Technische Universität München

Zentrum Mathematik

Probabilistic Analysis of Multivariate GARCH Models

Projektarbeit

von

Florian Andreas Fuchs

Themensteller: Dr. Robert Stelzer
Abgabetermin: 11. September 2009

Hiermit erkläre ich, dass ich die Projektarbeit selbstständig angefertigt und nur die angegebenen Quellen verwendet habe.

Garching, den 11. September 2009

Contents

Introduction

The origin of time series analysis can be found within the development of ARMA (autoregressive moving average) processes. However, the linearity supposed for such processes is a constraint which gave rise to the design of non-linear models. The most simple ones are Markov models in the form

$$X_t = F(X_{t-1}, e_t) \tag{1}$$

where $(X_t)_{t\in\mathbb{N}}$ is a sequence of random variables with values in a topological state space S equipped with its Borel σ-field $\mathcal{B}(S)$, F maps $S \times \mathbb{R}^d$ into S and $(e_t)_{t\in\mathbb{N}^*}$ is an $\mathbb{R}^d$-valued i.i.d. sequence of random variables independent of X_0.

It is often helpful for modelization if one can find strictly stationary solutions to the equation (1), i.e. if one can find invariant probabilities for the Markov chain.

In general, there are two well-known concepts to deduce stationarity. The first method is to show certain Lipschitz properties for the map F. More precisely, if there is a norm $\|\cdot\|$ on S and a real function α such that, for all $(x, y) \in S^2$,

$$\|F(x, e_t) - F(y, e_t)\| \leq \alpha(e_t)\, \|x - y\| \tag{2}$$

with $\mathbb{E}[\alpha(e_t)] < 1$ and $\mathbb{E}\left[\|F(0, e_t)\|\right] < \infty$, then there is a strictly stationary solution for the model (cf. [7]).

Another approach, developed in the book of Meyn and Tweedie [12], is based upon the irreducibility of the Markov chain $(X_t)_{t\in\mathbb{N}}$ and a so-called Lyapounov function $V \geq 1$ which satisfies

$$\mathbb{E}[V(X_t)|X_{t-1} = x] \leq \alpha V(x) + b\mathbb{1}_C(x) \tag{3}$$

where C is a suitable set on which V is bounded and $\alpha < 1,\ b < \infty$ are positive constants. This function permits to control the behavior of the chain and yields the existence of a strictly stationary solution for (1), too. Since our aim in this project is to study stationarity of multivariate GARCH models, where the first approach is not applicable, we will concentrate on the second method (3) during our work.

Our proceeding will be the following:

The first chapter will recall important definitions and results which we will use in this project thesis. It consists of two parts where the first one considers the principal notions and results for Markov chains on general state spaces and establishes a Foster - Lyapounov condition which ensures stationarity and ergodicity for irreducible and aperiodic chains. Markov chains on general state spaces in discrete time can be seen as an analogy to chains on discrete state spaces where the future value of the chain only depends on the present state and transitions are specified by stochastic matrices. The "only" difference is that we cannot use these matrices any longer but have to switch to so-called transition probability kernels, an appropriate "generalization" of stochastic matrices.

The second part of the first chapter will concentrate on algebraic geometry and in particular on algebraic varieties and regular maps. The motivation to introduce these concepts is clear: since we will use the Foster - Lyapounov method described above in (3) for the study of stationarity of multivariate GARCH models and since this approach requires irreducibility which in general cannot be shown for such processes on the whole state space, we will restrict the chain to a suitable subspace on which the process is irreducible. It will turn out that this subspace is an algebraic variety. Since we are going to use the Zariski topology in addition to the standard topology, we note that if not otherwise stated we always consider the "normal" topology.

In the second chapter, in view of an application to GARCH processes, we study strict stationarity and ergodicity of semi-polynomial Markov chains which are defined by (1) whereas F is a composition of a regular map and a diffeomorphism. In particular, we will construct using the Zariski topology a subspace on which these chains are irreducible. Therefore Chapter two will use the methods of algebraic geometry introduced in Chapter one. The chapter can be considered a generalization of Mokkadem [13].

The third chapter will start with a repetition of vec and BEKK models for multivariate GARCH processes. In the main part of this chapter we are going to prove that a GARCH process in the BEKK representation satisfying some additional properties is positive Harris recurrent and geometrically ergodic on the Zariski closure W of an orbit of a well chosen point T and that the strictly stationary solution is β - mixing. To show this we will suppose that the distribution of every e_t of the i.i.d. innovation sequence $(e_t)_{t\in\mathbb{N}^*}$ is absolutely continuous with respect to the Lebesgue measure, that the point zero is in the interior of the domain of positivity of the density of the distribution of e_t and that the spectral radius of the matrix $\sum_{i=1}^{q} A_i + \sum_{j=1}^{p} B_j$ is less than 1, where the matrices A_i and B_j occur in the vech representation of a BEKK model (i.e. the conditional covariance matrix Σ_t of the GARCH process can be written as $\text{vech}(\Sigma_t) = \text{vech}(C) + \sum_{i=1}^{q} A_i \text{vech}(X_{t-i}X_{t-i}^t) + \sum_{j=1}^{p} B_j \text{vech}(\Sigma_{t-j})$). Moreover, we will obtain the uniqueness of the strictly stationary solution. On the other hand, we will also show that whenever there exists a weak stationary solution the spectral radius

of $\sum_{i=1}^{q} A_i + \sum_{j=1}^{p} B_j$ has to be less than 1.
The main idea and structure of this work can be found in Boussama [3]. However, we changed subtle details in our proceeding so that attentive readers will easily find the difference to the work of Boussama.

Chapter 1

Preliminaries

In this first chapter, we shall summarize some basic properties concerning recurrence and ergodicity of Markov chains on general state spaces and important results from algebraic geometry which we will use for the study of autoregressive processes defined by a composition of a regular map and a diffeomorphism in Chapter two and in particular for the study of multivariate GARCH models in Chapter three.

1.1 Markov Chains

We recall well-known notations and concepts for Markov chains which can be found for example in Meyn and Tweedie [12]. By $\mathbb{N}$ we will denote the set of natural numbers including zero and for the natural numbers without zero we shall use $\mathbb{N}^*$.
Let $(X_t)_{t\in\mathbb{N}}$ be a Markov chain with values in $(S, \mathcal{B}(S))$ where S is a topological state space, i.e. a set equipped with a locally compact, separable, metrizable topology, and $\mathcal{B}(S)$ is the Borel σ-field on S. By $P = \{P(x, A) : x \in S,\ A \in \mathcal{B}(S)\}$ we denote its **transition probability kernel** (or also called **Markov transition function**), i.e. P fulfills the following two properties

(i) $P(\,\cdot\,, A) : S \to \mathbb{R}_{\geq 0}$ is $(\mathcal{B}(S), \mathcal{B}(\mathbb{R}_{\geq 0}))$ - measurable for all $A \in \mathcal{B}(S)$

(ii) $P(x, \cdot)$ is a probability measure on $(S, \mathcal{B}(S))$ for all $x \in S$.

Set $S^{\mathbb{N}} = \prod_{i=0}^{\infty} S_i$ and $\mathcal{B}(S)^{\mathbb{N}} = \bigvee_{i=0}^{\infty} \mathcal{B}(S_i)$ where $S_i = S$ for all $i \in \mathbb{N}$. As shown, for example, in [12] Theorem 3.4.1, there exists for any initial measure μ on $(S, \mathcal{B}(S))$ and any transition probability kernel P a probability measure $\mathbb{P}_\mu$ on $(S^{\mathbb{N}}, \mathcal{B}(S)^{\mathbb{N}})$ such that $\mathbb{P}_\mu(B)$ is the probability of the event $\{(X_t)_{t\in\mathbb{N}} \in B\}$ for all B in $\mathcal{B}(S)^{\mathbb{N}}$ and for all $A_i \in \mathcal{B}(S)$,

$i = 0, 1, \ldots, n$ and any $n \in \mathbb{N}$

$$\mathbb{P}_\mu(A_0 \times A_1 \times \ldots \times A_n \times S \times S \times \ldots) = \int_{y_0 \in A_0} \cdots \int_{y_{n-1} \in A_{n-1}} \mu(dy_0) P(y_0, dy_1) \cdots P(y_{n-1}, A_n).$$

We shall call $\mathbb{P}_\mu$ the distribution of the Markov chain $(X_t)_{t\in\mathbb{N}}$ with initial distribution μ and use $\mathbb{P}_\mu((X_t)_{t\in\mathbb{N}} \in B)$ instead of $\mathbb{P}_\mu(B)$ as an abuse of notation. By $\mathbb{P}_x$ we denote $\mathbb{P}_{\delta_x}$, where δ_x is the Dirac measure in x for $x \in S$.
The n-step transition probability kernel is inductively defined for $x \in S,\ A \in \mathcal{B}(S)$ by

$$P^0(x, A) = \delta_x(A)$$
$$P^n(x, A) = \int_S P(x, dy) P^{n-1}(y, A), \quad n \geq 1.$$

We write P^n for the n-step transition probability kernel $\{P^n(x, A) : x \in S,\ A \in \mathcal{B}(S)\}$. The kernel P^n operates naturally on bounded measurable functions f from the left and on σ - finite measures μ on $(S, \mathcal{B}(S))$ from the right via

$$P^n f(x) = \int_S f(y) P^n(x, dy), \quad x \in S,$$
$$\mu P^n(A) = \int_S \mu(dx) P^n(x, A), \quad A \in \mathcal{B}(S),$$

and we shall use the notation $P^n f,\ \mu P^n$ to denote these operations.
If we set $\mathbb{E}_\mu$ the expectation operation corresponding to $\mathbb{P}_\mu$, we get

$$P^n f(x) = \mathbb{E}_x[f(X_n)] \quad \text{and} \quad \mu P^n(A) = \mathbb{P}_\mu(X_n \in A).$$

1.1.1 Strict Stationarity and Stationarity

We want to introduce two concepts of stationarity which we will use during this work. As in Brockwell and Davis [4] we define:

Definition 1.1.1. *Let $(X_t)_{t\in\mathbb{Z}}$ be a stochastic process with values in $(S, \mathcal{B}(S))$.*
If, for all $n \in \mathbb{N}^*$ and all $t_1, \ldots, t_n \in \mathbb{Z},\ h \in \mathbb{Z}$,

$$(X_{t_1}, \ldots, X_{t_n})^t \stackrel{d}{=} (X_{t_1+h}, \ldots, X_{t_n+h})^t$$

then the stochastic process is called ***strictly stationary****.*

Definition 1.1.2. *Let $(X_t)_{t\in\mathbb{Z}}$ be a stochastic process with values in $(S, \mathcal{B}(S))$ and $\|\cdot\|$ a norm on S. If*

(i) $\mathbb{E}[\|X_t\|^2] < \infty \quad \forall t \in \mathbb{Z}$

(ii) $\mathbb{E}[X_t] = m$ *for some $m \in S$ and all $t \in \mathbb{Z}$*

(iii) $\gamma_X(s,t) := \mathrm{Cov}(X_s, X_t) = \gamma_X(s+h, t+h) \quad \forall s,t \in \mathbb{Z}, h \in \mathbb{Z}$

*then the process is called **stationary**.*

Remark 1.1.3.

(i) Strict stationarity means that the distribution of the process is invariant under shifts in the time domain $\mathbb{Z}$.

(ii) One can show easily that, if $(X_t)_{t\in\mathbb{Z}}$ is strictly stationary and has finite second order moments, then $(X_t)_{t\in\mathbb{Z}}$ is also stationary.

(iii) Stationary stochastic processes are also often called **weakly stationary** or **covariance stationary**.

1.1.2 Invariant Measures

We consider again a Markov chain $(X_t)_{t\in\mathbb{N}}$ on $(S, \mathcal{B}(S))$ with transition probability kernel P.

Definition 1.1.4. *A σ - finite measure π on $(S, \mathcal{B}(S))$ with the property*

$$\pi(A) = \int_S \pi(dx)P(x,A) \quad \forall A \in \mathcal{B}(S) \qquad (i.e.\ \pi = \pi P)$$

*will be called **P - invariant**.*
*If there exists a P - invariant probability measure π on $(S, \mathcal{B}(S))$, we shall call the Markov chain **positive**.*

In the case where the transition probability kernel P of the Markov chain $(X_t)_{t\in\mathbb{N}}$ possesses an invariant probability measure π (i.e. if the Markov chain is positive), the chain with initial distribution π fulfills

$$(X_{t_1}, \dots, X_{t_n})^t \stackrel{d}{=} (X_{t_1+h}, \dots, X_{t_n+h})^t$$

for all $n \in \mathbb{N}^*$ and all $t_1, \dots, t_n \in \mathbb{N}$, $h \in \mathbb{N}$. We denote by $(X_t)_{t\in\mathbb{Z}}$ its extension on $\mathbb{Z}$ which is a strictly stationary process in terms of Definition 1.1.1.

1.1.3 Irreducibility, Small Sets and Aperiodic Chains

We introduce the fundamental concept of irreducibility for the study of Markov chains. Intuitively the idea is to express a class of regions in S which will be "visited" by the chain, independently of the starting point.
If we denote for $A \in \mathcal{B}(S)$ by

$$\tau_A := \min\{t \in \mathbb{N}^* : X_t \in A\} \quad \text{(where } \min(\emptyset) := \infty \text{ by convention)}$$

the **first return time** to A and by

$$L(x, A) := \mathbb{P}_x(\tau_A < \infty) = \mathbb{P}_x((X_t)_{t\in\mathbb{N}^*} \text{ ever enters } A)$$

the **return time probability** to the set A starting from the state $x \in S$, then we can give the following formal definition of irreducibility.

Definition 1.1.5. *We call $(X_t)_{t\in\mathbb{N}}$ φ* ***- irreducible*** *if there exists a non-trivial measure φ on $(S, \mathcal{B}(S))$ such that, for all $x \in S$ and all $A \in \mathcal{B}(S)$, whenever $\varphi(A) > 0$, we also have $L(x, A) > 0$.*
φ is called ***irreducibility measure****.*

Instead of $L(x, A) > 0$ one can require equivalently that there exists an $n \in \mathbb{N}^*$ such that $P^n(x, A) > 0$.
It is clear that if $(X_t)_{t\in\mathbb{N}}$ is φ - irreducible and φ' is a measure absolutely continuous with respect to φ (denoted by $\varphi' \prec \varphi$), $(X_t)_{t\in\mathbb{N}}$ is also φ' - irreducible.
One is interested in finding a "maximal" irreducibility measure which defines the range of the chain much more completely than some of the other more arbitrary irreducibility measures one may construct initially. That there exists such a "maximal" irreducibility measure is ensured by the following proposition:

Proposition 1.1.6. *If $(X_t)_{t\in\mathbb{N}}$ is φ - irreducible for some measure φ, then there exists a probability measure ψ on $(S, \mathcal{B}(S))$ such that*

(i) $(X_t)_{t\in\mathbb{N}}$ is ψ - irreducible

(ii) for all measures φ' on $(S, \mathcal{B}(S))$ the chain $(X_t)_{t\in\mathbb{N}}$ is φ' - irreducible if and only if $\psi \succ \varphi'$

(iii) the probability measure ψ is equivalent to

$$\psi'(A) := \int_S \varphi'(dy) \sum_{n=0}^{\infty} P^n(y, A) 2^{-(n+1)}$$

for any finite irreducibility measure φ'.

Proof. See for instance [12] Proposition 4.2.2. □

We will consistently use ψ to denote an arbitrary maximal irreducibility measure for $(X_t)_{t\in\mathbb{N}}$. By $\mathcal{B}^+(S) := \{A \in \mathcal{B}(S) : \psi(A) > 0\}$ we denote the sets of positive ψ - measure. Since two arbitrary maximal irreducibility measures are always equivalent, $\mathcal{B}^+(S)$ is uniquely defined.
In the following we introduce small sets which will be very useful to check whether a Markov chain is Harris recurrent, a property of Markov chains which we will consider later on.

Definition 1.1.7. *Let $C \in \mathcal{B}(S)$. If there exists an $n \in \mathbb{N}^*$ and a non-trivial measure ν_n on $(S, \mathcal{B}(S))$ such that, for all $A \in \mathcal{B}(S)$ and all $x \in C$,*

$$P^n(x, A) \geq \nu_n(A),$$

then C is said to be a ***small set****.*

One can show that there exists a countable collection C_i of small sets in $\mathcal{B}(S)$ such that $S = \cup_{i=0}^{\infty} C_i$ if $(X_t)_{t\in\mathbb{N}}$ is ψ - irreducible (see Proposition 5.2.4 (ii) [12]). Using this existence of small sets, where $C_1 \in \mathcal{B}^+(S)$ without loss of generality, one can show that ψ - irreducible chains have a finite periodic breakup into cyclic sets.

Theorem 1.1.8. *Suppose that $(X_t)_{t\in\mathbb{N}}$ is a ψ - irreducible Markov chain on $(S, \mathcal{B}(S))$. Then there exist $d \in \mathbb{N}^*$ and disjoint sets $D_1, \dots, D_d \in \mathcal{B}(S)$ such that*

(i) $\forall i = 1, \dots, d \quad \forall z \in D_i \qquad P(z, D_{(i \bmod d)+1}) = 1$

(ii) $\psi((\cup_{i=1}^{d} D_i)^c) = 0.$

Proof. See for instance [12] Theorem 5.4.4. □

The largest such d is called the **period** of $(X_t)_{t\in\mathbb{N}}$. When the period is 1, the chain is said to be **aperiodic**.

1.1.4 Petite Sets

Meyn and Tweedie [12] give a generalization of small sets, the so-called petite sets.
Let therefore $a = \{a(n)\}$ be a probability measure on $\mathbb{N}$ and consider the Markov chain $(X_t^a)_{t\in\mathbb{N}}$ with transition probability kernel

$$K_a(x, A) := \sum_{n=0}^{\infty} P^n(x, A)a(n), \quad x \in S, \ A \in \mathcal{B}(S).$$

It is obvious that K_a is indeed a transition kernel, so that the chain $(X_t^a)_{t\in\mathbb{N}}$ is well-defined (cf. [12] Theorem 3.4.1). $(X_t^a)_{t\in\mathbb{N}}$ is called K_a - chain with sampling distribution a.

Definition 1.1.9. *We will call a set $C \in \mathcal{B}(S)$ ν_a* ***- petite*** *if the sampled chain satisfies the bound*

$$K_a(x, B) \geq \nu_a(B)$$

for all $x \in C,\ B \in \mathcal{B}(S)$, where ν_a is a non-trivial measure on $(S, \mathcal{B}(S))$.

From the definitions we see that a small set is petite, with the sampling distribution a taken as δ_m for some $m \in \mathbb{N}^*$. Formally spoken, if $C \in \mathcal{B}(S)$ is ν_m - small then C is ν_{δ_m} - petite. Hence the property of being a small set is in general stronger than the property of being petite.
However there are several special cases where the two concepts coincide. One important case is the following:

Theorem 1.1.10. *If the Markov chain $(X_t)_{t\in\mathbb{N}}$ is ψ - irreducible and aperiodic then every petite set is small.*

Proof. See for instance [12] Theorem 5.5.7. □

1.1.5 Feller Chains

As our focus in Chapter three will lie on multivariate GARCH models which are an instance of so-called Feller chains, we briefly introduce this concept. Furthermore, we give a condition for a Feller chain such that every compact set is petite.
Recall that a function h from S to $\mathbb{R}$ is lower semicontinuous if

$$\forall x \in S \qquad \liminf_{y\to x} h(y) \geq h(x).$$

Definition 1.1.11. *If $P(\,\cdot\,, O)$ is lower semicontinuous for any open set $O \in \mathcal{B}(S)$, then the chain $(X_t)_{t\in\mathbb{N}}$ with transition probability kernel P is called* ***Feller chain****.*

Denoting the class of bounded continuous functions from S to $\mathbb{R}$ by $\mathcal{C}_b(S)$, the Feller property is frequently defined by requiring that the transition probability kernel P maps $\mathcal{C}_b(S)$ to $\mathcal{C}_b(S)$. That this is indeed equivalent to Definition 1.1.11 is shown, for example, in [12] Proposition 6.1.1 (i).

Example 1.1.12. Let $(X_t)_{t\in\mathbb{N}}$ be a Markov chain with values in $(S, \mathcal{B}(S))$ defined by the equation

$$X_{t+1} = F(X_t, \eta_{t+1}), \quad t \in \mathbb{N},$$

where

(i) F is a continuous map from $S \times \mathbb{R}^d$ into S

(ii) $(\eta_t)_{t\in\mathbb{N}^*}$ is a sequence of i.i.d. random variables with distribution Γ.

We claim that this chain $(X_t)_{t\in\mathbb{N}}$ is a Feller chain.
Indeed, for each fixed $y \in \mathbb{R}^d$, the mapping $x \mapsto F(x, y)$ is continuous. Thus, whenever $h : S \to \mathbb{R}$ is bounded and continuous (i.e. $h \in \mathcal{C}_b(S)$), $h \circ F(\cdot, y)$ is also bounded and continuous for each fixed $y \in \mathbb{R}^d$. It follows from the Dominated Convergence Theorem that

$$\begin{aligned} Ph(x) &= \mathbb{E}[h \circ F(x, \eta_1)] \\ &= \int_{\mathbb{R}^d} \Gamma(d\eta) h \circ F(x, \eta) \end{aligned}$$

is a continuous function of $x \in S$. Clearly Ph is also bounded, i.e. $Ph \in \mathcal{C}_b(S)$. Thus the chain $(X_t)_{t\in\mathbb{N}}$ is a Feller chain.

We recall that supp ψ, the support of the probability measure ψ, is the smallest closed subset of S with measure 1.
It is possible to find auxiliary conditions under which all compact sets are petite for a Feller chain. More precisely, one has the following proposition:

Proposition 1.1.13. *Suppose that the Markov chain $(X_t)_{t\in\mathbb{N}}$ is ψ - irreducible. If $(X_t)_{t\in\mathbb{N}}$ has the Feller property and* supp ψ *has non-empty interior, then all compact subsets of S are petite.*

Proof. See for instance [12] Proposition 6.2.8 (ii). □

Using this proposition, one may deduce immediately the following corollary:

Corollary 1.1.14. *Suppose that the Markov chain $(X_t)_{t\in\mathbb{N}}$ is ψ - irreducible and aperiodic. If $(X_t)_{t\in\mathbb{N}}$ has the Feller property and* supp ψ *has non-empty interior, then all compact subsets of S are small.*

Proof. Apply Theorem 1.1.10 and Proposition 1.1.13. □

1.1.6 Transience, Recurrence and Harris Recurrence

For developing the dichotomy theorem concerning transience and recurrence of Markov chains, respectively, our focus will lie upon the behavior of the **occupation time**

$$\eta_A := \sum_{t=1}^{\infty} \mathbb{1}_{\{X_t \in A\}}$$

which counts the number of visits to a set $A \in \mathcal{B}(S)$. We give the following definition for transience and recurrence of sets:

Definition 1.1.15. *The set $A \in \mathcal{B}(S)$ is called **recurrent** if $\mathbb{E}_x[\eta_A] = \infty$ for all $x \in A$. It is called **uniformly transient** if there exists $M < \infty$ such that $\mathbb{E}_x[\eta_A] \leq M$ for all $x \in A$.*

With these definitions one can show the following theorem:

Theorem 1.1.16. *Suppose that $(X_t)_{t\in\mathbb{N}}$ is ψ - irreducible. Then either*

*(i) every set in $\mathcal{B}^+(S)$ is recurrent, in which case the chain is said to be **recurrent**, or*

*(ii) there is a countable cover of S with uniformly transient sets, in which case we call the chain **transient**.*

Furthermore every small set is uniformly transient, when (ii) holds.

Proof. See for instance [12] Theorem 8.0.1. □

There is a well-known stronger concept of recurrence which is called Harris recurrence. We denote by

$$Q(x, A) := \mathbb{P}_x(\eta_A = \infty)$$

the probability that the chain, starting in $x \in S$, visits the set $A \in \mathcal{B}(S)$ infinitely often.

Definition 1.1.17. *Let $A \in \mathcal{B}(S)$. Then A is **Harris recurrent** if $Q(x, A) = 1$ for all $x \in A$. The chain $(X_t)_{t\in\mathbb{N}}$ is called **Harris recurrent** if it is ψ - irreducible and every set in $\mathcal{B}^+(S)$ is Harris recurrent.*

Remark 1.1.18.

(i) Let $(X_t)_{t\in\mathbb{N}}$ be Harris recurrent. Then the chain is also recurrent, since $Q(x, A) = 1$ implies $\mathbb{E}_x[\eta_A] = \infty$. In this sense Harris recurrence is a stronger property than recurrence.

(ii) It is obvious that $Q(x,A) \leq L(x,A)$ holds for all $x \in S,\ A \in \mathcal{B}(S)$. In fact, one can even show that the definition of Harris recurrence in terms of Q is identical to a similar definition in terms of L. Even if the latter one is often used the one above illustrates better the difference between recurrence and Harris recurrence.

As announced in subsection 1.1.3, the existence of a "special" small set can imply Harris recurrence for the chain. More precisely, one has

Proposition 1.1.19. *Suppose that $(X_t)_{t\in\mathbb{N}}$ is ψ - irreducible. If there exists a small set $C \in \mathcal{B}(S)$ such that $L(x,C) = 1$ for all $x \in S$, then the chain is Harris recurrent.*

Proof. See for instance [12] Proposition 9.1.7 (ii). □

1.1.7 Ergodicity

If μ is a signed measure on $(S,\mathcal{B}(S))$, then the total variation norm $\|\mu\|_{var}$ is defined as

$$\|\mu\|_{var} := \sup_{f:|f|\leq 1} |\mu(f)| = \sup_{A\in\mathcal{B}(S)} \mu(A) - \inf_{A\in\mathcal{B}(S)} \mu(A)$$

where $\mu(f) := \int_S \mu(dx)f(x)$.

Definition 1.1.20. *A Markov chain $(X_t)_{t\in\mathbb{N}}$ with transition probability kernel P is called **ergodic** (resp. **geometrically ergodic**) if $(X_t)_{t\in\mathbb{N}}$ is positive Harris recurrent with invariant probability measure π and*

$$\forall x \in S \qquad \lim_{n\to\infty} \|P^n(x,\cdot) - \pi\|_{var} = 0$$

(resp. $\forall x \in S \quad \lim_{n\to\infty} \|P^n(x,\cdot) - \pi\|_{var} = o(\rho^n)$ *for some* $\rho < 1$ *independent of* x*).*

Theorem 1.1.21.
If $(X_t)_{t\in\mathbb{N}}$ is positive Harris recurrent and aperiodic, then it is ergodic.

Proof. See for instance [12] Theorem 13.3.3. □

1.1.8 β - Mixing

Let $(X_t)_{t\in\mathbb{Z}}$ be a strictly stationary process. We set $\mathcal{F}^k := \sigma(X_k, X_{k+1}, X_{k+2}, \ldots)$ for $k \in \mathbb{Z}$ and $\mathcal{F}_0 := \sigma(\ldots, X_{-2}, X_{-1}, X_0)$. We consider the following mixing condition:

Definition 1.1.22.

(i) We call $\beta_k := \mathbb{E}\left[\ \sup\left\{|\mathbb{P}(B|\mathcal{F}_0) - \mathbb{P}(B)| : B \in \mathcal{F}^k\right\}\right]$ *the* β ***- mixing coefficient*** *(in [5] called coefficient of complete regularity).*

(ii) If $\beta_k \to 0$ *as* $k \to \infty$*, then* $(X_t)_{t\in\mathbb{Z}}$ *is called* β ***- mixing****.*

Intuitively, the β - mixing property of a strictly stationary process expresses a sort of "asymptotic independence" of the process.

1.1.9 Criterion for Ergodicity and β - Mixing

The central theorem of this subsection provides a criterion for ergodicity of a Markov chain with transition probability kernel P. We need the following well-known

"Foster - Lyapounov" - condition:
There exists a small set C, positive constants $\alpha < 1,\ b < \infty$ and a function $V \geq 1$ bounded on C such that

$$\forall x \in S \qquad PV(x) \leq \alpha \cdot V(x) + b \cdot \mathbb{1}_C(x). \tag{FL}$$

At first, ergodicity under the condition (FL) was studied in [19]. In [12] one can find results for ergodicity under more general conditions. Nevertheless the following result will be sufficient for this work.

Theorem 1.1.23. *Let* $(X_t)_{t\in\mathbb{N}}$ *be a* ψ *- irreducible Markov chain on* $(S, \mathcal{B}(S))$ *with transition probability kernel* P*. If the chain is aperiodic and the (FL) - condition holds, then* $(X_t)_{t\in\mathbb{N}}$ *is positive Harris recurrent, geometrically ergodic and the strictly stationary process* $(X_t)_{t\in\mathbb{Z}}$ *is geometrically* β *- mixing. Furthermore,* $\pi(V) < \infty$*.*

Proof.

- **Harris recurrence:**
  We set $V' := V - 1,\ f := 1 - \alpha,\ s := b \cdot \mathbb{1}_C$. Since (FL) holds, $V',\ f,$ and s are non-negative functions satisfying

$$\begin{aligned} PV'(x) &= PV(x) - 1 \leq \alpha V(x) + s(x) - 1 \\ &= \alpha V'(x) - f(x) + s(x) \leq V'(x) - f(x) + s(x). \end{aligned}$$

Hence, using Theorem 14.2.2 [12], we have for all $x \in S$

$$\mathbb{E}_x[\tau_C] = \frac{1}{1-\alpha}\mathbb{E}_x\left[\sum_{k=0}^{\tau_C-1} f(X_k)\right] \leq \frac{1}{1-\alpha}\left(V(x) + \underbrace{\mathbb{E}_x\left[\sum_{k=0}^{\tau_C-1} s(X_k)\right]}_{=b\cdot\mathbb{1}_C(x)}\right) < \infty. \quad (1.1)$$

Thus $L(x,C) = 1$ for all $x \in S$. We use Proposition 1.1.19 to conclude.

- **Positivity:**
  Since the chain is in particular recurrent (see Remark 1.1.18 (i)) and

$$\sup_{x\in C}\mathbb{E}_x[\tau_C] \overset{(1.1)}{\leq} \frac{1}{1-\alpha}\Big(\underbrace{\sup_{x\in C} V(x)}_{<\infty \text{ (FL)}} + b\Big) < \infty$$

  we can use Theorem 10.0.1 [12] to show the existence of a P - invariant probability measure π, i.e. the chain is positive.

- **Ergodicity:**
  The chain is aperiodic by assumption and thus it is also geometrically ergodic with

$$\forall x \in S \quad \forall n \in \mathbb{N}^* \qquad \|P^n(x,\cdot) - \pi\|_{var} \leq R \cdot r^n \cdot V(x) \quad (1.2)$$

  for some positive constants $R < \infty,\ r < 1$ (see for instance Theorem 19.1.3 [12]). To see $\pi(V) = \int_S \pi(dx)V(x) < \infty$, we refer to [20] Theorem 3, where we have to choose $f := V,\ A := C$ and $\delta := 1-\alpha$.

- **β - Mixing:**
  According to Proposition 1 in [5], the coefficient of β - mixing of the strictly stationary process $(X_t)_{t\in\mathbb{Z}}$ is given by

$$\beta_k = \int_S \pi(dx)\left\|P^k(x,\cdot) - \pi\right\|_{var}.$$

  Then (1.2) yields $\beta_k \leq R \cdot r^k \cdot \pi(V)$. Since $\pi(V) < \infty,\ R < \infty$ and $r < 1$, we obtain that the strictly stationary process is geometrically β - mixing.

□

Remark 1.1.24.
The (FL) - condition is often considered a state dependent drift condition since:

- $\forall x \notin C \quad PV(x) \leq \underbrace{\alpha}_{<1} \cdot V(x) \rightsquigarrow$ drift towards the small set C

- $\forall x \in C \quad PV(x) \leq \alpha V(x) + b \leq \alpha \sup_{x\in C} V(x) + b < \infty \rightsquigarrow PV(\cdot)$ bounded on C

1.2 Algebraic Geometry

In this section we are going to give an introduction to real algebraic geometry and highlight some important results which we will use in Chapter two. For a more detailled exposure we refer to the book of Benedetti and Risler [1].

1.2.1 Semi-algebraic and Algebraic Sets

We denote by $\mathbb{R}[X_1, \ldots, X_n]$ the polynomial ring in n variables formed from the set of polynomials in the variables $X_1, \ldots, X_n$ with coefficients in the field $\mathbb{R}$.

Definition 1.2.1. *A subset $V \subseteq \mathbb{R}^n$ is called* ***semi-algebraic*** *if it admits some representation of the form*

$$V = \bigcup_{i=1}^{s} \bigcap_{j=1}^{r_i} \{x \in \mathbb{R}^n : P_{i,j}(x) \ \sim_{ij} \ 0\} ,$$

where, for all $i = 1, \ldots, s$ and $j = 1, \ldots, r_i$,

(1) $\sim_{ij} \in \{>, =, <\}$

(2) $P_{i,j}(X) \in \mathbb{R}[X], \ X = (X_1, \ldots, X_n)$.

It is easy to see that the semi-algebraic sets of $\mathbb{R}^n$ constitute the smallest family of subsets of $\mathbb{R}^n$ such that

(i) it contains all the sets of the form

$$\{x \in \mathbb{R}^n : P(x) \geq 0, \ P(X) \in \mathbb{R}[X]\}$$

(ii) it is closed with respect to the set-theoretic operations of finite union, finite intersection and complementation.

Recall that a subset I of a commutative ring A is said to be an ideal of A if it is an additive subgroup of A and $ax \in I$ for all $a \in A$ and $x \in I$.
We will henceforth focus on special semi-algebraic sets, the algebraic sets.

Definition 1.2.2. *A subset $V \subseteq \mathbb{R}^n$ is called **algebraic** if it can be represented as*

$$V = \{x \in \mathbb{R}^n : P_1(x) = \ldots = P_k(x) = 0\}$$

where $k \in \mathbb{N}^*$ and $P_i(X) \in \mathbb{R}[X_1, \ldots, X_n]$ for all $i = 1, \ldots, k$.
If I is the ideal generated by $P_1, \ldots, P_k$, denoted by $I = (P_1, \ldots, P_k)$, we set $V(I) := V$.

Remark 1.2.3.
As an important difference to complex algebraic geometry, real algebraic sets can be represented by one single polynomial, namely, if $V = \{P_1 = \ldots = P_k = 0\}$, then we can take $P := P_1^2 + \ldots + P_k^2$.

Definition 1.2.4.
Let V be an algebraic set. We define $I(V) := \{P \in \mathbb{R}[X_1, \ldots, X_n] : P(x) = 0 \quad \forall x \in V\}$.
*Clearly $I(V)$ is an ideal of $\mathbb{R}[X_1, \ldots, X_n]$ and it is called the **ideal of V**.*

Note that $I(V)$ is not necessarily the ideal generated by the polynomials $P_1, \ldots, P_k$ which define V (cf. Example 1.2.14 (i)). In [17] it has been shown that this is the case if and only if, for every $l \in \mathbb{N}$ and every $Q_1, \ldots, Q_l \in \mathbb{R}[X_1, \ldots, X_n]$, $Q_1^2 + \ldots + Q_l^2 \in (P_1, \ldots, P_k)$ implies that each $Q_i \in (P_1, \ldots, P_k)$.
The Hilbert Basis Theorem (cf. for example [11]) says that $\mathbb{R}[X_1, \ldots, X_n]$ is Noetherian. Intuitively a ring is called Noetherian if there is no infinite nesting of strictly ascending substructures possible. Hence every ideal of $\mathbb{R}[X_1, \ldots, X_n]$ is finitely generated, i.e. for every ideal I of $\mathbb{R}[X_1, \ldots, X_n]$ there exist $k \in \mathbb{N}^*$ and polynomials $P_1, \ldots, P_k \in I$ such that $I = (P_1, \ldots, P_k)$. In particular, for all algebraic sets $V \subseteq \mathbb{R}^n$, $I(V)$ is finitely generated, i.e. there exist $P_1, \ldots, P_k \in I(V)$ such that, for all $P \in I(V)$, there exist $Q_1, \ldots, Q_k \in \mathbb{R}[X_1, \ldots, X_n]$ satisfying

$$P = Q_1 P_1 + \ldots + Q_k P_k.$$

We adopt the notation $I(V) = (P_1, \ldots, P_k)$ and we say that $I(V)$ is generated by $P_1, \ldots, P_k$.
Now we want to develop the so-called "descending chain condition", a consequence of the Hilbert Basis Theorem, which will allow us to introduce the "Zariski topology".

Proposition 1.2.5 ("descending chain condition")**.**
Let $(V_i)_{i \in \mathbb{N}^*}$ be a descending sequence of algebraic sets, i.e. $V_1 \supseteq V_2 \supseteq V_3 \supseteq \ldots$. Then the sequence stabilizes, i.e. there exists some $l \in \mathbb{N}^*$ with $V_k = V_l$ for all $k \geq l$.

Proof. $V_1 \supseteq V_2 \supseteq V_3 \supseteq \ldots$ implies by definition $I(V_1) \subseteq I(V_2) \subseteq I(V_3) \subseteq \ldots$.
Using the fact that $\mathbb{R}[X_1, \ldots, X_n]$ is Noetherian, there must exist some $l \in \mathbb{N}^*$ with

$I(V_l) = I(V_{l+1}) = \ldots$.
Since, for arbitrary algebraic sets V and W, the ideal of V contains the ideal of W if and only if V is a subset of W (cf. [11], p. 381), we have for all $k \in \mathbb{N}^*$

$$I(V_k) = I(V_{k+1}) \Leftrightarrow V_k = V_{k+1}.$$

Hence $V_l = V_{l+1} = \ldots$. □

Theorem 1.2.6 ("Zariski topology").
The algebraic sets in $\mathbb{R}^n$ are the closed sets of a topology on $\mathbb{R}^n$ called the ***"Zariski topology"****.*

Proof.
Recall that a family T of subsets of $\mathbb{R}^n$ is said to be a topology on $\mathbb{R}^n$ if it satisfies the following properties:

(i) $\emptyset \in T$ and $\mathbb{R}^n \in T$.

(ii) If $O_1 \in T$ and $O_2 \in T$, then $O_1 \cap O_2 \in T$.

(iii) If $O_i \in T$ for all $i \in I$ where I is an arbitrary index set, then $\cup_{i \in I} O_i \in T$.

The open sets in $\mathbb{R}^n$ are then defined to be the members of T. Since a subset of $\mathbb{R}^n$ is said to be closed if its complement is open, i.e. if its complement is in T, it is clear that the only non-trivial fact to show for the claim of the proof is that the intersection of any infinite family of Zariski closed sets is Zariski closed. Since finite intersections of algebraic sets are again algebraic sets, we deduce with Proposition 1.2.5 that any infinite intersection of Zariski closed sets is actually a finite intersection and thus Zariski closed. □

Remark 1.2.7.

(i) A topological space is said to be a Hausdorff space if distinct points have disjoint neighborhoods. The Zariski topology is not Hausdorff.

(ii) Every Zariski closed set in $\mathbb{R}^n$ is also closed in the usual topology because polynomial functions are continuous with respect to the usual topology. Thus, the usual topology is "stronger" than the Zariski topology.

(iii) We define the Zariski closure of a set A by ${}^{Z}\overline{A} := \bigcap_{\substack{B \text{ Zariski closed} \\ B \supseteq A}} B$.

From the "descending chain condition" (Proposition 1.2.5) it is easy to deduce a canonical decomposition of any algebraic set V into "irreducible components".

Definition 1.2.8. *An algebraic set $V \subseteq \mathbb{R}^n$ is said to be **irreducible** if it cannot be decomposed as $V = V_1 \cup V_2$, where both V_1 and V_2 are algebraic sets and $V_1 \neq V$ and $V_2 \neq V$.*
*If V is an irreducible algebraic set, it is also called **algebraic variety**.*

Proposition 1.2.9. *Any algebraic set $V \subseteq \mathbb{R}^n$ can be decomposed into a finite union of irreducible algebraic sets*

$$V = V_1 \cup \ldots \cup V_s.$$

If $V_i \not\subseteq V_j$ whenever $i \neq j$, then this decomposition is unique (up to reordering the indices). It is called the decomposition of V into irreducible components: every V_i is an irreducible component of V.

Proof. See for instance [1] Proposition 3.1.5. □

1.2.2 Regular Points and Dimension of Algebraic Varieties

We shall introduce the regular and the singular part of an algebraic variety.
To do this we specify the subtle notion of a regular point of an algebraic variety.

Definition 1.2.10. *Let $V \subseteq \mathbb{R}^n$ be an algebraic variety with $I(V) = (P_1, \ldots, P_k)$. We call*

$$\rho(V) := \sup_{x \in V} rank \left(\frac{\partial P_i}{\partial x_j}(x) \right)_{\substack{1 \leq i \leq k \\ 1 \leq j \leq n}}$$

*the **rank of V**.*

Note that $\rho(V)$ does not depend on the choice of the generators $P_1, \ldots, P_k$ of $I(V)$. If $P_{k+1} = \sum_{i=1}^{k} Q_i P_i$, then

$$rank \left(\frac{\partial P_i}{\partial x_j}(x) \right)_{\substack{1 \leq i \leq k \\ 1 \leq j \leq n}} = rank \left(\frac{\partial P_i}{\partial x_j}(x) \right)_{\substack{1 \leq i \leq k+1 \\ 1 \leq j \leq n}} \quad \forall x \in V$$

since the last row of the second matrix is a linear combination of the old rows. Hence Definition 1.2.10 is well posed.

Definition 1.2.11. *Let $V \subseteq \mathbb{R}^n$ be an algebraic variety with $I(V) = (P_1, \ldots, P_k)$.*

*(i) Then $x_0 \in V$ is said to be a **regular point of V** if*

$$\rho(V) = rank \left(\frac{\partial P_i}{\partial x_j}(x_0) \right)_{\substack{1 \leq i \leq k \\ 1 \leq j \leq n}}.$$

*Otherwise x_0 is called **singular point of V**.*

(ii) *We write $\mathcal{R}(V)$ to denote the set of regular points of V and $\mathcal{S}(V)$ for the set of singular points.*

(iii) *V is called a* ***smooth algebraic variety*** *if $V = \mathcal{R}(V)$.*

Remark 1.2.12.

(i) $\mathcal{R}(V)$ is a non-empty analytic submanifold of $\mathbb{R}^n$ (cf. [1] Proposition 3.2.9).

(ii) $\mathcal{S}(V)$ is an algebraic set properly contained in V (cf. [1] Proposition 3.2.4).

(iii) Note that, for every algebraic set $V \subseteq \mathbb{R}^n$, $\mathcal{R}(V)$ is a semi-algebraic set, because $\mathcal{S}(V)$ is algebraic. In fact, it is a Zariski open set of V (cf. [1] Remarks 3.4.7).

We are now able to define the dimension of an algebraic variety.

Definition 1.2.13. *Let $V \subseteq \mathbb{R}^n$ be an algebraic variety. The* ***(algebraic) dimension*** *of V is defined as*

$$\dim V := n - \rho(V).$$

The dimension of a semi-algebraic variety (i.e. a semi-algebraic set whose Zariski closure is irreducible) is the dimension of its Zariski closure.

One can show that the algebraic dimension coincides with the analytic dimension of $\mathcal{R}(V)$.

Example 1.2.14.

(i) Consider $V = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 0\} = \{(0, 0)\}$.
Then $I(V) = (X, Y) \neq (X^2 + Y^2)$ which shows that the ideal of V is not necessarily the ideal generated by the polynomials which define V (cf. comment after Definition 1.2.4).

(ii) Let $V = \{(x, y) \in \mathbb{R}^2 : x(y^2 + x^2 - x^3) = 0\}$. Then we see that V is reducible since $V = \{(x, y) \in \mathbb{R}^2 : x = 0\} \cup \{(x, y) \in \mathbb{R}^2 : y^2 = x^3 - x^2\} =: V_1 \cup V_2$. V_1 and V_2 are the irreducible components of V (cf. Figure 1.1).

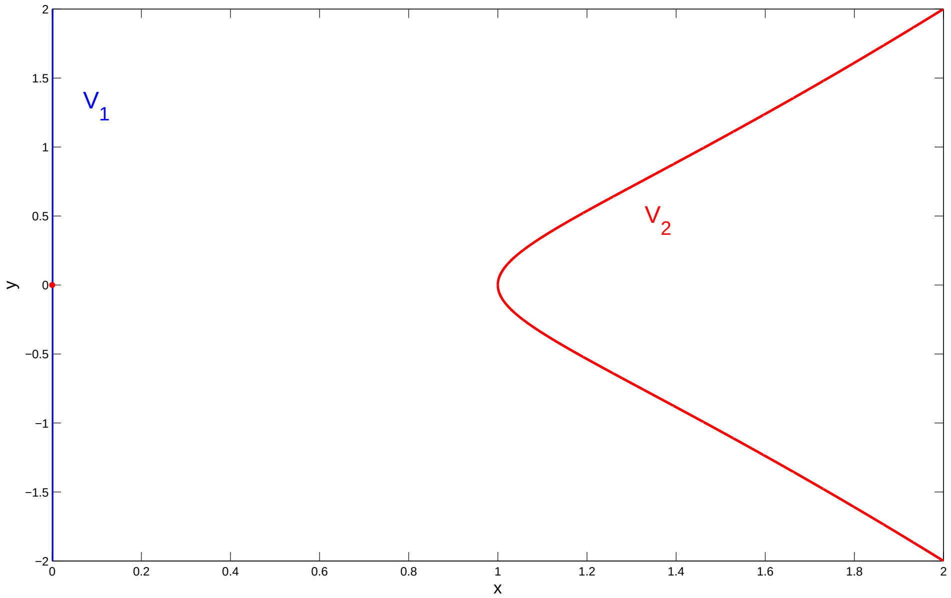

Figure 1.1: Algebraic set $V = \{(x, y) \in \mathbb{R}^2 : x(y^2 + x^2 - x^3) = 0\}$

(iii) Consider the second component V_2 of V in (ii). We define $P := Y^2 + X^2 - X^3$. V_2 is an algebraic variety and its ideal is generated by P, i.e. $I(V_2) = (P)$. We obtain $\rho(V_2) = \sup_{(x,y)\in V_2} rank\,(2x - 3x^2, 2y) = 1$ and $\mathcal{S}(V) = \{(0, 0)\}$. Furthermore the dimension of V_2 is, as expected, $\dim V_2 = 1$.

Proposition 1.2.15.
If $V \subseteq W$ are two algebraic varieties, then: $\dim V = \dim W \Leftrightarrow V = W$.

Proof. See for instance [1] Corollary 3.4.5. □

As an immediate consequence of this proposition we obtain

Corollary 1.2.16. *If $V_1 \subseteq V_2 \subseteq V_3 \subseteq \ldots$ is an ascending sequence of algebraic varieties in $\mathbb{R}^n$, then there exists $l \in \mathbb{N}^*$ such that $V_k = V_l$ for all $k \geq l$.*

1.2.3 Regular Maps

In the following we introduce a natural class of maps such that preimages of algebraic sets are again algebraic, i.e. maps which are continuous with respect to the Zariski topology.

Definition 1.2.17. *Let $V \subseteq \mathbb{R}^n$ and $W \subseteq \mathbb{R}^m$ be algebraic varieties. Then $f : V \to W$ is said to be a **regular map**, if all its components $(f_i)_{1\leq i\leq m}$ are regular functions, i.e., for* *all $i = 1, \dots, m$, there exist $P_i, Q_i \in \mathbb{R}[X_1, \dots, X_n]$ such that*

$$V \cap \{x \in \mathbb{R}^n : Q_i(x) = 0\} = \emptyset \quad \text{and} \quad f_i(x) = \frac{P_i(x)}{Q_i(x)} \quad \forall x \in V.$$

Proposition 1.2.18. *Let $V \subseteq \mathbb{R}^n$, $W \subseteq \mathbb{R}^m$ be algebraic varieties and $f : V \to W$ a regular map. Then f is continuous with respect to the Zariski topology.*

Proof. Let Z be a Zariski closed set in W. Using Remark 1.2.3 there exists a polynomial $P \in \mathbb{R}[X_1, \dots, X_m]$ such that

$$Z = \{y \in W : P(y) = 0\}$$

and thus

$$f^{-1}(Z) = \{x \in V : P(f(x)) = 0\}.$$

Since f is a regular map, there exist two polynomials $G, H \in \mathbb{R}[X_1, \dots, X_n]$ such that $P(f(x)) = G(x)/H(x)$ and $H(x) \neq 0$ for all $x \in V$.
Hence

$$f^{-1}(Z) = \{x \in V : G(x) = 0\},$$

i.e. $f^{-1}(Z)$ is a Zariski closed set in V. □

Before we consider the concept of smoothness for regular maps, we cite the basic theorem of Tarski-Seidenberg.

Theorem 1.2.19 (Tarski-Seidenberg)**.**
The image of an algebraic variety under a regular map is a semi-algebraic set.

Proof. See for instance [1] Theorem 2.3.4. □

Definition 1.2.20. *Let $V \subseteq \mathbb{R}^n$, $W \subseteq \mathbb{R}^m$ be algebraic varieties and $f : V \to W$ a regular map. Then we say that f is* ***smooth*** *at $x_0 \in V$ if*

(i) $x_0 \in \mathcal{R}(V)$

(ii) $f(x_0) \in \mathcal{R}(W)$ *and*

(iii) the restriction of Df to $\mathcal{R}(V)$ has rank $\dim W$ at the point x_0.

Otherwise x_0 is said to be a ***critical point*** *of f and its image is called a* ***singular value*** *of f.*

There is an easy criterion to check whether a given point is a smooth point, namely the Jacobian criterion (cf. [14] p. 42, cf. also [13] (A21)).

Theorem 1.2.21 (Jacobian criterion)**.** *Let* $V \subseteq \mathbb{R}^n$, $W \subseteq \mathbb{R}^m$ *be algebraic varieties and* $f : V \to W$ *a regular map. Suppose that* $\dim V = p$, $\dim W = q$, $x \in \mathcal{R}(V)$ *and* $f(x) \in \mathcal{R}(W)$*. If* $I(V) = (P_1, \ldots, P_k)$*, then* f *is smooth at* x *if and only if*

$$rank \left(\left(\frac{\partial P_i}{\partial x_j}(x) \right)_{\substack{1 \le j \le n \\ 1 \le i \le k}} \middle| \left(\frac{\partial f_i}{\partial x_j}(x) \right)_{\substack{1 \le j \le n \\ 1 \le i \le m}} \right) = n - p + q$$

Definition 1.2.22. *Let* $V \subseteq \mathbb{R}^n$, $W \subseteq \mathbb{R}^m$ *be algebraic varieties and* $f : V \to W$ *a regular map.* f *is said to be* ***dominating*** *if* W *is the Zariski closure of* $f(V)$*.*

Proposition 1.2.23 (cf. [13] (A23))**.**
Let $V \subseteq \mathbb{R}^n$, $W \subseteq \mathbb{R}^m$ *be algebraic varieties and* $f : V \to W$ *a regular map. Then the following statements are equivalent:*

(i) f *is dominating.*

(ii) $f(V)$ *contains a non-emtpy open set of* $\mathcal{R}(W)$*.*

(iii) f *has a smooth point.*

Note that if f is a dominating regular map, one has necessarily $\dim V \geq \dim W$.

Chapter 2

Autoregressive Processes defined by a Composition of a Regular Map and a Diffeomorphism

2.1 Introduction

In view of an application to multivariate GARCH processes, which we will present in Section 3.3 in an autoregressive manner, we are going to study in this chapter Markov chains which can be written in the form $X_{t+1} = F(X_t, e_t)$ where F is an appropriate map and $(e_t)_{t\in\mathbb{N}}$ is an i.i.d. innovation sequence (cf. Introduction (1)).
In the case where F is a regular map, Mokkadem showed mixing, ergodicity and recurrence properties for these chains in [13]. His results apply for example to the well-known ARMA (autoregressive moving average) models.
However, for an application to GARCH processes, it will turn out that we have to consider F more general, namely allow F to be a composition of a regular map and a diffeomorphism, i.e. the chain can be written as

$$X_{t+1} = F(X_t, e_t) = \varphi(X_t, f_{X_t}(e_t))$$

where φ is a regular map, $f_z(\,\cdot\,)$ is a diffeomorphism for every possible state z of the chain and $(e_t)_{t\in\mathbb{N}}$ is an i.i.d. innovation sequence. We shall call these chains semi-polynomial Markov chains.

More precisely, we will consider in the next two sections the following general problem:

Let V be an algebraic variety (cf. Definition 1.2.8) contained in U where U is an open subset of $\mathbb{R}^n$ and let $F : U \times \mathbb{R}^m \to \mathbb{R}^n$ be a C^1 - map such that there exist a C^1 - map $f : U \times \mathbb{R}^m \to \mathbb{R}^m$ and a map $\varphi : U \times \mathbb{R}^m \to \mathbb{R}^n$ satisfying:

(2.1) $\varphi(V \times \mathbb{R}^m) \subseteq V$

(2.2) for all $z \in U$, the map $f_z(\cdot) = f(z, \cdot)$ is a C^1 - diffeomorphism from $\mathbb{R}^m$ onto $\mathbb{R}^m$ and the map $U \times \mathbb{R}^m \to \mathbb{R}^m$, $(z, y) \mapsto f_z^{-1}(y)$ is continuous in (z, y) where $f_z^{-1}(\cdot)$ denotes the inverse map of $f_z(\cdot)$

(2.3) the map $\varphi(z, y)$ is regular in (z, y) (cf. Definition 1.2.17)

(2.4) for all $(z, y) \in U \times \mathbb{R}^m$, $F(z, y) = \varphi(z, f_z(y))$.

2.2 Properties of the Image Measure

In general, the image of $\mathbb{R}^m$ under $F_z(\cdot) = F(z, \cdot)$ is a semi-algebraic set in $\mathbb{R}^n$ with dimension less than n (Theorem 1.2.19 Tarski-Seidenberg). Thus the Lebesgue measure of this image is often zero.
That is why we need to work with Hausdorff measures (for more details concerning these measures see for example [6]). In particular, we will suppose that the algebraic variety V is furnished with a regular measure μ_V obtained by equipping the regular set $\mathcal{R}(V)$ of V with an appropriate Hausdorff measure which is extended by zero to the singular set $\mathcal{S}(V)$.
Our first question is, what can we say about the image measure of Γ under the map $F_z(\cdot)$ where Γ is a measure absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^m$ with density γ?
For $z \in V$ we denote by Γ_z this image measure in V. Furthermore we define the domain of positivity of the density γ by $E := \{y \in \mathbb{R}^m : \gamma(y) > 0\}$. Note that the notion of smooth points (cf. Definiton 1.2.20) also makes sense for general C^1 - maps.

Theorem 2.2.1. *Suppose that* $z_0 \in V$ *and* $F_{z_0}(\cdot)$ *has a smooth point in* $\mathbb{R}^m$. *Then* Γ_{z_0} *is absolutely continuous with respect to the measure* μ_V *and its density has* $F_{z_0}(E)$ *as domain of positivity.*

To prove this we cite a lemma shown by Mokkadem [13].

Lemma 2.2.2. *Let* $z_0 \in V$ *and* $\varphi_{z_0}(\cdot) = \varphi(z_0, \cdot)$ *be dominating from* $\mathbb{R}^m$ *to* V *(cf. Definition 1.2.22). Suppose that* Γ *is a measure absolutely continuous with respect to the Lebesgue measure on* $\mathbb{R}^m$ *with* E *as the density's domain of positivity. Then* *(i) the image measure $\varphi_{z_0}(\Gamma)$ in the variety V is absolutely continuous with respect to μ_V and its density has the domain of positivity $\varphi_{z_0}(E)$*

(ii) for all $A \in \mathcal{B}(V)$, $\liminf_{\substack{z\to z_0\\ z\in V}} \varphi_z(\Gamma)(A) \geq \varphi_{z_0}(\Gamma)(A)$.

Proof. See [13] Theorem 3.1 and Theorem 3.2. □

Proof of Theorem 2.2.1. First we denote by Γ'_{z_0} the image measure $f_{z_0}(\Gamma)$. We are going to show that Γ'_{z_0} is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^m$ in order to apply Lemma 2.2.2 (i) with the measure Γ'_{z_0} and the regular map $\varphi_{z_0}(\cdot)$. In fact, one obtain immediately by the Density Transformation Theorem that Γ'_{z_0} is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^m$ with density

$$\gamma'_{z_0}(y) := \gamma(f_{z_0}^{-1}(y)) \cdot \frac{1}{\left|\det Df_{z_0}(f_{z_0}^{-1}(y))\right|}\ , \quad y \in \mathbb{R}^m. \tag{2.5}$$

The density's domain of positivity is given by

$$\left\{y \in \mathbb{R}^m : \gamma'_{z_0}(y) > 0\right\} = \left\{y \in \mathbb{R}^m : f_{z_0}^{-1}(y) \in E\right\} = f_{z_0}(E).$$

To apply Lemma 2.2.2 we must ensure that $\varphi_{z_0}(\cdot)$ is dominating.
In fact, by assumption there exists $x_0 \in \mathbb{R}^m$ such that $F_{z_0}(x_0) = \varphi_{z_0}(y_0)$ is a regular point of V and $rank(DF_{z_0}(x_0)) = \dim V$ where $y_0 = f_{z_0}(x_0)$. Since $F_{z_0}(\cdot) = \varphi_{z_0}(f_{z_0}(\cdot))$ we have $DF_{z_0}(x_0) = D\varphi_{z_0}(y_0) \cdot Df_{z_0}(x_0)$. Due to the C^1 - diffeomorphism $f_{z_0}(\cdot)$, the linear map $Df_{z_0}(x_0)$ is invertible. Thus $rank(D\varphi_{z_0}(y_0)) = \dim V$. Since $\varphi_{z_0}(y_0) \in \mathcal{R}(V)$ (see above) and $y_0 \in \mathcal{R}(\mathbb{R}^m)$ ($\mathbb{R}^m$ is a smooth algebraic variety, i.e. $\mathcal{R}(\mathbb{R}^m) = \mathbb{R}^m$), the regular map $\varphi_{z_0}(\cdot)$ is smooth at y_0 and hence dominating due to Proposition 1.2.23. □

Proposition 2.2.3. *Let $z_0 \in U$. Then, for every $\epsilon > 0$, there exists $\alpha > 0$ such that*

$$\left|\Gamma'_z(B) - \Gamma'_{z_0}(B)\right| = |f_z(\Gamma)(B) - f_{z_0}(\Gamma)(B)| < \epsilon$$

for all $B \in \mathcal{B}(\mathbb{R}^m)$ and every $z \in U$ with $\|z - z_0\| < \alpha$.

Proof. The image measure $\Gamma'_z = f_z(\Gamma)$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^m$ with density γ'_z given by equation (2.5).
Let $\epsilon > 0$ and $B \in \mathcal{B}(\mathbb{R}^m)$. The space of real-valued continuous functions on $\mathbb{R}^m$ with compact support is dense in the L^1 sense in the space of all Lebesgue-integrable functions on $\mathbb{R}^m$. Thus, there exists a continuous function $\tilde{\gamma} : \mathbb{R}^m \to \mathbb{R}$ with compact support K such that

$$\int_{\mathbb{R}^m} |\gamma(x) - \tilde{\gamma}(x)|\ dx < \frac{\epsilon}{3}. \tag{2.6}$$

Hence,

$$\begin{aligned}
\left|\Gamma'_z(B) - \Gamma'_{z_0}(B)\right| &= \left| \int\limits_B \gamma'_z(y)\, dy - \int\limits_B \gamma'_{z_0}(y)\, dy \right| \\
&\leq \int\limits_{\mathbb{R}^m} \left|\gamma(f_z^{-1}(y)) - \tilde{\gamma}(f_z^{-1}(y))\right| \cdot \left|\det Df_z^{-1}(y)\right|\, dy \\
&+ \int\limits_{\mathbb{R}^m} \Big|\tilde{\gamma}(f_z^{-1}(y)) \left|\det Df_z^{-1}(y)\right| - \tilde{\gamma}(f_{z_0}^{-1}(y)) \left|\det Df_{z_0}^{-1}(y)\right|\Big|\, dy \\
&+ \int\limits_{\mathbb{R}^m} \left|\tilde{\gamma}(f_{z_0}^{-1}(y)) - \gamma(f_{z_0}^{-1}(y))\right| \cdot \left|\det Df_{z_0}^{-1}(y)\right|\, dy \\
&=: I_1 + I_2 + I_3.
\end{aligned}$$

With (2.6) we obtain immediately by substitution $I_1 < \epsilon/3$ and $I_3 < \epsilon/3$.
We claim that $I_2 \to 0$ as $z \to z_0$ due to the Dominated Convergence Theorem.
Indeed, $\tilde{\gamma}$ is bounded on $\mathbb{R}^m$ by $\sup\tilde{\gamma}$. On the other hand, for all $r > 0$ such that $\overline{B(z_0, r)} := \{z \in \mathbb{R}^n : \|z - z_0\| \leq r\} \subseteq U$, the set

$$C := \left\{(z, y) \in U \times \mathbb{R}^m : \|z - z_0\| \leq r \text{ and } f_z^{-1}(y) \in K\right\}$$

is a compact set in $U \times \mathbb{R}^m$ since the map $\psi : U \times \mathbb{R}^m \to U \times \mathbb{R}^m,\ \psi(z, y) := (z, f_z(y))$ is continuous and $C = \psi(\overline{B(z_0, r)} \times K)$. Thus, there is a real number $b > 0$ such that, for all $(z, y) \in C$, $\left|\det Df_z^{-1}(y)\right| < b$. The map $\tilde{\gamma}(f_z^{-1}(y)) \cdot \left|\det Df_z^{-1}(y)\right|$ is hence bounded on C by $b \cdot \sup\tilde{\gamma}$.
Let C_1 be the projection of C on $\mathbb{R}^m$ and suppose without loss of generality $\|z - z_0\| \leq r$. Then, for all $y \notin C_1$, we have $\tilde{\gamma}(f_z^{-1}(y)) = \tilde{\gamma}(f_{z_0}^{-1}(y)) = 0$ which implies

$$I_2 = \int\limits_{C_1} \Big|\tilde{\gamma}(f_z^{-1}(y)) \left|\det Df_z^{-1}(y)\right| - \tilde{\gamma}(f_{z_0}^{-1}(y)) \left|\det Df_{z_0}^{-1}(y)\right|\Big|\, dy.$$

This integrand is dominated by $2b\sup\tilde{\gamma}$ and converges pointwise to zero if z converges to z_0 (cf. (2.2)). Since b and $\sup\tilde{\gamma}$ are finite constants and C_1 is compact the dominant $2b\sup\tilde{\gamma}$ is integrable over C_1. Hence, we can apply the Dominated Convergence Theorem and get $I_2 \to 0$ as $z \to z_0$, i.e. there exists $0 < \alpha < r$ such that $I_2 < \epsilon/3$ for all $z \in U$ with $\|z - z_0\| < \alpha$. □

Theorem 2.2.4. *Suppose that $z_0 \in V$ and $F_{z_0}(\cdot)$ has a smooth point in $\mathbb{R}^m$. Then*

$$\liminf_{\substack{z \to z_0 \\ z \in V}} \Gamma_z(A) \geq \Gamma_{z_0}(A) \tag{2.7}$$

for every $A \in \mathcal{B}(V)$.

Proof. Let $A \in \mathcal{B}(V)$ and $\epsilon > 0$. Since $\varphi_{z_0}(\cdot)$ is dominating and $\Gamma'_{z_0} = f_{z_0}(\Gamma)$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^m$ (cf. proof of Theorem 2.2.1), Lemma 2.2.2 (ii) yields a neighborhood V_0 of z_0 in V such that for all $z \in V_0$

$$\varphi_z(\Gamma'_{z_0})(A) \geq \varphi_{z_0}(\Gamma'_{z_0})(A) - \frac{\epsilon}{2}$$

which is equivalent to

$$\forall z \in V_0 \qquad f_{z_0}(\Gamma)(\varphi_z^{-1}(A)) \geq f_{z_0}(\Gamma)(\varphi_{z_0}^{-1}(A)) - \frac{\epsilon}{2}. \tag{2.8}$$

Due to Proposition 2.2.3, there exists $\alpha > 0$ such that, for all $B \in \mathcal{B}(\mathbb{R}^m)$ and every $z \in U$ with $\|z - z_0\| < \alpha$,

$$f_z(\Gamma)(B) \geq f_{z_0}(\Gamma)(B) - \frac{\epsilon}{2}.$$

We choose $B = \varphi_z^{-1}(A)$ and deduce for every $z \in U$, $\|z - z_0\| < \alpha$,

$$f_z(\Gamma)(\varphi_z^{-1}(A)) \geq f_{z_0}(\Gamma)(\varphi_z^{-1}(A)) - \frac{\epsilon}{2}. \tag{2.9}$$

With (2.8) and (2.9) we obtain for all $z \in V_0$ with $\|z - z_0\| < \alpha$

$$f_z(\Gamma)(\varphi_z^{-1}(A)) \geq f_{z_0}(\Gamma)(\varphi_{z_0}^{-1}(A)) - \epsilon.$$

Since $\varphi_z(\Gamma'_z) = F_z(\Gamma) = \Gamma_z$ this is equivalent to

$$\forall z \in V_0 \cap \{z \in \mathbb{R}^n : \|z - z_0\| < \alpha\} \qquad \Gamma_z(A) \geq \Gamma_{z_0}(A) - \epsilon.$$

This shows (2.7) since $\epsilon > 0$ can be chosen arbitrarily small. □

2.3 Semi-polynomial Markov Chains

As announced at the beginning of this chapter, we consider now a Markov chain $(X_t)_{t\in\mathbb{N}}$ satisfying an equation $X_{t+1} = F(X_t, e_t)$ where we will assume $(e_t)_{t\in\mathbb{N}}$ to be a sequence of i.i.d. random variables and F to be a C^1 - map satisfying (2.1) - (2.4). Such Markov chains will be called **semi-polynomial**.

In Theorem 2.3.8 we will show Harris recurrence, ergodicity and β - mixing for semi-polynomial Markov chains.

2.3.1 Model and Assumptions

Let V be an algebraic variety contained in $\mathbb{R}^n$ and $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ a C^1 - map satisfying the assumptions (2.1) - (2.4) of Section 2.1 (i.e. here $U = \mathbb{R}^n$).
Suppose $(e_t)_{t\in\mathbb{N}}$ to be a sequence of i.i.d. random variables in $\mathbb{R}^m$ and $(X_t)_{t\in\mathbb{N}}$ to be a semi-polynomial Markov chain defined by

$$X_{t+1} = F(X_t, e_t), \qquad t \in \mathbb{N}. \tag{2.10}$$

Concerning the sequence $(e_t)_{t\in\mathbb{N}}$ we make the following additional assumption:

(A1) Every e_t has the distribution Γ which is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^m$ with density γ. Let $E := \{y \in \mathbb{R}^m : \gamma(y) > 0\}$ denote the domain of positivity of γ.

We define for all $k \in \mathbb{N}^*$

$$F^k(z, y_1, \ldots, y_k) := F(F^{k-1}(z, y_1, \ldots, y_{k-1}), y_k)$$

where $z \in \mathbb{R}^n,\ (y_1, \ldots, y_k) \in (\mathbb{R}^m)^k$.
With this notation we introduce for $z \in V$ the orbit

$$S_z := \bigcup_{k\in\mathbb{N}^*} \left\{F^k(z, y_1, \ldots, y_k) : y_1, \ldots, y_k \in E\right\}.$$

To prove the desired properties for semi-polynomial Markov chains we assume:

(A2) There is a point $a \in \operatorname{int}(E)$ and a point $T \in V$ such that, for all $z \in V$, the sequence $(X_t^z)_{t\in\mathbb{N}}$ defined by $X_0^z = z$ and $X_t^z = F(X_{t-1}^z, a)$ for $t \geq 1$ converges to the point T.

T is called **attracting point** of the chain $(X_t)_{t\in\mathbb{N}}$.
We set $W := {}^Z\overline{S_T}$ the Zariski closure of the orbit S_T. To show uniqueness of the strictly stationary solution we need the assumption:

(A3) Any strictly stationary solution to (2.10) takes its values in the algebraic variety W.

Remark 2.3.1.

(i) If (A2) is satisfied, then T is a fixed point of $F(\,\cdot\,, a)$ since F is continuous.

(ii) If for some $a \in \operatorname{int}(E)$ the map $F(\,\cdot\,, a)$ is Lipschitz continuous with Lipschitz constant smaller than 1, then the Banach Fixed Point Theorem shows that there exists a point $T \in V$ such that (A2) is satisfied.

(iii) It is obvious that W is an algebraic set since it is the Zariski closure of S_T. In fact, it is even irreducible (cf. upcoming Section 2.3.2).

2.3.2 Algebraic Variety of States

In this subsection we suppose that the assumptions (A1) and (A2) hold. We will show that W, defined as above, is indeed an algebraic variety which we will call the Markov chain's **algebraic variety of states**.
Let $(D_k)_{k\in\mathbb{N}^*}$ be the sequence of subsets of $\mathbb{R}^n$ defined by

$$D_k := F^k(T, E^k).$$

Since $F(T,a) = T$ (cf. Remark 2.3.1 (i)) we obtain

$$D_k = F^k(F(T,a), E^k) = F^{k+1}(T, \underbrace{\{a\} \times E^k}_{\subseteq E^{k+1}}) \subseteq D_{k+1},$$

i.e. the sequence $(D_k)_{k\in\mathbb{N}^*}$ is an ascending sequence of subsets of $\mathbb{R}^n$.
We set $W_k := {}^Z\overline{F^k(T,(\mathbb{R}^m)^k)}$. Then we have $W_k = {}^Z\overline{\varphi^k(T,(\mathbb{R}^m)^k)}$ (defining φ^k analogously to F^k) since $f_T(\,\cdot\,)$ is a C^1 - diffeomorphism.

Lemma 2.3.2. *For all $k \in \mathbb{N}^*$ we have $W_k = {}^Z\overline{D_k}$.*

Proof. To this end consider the map

$$f_T^{(k)} : (\mathbb{R}^m)^k \to (\mathbb{R}^m)^k, \qquad (y_1,\dots,y_k) \mapsto (x_1,\dots,x_k)$$

where $x_1 = f_T(y_1),\ x_2 = f_{F(T,y_1)}(y_2), \dots,\ x_k = f_{F^{k-1}(T,y_1,\dots,y_{k-1})}(y_k)$.
Due to the properties of f and F (in particular (2.2)), it is clear that $f_T^{(k)}$ is bijective, continuous and its inverse is continuous as well, i.e. $f_T^{(k)}$ is a homeomorphism.
Assumption (A2) implies that E^k contains an open ball of $(\mathbb{R}^m)^k$. Thus, since $f_T^{(k)}$ is homeomorphic, $f_T^{(k)}(E^k)$ contains an open ball of $(\mathbb{R}^m)^k$.
Note that Proposition 1.2.15 also holds if the algebraic variety V in 1.2.15 is only an algebraic set (see [1] Corollary 3.4.5 and concerning the definition of the dimension of an algebraic set see [1] p. 135). Hence, 1.2.15 implies ${}^Z\overline{f_T^{(k)}(E^k)} = (\mathbb{R}^m)^k$.
This shows

$$W_k = {}^Z\overline{F^k(T,(\mathbb{R}^m)^k)} = {}^Z\overline{\varphi^k(T,(\mathbb{R}^m)^k)} = {}^Z\overline{\varphi^k\left(T, {}^Z\overline{f_T^{(k)}(E^k)}\right)}.$$

Since $\varphi^k(T,\cdot)$ is regular (cf. (2.3)), $\varphi^k(T,\cdot)$ is continuous with respect to the Zariski topology due to Proposition 1.2.18.

Hence,

$$W_k = {}^Z\overline{\varphi^k\left(T, f_T^{(k)}(E^k)\right)} = {}^Z\overline{F^k(T, E^k)} = {}^Z\overline{D_k}$$

which proves W_k to be the Zariski closure of D_k. □

Lemma 2.3.3. *W_k is irreducible for all $k \in \mathbb{N}^*$.*

Proof. If we suppose that there is $k \in \mathbb{N}^*$ such that $W_k = V_1 \cup V_2$ where V_1 and V_2 are algebraic sets with $V_1 \subsetneq W_k$ and $V_2 \subsetneq W_k$, then

$$(\mathbb{R}^m)^k = \left(\varphi_T^k\right)^{-1}(W_k) = \underbrace{\left(\varphi_T^k\right)^{-1}(V_1)}_{(*)} \cup \underbrace{\left(\varphi_T^k\right)^{-1}(V_2)}_{(**)}$$

where $\varphi_T^k(\,\cdot\,) = \varphi^k(T, \,\cdot\,)$.
$(*)$ and $(**)$ are algebraic sets because V_1 and V_2 are algebraic sets and $\varphi_T^k(\,\cdot\,)$ is continuous with respect to the Zariski topology (see proof of Lemma 2.3.2). Since $\left(\varphi_T^k\right)^{-1}(V_i) \subsetneq (\mathbb{R}^m)^k$ for $i = 1, 2$ (otherwise $W_k = {}^Z\overline{\varphi^k(T, (\mathbb{R}^m)^k)} \subseteq {}^Z\overline{V_i} = V_i$ which would be a contradiction to $V_i \subsetneq W_k$), this would prove $(\mathbb{R}^m)^k$ to be reducible which is a contradiction. □

Proposition 2.3.4. *There exists $l \in \mathbb{N}^*$ such that $W_k = W_l$ for all $k \geq l$ and $W = W_l$. In particular, W is an algebraic variety.*

Proof. Lemma 2.3.2 and Lemma 2.3.3 show that $(W_k)_{k \in \mathbb{N}^*}$ is an ascending sequence of algebraic varieties and due to Corollary 1.2.16 there exists $l \in \mathbb{N}^*$ such that $W_k = W_l$ for all $k \geq l$.
We then observe that

$$S_T = \bigcup_{k \in \mathbb{N}^*} \underbrace{\left\{F^k(T, y_1, \ldots, y_k) : y_1, \ldots, y_k \in E\right\}}_{=F^k(T, E^k)} = \bigcup_{k \in \mathbb{N}^*} D_k.$$

Since

$${}^Z\overline{\bigcup_{k \in \mathbb{N}^*} D_k} \subseteq {}^Z\overline{\underbrace{\bigcup_{k \in \mathbb{N}^*} {}^Z\overline{D_k}}_{=\bigcup_{k \in \mathbb{N}^*} W_k = W_l}} = {}^Z\overline{W_l} = W_l$$

and $W_l = {}^Z\overline{D_l} \subseteq {}^Z\overline{\cup_{k \in \mathbb{N}^*} D_k}$ we obtain $W = {}^Z\overline{S_T} = W_l$. □

Lemma 2.3.5. *For all $k \in \mathbb{N}^*$ we have $F^k(W, (\mathbb{R}^m)^k) \subseteq W$. Hence the Markov chain can be restricted to the variety of states W.*

Proof. With the definition of the subsets D_k and W_k, respectively, one has

$$\forall k \in \mathbb{N}^* \qquad F(D_k, \mathbb{R}^m) = F(\underbrace{F^k(T, E^k)}_{\subseteq F^k(T, (\mathbb{R}^m)^k)}, \mathbb{R}^m) \subseteq F^{k+1}(T, (\mathbb{R}^m)^{k+1}) \subseteq W_{k+1} \subseteq W. \quad (2.11)$$

The continuity of regular maps with respect to the Zariski topology and the regularity of φ yield

$$\begin{aligned} F(W,\mathbb{R}^m) &= \varphi(W, \underbrace{f_W(\mathbb{R}^m)}_{=\mathbb{R}^m}) = \varphi\left({}^{Z}\overline{D_l}, \mathbb{R}^m\right) \\ &\subseteq {}^{Z}\overline{\varphi\left({}^{Z}\overline{D_l}, \mathbb{R}^m\right)} = {}^{Z}\overline{\varphi(D_l, \mathbb{R}^m)} \\ &= {}^{Z}\overline{F(D_l, \mathbb{R}^m)} \overset{(2.11)}{\subseteq} W. \end{aligned}$$

By induction we have $F^k(W, (\mathbb{R}^m)^k) \subseteq W$ for all $k \in \mathbb{N}^*$. Hence we can restrict the Markov chain to the variety of states W. □

Proposition 2.3.6. *For all $A \in \mathcal{B}(W)$ and all $k \geq l$*

$$\liminf_{\substack{z \to T \\ z \in W}} P^k(z, A) \geq P^k(T, A),$$

where P^k is the k-step transition probability kernel of the Markov chain $(X_t)_{t \in \mathbb{N}}$.

Proof. Since $\varphi_T^k(\,\cdot\,) = \varphi^k(T, \,\cdot\,)$ is regular and dominating for all $k \geq l$ (since ${}^{Z}\overline{\varphi^k(T, (\mathbb{R}^m)^k)} = W_k = W$ for all $k \geq l$, cf. Proposition 2.3.4), Proposition 1.2.23 implies that $\varphi_T^k(\,\cdot\,)$ has a smooth point. As in the proof of Theorem 2.2.1 this shows that the map $F^k(T, \,\cdot\,)$ has a smooth point in $(\mathbb{R}^m)^k$.
For, let $x_0 \in (\mathbb{R}^m)^k$ be the smooth point of $\varphi_T^k(\,\cdot\,)$, i.e. $\varphi^k(T, x_0) \in \mathcal{R}(W)$ and *rank* $\left(D\varphi_T^k(x_0)\right) = \dim W$. Then $F^k\left(T, \left(f_T^{(k)}\right)^{-1}(x_0)\right) = \varphi^k(T, x_0)$ and

$$DF^k(T, \,\cdot\,)\left(\left(f_T^{(k)}\right)^{-1}(x_0)\right) = D\varphi_T^k(x_0) \cdot Df_T^{(k)}\left(\left(f_T^{(k)}\right)^{-1}(x_0)\right)$$

where the linear map $Df_T^{(k)}(x)$ is invertible for all $x \in (\mathbb{R}^m)^k$ (to this end note that $f_T^{(k)}$ is not only continuous but also differentiable and that $Df_T^{(k)}$ is a block matrix with lower triangle structure where the blocks on the diagonal are invertible, respectively). Hence the matrix on the left hand side has also rank $\dim W$ and $\left(f_T^{(k)}\right)^{-1}(x_0)$ is a smooth point of $F^k(T, \,\cdot\,)$.
Finally, note that $P^k(z, A) = F^k(z, \otimes_{i=1}^k \Gamma)(A)$ where Γ is the distribution of every e_t (cf. (A1)) and we conclude with Theorem 2.2.4. □

2.3.3 Harris Recurrence, Ergodicity and β - Mixing

In this subsection we will prove the promised properties of semi-polynomial Markov chains under the (FL) - condition. First we show irreducibility on the algebraic variety of states.

After that we verify aperiodicity and apply Theorem 1.1.23.

Proposition 2.3.7. *Suppose that (A1) and (A2) hold. Then the Markov chain* $(X_t)_{t\in\mathbb{N}}$ *defined by equation (2.10) is* ψ *- irreducible and aperiodic on the state space* $(W, \mathcal{B}(W))$*. Moreover, the support of* ψ *has non-empty interior.*

Proof.
(1) Due to Proposition 2.3.6 we have for all $A \in \mathcal{B}(W)$

$$\liminf_{\substack{z\to T\\ z\in W}} P^l(z, A) \geq P^l(T, A). \tag{2.12}$$

We define the probability measure ν on the state space $(W, \mathcal{B}(W))$ by

$$\nu(A) := P^l(T, A), \qquad A \in \mathcal{B}(W).$$

Then, for every $A \in \mathcal{B}(W)$ with $\nu(A) \neq 0$, there exists due to (2.12) a neighborhood W_1 of T in W such that

$$\forall z \in W_1 \qquad P^l(z, A) \geq \frac{\nu(A)}{2}. \tag{2.13}$$

(2) Let $K = \{z_1, \dots, z_r\} \subseteq W$ for some $r \in \mathbb{N}^*$. We are going to show that there is a $q \in \mathbb{N}^*$ such that

$$\forall i \in \{1, \dots, r\} \qquad P^q(z_i, W_1) > 0.$$

To this end, consider for $i = 1, \dots, r$ the sequences $(X_t^{z_i})_{t\in\mathbb{N}}$ defined by

$$X_0^{z_i} = z_i \qquad \text{and} \qquad X_t^{z_i} = F(X_{t-1}^{z_i}, a),\ t \geq 1$$

where $a \in \text{int}(E)$ (cf. (A2)).
Due to assumption (A2) there is $q \in \mathbb{N}^*$ such that

$$\forall i \in \{1, \dots, r\} \qquad X_q^{z_i} = F^q(z_i, a, \dots, a) \in W_1.$$

Since $F^q : W \times (\mathbb{R}^m)^q \to W$ is continuous, there exists for every $i \in \{1, \dots, r\}$ a neighborhood U_i of $(z_i, a, \dots, a)$ in $W \times (\mathbb{R}^m)^q$ such that

$$\forall (y, y_1, \dots, y_q) \in U_i \qquad F^q(y, y_1, \dots, y_q) \in W_1.$$

Then, for all $i \in \{1, \dots, r\}$, U_i contains $U_i' \times U^i_{(a,\dots,a)}$ where U_i' and $U^i_{(a,\dots,a)}$ are suitable neighborhoods of z_i in W and $(a, \dots, a)$ in $(\mathbb{R}^m)^q$, respectively.
We define $U_{(a,\dots,a)} := \bigcap_{i=1}^{r} U^i_{(a,\dots,a)}$ which is also a neighborhood of $(a, \dots, a)$ in $(\mathbb{R}^m)^q$ since $U^i_{(a,\dots,a)}$ is a neighborhood of $(a, \dots, a)$ in $(\mathbb{R}^m)^q$ for every $i \in \{1, \dots, r\}$. Then we have in

particular

$$\forall i \in \{1, \ldots, r\} \quad \forall (y_1, \ldots, y_q) \in U_{(a,\ldots,a)} \qquad F^q(z_i, y_1, \ldots, y_q) \in W_1.$$

Since $U_{(a,\ldots,a)}$ contains $U_a \times \ldots \times U_a$ where U_a is an appropriate neighborhood of a in $\mathbb{R}^m$, we deduce for all $i \in \{1, \ldots, r\}$:

$$\begin{aligned} P^q(z_i, W_1) &\geq \mathbb{P}((e_1, \ldots, e_q) \in U_{(a,\ldots,a)}) \\ &\geq \mathbb{P}(e_1 \in U_a)^q = \Gamma(U_a)^q. \end{aligned} \tag{2.14}$$

(3) Let $A \in \mathcal{B}(W)$ with $\nu(A) \neq 0$. W_1 denotes as in (1) the neighborhood of T in W such that $P^l(z, A) \geq \nu(A)/2$ for all $z \in W_1$. Using the Chapman - Kolmogorov equation (cf. [12] Theorem 3.4.2), we obtain for every $i \in \{1, \ldots, r\}$

$$\begin{aligned} P^{q+l}(z_i, A) &= \int_W P^q(z_i, dy) P^l(y, A) \geq \int_{W_1} P^q(z_i, dy) P^l(y, A) \\ &\overset{(2.13)}{\geq} \frac{\nu(A)}{2} \cdot \int_{W_1} P^q(z_i, dy) = \frac{\nu(A)}{2} \cdot P^q(z_i, W_1) \\ &\overset{(2.14)}{\geq} \frac{\Gamma(U_a)^q}{2} \cdot \nu(A). \end{aligned}$$

Due to assumption (A2) U_a contains an open set of E. Thus $\Gamma(U_a) > 0$. This implies that the chain $(X_t)_{t\in\mathbb{N}}$ is ν - irreducible (and thus also ψ - irreducible due to Proposition 1.1.6).

(4) To show aperiodicity, we suppose the chain to be periodic with period d. Due to Theorem 1.1.8 there exist disjoint sets $D_1, \ldots, D_d \in \mathcal{B}(W)$ such that

(i) $\forall i = 1, \ldots, d \quad \forall z \in D_i \qquad P(z, D_{(i \mod d)+1}) = 1$

(ii) $\psi((\cup_{i=1}^d D_i)^c) = 0$.

Since $\psi \succ \nu$ (cf. Proposition 1.1.6 (ii)), $(\cup_{i=1}^d D_i)^c$ is also a ν - null set. Obviously there must be a set D_i with positive ν - measure, let this set be D_1 without loss of generality. Let $x \in D_1$ and $y \in D_d$. For $K := \{x, y\}$ we have just shown in step (3) that

$$P^{q+l}(x, D_1) > 0 \qquad \text{and} \qquad P^{q+l}(y, D_1) > 0$$

for some $q \in \mathbb{N}^*$. Hence, the integers $q + l$ and $q + l - 1$ are divisible by d. Consequently $d = 1$.

(5) We have shown in step (3) that $(X_t)_{t\in\mathbb{N}}$ is $P^l(T, \cdot)$ - irreducible. Since $F^l(T, \cdot)$

has a smooth point in $(\mathbb{R}^m)^l$ (cf. proof of Proposition 2.3.6) Theorem 2.2.1 implies that $P^l(T,\cdot) = F^l(T, \otimes_{i=1}^l \Gamma)$ is absolutely continuous with respect to the measure μ_W. Hence, int(supp $P^l(T,\cdot)) \neq \emptyset$ and we also obtain that int(supp $\psi) \neq \emptyset$. □

We can now state the main result of this chapter:

Theorem 2.3.8. *Suppose (A1) and (A2) to be valid. If in addition the (FL) - condition holds, then the Markov chain* $(X_t)_{t\in\mathbb{N}}$ *defined by equation (2.10) is positive Harris recurrent and geometrically ergodic on the algebraic variety of states* W*. Furthermore, the strictly stationary process* $(X_t)_{t\in\mathbb{Z}}$ *is geometrically* β *- mixing and* $\pi(V) < \infty$*.*

Proof. Due to Proposition 2.3.7 and the assumptions (A1) and (A2), $(X_t)_{t\in\mathbb{N}}$ is ψ - irreducible and aperiodic on $(W, \mathcal{B}(W))$. We conclude by using Theorem 1.1.23. □

Theorem 2.3.9. *Suppose the setting of Theorem 2.3.8 and assume in addition that (A3) holds. Then the strictly stationary process is unique.*

Proof. If there is another P - invariant probability measure π', then $\text{supp}(\pi') \subseteq W$ due to (A3). Since the chain $(X_t)_{t\in\mathbb{N}}$ is recurrent on W (Theorem 2.3.8), it has at most one P - invariant probability measure on $(W, \mathcal{B}(W))$ (cf. [12] Theorem 10.4.4). Therefore the strictly stationary solution is unique. □

Chapter 3

Multivariate GARCH Models

3.1 Introduction and Notations

Multivariate GARCH (generalized autoregressive conditionally heteroskedastic) models have been studied intensively in recent years.
Our focus in this chapter will be the discussion of strict stationarity and geometric ergodicity of these processes as it can be found in [3]. For these purposes we are going to deal with the vec and BEKK models. The latter have been analyzed in detail in [9].
The central section of this chapter presents multivariate GARCH models in an autoregressive manner which will allow us to use the results of Chapter two. To obtain the properties of positivity, Harris recurrence and geometric ergodicity we then verify the key assumptions (A2) and (A3) (cf. Section 2.3.1) and construct a function that fulfills the (FL) - condition (cf. Section 1.1.9).
Regarding notation we denote the set of real $n \times d$ matrices by $M_{n\times d}(\mathbb{R})$, the vector space of real $d \times d$ matrices by $M_d(\mathbb{R})$, the linear subspace of symmetric matrices by $\mathbb{S}_d$, the positive semidefinite cone by $\mathbb{S}_d^+$ and the strictly positive definite matrices by $\mathbb{S}_d^{++}$. For a positive definite and positive semidefinite matrix $A \in \mathbb{S}_d$ we also write $A > 0$ and $A \geq 0$, respectively. The transpose of a matrix $A \in M_{n\times d}(\mathbb{R})$ will be denoted by A^t.
Every matrix $A \in M_{n\times d}(\mathbb{R})$ can be considered a vector in $\mathbb{R}^{nd}$ using the bijective vec transformation which stacks the columns of a matrix below one another beginning with the left one. In the case of symmetric matrices, one often uses the vech transformation which maps $\mathbb{S}_d$ bijectively to $\mathbb{R}^{d(d+1)/2}$ by stacking the lower triangular portion of a matrix. Finally, we denote for two matrices $A = (a_{i,j}) \in M_{n\times d}(\mathbb{R})$ and $B = (b_{k,l}) \in M_{r\times m}(\mathbb{R})$ the tensor (Kronecker) product by $A \otimes B$. $A \otimes B$ is an element of $M_{nr\times dm}(\mathbb{R})$ defined by

$$A \otimes B = \begin{pmatrix} a_{1,1}B & \cdots & a_{1,d}B \\ \vdots & \ddots & \vdots \\ a_{n,1}B & \cdots & a_{n,d}B \end{pmatrix}.$$

3.2 The vec and BEKK Models

The well-known single-dimensional GARCH(p, q) model introduced in Bollerslev [2] generalizes the ARCH(q) model initially presented in Engle [8].
A single-dimensional **GARCH(p, q) process** is defined via an i.i.d. sequence $(\epsilon_n)_{n\in\mathbb{N}^*}$ and the equations

$$X_n = \sqrt{\sigma_n^2}\epsilon_n, \tag{3.1}$$

$$\sigma_n^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i X_{n-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{n-j}^2 \tag{3.2}$$

for $n \in \mathbb{N}^*$. Moreover, the initial values $\sigma_0^2, \sigma_{-1}^2, \ldots, \sigma_{1-p}^2$ and the parameters $\alpha_1, \ldots, \alpha_q$ and $\beta_1, \ldots, \beta_p$ are nonnegative and $\alpha_0 > 0$. The process σ^2 is the latent **conditional variance process** of the GARCH(p, q) process $(X_n)_{n\in\mathbb{N}^*}$.
When one moves from a single-dimensional to a d-dimensional GARCH process, the variance process σ^2 becomes a $d \times d$ covariance matrix process Σ. To specify the process one can use the **vec model** (see [9]) which is given by

$$X_n = \Sigma_n^{1/2}\epsilon_n, \tag{3.3}$$

$$\text{vec}(\Sigma_n) = \text{vec}(C) + \sum_{i=1}^{q} \tilde{A}_i \ \text{vec}(X_{n-i}X_{n-i}^t) + \sum_{j=1}^{p} \tilde{B}_j \ \text{vec}(\Sigma_{n-j}) \tag{3.4}$$

for $n \in \mathbb{N}^*$ where $(\epsilon_n)_{n\in\mathbb{N}^*}$ is now an $\mathbb{R}^d$-valued i.i.d. sequence and $\Sigma_n^{1/2}$ denotes the unique positive semidefinite matrix whose square is Σ_n, i.e. $\Sigma_n^{1/2} \in \mathbb{S}_d^+$ and $\Sigma_n^{1/2}\Sigma_n^{1/2} = \Sigma_n$. To ensure the positive semidefiniteness of the covariance matrix process Σ the initial values $\Sigma_0, \ldots, \Sigma_{1-p}$ and C have to be positive semidefinite and $\tilde{A}_1, \ldots, \tilde{A}_q$ as well as $\tilde{B}_1, \ldots, \tilde{B}_p$ need to be $d^2 \times d^2$ matrices mapping the vectorized positive semidefinite matrices into themselves.
The restriction on the linear operators $\tilde{A}_i$ and $\tilde{B}_j$ necessary to ensure positive semidefiniteness gave rise to the so-called **BEKK model** (see again [9]) which automatically

ensures positive semidefiniteness:

$$X_n = \Sigma_n^{1/2} \epsilon_n, \tag{3.5}$$

$$\Sigma_n = C + \sum_{i=1}^{q} \sum_{k=1}^{l_i} \bar{A}_{i,k} X_{n-i} X_{n-i}^t \bar{A}_{i,k}^t + \sum_{j=1}^{p} \sum_{r=1}^{s_j} \bar{B}_{j,r} \Sigma_{n-j} \bar{B}_{j,r}^t, \tag{3.6}$$

where $\bar{A}_{i,k}$ and $\bar{B}_{j,r}$ are now arbitrary elements of $M_d(\mathbb{R})$.
The BEKK model is equivalent to the vec model with $\tilde{A}_i = \sum_{k=1}^{l_i} \bar{A}_{i,k} \otimes \bar{A}_{i,k},\ i = 1, \dots, q$, and $\tilde{B}_j = \sum_{r=1}^{s_j} \bar{B}_{j,r} \otimes \bar{B}_{j,r},\ j = 1, \dots, p$ (cf. Lemma 3.2.1 (i)).
The question of which vec models are representable in the BEKK form is adressed in [18].
If we take the symmetry of the matrices Σ_n into account, we can write the vec model also in the **vech representation**:

$$X_n = \Sigma_n^{1/2} \epsilon_n, \tag{3.7}$$

$$\text{vech}(\Sigma_n) = \text{vech}(C) + \sum_{i=1}^{q} A_i \ \text{vech}(X_{n-i} X_{n-i}^t) + \sum_{j=1}^{p} B_j \ \text{vech}(\Sigma_{n-j}), \tag{3.8}$$

where $A_i = H_d \tilde{A}_i K_d^t,\ B_j = H_d \tilde{B}_j K_d^t,\ i = 1, \dots, q,\ j = 1, \dots, p$ and the matrices H_d and K_d are given by the following lemma.

Lemma 3.2.1. *Let $A, B, C \in M_d(\mathbb{R})$. Then*

(i) $\text{vec}(ABC) = (C^t \otimes A)\ \text{vec}(B)$

(ii) $(A \otimes B)^t = A^t \otimes B^t$

(iii) there exist unique $K_d, H_d \in M_{\frac{d(d+1)}{2} \times d^2}(\mathbb{R})$ such that

$$\text{vech}(D) = H_d \ \text{vec}(D),\ \ \text{vec}(D) = K_d^t \ \text{vech}(D) \quad \textit{and} \quad H_d K_d^t = \text{Id}_{d(d+1)/2}$$

for every $D \in \mathbb{S}_d$.

Proof. (i) and (ii) are easy to show and (iii) is clear since the vec and vech operator are both linear (cf. also Theorem A.1.3 [15]). □

Remark 3.2.2.
Suppose $(X_n)_{n \in \mathbb{N}^*}$ to be a GARCH(p, q) process in the BEKK representation. We define $\mathcal{F}_n := \sigma(X_{1-q}, X_{2-q}, \dots, X_0, \dots, X_n)$ for $n \in \mathbb{N}$. Moreover, we assume $\mathbb{E}[\epsilon_1] = 0$ and $\mathbb{E}[\epsilon_1 \epsilon_1^t] = \text{Id}_d$ (where Id_d denotes the identity matrix in $M_d(\mathbb{R})$). Finally, ϵ_n is supposed to be independent of $\mathcal{F}_{n-1}$ for every $n \in \mathbb{N}^*$.

It is clear that then the conditional expectation of X_n is $\mathbb{E}[X_n|\mathcal{F}_{n-1}] = 0$ and the conditional covariance matrix of X_n is $\mathbb{E}[X_n X_n^t|\mathcal{F}_{n-1}] = \Sigma_n$ for all $n \in \mathbb{N}^*$. We can now explain one of the advantages of this BEKK representation, namely the stability of the model under invertible linear transformations.
For, let $P \in M_d(\mathbb{R})$ be regular and set $\tilde{X}_n := PX_n$ for all $n \geq 1-q$. Then we obtain:

- $\tilde{\mathcal{F}}_n := \sigma(\tilde{X}_{1-q}, \tilde{X}_{2-q}, \ldots, \tilde{X}_0, \ldots, \tilde{X}_n) = \mathcal{F}_n \quad \forall n \in \mathbb{N}$
- $\mathbb{E}[\tilde{X}_n|\tilde{\mathcal{F}}_{n-1}] = \mathbb{E}[PX_n|\mathcal{F}_{n-1}] = P\mathbb{E}[X_n|\mathcal{F}_{n-1}] = 0 \quad \forall n \in \mathbb{N}^*$
- $\tilde{\Sigma}_n := \mathbb{E}[\tilde{X}_n\tilde{X}_n^t|\tilde{\mathcal{F}}_{n-1}] = \mathbb{E}[PX_nX_n^tP^t|\mathcal{F}_{n-1}] = P\mathbb{E}[X_nX_n^t|\mathcal{F}_{n-1}]P^t = P\Sigma_nP^t$
  $= \tilde{C} + \sum_{i=1}^q \sum_{k=1}^{l_i} \tilde{A}_{i,k}\tilde{X}_{n-i}\tilde{X}_{n-i}^t\tilde{A}_{i,k}^t + \sum_{j=1}^p \sum_{r=1}^{s_j} \tilde{B}_{j,r}\tilde{\Sigma}_{n-j}\tilde{B}_{j,r}^t \quad \forall n \in \mathbb{N}^*$,
  where $\tilde{C} = PCP^t,\ \tilde{A}_{i,k} = PA_{i,k}P^{-1},\ \tilde{B}_{j,r} = PB_jP^{-1}$.

Note that $\tilde{C}$ and the new initial values $\tilde{\Sigma}_0 = P\Sigma_0P^t, \ldots, \tilde{\Sigma}_{1-p} = P\Sigma_{1-p}P^t$ are positive semidefinite. Hence $(\tilde{X}_n)_{n\in\mathbb{N}^*}$ is also a GARCH(p,q) process representable in the BEKK form.

In the following we make some additional assumptions for a GARCH process. More precisely, we define:

Definition 3.2.3. *Suppose* $(X_n)_{n\in\mathbb{N}^*}$ *to be a* GARCH(p,q) *process in the BEKK representation (cf. (3.5), (3.6)) satisfying*

(i) C is positive definite and the initial values $\Sigma_0, \ldots, \Sigma_{1-p}$ are positive semidefinite

(ii) the sequence $(\epsilon_n)_{n\in\mathbb{N}^*}$ is an $\mathbb{R}^d$-valued i.i.d. sequence with distribution Γ, $\mathbb{E}[\epsilon_1] = 0$ and $\mathbb{E}[\epsilon_1\epsilon_1^t] = \mathrm{Id}_d$

(iii) for every $n \in \mathbb{N}^*$, ϵ_n is independent of $\mathcal{F}_{n-1} = \sigma(X_{1-p}, X_{2-p}, \ldots, X_0, \ldots, X_{n-1})$.

We shall call $(X_n)_{n\in\mathbb{N}^*}$ **GARCH**(p,q) ***process of type*** Γ.

Remark 3.2.4.

(i) Supposing C to be positive definite ensures that the conditional covariance matrix Σ_n of X_n is also positive definite for every $n \in \mathbb{N}^*$.

(ii) In the representation $X_n = \Sigma_n^{1/2}\epsilon_n$ one could also choose another transformation $G(\Sigma_n)$ of the conditional covariance matrix Σ_n instead of taking the square root. Boussama [3] for example suggests to take $G(\Sigma_n)$ as a lower triangular matrix. However, the exact form of this transformation does not have any impact on our results concerning stationarity and ergodicity provided that G is an appropriate "smooth" transformation such that we can fit the GARCH process to the setting of Chapter two (cf. Lemma 3.2.5).

Lemma 3.2.5. *For every $\Sigma \in \mathbb{S}_d^{++}$ there is a unique square root $\Sigma^{1/2} \in \mathbb{S}_d^{++}$ such that*

$$\Sigma^{1/2} \cdot \Sigma^{1/2} = \Sigma.$$

Moreover, the application which maps Σ to $\Sigma^{1/2}$ is a C^1 - diffeomorphism on $\mathbb{S}_d^{++}$.

Proof. We define $f_d : \mathbb{S}_d^{++} \to \mathbb{S}_d^{++}$, $X \mapsto X \cdot X$. Since every $X \in \mathbb{S}_d^{++}$ has full rank f_d is well-defined. We will show that f_d is bijective and that, for every $X \in \mathbb{S}_d^{++}$, the differential $df_d(X)$ is a linear homeomorphism.
To this end, we first recall the well-known result that a positive definite matrix has a unique positive definite square root (cf. [10], Theorem 7.2.6). Hence f_d is bijective.
Let $X \in \mathbb{S}_d^{++}$. The differential of f_d at the point X is given by

$$\forall H \in \mathbb{S}_d \qquad df_d(X).H = \frac{\partial}{\partial t} f_d(X + tH)\bigg|_{t=0} = HX + XH.$$

Our aim is to show that $H = 0$ whenever $df_d(X).H = 0$. In fact, this is a simple consequence of Theorem 1 [16] where the solutions X of the general matrix quadratic equation $0 = A + BX + XB^t - XCX$ for fixed $A, B, C \in M_d(\mathbb{R})$ have been analyzed. □

3.3 Stationarity of Multivariate GARCH Models

In this section we present multivariate GARCH models of type Γ in an autoregressive manner. We then specify assumptions such that the key assumptions (A1) - (A3) of Section 2.3.1 are satisfied. Finally, we construct a function for which the (FL) - condition of Section 1.1.9 holds and apply the results of Chapter two to obtain positivity, Harris recurrence, geometric ergodicity and β - mixing of the strictly stationary solution for multivariate GARCH models of type Γ.

3.3.1 Autoregressive Representation

Suppose $(X_n)_{n \in \mathbb{N}^*}$ to be a GARCH(p, q) process of type Γ with values in $\mathbb{R}^d$. Then we can consider its vech representation (cf. (3.8))

$$\mathrm{vech}(\Sigma_n) = \mathrm{vech}(C) + \sum_{i=1}^{q} A_i \; \mathrm{vech}(X_{n-i} X_{n-i}^t) + \sum_{j=1}^{p} B_j \; \mathrm{vech}(\Sigma_{n-j}).$$

We denote by $\mathscr{C}$ the vector in $\left(\mathbb{R}^{d(d+1)/2}\right)^p \times \left(\mathbb{R}^d\right)^q$ defined by

$$\mathscr{C} := (\mathrm{vech}(C)^t, 0, 0, \ldots, 0)^t.$$

With $(X_n)_{n\in\mathbb{N}^*}$ we associate the process $(Y_n)_{n\in\mathbb{N}^*}$ in $\left(\mathbb{R}^{d(d+1)/2}\right)^p \times \left(\mathbb{R}^d\right)^q$ as follows:

$$Y_n := \left(\text{vech}(\Sigma_n)^t, \text{vech}(\Sigma_{n-1})^t, \ldots, \text{vech}(\Sigma_{n-p+1})^t, X_n^t, X_{n-1}^t, \ldots, X_{n-q+1}^t\right)^t.$$

Thus we can write

$$Y_n = \mathscr{C} + \begin{pmatrix} \sum_{i=1}^{q} A_i \text{ vech}(X_{n-i}X_{n-i}^t) + \sum_{j=1}^{p} B_j \text{ vech}(\Sigma_{n-j}) \\ \text{vech}(\Sigma_{n-1}) \\ \vdots \\ \text{vech}(\Sigma_{n-p+1}) \\ X_n \\ X_{n-1} \\ \vdots \\ X_{n-q+1} \end{pmatrix}.$$

One can easily see that Y_n is a regular map of Y_{n-1} and X_n, i.e.

$$Y_n = \varphi(Y_{n-1}, X_n),$$

where $\varphi : \left(\left(\mathbb{R}^{d(d+1)/2}\right)^p \times \left(\mathbb{R}^d\right)^q\right) \times \mathbb{R}^d \to \left(\mathbb{R}^{d(d+1)/2}\right)^p \times \left(\mathbb{R}^d\right)^q$ is a regular map (cf. Definition 1.2.17).

Since $X_n = \Sigma_n^{1/2}\epsilon_n$ (cf. (3.7)) and since the map $G : \mathbb{S}_d^{++} \to \mathbb{S}_d^{++},\ \Sigma \mapsto G(\Sigma) = \Sigma^{1/2}$ is a C^1 - diffeomorphism (cf. Lemma 3.2.5) we obtain that

$$\begin{aligned} f : U \times \mathbb{R}^d &\to \mathbb{R}^d \\ (Y_{n-1}, \epsilon_n) &\mapsto G(\Sigma_n)\epsilon_n = X_n \end{aligned}$$

is a C^1 - map from $U \times \mathbb{R}^d$ into $\mathbb{R}^d$ where U is the open set in $\left(\mathbb{R}^{d(d+1)/2}\right)^p \times \left(\mathbb{R}^d\right)^q$ defined by

$$U := \underbrace{\text{vech}(\mathbb{S}_d^{++}) \times \ldots \times \text{vech}(\mathbb{S}_d^{++})}_{p} \times \underbrace{\mathbb{R}^d \times \ldots \times \mathbb{R}^d}_{q}.$$

Due to the assumption that C is positive definite (note that $(X_n)_{n\in\mathbb{N}^*}$ is assumed to be a GARCH process of type Γ), every Σ_n and $\Sigma_n^{1/2}$ is also positive definite (cf. Remark 3.2.4 (i)) and thus $G(\Sigma_n) = \Sigma_n^{1/2}$ is in particular regular. Hence, for every $Y \in U$, the map $f_Y(\cdot) = f(Y, \cdot)$ is linear bijective from $\mathbb{R}^d$ onto $\mathbb{R}^d$, i.e. $f_Y(\cdot)$ is a C^1 - diffeomorphism. Moreover the map $U \times \mathbb{R}^d \to \mathbb{R}^d,\ (Y, \epsilon) \mapsto f_Y^{-1}(\epsilon)$ is continuous in (Y, ϵ) where $f_Y^{-1}(\cdot)$ denotes the inverse of $f_Y(\cdot)$.

Altogether we are in the same situation as in Chapter two. We shall denote Y_n in its autoregressive representation

$$Y_n = F(Y_{n-1}, \epsilon_n) := \varphi(Y_{n-1}, \underbrace{f_{Y_{n-1}}(\epsilon_n)}_{=X_n})$$

where F is obviously a C^1 - map from $U \times \mathbb{R}^d$ into U.

3.3.2 Some Results from Linear Algebra

In this subsection we will show some results from linear algebra which will be necessary to establish the (FL) - condition for multivariate GARCH models.
Let $n \in \mathbb{N}^*$ and $(F_i)_{1 \le i \le n}$ be a family of elements of $M_d(\mathbb{R})$. We introduce the map

$$\xi : M_d(\mathbb{R}) \to M_d(\mathbb{R})$$
$$\forall M \in M_d(\mathbb{R}) \qquad \xi(M) := \sum_{i=1}^{n} F_i M F_i^t.$$

This map is obviously linear. We can consider ξ a linear map from $\mathbb{R}^{d^2}$ into $\mathbb{R}^{d^2}$ using the vec operator as follows (cf. Lemma 3.2.1 (i)):

$$\mathrm{vec}(\xi(M)) = \left(\sum_{i=1}^{n} F_i \otimes F_i \right) \mathrm{vec}(M).$$

Thus the transformation matrix of ξ is given by

$$F := \sum_{i=1}^{n} F_i \otimes F_i.$$

Note that we have $\xi(\mathbb{S}_d) \subseteq \mathbb{S}_d$, i.e. the symmetric $d \times d$ matrices are mapped into themselves by ξ. We denote $\tilde{\xi}$ the restriction of ξ to the linear subspace $\mathbb{S}_d$. Using again Lemma 3.2.1, we obtain, for all $M \in \mathbb{S}_d$,

$$\begin{aligned} \mathrm{vech}(\tilde{\xi}(M)) &= \mathrm{vech}(\xi(M)) \\ &= H_d \ \mathrm{vec}(\xi(M)) \\ &= H_d F \ \mathrm{vec}(M) \\ &= H_d F K_d^t \ \mathrm{vech}(M). \end{aligned}$$

Since we can identify $\mathbb{S}_d$ via the vech operator with $\mathbb{R}^{d(d+1)/2}$, the transformation matrix of $\tilde{\xi}$ is given by

$$\tilde{F} := H_d F K_d^t.$$

We obtain the following lemma:

Lemma 3.3.1. *Let $C \in \mathbb{S}_d^{++}$. The following statements are equivalent:*

(i) The spectral radius of ξ is less than 1.

(ii) The spectral radius of $\tilde{\xi}$ is less than 1.

(iii) There is $\Sigma \in \mathbb{S}_d^{++}$ such that $\Sigma = C + \xi(\Sigma)$.

Proof.
(i) $\Rightarrow$ (ii): Obvious since $\tilde{\xi}$ is a restriction of ξ.
(ii) $\Rightarrow$ (iii): If the spectral radius of $\tilde{\xi}$ is less than 1, then the Neumann series $\sum_{n=0}^{\infty} \tilde{\xi}^n$ is convergent with respect to a suitable operator norm. We define

$$\Sigma := \sum_{n=0}^{\infty} \tilde{\xi}^n(C). \tag{3.9}$$

Clearly, Σ is symmetric. Moreover, for all $M \in \mathbb{S}_d^{+}$, we have $\tilde{\xi}(M) \in \mathbb{S}_d^{+}$ by the definition of $\tilde{\xi}$ and ξ, respectively. By recurrence we obtain that $\tilde{\xi}^n(M) \in \mathbb{S}_d^{+}$ for all $n \in \mathbb{N}^*$. Thus the matrix $\Sigma - \tilde{\xi}^0(C) = \Sigma - C$ is symmetric and positive semidefinite. This implies that Σ is positive definite.
Due to the definition of Σ (cf. (3.9)) and to the fact that ξ and $\tilde{\xi}$ coincide on $\mathbb{S}_d$, we deduce:

$$\begin{aligned}\Sigma &= \sum_{n=0}^{\infty} \xi^n(C) \\ &= C + \xi(\sum_{n=1}^{\infty} \xi^{n-1}(C)) \\ &= C + \xi(\Sigma).\end{aligned}$$

(iii) $\Rightarrow$ (i): Suppose that there exists $\Sigma \in \mathbb{S}_d^{++}$ such that $\Sigma = C + \xi(\Sigma)$. We denote the complex $d \times d$ matrices by $M_d(\mathbb{C})$ and the conjugate transpose of a vector $x \in \mathbb{C}^d$ by x^*. For every $P \in M_d(\mathbb{C})$ we define

$$\|P\|_\Sigma := \sup_{x \in \mathbb{C}^d,\ x^* \Sigma\, x = 1} |x^* P x|$$

which is a norm on $M_d(\mathbb{C})$ since $\Sigma \in \mathbb{S}_d^{++}$ (can be verified easily by checking the properties of a norm).
Then, for all $x \in \mathbb{C}^d$,

$$|x^* P x| \le \|P\|_\Sigma \,(x^* \Sigma x).$$

Since the unit sphere $\left\{x \in \mathbb{C}^d : \; x^*\Sigma x = 1\right\}$ is compact, there exists, for every $P \in M_d(\mathbb{C})$, a vector $x_p \in \mathbb{C}^d$ such that

$$\|P\|_\Sigma = |x_P^* P x_P| \quad \text{with} \quad x_P^* \Sigma x_P = 1.$$

Let now λ be an eigenvalue of ξ. Then there is an $M \in M_d(\mathbb{C}),\ M \neq 0$ such that

$$\lambda M = \xi(M) = \sum_{i=1}^n F_i M F_i^t.$$

For every $x \in \mathbb{C}^d$, we deduce

$$\begin{aligned}
|\lambda| \cdot |x^* M x| &= \left| \sum_{i=1}^n x^* F_i M F_i^t x \right| \\
&\leq \sum_{i=1}^n \left| \left(F_i^t x\right)^* M \left(F_i^t x\right) \right| \\
&\leq \|M\|_\Sigma \sum_{i=1}^n x^* F_i \Sigma F_i^t x \\
&= \|M\|_\Sigma \, x^* \underbrace{\left(\sum_{i=1}^n F_i \Sigma F_i^t \right)}_{=\xi(\Sigma)=\Sigma - C} x.
\end{aligned}$$

If we choose x_M such that $\|M\|_\Sigma = |x_M^* M x_M|$ and $x_M^* \Sigma x_M = 1$, we obtain (note that $\|M\|_\Sigma \neq 0$)

$$|\lambda| \leq 1 - x_M^* C x_M < 1.$$

Hence the spectral radius of ξ is less than 1. □

We consider now the families of matrices $(\bar{A}_{i,k}, \bar{B}_{j,r})$ and (A_i, B_j) which occur in the BEKK and vech representation of a GARCH(p, q) model, respectively.

Proposition 3.3.2. *The spectral radius of* $(\sum_{i=1}^q A_i + \sum_{j=1}^p B_j)$ *is less than* 1 *if and only if there exists* $\Sigma \in \mathbb{S}_d^{++}$ *such that*

$$\Sigma = C + \sum_{i=1}^q \sum_{k=1}^{l_i} \bar{A}_{i,k} \Sigma \bar{A}_{i,k}^t + \sum_{j=1}^p \sum_{r=1}^{s_j} \bar{B}_{j,r} \Sigma \bar{B}_{j,r}^t.$$

Proof. We introduce the linear map $\xi : M_d(\mathbb{R}) \to M_d(\mathbb{R})$ given by

$$\forall M \in M_d(\mathbb{R}) \qquad \xi(M) = \sum_{i=1}^{q}\sum_{k=1}^{l_i} \bar{A}_{i,k} M \bar{A}_{i,k}^t + \sum_{j=1}^{p}\sum_{r=1}^{s_j} \bar{B}_{j,r} M \bar{B}_{j,r}^t.$$

Then the transformation matrix of ξ is

$$F = \sum_{i=1}^{q}\sum_{k=1}^{l_i} \bar{A}_{i,k} \otimes \bar{A}_{i,k} + \sum_{j=1}^{p}\sum_{r=1}^{s_j} \bar{B}_{j,r} \otimes \bar{B}_{j,r} = \sum_{i=1}^{q} \tilde{A}_i + \sum_{j=1}^{p} \tilde{B}_j.$$

Note that the transformation matrix of $\tilde{\xi}$ (restriction of ξ to the linear subspace $\mathbb{S}_d$) is $H_d F K_d^t = \sum_{i=1}^q A_i + \sum_{j=1}^p B_j$.
Due to Lemma 3.3.1 the spectral radius of $(\sum_{i=1}^q A_i + \sum_{j=1}^p B_j)$ is less than 1 if and only if there is $\Sigma \in \mathbb{S}_d^{++}$ such that

$$\Sigma = C + \xi(\Sigma) = C + \sum_{i=1}^{q}\sum_{k=1}^{l_i} \bar{A}_{i,k} \Sigma \bar{A}_{i,k}^t + \sum_{j=1}^{p}\sum_{r=1}^{s_j} \bar{B}_{j,r} \Sigma \bar{B}_{j,r}^t.$$

□

Remark 3.3.3. One can equivalently state that the spectral radius of $(\sum_{i=1}^q A_i + \sum_{j=1}^p B_j)$ is less than 1 if and only if there exists $\Sigma \in \mathbb{S}_d^{++}$ such that

$$\Sigma = C + \sum_{i=1}^{q}\sum_{k=1}^{l_i} \bar{A}_{i,k}^t \Sigma \bar{A}_{i,k} + \sum_{j=1}^{p}\sum_{r=1}^{s_j} \bar{B}_{j,r}^t \Sigma \bar{B}_{j,r}$$

which we are going to use in the upcoming proof of Theorem 3.3.8.
Indeed, the transpose of F, where F is the transformation matrix of ξ (see proof above), is given by

$$F^t = \sum_{i=1}^{q}\sum_{k=1}^{l_i} \bar{A}_{i,k}^t \otimes \bar{A}_{i,k}^t + \sum_{j=1}^{p}\sum_{r=1}^{s_j} \bar{B}_{j,r}^t \otimes \bar{B}_{j,r}^t$$

(cf. Lemma 3.2.1 (ii)).
Then the spectral radius of $(\sum_{i=1}^q A_i + \sum_{j=1}^p B_j)$ is less than 1 if and only if the spectral radius of F is less than 1 (cf. Lemma 3.3.1). Since the spectral radiuses of F and F^t are equal we can use again Lemma 3.3.1 with the linear map ξ' where ξ' possesses the transformation matrix F^t.

In the following we consider the block matrix B defined by

$$B := \begin{pmatrix} B_1 & B_2 & \dots & B_{p-1} & B_p \\ I & 0 & \dots & 0 & 0 \\ 0 & I & \ddots & \vdots & 0 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & I & 0 \end{pmatrix} \in M_{p\frac{d(d+1)}{2}}(\mathbb{R})$$

where I denotes the identity matrix in $M_{\frac{d(d+1)}{2}}(\mathbb{R})$.

Proposition 3.3.4.

(i) If the spectral radius of the matrix $\sum_{j=1}^p B_j$ is less than 1, then the one of B is also less than 1.

(ii) If the spectral radius of the matrix $(\sum_{i=1}^q A_i + \sum_{j=1}^p B_j)$ is less than 1, then the one of $\sum_{j=1}^p B_j$ is also less than 1.

Proof.
(i) Suppose that the spectral radius of $\sum_{j=1}^p B_j$ is less than 1. Then there exists due to Lemma 3.3.1 a symmetric positive definite matrix $\tilde{\Sigma} \in \mathbb{S}_d^{++}$ such that

$$\tilde{\Sigma} = C + \sum_{j=1}^{p} \sum_{r=1}^{s_j} \bar{B}_{j,r} \tilde{\Sigma} \bar{B}_{j,r}^t. \tag{3.10}$$

Let λ be an eigenvalue of B associated with the eigenvector $h = (h_1^t, \dots, h_p^t)^t \in \left(\mathbb{R}^{\frac{d(d+1)}{2}}\right)^p$. Then

$$\lambda h_1 = \sum_{j=1}^{p} B_j h_j \qquad \text{and} \qquad \lambda h_j = h_{j-1} \ \text{ for } 2 \le j \le p.$$

Thus $h_p \neq 0$ (otherwise h would be zero) and

$$\lambda^p h_p = \lambda(\lambda^{p-1} h_p) = \lambda h_1 = \sum_{j=1}^{p} B_j h_j = \sum_{j=1}^{p} \lambda^{p-j} B_j h_p.$$

Let $M \in \mathbb{S}_d$ such that $\operatorname{vech}(M) = h_p$. Then

$$\lambda^p M = \sum_{j=1}^{p} \sum_{r=1}^{s_j} \lambda^{p-j} \bar{B}_{j,r} M \bar{B}_{j,r}^t.$$

We define the norm $\|\cdot\|_{\tilde{\Sigma}}$ on $M_d(\mathbb{C})$ as in the proof of Lemma 3.3.1 by

$$\|P\|_{\tilde{\Sigma}} := \sup_{x\in\mathbb{C}^d,\ x^*\tilde{\Sigma}\,x=1} |x^*Px|, \quad P \in M_d(\mathbb{C}).$$

Then, for all $x \in \mathbb{C}^d$,

$$\begin{aligned}|\lambda|^p \cdot |x^*Mx| &= \left|\sum_{j=1}^{p}\sum_{r=1}^{s_j} \lambda^{p-j} x^* \bar{B}_{j,r} M \bar{B}_{j,r}^t x\right| \\ &\leq \sum_{j=1}^{p}\sum_{r=1}^{s_j} |\lambda|^{p-j} \left|x^* \bar{B}_{j,r} M \bar{B}_{j,r}^t x\right| \\ &\leq \|M\|_{\tilde{\Sigma}} \sum_{j=1}^{p}\sum_{r=1}^{s_j} |\lambda|^{p-j} \,(x^* \bar{B}_{j,r} \tilde{\Sigma} \bar{B}_{j,r}^t x).\end{aligned}$$

If we assume that there is an eigenvalue λ of B with $|\lambda| \geq 1$, then we obtain, taking the vector x such that $x^*\tilde{\Sigma}x = 1$ and $|x^*Mx| = \|M\|_{\tilde{\Sigma}}$ and using (3.10), that

$$\begin{aligned}|\lambda|^p &\leq \sum_{j=1}^{p}\sum_{r=1}^{s_j} |\lambda|^{p-j} \,(x^* \bar{B}_{j,r} \tilde{\Sigma} \bar{B}_{j,r}^t x) \\ &\leq |\lambda|^{p-1} \left[x^* \left(\sum_{j=1}^{p}\sum_{r=1}^{s_j} \bar{B}_{j,r} \tilde{\Sigma} \bar{B}_{j,r}^t\right) x\right] \\ &= |\lambda|^{p-1} \left[x^* \left(\tilde{\Sigma} - C\right) x\right] = |\lambda|^{p-1}\,(1 - x^*Cx).\end{aligned}$$

Since C is symmetric positive definite, one has $x^*Cx > 0$. Hence $|\lambda|^p < |\lambda|^{p-1}$, i.e. $|\lambda| < 1$ which is a contradiction.
Thus the spectral radius of B has to be less than 1.
(ii) Suppose that the spectral radius of the matrix $(\sum_{i=1}^{q} A_i + \sum_{j=1}^{p} B_j)$ is less than 1. Then, due to Proposition 3.3.2, there exists $\Sigma \in \mathbb{S}_d^{++}$ such that

$$\Sigma = C + \sum_{i=1}^{q}\sum_{k=1}^{l_i} \bar{A}_{i,k} \Sigma \bar{A}_{i,k}^t + \sum_{j=1}^{p}\sum_{r=1}^{s_j} \bar{B}_{j,r} \Sigma \bar{B}_{j,r}^t.$$

We set $\tilde{C} := C + \sum_{i=1}^{q}\sum_{k=1}^{l_i} \bar{A}_{i,k} \Sigma \bar{A}_{i,k}^t$. Clearly, $\tilde{C}$ is symmetric positive definite and

$$\Sigma = \tilde{C} + \sum_{j=1}^{p}\sum_{r=1}^{s_j} \bar{B}_{j,r} \Sigma \bar{B}_{j,r}^t.$$

Using again Proposition 3.3.2 we deduce that the spectral radius of $\sum_{j=1}^{p} B_j$ is less than

1. □

Remark 3.3.5. The matrix $\tilde{\Sigma}$ in (3.10) is the limit of a Neumann series (cf. proof of Lemma 3.3.1). With some background of functional analysis one may show that

$$\mathrm{vech}(\tilde{\Sigma}) = \left(I - \sum_{j=1}^{p} B_j \right)^{-1} \mathrm{vech}(C)$$

where I is again the identity matrix of $M_{\frac{d(d+1)}{2}}(\mathbb{R})$.

After this preparatory work we will now prove stationarity for multivariate GARCH processes of type Γ. We are going to specify assumptions such that we can then show that (A1) - (A3), necessary for Theorem 2.3.8 and Theorem 2.3.9, hold.

3.3.3 Verification of Assumption (A2)

We are going to consider the following two assumptions:

(H1) The distribution of every ϵ_n is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^d$. By E we denote the domain of positivity of its density γ.

Note that (H1) coincides with the assumption (A1).

(H2) The point zero is in the interior of E.

We will suppose throughout that (H1) holds.

Proposition 3.3.6. *Suppose that (H2) holds. If the spectral radius of the matrix* $\sum_{j=1}^{p} B_j$ *is less than* 1, *then (A2) (cf. Section 2.3.1) holds.*

Proof. Let U be the open set in $\left(\mathbb{R}^{d(d+1)/2}\right)^p \times \left(\mathbb{R}^d\right)^q$ defined as in Section 3.3.1. For arbitrary $y \in U$ we define the sequence $(Y_n^y)_{n\in\mathbb{N}}$ by

$$Y_0^y = y \qquad \text{and} \qquad Y_n^y = F(Y_{n-1}^y, 0),\ n \geq 1$$

where F is the C^1 - map from $U \times \mathbb{R}^d$ into U introduced at the end of Section 3.3.1.
We denote by X_n^y and $\mathrm{vech}(\Sigma_n^y)$ the associated values of X_n and $\mathrm{vech}(\Sigma_n)$. Since, by definition, $X_n = G(\Sigma_n)\epsilon_n = \Sigma_n^{1/2}\epsilon_n$, we obtain that $X_n^y = 0$ for all $n \geq 1$. Due to (3.8), Σ_n^y can be written, for every $n > q$, as

$$\mathrm{vech}(\Sigma_n^y) = \mathrm{vech}(C) + \sum_{j=1}^{p} B_j\ \mathrm{vech}(\Sigma_{n-j}^y).$$

Thus, for all $n > q$ and for all $y \in U$,

$$Y_n^y = \mathscr{C} + \tilde{B} Y_{n-1}^y \tag{3.11}$$

where $\tilde{B}$ is the matrix defined by

$$\tilde{B} := \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix} \in M_{p\frac{d(d+1)}{2}+qd}(\mathbb{R})$$

with B the block matrix as in Section 3.3.2

$$B = \begin{pmatrix} B_1 & B_2 & \dots & B_{p-1} & B_p \\ I & 0 & \dots & 0 & 0 \\ 0 & I & \ddots & \vdots & 0 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & I & 0 \end{pmatrix} \in M_{p\frac{d(d+1)}{2}}(\mathbb{R}).$$

Due to (3.11), the assumption (A2) is satisfied with $a = 0$ if the spectral radius of B is less than 1. This is the case, since the spectral radius of $\sum\limits_{j=1}^{p} B_j$ is supposed to be less than 1 (cf. Proposition 3.3.4 (i)). □

Hence, for all $y \in U$, the sequence $(Y_n^y)_{n\in\mathbb{N}}$ converges to the unique fixed point T defined by

$$T = \mathscr{C} + \tilde{B} T. \tag{3.12}$$

Using Lemma 3.3.1 and the fact that the spectral radius of $\sum_{j=1}^{p} B_j$ is assumed to be less than 1, there is $\tilde{\Sigma} \in \mathbb{S}_d^{++}$ (cf. (3.10)) such that

$$\tilde{\Sigma} = C + \sum_{j=1}^{p} \sum_{r=1}^{s_j} \bar{B}_{j,r} \tilde{\Sigma} \bar{B}_{j,r}^t.$$

It is then easy to see that T can be written as

$$T = \left(\underbrace{\mathrm{vech}(\tilde{\Sigma})^t, \dots, \mathrm{vech}(\tilde{\Sigma})^t}_{p}, \underbrace{0, \dots, 0}_{qd} \right)^t.$$

We set $\mathscr{C}_1 := (\mathrm{vech}(C)^t, 0, \dots, 0)^t \in \left(\mathbb{R}^{d(d+1)/2}\right)^p$. Then (3.12) yields

$$\sigma = \mathscr{C}_1 + B\sigma \tag{3.13}$$

where $\sigma := \left(\mathrm{vech}(\tilde{\Sigma})^t, \ldots, \mathrm{vech}(\tilde{\Sigma})^t\right)^t \in \left(\mathbb{R}^{d(d+1)/2}\right)^p$.

3.3.4 Verification of Assumption (A3)

In this subsection we will study the algebraic variety of states of the process $(Y_n)_{n\in\mathbb{N}^*}$. In particular, we can derive from a formula shown in Theorem 3.3.7 that all strictly stationary solutions, if they exist at all, have their values in the algebraic variety of states, i.e. we show that (A3) holds provided that (H1) and (H2) are satisfied.
As introduced in Section 2.3.1 the algebraic variety of states is defined as the Zariski closure of the orbit S_T. If (H2) is satisfied, then E contains an open set of $\mathbb{R}^d$ and we obtain with the same arguments as in Section 2.3.2 that

$$\begin{aligned} W = {}^Z\overline{S_T} &= {}^Z\overline{\bigcup_{n\in\mathbb{N}^*} F^n\left(T, E^n\right)} \\ &= {}^Z\overline{\bigcup_{n\in\mathbb{N}^*} F^n\left(T, (\mathbb{R}^d)^n\right)} \\ &= {}^Z\overline{\bigcup_{n\in\mathbb{N}^*} \varphi^n\left(T, (\mathbb{R}^d)^n\right)}. \end{aligned}$$

Let $n \in \mathbb{N}^*$ and consider $y(n) \in \varphi^n\left(T, (\mathbb{R}^d)^n\right)$ given by

$$y(n) = \varphi^n(T, x_1, \ldots, x_n)$$

where $x_1, \ldots, x_n \in \mathbb{R}^d$.
We define $x(n)$ and $\sigma(n)$ by the coordinates of $y(n)$ as follows:

$$\begin{aligned} & x(n) = \left(\mathrm{vech}(x_n x_n^t)^t, \ldots, \mathrm{vech}(x_{n-q+1}x_{n-q+1}^t)^t\right)^t \\ \text{and} \quad & \sigma(n) = \left(\mathrm{vech}(\sigma_n)^t, \ldots, \mathrm{vech}(\sigma_{n-p+1})^t\right)^t. \end{aligned}$$

That is, $y(n) = \left(\sigma(n)^t, x_n^t, \ldots, x_{n-q+1}^t\right)^t$. Then

$$\sigma(n+1) = \mathscr{C}_1 + Ax(n) + B\sigma(n) \tag{3.14}$$

where $\mathscr{C}_1$ and B are defined in Section 3.3.3 and A is given by

$$A := \begin{pmatrix} A_1 & A_2 & \ldots & A_q \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 \end{pmatrix} \in M_{p\frac{d(d+1)}{2} \times q\frac{d(d+1)}{2}}(\mathbb{R}).$$

Iterating (3.14) and due to $\sigma(0) = \sigma = \left(\mathrm{vech}(\tilde{\Sigma})^t, \ldots, \mathrm{vech}(\tilde{\Sigma})^t\right)^t$ (since $y(0) = T$) we deduce

$$\begin{aligned}\sigma(n) &= \underbrace{\sum_{i=0}^{n-1} B^i \mathscr{C}_1 + B^n \sigma}_{\overset{(3.13)}{=}\sigma} + \sum_{i=1}^{n-1} B^{i-1} A x(n-i) \\ &= \sigma + \sum_{i=1}^{n-1} B^{i-1} A x(n-i).\end{aligned}$$

This yields

$$\mathrm{vech}(\sigma_n) = \mathrm{vech}(\tilde{\Sigma}) + \sum_{i=1}^{n-1} K_i \mathrm{vech}(x_{n-i} x_{n-i}^t)$$

where, for all $i \in \mathbb{N}^*$, K_i is defined by

$$K_i := \left[B^{i-1}A\right]_{1,1} + \left[B^{i-2}A\right]_{1,2} + \ldots + \left[B^{i-q}A\right]_{1,q}$$

with convention $B^0 := I$ (where I is the identity matrix in $M_{pd(d+1)/2}(\mathbb{R})$), $B^i := 0$ if $i < 0$ and $[M]_{1,j}$ is the $d(d+1)/2 \times d(d+1)/2$ block from lines 1 to $d(d+1)/2$ and from columns $(j-1)d(d+1)/2$ to $jd(d+1)/2$ of M.

Thus, W is the Zariski closure of the orbit

$$\begin{aligned}S_T &= \bigcup_{n \in \mathbb{N}^*} \left\{y(n) : \ x_1, \ldots, x_n \in \mathbb{R}^d\right\} \\ &= \bigcup_{n \in \mathbb{N}^*} \Bigg\{ T + \Bigg(\sum_{i=1}^{n-1} \left(K_i \mathrm{vech}(x_{n-i} x_{n-i}^t)\right)^t, \ldots, \sum_{i=1}^{n-p} \left(K_i \mathrm{vech}(x_{n-p+1-i} x_{n-p+1-i}^t)\right)^t, \\ &\qquad x_n^t, x_{n-1}^t, \ldots, x_{n-q+1}^t \Bigg)^t \ : \ x_1, \ldots, x_n \in \mathbb{R}^d \Bigg\}\end{aligned}$$

where $x_{1-q} = x_{2-q} = \ldots = x_0 = 0$ (since $y(0) = T = (\sigma^t, 0, \ldots, 0)^t$).

In particular, this implies $\varphi(W \times \mathbb{R}^d) \subseteq W$ (cf. (2.1)). For, $\varphi(S_T \times \mathbb{R}^d) \subseteq S_T$ yields

$$\varphi(W \times \mathbb{R}^d) = \varphi\left(\overline{{}^Z S_T} \times \mathbb{R}^d\right) \subseteq \overline{{}^Z \varphi\left(\overline{{}^Z S_T} \times \mathbb{R}^d\right)} = \overline{{}^Z \varphi(S_T \times \mathbb{R}^d)} \subseteq \overline{{}^Z S_T} = W$$

since φ is a regular map and thus continuous with respect to the Zariski topology.

Theorem 3.3.7. *Suppose that (H2) holds and that there is a strictly stationary solution $(X_n)_{n\in\mathbb{Z}}$ for the GARCH(p, q) model of type Γ, then the process $(Y_n)_{n\in\mathbb{Z}}$ takes its values in the algebraic variety of states W. Moreover, one has*

$$\mathrm{vech}(\Sigma_n) = \left(I - \sum_{j=1}^{p} B_j\right)^{-1} \mathrm{vech}(C) + \sum_{i=1}^{\infty} K_i \ \mathrm{vech}(X_{n-i}X_{n-i}^t)$$

where I denotes the identity matrix in $M_{\frac{d(d+1)}{2}}(\mathbb{R})$.

Proof. Let $(X_n)_{n\in\mathbb{Z}}$ be a strictly stationary solution of the GARCH(p,q) model of type Γ with conditional covariance matrices Σ_n. We denote by $X(n)$ and $\Sigma(n)$ the following random vectors:

$$X(n) = \left(\mathrm{vech}(X_nX_n^t)^t, \ldots, \mathrm{vech}(X_{n-q+1}X_{n-q+1}^t)^t\right)^t$$

$$\text{and} \qquad \Sigma(n) = \left(\mathrm{vech}(\Sigma_n)^t, \ldots, \mathrm{vech}(\Sigma_{n-p+1})^t\right)^t.$$

Since $\Sigma(n) = \mathscr{C}_1 + AX(n-1) + B\Sigma(n-1)$ (cf. (3.14)), iterating yields

$$\Sigma(n) = \sum_{i=0}^{k-1} B^i\mathscr{C}_1 + B^k\Sigma(n-k) + \sum_{i=1}^{k} B^{i-1}AX(n-i) \tag{3.15}$$

for all $k \in \mathbb{N}$.
For any vectors $M = (\mathrm{vech}(M_1)^t, \ldots, \mathrm{vech}(M_p)^t)^t$ and $N = (\mathrm{vech}(N_1)^t, \ldots, \mathrm{vech}(N_p)^t)^t$ in $(\mathbb{R}^{d(d+1)/2})^p$, let us denote $M \geq N$ if and only if $M_1 \geq N_1, \ldots, M_p \geq N_p$ (where, for all $M_i, N_i \in \mathbb{S}_d$, $M_i \geq N_i \Leftrightarrow M_i - N_i \geq 0 \Leftrightarrow M_i - N_i$ positive semidefinite (cf. Introduction Chapter three)). This defines a partial order on $(\mathbb{R}^{d(d+1)/2})^p$.
Then (3.15) yields $\Sigma(n) \geq \sum_{i=0}^{k-1} B^i\mathscr{C}_1$. Since $\Sigma(n)$ is finite the series $\sum_{i=0}^{k-1} B^i\mathscr{C}_1$ converges as $k \to \infty$ (see for instance [21] for further details concerning partially ordered topological spaces; in particular the Corollary after Lemma 5 proves that our series must converge). Setting $\tilde{\sigma} := \sum_{i=0}^{\infty} B^i\mathscr{C}_1$, it is easy to see that $\tilde{\sigma} = \mathscr{C}_1 + B\tilde{\sigma}$. Using the definition of B and $\mathscr{C}_1$, respectively, we obtain that $\tilde{\sigma} = (\sigma_1^t, \sigma_1^t, \ldots, \sigma_1^t)^t$ for some $\sigma_1 \in \mathbb{R}^{d(d+1)/2}$ which fulfills $\sigma_1 = \mathrm{vech}(C) + \sum_{j=1}^{p} B_j\sigma_1$. One may then verify that $\sigma_1 = \mathrm{vech}(\Sigma_1)$ for some $\Sigma_1 \in \mathbb{S}_d^{++}$ and hence that the spectral radius of $\sum_{j=1}^{p} B_j$ is less than 1 (cf. Proposition 3.3.2). Due to Proposition 3.3.4 (i) we obtain that the spectral radius of B is also less than 1. Thus $\tilde{\sigma} = \sigma$.
Next, since the spectral radius of B is less than 1, the sequence $(B^k)_{k\in\mathbb{N}}$ converges to zero as $k \to \infty$. The random vectors $(\Sigma(n-k))_{k\in\mathbb{N}}$ have a constant law because $(X_n)_{n\in\mathbb{Z}}$ is supposed to be a strictly stationary solution of the GARCH model. Thus $B^k\Sigma(n-k)$ converges to zero in probability when $k \to \infty$.

With an analog argument as for $\sum_{i=0}^{k-1} B^i \mathscr{C}_1$ one can see that $\sum_{i=1}^{k} B^{i-1}AX(n-i)$ converges almost surely as $k \to \infty$. Hence, taking the limit of (3.15) yields

$$\Sigma(n) = \sigma + \sum_{i=1}^{\infty} B^{i-1}AX(n-i) \qquad \text{a.s.}$$

Using the matrices K_i, defined during the investigation of the variety of states W, we obtain

$$\mathrm{vech}(\Sigma_n) = \mathrm{vech}(\tilde{\Sigma}) + \sum_{i=1}^{\infty} K_i \ \mathrm{vech}(X_{n-i}X_{n-i}^t) \qquad \text{a.s.}$$

This shows that $(Y_n)_{n\in\mathbb{Z}}$ takes its values in the variety W. Note that the strictly stationary solution is causal. To finish the proof we refer to Remark 3.3.5 from which we obtain $\mathrm{vech}(\tilde{\Sigma}) = (I - \sum_{j=1}^{p} B_j)^{-1} \ \mathrm{vech}(C)$. □

3.3.5 Foster - Lyapounov Condition (FL)

We will derive a function V satisfying the (FL) - condition of Section 1.1.9 for multivariate GARCH processes of type Γ provided that the spectral radius of $(\sum_{i=1}^{q} A_i + \sum_{j=1}^{p} B_j)$ is less than 1. That is, we prove the following theorem:

Theorem 3.3.8. *Suppose that the spectral radius of the matrix* $(\sum_{i=1}^{q} A_i + \sum_{j=1}^{p} B_j)$ *is less than* 1. *Then there exist a function* $V \geq 1$ *and positive constants* $\alpha < 1,\ b < \infty$ *as well as a compact set* K *such that the (FL) - condition is satisfied on* K.

Proof. For convenience we suppose without loss of generality that in the BEKK representation (3.6) $l_i = s_j = 1$ for all $i = 1, \ldots, q$ and $j = 1, \ldots, p$. We shall then denote the matrices by $\bar{A}_i$ and $\bar{B}_j$ instead of $\bar{A}_{i,1}$ and $\bar{B}_{j,1}$. That is, we have

$$\Sigma_n = C + \sum_{i=1}^{q} \bar{A}_i X_{n-i} X_{n-i}^t \bar{A}_i^t + \sum_{j=1}^{p} \bar{B}_j \Sigma_{n-j} \bar{B}_j^t. \tag{3.16}$$

If the spectral radius of the matrix $(\sum_{i=1}^{q} A_i + \sum_{j=1}^{p} B_j)$ is less than 1, then, due to Proposition 3.3.2 and Remark 3.3.3, there exists $\Sigma \in \mathbb{S}_d^{++}$ such that

$$\Sigma = C + \sum_{i=1}^{q} \bar{A}_i^t \Sigma \bar{A}_i + \sum_{j=1}^{p} \bar{B}_j^t \Sigma \bar{B}_j.$$

We define the map $V : U \to [1, \infty)$ by

$$V(Y_n) := \mathrm{tr}(V_1 \Sigma_n) + \ldots + \mathrm{tr}(V_p \Sigma_{n-p+1}) + X_n^t V_{p+1} X_n + \ldots + X_{n-q+1}^t V_{p+q} X_{n-q+1} + 1$$

where $\mathrm{tr}(\cdot)$ denotes the trace of a matrix and the $d \times d$ matrices $(V_i)_{1\le i\le p+q}$ are given by

$$V_k := \frac{p-k+1}{p+q}C + \sum_{j=k}^{p} \bar{B}_j^t \Sigma \bar{B}_j, \quad 1 \le k \le p$$
$$V_{p+k} := \frac{q-k+1}{p+q}C + \sum_{i=k}^{q} \bar{A}_i^t \Sigma \bar{A}_i, \quad 1 \le k \le q.$$

With $y = \left(\mathrm{vech}(\Sigma_{n-1})^t, \ldots, \mathrm{vech}(\Sigma_{n-p})^t, X_{n-1}^t, \ldots, X_{n-q}^t\right)^t \in U$ we obtain

$$\begin{aligned}
&\mathbb{E}[V(Y_n)|Y_{n-1} = y] \\
&\quad = \mathbb{E}\left[\mathrm{tr}(V_1\Sigma_n) + X_n^t V_{p+1} X_n | Y_{n-1} = y\right] + \mathrm{tr}(V_2\Sigma_{n-1}) + \ldots + \mathrm{tr}(V_p\Sigma_{n-p+1}) \\
&\qquad + X_{n-1}^t V_{p+2} X_{n-1} + \ldots + X_{n-q+1}^t V_{p+q} X_{n-q+1} + 1. \qquad (3.17)
\end{aligned}$$

Using (3.16) for Σ_n, we deduce for the first term at the right hand side

$$\begin{aligned}
&\mathbb{E}\left[\mathrm{tr}(V_1\Sigma_n) + X_n^t V_{p+1} X_n | Y_{n-1} = y\right] \\
&\quad = \mathbb{E}\left[X_n^t V_{p+1} X_n | Y_{n-1} = y\right] + \mathrm{tr}(V_1 C) \\
&\qquad + \mathrm{tr}(V_1 \bar{A}_1 X_{n-1} X_{n-1}^t \bar{A}_1^t) + \ldots + \mathrm{tr}(V_1 \bar{A}_q X_{n-q} X_{n-q}^t \bar{A}_q^t) \\
&\qquad + \mathrm{tr}(V_1 \bar{B}_1 \Sigma_{n-1} \bar{B}_1^t) + \ldots + \mathrm{tr}(V_1 \bar{B}_p \Sigma_{n-p} \bar{B}_p^t) \\
&\quad = \mathbb{E}\left[X_n^t V_{p+1} X_n | Y_{n-1} = y\right] + \mathrm{tr}(V_1 C) + X_{n-1}^t \bar{A}_1^t V_1 \bar{A}_1 X_{n-1} + \ldots + X_{n-q}^t \bar{A}_q^t V_1 \bar{A}_q X_{n-q} \\
&\qquad + \mathrm{tr}(\bar{B}_1^t V_1 \bar{B}_1 \Sigma_{n-1}) + \ldots + \mathrm{tr}(\bar{B}_p^t V_1 \bar{B}_p \Sigma_{n-p}).
\end{aligned}$$

Since $X_n = \Sigma_n^{1/2}\epsilon_n$, $\Sigma_n^{1/2}\Sigma_n^{1/2} = \Sigma_n$ and $\mathbb{E}[\epsilon_n\epsilon_n^t] = \mathrm{Id}_d$ (since $(X_n)_{n\in\mathbb{N}^*}$ is a GARCH process of type Γ), we obtain

$$\begin{aligned}
&\mathbb{E}\left[X_n^t V_{p+1} X_n | Y_{n-1} = y\right] = \mathbb{E}\left[\mathrm{tr}(X_n (V_{p+1} X_n)^t) | Y_{n-1} = y\right] \\
&\quad = \mathrm{tr}\left(\mathbb{E}\left[X_n X_n^t V_{p+1} | Y_{n-1} = y\right]\right) = \mathrm{tr}\left(\mathbb{E}\left[X_n X_n^t | Y_{n-1} = y\right] V_{p+1}\right) = \mathrm{tr}(\Sigma_n V_{p+1}) \\
&\quad = \mathrm{tr}(V_{p+1} C) + \mathrm{tr}(V_{p+1} \bar{A}_1 X_{n-1} X_{n-1}^t \bar{A}_1^t) + \ldots + \mathrm{tr}(V_{p+1} \bar{A}_q X_{n-q} X_{n-q}^t \bar{A}_q^t) \\
&\qquad + \mathrm{tr}(V_{p+1} \bar{B}_1 \Sigma_{n-1} \bar{B}_1^t) + \ldots + \mathrm{tr}(V_{p+1} \bar{B}_p \Sigma_{n-p} \bar{B}_p^t) \\
&\quad = \mathrm{tr}(V_{p+1} C) + X_{n-1}^t \bar{A}_1^t V_{p+1} \bar{A}_1 X_{n-1} + \ldots + X_{n-q}^t \bar{A}_q^t V_{p+1} \bar{A}_q X_{n-q} \\
&\qquad + \mathrm{tr}(\bar{B}_1^t V_{p+1} \bar{B}_1 \Sigma_{n-1}) + \ldots + \mathrm{tr}(\bar{B}_p^t V_{p+1} \bar{B}_p \Sigma_{n-p}).
\end{aligned}$$

Hence, (3.17) can be rewritten as

$$\begin{aligned}
&\mathbb{E}[V(Y_n)|Y_{n-1}=y] \\
&\quad= \operatorname{tr}\left[\left(\bar{B}_1^t(V_1+V_{p+1})\bar{B}_1+V_2\right)\Sigma_{n-1}\right]+\ldots+\operatorname{tr}\left[\left(\bar{B}_{p-1}^t(V_1+V_{p+1})\bar{B}_{p-1}+V_p\right)\Sigma_{n-p+1}\right] \\
&\qquad+\operatorname{tr}\left[\bar{B}_p^t(V_1+V_{p+1})\bar{B}_p\Sigma_{n-p}\right]+X_{n-1}^t\left(\bar{A}_1^t(V_1+V_{p+1})\bar{A}_1+V_{p+2}\right)X_{n-1}+\ldots \\
&\qquad+X_{n-q+1}^t\left(\bar{A}_{q-1}^t(V_1+V_{p+1})\bar{A}_{q+1}+V_{p+q}\right)X_{n-q+1}+X_{n-q}^t\bar{A}_q^t(V_1+V_{p+1})\bar{A}_qX_{n-q} \\
&\qquad+\operatorname{tr}[(V_1+V_{p+1})C]+1.
\end{aligned}$$

By definition of V_i, we deduce

$$\begin{aligned}
&\bar{B}_k^t(V_1+V_{p+1})\bar{B}_k+V_{k+1}=V_k-\frac{C}{p+q}, \quad 1\leq k\leq p-1 \\
&\bar{B}_p^t(V_1+V_{p+1})\bar{B}_p=V_p-\frac{C}{p+q} \\
&\bar{A}_k^t(V_1+V_{p+1})\bar{A}_k+V_{p+k+1}=V_{p+k}-\frac{C}{p+q}, \quad 1\leq k\leq q-1 \\
&\bar{A}_q^t(V_1+V_{p+1})\bar{A}_q=V_{p+q}-\frac{C}{p+q}.
\end{aligned}$$

Furthermore, V_k is symmetric positive definite for all $k=1,\ldots,p+q$ which implies that $V_k-\frac{C}{p+q}$ is symmetric positive semidefinite for all $k=1,\ldots,p+q$.
Consider the nonnegative constants $(\alpha_k)_{1\leq k\leq p+q}$ defined by

$$\alpha_k:=\max\left\{x^t\left(V_k-\frac{C}{p+q}\right)x:\ x\in\mathbb{R}^d,\ x^tV_kx=1\right\}.$$

Since the maximum is calculated over the unit sphere with respect to the induced norm by V_k and since this unit sphere is compact there exists $x_k\in\mathbb{R}^d$ such that $x_k^tV_kx_k=1$ and

$$\begin{aligned}
\alpha_k&=x_k^t\left(V_k-\frac{C}{p+q}\right)x_k \\
&=1-x_k^t\frac{C}{p+q}x_k.
\end{aligned}$$

The matrices V_k, $k=1,\ldots,p+q$, and $\frac{C}{p+q}$ are positive definite which yields $0\leq\alpha_k<1$ for all $k=1,\ldots,p+q$.
Setting $\alpha_0:=\max\{\alpha_k:\ k=1,\ldots,p+q\}$ we obtain $0\leq\alpha_0<1$ and

$$\forall k\in\{1,\ldots,p+q\}\qquad V_k-\frac{C}{p+q}\leq\alpha_0V_k.$$

Hence, for all $M \in \mathbb{S}_d^{++}$ and all $k \in \{1, \dots, p+q\}$,

$$\mathrm{tr}\left[\left(V_k - \frac{C}{p+q}\right) M\right] \leq \alpha_0 \ \mathrm{tr}(V_k M).$$

We deduce

$$\mathbb{E}[V(Y_n)|Y_{n-1} = y] \leq \alpha_0 V(y) + \mathrm{tr}(\Sigma C) + 1 - \alpha_0.$$

If we choose $\alpha := (\alpha_0 + 1)/2 \in [1/2, 1)$ and $b := \mathrm{tr}(\Sigma C) + 1 - \alpha_0 \in (0, \infty)$, then the (FL) - condition is satisfied on the compact set K given by

$$K := \left\{ x \in W : \ V(x) \leq \frac{b}{\alpha - \alpha_0} \right\}.$$

□

3.3.6 Harris Recurrence, Ergodicity and β - Mixing

Theorem 3.3.9.

(i) Suppose that (H1) and (H2) hold and that the spectral radius of $(\sum_{i=1}^q A_i + \sum_{j=1}^p B_j)$ is less than 1, then the Markov chain $(Y_n)_{n\in\mathbb{N}^*}$ is positive Harris recurrent and geometrically ergodic on the state space $(W, \mathcal{B}(W))$. Moreover, the strictly stationary solution $(X_n)_{n\in\mathbb{Z}}$ of the GARCH(p,q) model of type Γ associated with $(Y_n)_{n\in\mathbb{Z}}$ is unique and geometrically β - mixing.

(ii) If there exists a stationary solution for the GARCH(p,q) model of type Γ, then the spectral radius of the matrix $(\sum_{i=1}^q A_i + \sum_{j=1}^p B_j)$ is less than 1.

Proof.
(i) Suppose that (H1) and (H2) are satisfied. If the spectral radius of $(\sum_{i=1}^q A_i + \sum_{j=1}^p B_j)$ is less than 1, Theorem 3.3.8 ensures the existence of a function V which fulfills the (FL) - condition on a compact set K.
Since, due to Proposition 3.3.4 (ii), the spectral radius of $\sum_{j=1}^p B_j$ is also less than 1, Proposition 3.3.6 and Theorem 3.3.7 imply that (A2) and (A3) hold. Using then Proposition 2.3.7 we deduce that $(Y_n)_{n\in\mathbb{N}^*}$ is ψ - irreducible and aperiodic on the state space $(W, \mathcal{B}(W))$. Moreover, $(Y_n)_{n\in\mathbb{N}^*}$ has the Feller property (cf. Example 1.1.12) and supp ψ has non-empty interior (see again Proposition 2.3.7). Thus, Corollary 1.1.14 shows that K is small.
We can conclude with Theorem 2.3.8 and Theorem 2.3.9.
(ii) We now assume that there is a stationary solution for the GARCH(p,q) model of type

Γ. Then $\Sigma := \mathbb{E}[X_n X_n^t]$ is well-defined.
Since $\Sigma = \mathbb{E}[\Sigma_n]$, taking the expectation in (3.6) on both sides yield

$$\Sigma = C + \sum_{i=1}^{q} \sum_{k=1}^{l_i} \bar{A}_{i,k} \Sigma \bar{A}_{i,k}^t + \sum_{j=1}^{p} \sum_{r=1}^{s_j} \bar{B}_{j,r} \Sigma \bar{B}_{j,r}^t.$$

Due to Proposition 3.3.2 the spectral radius of the matrix $(\sum_{i=1}^q A_i + \sum_{j=1}^p B_j)$ has to be less than 1. □

Remark 3.3.10. Suppose that (H1) and (H2) hold and that the spectral radius of $(\sum_{i=1}^q A_i + \sum_{j=1}^p B_j)$ is less than 1, then Theorem 2.3.8 yields in addition $\pi(V) < \infty$ and hence

$$\forall n \in \mathbb{Z} \qquad \mathbb{E}[X_n^t V_{p+1} X_n] \leq \mathbb{E}[V(Y_n)] = \pi(V) < \infty$$

by definition of V (cf. proof of Theorem 3.3.8). This shows that $X_n \in L^2$ for all $n \in \mathbb{Z}$.
Since $\mathbb{E}[\Sigma_n] = \mathbb{E}[X_n X_n^t]$ (cf. Remark 3.2.2), we deduce $\mathbb{E}[\Sigma_n] < \infty$. One may verify that this implies $\Sigma_n \in L^1$ for all $n \in \mathbb{Z}$.
The strictly stationary solution in Theorem 3.3.9 (i) is hence also weakly stationary.

Example 3.3.11.
We give an easy example where the algebraic variety W differs from the whole space.
Consider the following bivariate GARCH(1, 1) model:

$$\text{vech}(\Sigma_n) = \text{vech}(C) + A\ \text{vech}(X_{n-1} X_{n-1}^t) + B\ \text{vech}(\Sigma_{n-1}) \tag{3.18}$$

where A and B are two 3×3 matrices such that the spectral radius of $A + B$ is less than 1 and $BA = 0$.
The attracting point T of this chain is defined by $T = \mathscr{C} + \tilde{B} T$ (cf. (3.12)).
Starting from the initial point T (i.e. $\text{vech}(\Sigma_0) = \text{vech}(\tilde{\Sigma})$ and $X_0 = 0$), we note that $\text{vech}(\Sigma_0) = \text{vech}(C) + B\ \text{vech}(\Sigma_0)$ and obtain by iterating (3.18)

$$\text{vech}(\Sigma_n) = \text{vech}(\Sigma_0) + A\ \text{vech}(X_{n-1} X_{n-1}^t).$$

Let f be the regular map from $\mathbb{R}^4$ into $\mathbb{R}^5$ given by

$$(x_1, x_2, x_3, x_4) \mapsto f(x_1, x_2, x_3, x_4) := T + \begin{pmatrix} A\ (x_1^2, x_1 x_2, x_2^2)^t \\ x_3 \\ x_4 \end{pmatrix}.$$

Then W is the Zariski closure of the semi-algebraic set $f(\mathbb{R}^4)$ and W has to be strictly contained in $\mathbb{R}^5$ since $\dim f(\mathbb{R}^4) \leq 4 = \dim \mathbb{R}^4$.

Bibliography

[1] R. Benedetti and J.-J. Risler, *Real algebraic and semi-algebraic sets*, Hermann, Paris, 1990.

[2] T. Bollerslev, *Generalized Autoregressive Conditional Heteroskedasticity*, Journal of Econometrics **31** (1986), 307–327.

[3] F. Boussama, *Ergodicité, mélange et estimation dans les modèles GARCH*, Ph.D. thesis, Université Paris 7, 1998.

[4] P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods*, 2nd ed., Springer, New York, 1991.

[5] Y. A. Davydov, *Mixing conditions for Markov chains*, Theory of Probability and its Applications **18** (1973), no. 2, 312–328.

[6] J. Dieudonné, *Éléments d'analyse*, 2nd ed., t. III, Gauthier-Villars, Paris/Bruxelles/Montréal, 1974.

[7] M. Duflo, *Méthodes récursives aléatoires*, Masson, 1990.

[8] R. F. Engle, *Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation*, Econometrica **50** (1982), 987–1007.

[9] R. F. Engle and K. F. Kroner, *Multivariate Simultaneous Generalized ARCH*, Econometric Theory **11** (1995), 122–150.

[10] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, MA, 1985.

[11] S. Lang, *Algebra*, Rev. 3rd ed., Springer, 2002.

[12] S. P. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability*, Springer, London, 1993.

[13] A. Mokkadem, *Propriétés de mélange des processus autorégressifs polynomiaux*, Annales de l'I. H. P. **26** (1990), no. 2, 219–260.

[14] D. Mumford, *Algebraic Geometry I, Complex Projective Varieties*, Springer-Verlag, Berlin, 1976.

[15] D. F. Nicholls and B. G. Quinn, *Random Coefficient Autoregressive Models: An Introduction*, Lecture Notes in Statistics, vol. 11, Springer-Verlag, New York, 1982.

[16] J. E. Potter, *Matrix Quadratic Solutions*, SIAM Journal on Applied Mathematics **14** (1966), no. 3, 496–501.

[17] J.-J. Risler, *Le théorème des zéros en géométries algébrique et analytique réelles*, Bulletin de la Société Mathématique de France **104** (1976), 113–127.

[18] R. Stelzer, *On the relation between the vec and BEKK multivariate GARCH models*, Econometric Theory **24** (2008), 1131–1136.

[19] R. L. Tweedie, *Sufficient conditions for ergodicity and recurrence of Markov chains on a general state space*, Stochastic Processes and their Applications **3** (1975), no. 4, 385–403.

[20] ———, *The Existence of Moments for Stationary Markov Chains*, Journal of Applied Probability **20** (1983), no. 1, 191–196.

[21] L. E. Ward, *Partially Ordered Topological Spaces*, Proceedings of the American Mathematical Society **5** (1954), no. 1, 144–161.