoid the problem of considering the continuous time limit of the volatility process one period ahead. Second we shall see that, in the diffusion limit, volatility is a stochastic process with smooth, differentiable paths, thus capturing the idea of predictability even over very short time horizons.

DEFINITION 4.1: *Modified Semi-Strong GARCH*
We say that Y_k, $k = 1, 2, \ldots$ is a modified semi-strong GARCH processes if it satisfies the following condition:

$$E((Y_k - E(Y_k|F_{k-1}))^2|F_{k-1}) = \sigma^2_k$$

where $F_{k-1} = \sigma(\sigma^2_0, Y_0, Y_1, \ldots Y_{k-1})$ and

$$\sigma^2_k = \omega_0 + \sum_{i=1}^{p}\omega_{1i}\sigma^2_{k-i} + \sum_{j=1}^{q}\omega_{2j}Y^2_{k-j}$$

The reason for the term "modified semi-strong GARCH" is now immediate: in the case $E(Y_k|F_{k-1}) = 0$ and $Y_k = \sigma_k \epsilon_k$, we would collapse into the class of semi-strong GARCH, as defined in Drost and Nijman (1993).

We now consider the following, first order Markovian, modified semi-strong GARCH process

$$Y_k = \gamma Y_{k-1} + \sigma_k \epsilon_k$$
$$\sigma_k^2 = \omega_0 + \omega_1 \sigma_{k-1}^2 + \omega_2 Y_{k-1}^2$$

with $\alpha \in (0, 1]$, furthermore we assume $\epsilon_k \sim iidN(0, 1)$; we consider the following approximation schemes

$$Y_{kh} - Y_{(k-1)h} = (\gamma_h - 1)Y_{(k-1)h} + \sigma_{kh}\epsilon_{kh}$$
$$\sigma_{kh}^2 - \sigma_{(k-1)h}^2 = \omega_{0h} + (\omega_{1h} - 1)\sigma_{(k-1)h}^2 + \omega_{2h} Y_{(k-1)h}^2$$

If we fix $\gamma_h = \gamma = 1$ we can think at Y_k as at cumulative returns on some financial asset; for $0 < \gamma < 1$ we can think at Y_k as at some generic financial or macroeconomic variable. We shall rely on the following parameterization, as a function of h,

(29) $$\lim_{h \to 0} h^{-1}(\gamma_h - 1) = \delta, \delta < 0$$

(30) $$\omega_{0h} = h\omega_0$$

(31) $$\lim_{h \to 0} h^{-1}(\omega_{1h} - 1) = \theta < 0$$

(32) $$\omega_{2h} = h\omega_2$$

We now consider the first two conditional moments, let $F_{(k-1)h} = \sigma(Y_{(k-1)h}, \sigma_{(k-1)h}^2)$,

$$h^{-1}E((Y_{kh} - Y_{(k-1)h})|F_{(k-1)h}) = h^{-1}(\gamma_h - 1)Y_{(k-1)h}$$
$$h^{-1}E((Y_{kh} - Y_{(k-1)h})^2|F_{(k-1)h}) = h^{-1}(\gamma_h - 1)^2 Y_{(k-1)h}^2 + \sigma_{kh}^2$$
$$= \omega_{0h} + \omega_{1h}\sigma_{(k-1)h}^2 + \omega_{2h} Y_{(k-1)h} + O_p(h) = \sigma_{(k-1)h}^2 + O_p(h)$$

$$h^{-1}E((\sigma_{kh}^2 - \sigma_{(k-1)h}^2)|F_{(k-1)h}) = h^{-1}\omega_{0h} + h^{-1}(\omega_{1h} - 1)\sigma_{(k-1)h}^2 + h^{-1}\omega_{2h}Y_{(k-1)h}^2$$

$$h^{-1}E((\sigma_{kh}^2 - \sigma_{(k-1)h}^2)^2|F_{(k-1)h}) = h^{-1}\omega_{0h}^2 + h^{-1}(\omega_{1h} - 1)^2\sigma_{(k-1)h}^4 + h^{-1}\omega_{2h}^2Y_{(k-1)h}^4 + 2h^{-1}\omega_{0h}(\omega_{1h} - 1)\sigma_{(k-1)h}^2 + 2h^{-1}\omega_{0h}\omega_{2h}Y_{(k-1)h}^2 + 2h^{-1}(\omega_{2h} - 1)\omega_{2h}\sigma_{(k-1)h}^2Y_{(k-1)h}^2 = O_p(h)$$

Note that $O_p(h)$ terms above "become" $O(h)$ terms when we condition on a given value of $Y_{(k-1)h}$ and $\sigma_{(k-1)h}$. Thus we note that the conditional second moment of $(\sigma_{kh}^2 - \sigma_{(k-1)h}^2)$ vanishes at $h \to 0$.

If we would replace (32) with (18), then the second conditional moment would not vanish as $h \to 0$, but the first conditional moment would diverge at rate $h^{-1/2}$.

Let $Y_t^h = Y_{kh}$, $\sigma_t^{2,h} = \sigma_{kh}^2$, $kh \leq t < (k+1)h$, we have

PROPOSITION 4.2

If as $h \to 0$, $(Y_0^h, \sigma_0^{2,h}) \to (Y_0, \sigma_0^2)$, or $P_0^h \to P_0$, then under the parameterization in (29)-(32), as $h \to 0$, $(Y_t^h, \sigma_t^{2,h}) \Rightarrow (Y_t, \sigma_t^2)$, where (Y_t, σ_t^2) is a diffusion process solution to

$$dY_t = \delta Y_t dt + \sigma_t dW_t \tag{33}$$

$$d\sigma_t^2 = (\omega_0 + \theta\sigma_t^2 + \omega_2 Y_t^2)dt \tag{34}$$

□

Thus we still obtain a degenerate diffusion limit, in the sense of two state variables driven by only one Brownian motion. However, because of Y_t^2 on the right hand side of (34), σ_t^2 is no longer F_0-measurable, but instead it is a F_t-measurable process with paths of bounded variation and so differentiable. As we shall see in Section 6, the support properties of the diffusion in (33)(34) are dramatically different from the support properties of the degenerate diffusion in (12)-(13).

Finally it is easy to see that there is an Euler approximation to the diffusion in (33)-(34)

that is the discrete approximation scheme we began from. In fact we have

$$\begin{aligned} Y_{kh} - Y_{(k-1)h} &= \delta h Y_{(k-1)h} + \sigma_{(k-1)h}\epsilon_{kh} \\ &= \delta h Y_{(k-1)h} + \sigma_{kh}\epsilon_{kh} + O_p(\sqrt{h}) \\ &= (\gamma_h - 1) Y_{(k-1)h} + \sigma_{kh}\epsilon_{kh} + O_p(\sqrt{h}) \\ \sigma^2_{kh} - \sigma^2_{(k-1)h} &= \omega_0 h + \theta h \sigma^2_{(k-1)h} + \omega_2 h Y^2_{(k-1)h} \end{aligned}$$

The desired result just follows setting $\theta h = (\omega_{1h} - 1)$.

5 Support, Invariant Measures and Ergodic Properties of Degenerate Diffusions

Our last objective is to analyze and compare the support of the transition functions associated with the different diffusion limits we obtained. In this section we give some definitions and facts that will be used in the sequel. For the moment with the term non-degenerate diffusions we mean a diffusion having a strictly positive definite covariance matrix, and with the term degenerate diffusion we mean a diffusion having a singular covariance matrix; we shall become more precise below. Hereafter we assume that the state space is S^m, an open subset of R^m; nevertheless all the facts and definitions given below hold also when the state space is a smooth manifold of dimension m; simply we want to keep things as simple as possible.

Let P_x be the probability law of the diffusion $X = (x_t, t \geq 0)$, starting at x at time 0 and evolving on a probability space (Ω, F, P), and let $P_t(x,.)$ denote the associated transition function; it is known (e.g. Kunita 1976) that, in the non-degenerate case, the support of P_x and of $P_t(x,.)$, $\forall t > 0, \forall x \in S^m$, is the entire state space S^m. The situation is instead much more complex in the degenerate case, as, intuitively, not all points in the state space can be reached at any time, with positive probability, from arbitrary initial conditions. Thus the first task is to define the support of P_x and of $P_t(x,.)$ in the degenerate case. Basically we shall appeal to the support theorem of Strook and Varadhan (1972), stating that the support of the underlying diffusion coincides with the closure of all the paths of an associated deterministic control system, varying the control functions. The problem thus collapses to that of finding the sets of points reachable (accessible) of the associated deterministic control system. Before stating the relevant support theorems, we need to introduce some background material.

Given an Ito diffusion solution to the following stochastic differential equation,

$$dx_t = f(x_t)dt + g(x_t)dW_t \tag{35}$$

where f is $m \times 1$ and g is $m \times r$, $r \leq m$, and $W = (W_{it}, i = 1, \ldots r, t \geq 0)$ is a r-dimensional standard Brownian motion, we can always rewrite (35) in the Fisk-Stratonovich format, (e.g. Clark, 1973, p.141)

$$dx_t = a(x_t)dt + g(x_t) \circ dW_t, \tag{36}$$

where $a^i = f^i - \frac{1}{2}\sum_{j=1}^{r}\sum_{k=1}^{m} g^{jk}\frac{\partial g^{ik}}{\partial x_j}, i = 1, \ldots m$, $\circ$ denotes the Fisk-Stratonovich integral, e.g. $\int_0^t Y_s \circ dX_s = \int_0^t Y_s dX_s - [Y, X]_t$, with $[Y, X]_t$ denoting the cross variation process, defined as the L^2-limit of $\sum_{k=1}^{[2^n t]}(X_{t_k} - X_{t_{k-1}})(Y_{t_k} - Y_{t_{k-1}})$, $t_k - t_{k-1} = 2^{-n}$.

Now let,

$$W_t^n = W_{k/n} + n(t - \frac{k}{n})(W_{(k+1)/n} - W_{k/n}), k/n \leq t < (k+1)/n$$

and consider the process x_t^n defined as

$$\frac{dx_t^n}{dt} = a(x_t^n) + g(x_t^n)\frac{dW_t^n}{dt} \tag{37}$$

with a and g defined as in (36).

It is known, e.g. Kunita (1976, p.174), that, as $n \to \infty$, x_t^n converges almost surely to x_t, uniformly over finite sets in $[0, \infty)$. Now, following Kunita (1976), we can replace $\frac{dW_t^n}{dt}$ with piecewise constant functions $u_t = (u_t^1, \ldots u_t^r)$ and obtain the following system of ordinary differential equations,

$$\frac{d\phi_t}{dt} = a(\phi_t) + g(\phi_t)u_t \tag{38}$$

with a and g defined as in (36); the piecewise constant functions u_t are called controls. By varying the controls we obtain different paths for ϕ_t. We shall see below that the trajectories of the deterministic system (38) determines the support of the diffusion in (36), and so in (35), and the support of the associated transition function. Below for any set A, with clA we denote the closure of A and with $intA$ we denote the interior of A.

DEFINITION 5.1- *Invariant Control Sets* (from Arnold and Kliemann, 1987, def.1.1)

A set $C \in S^m$, $C \neq \emptyset$ is said to be an invariant control set for (38), if

$$clO^+(x) = clC$$

$\forall x \in C$, with C being maximal with respect to the inclusion, and $O^+(x)$ is the set of points reachable from x forward in time,

$$O^+(x) = \bigcup_{t>0} O^+(t,x)$$

where

$$O^+(t,x) = \{y : \exists u : R^+ \to R^r, s.t. y = \phi(t,u,x)\},$$

by varying the admissible controls u. □

Thus $O^+(t,x)$ is the set of points reachable at time t, starting from x at time 0; and $O^+(x)$ is the set of points reachable forward in time, starting from x at time 0. Heuristically any invariant control set is characterized by the fact that, if we start at time 0 from $x \in C$, then all points in C can be reached forward in time, by properly varying the controls.

THEOREM 5.2 - *Support Theorem* (from Arnold and Kliemann, 1987, p.44 and from Kunita 1976, p.164)

The original formulation of the support theorem is given in Strook and Varadhan (1972); however here we state the formulation of the theorem stated in the sources cited above.

$$(i) suppP_x = cl\{\phi(.,u,x)\}$$

where u is an admissible control and P_x is the probability law of the diffusion $(x_t, t \geq 0)$ in (35)-(36), starting from x at time zero; $\forall x \in S^m$,

$$(ii) suppP_t(x,.) = clO^+(t,x)$$

and

$$(iii) suppU^\alpha(x,.) = clO^+(x),$$

where $U^\alpha(x,.) = \int_0^\infty e^{-\alpha t} P_t(x,.)dt$, $\alpha > 0$, with $P_t(x,.)$ being the transition function of the underlying diffusion. □

Thus the support of P_x, the probability law of $(x_t, t \geq 0)$, is the closure of all the paths of the determistic system in (38), obtained by varying the controls; the support of the transition function at time t, instead, is given by the closure of all the points reachable at time t starting from x at time 0; finally the support of $U^\alpha(x, .)$ (the α-potential or Green measure of the diffusion), is given by the set of all points, reachable forward in time, by the tractories of the associated deterministic system. Intuitively we are now interested in checking whether $O^+(t, x)$ and $O^+(x)$ are non empty and in particular whether they are thick, i.e. with non empty interior.

It is known (e.g. Kunita 1976) that m-dimensional diffusions having a covariance matrix of rank m are characterized by the fact that $intO^+(t, x) = O^+(t, x), \forall t > 0$, and $intO^+(x) = O^+(x)$, furthermore $\forall t > 0, O^+(t, x) = O^+(x) = S^m$, where S^m is the state space. We shall see below that in the degenerate case, the conditions $intO^+(t, x) = O^+(t, x), \forall t > 0$, and $O^+(t, x) = S^m$ do no longer hold and we shall find conditions assuring that instead $intO^+(x) \neq \emptyset$. Before characterizing the support of the transition function and of the probability law of degenerate diffusions, we need some further terminology, and in particular we need to introduce the notions of Lie algebra and of dimension of a Lie algebra.

Let $c = gg\prime$, where g is defined as in (35); we define

$$L = \sum_{i=1}^{m} a_i \frac{\partial}{dx_i} + \frac{1}{2} \sum_{j=1}^{r} \left(\sum_{i=1}^{m} c^{ij} \frac{\partial^2}{dx_i dx_j} \right) = X_0 + \frac{1}{2} \sum_{j=1}^{r} X_j^2$$

where $X_0, X_j, j = 1, 2, \ldots r$ are first order differential operator with C^∞ coefficients and so are vector fields.

DEFINITION 5.3 - *Lie Brackets* (from Hermann, 1968, p.35).
Given two vector fields, say $X_0 = \sum_{i=1}^{m} a_i \frac{\partial}{\partial x_i}$ *and* $X_1 = \sum_{i=1}^{m} g^{i1} \frac{\partial}{\partial x_i}$, *we define the Lie (or Jacobi) bracket* $[X_0, X_1]$ *as the* $m \times 1$ *vector whose* i*-th element is given by*

$$\left(\sum_{j=1}^{m} \frac{\partial g^{i1}}{\partial x_j} a_j - \sum_{j=1}^{m} \frac{\partial a_i}{\partial x_j} g^{j1} \right) \frac{\partial}{\partial x_i}$$

□

DEFINITION 5.4- *Lie Algebra* (from Hermann, p.35, 1968)
A Lie algebra is a vector space with multiplication $(X_0, X_i) \rightarrow [X_0, X_i]$, *defined* $\forall i$, *satisfying*

the following properties

$$(i)[X_0, X_i] = -[X_i, X_0], \forall i$$

$$(ii)[X_0, [X_i, X_j]] = [[X_0, X_i], X_j] + [X_i, [X_0, X_j]]$$

for given constant $c_i, i = 1, 2, 3, 4,$

$$(iii)[c_1X_1 + c_2X_2, c_3X_3 + c_4X_4] = c_1c_3[X_1, X_3] + c_2c_3[X_2, X_3] + c_1c_4[X_1, X_4] + c_2c_4[X_2, X_4]$$

□

For example in the case of two vector fields, say X_0 and X_1, the Lie algebra generated by X_0 and X_1, $LA(X_0, X_1)$ is given by finite sums of terms like

$$X_1, [X_0, X_1], [X_1, [X_0, X_1]], [X_1, [[X_1, X_0], X_1]] \ldots$$

DEFINITION 5.5- *Dimension of a Lie Algebra*

The dimension of $LA(X_0, X_1, \ldots X_r)$ *is given by the rank of the matrix whose columns are the elements of the Lie algebra. For example in the two-vector fields case, the dimension of* $LA(X_0, X_1)$ *is given by (see Brockett, 1976, p.65),*

$$rank\,(X_1, [X_0, X_1], [X_1, [X_0, X_1]], [X_1, [[X_0, X_1], X_1]] \ldots)$$

□

Basically we keep adding terms, i.e. additional Lie brackets terms, until we are able to evaluate the rank.

Note that the rank may depend on the evaluation point x; for example let $X_0(x) = \sum_{i=1}^{m} a_i(x)$, then the dimension of $LA(X_0, X_1, \ldots X_r)(x)$ may vary with x; so that we may have degeneracies (singularities) at some particular x.

DEFINITION 5.6- *Non-degenerate, Degenerate Diffusions, Hypoellipticity* (from Arnold and Kliemann, 1987, p.44-45)

Let $X_0, X_1, \ldots X_r$ *be the vector fields characterizing the operator* L*, i.e. the generator of the diffusion.*

(i) If $dimLA(X_1, \ldots X_r)(x) = m, \forall x \in S^m$*, then the diffusion is said to be non-degenerate. Note*

that diffusions with non-singular covariance matrix, at any point in the state space, are always non-degenerate.

(ii) If $dimLA(X_1, \ldots X_r)(x) < m$*, for some* $x \in S^m$*, then the diffusion is said to be degenerate.*

(iii) If $dimLA(X_0, X_1, \ldots X_r)(x) = m, \forall x \in S^m$*, then the generator of the diffusion L is said to be hypoelliptic.* □

FACT 5.7 (Kliemann, 1987, p.691)

If $dimLA(X_0, X_1, \ldots X_r)(x) = m, \forall x \in S^m$*, and so the generator L is an hypoelliptic operator, then*

$$intO^+(x) \neq \emptyset, \forall x \in S^m$$

□

Note that instead $O^+(t, x)$ may have empty interior $\forall t$.

FACT 5.8 (from Kliemann, 1987, lemma 2.1)

Let C be an invariant control set; if $dimLA(X_0, X_1 \ldots X_r)(x) = m, \forall x \in S^m$*, then (i) C is closed in* S^m*, (ii)* $intC \neq \emptyset$*, (iii)* $O^+(x) = intC, \forall x \in intC$*.* □

It should be noted from the two facts above, that the main difference between non-degeneracy and hypoellipticity is that in the former case $intO^+(t, x) \neq \emptyset, \forall x \in S^m, \forall t > 0$, while in the latter case $intO^+(x) = int\left(\bigcup_{t>0} O^+(t, x)\right) \neq \emptyset$. Recalling the support theorem, this implies that, in the hypoelliptic case, the transition function, at any given time, can be supported on some subset of the state space which has empty interior; that is it may be supported on some lower dimensional subset of the state space, as for example a line or a curve in R^2. An open problem is that, unless the state space is compact, hypoellipticity does not guarantee the existence of at least one invariant control set (see Kliemann, 1987, lemma 2.2). However if an invariant control set exists, then hypoellipticity implies several nice properties.

Let μ be an invariant probability measure for the diffusion $(x_t; t \geq 0)$, that is

$$\mu(.) = \int P_t(x, .)\mu(dx), \forall t \geq 0$$

where $P_t(x, .)$ is the underlying transition function. Hereafter we state two facts from Arnold and Kliemann 1987; it should be pointed out that Assumption A in their paper is implied by hypoellipticity.

FACT 5.9 - *Uniqueness of Invariant Measures* (from Arnold and Kliemann, 1987,cor.2.1)
If $dimLA(X_0, X_1, \ldots X_r)(x) = m, \forall x \in S^m$, *then invariant probability measures are unique over invariant control sets. It is also true that*

$$P_x(x_t \in C, \forall t \geq 0) = 1, \forall x \in C$$

□

As for the ergodic properties, we have
FACT 5.10 (from Arnold and Kliemann, 1987, theorem 6.1)
If $dimLA(X_0, X_1, \ldots X_r)(x) = m, \forall x \in S^m$, *then* $\forall x \in intC$, *and for any* $f \in L^1(\mu)$,

$$P_x\left(\lim_{T\to\infty} \frac{1}{T} \int_0^T f(x_t)dt = \mu(f)\right) = 1$$

□

6 The Support of the GARCH Degenerate Diffusion Limit

As mentioned in the previous section, we know that both the support of the probability measure underlying the diffusion process and the support of the associated transition function, in the non-degenerate case, is the entire state space $R \times R^+$. We also know, from Nelson's paper, that, as the time span goes to infinity, the volatility process is distribuited according to an inverted gamma distribution.

As for the diffusion limit in (12)-(13), it is easy to see that, for given initial condition, the solution σ_t^2 to (13) is given by,

$$\sigma_t^2 = e^{\omega_2+\theta}\sigma_0^2 - \omega_0(\omega_2+\theta)^{-1}$$

with $\theta < 0, \omega_0, \omega_2 > 0$. We observe that σ_t^2 has a well defined deterministic limit, as $t \to \infty$, for $\theta + \omega_2 < 0$. If $\theta = (\omega_1 - 1)$, then such a condition is the usual condition, $\omega_1 + \omega_2 < 1$. We also note that σ_t^2 is F_0-measurable. Thus we can write

$$Y_t = \int_0^t \sigma_t^2 dW_s$$

so that Y_t is a Gaussian process, with time varying, but deterministic (in the case of σ_0^2 deterministic) variance. Let $x = (Y, \sigma^2)$ and let P_x denote the probability law of the diffusion starting from x at time 0, also let $P_t(x, .)$ be the transition function, starting from x at time 0; now it is easy to see that: (i) the support of P_x is a line in $S^2 = R \times R^+$, (ii) the support of $P_t(x, .)$, $\forall t > 0$ is still a line in S^2.

Thus the only non trivial case to analyze, is the degenerate diffusion in (33)-(34). Let's consider again

$$dY_t = \delta Y_t dt + \sigma_t dW_t \tag{39}$$

$$d\sigma_t^2 = (\omega_0 + \theta\sigma_t^2 + \omega_2 Y_t^2)dt, \tag{40}$$

with $\theta < 0$ and $\delta < 0$. The Fisk-Stratonovich representation of the SDE in (34)-(35) is simply

$$dY_t = \delta Y_t dt + \sigma_t \circ dW_t \tag{41}$$

$$d\sigma_t^2 = (\omega_0 + \theta\sigma_t^2 + \omega_2 Y_t^2)dt \tag{42}$$

Following the procedure described in the previous section, we know that the support of (41)(42) can be studied via the analysis of the trajectories of the following associated deterministic control system

$$\frac{d\phi_{1t}}{dt} = \delta\phi_{1t} + \sqrt{\phi_{2t}}u_t$$

$$\frac{d\phi_{2t}}{dt} = (\omega_0 + \theta\phi_{2t} + \omega_2\phi_{1t}^2)$$

The generator of the diffusion in (41)(42) is given by

$$L_x = \delta y\frac{\partial}{\partial y} + (\omega_0 + \theta\sigma^2 + \omega_2 y^2)\frac{\partial}{\partial\sigma^2} + \frac{1}{2}\sigma^2\frac{\partial^2}{\partial\sigma^2\partial\sigma^2} = X_0 + \frac{1}{2}X_1^2$$

where $x = (y, \sigma^2)$.

Note that in our case the state space is given by $S^2 = R^+ \times R$, which is an open subset of R^2.

PROPOSITION 6.1

If $\omega_2 > 0, \sigma^2 > 0$, *then*

$$(i) dim LA(X_1)(x) = 1, \forall x \in S^2$$

and

$$(ii) dimLA(X_0, X_1)(x) = 2, \forall x \in S^2$$

where $x = (y, \sigma^2)\prime$ *is the starting value at time* 0. □

PROOF:

(i)

$$dimLA(X_1) = dimLA(X_1, 0) = rank \begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix}$$

(ii)

It follows from Definition 5.3, recalling that here $g^{11} = \sigma$, $g^{12} = g^{21} = g^{22} = 0$, $a_1 = \delta y$, $a_2 = (\omega_0 + \theta\sigma^2 + \omega_2 y^2)$, $x = (x_1, x_2)\prime = (y, \sigma^2)\prime$; in fact

$$[X_0, X_1] = \begin{pmatrix} -(\sigma\delta - \frac{\omega_0+\theta\sigma^2+\omega_2 y^2}{2\sigma})\frac{\partial}{\partial y} \\ 2\omega_2 y\sigma\frac{\partial}{\partial\sigma^2} \end{pmatrix}$$

Set $a = -\sigma\delta + \frac{\omega_0+\theta\sigma^2+\omega_2 y^2}{2\sigma}$, now

$$[X_1, [X_0, X_1]] = \begin{pmatrix} -(\frac{a}{2\sigma} - \sigma\frac{\partial a}{\partial y})\frac{\partial}{\partial y} \\ -(2\omega_2\sigma^2)\frac{\partial}{\partial\sigma^2} \end{pmatrix}$$

$$dimLA(X_0, X_1) = rank\,(X_1, [X_0, X_1], [X_1, [X_1, X_0]])$$

$$= rank \begin{pmatrix} \sigma & a & -(\frac{a}{2\sigma} + \sigma\frac{\partial a}{\partial y}) \\ 0 & 2\omega_2 y\sigma & -2\omega_2\sigma^2 \end{pmatrix} = 2, \omega_2 > 0, \sigma > 0, \delta < 0$$

COROLLARY 6.2

$$intO^+(x) \neq \emptyset, \forall x \in S^2$$

where $x = (y, \sigma^2)$, □.

PROOF: Immediate from Fact 5.7

From the statements above we know that, starting at any point x, at time zero, the set of points reachable forward in time is a non empty set in S^2. On the other hand we know, from the support theorem, that $suppP_t(x,.) = clO^+(t,x)$, $\forall t > 0, \forall x \in S^2$, but hypollepiticity does not guarantee that $intO^+(t,x) \neq \emptyset, \forall t > 0$, thus $\forall t > 0$, the transition function may be supported on some subset of the state space having empty interior; for example it can be supported on a straight line or on a curve in S^2. This is consistent with the fact that in discrete time the probability measure of the couple (Y_k, σ_k^2), given Y_{k-1}, σ_{k-1}^2, is such that

$$suppP(Y_k, \sigma_k^2 | Y_{k-1}, \sigma_{k-1}^2) = suppP(Y_k, \sigma_k^2 | \sigma_0^2, Y_0, Y_1, \dots Y_{k-1})$$

which is given by a line across the first and second (or first and fourth) quadrants. In fact σ_k^2 is F_{k-1}-measurable, with $F_{k-1} = \sigma(\sigma_0^2, Y_0, Y_1, \dots Y_{k-1})$, so given the information available at time $k-1$, the support of (Y_k, σ_k^2) is a line in the state space and so has zero Lebesgue measure in S^2. On the other hand $suppP(Y_k, \sigma_k^2 | Y_0, \sigma_0^2)$ generally is not a straight in line in S^2, but can be a generic subset in S^2 with possible empty interior, for example a curve in S^2. However, from the support theorem we also know that $suppU^\alpha(x, .) = O^+(x), \alpha > 0$ and that $intO^+(x) \neq \emptyset$, $\forall x \in S^2$; thus from any arbitrary point we start at time zero, forward in time, we can reach, with positive probability, some subset of the state space with non-empty interior; this implies, by Theorem 5.2(iii), that the Green measure, or α-potential of the diffusion $\int_0^\infty e^{-\alpha t} P_t(x, .)$ is supported on a two-dimensional subset of S^2 with non-empty interior.

As the state space is non compact, the sufficient conditions for the existence of at least one invariant control set are not satisfied (see Kliemann 1987, lemma 2.2). In any case, even if there would be an invariant control set (or more invariant control sets) and a unique invariant measure on them, from Fact 5.10, we just could say that

$$P_x\{\lim_{T\to\infty} \int_0^T f(x_t)dt = \mu(f)\} = 1$$

$\forall f \in L^1(\mu), \forall x \in C$, with μ denoting the (unique) invariant measure on C.

Thus the only ergodic result we can hope for is a law of large numbers over invariant control sets. This is a much weaker result than that obtained by Nelson (1990) in the non-degenerate case. In fact, in his framework, the volatility process, in the continuous limit, follows a constant elasticity of variance diffusion; as the time span approaches infinity the corresponding transition density approaches, in the total variation norm, an inverted gamma density.

7 Appendix

PROOF OF PROPOSITION 2.1

The conditions for the existence of a unique strong solution (e.g. conditions 10.6 and 10.7 in Chung and Williams, 1990) are satisfied. This implies that there is a unique weak solution. We note that $\forall y, \sigma^2$

$$h^{-1}E((Y_{kh} - Y_{(k-1)h})^4 | Y_{(k-1)h} = y, \sigma^2_{(k-1)h} = \sigma^2) = o(h)$$

and

$$h^{-1}E((\sigma^2_{kh} - \sigma^2_{(k-1)h})^4 | Y_{kh} = y, \sigma^2_{(k-1)h} = \sigma^2) = o(h)$$

Thus, from Arnold (1974, p.40), we know that it will suffice to check conditions (2.4)-(2.6) in Strook and Varadhan (Ch. 11, 1979) over the entire state space. Given the first two conditional moments in (5)-(8) and the parameterization in (9)-(11), and given that the size of the jumps approaches zero as $h \to 0$, the result then follows from theorem 11.23 in Strook and Varadhan (1979).

PROOF OF PROPOSITION 2.2

By the same argument as the proof of Proposition 2.1.

PROOF OF PROPOSITION 4.2

The generator associated to the diffusion in (33)(34) is give by

$$Lf(x) = \sum_{i=1}^{2} a_i(x) \frac{\partial f}{\partial x_i} + \frac{1}{2} \sum_{i=1}^{2} \sum_{j=1}^{2} c^{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} f$$

for all $f \in C^2$,$x = (y, \sigma^2)\prime$, where C^2 denotes the class of twice continuous differentiable functions, and $a_1 = \delta y$, $\delta < 0$, $a_2 = (\omega_0 + \theta\sigma^2 + \omega_2 y^2)$, $c = gg\prime$ and $g^{11} = \sqrt{\sigma^2} = \sigma$, $g^{ij} = 0 \forall i, j \neq 1$. As a_i and g^{ij}, $\forall i, j$, are locally bounded, by theorem 10.21 in Strook and Varadhan, the martingale problem associated with the operator L is well posed, and so there exists a unique weak solution, if we can show that the conditions for non-explosion are satisfied. Hereafter $||x||$ denotes the Euclidean norm of x. Thus it suffices to show that for any $\phi \in C^{1,2}([0,T], R \times R^+)$, such that

$$A1 \lim_{||x|| \to \infty} \inf_{0 \leq t \leq T} \phi(t, x) = \infty,$$

the following holds

$$A2 \left(\frac{\partial}{\partial t} + L_t \right) \phi \leq \lambda \phi$$

on $[0, T] \times R \times R^+$, where $\lambda = \lambda_T > 0$. A2 assures that the diffusion does not run to infinity in finite time.

We can set $\phi(x) = y^2 + \sigma^2 + K$, with $0 < K < \infty$. Now $\|x\| \leq |y| + \sigma^2$, so when the LHS approaches infinity the RHS also does. Now whenever $|y| + \sigma^2 \to \infty$, $\phi(x) \to \infty$; thus $\phi(x)$ satisfies A1. Now

$$L\phi(x) = 2\delta y^2 + (\omega_0 + \theta\sigma^2 + \omega_2 y^2) + \sigma^2 \leq \lambda\phi(x), \forall x$$

for $0 < \lambda < \infty$, so A2 is satisfied too.

It remains to check the moment conditions (2.4)-(2.6) in Strook and Varadhan (1979, p.268); the result will then follow from theorem 11.2.3 of Strook and Varadhan. First we want to show that $h^{-1}E((y_{kh} - y_{(k-1)h})^m | y_{(k-1)h} = y, \sigma^2_{(k-1)h} = \sigma^2) = o(1)$ and $h^{-1}E((\sigma^2_{kh} - \sigma^2_{(k-1)h})^m | y_{(k-1)h} = y, \sigma^2_{(k-1)h} = \sigma^2) = o(1)$, for some $m > 2$, in fact in this case, because of Arnold (1974, p.40), it will suffice to check (2.4)-(2.6), in Ch. 11 of Strook and Varadhan, over all the state space. Now given (29)-(32) it is immediate to see that

$$h^{-1}E((y_{kh} - y_{(k-1)h})^4 | y_{(k-1)h} = y, \sigma_{(k-1)h} = \sigma^2) = o(h) = o(1)$$

and

$$h^{-1}E((\sigma_{kh} - \sigma_{(k-1)h})^4 | y_{(k-1)h} = y, \sigma_{(k-1)h} = \sigma^2) = o(h) = o(1) \quad .$$

Recalling the first two conditional moments computed below equation (32), and recalling that the jumps in the right continuous approximation are of order h, we know that the result follows from theorem 11.2.3 in Strook and Varadhan (1979).

8 References

Arnold L. (1974), *Stochastic Differential Equations: Theory and Applications* , Wiley, New York.

Arnold L. and W. Kliemann (1987), On unique ergodicity for degenerate diffusions, *Stochastics* v.21, 41-61.

Bollerslev T., R.F. Engle & D.B. Nelson (1994), ARCH models, in *Handbook of Econometrics*, v. iv, ed. by R.F. Engle and D.L. McFadden, Elsevier.

Brockett R.W. (1976), Nonlinear systems and differential geometry, *Proceedings of the IEEE*, v.64, 61-72.

Clark J.M.C. (1973), An introduction to stochastic differential equations on manifolds, in *Geometric Methods in Systems Theory*, ed. by R.W. Brockett and D.Q. Mayne.

Chung K.L. and R.J. Williams (1990), *Introduction to Stochastic Integration*, Birkhauser, Boston, Second Edition

Ding Z., C.W.J. Garanger & R.F. Engle (1993), A long memory property of stock market return and a new model, *Journal of Empirical Finance*, v.1, 83-106.

Drost F.C. & T.E. Nijman (1992), Temporal aggregation of GARCH processes, *Econometrica*, v.61, 909-927.

Drost F.C. & J.C. Werker (1996), Closing the GARCH gap: continuous GARCH modelling, *Journal of Econometrics*, v.74, 31-57.

Engle R.F. (1982), Autoregressive, conditional heteroskedasticity with estimates of the variance of U.K. inflation, *Econometrica*, 50, 987-1008.

Ethier S.N. & T.G. Kurtz (1986), *Markov Processes: characterization and convergence*, Wiley, New York.

Fornari F. & A. Mele (1996), Weak convergence and distributional assumptions for a general class of nonlinear ARCH models, *Econometric Reviews*, forthcoming.

Harrison J.M., R. Pitbladdo & S.M. Schaefer (1984), Continuous time processes in frictionless markets have infinite variation, *Journal of Business*, v.57, 353-365.

Hermann R. (1968), *Differential geometry and the calculus of variation*, Academic Press, New York.

Hobson D.G. & Rogers L.C.G. (1994), Models of Endogenous stochastic volatility, University of

Cambridge and University of Bath.

Ichihara K. & H. Kunita (1974), A classification of seconder order degenerate elliptic operators and its probabilistic characterization *Z. Wahrscheinlichkeistheorie verw. Gebiete*, v.30, 235-254.

Journal of Econometrics, (1996), v.74, issue on ARCH and long memory models.

Karatzas I. & Shreve S.E. (1991) *Brownian Motion and Stochastic Calculus*, Springer Verlag, Second Edition, New York.

Kliemann W. (1987), Recurrence and Invariant measures for degenerate diffusions, *Annals of Probability*, v.15, 690-707.

Kunita H. (1976), Support of diffusion processes and controllability problems, in *Proceeding of the International Symposium on Stochastic Differential Equations, Kyoto 1976*, ed. K. Ito, Wiley, New York.

Meddahi N. & E. Renault (1996) Aggregation and marginalization of GARCH and stochastic volatility models, Working Paper, GREMAQ and Universite' de Toulouse

Nelson D.B. (1990), ARCH models as diffusion approximations, *Journal of Econometrics*, v.45, 7-38.

Strook D.W. & S.R.S. Varadhan (1972), On the support of diffusion processes with applications to the strong maximum principle, *Proceedings of the 6-th Berkeley Symposium on Mathematical Statistics and Probability*, v.3, 333-368.

Strook D.W. & S.R.S. Varadhan (1979), *Multidimensional Diffusion Processes*, Springer and Verlag, Berlin