Weak Approximation of Stochastic Delay Differential Equations with Bounded Memory by Discrete Time Series

DISSERTATION

zur Erlangung des akademischen Grades
doctor rerum naturalium
(Dr. rer. nat.)
im Fach Mathematik

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät II
der Humboldt-Universität zu Berlin

von
Herrn Dipl.-Math.RobertLorenz
geboren am 9. Juni 1973 in Schwedt/Oder

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Hans Jürgen Prömel

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät II:
Prof. Dr. Uwe Küchler

Gutachter:

1. Prof. Dr. Uwe Küchler
2. Prof. Dr. Evelyn Buckwar
3. Prof. Dr. Hans-Michael Dietz

eingereicht am:	25. Oktober 2005
Tag der mündlichen Prüfung:	20. März 2006

Abstract

Consider the stochastic delay differential equation (SDDE) with length of memory r

$$dX(t) = b(X_t)dt + \sigma(X_t)dB(t),$$

which has a unique weak solution. Here B is a Brownian motion, b and σ are continuous, locally bounded functions defined on the space $C[-r, 0]$, and X_t denotes the segment of the values of $X(u)$ for time points u in the interval $[t, t-r]$. Our aim is to construct a sequence of discrete time series X^h of higher order, such that X^h converges weakly to the solution X of the stochastic differential delay equation as h tends to zero.

On the other hand we shall establish under which conditions a given sequence of time series X^h of higher order converges weakly to the weak solution X of a stochastic differential delay equation.

As an illustration we shall derive a weak limit of a sequence of GARCH processes of higher order. This limit tends out to be the weak solution of a stochastic differential delay equation.

Keywords:
stochastic delay differential equations, weak approximation, discrete time series, GARCH processes

Zusammenfassung

Wir betrachten die stochastische Differentialgleichung mit Gedächtnis (SDDE) mit Gedächtnislänge r

$$dX(t) = b(X_t)dt + \sigma(X_t)dB(t)$$

mit eindeutiger schwacher Lösung . Dabei ist B eine Brownsche Bewegung, b und σ sind stetige, lokal beschränkte Funktionen mit Definitionsbereich $C[-r, 0]$, und X_t bezeichnet das Segment der Werte von $X(u)$ für Zeitpunkte u im Intervall $[t, t-r]$. Unser Ziel ist eine Folge von diskreten Zeitreihen X^h höherer Ordung zu konstruieren, so dass mit h gegen 0 die Zeitreihen X^h schwach gegen die Lösung X der stochastischen Differentialgleichung mit Gedächtnis konvergieren.

Desweiteren werden wir Bedingungen angeben, unter denen eine gegeben Folge von Zeitreihen X^h höherer Ordung schwach gegen die Lösung X einer stochastischen Differentialgleichung mit Gedächtnis konvergiert.

Als ein Beispiel werden wir den schwachen Grenzwert einer Folge von diskreten GARCH-Prozessen höherer Ordnung ermitteln. Dieser Grenzwert wird sich als schwache Lösung einer stochastischen Differentialgleichung mit Gedächtnis herausstellen.

Schlagwörter:
stochastische Differentialgleichungen mit Gedächtnis, schwache Approximation, diskrete Zeitreihen, GARCH-Prozesse

Contents

Chapter 1

Introduction

Convergence of stochastic processes

Let $X := (X(t), t \geq 0)$ be a continuous stochastic process. Often one is interested in the distribution of certain functionals of the process, for instance of $\phi(X) = \max_{0 \leq t \leq T} X(t)$. In general it is difficult to determine the distribution of $\phi(X)$. One way to tackle this problem is to consider an appropriate sequence of processes X_n converging weakly to X. Sometimes the distribution of the functional under consideration can be determined much easier for every X_n, and the distribution for X can be obtained as a limit distribution. This is exactly the procedure of Donsker's invariance principle. Along this line we shall establish convergence results for weak solutions of stochastic delay differential equations.

Stochastic delay differential equations

Stochastic delay differential equations (SDDE's) have become widespread in the last 30 years. Phenomena of time delay occur in many different areas of the real world. Stochastic delay differential equations are their mathematical reflection. A description without time delay is nowadays not to think of. In physics it is the time of transportation of particles or information from one system to another. In financial mathematics it is the time to react on developments in financial markets. In econometrics time delay corresponds to the reaction time of the client to behave in a certain way. A first survey on the theory of SDDE's is presented in Mohammed [21] and Mao [19]. We shall consider SDDE's of the kind

$$\begin{cases} X(t) &= \xi(t), \quad -r \leq t \leq 0 \\ dX(t) &= b(X_t)\,dt + \sigma(X_t)\,dB(t), \quad t \geq 0. \end{cases} \tag{1.0.1}$$

Here B is a Brownian motion, b and σ are measurable, locally bounded functions defined on $C[-r, 0]$, and ξ is a deterministic, continuous function on $[-r, 0]$. Furthermore X_t is the segment $(X(t+u))_{-r \leq u \leq 0}$, where $r \geq 0$ denotes the length of memory. We exclude the case $r = \infty$ which has been treated in Riedle [25]. In the case $r = 0$ the system (1.0.1) is a stochastic ordinary differential equation. The aim of this work is to approximate weakly the solution of (1.0.1). We shall construct processes X_n which will converge weakly to the process X, where X is the unique weak solution of the system (1.0.1). This will be done firstly if the coefficients b and σ are continuous and bounded, secondly if they are continuous and locally bounded, finally we admit discontinuity points for the coefficients b and σ. The approximating processes are constructed in a first step as autoregressive time series $(X^{(h)}_{mh})_{m\in\mathbb{N}_0}$ on a time grid $\{mh : m \in \mathbb{N}_0\}$ for $h > 0$. The quantity h is called step length. In a second step $X^{(h)}$ is extended to a continuous process by linear interpolation. To indicate the correspondence to the step length h, we shall denote the approximating processes by $X^{(h)}$ rather than by X_n.

History of convergence results for stochastic processes

One of the first results on convergence of stochastic processes is the famous Donsker theorem, see Billingsley [2], Theorem 10.1. If $\{\epsilon_k\}$ is a sequence of i.i.d. centered random variables with variance 1, then the sequence of processes defined by

$$S_n(t) := \frac{1}{\sqrt{n}} \sum_{k=1}^{[nt]} \epsilon_k, \quad 0 \le t \le T$$

converges weakly to a Brownian motion on $[0, T]$ as n tends to infinity. The Brownian motion is a special case of a Markov diffusion with vanishing drift coefficient and diffusion coefficient 1.

A general result on convergence of stochastic processes to a Markov diffusion is presented in Stroock and Varadhan [28]. For each $h > 0$ let us be given a d-dimensional Markov chain $X^{(h)} = (X^{(h)}_0, X^{(h)}_h, X^{(h)}_{2h}, \dots)$ with transition probabilities

$$p^{(h)}_{kh}(X^{(h)}_{kh}, A) = P(X^{(h)}_{(k+1)h} \in A | X^{(h)}_{kh}), \quad A \in \mathbb{B}^d. \tag{1.0.2}$$

In terms of the transition probabilities the following quantities are defined for $t \ge 0$ and $x \in \mathbb{R}^d$

$$\begin{aligned} a^{(h)}(t,x) &:= \frac{1}{h} \int_{\mathbb{R}^d} (y-x) p^{(h)}_{[\frac{t}{h}]h}(x, dy) \\ b^{(h)}(t,x) = \sigma^{(h)2}(t,x) &:= \frac{1}{h} \int_{\mathbb{R}^d} (y-x)(y-x)^T p^{(h)}_{[\frac{t}{h}]h}(x, dy). \end{aligned}$$

A stochastic process $X^{(h)}(t)$ in continuous time is constructed by linear interpolation. If there exist functions a and σ such that

$$a^{(h)}(t,x) \xrightarrow[h\to 0]{} a(t,x), \quad \sigma^{(h)2}(t,x) \xrightarrow[h\to 0]{} \sigma^2(t,x), \quad t \ge 0, \quad x \in \mathbb{R}^d$$

uniformly on compact sets of $\mathbb{R}_+ \times \mathbb{R}^d$, then the sequence of processes $X^{(h)}(t)$ converges weakly to a Markov diffusion X with coefficients a and σ. This means, the process X

is the weak solution of the stochastic ordinary differential equation

$$dX(t) = a(t, X(t))\, dt + \sigma(t, X(t))\, dB(t), \quad t \geq 0, \tag{1.0.3}$$

where B denotes a Brownian motion. The proof in Stroock and Varadhan [28] uses that the weak solution of (1.0.3) solves a martingale problem. The same result in the one-dimensional case can also be found in Gichman and Skorochod [6]. Here the proof is performed by analysis of finite-dimensional distributions via characteristic functions.

In Jacod and Shiryaev [10] it is studied when a sequence of semimartingales converges weakly. The theory in this book provides results for approximating processes $X_n(t)$ being piecewise constant processes. Here, in contrast to Stroock and Varadhan [28], the underlying space is the space of right-continuous functions with left-hand limits (cadlag-space).

A general guideline

Including the introduction this thesis consists of three chapters. As a first result in Chapter two we will show in generalization of Stroock and Varadhan [28] that every weak solution of a stochastic delay differential equation (SDDE) corresponds to one and only one solution of a martingale problem. To approximate weakly a given SDDE with a unique weak solution X, we will construct a sequence of continuous autoregressive schemes $(X_{mh}^{(h)})_{m\in\mathbb{N}_0}$ of higher order. We shall do it in such a way that the order of $(X_{mh}^{(h)})_{m\in\mathbb{N}_0}$ increases to infinity as the step length h tends to zero. The schemes $(X_{mh}^{(h)})_{m\in\mathbb{N}_0}$ are extended to continuous processes $X^{(h)}(t)$ by linear interpolation. As a main result we will give conditions under which the sequence $\{X^{(h)}(t)\}$ converges weakly to the solution X of the given SDDE as h tends to zero. The occurrence of time delay in the weak limit can be explained by the unboundedly increasing order of the autoregressive schemes $(X_{mh}^{(h)})_{m\in\mathbb{N}_0}$ as h tends to zero. There is a series of applications. We shall use the main result to establish the weak limit for a given sequence of autoregressive schemes with unboundedly increasing order. We shall illustrate the procedure for a sequence of GARCH-processes $(X^{(h)}, \rho^{(h)2})$. We will give conditions on the coefficients of the GARCH-processes under which the sequence of processes $(X^{(h)}, \rho^{(h)2})$ converges weakly as h tends to zero. The limit process will be the weak solution (X, ρ^2) of a SDDE. To emphasize the importance of the assumptions in the main result, we shall give two counterexamples where the sequence $\{X^{(h)}\}$ does not converge weakly to the solution process X. As a further application we shall use the main result to prove the existence of a weak solution for a certain class of SDDE's. Every weak solution Y of an SDDE is a semimartingale. We shall approximate the solution Y by a sequence of piecewise constant processes $Y^{(h)}$. As mentioned before, the processes $Y^{(h)}$ take values in the cadlag-space. We shall give conditions under which the sequence $\{Y^{(h)}\}$ converges weakly to the solution process Y. The proof of the convergence result we shall perform with the help of semimartingale theory.

Chapter three deals with a special class of stationary Gaussian processes. We shall consider piecewise constant ARMA($p^{(h)}+1, q^{(h)}$)-processes of unboundedly increasing order of the kind

$$\begin{cases} Y^{(h)}_{(m+1)h} & = & Y^{(h)}_{mh} + \sum_{j=0}^{p^{(h)}} a^{(h)}_j Y^{(h)}_{(m-j)h} h + \sum_{i=0}^{q^{(h)}} \sigma^{(h)}_i \sqrt{h} \epsilon_{m+1-i}, \quad m \in \mathbb{Z} \\ Y^{(h)}_t & = & Y^{(h)}_{[\frac{t}{h}]h}, \quad t \in \mathbb{R}. \end{cases} \tag{1.0.4}$$

We will study under which conditions on the coefficients the sequence $\{Y^{(h)}\}$ converges weakly. In this case we shall get a stationary Gaussian process as a limit. It turns out to be the stationary solution of a stochastic equation of the kind

$$dY(t) = \int_{-r}^{0} Y(t+u)\, da(u)\, dt + dZ(t), \quad t \geq 0, \tag{1.0.5}$$

where Z is a certain mixture of a Brownian motion process B, indeed, the driving force Z may be represented in the form

$$Z(t) = \int_{-q}^{0} [B(t+u) - B(u)]\, d\sigma(u), \quad t \geq 0. \tag{1.0.6}$$

The time delay in the drift occurs because of the unbounded increase of the number of coefficients $\{a^{(h)}_j\}$. If $q^{(h)} = 0$ and $\sigma^{(h)}_0 = 1$, then the driving force Z in (1.0.6) is a Brownian motion. This case has been studied in Section 2.5 of Reiß [23]. The occurrence of the new kind of driving force Z in (1.0.6) is explained by the unbounded increase of the number of coefficients $\{\sigma^{(h)}_i\}$. Based on the stationarity of the underlying processes we will prove convergence of finite-dimensional distributions by using spectral densities. We will study the process Z in (1.0.6) and the stationary solution process Y in (1.0.5) in detail. We will give explicit representations of Y in terms of the underlying Brownian motion B and an Ornstein-Uhlenbeck process X.

Chapter 2

Stochastic Delay Differential Equations Driven by a Brownian Motion

2.1 Introduction

Let $0 \leq r < \infty$ denote the length of memory. Furthermore let $B = (B_1, \ldots, B_n)$ denote an n-dimensional Brownian motion for $n \in \mathbb{N}$. Consider the following stochastic differential delay equation with values in $\mathbb{R}^d$ for $d \in \mathbb{N}$

$$\begin{cases} X_0 &= \xi \\ dX(t) &= b(X_t)\,dt + \sigma(X_t)\,dB(t), \quad t \geq 0, \end{cases} \tag{2.1.1}$$

or in coordinate form

$$\begin{cases} (X_0)_i &= \xi_i, \quad i = 1, \ldots, d \\ dX_i(t) &= b_i(X_t)\,dt + \sum_{j=1}^{n} \sigma_{ij}(X_t)\,dB_j(t), \quad i = 1, \ldots, d, \quad t \geq 0. \end{cases}$$

Here

$$b : C([-r, 0]; \mathbb{R}^d) \longrightarrow \mathbb{R}^d, \quad \sigma : C([-r, 0]; \mathbb{R}^d) \longrightarrow \mathcal{M}(\mathbb{R}^{d \times n}), \quad d, n \in \mathbb{N}$$

are measurable functions, and X_t denotes the segment $(X(t + u))_{-r \leq u \leq 0}$ of X at time t. Note that $X_t \in C([-r, 0]; \mathbb{R}^d)$. For $x \in C([-r, 0]; \mathbb{R}^d)$ define the norm

$$\|x\|_\infty := \sup_{-r \leq u \leq 0} |x(u)|.$$

Then $C([-r, 0]; \mathbb{R}^d)$ is a Polish space. We assume for the initial segment X_0 that it is deterministic: $X_0 = \xi$ for a function $\xi \in C([-r, 0]; \mathbb{R}^d)$. One can consider strong and weak solutions for (2.1.1). In this chapter we will deal with weak solutions only.

2.1.1 Definition. *Let $\xi \in C([-r,0];\mathbb{R}^d)$ be an initial segment. A weak solution of (2.1.1) with start in ξ is a sextuple $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, Q, B, X)$ such that $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, Q)$ is a filtered probability space, and B and X are processes defined on this space satisfying the following four conditions.*

1. *B is a continuous n-dimensional Brownian motion martingale, and X is a continuous adapted d-dimensional process.*
2. *$X_0 = \xi$ Q-a.s.*
3. *$\int_0^t |b(X_s)| + \|\sigma(X_s)\|\, ds < \infty$ Q-a.s. for all $t \geq 0$.*
4. *Equation (2.1.1) holds Q-a.s. for all $t \geq 0$.*

We say that weak existence holds for the SDDE (2.1.1) if there is a weak solution of (2.1.1) with start in ξ. We say that weak uniqueness holds for the SDDE (2.1.1) if all weak solutions of (2.1.1) with start in ξ have the same law.

2.2 The Martingale Problem on $C[-r,\infty)$

The aim of this section is to formulate the martingale problem on the function space $C([-r,\infty);\mathbb{R}^d)$ and to establish its connection with weak solutions of (2.1.1). Let $\Omega := C([-r,\infty);\mathbb{R}^d)$ denote the space of all continuous $\mathbb{R}^d$-valued functions on $[-r,\infty)$. We denote a generic element of Ω by m rather than by ω. Define the coordinate projection by

$$X^\circ(t)(m) := m(t), \quad t \geq -r, \quad m \in \Omega$$

and the segment projection by

$$X_t^\circ(m) := (m(t+u))_{-r\leq u\leq 0}, \quad t \geq 0, \quad m \in \Omega.$$

The natural filtration $(\mathcal{M}_t)_{t\geq -r}$ on Ω is defined by

$$\mathcal{M}_t := \sigma(X^\circ(u) : -r \leq u \leq t), \quad t \geq -r.$$

Finally define the following σ-algebra on Ω

$$\mathcal{M} := \bigvee_{t\geq -r} \mathcal{M}_t.$$

Then $(\Omega, \mathcal{M})$ is a measurable space. It is called the canonical space. The coordinate and segment projection are measurable functions on Ω with respect to the Borel-σ-algebra of $\mathbb{R}^d$ and of $C([-r,0];\mathbb{R}^d)$. Define a metric on Ω by

$$d(m_1,m_2) := \sum_{T=1}^{\infty} \frac{1}{2^T} \sup_{-r\leq u\leq T} (|m_1(u) - m_2(u)| \wedge 1), \quad m_1, m_2 \in \Omega.$$

Then Ω is d-complete, and the σ-algebra $\mathcal{M}$ equals the σ-algebra generated by the d-open sets. This is the well-known Skorochod topology. In the following lemma we shall give a property of the Skorochod topology regarding unbounded time intervals.

2.2.1 Lemma. *Let $\{P_n\}$ be a sequence of probability measures on $C([-r,\infty);\mathbb{R}^d)$. If the marginal distributions $(P_n)_T$ on $C([-r,T];\mathbb{R}^d)$ converge weakly to Q_T for all $T>0$ in the Skorochod topology on $C([-r,T];\mathbb{R}^d)$, and Q_T on $C([-r,T];\mathbb{R}^d)$ is a marginal of a probability measure Q on $C([-r,\infty);\mathbb{R}^d)$ for all $T>0$, then the sequence $\{P_n\}$ converges weakly to Q in the Skorochod topology on $C([-r,\infty);\mathbb{R}^d)$.*

Proof. Recall the notation $\Omega=(C[-r,\infty);\mathbb{R}^d)$. Let f be a real-valued bounded, uniformly continuous function on Ω. Then for all $\epsilon>0$ there exists a $\delta>0$ such that

$$|f(m)-f(\overline{m})|<\frac{\epsilon}{2\|f\|_\infty},\quad d(m,\overline{m})<\delta,\quad m,\overline{m}\in\Omega.$$

Choose $T\in\mathbb{N}$ such large that $\sum_{k\geq T}\frac{1}{2^k}<\delta$. Let $\pi_T m$ denote the projection of $m\in\Omega$ onto the finite interval $[-r,T]$. Define the continuous functions $m_T(u):=m(u)1_{\{u\leq T\}}+m(T)1_{\{u\geq T\}}, u\geq -r$ and $\phi(\pi_T m):=f(m_T)$. Then it follows by transformation of measures that

$$\left|\int_{C([-r,\infty);\mathbb{R}^d)} f(m)\,dP_n(m)-\int_{C([-r,T];\mathbb{R}^d)}\phi(\pi_T m)\,d(P_n)_T(\pi_T m)\right| < \frac{\epsilon}{2}$$
$$\left|\int_{C([-r,\infty);\mathbb{R}^d)} f(m)\,dQ(m)-\int_{C([-r,T];\mathbb{R}^d)}\phi(\pi_T m)\,dQ_T(\pi_T m)\right| < \frac{\epsilon}{2}.$$

Since by assumption the sequence $(P_n)_T$ converges weakly to Q_T, the proof is finished. □

Now fix a measurable vector function $b=(b_i)_{1\leq i\leq d}$ from $C([-r,0];\mathbb{R}^d)$ to $\mathbb{R}^d$ and a measurable matrix function $a=(a_{ij})_{1\leq i,j\leq d}$ from $C([-r,0];\mathbb{R}^d)$ to $\mathcal{M}^+(\mathbb{R}^{d\times d})$. Let $C_0^\infty(\mathbb{R}^d)$ denote the space of the real-valued, infinitely often differentiable functions on $\mathbb{R}^d$ with compact support. Define for all functions $f\in C_0^\infty(\mathbb{R}^d)$ the operator

$$(L_{b,a}f)(x):=\sum_{i=1}^d b_i(x)\frac{\partial f}{\partial x_i}(x(0))+\frac{1}{2}\sum_{i,j=1}^d a_{ij}(x)\frac{\partial^2 f}{\partial x_i\partial x_j}(x(0)),\quad x\in C([-r,0];\mathbb{R}^d). \tag{2.2.1}$$

The domain of this operator for a fixed function $f\in C_0^\infty(\mathbb{R}^d)$ is $C([-r,0];\mathbb{R}^d)$, and it takes values in $\mathbb{R}$. Furthermore fix an $\mathbb{R}^d$-valued continuous function ξ on $[-r,0]$. We shall now give a definition of a martingale problem for the operator in (2.2.1).

2.2.2 Definition. *A probability measure Q_ξ on $(\Omega,\mathcal{M})$ solves the martingale problem associated with b and a with start in ξ if*

1. $Q_\xi(X_0^\circ=\xi)=1$.

2. *$f(X^\circ(t)) - \int_0^t (L_{b,a}f)(X_s^\circ)\, ds, t \geq 0$ is a $(\mathcal{M}_t, Q_\xi)$-martingale for all $f \in C_0^\infty(\mathbb{R}^d)$.*

A probability measure Q_ξ on $(\Omega, \mathcal{M})$ solves the local martingale problem associated with b and a with start in ξ if

1. *$Q_\xi(X_0^\circ = \xi) = 1$.*
2. *$f(X^\circ(t)) - \int_0^t (L_{b,a}f)(X_s^\circ)\, ds, t \geq 0$ is a local $(\mathcal{M}_t, Q_\xi)$-martingale for all $f \in C_0^\infty(\mathbb{R}^d)$.*

This formulation of the martingale problem differs from martingale problems in other literature. In Stroock and Varadhan [28] the operator $(L_{b,a}f)(x)$ has the form

$$(L_{b,a}f)(x) = \sum_{i=1}^d b_i(x) \frac{\partial f}{\partial x_i}(x) + \frac{1}{2} \sum_{i,j=1}^d a_{ij}(x) \frac{\partial^2 f}{\partial x_i \partial x_j}(x), \quad x \in \mathbb{R}^d.$$

This is the operator in (2.2.1) in the case $r = 0$. In Karatzas and Shreve [12] for $t \geq 0$ a time-dependent operator $L_{b,a}^t f$ is defined for time-dependent coefficients b^t and a^t with increasing delay:

$$(L_{b^t,a^t}^t f)(x) = \sum_{i=1}^d b_i^t(x) \frac{\partial f}{\partial x_i}(x(t)) + \frac{1}{2} \sum_{i,j=1}^d a_{ij}^t(x) \frac{\partial^2 f}{\partial x_i \partial x_j}(x(t)), \quad x \in C([0,\infty); \mathbb{R}^d).$$

The formulation of the martingale problem with this operator does not take into account an initial segment $\xi \in C([-r,0]; \mathbb{R}^d)$ for $r > 0$.

Now we shall come to the relation between the martingale problem of Definition 2.2.2 and weak solutions of stochastic delay differential equations. Firstly assume that there exists a weak solution of (2.1.1) with start in ξ. Then there exists a sextuple $(\Omega, \mathcal{F}, (\mathcal{F}_t), P, B, X)$ such that

$$X_i(t) = \xi_i(0) + \int_0^t b_i(X_s)\, ds + \sum_{j=1}^n \int_0^t \sigma_{ij}(X_s)\, dB_j(s), \quad t \geq 0, \quad i = 1, \dots, d$$

holds a.s., or equivalently

$$dX_i(t) = b_i(X_t)\, dt + \sum_{j=1}^n \sigma_{ij}(X_t)\, dB_j(t), \quad t \geq 0, \quad i = 1, \dots, d. \tag{2.2.2}$$

Define the probability measure

$$Q_\xi(A) := P(X \in A), \quad A \in \mathcal{M}.$$

We shall show that Q_ξ solves the local martingale problem of Definition 2.2.2 for the coefficients b and $a=\sigma\sigma^T$, where T denotes the transpose. For $f \in C_0^\infty(\mathbb{R}^d)$ it follows from Itô's formula that a.s.

$$\begin{aligned} f(X(t)) = f(X(0)) &+ \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(X(s))\, dX_i(s) \\ &+ \frac{1}{2}\sum_{i,j=1}^d \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X(s))\, d<X_i,X_j>(s), \quad t\geq 0. \end{aligned}$$

Using (2.2.2) and that $d<X_i,X_j>(s) = (\sigma\sigma^T)_{ij}(X_s)\,ds$ we see that

$$\begin{aligned} M(t) &:= f(X(t)) - f(X(0)) - \sum_{i=1}^d \int_0^t b_i(X_s)\frac{\partial f}{\partial x_i}(X(s))\, ds \\ &- \frac{1}{2}\sum_{i,j=1}^d \int_0^t (\sigma\sigma^T)_{ij}(X_s)\frac{\partial^2 f}{\partial x_i \partial x_j}(X(s))\, ds \\ &= f(X(t)) - f(X(0)) - \int_0^t (L_{b,a}f)(X_s)\, ds, \quad t\geq 0 \end{aligned}$$

is a local $(\mathcal{F}_t, P)$-martingale. Then by transformation of measures it holds that

$$f(X^\circ(t)) - \int_0^t (L_{b,a}f)(X_s^\circ)\, ds, \quad t\geq 0$$

is a local $(\mathcal{M}_t, Q_\xi)$-martingale on the canonical space $\Omega = C([-r,\infty);\mathbb{R}^d)$. This shows that Q_ξ, the distribution of the solution process, solves the local martingale problem. Since every distribution of the solution process solves the local martingale problem, it holds that from uniqueness of the local martingale problem it follows weak uniqueness for the system (2.1.1). We are also interested in the other direction. We would like to establish that from weak uniqueness for the system (2.1.1) it follows uniqueness of the local martingale problem. At first we need a result on weak existence.

2.2.3 Theorem. *Assume that a probability measure Q on $(\Omega,\mathcal{M})$ solves the local martingale problem for the coefficients b and $a=\sigma\sigma^T$ with start in ξ. Then there exists a weak solution (X,B) of (2.1.1) such that Law $(X(t): t\geq -r) = Q$.*

Proof. For $t\geq 0$ and $m \in C([-r,\infty);\mathbb{R}^d)$ define the time-dependent, measurable, $(\mathcal{M}_t)$-adapted processes

$$\begin{aligned} b^t(m) &:= \begin{cases} b((\xi(w))_{t-r\leq w\leq 0}, (m(v))_{0\leq v\leq t})\cdot 1_{\{m(0)=\xi(0)\}} & ,0\leq t\leq r \\ b(m_t) & ,r\leq t \end{cases} \\ a^t(m) &:= \begin{cases} a((\xi(w))_{t-r\leq w\leq 0}, (m(v))_{0\leq v\leq t})\cdot 1_{\{m(0)=\xi(0)\}} & ,0\leq t\leq r \\ a(m_t) & ,r\leq t. \end{cases} \end{aligned}$$

Then the restriction $\tilde{Q}$ on $C([0,\infty);\mathbb{R}^d)$, derived from Q on $C([-r,\infty);\mathbb{R}^d)$, solves the local martingale problem for the coefficients b^t and a^t with start in $\xi(0)$ for the operator

$$(L^t_{b^t,a^t}f)(x) = \sum_{i=1}^d b_i^t(x)\frac{\partial f}{\partial x_i}(x(t)) + \frac{1}{2}\sum_{i,j=1}^d a_{ij}^t(x)\frac{\partial^2 f}{\partial x_i \partial x_j}(x(t)), \quad x \in C([0,\infty);\mathbb{R}^d).$$

Now we can apply Proposition 5.4.6 in Karatzas and Shreve [12]. By this proposition there exists a sextuple $(\Omega, \mathcal{F}, P, (\mathcal{F}_t), X, B)$ such that

$$\begin{cases} X(0) &= \xi(0) \\ X(t) &= X(0) + \int_0^t b^s(X)\,ds + \int_0^t \sigma^s(X)\,dB(s), \quad \sigma^t(\sigma^t)^T = a^t, \quad t \geq 0 \end{cases}$$

holds P-a.s and Law $(X(t) : t \geq 0) = \tilde{Q}$. Define the initial segment

$$X(u,\omega) := \xi(u), \quad -r \leq u \leq 0, \quad \omega \in \Omega.$$

Then for the sextuple $(\Omega, \mathcal{F}, P, (\mathcal{F}_t), X, B)$ it holds that

$$\begin{cases} X_0 &= \xi \\ X(t) &= X(0) + \int_0^t b(X_s)\,ds + \int_0^t \sigma(X_s)\,dB(s), \quad t \geq 0 \end{cases}$$

a.s. Furthermore it holds that Law $(X(t) : t \geq -r) = Q$. This completes the proof of the theorem. □

2.2.4 Corollary. *Weak uniqueness for the system (2.1.1) is equivalent to uniqueness of the local martingale problem.*

Proof. It suffices for the proof to assume that weak uniqueness holds for the system (2.1.1). Let Q_1 and Q_2 solve the local martingale problem. By Theorem 2.2.3 there exist solutions (X^1, B^1) and (X^2, B^2) such that

$$\text{Law}(X^1(t) : t \geq -q) = Q_1, \quad \text{Law}(X^2(t) : t \geq -q) = Q_2.$$

But by weak uniqueness the laws of X^1 and X^2 are the same, hence it follows that $Q_1 = Q_2$. □

2.3 The Martingale Problem in Discrete Time

In this section we shall construct a martingale problem on the canonical space in discrete time starting from transition probabilities. The setting is the following. For $R \in \mathbb{N}_0$ define the measurable space $(\Omega, \mathcal{M})$ by

$$\mathcal{N} := \{-R; \ldots; -1; 0; 1; 2; \ldots\}, \quad \Omega =: (\mathbb{R}^d)^{\mathcal{N}}, \quad \mathcal{M} := \bigotimes_{i\in\mathcal{N}} \mathcal{B}^d.$$

Define also the following measurable functions and sub-σ-algebras

$$X_n^\circ(\omega) := \omega(n), \quad \omega \in \Omega, \quad \mathcal{M}_n := \sigma(X_k^\circ : -R \le k \le n), \quad n \ge -R.$$

Furthermore consider a function p

$$\begin{aligned} p : (\mathbb{R}^d)^{R+1} \times \mathcal{B}^d &\longrightarrow [0,1] \\ (x_0, \dots, x_{-R}; A) &\mapsto p(x_0, \dots, x_{-R}; A) \end{aligned}$$

such that

1. $(x_0, \dots, x_{-R}) \mapsto p(x_0, \dots, x_{-R}; A)$ is measurable for all $A \in \mathcal{B}^d$.
2. $A \mapsto p(x_0, \dots, x_{-R}; A)$ is a probability measure on $\mathcal{B}^d$ for all $(x_0, \dots, x_{-R}) \in (\mathbb{R}^d)^{R+1}$.

The function p is called transition probability of order $(R+1)$. Define for all $f \in C_0^\infty(\mathbb{R}^d)$ the operator

$$(Af)(x_0, \dots, x_{-R}) := \int_{\mathbb{R}^d} (f(z) - f(x_0)) p(x_0, \dots, x_{-R}; dz), \quad (x_0, \dots, x_{-R}) \in (\mathbb{R}^d)^{R+1}.$$

Note that the integrand depends only on x_0, whereas the past values $(x_0, x_1, \dots, x_{-R})$ occur in the transition probability p. Fix an initial condition $\xi = (\xi_0, \dots, \xi_{-R}) \in (\mathbb{R}^d)^{R+1}$. We shall now give an analogue to Definition 2.2.2 in discrete time, in generalization to the case $R = 0$ in Stroock and Varadhan [28].

2.3.1 Definition. *A probability measure P_ξ on $(\Omega, \mathcal{F})$ solves the martingale problem associated with the operator A with start in $\xi = (\xi_0, \dots, \xi_{-R})$ if*

1. $P_\xi(X_k^\circ = \xi_k, -R \le k \le 0) = 1$.
2. $f(X_n^\circ) - \sum_{k=0}^{n-1} (Af)(X_k^\circ, \dots, X_{k-R}^\circ), n \in \mathbb{N}_0$ *is a $(\mathcal{M}_n, P_\xi)$-martingale for all* $f \in C_0^\infty(\mathbb{R}^d)$ $(\sum_{k=0}^{-1} := 0)$.

In this setting the martingale problem in discrete time is always well-posed, that means, there always exists a solution of the martingale problem, and this solution is unique. The following lemma constructs the solution. The proof is elementary and therefore omitted.

2.3.2 Lemma. *P_ξ solves the martingale problem for the operator A if and only if*

$$\begin{aligned} &P_\xi(X_{-R}^\circ \in \Gamma_{-R}, \dots, X_0^\circ \in \Gamma_0, X_1^\circ \in B_1, \dots, X_n^\circ \in B_n) \qquad (2.3.1) \\ &= 1_{\Gamma_{-R}}(\xi_{-R}) \cdot \dots \cdot 1_{\Gamma_0}(\xi_0) \int_{B_1} \dots \int_{B_n} p(x_{n-1}, \dots, x_{n-1-R}; dx_n) \dots p(\xi; dx_1) \end{aligned}$$

for all $\Gamma_{-R}, \dots, \Gamma_0, B_1, \dots, B_n, n \in \mathbb{N}$ with $x_k := \xi_k$ for $-R \le k \le 0$.

2.4 The Main Results

Regarding convergence of stochastic processes we shall follow the notations in Billingsley [2]. In this section we shall formulate and prove our main convergence results. All applications will rely on these results. Our setting is the following. Fix a real number $0 \le r < \infty$. We shall consider step lengths $h > 0$ for which

$$r^{(h)} := \frac{r}{h} \in \mathbb{N}_0. \tag{2.4.1}$$

In the sequel we shall always write "for $h > 0$" meaning "for all $h > 0$ such that $r/h \in \mathbb{N}_0$". For $h > 0$ we shall consider $\mathbb{R}^d$-valued series $\{X^{(h)}_{kh} : k \ge -r^{(h)}\}$ in discrete time with step length h given on any probability space $(\Omega, \mathcal{F}, P)$. Define for $h > 0$ the sub-σ-algebras

$$\mathcal{F}^{(h)}_{mh} := \sigma(X^{(h)}_{kh} : -r^{(h)} \le k \le m), \quad m \ge -r^{(h)}.$$

For $h > 0$ we are also given an initial function $\xi^{(h)} \in C([-r, 0]; \mathbb{R}^d)$. Starting with the series $\{X^{(h)}_{kh} : k \ge -r^{(h)}\}$ we assume that for $h > 0$ a stochastic process $X^{(h)}$ satisfies the following three conditions.

1. $P(X^{(h)}_{ih} = \xi^{(h)}(ih), -r^{(h)} \le i \le 0) = 1$.
2. $X^{(h)}$ is interpolated linearly between two discrete points $(mh, X^{(h)}_{mh})$ and $((m+1)h, X^{(h)}_{(m+1)h})$ for $m \ge -r^{(h)}$.
3. $P(X^{(h)}_{k+1} \in \Gamma | \mathcal{F}^{(h)}_{kh}) = p^{(h)}(X^{(h)}_{kh}, \dots, X^{(h)}_{(k-r^{(h)})h}; \Gamma), \quad k \in \mathbb{N}_0$.

Here

$$p^{(h)} : (\mathbb{R}^d)^{r^{(h)}+1} \times \mathcal{B}^d \longrightarrow [0, 1]$$

is a transition probability of order $(r^{(h)} + 1)$. It follows that $\{X^{(h)}_{kh}\}_{k \ge -r^{(h)}}$ is a discrete homogeneous Markov chain of rank $(r^{(h)} + 1)$ with transition probability $p^{(h)}$ which does not depend on $k \in \mathbb{N}_0$ and with start in

$$(\xi^{(h)}(-r), (\xi^{(h)}(-r + h)), \dots, \xi^{(h)}(0)).$$

The domain of the transition probabilities $p^{(h)}$ depends on h. For our purposes it is necessary to have them defined on a common domain. Define for $x \in C([-r, 0]; \mathbb{R}^d)$

$$p^{(h)}(x; \Gamma) := p^{(h)}(x(0), x(-h), x(-2h), \dots, x(-r); \Gamma), \quad \Gamma \in \mathbb{B}^d.$$

By this definition we have constructed transition probabilities

$$p^{(h)} : C([-r, 0]; \mathbb{R}^d) \times \mathcal{B}^d \longrightarrow [0, 1].$$

In slight abuse we have used the same notation $p^{(h)}$ for transition probabilities with domain $(\mathbb{R}^d)^{r^{(h)}+1}$ and $C([-r, 0]; \mathbb{R}^d)$. For simplification of notation we shall always write

$$l^{(h)}_{mh} X^{(h)} := l^{(h)}(X^{(h)}_{mh}, X^{(h)}_{(m-1)h}, \dots X^{(h)}_{(m-r^{(h)})h}), \quad m \in \mathbb{N}_0,$$

where $l^{(h)}$ stands for the procedure of linear interpolation. That means,

$$l^{(h)}(x(0), x(-h), \dots, x(-r)), \quad x \in C([-r, 0]; \mathbb{R}^d)$$

is a linearly interpolated, continuous function on $[-r, 0]$ with values $x(-ih)$ at time points $(-ih)$ for $0 \leq i \leq r^{(h)}$. Correspondingly, $l_{mh}^{(h)} X^{(h)}$ is a linearly interpolated, continuous function on $[mh - r, mh]$. Then it holds for every $x \in C([-r, 0]; \mathbb{R}^d)$ that

$$\Big(X_{mh}^{(h)} = x(0), X_{(m-1)h}^{(h)} = x(-h), \dots, X_{(m-r^{(h)})h}^{(h)} = x(-r)\Big) \Longleftrightarrow l_{mh}^{(h)} X^{(h)} = l^{(h)} x,$$

where $l^{(h)} x := l^{(h)}(x(0), x(-h), \dots, x(-r))$ for $x \in C([-r, 0]; \mathbb{R}^d)$. As a consequence of the condition

$$P(X_{k+1}^{(h)} \in \Gamma | \mathcal{F}_{kh}^{(h)}) = p^{(h)}(X_{kh}^{(h)}, \dots, X_{(k-r^{(h)})h}^{(h)}; \Gamma), \quad k \in \mathbb{N}_0$$

it holds for every integrable function g and for $P_{l_{mh}^{(h)} X^{(h)}}$-almost all $x \in C([-r, 0]; \mathbb{R}^d)$ that

$$E(g(X_{(m+1)h}^{(h)} - X_{mh}^{(h)}) | l_{mh}^{(h)} X^{(h)} = l^{(h)} x) = \int_{\mathbb{R}^d} g(z - x(0)) p^{(h)}(x; dz), \quad m \in \mathbb{N}_0.$$

Define further for each $x \in C([-r, 0]; \mathbb{R}^d)$ and each $\epsilon > 0$

$$\begin{aligned}
b^{(h)}(x) &:= \frac{1}{h} \int_{|z-x(0)| \leq 1} (z - x(0)) p^{(h)}(x; dz) \\
a^{(h)}(x) &:= \frac{1}{h} \int_{|z-x(0)| \leq 1} (z - x(0))(z - x(0))^T p^{(h)}(x; dz) \\
\triangle_\epsilon^{(h)}(x) &:= \frac{1}{h} \int_{|z-x(0)| > \epsilon} p^{(h)}(x; dz),
\end{aligned}$$

where M^T denotes the transpose of a matrix M (here a column vector). Note that the first two integrals are taken over a bounded domain of $\mathbb{R}^d$ since the integrals over the whole $\mathbb{R}^d$ need not exist. Those quantities have the following representation. If a truncation function ϕ on $\mathbb{R}^d$ is defined by $\phi(x) := x 1_{\{|x| \leq 1\}}$, then for all $m \in \mathbb{N}_0$ it holds for $P_{l_{mh}^{(h)} X^{(h)}}$-almost all $x \in C([-r, 0]; \mathbb{R}^d)$ that

$$\begin{aligned}
b^{(h)}(x) &= \frac{1}{h} E(\phi(X_{(m+1)h}^{(h)} - X_{mh}^{(h)}) | l_{mh}^{(h)} X^{(h)} = l^{(h)} x) \\
a^{(h)}(x) &= \frac{1}{h} E(\phi(X_{(m+1)h}^{(h)} - X_{mh}^{(h)}) \phi(X_{(m+1)h}^{(h)} - X_{mh}^{(h)})^T | l_{mh}^{(h)} X^{(h)} = l^{(h)} x) \\
\triangle_\epsilon^{(h)}(x) &= \frac{1}{h} P(|X_{(m+1)h}^{(h)} - X_{mh}^{(h)})| > \epsilon | l_{mh}^{(h)} X^{(h)} = l^{(h)} x).
\end{aligned}$$

It follows that $b^{(h)}$ is the truncated conditional expectation vector divided by the step length h. The matrix $a^{(h)}$ is the truncated conditional second moment matrix divided by h. Finally for each $\epsilon > 0$, $\triangle_\epsilon^{(h)}$ is the conditional probability to jump into a next state with distance greater than ϵ, also divided by h.

Our aim is to let h tend to zero. We are interested in the case that the quantities $b^{(h)}(x)$ and $a^{(h)}(x)$ behave properly as h tends to zero. Therefore we impose the following convergence conditions. Assume that there exist measurable functions b and a such that for every compact subset K of the Polish space $C([-r,0];\mathbb{R}^d)$

$$\sup_{x\in K} |b^{(h)}(x) - b(x)| \underset{h\to 0}{\longrightarrow} 0 \tag{2.4.2}$$

$$\sup_{x\in K} \|a^{(h)}(x) - a(x)\| \underset{h\to 0}{\longrightarrow} 0 \tag{2.4.3}$$

$$\sup_{x\in K} \triangle_\epsilon^{(h)}(x) \underset{h\to 0}{\longrightarrow} 0. \tag{2.4.4}$$

Finally define for each $f \in C_0^\infty(\mathbb{R}^d)$ and for each $x \in C([-r,0];\mathbb{R}^d)$ the operator

$$(A^{(h)}f)(x) := \int_{\mathbb{R}^d} (f(z) - f(x(0))p^{(h)}(x;\, dz).$$

The following purely analytical lemma shows that the operators $A^{(h)}f$ per unit of step length tend to the operator $L_{b,a}f$ in (2.2.1).

2.4.1 Lemma. *Under conditions (2.4.2), (2.4.3) and (2.4.4) it holds for every compact subset K of $C([-r,0];\mathbb{R}^d)$ that*

$$\sup_{x\in K} |\frac{1}{h}(A^{(h)}f)(x) - (L_{b,a}f)(x)| \underset{h\to 0}{\longrightarrow} 0.$$

Proof. Define for each $f \in C_0^\infty(\mathbb{R}^d)$ the function

$$H(z,x) := \sum_{i=1}^d (z-x)_i \frac{\partial f}{\partial z_i}(z) + \frac{1}{2}\sum_{i,j=1}^d (z-x)_i(z-x)_j \frac{\partial^2 f}{\partial z_i \partial z_j}(z), \quad x,z \in \mathbb{R}^d.$$

Then by Taylor's formula there is a constant $C_f < \infty$ such that

$$|f(z) - f(x) - H(z,x)| \le C_f |z-x|^3 \quad \forall x,z \in \mathbb{R}^d.$$

Let $L^{(h)}$ be the corresponding operator for the coefficients $b^{(h)}$ and $a^{(h)}$

$$(L^{(h)}f)(x) = \sum_{i=1}^d b_i^{(h)}(x)\frac{\partial f}{\partial x_i}(x(0)) + \frac{1}{2}\sum_{i,j=1}^d a_{ij}^{(h)}(x)\frac{\partial^2 f}{\partial x_i \partial x_j}(x(0)), \quad x \in C([-r,0];\mathbb{R}^d).$$

Then we obtain

$$\begin{aligned}\frac{1}{h}(A^{(h)}f)(x)-(L^{(h)}f)(x) &= \frac{1}{h}\int_{|z-x(0)|>1} f(z)-f(x(0))p^{(h)}(x;\,dz)\\ &+ \frac{1}{h}\int_{|z-x(0)|\le 1} f(z)-f(x(0))-H(z,x(0))p^{(h)}(x;\,dz),\end{aligned}$$

which gives

$$\begin{aligned}|\frac{1}{h}(A^{(h)}f)(x)-(L^{(h)}f)(x)| &\le \frac{1}{h}\int_{|z-x(0)|>1} |f(z)-f(x(0))|p^{(h)}(x;\,dz) \quad (2.4.5)\\ &+ \frac{1}{h}\int_{\epsilon<|z-x(0)|\le 1} C_f|z-x(0)|^3p^{(h)}(x;\,dz) \quad (2.4.6)\\ &+ \frac{1}{h}\int_{|z-x(0)|\le\epsilon} C_f|z-x(0)|^3p^{(h)}(x;\,dz), \quad (2.4.7)\end{aligned}$$

where $0<\epsilon<1$ is arbitrary. The expression in (2.4.5) is lower or equal than

$$\frac{1}{h}\int_{|z-x(0)|>1} 2\|f\|_\infty p^{(h)}(x;\,dz) = 2\|f\|_\infty \triangle_1^{(h)}(x), \quad x\in C([-r,0];\mathbb{R}^d),$$

which tends to zero uniformly on compacts of $C([-r,0];\mathbb{R}^d)$ by condition (2.4.4). The term in (2.4.6) is not greater than

$$\frac{1}{h}\int_{|z-x(0)|>\epsilon} C_f p^{(h)}(x;\,dz) = C_f\triangle_\epsilon^{(h)}(x), \quad x\in C([-r,0];\mathbb{R}^d),$$

which also tends to zero uniformly on compacts of $C([-r,0];\mathbb{R}^d)$ by condition (2.4.4). Finally, the expression in (2.4.7) does not extend

$$\frac{\epsilon}{h}\int_{|z-x(0)|\le 1} C_f|z-x(0)|^2p^{(h)}(x;\,dz) = \epsilon C_f\sum_{i=1}^d a_{ii}^{(h)}(x) \le \epsilon C_f K_a, \quad x\in C([-r,0];\mathbb{R}^d),$$

where the constant K_a depends on the uniform upper bound for $a_{ii}^{(h)}$ on the compact set K. Since ϵ was arbitrary, it follows that

$$\sup_{x\in K}|\frac{1}{h}(A^{(h)}f)(x)-(L^{(h)}f)(x)| \xrightarrow[h\to 0]{} 0.$$

Obviously, for fixed $f\in C_0^\infty(\mathbb{R}^d)$ the operators $(L^{(h)}f)(x)$ converge to $(L_{b,a}f)(x)$ uniformly on compacts of $C([-r,0];\mathbb{R}^d)$. Therefore the lemma has been shown. □

For the moment we will confine our interest to the case that the limit functions b and a are continuous and bounded. Recall the truncation function $\phi(x)=x1_{\{|x|\le 1\}}$ for $x\in\mathbb{R}^d$.

2.4.2 Theorem. *Assume that for $h > 0$ we are given a time series $(X^{(h)}_{mh})_{m\geq -r^{(h)}}$ such that with*

$$\begin{aligned} b^{(h)}(x) &= \frac{1}{h}E(\phi(X^{(h)}_{(m+1)h} - X^{(h)}_{mh})|l^{(h)}_{mh}X^{(h)} = l^{(h)}x) \\ a^{(h)}(x) &= \frac{1}{h}E(\phi(X^{(h)}_{(m+1)h} - X^{(h)}_{mh})\phi(X^{(h)}_{(m+1)h} - X^{(h)}_{mh})^T|l^{(h)}_{mh}X^{(h)} = l^{(h)}x) \\ \triangle^{(h)}_\epsilon(x) &= \frac{1}{h}P(|X^{(h)}_{(m+1)h} - X^{(h)}_{mh})| > \epsilon|l^{(h)}_{mh}X^{(h)} = l^{(h)}x) \end{aligned}$$

there exist continuous, bounded functions b and a such that for every compact set K of $C([-r,0];\mathbb{R}^d)$

$$\begin{aligned} \sup_{x\in K}|b^{(h)}(x) - b(x)| &\xrightarrow[h\to 0]{} 0 \\ \sup_{x\in K}\|a^{(h)}(x) - a(x)\| &\xrightarrow[h\to 0]{} 0. \end{aligned}$$

Assume that in addition

$$\sup_{h>0}\sup_{x\in C([-r,0];\mathbb{R}^d)}|b^{(h)}(x)| + \|a^{(h)}(x)\| < \infty.$$

Furthermore suppose that instead of (2.4.4) for every $\epsilon > 0$

$$\sup_{x\in C([-r,0];\mathbb{R}^d)}\triangle^{(h)}_\epsilon(x) \xrightarrow[h\to 0]{} 0. \tag{2.4.8}$$

The time series $(X^{(h)}_{mh})_{m\geq -r^{(h)}}$ is extended to a continuous process $X^{(h)}$ by linear interpolation. If $\xi^{(h)} \xrightarrow[h\to 0]{} \xi$, then the laws of $\{X^{(h)} : h > 0\}$ are tight, and every limit point solves the martingale problem associated with the functions b and a with start in ξ.

Proof. Firstly, we shall establish the tightness of the sequence

$$\{P^{(h)} : h > 0\} := \{\text{Law}(X^{(h)}) : h > 0\}.$$

Since by assumption $\xi^{(h)} \xrightarrow[h\to 0]{} \xi$, it suffices to prove tightness for the restrictions of $P^{(h)}$, which are given on $C([-r,\infty);\mathbb{R}^d)$, to $C([0,\infty);\mathbb{R}^d)$. It follows from Lemma 2.3.2 that

$$f(X^\circ_{nh}) - \sum_{k=0}^{n-1}(A^{(h)}f)(X^\circ_{kh},\ldots,X^\circ_{(k-r^{(h)})h}), \quad n\in\mathbb{N}_0$$

is a $(\mathcal{M}_{kh}, P^{(h)})$-martingale in discrete time. Using the arguments in the proof of Lemma 2.4.1 and the uniform boundedness of $b^{(h)}$ and $a^{(h)}$, it is easy to establish that

$$\sup_{h>0}\sup_{x\in C([-r,0];\mathbb{R}^d)}|\frac{1}{h}(A^{(h)}f)(x)| \leq D_f < \infty,$$

where the constant D_f depends on the bounds of f and its first two derivatives. Then, using the just established martingale property, a simple calculation shows that

$$f(X^\circ_{nh}) + D_f(n-1)h, \quad n \in \mathbb{N}_0$$

is a nonnegative $(\mathcal{M}_{kh}, P^{(h)})$-submartingale in discrete time. Now we can apply Theorem 1.4.11 in Stroock and Varadhan [28]. By this theorem $\{P^{(h)} : h > 0\}$ is tight on $C([0,\infty);\mathbb{R}^d)$ if for every $\epsilon > 0$ and $T > 0$

$$\sum_{0 \le jh \le T} P^{(h)}(|X^\circ_{(j+1)h} - X^\circ_{jh}| \ge \epsilon) \underset{h\to 0}{\longrightarrow} 0$$

$$\sup_{h>0} P^{(h)}(|X^\circ_0| \ge l) \underset{l\to\infty}{\longrightarrow} 0.$$

But in our case we have that

$$P^{(h)}(|X^\circ_{(j+1)h} - X^\circ_{jh}| \ge \epsilon) = E^{(h)}(h\triangle^{(h)}_\epsilon(l^{(h)}_{jh}X^\circ)) \le \sup_{x\in C([-r,0];\mathbb{R}^d)} h\triangle^{(h)}_\epsilon(l^{(h)}x),$$

so that we obtain

$$\sum_{0\le jh\le T} P^{(h)}(|X^\circ_{(j+1)h} - X^\circ_{jh}| \ge \epsilon) \le (T+1) \sup_{x\in C([-r,0];\mathbb{R}^d)} \triangle^{(h)}_\epsilon(x) \underset{h\to 0}{\longrightarrow} 0$$

by assumption (2.4.8). This proves the tightness of the family $\{P^{(h)} : h > 0\}$ on $C([0,\infty);\mathbb{R}^d)$ and thus on $\Omega = C([-r,\infty);\mathbb{R}^d)$.

Next we shall establish that every limit point of $\{P^{(h)} : h > 0\}$ solves the martingale problem. Fix two time points $0 \le t_1 < t_2$ and a bounded, $\mathcal{M}_{t_1}$-measurable, continuous function Φ on $C([-r,\infty);\mathbb{R}^d)$. Then it was already established that for each $h > 0$

$$E^{(h)}([f(X^\circ_{[\frac{t_2}{h}]h}) - f(X^\circ_{[\frac{t_1}{h}]h}) - \sum_{i=[\frac{t_1}{h}]}^{[\frac{t_2}{h}]-1} (A^{(h)}f)(l^{(h)}_{ih}X^\circ)]\Phi) = 0,$$

or equivalently

$$E^{(h)}([Z^{(h)} \circ X^\circ]\Phi) = 0, \quad X^\circ(m) = m, \quad m \in C([-r,\infty);\mathbb{R}^d)$$

with

$$Z^{(h)}(m) := f(m([\frac{t_2}{h}]h)) - f(m([\frac{t_1}{h}]h)) - \int_{[\frac{t_1}{h}]h}^{[\frac{t_2}{h}]h} (\frac{A^{(h)}}{h}f)(l^{(h)}m_{[\frac{u}{h}]h})\,du$$

for $m \in C([-r,\infty);\mathbb{R}^d)$. Let for $h > 0$

$$w_m(h) := \sup_{|s-t|\le h} |m(s) - m(t)|, \quad m \in C([-r,\infty);\mathbb{R}^d)$$

denote the modulus of continuity of the function m. By the Arzelá-Ascoli theorem a subset K of $C([-r,\infty);\mathbb{R}^d)$ is precompact if and only if

$$\sup_{m\in K}\sup_{s\geq -r}|m(s)|<\infty, \quad \sup_{m\in K}|w_m(h)| \xrightarrow[h\to 0]{} 0.$$

Let K be a compact subset of $C([-r,\infty);\mathbb{R}^d)$. We assume without loss of generality that for the initial conditions $\xi^{(h)}$ it holds that

$$\overline{\{{}^s\xi^{(h)}:0\leq s\leq r, h>0\}}\subset K, \quad {}^s\xi^{(h)}(u):=\xi^{(h)}(s+u), \quad -r\leq u\leq 0$$

since we have by assumption that $\xi^{(h)} \xrightarrow[h\to 0]{} \xi$. Then it follows from Arzelá-Ascoli and uniform continuity of f on compacts J of $\mathbb{R}^d$ that

$$\sup_{m\in K}|f(m([\frac{t}{h}]h))-f(m(t))|\leq \sup_{x,y\in J,|x-y|\leq \sup\limits_{m\in K} w_m(h)}|f(x)-f(y)| \xrightarrow[h\to 0]{} 0.$$

So we have shown that

$$f(m([\frac{t_i}{h}]h)) \xrightarrow[h\to 0]{} f(m(t_i)), \quad i=1,2$$

uniformly on compacts of $C([-r,\infty);\mathbb{R}^d)$. Since $\sup_{x\in C([-r,0];\mathbb{R}^d)}|(\frac{A^{(h)}}{h}f)(x)|<\infty$, we have that

$$\sup_{m\in C([-r,\infty);\mathbb{R}^d)}\left|\int_{[\frac{t_1}{h}]h}^{[\frac{t_2}{h}]h}(\frac{A^{(h)}}{h}f)(l^{(h)}m_{[\frac{u}{h}]h})\,du-\int_{t_1}^{t_2}(\frac{A^{(h)}}{h}f)(l^{(h)}m_{[\frac{u}{h}]h})\,du\right| \xrightarrow[h\to 0]{} 0.$$

Next one checks easily with the Arzelá-Ascoli theorem that the set

$$A=:\overline{\bigcup_{t_1\leq u\leq t_2}\{l^{(h)}m_{[\frac{u}{h}]h}:m\in K,h>0\}\cup\{m_u:m\in K\}}\subset C([-r,0];\mathbb{R}^d)$$

is compact. Therefore we obtain

$$\sup_{m\in K}\sup_{t_1\leq u\leq t_2}|(L_{b,a}f)(l^{(h)}m_{[\frac{u}{h}]h})-(L_{b,a}f)(m_u)|\leq$$
$$\sup_{x,x'\in A,\|x-x'\|_\infty\leq \sup_{m\in K}w_m(h)}|(L_{b,a}f)(x)-(L_{b,a}f)(x')| \xrightarrow[h\to 0]{} 0$$

by uniform continuity of b and a on the compact set A and by uniform continuity of the derivatives of f. Furthermore we see that

$$\sup_{m\in K}\sup_{t_1\leq u\leq t_2}|(\frac{A^{(h)}}{h}f)(l^{(h)}m_{[\frac{u}{h}]h})-(L_{b,a}f)(l^{(h)}m_{[\frac{u}{h}]h})|\leq$$
$$\sup_{x\in A}|(\frac{A^{(h)}}{h}f)(x)-(L_{b,a}f)(x)| \xrightarrow[h\to 0]{} 0,$$

where we have used Lemma 2.4.1 for the convergence in the last line. Thereby we have shown that $Z^{(h)} \xrightarrow[h\to 0]{} Z$ uniformly on compacts of $C([-r,\infty);\mathbb{R}^d)$ and boundedly with

$$Z(m) := f(m(t_2)) - f(m(t_1)) - \int_{t_1}^{t_2} (L_{b,a}f)(m_u)\,du, \quad m \in C([-r,\infty);\mathbb{R}^d).$$

Note that the function $Z(m)$ is bounded and continuous, since b and a were assumed to be bounded and continuous. Now we are able to complete the proof. We have that

$$0 = E^{(h)}(Z^{(h)}\Phi) = E^{(h)}((Z - Z^{(h)})\Phi) + E^{(h)}(Z\Phi). \tag{2.4.9}$$

The first summand tends to zero. Indeed, since $\{P^{(h)} : h > 0\}$ is tight, for every $\epsilon > 0$ there is a compact set B such that

$$P^{(h)}(B^C) \le \frac{\epsilon}{2M\|\Phi\|_\infty}, \quad |Z^{(h)}(m)| \le M, \quad |Z(m)| \le M$$

for all small h. Then the inequality

$$|E^{(h)}((Z - Z^{(h)})\Phi)| \le \|\Phi\|_\infty \sup_{m\in A} |Z(m) - Z^{(h)}(m)| + E^{(h)}(2M\|\Phi\|_\infty 1_{A^C})$$

shows that the first summand in (2.4.9) tends to zero as h to zero. Since $Z\Phi$ is bounded and continuous, the second summand in (2.4.9) tends to $E^Q(Z\Phi)$ by definition of weak convergence, where Q is an arbitrary limit point of $\{P^{(h)} : h > 0\}$. So we have shown that

$$E^Q(Z\Phi) = 0$$

for all $\mathcal{M}_{t_1}$-measurable, bounded, continuous functions Φ, thus for all $\mathcal{M}_{t_1}$-measurable, bounded functions Φ. Since it holds that $Q(X_0^\circ = \xi) = 1$, any limit point of $\{P^{(h)} : h > 0\}$ solves the martingale problem associated with b and a with start in ξ. □

2.4.3 Remark. *Assume that the obtained limit functions b and $a = \sigma\sigma^T$ are such that for the stochastic delay differential equation (2.1.1) with coefficients b and σ weak existence and weak uniqueness hold. Then the sequence of the laws of $X^{(h)}$ converges weakly to the law Q of the solution process with coefficients b and σ. This comes from the fact that by Corollary (2.2.4) there exists exactly one solution Q to the local martingale problem if weak uniqueness holds. Every convergent subsequence of the laws of $X^{(h)}$ converges weakly to any solution of the martingale problem, that is to Q. Therefore the whole sequence $\{Law(X^{(h)}) : h > 0\}$ converges weakly to Q.*

We would like to loosen the assumptions of continuity of the limit coefficients. Yan [31] treats the case of not necessarily continuous coefficients of stochastic ordinary differential equations in the framework of strong approximation. We shall do it for a certain class of stochastic delay differential equations in the framework of weak approximation. The crucial point in the proof of our next theorem with not necessarily continuous coefficients is a result for preservation of weak convergence under not necessarily continuous mappings, see Theorem 5.5 in Billingsley [2]: If a sequence of measures $Q^{(h)}$ converges weakly to the measure Q, and there are uniformly bounded functions $g^{(h)}$ and g such that

$$Q(x : g^{(h)}(x^{(h)}) \underset{h\to 0}{\not\longrightarrow} g(x), x^{(h)} \underset{h\to 0}{\longrightarrow} x) = 0,$$

then it holds that

$$\int g^{(h)}(x)\, dQ^{(h)}(x) \underset{h\to 0}{\longrightarrow} \int g(x)\, dQ(x).$$

2.4.4 Theorem. *Assume that for $h > 0$ we are given a time series $(X^{(h)}_{mh})_{m\geq -r^{(h)}}$ such that with*

$$\begin{aligned} b^{(h)}(x) &= \frac{1}{h} E(\phi(X^{(h)}_{(m+1)h} - X^{(h)}_{mh}) | l^{(h)}_{mh} X^{(h)} = l^{(h)} x) \\ a^{(h)}(x) &= \frac{1}{h} E(\phi(X^{(h)}_{(m+1)h} - X^{(h)}_{mh}) \phi(X^{(h)}_{(m+1)h} - X^{(h)}_{mh})^T | l^{(h)}_{mh} X^{(h)} = l^{(h)} x) \\ \triangle^{(h)}_{\epsilon}(x) &= \frac{1}{h} P(|X^{(h)}_{(m+1)h} - X^{(h)}_{mh})| > \epsilon | l^{(h)}_{mh} X^{(h)} = l^{(h)} x) \end{aligned}$$

there exist bounded measurable functions b and a such that

$$\begin{aligned} b^{(h)}(x^{(h)}) &\underset{h\to 0}{\longrightarrow} b(x) \\ a^{(h)}(x^{(h)}) &\underset{h\to 0}{\longrightarrow} a(x) \end{aligned}$$

for any sequence $x^{(h)}$ approximating x in the points where b and a are continuous. Assume that in addition

$$\sup_{h>0} \sup_{x\in C([-r,0];\mathbb{R}^d)} |b^{(h)}(x)| + \|a^{(h)}(x)\| < \infty.$$

Furthermore suppose that for every $\epsilon > 0$

$$\sup_{x\in C([-r,0];\mathbb{R}^d)} \triangle^{(h)}_{\epsilon}(x) \underset{h\to 0}{\longrightarrow} 0.$$

The time series $(X^{(h)}_{mh})_{m\geq -r^{(h)}}$ is extended to a continuous process $X^{(h)}$ by linear interpolation. If $\xi^{(h)} \underset{h\to 0}{\longrightarrow} \xi$, then the laws of $\{X^{(h)} : h > 0\}$ are tight. Assume that every limit point Q of $\{Law(X^{(h)}) : h > 0\}$ satisfies for each $T > 0$

$$Q\left(\int_0^T 1(X^{\circ}_u \in D_b \cup D_a)\, du = 0\right) = 1, \tag{2.4.10}$$

where D_b and D_a denote the sets of discontinuity points of the functions b and a. Then every limit point Q of $\{Law(X^{(h)}) : h > 0\}$ solves the martingale problem associated with the functions b and a with start in ξ.

Proof. The sequence

$$\{P^{(h)} : h > 0\} = \{\text{Law}(X^{(h)}) : h > 0\}.$$

is tight according to the proof of the preceding theorem. We emphasize once more that we did not use continuity of the functions b and a to prove tightness. For $m \in C([-r,\infty);\mathbb{R}^d)$ and fixed time points $t_1 \le t_2$ recall the notations

$$Z^{(h)}(m) = f(m([\frac{t_2}{h}]h)) - f(m([\frac{t_1}{h}]h)) - \int_{[\frac{t_1}{h}]h}^{[\frac{t_2}{h}]h} (\frac{A^{(h)}}{h} f)(l^{(h)} m_{[\frac{u}{h}]h})\, du$$

and

$$Z(m) = f(m(t_2)) - f(m(t_1)) - \int_{t_1}^{t_2} (L_{b,a} f)(m_u)\, du$$

from the proof of the preceding theorem. Our aim is to show that

$$0 = E^{(h)}(Z^{(h)}\Phi) \xrightarrow[h\to 0]{} E^Q(Z\Phi)$$

for every weak limit Q of the sequence $\{P^{(h)} : h > 0\}$ for every bounded, continuous, $\mathcal{M}_{t_1}$-measurable function Φ. Then Q will solve the martingale problem for b and a. Fix a sequence of functions $m^{(h)} \in C([-r,\infty);\mathbb{R}^d)$ and a function $m \in C([-r,\infty);\mathbb{R}^d)$ with

$$m^{(h)} \xrightarrow[h\to 0]{} m.$$

Since $(\frac{A^{(h)}}{h} f)(x)$ remains bounded in x and h we see that

$$\left| \int_{[\frac{t_1}{h}]h}^{[\frac{t_2}{h}]h} (\frac{A^{(h)}}{h} f)(l^{(h)} m^{(h)}_{[\frac{u}{h}]h})\, du - \int_{t_1}^{t_2} (\frac{A^{(h)}}{h} f)(l^{(h)} m^{(h)}_{[\frac{u}{h}]h})\, du \right| \xrightarrow[h\to 0]{} 0.$$

Since by assumption

$$\sup_{x \in C([-r,0];\mathbb{R}^d)} \triangle_\epsilon^{(h)}(x) \xrightarrow[h\to 0]{} 0,$$

we have by the proof of Lemma 2.4.1 that for each compact set $K \subset C([-r,0];\mathbb{R}^d)$

$$\sup_{x\in K} |(\frac{A^{(h)}}{h} f)(x) - (L^{(h)} f)(x)| \xrightarrow[h\to 0]{} 0$$

with

$$(L^{(h)} f)(x) = \sum_{i=1}^d b_i^{(h)}(x) \frac{\partial f}{\partial x_i}(x(0)) + \frac{1}{2} \sum_{i,j=1}^d a_{ij}^{(h)}(x) \frac{\partial^2 f}{\partial x_i \partial x_j}(x(0)), \quad x \in C([-r,0];\mathbb{R}^d).$$

In the proof of Lemma 2.4.1 we neither used continuity of the functions $b^{(h)}$ and $a^{(h)}$. Thereby we conclude that

$$\left| \int_{t_1}^{t_2} (\frac{A^{(h)}}{h} f)(l^{(h)} m^{(h)}_{[\frac{u}{h}]h})\, du - \int_{t_1}^{t_2} (L^{(h)} f)(l^{(h)} m^{(h)}_{[\frac{u}{h}]h})\, du \right| \xrightarrow[h \to 0]{} 0.$$

Fix a time point u with $t_1 \leq u \leq t_2$ and an index l with $1 \leq l \leq d$. Consider the coefficient b_l, all other are treated in the same way. If b_l is continuous in the point $m_u \in C([-r, 0]; \mathbb{R}^d)$, then it holds by the assumption of the theorem that

$$b_l^{(h)}(l^{(h)} m^{(h)}_{[\frac{u}{h}]h}) \xrightarrow[h \to 0]{} b_l(m_u).$$

If

$$\lambda^{b_l}_{t_1,t_2}(m) := \lambda(\{t_1 \leq u \leq t_2 : b_l \text{ is not continuous in } m_u\}) = 0,$$

then

$$\left| \int_{t_1}^{t_2} b_l^{(h)}(l^{(h)} m^{(h)}_{[\frac{u}{h}]h}) f_l'(m([\frac{u}{h}]h))\, du - \int_{t_1}^{t_2} b_l(m_u) f_l'(m(u))\, du \right| \xrightarrow[h \to 0]{} 0$$

by dominated convergence. Thus we have that

$$Z^{(h)}(m^{(h)}) \xrightarrow[h \to 0]{} Z(m),$$

and hence by boundedness of Z and Φ and continuity of Φ

$$Z^{(h)}(m^{(h)}) \Phi(m^{(h)}) \xrightarrow[h \to 0]{} Z(m)\Phi(m),$$

provided that

$$\lambda_{t_1,t_2}(m) := \lambda(\{t_1 \leq u \leq t_2 : m_u \in D_b \cup D_a\}) = 0.$$

But by assumption (2.4.10) we have for each $T > 0$ that

$$Q\left(\int_0^T 1(X_u^\circ \in D_b \cup D_a)\, du = 0 \right) = 1,$$

and thus $\lambda_{t_1,t_2}(m) = 0$ Q-a.s. Now we can apply Theorem 5.5. in Billingsley [2] and deduce that

$$0 = E^{(h)}(Z^{(h)}\Phi) \xrightarrow[h \to 0]{} E^Q(Z\Phi) = 0$$

as required. □

2.4.5 Remark. *In the case $r = 0$ a condition of the type*

$$Q\left(\int_0^T 1(X_u^\circ \in D_b \cup D_a)\, du = 0 \right) = 1$$

appears in the framework of strong approximation in Yan [31]. Practically, this condition may be difficult to verify since the law of the limit point Q is beforehand essentially unknown.

We seek therefore for suitable conditions on the sequence $\{P^{(h)}\} = \{\text{Law}(X^{(h)})\}$ itself which will imply relation (2.4.10). To this end we assume from now on that the limit functions b and a belong to the following special class of vector and matrix-valued functions ϕ with domain $C([-r,0];\mathbb{R}^d)$

$$\phi(x) = f((x(u_i))_{i\in I_\phi}), \quad x \in C([-r,0];\mathbb{R}^d) \tag{2.4.11}$$

for a finite index set $I_\phi \subset [-r,0]$ and a measurable function f with domain $(\mathbb{R}^d)^{I_\phi}$. We demand that

$$\{y \in (\mathbb{R}^d)^{I_\phi} : f \text{ is not continuous in } y\} \subset (\mathbb{R}^d)^{I_\phi}$$

is an at most countable set.

2.4.6 Example. *The one-dimensional function*

$$\phi_1(x) := sgn(x(-r)), \quad x \in C[-r,0]$$

belongs to the class in (2.4.11), since zero is the only discontinuity point of the function $f(y) = sgn(y)$ for $y \in \mathbb{R}$. However, regarded as a subset of $C[-r,0]$ the set of discontinuities of ϕ_1 is uncountable. Indeed, the set D_{ϕ_1} of discontinuity points consists of all functions $x \in C[-r,0]$ with $x(-r) = 0$, and there are uncountably many such functions x on the space $C[-r,0]$. The quadratic variation function

$$\phi_2(x) :=< x >= \lim \sum_i [x(u_{i+1}) - x(u_i)]^2, \quad x \in C[-r,0]$$

is not continuous in any point and does not belong to the class in (2.4.11).

Fix a function ϕ which belongs to the class in (2.4.11). We need one more technical restriction on the initial condition ξ with respect to ϕ. For each $0 \le s \le r$ define

$$A^s_\xi := \{x \in C([-r,0],\mathbb{R}^d) : x(u) = \xi(t+u), -r \le u \le -s\}.$$

This is the set of all continuous functions which equal the by s shifted initial condition ξ on the subinterval $[-r,-s]$. Now we impose the following condition

$$\lambda\{0 \le s \le r : A^s_\xi \cap D_\phi = \emptyset\} = 0. \tag{2.4.12}$$

Under this condition one has the following implication for time points T up to the length of memory r

$$Q(X^\circ_0 = \xi) = 1 \Longrightarrow Q\left(\int_0^T 1(X^\circ_u \in D_\phi)\, du = 0\right) = 1, \quad 0 \le T \le r.$$

2.4.7 Example. *To understand the nature of condition (2.4.12) better take again*

$$\phi(x) = sgn(x(-r)), \quad x \in C[-r, 0].$$

For

$$\xi_1(u) := u - r, \quad -r \leq u \leq 0$$

we have that

$$\begin{aligned} A^s_{\xi_1} \cap D_\phi &= \emptyset, \quad 0 < s \leq r \\ A^s_{\xi_1} \cap D_\phi &\neq \emptyset, \quad s = 0, \end{aligned}$$

hence condition (2.4.12) is met by ξ_1. On the contrary, for $\xi_2 :\equiv 0$ it follows that

$$Q\left(\int_0^T 1(X^\circ_u \in D_\phi)\, du = T\right) = 1, \quad T > 0$$

if Q starts in ξ_2, hence condition (2.4.12) is not met by ξ_2.

Now let Q be a limit point of $\{P^{(h)} : h > 0\}$ with start in ξ. Assume that the functions b and a belong to the class of functions ϕ in (2.4.11), and that they fulfil condition (2.4.12) for the initial condition ξ. Then we know for time points T up to the length of memory r that

$$Q\left(\int_0^T 1(X^\circ_u \in D_b \cup D_a)\, du = 0\right) = 1, \quad T \leq r.$$

Next we consider the case $T > r$. We obtain by Fubini

$$\begin{aligned} E^Q\left(\int_0^T 1(X^\circ_u \in D_b \cup D_a)\, du\right) &= \int_{C([-r,\infty);\mathbb{R}^d)} \left(\int_0^T 1(m_u \in D_b \cup D_a)\, du\right) dQ(m) \\ &= \int_0^T \left(\int_{C([-r,\infty);\mathbb{R}^d)} 1(m_u \in D_b \cup D_a)\, dQ(m)\right) du \\ &= \int_0^T Q(X^\circ_u \in D_b \cup D_a)\, du. \qquad (2.4.13) \end{aligned}$$

We analyze the integrand further. We fix one component b_l and use the special structure of b_l

$$b_l(x) = f_l((x(u_i))_{i \in I_{b_l}}), \quad x \in C[-r, 0].$$

We have

$$\begin{aligned} Q(X^\circ_u \in D_{b_l}) &= Q((X^\circ(u + u_i))_{i \in I_{b_l}} \in D_{f_l}) \\ &= \sum_{i \in I_{b_l}} \sum_k Q(X^\circ(u + u_i) = y^k_i), \qquad (2.4.14) \end{aligned}$$

where y_i^k denotes the i-th component of a discontinuity point $y^k \in (\mathbb{R}^d)^{I_{b_l}}$ of the function f_l. By assumption there are at most countably many y^k. If we know that for each $u > 0$ and $y \in \mathbb{R}^d$

$$Q(X^\circ(u) = y) = 0,$$

then it holds by (2.4.13) in combination with (2.4.14) that for all $T > 0$

$$Q\left(\int_0^T 1(X_u^\circ \in D_b \cup D_a)\, du = 0\right) = 1.$$

For each $\delta > 0$ the set $\{m : m(u) \in B_\delta(y)\} \subset C([-r, \infty); \mathbb{R}^d)$, where $B_\delta(y)$ denotes the open ball with radius δ around the point y, is open. Hence it is by Portmanteau's theorem enough to assume that for each $y \in \mathbb{R}^d$ and time point $u > 0$

$$\varlimsup_{h\to 0} P^{(h)}(X^\circ(u) \in B_\delta(y)) \xrightarrow[\delta\to 0]{} 0$$

to ensure that $Q(X^\circ(u) = y) = 0$ for every limit point Q of the sequence $\{P^{(h)} : h > 0\}$. Intuitively speaking, it is forbidden that the sequence of measures $P^{(h)}$ runs into a fix state with positive probability. For later reference we formulate our considerations in a lemma.

2.4.8 Lemma. *If the functions b and σ belong to the class of vector and matrix-valued functions ϕ with domain $C([-r, 0]; \mathbb{R}^d)$ which belong to the class of functions ϕ of the kind*

$$\phi(x) = f((x(u_i))_{i\in I_\phi}), \quad x \in C([-r, 0]; \mathbb{R}^d)$$

with at most countably many discontinuities of the function f, and it holds for the initial condition ξ that

$$\lambda\{0 \le s \le r : A_\xi^s \cap (D_b \cup D_a)\} = \emptyset\} = 0,$$

where for each $0 \le s \le r$

$$A_\xi^s = \{x \in C([-r, 0]; \mathbb{R}^d) : x(u) = \xi(t + u), -r \le u \le -s\},$$

then the following implication holds: If a sequence $\{P^{(h)} : h > 0\}$ of probability measures on $C([-r, \infty); \mathbb{R}^d)$ is tight and one has for all states $y \in \mathbb{R}^d$ and time points $u > 0$ that

$$\varlimsup_{h\to 0} P^{(h)}(X^\circ(u) \in B_\delta(y)) \xrightarrow[\delta\to 0]{} 0,$$

then it holds for all $T > 0$ that

$$Q\left(\int_0^T 1(X_u^\circ \in D_b \cup D_a)\, du = 0\right) = 1$$

for every limit point Q of the sequence $\{P^{(h)} : h > 0\}$.

Our next step is to replace boundedness of the limit functions b and a by local boundedness. To handle the case of local boundedness we shall perform a localization procedure. This is extremely technical but unavoidable. This is done in the case $r = 0$ in Stroock and Varadhan [28]. We shall do it also for the case $r > 0$. It is necessary to introduce a martingale problem for time points $s \geq 0$.

2.4.9 Definition. *A probability measure $P_{s,\xi}$ on $(\Omega, \mathcal{M})$ solves the martingale problem associated with b and a after time s if*

1. $P_{s,\xi}(\{X_s^\circ = \xi, X^\circ(u) = \xi(s-r), -r \leq u \leq s-r\}) = 1.$
2. $f(X^\circ(t)) - \int_s^t (L_{b,a}f)(X_u^\circ)\, du, t \geq s$ *is a* $(\mathcal{M}_t, P_\xi)$*-martingale for all* $f \in C_0^\infty(\mathbb{R}^d)$.

Let $s \geq 0$ be given and define the sub-σ-algebra

$$\mathcal{M}^{s-r} =: \sigma(X_u^\circ : u \geq s-r).$$

Consider a probability measure Q on $(\Omega, \mathcal{M}^{s-r})$ with the property

$$Q(X_s^\circ = \eta_s) = 1$$

for a given function η in $C([-r,s];\mathbb{R}^d)$. The following lemma states that Q can be uniquely continued onto $\mathcal{M}$.

2.4.10 Lemma. *There is a unique probability measure $\delta_\eta \otimes_{[s-r,s]} Q$ on $(\Omega, \mathcal{M})$ such that*

1. $(\delta_\eta \otimes_{[s-r,s]} Q)(X^\circ(u) = \eta(u) : -r \leq u \leq s) = 1.$
2. $(\delta_\eta \otimes_{[s-r,s]} Q)(A) = Q(A)$ *for all* $A \in \mathcal{M}^{s-r}$.

Proof. For m in $C([-r,\infty);\mathbb{R}^d)$ consider the map

$$m \mapsto \Phi(m) := (m(u))_{u \geq s-r}$$

with values in $C([s-r,\infty);\mathbb{R}^d)$. Let δ_η be the point mass on $C([-r,s];\mathbb{R}^d)$ and define

$$\tilde{Q}(\Gamma_1 \times \Gamma_2) := \delta_\eta(\Gamma_1)(Q \circ \Phi^{-1})(\Gamma_2), \quad \Gamma_1 \in \mathcal{M}_s, \Gamma_2 \in \mathcal{M}^{s-r}$$

on $\tilde{X} := C([-r,s];\mathbb{R}^d) \times C([s-r,\infty);\mathbb{R}^d)$. If one sets

$$X := \{(\alpha,\beta) \in C([-r,s];\mathbb{R}^d) \times C([s-r,\infty);\mathbb{R}^d) : \alpha_s = \beta_s\},$$

then X is a measurable subspace of $\tilde{X}$ and it holds by construction that $\tilde{Q}(X) = 1$. Now define $\Xi : X \longrightarrow C([-r,\infty);\mathbb{R}^d)$ by

$$\Xi((\alpha,\beta))(u) := \begin{cases} \alpha(u), & -r \leq u < s \\ \beta(u), & u \geq s. \end{cases}$$

Then the probability measure defined by

$$(\delta_\eta \otimes_{[s-r,s]} Q)(A) := \tilde{Q}(x : \Xi(x) \in A), \quad A \in \mathcal{M}$$

is the unique probability measures with the desired properties. □

Now consider a finite stopping time on $\Omega = C([-r,\infty);\mathbb{R}^d)$ and a mapping $m \mapsto Q_m$ into probability measures on $(\Omega, \mathcal{M})$ with

1. $m \mapsto Q_m(A)$ is $\mathcal{M}_\tau$-measurable for all $A \in \mathcal{M}$.
2. $Q_m(\tilde{m} : \tilde{m}_{\tau(m)} = m_{\tau(m)}) = 1$ for all $m \in \Omega$.

According to the preceding lemma it is possible to define for each $m \in \Omega$ the probability measure

$$\delta_m \otimes_{\tau(m)} Q_m := \delta_{\{m(u): -r \le u \le \tau(m)\}} \otimes_{[\tau(m)-r,\tau(m)]} (Q_m|_{\mathcal{M}^{\tau(m)-r}}),$$

where $Q_m|_{\mathcal{M}^{\tau(m)-r}}$ denotes the restriction of Q_m on $C([-r,\infty);\mathbb{R}^d)$ to $C([\tau(m)-r,\infty);\mathbb{R}^d)$. The next lemma shows that the family $\{\delta_m \otimes_{\tau(m)} Q_m : m \in \Omega\}$ serves as a a regular conditional distribution of a probability measure $P \otimes_{\tau(\cdot)} Q.$ on $(\Omega, \mathcal{M})$.

2.4.11 Lemma. *Given a probability measure P on $(\Omega, \mathcal{M})$, there is a unique probability measure $P \otimes_{\tau(\cdot)} Q.$ on $(\Omega, \mathcal{M})$ such that $P \otimes_{\tau(\cdot)} Q.$ equals P on $\mathcal{M}_\tau$ and $\{\delta_m \otimes_{\tau(m)} Q_m : m \in \Omega\}$ is a regular conditional distribution of $P \otimes_{\tau(\cdot)} Q.|\mathcal{M}_\tau$.*

Proof. Set for each $\Gamma \in \mathcal{M}$

$$(P \otimes_{\tau(\cdot)} Q.)(\Gamma) := \int_\Omega (\delta_m \otimes_{\tau(m)} Q_m)(\Gamma)\, dP(m).$$

For sets $B \in \mathcal{M}_t$ and $\Gamma \in \mathcal{M}$ we have by construction that for $\tau(m) \le t$

$$(\delta_m \otimes_{\tau(m)} Q_m)(B \cap \Gamma) = 1_B(m)(\delta_m \otimes_{\tau(m)} Q_m)(\Gamma),$$

and hence for $A \in \mathcal{M}_\tau$

$$(\delta_m \otimes_{\tau(m)} Q_m)(A \cap \tau \le t \cap \Gamma)1_{\{\tau(m)\le t\}} = 1_{\{A\cap\tau\le t\}}(m)(\delta_m \otimes_{\tau(m)} Q_m)(\Gamma)1_{\{\tau(m)\le t\}}.$$

Since τ is finite, we obtain by letting t to infinity

$$(\delta_m \otimes_{\tau(m)} Q_m)(A \cap \Gamma) = 1_A(m)(\delta_m \otimes_{\tau(m)} Q_m)(\Gamma), \quad A \in \mathcal{M}_\tau. \tag{2.4.15}$$

Therefore one sees setting $\Gamma = \Omega$ that for $A \in \mathcal{M}_\tau$

$$(P \otimes_{\tau(\cdot)} Q.)(A) = \int_\Omega (\delta_m \otimes_{\tau(m)} Q_m)(A)\, dP(m) = P(A),$$

which shows that $P \otimes_{\tau(\cdot)} Q.$ equals P on $\mathcal{M}_\tau$. Next we prove that $\{\delta_m \otimes_{\tau(m)} Q_m : m \in \Omega\}$ is a regular conditional distribution of $P \otimes_{\tau(\cdot)} Q.|\mathcal{M}_\tau$. Note that for each $\Gamma \in \mathcal{M}$ the map

$$m \mapsto (\delta_m \otimes_{\tau(m)} Q_m)(\Gamma)$$

is $\mathcal{M}_\tau$-measurable. Indeed, we have for $\Gamma = \{m : m(u_1) \in \Gamma_1, \ldots, m(u_n) \in \Gamma_n\}$ that

$$\begin{aligned}(\delta_m \otimes_{\tau(m)} Q_m)(\Gamma) &= 1_{[0,u_1]}(\tau(m))Q_m(\Gamma) \\ &+ \sum_{k=1}^{n-1} 1_{[u_k,u_{k+1})}(\tau(m))1_{\Gamma_1}(m(u_1)) \cdots 1_{\Gamma_k}(m(u_k)) \\ &\times Q_m(m : m(u_{k+1}) \in \Gamma_{k+1}, \ldots m(u_n) \in \Gamma_n) \\ &+ 1_{[u_n,\infty)}(\tau(m))1_{\Gamma_1}(m(u_1)) \cdots 1_{\Gamma_n}(m(u_n)),\end{aligned}$$

which is clearly $\mathcal{M}_\tau$-measurable. Finally by integrating (2.4.15) with respect to the measure $P \otimes_{\tau(\cdot)} Q.$ we obtain for $A \in \mathcal{M}_\tau$ and $\Gamma \in \mathcal{M}$

$$(P \otimes_{\tau(\cdot)} Q.)(A \cap \Gamma) = \int_A (\delta_m \otimes_{\tau(m)} Q_m)(\Gamma)\, d(P \otimes_{\tau(\cdot)} Q.)(m).$$

The lemma has been shown. □

From now on we shall assume uniqueness of the martingale problem of Definition 2.2.2. This is necessary to perform the localization procedure.

2.4.12 Lemma. *Let the martingale problem for the coefficients b and a be well-posed for every time $s \geq 0$ and for every initial condition ξ on $[s-r, s]$. Denote the family of solutions by $P_{s,\xi}$. Then for every $\Gamma \in \mathcal{M}$ the map*

$$(s, \xi) \mapsto P_{s,\xi}(\Gamma)$$

is measurable.

Proof. The measurability is a pure consequence of the uniqueness of the martingale problem. We refer the reader to Exercise 6.7.4. in Stroock and Varadhan [28], where the prove is given by a result on measurable inverses on Polish spaces. If $\mathcal{A}$ is the (measurable) set of the probability measures P which solve the martingale problem for b and a and start in some ξ, then $F(P) = \xi$ is a continuous map onto the Polish space $C([-r, 0]; \mathbb{R}^d)$. It is one to one since the solution of the martingale problem is unique. A one to one measurable map from any subset $\mathcal{A}$ of a Polish space onto another Polish space has a measurable inverse by a theorem on Polish spaces. □

The following theorem compares solutions of two martingale problems with different coefficients.

2.4.13 Theorem. *Let the martingale problem for the coefficients b and a be well-posed for every time $s \geq 0$ and for every initial condition ξ on $[s-r, s]$. Denote the family of solutions by $P_{s,\xi}$. Assume that there is a second set of coefficients $\bar{b}, \bar{a}$ such that $\bar{b} = b$ and $\bar{a} = a$ on some open bounded set $G \in C([-r, 0]; \mathbb{R}^d)$. Then for $\xi \in G$ it holds for any solution $\bar{P}$ of the martingale problem for $\bar{b}$ and $\bar{a}$ with initial condition ξ: $\bar{P}$ equals $P_{0,\xi}$ on $\mathcal{M}_\tau$, where $\tau = \inf\{t \geq 0 : m_t \notin G\}$.*

Proof. For $f \in C_0^\infty(\mathbb{R}^d)$ define

$$\theta(t,m) := f(m(t)) - \int_0^t (L_{b,a}f)(m_u)\,du, \quad t \geq 0, \quad m \in C([-r,\infty);\mathbb{R}^d).$$

We have by assumption that

$$f(m(t)) - \int_0^t (L_{\bar{b},\bar{a}}f)(m_u)\,du, \quad t \geq 0$$

is a $(\mathcal{M}_t, \bar{P})$-martingale. Hence the stopped process

$$\begin{aligned} f(m(t\wedge\tau)) \quad - \quad & \int_0^{t\wedge\tau(m)} (L_{\bar{b},\bar{a}}f)(m_u)\,du \\ = \quad & f(m(t\wedge\tau)) - \int_0^{t\wedge\tau(m)} (L_{b,a}f)(m_u)\,du = \theta(t\wedge\tau(m),m), \quad t \geq 0 \end{aligned}$$

is a $(\mathcal{M}_t, \bar{P})$-martingale. Now consider the family of probability measures

$$Q_m = P_{\tau(m),m_{\tau(m)}}, \quad m \in C([-r,\infty);\mathbb{R}^d).$$

By the preceding lemma the map $m \mapsto Q_m$ is measurable. Here we used uniqueness of the martingale problem. By construction the probability measure $\bar{P} \otimes_{\tau(\cdot)} P_{\tau(\cdot),\cdot}$ equals $\bar{P}$ on $\mathcal{M}_\tau$. Since $\theta(t\wedge\tau)$ is $\mathcal{M}_\tau$-measurable, $(\theta(t\wedge\tau), t \geq 0)$ is also a $(\mathcal{M}_t, \bar{P} \otimes_{\tau(\cdot)} P_{\tau(\cdot),\cdot})$-martingale. Furthermore, for $m \in C([-r,\infty);\mathbb{R}^d)$ we have by definition that

$$f(\tilde{m}(t)) - \int_{\tau(m)}^t (L_{b,a}f)(\tilde{m}_u)\,du, \quad t \geq \tau(m)$$

is a $(\mathcal{M}_t, P_{\tau(m),m_{\tau(m)}})$-martingale. This can be restated that for $m \in C([-r,\infty);\mathbb{R}^d)$

$$\begin{aligned} & f(\tilde{m}(t)) - \int_0^t (L_{b,a}f)(\tilde{m}_u)\,du \quad - \\ & \left(f(\tilde{m}(t\wedge\tau(m))) - \int_0^{t\wedge\tau(m)} (L_{b,a}f)(\tilde{m}_u)\,du \right) \quad = \quad \theta(t,\tilde{m}) - \theta(t\wedge\tau(m),\tilde{m}), \quad t \geq 0 \end{aligned}$$

is a $(\mathcal{M}_t, P_{\tau(m),m_{\tau(m)}})$-martingale with start in ξ. Then it follows from Theorem 1.2.10 in Stroock and Varadhan [28] that $(\theta(t), t \geq 0)$ is a $(\mathcal{M}_t, \bar{P} \otimes_{\tau(\cdot)} P_{\tau(\cdot),\cdot})$-martingale. Once again by uniqueness of the martingale problem we deduce that $\bar{P} \otimes_{\tau(\cdot)} P_{\tau(\cdot),\cdot} = P_{0,\xi}$. But by construction $\bar{P} \otimes_{\tau(\cdot)} P_{\tau(\cdot),\cdot}$ equals $\bar{P}$ on $\mathcal{M}_\tau$, and thus $\bar{P}$ equals $P_{0,\xi}$ on $\mathcal{M}_\tau$. The proof of the theorem is complete. □

We are now able to prove a convergence theorem for not necessarily bounded limit functions.

2.4.14 Theorem. *Assume that for $h > 0$ we are given a time series $(X_{mh}^{(h)})_{m\geq -r^{(h)}}$ such that with*

$$\begin{aligned} b^{(h)}(x) &= \frac{1}{h}E(\phi(X_{(m+1)h}^{(h)} - X_{mh}^{(h)})|l_{mh}^{(h)}X^{(h)} = l^{(h)}x) \\ a^{(h)}(x) &= \frac{1}{h}E(\phi(X_{(m+1)h}^{(h)} - X_{mh}^{(h)})\phi(X_{(m+1)h}^{(h)} - X_{mh}^{(h)})^t|l_{mh}^{(h)}X^{(h)} = l^{(h)}x) \\ \triangle_\epsilon^{(h)}(x) &= \frac{1}{h}P(|X_{(m+1)h}^{(h)} - X_{mh}^{(h)})| > \epsilon|l_{mh}^{(h)}X^{(h)} = l^{(h)}x) \end{aligned}$$

there exist continuous, locally bounded functions b and a such that for every compact set K of $C([-r,0];\mathbb{R}^d)$

$$\begin{aligned} \sup_{x\in K}|b^{(h)}(x) - b(x)| &\xrightarrow[h\to 0]{} 0 \\ \sup_{x\in K}\|a^{(h)}(x) - a(x)\| &\xrightarrow[h\to 0]{} 0. \end{aligned}$$

In addition assume that the following uniform local boundedness condition holds

$$\sup_{h>0}\sup_{\|x\|_\infty\leq R}|b^{(h)}(x)| + \|a^{(h)}(x)\| < \infty \quad \forall R > 0. \tag{2.4.16}$$

Assume furthermore that for every $\epsilon > 0$

$$\sup_{\|x\|_\infty\leq R}\triangle_\epsilon^{(h)}(x) \xrightarrow[h\to 0]{} 0 \quad \forall R > 0. \tag{2.4.17}$$

The time series $(X_{mh}^{(h)})_{m\geq -r^{(h)}}$ is extended to a continuous process $X^{(h)}$ by linear interpolation. Let the martingale problem for the functions b and a be well-posed for every initial condition. If $\xi^{(h)} \xrightarrow[h\to 0]{} \xi$, then the laws of $\{X^{(h)} : h > 0\}$ converge weakly to Q_ξ, where Q_ξ denotes the unique solution of the martingale problem for b and a with start in ξ.

Proof. Denote as in the preceding proofs

$$\{P^{(h)} : h > 0\} = \{\text{Law}(X^{(h)}) : h > 0\}.$$

For each $k \in \mathbb{N}$ choose a continuous function Ψ_k such that $0 \leq \Psi_k \leq 1$, $\Psi_k \equiv 1$ on $\{\|x\|_\infty \leq k\}$ and $\Psi_k \equiv 0$ on $\{\|x\|_\infty > k+1\}$. Such a function exists since the map $x \mapsto \|x\|_\infty$ is continuous, and we can choose a map of the form $\Psi_k(x) = \Phi_k(\|x\|_\infty)$. Now set for each $k \in \mathbb{N}$

$$p_k^{(h)}(x;\Gamma) := \Psi_k(x)p^{(h)}(x;\Gamma) + [1 - \Psi_k(x)]1_\Gamma(x(0)), \quad x \in C([-r,0];\mathbb{R}^d), \quad \Gamma \in \mathbb{B}^d,$$

where $p^{(h)}$ denotes the transition probability for $X^{(h)}$. Then for every measurable, integrable function f it holds that

$$\int_{\mathbb{R}^d} f(z)p_k^{(h)}(x;\,dz) = \Psi_k(x)\int_{\mathbb{R}^d} f(z)p^{(h)}(x;\,dz) + [1 - \Psi_k(x)]f(x(0)).$$

Furthermore we see that

$$\begin{aligned}\triangle_{k,\epsilon}^{(h)}(x) : &= \frac{1}{h}\int_{|z-x(0)|>\epsilon} p_k^{(h)}(x;\, dz)\\ &= \Psi_k(x)\frac{1}{h}\int_{|z-x(0)|>\epsilon} p^{(h)}(x;\, dz) \xrightarrow[h\to 0]{} 0\end{aligned}$$

uniformly on $C([-r,0];\mathbb{R}^d)$, since the last expression converges to zero uniformly on the set $\{\|x\|_\infty \le k{+}1\}$ in view of assumption (2.4.17), and $\Psi_k(x)$ vanishes for $\|x\|_\infty > k{+}1$. Then it follows from Theorem 2.4.2 that the sequence of the corresponding probability measures $\{P_k^{(h)} : h > 0\}$, derived from $p_k^{(h)}$, is tight. Furthermore define for $k \in \mathbb{N}$ the quantities $b_k^{(h)}$ in terms of $p_k^{(h)}$

$$\begin{aligned}b_k^{(h)}(x) &:= \frac{1}{h}\int_{|z-x(0)|\le 1}(z-x(0))p_k^{(h)}(x;\, dz)\\ &= \Psi_k(x)\frac{1}{h}\int_{|z-x(0)|\le 1}(z-x(0))p^{(h)}(x;\, dz)\\ &= \Psi_k(x)b^{(h)}(x), \quad x \in C([-r,0];\mathbb{R}^d).\end{aligned}$$

Then, using the assumptions of the theorem, $b_k^{(h)}$ remains uniformly bounded in h and x. The same holds for

$$\begin{aligned}a_k^{(h)}(x) &:= \frac{1}{h}\int_{|z-x(0)|\le 1}(z-x(0))(z-x(0))'p_k^{(h)}(x;\, dz)\\ &= \Psi_k(x)\frac{1}{h}\int_{|z-x(0)|\le 1}(z-x(0))(z-x(0))'p^{(h)}(x;\, dz)\\ &= \Psi_k(x)a^{(h)}(x), \quad x \in C([-r,0];\mathbb{R}^d).\end{aligned}$$

Since

$$b_k^{(h)}(x) \xrightarrow[h\to 0]{} \Psi_k(x)b(x)$$

and

$$a_k^{(h)}(x) \xrightarrow[h\to 0]{} \Psi_k(x)a(x)$$

uniformly on compacts of $C([-r,0];\mathbb{R}^d)$, and the limits on the right side are continuous and bounded, it follows from Theorem 2.4.2 that for fixed k every limit point of $\{P_k^{(h)} : h > 0\}$ solves the martingale problem for $\Psi_k(\cdot\,)b(\cdot\,)$ and $\Psi_k(\cdot\,)a(\cdot\,)$. Now define for each $k \in \mathbb{N}$ the open, bounded set

$$G_k =: \{x : \|x\|_\infty < k\}$$

and the stopping time

$$\tau_k(m) := \inf_{t\ge 0}\{m_t \notin G_k\} = \inf_{t\ge 0}\{|m(t)| \ge k\} \nearrow \infty, \quad m \in C([-r,\infty);\mathbb{R}^d).$$

Since $\Psi_k(\cdot)b(\cdot) = b(\cdot)$ and $\Psi_k(\cdot)a(\cdot) = a(\cdot)$ on G_k, it follows from Theorem 2.4.13 that any limit point Q_k of $\{P_k^{(h)} : h > 0\}$ equals P_ξ on $\mathcal{M}_{\tau_k}$. Here we used the assumption that the martingale problem is well-posed for every initial condition. Now let A be a set in $\mathcal{M}_{\tau_k}$. Then it holds for $s \in \mathbb{N}$ that

$$A \cap \{\tau_k \leq sh\} \in \mathcal{M}_{sh},$$

and therefore

$$E_k^{(h)}(1_{A\cap\tau_k\leq sh})$$

is an s-fold integral over $p_k^{(h)}$. On $\{\tau_k(m) \leq sh\}$ it holds for $x = m_u$ that $\|x\|_\infty < k$ for $u \leq sh$, so that $p_k^{(h)}(x;\Gamma)$ equals $p^{(h)}(x;\Gamma)$ on this set. Therefore we have that

$$E_k^{(h)}(1_{A\cap\tau_k\leq sh}) = E^{(h)}(1_{A\cap\tau_k\leq sh}).$$

Letting s to infinity on both sides, we obtain by dominated convergence

$$E_k^{(h)}(1_A) = E^{(h)}(1_A).$$

Therefore $P_k^{(h)}$ equals $P^{(h)}$ on $\mathcal{M}_{\tau_k}$ for all $h > 0$. We remind the reader of Lemma 11.1.1 in Stroock and Varadhan [28]: If

1. $P_k^{(h)} = P^{(h)}$ on $\mathcal{M}_{\tau_k}$ and $\tau_k \nearrow \infty$,
2. $\{P_k^{(h)} : h > 0\}$ is tight with limit point Q_k and $Q_k = Q_\xi$ on $\mathcal{M}_{\tau_k}$,

then the sequence $\{P^{(h)} : h > 0\}$ converges weakly to Q_ξ. It suffices to apply this lemma to obtain the desired weak convergence $P^{(h)} \xrightarrow[h\to 0]{} P_\xi$. The theorem has been shown. □

2.4.15 Remark. *We used in the proof that the coefficients b and a are locally bounded. Note that the continuity of the coefficients on the space $C([-r,0];\mathbb{R}^d)$ does not imply their local boundedness. This comes from the fact that the space $C([-r,0];\mathbb{R}^d)$, in contrast to $\mathbb{R}^d$, is not locally compact.*

The terms $b^{(h)}$ and $a^{(h)}$, used in all preceding theorems, are truncated expected values. In practical it may be difficult to compute integrals over a bounded domain. We present therefore alternative conditions which contain integrals over $\mathbb{R}^d$, but which are more restrictive. Assume that for one $\delta > 0$

$$\triangle_{\delta,i}^{(h)*}(x) := \frac{1}{h}\int_{\mathbb{R}^d} |z - x(0)|_i^{2+\delta} p^{(h)}(x;\, dz), \quad x \in C([-r,0];\mathbb{R}^d)$$

remains finite. Then for $x \in C([-r,0];\mathbb{R}^d)$ the following expressions are well-defined

$$\begin{aligned} b^{(h)*}(x) &:= \frac{1}{h}\int_{\mathbb{R}^d} (z - x(0))p^{(h)}(x;\, dz) \\ a^{(h)*}(x) &:= \frac{1}{h}\int_{\mathbb{R}^d} (z - x(0))(z - x(0))^T p^{(h)}(x;\, dz). \end{aligned}$$

Now we are able to formulate the next theorem with continuous coefficients. This version will play the most important role for applications in this thesis.

2.4.16 Theorem. *Assume that for $h > 0$ we are given a time series $(X^{(h)}_{mh})_{m\geq -r^{(h)}}$ such that with*

$$\begin{aligned} b^{(h)*}(x) &= \frac{1}{h}E((X^{(h)}_{(m+1)h} - X^{(h)}_{mh})|l^{(h)}_{mh}X^{(h)} = l^{(h)}x) \\ a^{(h)*}(x) &= \frac{1}{h}E((X^{(h)}_{(m+1)h} - X^{(h)}_{mh})(X^{(h)}_{(m+1)h} - X^{(h)}_{mh})^t|l^{(h)}_{mh}X^{(h)} = l^{(h)}x) \\ \triangle^{(h)*}_{\delta,i}(x) &= \frac{1}{h}E(|X^{(h)}_{(m+1)h} - X^{(h)}_{mh})|^{2+\delta}_i|l^{(h)}_{mh}X^{(h)} = l^{(h)}x) \end{aligned}$$

there exist continuous and locally bounded functions b and a such that

$$b^{(h)*}(x) \underset{h\to 0}{\longrightarrow} b(x) \tag{2.4.18}$$

$$a^{(h)*}(x) \underset{h\to 0}{\longrightarrow} a(x) \tag{2.4.19}$$

uniformly on compacts of $C([-r, 0]; \mathbb{R}^d)$ and locally uniformly boundedly. Furthermore assume that

$$\sup_{\|x\|_\infty \leq R} \triangle^{(h)*}_{\delta,i}(x) \underset{h\to 0}{\longrightarrow} 0 \quad \forall R > 0, i = 1, \dots, d. \tag{2.4.20}$$

The time series $(X^{(h)}_{mh})_{m\geq -r^{(h)}}$ is extended to a continuous process $X^{(h)}$ by linear interpolation. Let the martingale problem for the functions b and a be well-posed for every initial condition. If $\xi^{(h)} \underset{h\to 0}{\longrightarrow} \xi$, then the laws of $\{X^{(h)} : h > 0\}$ converge weakly to Q_ξ, where Q_ξ denotes the unique solution of the martingale problem for b and a with start in ξ.

Proof. It suffices to show that conditions (2.4.20), (2.4.18) and (2.4.19) imply the conditions of Theorem 2.4.14 and then to apply this theorem. We have for $x \in C([-r, 0]; \mathbb{R}^d)$ that

$$\begin{aligned} \triangle^{(h)}_\epsilon(x) &= \frac{1}{h}\int_{|z-x(0)|>\epsilon} p^{(h)}(x; dz) \leq \frac{1}{h}\int_{\mathbb{R}^d} \frac{|z - x(0)|^{2+\delta}}{\epsilon^{2+\delta}} p^{(h)}(x; dz) \\ &= \frac{1}{h}\frac{1}{\epsilon^{2+\delta}}\int_{\mathbb{R}^d} \left(\sum_{i=1}^d |(z - x(0))_i|^2\right)^{\frac{2+\delta}{2}} p^{(h)}(x; dz) \\ &\leq \frac{1}{\epsilon^{2+\delta}}\left[\sum_{i=1}^d (\triangle^{(h)*}_{\delta,i}(x))^{\frac{2}{2+\delta}}\right]^{\frac{2+\delta}{2}} \underset{h\to 0}{\longrightarrow} 0 \end{aligned}$$

uniformly on bounded sets of $C([-r,0];\mathbb{R}^d)$ by condition (2.4.20). Furthermore we obtain by Hölder

$$\frac{1}{h}\int_{|z-x(0)|>1} |z-x(0)|_i \, p^{(h)}(x;\,dz) \leq$$

$$\left(\frac{1}{h}\int_{\mathbb{R}^d} |z-x(0)|_i^{2+\delta} p^{(h)}(x;\,dz)\right)^{\frac{1}{2+\delta}} \left(\frac{1}{h}\int_{|z-x(0)|>1} p^{(h)}(x;\,dz)\right)^{\frac{1+\delta}{2+\delta}},$$

which tends to zero uniformly on compacts of $C([-r,0];\mathbb{R}^d)$ by condition (2.4.20). Finally, we obtain again by Hölder

$$\frac{1}{h}\int_{|z-x(0)|>1} |z-x(0)|_i^2 \, p^{(h)}(x;\,dz) \leq$$

$$\left(\frac{1}{h}\int_{\mathbb{R}^d} |z-x(0)|_i^{2+\delta} p^{(h)}(x;\,dz)\right)^{\frac{2}{2+\delta}} \left(\frac{1}{h}\int_{|z-x(0)|>1} p^{(h)}(x;\,dz)\right)^{\frac{\delta}{2+\delta}},$$

which tends to zero uniformly on compacts of $C([-r,0];\mathbb{R}^d)$ by condition (2.4.20). It is obvious that the limit relations just shown imply the desired result. □

We end this section with a generalization on initial conditions. In all convergence theorems we considered deterministic initial conditions $\xi^{(h)}$ of the processes $X^{(h)}$. It is useful to replace them by random ones $\rho^{(h)}$. Denote by $X^{(h)}_{\rho^{(h)}}$ the time series which has $\rho^{(h)}$ as initial condition. If the martingale problem for b and a is well-posed for every initial condition ξ (denote the solution by Q_ξ), and μ is a distribution on $C([-r,0];\mathbb{R}^d)$, then the well-defined probability measure

$$Q(\Gamma) := \int_{C([-r,0];\mathbb{R}^d)} Q_\xi(\Gamma)\, d\mu(\xi), \quad \Gamma \in \mathcal{M}$$

solves the martingale problem for b and a with initial distribution μ. Assume that we have shown for every sequence of deterministic initial conditions $\xi^{(h)}$, tending to ξ, that $\{X^{(h)}_{\xi^{(h)}} : h > 0\}$ converges weakly to Q_ξ, where Q_ξ starts in ξ. If we know that the distributions of $\rho^{(h)}$ converge to a probability measure μ on $C([-r,0];\mathbb{R}^d)$, under which conditions can we deduce that the sequence $\{X^{(h)}_{\rho^{(h)}}\}$ converges in distribution to Q_μ, where the initial law of Q_μ is μ? The next theorem answers the question. We introduce the notation

$$\epsilon := \{\epsilon_i : i \in \mathbb{N}\}$$

for an arbitrary random sequence $\{\epsilon_i : i \in \mathbb{N}\}$, which is typically the driving force of the process $X^{(h)}$.

2.4.17 Theorem. *Let for $h > 0$ the process $X^{(h)}_{\rho^{(h)}}$ be defined in terms of the initial condition $\rho^{(h)}$ and the driving force ϵ*

$$X^{(h)}_{\rho^{(h)}} = \Phi^{(h)}(\rho^{(h)}, \epsilon),$$

where $(\rho^{(h)}, \epsilon)$ *are assumed to be independent. Furthermore assume that* $\xi \mapsto Q_\xi$ *is continuous in the initial condition:*

$$E_{\xi^{(h)}}(f) \xrightarrow[h\to 0]{} E_\xi(f) \quad \textit{for} \quad \xi^{(h)} \xrightarrow[h\to 0]{} \xi$$

for every bounded, continuous function f*. If*

$$\textit{Law}(X^{(h)}_{\xi^{(h)}}) \xrightarrow[h\to 0]{} Q_\xi \quad \textit{for} \quad \xi^{(h)} \xrightarrow[h\to 0]{} \xi$$

and $\textit{Law}(\rho^{(h)}) \xrightarrow[h\to 0]{} \mu$ *on* $C([-r,0];\mathbb{R}^d)$*, then the laws of* $\{X^{(h)}_{\rho^{(h)}} : h > 0\}$ *converge weakly to* Q_μ*.*

Proof. The assumed independence of $(\rho^{(h)}, \epsilon)$ yields for every bounded measurable f the relation

$$E_{X^{(h)}_{\rho^{(h)}}}(f) = \int\limits_{C([-r,0];\mathbb{R}^d)} E^{(h)}_\xi(f)\, d\mu^{(h)}(\xi),$$

where $\mu^{(h)}$ denotes the law of $\rho^{(h)}$. Our aim is to show that for every bounded, continuous function f

$$E_{X^{(h)}_{\rho^{(h)}}}(f) = \int\limits_{C([-r,0];\mathbb{R}^d)} E^{(h)}_\xi(f)\, d\mu^{(h)}(\xi) \xrightarrow[h\to 0]{} \int\limits_{C([-r,0];\mathbb{R}^d)} E_\xi(f)\, d\mu(\xi) = E_Q(f).$$

By assumption of the theorem the sequence $\{\mu^{(h)} : h > 0\}$ converges weakly to μ. Since by boundedness of f the integrands are uniformly bounded in h, it suffices by Theorem 5.5 in Billingsley [2] to show that

$$E^{(h)}_\xi(f) \xrightarrow[h\to 0]{} E_\xi(f) \tag{2.4.21}$$

uniformly on compacts of $C([-r,0];\mathbb{R}^d)$. For a sequence $\xi^{(h)}$ converging to ξ combine the assumed relations

$$E^{(h)}_{\xi^{(h)}}(f) \xrightarrow[h\to 0]{} E_\xi(f), \quad E_{\xi^{(h)}}(f) \xrightarrow[h\to 0]{} E_\xi(f)$$

to obtain uniform convergence in (2.4.21). The theorem has been shown. □

2.5 Applications

2.5.1 Approximation of a Given Stochastic Delay Diff. Equation

Assume that we are given the $d-$dimensional autonomous SDDE with length of memory $r \geq 0$

$$\begin{cases} X_0 &= \xi \\ dX(t) &= b(X_t)\,dt + \sigma(X_t)\,dB(t), \quad t \geq 0 \end{cases} \tag{2.5.1}$$

with deterministic initial condition ξ on $C([-r,0];\mathbb{R}^d)$, where

$$X_t = X(t+u), \quad -r \le u \le 0$$

denotes the function segment. To ensure that we deal with real martingales instead of local martingales, for this and all other delay equations in this subsection we shall assume the following integrability condition

$$E\left|\int_0^t \sigma(X_s)\,dB(s)\right| < \infty, \quad t \ge 0. \tag{2.5.2}$$

We shall now approximate this system weakly by a sequence of autoregressive time series $(X^{(h)}_{mh})_{m\ge -r^{(h)}}$. Thereby consider only those h for which $r^{(h)} = r/h \in \mathbb{N}_0$. Recall the notation $l^{(h)}_{mh}X^{(h)} := l^{(h)}(X^{(h)}_{mh},\ldots,X^{(h)}_{(m-r^{(h)})h})$.

2.5.1 Theorem. *Assume that weak existence and weak uniqueness hold for the system (2.5.1) for every initial condition, where the coefficients b and σ are assumed to be locally bounded and continuous. Let*

$$\epsilon = \{\epsilon_{m,j} : m \in \mathbb{N}, j = 0,\ldots,n\}$$

be a sequence of i.i.d. variables on some probability space with

$$E(\epsilon_1) = 0, \quad E(\epsilon_1\epsilon_1^T) = I_n, \quad E(|\epsilon_{1,1}|^{2+\delta}) < \infty$$

for some $\delta > 0$, where I_n denotes the identity matrix in dimension n. Let for $h > 0$ the discrete $d-$dimensional time series $(X^{(h)}_{mh})_{m\ge -r^{(h)}}$ be defined by

$$\begin{cases} X^{(h)}_{mh} &= \xi^{(h)}(mh), \quad m = -r^{(h)},\ldots,0 \\ X^{(h)}_{(m+1)h} &= X^{(h)}_{mh} + b^{(h)}(l^{(h)}_{mh}X^{(h)})h + \sigma^{(h)}(l^{(h)}_{mh}X^{(h)})\sqrt{h}\epsilon_{m+1}, \quad m \in \mathbb{N}_0 \end{cases}$$

for some functions $b^{(h)}$ and $\sigma^{(h)}$ with domain $C([-r,0];\mathbb{R}^d)$. The scheme written out means that for $i = 1,\ldots,d$

$$X^{i,(h)}_{(m+1)h} = X^{i,(h)}_{mh} + b^{(h)}_i(l^{(h)}_{mh}X^{(h)})h + \sum_{j=1}^n \sigma^{(h)}_{i,j}(l^{(h)}_{mh}X^{(h)})\sqrt{h}\epsilon_{m+1,j}, \quad m \in \mathbb{N}_0.$$

The time series $(X^{(h)}_{mh})_{m\ge -r^{(h)}}$ is extended to a continuous process $X^{(h)}$ by linear interpolation. If

$$\xi^{(h)} \xrightarrow[h\to 0]{} \xi$$

and

$$\begin{aligned} b^{(h)}(x) &\xrightarrow[h\to 0]{} b(x) \\ \sigma^{(h)}(x) &\xrightarrow[h\to 0]{} \sigma(x) \end{aligned}$$

uniformly on compacts of $C([-r,0];\mathbb{R}^d)$ and uniformly locally bounded, then the processes $X^{(h)}$ converge weakly to X, where X is the unique weak solution of the system (2.5.1) with initial value ξ.

Proof. The proof is a straightforward application of Theorem 2.4.16. Define the σ-algebra

$$\mathcal{F}_{kh}^{(h)} := \sigma(X_t^{(h)} : -r \le t \le kh), \quad k \ge -r^{(h)}.$$

Then we have almost surely by independence of the sequence ϵ that

$$\begin{aligned} P(X_{(m+1)h}^{(h)} \in \Gamma | \mathcal{F}_{mh}^{(h)}) &= P(\begin{pmatrix} b^{(h)}(l_{mh}^{(h)} X^{(h)})h \\ +\sigma^{(h)}(l_{mh}^{(h)} X^{(h)})\sqrt{h}\epsilon_{m+1} \end{pmatrix} \in \Gamma - \{X_{mh}^{(h)}\} | \mathcal{F}_{mh}) \\ &= p^{(h)}(l_{mh}^{(h)} X^{(h)}; \Gamma), \quad m \in \mathbb{N}_0, \end{aligned}$$

where for $x \in C([-r,0];\mathbb{R}^d)$ and $\Gamma \in \mathbb{B}^d$ the transition probability $p^{(h)}$ is defined by

$$p^{(h)}(x;\Gamma) := P(b^{(h)}(l^{(h)}x)h + \sigma^{(h)}(l^{(h)}x)\sqrt{h}\epsilon_1 \in \Gamma - \{x(0)\}). \tag{2.5.3}$$

As a consequence of (2.5.3) it holds for every measurable, integrable g and for $P_{l_{mh}^{(h)} X^{(h)}}$-almost all $x \in C([-r,0];\mathbb{R}^d)$ that

$$\int_{\mathbb{R}^d} g(z - x(0)) p^{(h)}(x;\, dz) = E(g(b^{(h)}(l^{(h)}x)h + \sigma^{(h)}(l^{(h)}x)\sqrt{h}\epsilon_1)).$$

Therefore we obtain

$$\begin{aligned} \triangle_{\delta,i}^{(h)*}(x) &= \frac{1}{h} E(|b^{(h)}(l^{(h)}x)h + \sigma^{(h)}(l^{(h)}x)\sqrt{h}\epsilon_1|_i^{2+\delta}) \\ &\le \frac{1}{h}\left(|b_i^{(h)}(l^{(h)}x)|h + \sum_{j=1}^r E(|\sigma_{i,j}^{(h)}(l^{(h)}x)\sqrt{h}\epsilon_{1,j}|^{2+\delta})^{\frac{1}{2+\delta}}\right)^{2+\delta} \\ &= \left(|b_i^{(h)}(l^{(h)}x)|h^{\frac{1+\delta}{2+\delta}} + \sum_{j=1}^r |\sigma_{i,j}^{(h)}(l^{(h)}x)| \|\epsilon_{1,j}\|_{2+\delta} h^{\frac{\delta/2}{2+\delta}}\right)^{2+\delta} \xrightarrow[h \to 0]{} 0 \end{aligned}$$

uniformly on bounded sets of $C([-r,0];\mathbb{R}^d)$. Furthermore we have that

$$\begin{aligned} b^{(h)*}(x) &= E(b^{(h)}(l^{(h)}x) + \sigma^{(h)}(l^{(h)}x)(\sqrt{h}/h)\epsilon_1) = b^{(h)}(l^{(h)}x) \xrightarrow[h \to 0]{} b(x) \\ a^{(h)*}(x) &= E((b^{(h)}(l^{(h)}x)\sqrt{h} + \sigma^{(h)}(l^{(h)}x)\epsilon_1)(b^{(h)}(l^{(h)}x)\sqrt{h} + \sigma^{(h)}(l^{(h)}x)\epsilon_1)^T) \\ &= h(b^{(h)} \cdot b^{(h)T})(l^{(h)}x) + (\sigma^{(h)} \cdot \sigma^{(h)T})(l^{(h)}x) \xrightarrow[h \to 0]{} \sigma\sigma^T(x) \end{aligned}$$

uniformly on compact sets of $C([-r,0];\mathbb{R}^d)$ and uniformly locally bounded. This comes from the fact that for a compact set K of $C([-r,0];\mathbb{R}^d)$ the set

$$\tilde{K} := \overline{\{l^{(h)}x : x \in K, h > 0\}}$$

is compact, and it holds for example for the function b that

$$\sup_{x \in K} |b^{(h)}(l^{(h)}x) - b(x)| \le \sup_{x \in \tilde{K}} |b^{(h)}(l^{(h)}x) - b(l^{(h)}x)| + \sup_{x \in K} |b(l^{(h)}x) - b(x)| \xrightarrow[h \to 0]{} 0.$$

Due to the integrability condition and weak uniqueness for the system (2.5.1) the martingale problem for b and $a = \sigma\sigma^T$ is well-posed. Now we obtain the assertion by applying Theorem 2.4.16. □

2.5.2 Remark. *We imposed the integrability condition (2.5.2) to refer to solutions of the martingale problem instead of the local martingale problem. A version of Theorems 2.4.14 and 2.4.16 supposing uniqueness of the local martingale problem is straightforward.*

The process

$$Y_t^{(h)} = Y_{mh}^{(h)} + \sigma(Y_{mh}^{(h)})(B(t) - B(mh)), \quad mh < t \leq (m+1)h$$

is called Euler scheme in numerical mathematics. Note that $Y^{(h)}$ is not interpolated linearly between to time points mh and $(m+1)h$. Rather the values of the Brownian motion for all time points $t \geq 0$ enter in $Y^{(h)}$. Only the state y is discretized, the time t is not. The series

$$\begin{cases} Z_{(m+1)h}^{(h)} &= Z_{mh}^{(h)} + \sigma(Z_{mh}^{(h)})(B((m+1)h) - B(mh)) \\ Z_t^{(h)} &= Z_{[\frac{t}{h}]h}^{(h)} \end{cases}$$

is called a discretized Euler scheme. Then $Z^{(h)}$ is a random variable with values in the space of right-continuous functions which is defined in terms of B only at discrete time points. Thus the approximating processes $X^{(h)}$ in Theorem 2.5.1 resemble discretized Euler schemes with two modifications. Firstly, between two grid points it is interpolated linearly to construct random variables on the space of continuous functions. Secondly, the driving sequence ϵ need not necessarily be distributed normally, but it is only required that it is centered with variance 1 and has finite absolute $(2+\delta)$-moments for some $\delta > 0$.

The stochastic differential delay equation

$$\begin{cases} X_0 &= \xi \\ dX(t) &= b(X_t, t)\, dt + \sigma(X_t, t)\, dB(t), \quad t \geq 0 \end{cases} \tag{2.5.4}$$

has time-dependent coefficients b and σ and is therefore a generalization of system (2.5.1). We assume that b and σ are locally bounded and continuous in (x,t). Defining for each $x \in C([-r,0];\mathbb{R}^d)$ and $t \geq 0$ the time-dependent operator

$$(L_{b,a}^t f)(x) := \sum_{i=1}^d b_i(x,t) \frac{\partial f}{\partial x_i}(x(0)) + \frac{1}{2} \sum_{i,j=1}^d a_{ij}(x,t) \frac{\partial^2 f}{\partial x_i \partial x_j}(x(0)), \quad a = \sigma\sigma^T$$

one sees that

$$f(X^\circ(t)) - \int_0^t (L_{b,a}^u f)(X_u^\circ)\, du, \quad t \geq 0$$

is a $(\mathcal{M}_t, Q_\xi)$-martingale for all $f \in C_0^\infty(\mathbb{R}^d)$, where Q_ξ is the distribution of the solution process in (2.5.4). Now we shall approximate system (2.5.4) weakly. Let for $h > 0$ the d-dimensional time series $(X^{(h)}_{mh})_{m\geq -r^{(h)}}$ be defined by

$$\begin{cases} X^{(h)}_{mh} &= \xi^{(h)}(mh), \quad -r^{(h)} \leq m \leq 0 \\ X^{(h)}_{(m+1)h} &= X^{(h)}_{mh} + b^{(h)}(l^{(h)}_{mh}X^{(h)}, mh)h + \sigma^{(h)}(l^{(h)}_{mh}X^{(h)}, mh)\sqrt{h}\epsilon_{m+1}, \quad m \in \mathbb{N}_0, \end{cases}$$

where the sequence ϵ has the properties as above. The time series $(X^{(h)}_{mh})_{m\geq -r^{(h)}}$ is extended to a continuous process $X^{(h)}$ by linear interpolation. The transition probabilities with domain $C([-r,0];\mathbb{R}^d) \times \mathcal{B}^d$ for the schemes $(X^{(h)}_{mh})_{m\geq -r^{(h)}}$ become time-dependent

$$p^{(h)}_{mh}(x;\Gamma) = P(b^{(h)}(l^{(h)}x, mh)h + \sigma^{(h)}(l^{(h)}x, mh)\sqrt{h}\epsilon_1 \in \Gamma - \{x(0)\}).$$

The following quantities defined for $x \in C([-r,0];\mathbb{R}^d)$ and $t \geq 0$

$$\begin{aligned} b^{(h)*}(x,t) &:= \frac{1}{h}\int_{\mathbb{R}^d} (z - x(0)) p^{(h)}_{[\frac{t}{h}]h}(x;\, dz) \\ a^{(h)*}(x,t) &:= \frac{1}{h}\int_{\mathbb{R}^d} (z - x(0))(z - x(0))^T p^{(h)}_{[\frac{t}{h}]h}(x;\, dz) \end{aligned}$$

become time-dependent too. Repeating the arguments of the proof in Theorem 2.5.1 one sees that if

$$\begin{aligned} b^{(h)}(x,t) &\underset{h\to 0}{\longrightarrow} b(x,t) \\ \sigma^{(h)}(x,t) &\underset{h\to 0}{\longrightarrow} \sigma(x,t) \end{aligned}$$

uniformly on compacts of $C([-r,0];\mathbb{R}^d) \times \mathbb{R}_+$ and uniformly locally bounded, then

$$\begin{aligned} b^{(h)*}(x,t) &\underset{h\to 0}{\longrightarrow} b(x,t) \\ a^{(h)*}(x,t) &\underset{h\to 0}{\longrightarrow} (\sigma\sigma^T)(x,t) \end{aligned}$$

uniformly on compacts of $C([-r,0];\mathbb{R}^d) \times \mathbb{R}_+$ and uniformly locally bounded. If $\xi^{(h)} \underset{h\to 0}{\longrightarrow} \xi$, then it follows from the obvious time-dependent modification of Theorem 2.4.16 that the processes $X^{(h)}$ converge weakly to the solution process X in (2.5.4).

In the approximating time series $(X^{(h)}_{mh})_{m\geq -r^{(h)}}$ of Theorem 2.5.1 occur terms of the form

$$b^{(h)}(l^{(h)}(X^{(h)}_{mh}, \ldots, X^{(h)}_{(m-r^{(h)})h})), \quad \sigma^{(h)}(l^{(h)}(X^{(h)}_{mh}, \ldots, X^{(h)}_{(m-r^{(h)})h}))$$

which in principle may be difficult to compute, since the argument is a function on $C([-r,0];\mathbb{R}^d)$. This difficulty may be overcome if we have a delay equation with point delay. Then the coefficients have the cylindric structure

$$b(x) = \bar{b}(x(u_1), \ldots, x(u_n)), \quad \sigma(x) = \bar{\sigma}(x(u_1), \ldots, x(u_n)), \quad x \in C([-r,0];\mathbb{R}^d)$$

for time points $-r \leq u_n \ldots \leq u_1 \leq 0$ and functions $\bar{b}$ and $\bar{\sigma}$ from n variables in $\mathbb{R}^d$. In this case the terms

$$b^{(h)}(l^{(h)}(X^{(h)}_{mh}, \ldots, X^{(h)}_{(m-r^{(h)})h})), \quad \sigma^{(h)}(l^{(h)}(X^{(h)}_{mh}, \ldots, X^{(h)}_{(m-r^{(h)})h}))$$

may be replaced by

$$\bar{b}^{(h)}(X^{(h)}_{(m+[\frac{u_1}{h}])h}, \ldots, X^{(h)}_{(m+[\frac{u_n}{h}])h}), \quad \bar{\sigma}^{(h)}(X^{(h)}_{(m+[\frac{u_1}{h}])h}, \ldots, X^{(h)}_{(m+[\frac{u_n}{h}])h}).$$

If we demand that

$$\bar{b}^{(h)}(x) \xrightarrow[h\to 0]{} \bar{b}(x), \quad \bar{\sigma}^{(h)}(x) \xrightarrow[h\to 0]{} \bar{\sigma}(x), \quad x \in (\mathbb{R}^d)^n$$

uniformly on compacts of $(\mathbb{R}^d)^n$ we obtain for $x \in C([-r,0];\mathbb{R}^d)$

$$b^{(h)}(x) := \bar{b}^{(h)}(x([\frac{u_1}{h}]h), \ldots, x([\frac{u_n}{h}]h)) \xrightarrow[h\to 0]{} \bar{b}(x(u_1), \ldots, x(u_n)) = b(x)$$

uniformly on compacts of $C([-r,0];\mathbb{R}^d)$, respectively for σ. Therefore the sequence of processes $\{X^{(h)} : h > 0\}$ determined by $\bar{b}^{(h)}$ and $\bar{\sigma}^{(h)}$ converges weakly to the solution X of the SDDE with coefficients $\bar{b}$ and $\bar{\sigma}$.

Next we shall investigate the special case of linear coefficients. For simplification of notation we shall restrict ourselves to the one-dimensional case. Then there exist signed measures μ and ν such that

$$b(x) = \int_{-r}^{0} x(u)\, d\mu(u), \quad \sigma(x) = \int_{-r}^{0} x(u)\, d\nu(u), \quad x \in C[-r,0].$$

It is clear that the functions b and σ are continuous and locally bounded. The delay equation has then the form

$$\begin{cases} X_0 &= \xi \\ dX(t) &= \int_{-r}^{0} X(t+u)\, d\mu(u)\, dt + \int_{-r}^{0} X(t+u)\, d\nu(u)\, dB(t), \quad t \geq 0. \end{cases} \tag{2.5.5}$$

It is known that strong existence and strong uniqueness hold for the system (2.5.5) and that the integrability condition (2.5.2) holds. We assume that the measures μ and ν are approximated weakly by discrete measures $\mu^{(h)}$ and $\nu^{(h)}$. The approximating measures have mass only at time points $(-jh)$ for $j = 0, \ldots, r^{(h)}$. Define for $x \in C[-r,0]$ the quantities

$$b^{(h)}_j := \mu^{(h)}(\{-jh\}), \qquad b^{(h)}(x) := \int_{-r}^{0} x(u)\, d\mu^{(h)}(u) = \sum_{j=0}^{r^{(h)}} b^{(h)}_j x(-jh)$$

$$\sigma^{(h)}_j := \nu^{(h)}(\{-jh\}), \qquad \sigma^{(h)}(x) := \int_{-r}^{0} x(u)\, d\nu^{(h)}(u) = \sum_{j=0}^{r^{(h)}} \sigma^{(h)}_j x(-jh).$$

If the sequence of measures $\mu^{(h)}$ converges weakly to the measure μ, then we have by definition that

$$\int_{-r}^{0} x(u)\,d\mu^{(h)}(u) \xrightarrow[h\to 0]{} \int_{-r}^{0} x(u)\,d\mu(u)$$

for any fixed element x of $C[-r,0]$. The following lemma shows that this convergence is actually uniform on compact sets of $C[-r,0]$.

2.5.3 Lemma. *Let $\rho^{(h)}$ and ρ be signed measures on $[-r,0]$ such that the sequence $\rho^{(h)}$ converges weakly to ρ. Then it holds that*

$$\sup_{x\in K}\left|\int_{-r}^{0} x(u)\,d\rho^{(h)}(u) - \int_{-r}^{0} x(u)\,d\rho(u)\right| \xrightarrow[h\to 0]{} 0$$

for any compact set K of $C[-r,0]$.

Proof. Assume that the statement is wrong. Then there exists a compact set K of $C[-r,0]$, a sequence of functions $x^{(h)}$ in K and a number $\delta > 0$ such that

$$\left|\int_{-r}^{0} x^{(h)}(u)\,d\rho^{(h)}(u) - \int_{-r}^{0} x^{(h)}(u)\,d\rho(u)\right| > \delta, \quad \forall h > 0. \tag{2.5.6}$$

Since the set K is compact, there exists a subsequence h' of h such that

$$x^{(h')} \xrightarrow[h'\to 0]{} \tilde{x}$$

for an element $\tilde{x}$ of K. Since by assumed weak convergence of $\{\rho^{(h)}\}$

$$\sup_{h>0} \|\rho^{(h)}\|_{\mathrm{TV}} < \infty$$

we have furthermore that

$$\left|\int_{-r}^{0} x^{(h')}(u)\,d\rho^{(h')}(u) - \int_{-r}^{0} \tilde{x}(u)\,d\rho^{(h')}(u)\right| \le \|x^{(h')} - \tilde{x}\|_\infty \|\rho^{(h)}\|_{\mathrm{TV}} \xrightarrow[h'\to 0]{} 0.$$

Therefore we can estimate

$$\begin{aligned}
\left|\int_{-r}^{0} x^{(h')}(u)\,d\rho^{(h')}(u)\right. &- \left.\int_{-r}^{0} x^{(h')}(u)\,d\rho(u)\right| \\
&\le \left|\int_{-r}^{0} x^{(h')}(u)\,d\rho(u) - \int_{-r}^{0} \tilde{x}(u)\,d\rho(u)\right| \\
&\le \left|\int_{-r}^{0} \tilde{x}(u)\,d\rho(u) - \int_{-r}^{0} \tilde{x}(u)\,d\rho^{(h')}(u)\right| \\
&\le \left|\int_{-r}^{0} \tilde{x}(u)\,d\rho^{(h')}(u) - \int_{-r}^{0} x^{(h')}(u)\,d\rho^{(h')}(u)\right| \\
&\le \delta/3 + \delta/3 + \delta/3 \le \delta
\end{aligned}$$

for all sufficiently small h'. But this is a contradiction to the assumption (2.5.6). $\square$
This lemma shows that

$$b^{(h)}(x) \underset{h\to 0}{\longrightarrow} b(x), \quad \sigma^{(h)}(x) \underset{h\to 0}{\longrightarrow} \sigma(x), \quad x \in C[-r,0]$$

uniformly on compacts of $C[-r,0]$, if pointwise convergence is assumed. The next theorem follows now from Theorem 2.4.16 and Lemma 2.5.3.

2.5.4 Theorem. *Assume that we are given the linear SDDE*

$$\begin{cases} X_0 &= \xi \\ dX(t) &= \int_{-r}^{0} X(t+u)\,d\mu(u)\,dt + \int_{-r}^{0} X(t+u)\,d\nu(u)\,dB(t). \end{cases}$$

Let $\epsilon = \{\epsilon_m : m \in \mathbb{N}\}$ be a sequence of i.i.d. variables on some probability space with

$$E(\epsilon_1) = 0, \quad E(\epsilon_1^2) = 1, \quad E(|\epsilon_1|^{2+\delta}) < \infty$$

for some $\delta > 0$. Let for $h > 0$ the time series $(X^{(h)}_{mh})_{m\ge -r^{(h)}}$ be defined by

$$\begin{cases} X^{(h)}_{mh} &= \xi^{(h)}(mh), \quad m = -r^{(h)}, \dots, 0 \\ X^{(h)}_{(m+1)h} &= X^{(h)}_{mh} + \sum_{j=0}^{r^{(h)}} b^{(h)}_j X^{(h)}_{(m-j)h} h + \sum_{j=0}^{r^{(h)}} \sigma^{(h)}_j X^{(h)}_{(m-j)h} \sqrt{h}\epsilon_{m+1}, \quad m \in \mathbb{N}_0. \end{cases}$$

The time series $(X^{(h)}_{mh})_{m\ge -r^{(h)}}$ is extended to a continuous process $X^{(h)}$ by linear interpolation. Define discrete measures on $[-r,0]$ by

$$\mu^{(h)}(\{-jh\}) := b^{(h)}_j, \quad \nu^{(h)}(\{-jh\}) := \sigma^{(h)}_j, \quad j = 0, \dots, r^{(h)}.$$

If

$$\xi^{(h)} \underset{h\to 0}{\longrightarrow} \xi$$

and

$$\mu^{(h)} \Longrightarrow \mu, \quad \nu^{(h)} \Longrightarrow \nu$$

as weak convergence of measures, then the processes $X^{(h)}$ converge weakly to X, where X is the strong solution of the above linear SDDE with initial value ξ.

2.5.5 Remark. *Given an arbitrary measure ρ it is always possible to define a sequence of discrete measures $\rho^{(h)}$ which converges weakly to ρ. It is enough to set*

$$\rho^{(h)}(\{0\}) := 0, \quad \rho^{(h)}(\{-jh\}) := \rho[-jh, (-j+1)h).$$

Hence it is always possible to find an approximating sequence $X^{(h)}$ with the following property: In the evaluation of $X^{(h)}_{(m+1)h}$ enter directly the preceding values of $X^{(h)}$ itself and not the whole function segment

$$l^{(h)}(X^{(h)}_{mh}, \dots, X^{(h)}_{(m-r^{(h)})h})$$

as in Theorem 2.5.1.

Let us return to locally bounded, but not necessarily continuous coefficients. We shall focus on one-dimensional delay equations with discontinuous coefficients in the special case where they only depend on states at earlier times

$$\begin{aligned} b_\triangle(x) &= b_\triangle(\{x(u) : -r \le u \le \triangle\}), \quad \triangle > 0, \quad x \in C[-r,0] \\ \sigma_\triangle(x) &= \sigma_\triangle(\{x(u) : -r \le u \le \triangle\}), \quad \triangle > 0, \quad x \in C[-r,0]. \end{aligned}$$

This is a stringent assumption and does not include the case $r = 0$ of stochastic ordinary differential equations. We shall now formulate and prove an approximation theorem for SDDE's with coefficients $b_\triangle$ and $\sigma_\triangle$.

2.5.6 Theorem. *Assume that weak existence and weak uniqueness hold for system*

$$\left\{ \begin{aligned} X_0 &= \xi \\ dX(t) &= b_\triangle(X_t)\,dt + \sigma_\triangle(X_t)\,dB(t), \quad t \ge 0 \end{aligned} \right. \tag{2.5.7}$$

for every initial condition, where the coefficients $b_\triangle$ and $\sigma_\triangle$ are measurable, locally bounded and fulfil the requirements of Lemma 2.4.8 for the initial condition ξ. Furthermore it is assumed that $\sigma_\triangle$ is bounded away from zero. Let for $h > 0$ the discrete time series $(X^{(h)}_{mh})_{m \ge -r^{(h)}}$ be defined by

$$\left\{ \begin{aligned} X^{(h)}_{mh} &= \xi^{(h)}(mh), \quad m = -r^{(h)}, \dots, 0 \\ X^{(h)}_{(m+1)h} &= X^{(h)}_{mh} + b_\triangle(l^{(h)}_{mh} X^{(h)})h + \sigma_\triangle(l^{(h)}_{mh} X^{(h)})\sqrt{h}\epsilon_{m+1}, \quad m \in \mathbb{N}_0 \end{aligned} \right.$$

for a sequence $\{\epsilon_m : m \in \mathbb{N}\}$ of independent, standard Gaussian random variables on some probability space. The time series $(X^{(h)}_{mh})_{m \ge -r^{(h)}}$ is extended to a continuous process $X^{(h)}$ by linear interpolation. If $\xi^{(h)} \xrightarrow[h\to 0]{} \xi$, then $X^{(h)} \xrightarrow[h\to 0]{d} X$, where X is the unique weak solution of system (2.5.7) with initial value ξ.

Proof. Denote as in the preceding proofs

$$P^{(h)} := \mathrm{Law}(X^{(h)}).$$

For each $k \in \mathbb{N}$ choose a continuous function Ψ_k such that $0 \le \Psi_k \le 1$, $\Psi_k \equiv 1$ on $\{\|x\|_\infty \le k\}$ and $\Psi_k \equiv 0$ on $\{\|x\|_\infty > k+1\}$. At first consider for each $k \in \mathbb{N}$ schemes $X^{k,(h)}$ with the bounded coefficients $\Psi_k b_\triangle$ and $|\Psi_k \sigma_\triangle| \vee \alpha$, where α is a lower bound for the absolute value of $\sigma_\triangle$:

$$X^{k,(h)}_{(m+1)h} = X^{k,(h)}_{mh} + (\Psi_k b_\triangle)(l^{(h)}_{mh} X^{k,(h)})h + (|\Psi_k \sigma_\triangle| \vee \alpha)(l^{(h)}_{mh} X^{k,(h)})\sqrt{h}\epsilon_{m+1}, \quad m \in \mathbb{N}_0.$$

We have that

$$\frac{1}{h} E(X^{k,(h)}_{(m+1)h} - X^{k,(h)}_{mh} | l^{(h)}_{mh} X^{k,(h)} = l^{(h)} x) = (\Psi_k b_\triangle)(l^{(h)} x)$$

and

$$(\Psi_k b_\triangle)(l^{(h)}x^{(h)}) \xrightarrow[h\to 0]{} (\Psi_k b_\triangle)(x)$$

for any sequence $x^{(h)}$ approximating x in the points where $b_\triangle$ is continuous. For the second moments we obtain

$$\frac{1}{h}E(|X^{k,(h)}_{(m+1)h} - X^{k,(h)}_{mh}|^2 | l^{(h)}_{mh}X^{k,(h)} = l^{(h)}x) = h(\Psi_k b_\triangle)^2(l^{(h)}x) + (|\Psi_k\sigma_\triangle| \vee \alpha)^2(l^{(h)}x)$$

and

$$h(\Psi_k b_\triangle)^2(l^{(h)}x^{(h)}) + (|\Psi_k\sigma_\triangle| \vee \alpha)^2(l^{(h)}x^{(h)}) \xrightarrow[h\to 0]{} (\Psi_k^2\sigma_\triangle^2 \vee \alpha^2)(x)$$

for any sequence $x^{(h)}$ approximating x in the points where $\sigma_\triangle$ is continuous. The rescaled absolute $(2+\delta)$-moments for any $\delta > 0$ vanish in the limit:

$$\frac{1}{h}E(|X^{k,(h)}_{(m+1)h} - X^{k,(h)}_{mh}|^{2+\delta} | l^{(h)}_{mh}X^{k,(h)} = l^{(h)}x) \xrightarrow[h\to 0]{} 0$$

uniformly for all $x \in C[-r,0]$. It follows from Theorem 2.4.4 that the sequence

$$\{P^{(h)}_k : h > 0\} := \{\mathrm{Law}(X^{k,(h)}) : h > 0\}$$

is tight, and that every weak limit Q_k solves the martingale problem for $\Psi_k b_\triangle$ and $|\Psi_k\sigma_\triangle| \vee \alpha$ if for each $T > 0$

$$Q_k\left(\int_0^T 1(X^\circ_u \in D_{(\Psi_k b_\triangle)} \cup D_{(|\Psi_k\sigma_\triangle|\vee\alpha)})\, du = 0\right) = 1.$$

To show the last relation we will use the special structure of the coefficients $b_\triangle$ and $\sigma_\triangle$. By Lemma 2.4.8 it suffices to show that for all $x \in \mathbb{R}$ and $u > 0$

$$\varlimsup_{h\to 0} P^{(h)}_k(X^\circ(u) \in B_\delta(x)) \xrightarrow[\delta\to 0]{} 0.$$

At first we assume that $u > \triangle$. We have by construction that

$$\begin{aligned} X^{k,(h)}_{[\frac{u}{h}]h} = X^{k,(h)}_{([\frac{u}{h}]-[\frac{\triangle}{h}])h} &+ \sum_{i=[\frac{u}{h}]-[\frac{\triangle}{h}]}^{[\frac{u}{h}]-1} (\Psi_k b_\triangle)(l^{(h)}_{ih}X^{k,(h)})h \\ &+ \sum_{i=[\frac{u}{h}]-[\frac{\triangle}{h}]}^{[\frac{u}{h}]-1} (|\Psi_k\sigma_\triangle| \vee \alpha)(l^{(h)}_{ih}X^{k,(h)})\sqrt{h}\epsilon_{i+1}. \end{aligned}$$

Therefore conditioning on $\mathcal{F}_{([\frac{u}{h}]-[\frac{\triangle}{h}])h}$ yields

$$P(X^{k,(h)}_{[\frac{u}{h}]h} \in B_\delta(x)) = E(\int_{B_\delta(x)} \frac{1}{\sqrt{2\pi v^{k,(h)}}} \exp\left\{-\frac{(w-\mu^{k,(h)})^2}{2v^{k,(h)}}\right\} dw)$$

with

$$\mu^{k,(h)} = X^{k,(h)}_{([\frac{u}{h}]-[\frac{\triangle}{h}])h} + \sum_{i=[\frac{u}{h}]-[\frac{\triangle}{h}]}^{[\frac{u}{h}]-1} (\Psi_k b_\triangle)(l_{ih}^{(h)} X^{k,(h)})h$$

$$v^{k,(h)} = \sum_{i=[\frac{u}{h}]-[\frac{\triangle}{h}]}^{[\frac{u}{h}]-1} (\Psi_k^2 \sigma_\triangle^2 \vee \alpha^2)(l_{ih}^{(h)} X^{k,(h)})h.$$

Here we used the independence and normal distribution of the sequence ϵ, and that $b_\triangle$ and $\sigma_\triangle$ only depend on values up to time $-\triangle$. Now we can estimate

$$\begin{aligned} \lim_{h\to 0} P(X^{k,(h)}(u) \in B_\delta(x)) &= \lim_{h\to 0} P(X^{k,(h)}_{[\frac{u}{h}]h} \in B_\delta(x)) \\ &\leq \lim_{h\to 0} \int_{B_\delta(x)} \frac{1}{\sqrt{2\pi\alpha^2 h[\frac{\triangle}{h}]}}\, dw \underset{\delta\to 0}{\longrightarrow} 0. \end{aligned}$$

Since Ψ_k also takes the value zero, we had to truncate from above by α to obtain the convergence to zero. In the case $u \leq \triangle$ the random variable $X^{k,(h)}_{[\frac{u}{h}]h}$ is normally distributed with mean and variance depending on the initial condition $\xi^{(h)}$, and it follows in the same manner that

$$\lim_{h\to 0} P(X^{k,(h)}(u) \in B_\delta(x)) \underset{\delta\to 0}{\longrightarrow} 0.$$

Thus in view of Theorem 2.4.4 every limit point Q_k of $\{P_k^{(h)} : h > 0\}$ solves the martingale problem for $\Psi_k b_\triangle$ and $|\Psi_k \sigma_\triangle| \vee \alpha$. Hereby most part of the work is done. Define the stopping time

$$\tau_k(m) := \inf_{t\geq 0}\{|m(t)| \geq k\} \nearrow \infty, \quad m \in \Omega = [-r, \infty).$$

If one denotes by Q the unique law of the solution of

$$\begin{cases} X_0 &= \xi \\ dX(t) &= b_\triangle(X_t)\,dt + \sigma_\triangle(X_t)\,dB(t), \quad t \geq 0, \end{cases}$$

then by Theorem 2.4.13 Q equals Q_k on $\mathcal{M}_{\tau_k}$. This comes from assumed weak uniqueness and the fact that $|\sigma_\triangle| = |\sigma_\triangle| \vee \alpha$. But it also holds that $P_k^{(h)}$ equals $P^{(h)}$ on $\mathcal{M}_{\tau_k}$ using the definition of Ψ_k and using once more that α is a lower bound for $|\sigma_\triangle|$. Now it follows from Lemma 11.1.1 in Stroock and Varadhan [28] that $\{P^{(h)} : h > 0\}$ converges weakly to Q. The theorem has been shown. □

2.5.2 Existence of Weak Solutions

We shall illustrate in the proof of the next theorem how the obtained results may be used to prove weak existence for stochastic delay differential equations. For simplicity of notation we confine to the one-dimensional case.

2.5.7 Theorem. *Consider the autonomous stochastic delay differential equation*

$$\begin{cases} X_0 &= \xi \\ dX(t) &= b(X_t)\,dt + \sigma(X_t)\,dB(t), \quad t \geq 0. \end{cases} \tag{2.5.8}$$

If the coefficients b and σ are bounded and continuous, then there exists a weak solution of system (2.5.8).

Proof. The idea of the proof is to construct an approximating sequence of processes $X^{(h)}$ and to show that the distribution Q of one possible limit point of $\{\mathrm{Law}(X^{(h)}) : h > 0\}$ solves the martingale problem for the coefficients b and σ. Let $\epsilon = \{\epsilon_m : m \in \mathbb{N}\}$ be a sequence of i.i.d. variables on some probability space with

$$E(\epsilon_1) = 0, \quad E(\epsilon_1^2) = 1, \quad E(|\epsilon_1|^{2+\delta}) < \infty$$

for some $\delta > 0$. Let for $h > 0$ the d-dimensional time series $(X_{mh}^{(h)})_{m \geq -r^{(h)}}$ be defined by

$$\begin{cases} X_{mh}^{(h)} &= \xi(mh), \quad m = -r^{(h)}, \ldots, 0 \\ X_{(m+1)h}^{(h)} &= X_{mh}^{(h)} + b(l_{mh}^{(h)} X^{(h)})h + \sigma(l_{mh}^{(h)} X^{(h)})\sqrt{h}\epsilon_{m+1}, \quad m \in \mathbb{N}_0. \end{cases}$$

The time series $(X_{mh}^{(h)})_{m \geq -r^{(h)}}$ is extended to a continuous process $X^{(h)}$ by linear interpolation. Using the former notations we obtain

$$\begin{aligned} b^{(h)*}(x) &= E(b(l^{(h)}x) + \sigma(l^{(h)}x)(\sqrt{h}/h)\epsilon_1) = b(l^{(h)}x) \xrightarrow[h \to 0]{} b(x) \\ a^{(h)*}(x) &= E|b(l^{(h)}x)\sqrt{h} + \sigma(l^{(h)}x)\epsilon_1|^2 = hb^2(l^{(h)}x) + \sigma^2(l^{(h)}x) \xrightarrow[h \to 0]{} \sigma^2(x) \end{aligned}$$

uniformly on compact sets of $C([-r, 0]; \mathbb{R}^d)$ and uniformly locally bounded. Also we see that

$$\triangle_\delta^{(h)*}(x) = \frac{1}{h} E(|b(l^{(h)}x)h + \sigma(l^{(h)}x)\sqrt{h}\epsilon_1|^{2+\delta}) \xrightarrow[h \to 0]{} 0$$

uniformly on bounded sets of $C([-r, 0]; \mathbb{R}^d)$. By uniform boundedness of $b^{(h)*}$ and $a^{(h)*}$ the laws of the sequence $\{X^{(h)} : h > 0\}$ are tight, and every limit point Q solves the martingale problem for b and $a = \sigma^2$ according to Theorem 2.4.16. Then Q is one weak solution of system (2.5.8). □

We shall now treat the case of not necessarily continuous coefficients. In the next lemma we will prove weak existence for SDDE's with vanishing drift coefficients by construction of not discretized Euler schemes.

2.5.8 Lemma. *There exists a weak solution of the stochastic delay differential equation*

$$\begin{cases} X_0 &= \xi \\ dX(t) &= \sigma(X_t)\,dB(t), \quad t \geq 0 \end{cases}$$

if the coefficient σ with domain $C[-r,0]$ has the following structure

$$\sigma(x) = h(f_i(x(u_i))_{i\in I_\sigma}), \quad h : (\mathbb{R})^{I_\sigma} \to \mathbb{R}, \quad f_i : \mathbb{R} \to \mathbb{R}$$

for a continuous function h and Lebesgue measure zero of D_{f_i} for each $i \in I_\sigma \subset [-r,0]$ provided that σ is bounded and bounded away from zero, and the initial condition ξ fulfils the assumptions of Lemma 2.4.8 for the function σ.

Proof. We shall construct not discretized Euler schemes to prove weak existence. The arguments are taken from Yan [31]. Define for each $h > 0$ the in the initial condition ξ starting schemes

$$X_t^{(h)} = X_{mh}^{(h)} + \sigma(l_{mh}^{(h)} X^{(h)})(B(t) - B(mh))$$

for time points $mh < t \leq (m+1)h$, which for all $t \geq 0$ may be written as a stochastic integral:

$$X_t^{(h)} = \xi(0) + \int_0^t \sigma(l^{(h)} X_{[\frac{s}{h}]h}^{(h)})\,dB(s).$$

Since σ is bounded, it follows from earlier used arguments that the laws of the sequence $\{X^{(h)} : h > 0\}$ are tight on the space $C[-r,\infty)$. Then there exists a process X such that a subsequence of $\{X^{(h)}\}$, also denoted by $\{X^{(h)}\}$, converges weakly to X. By the almost sure representation theorem there exists a probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{P})$ and a sequence of processes $Y^{(h)}$ and Y with values in $C[-r,\infty)$ and Brownian motions $B^{(h)}$, all defined on $\bar{\Omega}$, such that

$$\mathrm{Law}(X^{(h)}) = \mathrm{Law}(Y^{(h)}) \quad \forall h > 0, \quad \mathrm{Law}(X) = \mathrm{Law}(Y), \quad Y^{(h)} \xrightarrow[h\to 0]{} Y \quad \text{a.s.},$$

and it holds a.s. on $\bar{\Omega}$ that

$$Y_t^{(h)} = \xi(0) + \int_0^t \sigma(l^{(h)} Y_{[\frac{s}{h}]h}^{(h)})\,dB^{(h)}(s), \quad t \geq 0.$$

Our aim is to show that the law of Y solves the martingale problem for σ. The process Y as almost sure limit of $\{Y^{(h)}\}$ is a continuous martingale with respect to its natural filtration, since each $Y^{(h)}$ is a continuous martingale, and for each $t \geq 0$ the random variables $Y_t^{(h)}$ are uniformly integrable by boundedness of σ. Furthermore it holds by Theorem 2.2 in Kurtz and Protter [16] that

$$\int_0^t Y_u^{(h)}\,dY_u^{(h)} \xrightarrow[h\to 0]{} \int_0^t Y(u)\,dY(u), \quad t \geq 0$$

in probability. Since for every continuous martingale Z

$$[Z](t) = Z^2(t) - 2\int_0^t Z(u)\,dZ(u),$$

it holds that $[Y^{(h)}]_t$ converges to $[Y](t)$ in probability, hence by uniform integrability of $[Y^{(h)}]_t$ also in L^1. Define for each $y \in C[-r,0]$

$$\sigma_1^2(y) := \liminf_{x\to y} \sigma_1^2(x), \quad y \in C[-r,0].$$

It holds for all $0 \le s_1 \le s_2$ that

$$\begin{aligned}
\bar{E}([Y](s_2) - [Y](s_1)) &= \lim_{h\to 0} \bar{E}([Y^{(h)}]_{s_2} - [Y^{(h)}]_{s_1}) \\
&= \lim_{h\to 0} \bar{E}\int_{s_1}^{s_2} \sigma^2(l^{(h)}Y^{(h)}_{[\frac{u}{h}]h})\,du \ge \bar{E}\int_{s_1}^{s_2} \underline{\lim}_{h\to 0}\,\sigma^2(l^{(h)}Y^{(h)}_{[\frac{u}{h}]h})\,du \\
&\ge \bar{E}\int_{s_1}^{s_2} \sigma_1^2(Y_u)\,du.
\end{aligned}$$

Now we use the special structure of σ. One sees for $t \ge r$, where $r \ge 0$ denotes the length of memory, and for α a lower bound of $|\sigma|$ that

$$\begin{aligned}
\bar{E}\int_r^t 1(Y_u \in D_\sigma)\,du &\le \sum_{i\in I_\sigma} \bar{E}\int_r^t 1(Y(u+u_i) \in D_{f_i})\,du \\
&= \sum_{i\in I_\sigma} \bar{E}\int_{r+u_i}^{t+u_i} 1(Y(u) \in D_{f_i})\,du \\
&\le \frac{1}{\alpha^2}\sum_{i\in I_\sigma} \bar{E}\int_{r+u_i}^{t+u_i} 1(Y(u) \in D_{f_i})\sigma_1^2(Y_u)\,du =: K.
\end{aligned}$$

Next we obtain by the occupation time formula

$$\begin{aligned}
K &\le \frac{1}{\alpha^2}\sum_{i\in I_\sigma} \bar{E}\int_{r+u_i}^{t+u_i} 1(Y(u) \in D_{f_i})\,d[Y]u \\
&= \frac{1}{\alpha^2}\sum_{i\in I_\sigma} \bar{E}\int_{D_{f_i}} L^x_{[r+u_i,t+u_i]}(Y)\,dx = 0,
\end{aligned}$$

since we assumed that f_i has Lebesgue measure zero, and the local time $L^x(Y)$ is finite for almost all x. Since Y starts in ξ, we have by assumption on the initial condition ξ that for all $t \le r$

$$\bar{E}\int_0^t 1(Y_u \in D_\sigma)\,du = 0.$$

We have achieved so far that $\int_0^t 1(Y_u \in D_a)\, du = 0$ a.s. for all $t > 0$. Then it holds by dominated convergence that

$$\begin{aligned} [Y^{(h)}]_t &= \int_0^t \sigma^2(l^{(h)} Y^{(h)}_{[\frac{u}{h}]h}) 1(Y_u \notin D_\sigma)\, du \\ &\xrightarrow[h\to 0]{} \int_0^t \sigma^2(Y_u) 1(Y_u \notin D_\sigma)\, du = \int_0^t \sigma^2(Y_u)\, du \end{aligned}$$

almost surely and in L^1. This implies that a.s. $[Y](t) = \int_0^t \sigma^2(Y_u)\, du$ first for fix $t \geq 0$ and then by monotonicity a.s. uniformly for all $t \geq 0$. Now for $f \in C_0^\infty(\mathbb{R})$ we can apply Itô's formula on the space $\bar{\Omega}$

$$f(Y(t)) = f(\xi(0)) + \int_0^t f'(Y(u))\, dY(u) + \frac{1}{2}\int_0^t f''(Y(u))\, d[Y](u), \quad t \geq 0.$$

Since Y is a martingale, the stochastic integral with respect to Y has expectation zero. Therefore Y solves the martingale problem for σ. The lemma has been shown. □

The next theorem relaxes the requirement for σ to be bounded away from zero and admits drift coefficients.

2.5.9 Theorem. *There exists a weak solution of the stochastic delay differential equation*

$$\begin{cases} X_0 &= \xi \\ dX(t) &= b(X_t)\, dt + \sigma(X_t)\, dB(t), \quad t \geq 0 \end{cases}$$

if b is bounded and measurable, and if σ with domain $C[-r, 0]$ has the following structure

$$\sigma(x) = h(f_i(x(u_i))_{i\in I_\sigma}), \quad h : (\mathbb{R})^{I_\sigma} \to \mathbb{R}, \quad f_i : \mathbb{R} \to \mathbb{R}$$

for a continuous function h and Lebesgue measure zero of D_{f_i} for each $i \in I_\sigma \subset [-r, 0]$ provided that σ is bounded and the initial condition ξ fulfils the assumptions of lemma 2.4.8 for the function σ.

Proof. By Lemma 2.5.8 for each $\alpha > 0$ there exists a weak solution of the system

$$\begin{cases} X_0^\alpha &= \xi \\ dX^\alpha(t) &= \sigma^\alpha(X_t^\alpha)\, dB(t), \quad t \geq 0. \end{cases}$$

for $\sigma^\alpha := |\sigma| \vee \alpha$. Then the function $\gamma := b/\sigma^\alpha$ is measurable and bounded and

$$M(t) := \exp\{\int_0^t \gamma(X_s^\alpha)\, dB(s) - \frac{1}{2}\int_0^t \gamma^2(X_s^\alpha)\, ds\}, \quad t \geq 0$$

is a martingale. Then by Girsanov's theorem there exists a weak solution of

$$\begin{cases} X_0^\alpha &= \xi \\ dX^\alpha(t) &= b(X_t^\alpha)\, dt + \sigma^\alpha(X_t^\alpha)\, dB(t), \quad t \geq 0. \end{cases}$$

The laws of the sequence $\{X^\alpha : \alpha > 0\}$ are tight. This follows from the tightness criterion of Kolmogorov, see page 474 in Revuz and Yor [24]. We have for each $\alpha > 0$ that

$$|X^\alpha(t) - X^\alpha(s)|^4 \leq 8 \left| \int_s^t b(X_u^\alpha)\, du \right|^4 + 8 \left| \int_s^t \sigma^\alpha(X_u^\alpha)\, dB(u) \right|^4 .$$

By the Burkholder-Davis-Gundy inequality the expectation of the second summand may be estimated from above

$$E \left| \int_s^t \sigma^\alpha(X_u^\alpha)\, dB(u) \right|^4 \leq CE \left| \int_s^t (\sigma^\alpha)^2(X_u^\alpha)\, du \right|^2 .$$

Using boundedness of b and σ we have for each $0 \leq s \leq t \leq T$ and for a constant β that

$$E[|X^\alpha(t) - X^\alpha(s)|^4] \leq \beta |t-s|^2,$$

from which tightness follows. Denote the laws of X^α by Q^α and the law of one limit point by Q. For each $\alpha > 0$ we have by the martingale property of Q^α that for $f \in C_0^\infty(\mathbb{R})$ and for every continuous, bounded, $\mathcal{M}_{t_1}$-measurable function Φ

$$E^{Q^\alpha}(Z^\alpha \Phi) = 0$$

with

$$Z^\alpha(m) := f(m(t_2)) - f(m(t_1)) + \int_{t_1}^{t_2} (L_{b,\sigma^\alpha} f)(m_u)\, du), \quad m \in C[-r,\infty).$$

Now note that

$$\sup_{x \in C[-r,0]} |(\sigma^\alpha)^2(x) - \sigma^2(x)| \leq \sup_{x \in C[-r,0]} |\alpha^2 - \sigma^2(x)| 1_{\{|\sigma(x)| \leq \alpha\}} \leq 2\alpha^2.$$

Therefore we conclude that

$$Z^\alpha(m) \xrightarrow[\alpha \to 0]{} f(m(t_2)) - f(m(t_1)) + \int_{t_1}^{t_2} (L_{b,\sigma} f)(m_u)\, du =: Z(m), \quad m \in C[-r,\infty)$$

uniformly on $C[-r,\infty)$, where Z^α and Z are uniformly bounded. Using this uniform convergence, the tightness of the laws of $\{X^\alpha : \alpha > 0\}$ and the uniform boundedness of Z^α and Z we conclude that $E^Q(Z\Phi) = 0$. Thus Q solves the martingale problem for b and σ, which is our desired result. □

2.5.3 A Continuous GARCH(p,1)-Model

We come now to another type of applications of Theorem 2.4.16. In Theorem 2.5.1 we were given the stochastic delay differential equation and constructed approximating discrete time series. One can also go the other way round. Now let us be given a sequence of discrete time series $Z^{(h)}$. We are interested in a weak limit if it should exist. We shall illustrate how to establish a limit process. In this subsection we deal with time series from financial mathematics, called GARCH(p,1)-models. They are defined as follows. Let $(\epsilon_i)_{i\in\mathbb{N}}$ be a sequence of real-valued i.i.d. variables on some probability space with

$$E(\epsilon_1)=0,\quad E(\epsilon_1^2)=1,\quad E(\epsilon_1^3)=0\quad,\quad E(\epsilon_1^4)=c^2+1,\quad E(|\epsilon_1|^{4+2\delta})<\infty$$

for some $\delta>0$ and for some constant $c\geq 0$. Define for each $h>0$ the following two-dimensional scheme

$$\begin{cases} X^{(h)}_{(m+1)h} &= X^{(h)}_{mh}+\rho^{(h)}_{mh}\sqrt{h}\epsilon_{m+1}\\ \rho^{(h)2}_{(m+1)h} &= v_0^{(h)}+\beta^{(h)}\rho^{(h)2}_{mh}+\sum_{j=1}^{p^{(h)}}\beta_j^{(h)}\rho^{(h)2}_{(m-j)h}+\alpha^{(h)}\rho^{(h)2}_{mh}h\epsilon^2_{m+1},\quad m\in\mathbb{N}_0. \end{cases} \tag{2.5.9}$$

In GARCH(p,q) the parameters p and q are defined by

$$\begin{aligned} p &:= 1+\max\{j:\beta_{j'}=0\quad\forall j'>j\}=p^{(h)}+1\\ q &:= 1+\max\{i:\alpha_{i'}=0\quad\forall i'>i\}=1. \end{aligned}$$

Then scheme (2.5.9) is a GARCH($p^{(h)}+1,1$)-process. The parameters

$$v_0^{(h)}>0,\quad\beta_j^{(h)}\geq 0,\quad\beta^{(h)}\geq 0,\quad\alpha^{(h)}\geq 0$$

ensure that $\rho^{(h)2}_{mh}$ is strictly positive for $m\in\mathbb{N}_0$. We assume that $p^{(h)}=p/h\in\mathbb{N}_0$. The first parameter $(p^{(h)}+1)$ of those GARCH-systems will tend to infinity as h tends to zero in the case $p>0$. We did not specify any initial condition yet. Set

$$X_0^{(h)}=x,\quad\rho^{(h)2}_{-mh}=\xi(-mh)^2,\quad m=0,\ldots,p^{(h)},$$

where ξ^2 is a strictly positive continuous function on $[-p,0]$ and $x\in\mathbb{R}$. As before we interpolate linearly between discrete time points to obtain continuous stochastic processes $Z^{(h)}=(X^{(h)},\rho^{(h)2})$. How do we have to choose sequences of parameters that the two-dimensional processes $Z^{(h)}$ converge weakly? This question is handled in the following way. First calculate the quantity

$$\frac{1}{h}E((Z^{(h)}_{(m+1)h}-Z^{(h)}_{mh})|\mathfrak{F}_{mh}),\quad m\in\mathbb{N}_0.$$

If $Z^{(h)}$ satisfies a.s

$$P(Z^{(h)}_{(m+1)h}\in\Gamma|\mathfrak{F}_{mh})=P(Z^{(h)}_{(m+1)h}\in\Gamma|\sigma(Z^{(h)}_{mh},\ldots,Z^{(h)}_{(m-p^{(h)})h})),\quad m\in\mathbb{N}_0,$$

then there exists a two-dimensional measurable function $b^{(h)}$ such that a.s.

$$\frac{1}{h}E((Z^{(h)}_{(m+1)h}-Z^{(h)}_{mh})|\mathfrak{F}_{mh})=b^{(h)}(Z^{(h)}_{mh},\ldots,Z^{(h)}_{(m-p^{(h)})h}),\quad m\in\mathbb{N}_0.$$

Next identify $b^{(h)}$ with a function of the space $C([-p,0];\mathbb{R}^2)$ such that

$$b^{(h)}(x) = b^{(h)}(Z^{(h)}_{mh},\dots,Z^{(h)}_{(m-p^{(h)})h})\Big|_{\{Z^{(h)}_{(m-i)h}=x(-ih):0\le i\le p^{(h)}\}}, \quad x\in C([-p,0];\mathbb{R}^2).$$

Finally, examine if there exists a limit of $b^{(h)}$ on $C([-p,0];\mathbb{R}^2)$ and check if the convergence is uniform on compact sets of $C([-p,0];\mathbb{R}^2)$. We shall now see how this procedure works for

$$Z^{(h)} = (X^{(h)}, \rho^{(h)2}).$$

Firstly, we obtain

$$\frac{1}{h}E(X^{(h)}_{(m+1)h} - X^{(h)}_{mh}|\mathfrak{F}_{mh}) = 0.$$

Now consider

$$\frac{1}{h}E(\rho^{(h)2}_{(m+1)h} - \rho^{(h)2}_{mh}|\mathfrak{F}_{mh}) = \frac{v_0^{(h)}}{h} + \sum_{j=1}^{p^{(h)}} \frac{\beta_j^{(h)}}{h}\rho^{(h)2}_{(m-j)h} + \left(\frac{\beta^{(h)}-1}{h} + \alpha^{(h)}\right)\rho^{(h)2}_{mh}$$

and identify with

$$b_2^{(h)}(x) = \frac{v_0^{(h)}}{h} + \sum_{j=1}^{p^{(h)}} \frac{\beta_j^{(h)}}{h}x(-jh) + \left(\frac{\beta^{(h)}-1}{h} + \alpha^{(h)}\right)x(0), \quad x\in C([-p,0];\mathbb{R}^2).$$

If we define a discrete nonnegative measure $\gamma^{(h)}$ on $[-p,0]$ by

$$\gamma^{(h)}(\{0\}) := 0, \quad \gamma^{(h)}(\{-jh\}) := \frac{\beta_j^{(h)}}{h}, \quad 1\le j\le p^{(h)},$$

then we obtain the following representation in terms of the measure $\gamma^{(h)}$

$$b_2^{(h)}(x) = \frac{v_0^{(h)}}{h} + \int_{-p}^{0} x(u)\,d\gamma^{(h)}(u) + \left(\frac{\beta^{(h)}-1}{h} + \alpha^{(h)}\right)x(0), \quad x\in C([-p,0];\mathbb{R}^2).$$

Now we see how the coefficients have to behave to obtain a limit. If

$$\frac{v_0^{(h)}}{h} \xrightarrow[h\to 0]{} v_0, \quad -\left(\frac{\beta^{(h)}-1}{h} + \alpha^{(h)}\right) \xrightarrow[h\to 0]{} \lambda, \quad \gamma^{(h)} \Longrightarrow \gamma,$$

then it holds for each $x\in C([-p,0];\mathbb{R}^2)$ that

$$b_2^{(h)}(x) \xrightarrow[h\to 0]{} b_2(x) = v_0 + \int_{-p}^{0} x(u)\,d\gamma(u) - \lambda x(0).$$

Recall that the notation $\gamma^{(h)} \Longrightarrow \gamma$ means that the sequence of measures $\gamma^{(h)}$ converges weakly to the (nonnegative) measure γ. We assume that $\gamma(\{0\}) = 0$. This convergence is even uniformly on compacts of $C([-p,0];\mathbb{R}^2)$ and locally uniformly bounded as we have shown in Lemma 2.5.3. We have obtained so far that the drift of the distribution limit has the form $(0,b_2)$, where b_2 is given above. The diffusion coefficients are treated in the same way. We investigate the behavior of

$$\frac{1}{h}E((Z^{(h)}_{(m+1)h}-Z^{(h)}_{mh})(Z^{(h)}_{(m+1)h}-Z^{(h)}_{mh})^T|\mathfrak{F}_{mh}).$$

Using $E(\epsilon_1^3)=0$ one sees that

$$\frac{1}{h}E((X^{(h)}_{(m+1)h}-X^{(h)}_{mh})(\rho^{(h)2}_{(m+1)h}-\rho^{(h)2}_{mh})|\mathfrak{F}_{mh})$$

vanishes, hence the mixed term a_{12} is zero. Furthermore we have that

$$\frac{1}{h}E((X^{(h)}_{(m+1)h}-X^{(h)}_{mh})^2|\mathfrak{F}_{mh})=\rho^{(h)2}_{mh}$$

and after some calculations

$$\frac{1}{h}E((\rho^{(h)2}_{(m+1)h}-\rho^{(h)2}_{mh})^2|\mathfrak{F}_{mh})=h\left(b_2^{(h)}(l^{(h)}_{mh}\rho^{(h)2})\right)^2+\alpha^{(h)2}h\rho^{(h)4}_{mh}E(\epsilon^4_{m+1}-1).$$

If one demands that

$$\alpha^{(h)}\sqrt{h}\xrightarrow[h\to 0]{}\alpha,$$

then

$$a^{(h)}_{22}(x):=h(b_2^{(h)}(l^{(h)}x))^2+c^2\alpha^{(h)2}hx(0)\xrightarrow[h\to 0]{}c^2\alpha^2x^2(0)=:a_{22}(x),\quad x\in C([-p,0];\mathbb{R}^2)$$

uniformly on compacts of $C([-p,0];\mathbb{R}^2)$ and uniformly locally bounded. At last one checks that

$$\begin{aligned}\frac{1}{h}E(|X^{(h)}_{(m+1)h}-X^{(h)}_{mh}|^{2+\delta}|\mathfrak{F}_{mh}) &\xrightarrow[h\to 0]{} 0\\ \frac{1}{h}E(|\rho^{(h)2}_{(m+1)h}-\rho^{(h)2}_{mh})|^{2+\delta}|\mathfrak{F}_{mh}) &\xrightarrow[h\to 0]{} 0\end{aligned}$$

uniformly on bounded sets of $C([-p,0];\mathbb{R}^2)$. Now by application of Theorem 2.4.16 we obtain the following result.

2.5.10 Theorem. *Under above assumptions the stochastic processes $(X^{(h)},\rho^{(h)2})$ in (2.5.9) converge weakly to (X,ρ^2), where (X,ρ^2) is the unique weak solution of*

$$\begin{cases} dX(t) &= \rho(t)\,dW(t)\\ d\rho^2(t) &= \left[v_0-\lambda\rho^2(t)+\int_{-p}^0\rho^2(t+u)\,d\gamma(u)\right]dt+c\alpha\rho^2(t)\,dB(t),\quad t\geq 0\end{cases}\tag{2.5.10}$$

with parameters

$v_0>0,\quad \lambda\in\mathbb{R},\quad \gamma$ *a nonnegative measure with* $\gamma(\{0\})=0,\quad \alpha\geq 0,\quad c\geq 0,$

driven by a two-dimensional Brownian motion (W,B).

Proof. We still have to check that weak existence and weak uniqueness hold for system (2.5.10). This is clear in the case $c\alpha = 0$, since we have a deterministic equation for ρ^2 in this case. Therefore we assume from now on that $c\alpha > 0$. First we shall show that $\rho^2(t)$ remains strictly positive a.s. if $\rho^2(0) > 0$ and $\rho^2(u) \geq 0$ for all $-p \leq u < 0$ a.s. Define the process

$$\begin{cases} \phi(0) &= c\alpha \\ d\phi(t) &= c\alpha\phi(t)\,dB(t), \quad t \geq 0. \end{cases}$$

Clearly, ϕ remains strictly positive a.s. and the process

$$Z(t) := \frac{\rho^2(t)}{\phi(t)}, \quad t \geq 0$$

is a.s. well-defined, and it lies in C^1 which may be checked by Itô's formula. Thus we have by partial integration and by definition that

$$\begin{aligned} \phi(t)\,dZ(t) + Z(t)\,d\phi(t) &= d\rho^2(t) \\ \phi(t)Z'(t)\,dt + c\alpha Z(t)\phi(t)\,dB(t) &= (v_0 + \lambda\rho^2(t) + \int_{-p}^{0} \rho^2(t+u)\,d\gamma(u))\,dt \\ &+ c\alpha\rho^2(t)\,dB(t), \end{aligned}$$

from which follows that

$$Z'(t) = \frac{1}{\phi(t)}\left(v_0 + \lambda\phi(t)Z(t) + \int_{-p}^{0} \phi(t+u)Z(t+u)\,d\gamma(u)\right). \tag{2.5.11}$$

Set

$$t_0 := \inf\{t > 0 : Z(t) = 0\}, \quad t_0 > 0 \quad (Z(0) > 0).$$

Assume that $t_0 < \infty$. By definition of t_0 we have that $Z'(t_0) \leq 0$. But on the other hand, it follows from (2.5.11) that $Z'(t_0) > 0$ a.s. since $Z(t_0) = 0$. The contradiction shows that Z remains positive a.s. Hence does ρ^2, and system (2.5.10) is well-defined. Note that strong existence and strong uniqueness holds for ρ^2. Since we have assumed deterministic initial conditions, ρ^2 lies in L^p for every p. Hence it is enough to choose a second Brownian motion W, which is independent of B, and set

$$X(t) = X(0) + \int_0^t \rho(u)\,dW(u),$$

which shows weak existence. Assume that there are two weak solutions (X, ρ^2) and $(\bar{X}, \bar{\rho}^2)$, driven by the two-dimensional Brownian motions (W, B) and $(\bar{W}, \bar{B})$ respectively. Firstly, we have $\rho^2 \overset{\mathrm{d}}{=} \bar{\rho}^2$ by strong uniqueness. Since X is a stochastic integral of ρ with respect to the Brownian motion W, and the same is true for $\bar{X}$ respectively, it follows that $(X, \rho^2) \overset{\mathrm{d}}{=} (\bar{X}, \bar{\rho}^2)$. □

If the measure γ is zero, then system (2.5.10) is the stochastic ordinary differential equation in Jeantheau [11]. If $\gamma \equiv 0$ and $\alpha = 0$, then system (2.5.10) is a stochastic ordinary differential equation which can be embedded in Pedersen [22].

In statistics it is interesting to know if there exists a stationary solution of stochastic differential equations. Therefore we shall investigate the question of existence and uniqueness of a stationary solution for the second component ρ^2 in system (2.5.10). Writing G for ρ^2 we have to study the SDDE

$$dG(t) = \left(v_0 - \lambda G(t) + \int_{-p}^{0} G(t+u)\, d\gamma(u)\right) dt + c\alpha G(t)\, dB(t), \quad t \geq 0. \tag{2.5.12}$$

We assume that $\alpha > 0$. In view of Theorem 15 in Itô and Nisio [9] there exists a stationary solution of system (2.5.12) if

$$\lambda > \|\gamma\| + \frac{1}{2}c^2\alpha^2, \qquad \|\gamma\| = \int_{-p}^{0} 1\, d\gamma(u).$$

Now assume that we have two stationary solutions G_1 and G_2. Then by linearity the difference $Z := G_1 - G_2$ solves

$$dZ(t) = \left(-\lambda Z(t) + \int_{-p}^{0} Z(t+u)\, d\gamma(u)\right) dt + c\alpha Z(t)\, dB(t).$$

Our aim is to establish under which conditions $E(\|Z_t\|_\infty^2) \xrightarrow[t\to\infty]{} 0$, where Z_t denotes the function segment. The following argumentation is taken from Riedle and Mao [26] and is therefore only presented roughly. Applying Itô's formula and using the estimation

$$|a||b| \leq \frac{1}{2}(a^2 + b^2)$$

one obtains for $\phi(t) := E(Z^2(t))$ and $s < t$

$$\begin{aligned} \frac{\phi(t) - \phi(s)}{t-s} &\leq (c^2\alpha^2 - 2\lambda + \|\gamma\|)\frac{1}{t-s}\int_s^t \phi(u)\, du \\ &+ \frac{1}{t-s}\int_s^t \int_{-p}^{0} \phi(v+u)\, d\gamma(v)\, du. \end{aligned}$$

Since ϕ is continuous, one obtains for the upper Dini-derivative D_+ by letting s tend to t

$$D_+\phi(t) \leq (c^2\alpha^2 - 2\lambda + \|\gamma\|)\phi(t) + \int_{-p}^{0} \phi(t+u)\, d\gamma(v).$$

Now consider the deterministic system

$$\begin{cases} x(u) = \phi(u), & -p \leq u \leq 0 \\ \dot{x}(t) = (c^2\alpha^2 - 2\lambda + \|\gamma\|)\phi(t) + \int_{-p}^{0} \phi(t+u)\, d\gamma(v). \end{cases}$$

The zeroes of the characteristic function

$$\chi(\mu) := \mu - \left((c^2\alpha^2 - 2\lambda + \|\gamma\|) + \int_{-p}^{0} e^{\mu v}\, d\gamma(v) \right)$$

all have negative real part if $\lambda > \|\gamma\| + (1/2)c^2\alpha^2$. In this case the deterministic theory tells us that $|x(t)| \xrightarrow[t\to\infty]{} 0$. Furthermore it holds that $\phi(t) \leq x(t)$. Therefore we deduce that $E(Z^2(t)) \xrightarrow[t\to\infty]{} 0$ and, by using standard estimates, it holds moreover that $E(\|Z_t\|_\infty^2) \xrightarrow[t\to\infty]{} 0$. Thus we have shown uniqueness of the stationary solution of system (2.5.12) if the parameters satisfy

$$\lambda > \|\gamma\| + \frac{1}{2}c^2\alpha^2.$$

This is exactly the condition under which Itô and Nisio [9] showed existence of a stationary solution of system (2.5.12). However, it is an unsolved problem to give an explicit description of the stationary distribution, even if γ is concentrated on the time point $\{-p\}$. This comes from the fact that in an Itô formula for $f(X(t), X(t-p))$ the Malliavin derivative $D(X(s))$ occurs, see page 10 in Buckwar and Winkler [4].

2.5.4 A Continuous GARCH(p,q)-Model

In generalisation to scheme (2.5.9) consider for $h > 0$ the series

$$\begin{cases} X^{(h)}_{(m+1)h} = X^{(h)}_{mh} + \rho^{(h)}_{mh}\sqrt{h}\epsilon_{m+1} \\ \rho^{(h)2}_{(m+1)h} = v^{(h)}_0 + \beta^{(h)}\rho^{(h)2}_{mh} + \sum_{j=1}^{p^{(h)}} \beta^{(h)}_j \rho^{(h)2}_{(m-j)h} \\ \qquad + \alpha^{(h)}\rho^{(h)2}_{mh} h\epsilon^2_{m+1} + \sum_{i=1}^{q^{(h)}} \alpha^{(h)}_i \rho^{(h)2}_{(m-i)h} h\epsilon^2_{m+1-i}, \quad m \in \mathbb{N}_0. \end{cases} \tag{2.5.13}$$

This is the general GARCH-model of order $(p^{(h)}+1, q^{(h)}+1)$. The parameters

$$v^{(h)}_0 > 0, \quad \beta^{(h)}_j \geq 0, \quad \beta^{(h)} \geq 0, \quad \alpha^{(h)}_i \geq 0, \quad \alpha^{(h)} \geq 0$$

ensure that $\rho^{(h)2}_{mh}$ is strictly positive for $m \in \mathbb{N}_0$. We assume that $p^{(h)} = p/h \in \mathbb{N}_0$ and $q^{(h)} = q/h \in \mathbb{N}_0$. The above system is equivalent to

$$\begin{cases} X^{(h)}_{(m+1)h} = X^{(h)}_{mh} + \rho^{(h)}_{mh}\sqrt{h}\epsilon_{m+1} \\ \rho^{(h)2}_{(m+1)h} = v^{(h)}_0 + \beta^{(h)}\rho^{(h)2}_{mh} + \sum_{j=1}^{p^{(h)}} \beta^{(h)}_j \rho^{(h)2}_{(m-j)h} \\ \qquad + \alpha^{(h)}\rho^{(h)2}_{mh} h\epsilon^2_{m+1} + \sum_{i=1}^{q^{(h)}} \alpha^{(h)}_i (X^{(h)}_{(m+1-i)h} - X^{(h)}_{(m-i)h})^2, \quad m \in \mathbb{N}_0 \end{cases}$$

for deterministic initial conditions $X_u^{(h)} = x(u)$ on $[-q,0]$ and $\rho_u^{(h)2} = \xi^2(u)$ on $[-p,0]$. Assume that $\alpha_i^{(h)} = 0$ with exception of finitely many i uniformly in h. Assume furthermore that

$$\frac{v_0^{(h)}}{h} \xrightarrow[h\to 0]{} v_0, \quad -\left(\frac{\beta^{(h)}-1}{h} + \alpha^{(h)}\right) \xrightarrow[h\to 0]{} \lambda, \quad \gamma^{(h)} \Longrightarrow \gamma, \quad \alpha^{(h)}\sqrt{h} \xrightarrow[h\to 0]{} \alpha,$$
$$\sup_{h>0} \frac{|\alpha_i^{(h)}|}{h} < \infty.$$

After some calculations one sees that in this case the stochastic processes $(X^{(h)}, \rho^{(h)2})$ in (2.5.13) converge weakly to (X,ρ^2), where (X,ρ^2) is the unique weak solution of

$$\begin{cases} dX(t) &= \rho(t)\,dW(t), \quad X_0 = x \\ d\rho^2(t) &= \left[v_0 + \lambda\rho^2(t) + \int_{-p}^0 \rho^2(t+u)\,d\gamma(u)\right] dt + c\alpha\rho^2(t)\,dB(t), \quad t \geq 0, \end{cases}$$

where $c = \sqrt{E|\epsilon_1|^4 - 1}$. This is the same system as in Theorem 2.5.10. Only finitely many $\alpha_i^{(h)}$ uniformly in h do not change the weak limit in comparison to GARCH$(p,1)$-models. This comes from the fact that the term

$$\sum_i \frac{\alpha_i^{(h)}}{h}(x(u_i^{(h)}+h) - x(u_i^{(h)}))^2, \quad x \in C[-r,0]$$

tends to zero as h to zero uniformly on compacts of $C[-r,0]$ if i remains in a finite index set. Now choose coefficients of the kind

$$\alpha_i^{(h)} = \alpha(-ih)h, \quad 0 \leq i \leq q^{(h)}$$

for a continuous function α on $[-q,0]$. Then in the drift term for σ^2 the following additional term occurs

$$\sum_{i=1}^{q^{(h)}} \alpha(-ih)(x((-i+1)h) - x(-ih))^2 \xrightarrow[h\to 0]{} \int_{-q}^0 \alpha(u)d[x](u) =: g(x), \quad x \in C[-r,0],$$

where $[x]$ denotes the quadratic variation of x. Here one has to be careful. Firstly, the limit $g(x)$ is infinite if x is a function of infinite quadratic variation. Furthermore, it is not locally bounded and not continuous in any point x. Finally, the convergence is not uniform on compact sets. So many conditions of Theorem 2.4.16 are not fulfilled. But if we formally write down the limit process, then we obtain

$$\begin{cases} dX(t) &= \rho(t)\,dW(t), \quad X_0 = x \quad \text{on} \quad [-q,0] \\ d\rho^2(t) &= \left[v_0 + \lambda\rho^2(t) + \int_{-p}^0 \rho^2(t+u)\,d\gamma(u) + \int_{-q}^0 \alpha(u)\,d[X_t](u)\right] dt \\ &+ \ c\alpha\rho^2(t)\,dB(t), \quad \rho_0^2 = \xi^2 \quad \text{on} \quad [-p,0]. \end{cases}$$

If $p = q$ and $d[x](v) = \xi^2(v)\,dv$ on $[-q,0]$, then this system is equivalent to

$$\left\{\begin{array}{rcl} dX(t) &=& \rho(t)\,dW(t), \quad X_0 = x \quad \text{on} \quad [-q,0] \\ d\rho^2(t) &=& \left[v_0 + \lambda\rho^2(t) + \int_{-p}^0 \rho^2(t+u)\,d\gamma(u) + \int_{-q}^0 \rho^2(t+u)\alpha(u)\,du\right] dt \\ &+& c\alpha\rho^2(t)\,dB(t), \quad \rho_0^2 = \xi^2 \quad \text{on} \quad [-q,0]. \end{array}\right. \tag{2.5.14}$$

We emphasize once more that the limit (X,ρ^2) in (2.5.14) is purely formal, and that by no means we are able to prove with our tools that the sequence $\{(X^{(h)},\rho^{(h)2}) : h > 0\}$ converges weakly to this limit. For a discussion for coefficients which have an even more general form than $\alpha_i^{(h)} = \alpha(-ih)h$ we refer to the next chapter.

2.5.5 Time Series with Fading Memory

In this subsection we will deal with a new appearance. We investigate what happens if the order of the autoregressive schemes $X^{(h)}$ remains constant for each $h > 0$. Then the length of memory shrinks to zero. Fix a number $R \in \mathbb{N}_0$. For $h > 0$ a one-dimensional time series $X^{(h)}$ is given by

$$\left\{\begin{array}{rcl} X_{mh}^{(h)} &=& \xi_i^{(h)}, \quad 0 \le i \le R, \quad \xi_i^{(h)} \in \mathbb{R} \\ X_{(m+1)h}^{(h)} &=& X_{mh}^{(h)} + b^{(h)}(X_{mh}^{(h)}, X_{(m-1)h}^{(h)}, \ldots, X_{(m-R)h}^{(h)})h \\ &+& \sigma^{(h)}(X_{mh}^{(h)}, X_{(m-1)h}^{(h)}, \ldots, X_{(m-R)h}^{(h)})\sqrt{h}\epsilon_{m+1}, \quad m \in \mathbb{N}_0. \end{array}\right.$$

Here $b^{(h)}$ and $\sigma^{(h)}$ are measurable functions with fixed domain $\mathbb{R}^{(R+1)}$ for all $h > 0$. By setting $Y_{mh}^{(h)} := (X_{mh}^{(h)}, X_{(m-1)h}^{(h)}, \ldots, X_{(m-R)h}^{(h)})$ we achieve that $\{Y_{mh}^{(h)} : m \in \mathbb{N}_0\}$ is an $\mathbb{R}^{(R+1)}$-valued Markov chain in discrete time with start in $\xi^{(h)}$. Its increments may be written in the form

$$Y_{(m+1)h}^{(h)} - Y_{mh}^{(h)} = \check{b}^{(h)}(Y_{mh}^{(h)})h + \check{\sigma}^{(h)}(Y_{mh}^{(h)})\sqrt{h}\epsilon_{m+1}, \quad m \in \mathbb{N}_0,$$

where the vector function $\check{b}^{(h)}$ and the matrix function $\check{\sigma}^{(h)}$ for $y \in \mathbb{R}^{(R+1)}$ are defined as follows

$$\begin{array}{rcl} \check{b}_0^{(h)}(y) &:=& b^{(h)}(y_0, \ldots, y_R), \quad \check{b}_i^{(h)}(y) := 0, \quad 1 \le i \le R \\ \check{\sigma}_{00}^{(h)}(y) &:=& \sigma^{(h)}(y_0, \ldots, y_R), \quad \check{\sigma}_{ij}^{(h)}(y) := 0, \quad i^2 + j^2 > 0, \quad 0 \le i,j \le R. \end{array}$$

Note that $Y_{mh}^{i,(h)} = X_{(m-i)h}^{(h)}$ for $0 \le i \le R$. We obtain for the conditional expectation of the second component $Y^{1,(h)}$ of $Y^{(h)}$

$$\frac{1}{h}E(Y_{(m+1)h}^{1,(h)} - Y_{mh}^{1,(h)} | \mathcal{F}_{mh}^{Y^{(h)}}) = \frac{1}{h}(Y_{(m+1)h}^{1,(h)} - Y_{mh}^{1,(h)}), \quad m \in \mathbb{N}_0,$$

since $Y_{(m+1)h}^{1,(h)} = Y_{mh}^{0,(h)}$ is $\mathcal{F}_{mh}^{Y^{(h)}}$-measurable. The second component of the drift takes the form

$$b^{1,(h)*}(y) = \frac{1}{h}(y_0 - y_1), \quad y \in \mathbb{R}^{(R+1)}.$$

Here we see that the drifts do not behave properly as h tends to zero. Our preceding theorems are for the sequence of Markov chains $Y^{(h)}$ not applicable. Therefore we shall drop again the lifting. Instead choose a real number $r \geq 0$ such that $R/h \leq r$ for all $h > 0$ which are lower than a given $h_0 > 0$. Then we can write the increments of $X^{(h)}$ also in the following form

$$X^{(h)}_{(m+1)h} - X^{(h)}_{mh} = \hat{b}^{(h)}(l_{mh}X^{(h)})h + \hat{\sigma}^{(h)}(l_{mh}X^{(h)})\sqrt{h}\epsilon_{m+1}, \quad m \in \mathbb{N}_0$$

with

$$\begin{aligned} \hat{b}^{(h)}(x) &= b^{(h)}(x(0), x(-h), \ldots, x(-Rh)), \quad x \in C[-r, 0] \\ \hat{\sigma}^{(h)}(x) &= \sigma^{(h)}(x(0), x(-h), \ldots, x(-Rh)), \quad x \in C[-r, 0]. \end{aligned}$$

If we now compute the conditional expectation of $X^{(h)}$ we obtain in the common fashion

$$\frac{1}{h}E(X^{(h)}_{(m+1)h} - X^{(h)}_{mh} | \mathcal{F}^{X^{(h)}}_{mh}) = b^{(h)}(X^{(h)}_{mh}, X^{(h)}_{(m-1)h}, \ldots, X^{(h)}_{(m-R)h}).$$

Then the drift takes the form

$$b^{(h)*}(x) = b^{(h)}(x(0), x(-h), \ldots, x(-Rh)), \quad x \in C[-r, 0].$$

Now there is an asymptotic behavior. It is easy to establish that

$$b^{(h)*}(x) \xrightarrow[h \to 0]{} b(x(0), x(0), \ldots, x(0))$$

uniformly on compacts of $C[-r, 0]$, if one assumes that $b^{(h)}$ tends to a continuous function b uniformly on compacts of $\mathbb{R}^{(R+1)}$. In the same manner we obtain for the second moments

$$a^{(h)*}(x) \xrightarrow[h \to 0]{} \sigma^2(x(0), x(0), \ldots, x(0))$$

uniformly on compacts of $C[-r, 0]$ if one assumes that $\sigma^{(h)}$ tends to a continuous function σ uniformly on compacts of $\mathbb{R}^{(R+1)}$. After those explanatory considerations we are able to formulate and prove a theorem for time series with fading memory.

2.5.11 Theorem. *Let $(\epsilon_m)_{m \in \mathbb{N}}$ be a sequence of i.i.d. variables on some probability space with*

$$E(\epsilon_1) = 0, \quad E(|\epsilon_1|^2) = 1, \quad E(|\epsilon_1|^{2+\delta}) < \infty$$

for some $\delta > 0$. Let for $h > 0$ the discrete time series $(X^{(h)}_{mh})_{m \geq -R}$ be defined by

$$\begin{cases} X^{(h)}_{mh} &= \xi^{(h)}_i, \quad 0 \leq i \leq R, \quad \xi^{(h)}_i \in \mathbb{R} \\ X^{(h)}_{(m+1)h} &= X^{(h)}_{mh} + b^{(h)}(X^{(h)}_{mh}, X^{(h)}_{(m-1)h}, \ldots, X^{(h)}_{(m-R)h})h \\ &+ \sigma^{(h)}(X^{(h)}_{mh}, X^{(h)}_{(m-1)h}, \ldots, X^{(h)}_{(m-R)h})\sqrt{h}\epsilon_{m+1}, \quad m \in \mathbb{N}_0. \end{cases}$$

The time series $(X^{(h)}_{mh})_{m\geq -R}$ *is extended to a continuous process* $X^{(h)}$ *by linear interpolation. If for the initial conditions*

$$\xi_i^{(h)} \xrightarrow[h\to 0]{} \xi, \quad 0 \leq i \leq R$$

and

$$\begin{aligned} b^{(h)}(x) &\xrightarrow[h\to 0]{} b(x) \\ \sigma^{(h)}(x) &\xrightarrow[h\to 0]{} \sigma(x) \end{aligned}$$

uniformly on compact sets of $\mathbb{R}^{(R+1)}$*, then the processes* $X^{(h)}$ *converge weakly to* X*, where* X *is assumed to be the weak unique solution of the system*

$$\begin{cases} X(0) &= \xi \\ dX(t) &= b(X(t), X(t), \ldots, X(t))\, dt + \sigma(X(t), X(t), \ldots, X(t))\, dB(t), \quad t \geq 0 \end{cases}$$

with initial value $\xi \in \mathbb{R}$ *and continuous coefficients* b *and* σ *with domain* $\mathbb{R}^{(R+1)}$.

Proof. The proof follows directly from Theorem 2.4.16 in its general form for coefficients with domain $C[-r, 0]$. If ϕ is an initial function on $[-r, 0]$ with $\phi(u) \equiv \xi$, then it holds for a sequence $\phi^{(h)}$ of continuous functions with $\phi^{(h)}(-ih) = \xi_i^{(h)}$ for $0 \leq i \leq R$ and $\phi^{(h)}(u) = \xi_R^{(h)}$ on $[-r, -Rh]$ that

$$\sup_{-r\leq u\leq 0} |\phi^{(h)}(u) - \phi(u)| \leq |\xi_R^{(h)} - \xi| + \max_{0\leq i\leq R} |\xi_i^{(h)} - \xi| \xrightarrow[h\to 0]{} 0,$$

since all $\xi_i^{(h)}$ tend to ξ by assumption. □

In the linear case we have the following scheme

$$\begin{cases} X^{(h)}_{mh} &= \xi_i^{(h)}, \quad 0 \leq i \leq R, \quad \xi_i^{(h)} \in \mathbb{R} \\ X^{(h)}_{(m+1)h} &= X^{(h)}_{mh} + \sum_{j=0}^{R} a_j^{(h)} X^{(h)}_{(m-j)h} h + \left(\sum_{i=0}^{R} \sigma_i^{(h)} X^{(h)}_{(m-i)h}\right) \sqrt{h}\epsilon_{m+1}. \end{cases}$$

According to Theorem 2.5.11 it is enough to assume that

$$\xi_i^{(h)} \xrightarrow[h\to 0]{} \xi, \quad 0 \leq i \leq R, \quad \sum_{j=0}^{R} a_j^{(h)} \xrightarrow[h\to 0]{} a, \quad \sum_{i=0}^{R} \sigma_i^{(h)} \xrightarrow[h\to 0]{} \alpha,$$

to ensure that the sequence $\{X^{(h)} : h > 0\}$ converges weakly to the strong unique solution of the system

$$\begin{cases} X(0) &= \xi \\ dX(t) &= aX(t)\, dt + \alpha X(t)\, dB(t), \quad t \geq 0. \end{cases}$$

In slight generalization one can consider coefficients of the form $b^{(h)}(x_0,\dots,x_{R^{(h)}})$, where $R^{(h)}$ tends to infinity, but it still holds that $R^{(h)}h \xrightarrow[h\to 0]{} 0$. Assume that

$$b^{(h)}(x(0),x(-h),\dots,x(-R^{(h)}h)) \xrightarrow[h\to 0]{} b(x(0))$$
$$\sigma^{(h)}(x(0),x(-h),\dots,x(-R^{(h)}h)) \xrightarrow[h\to 0]{} \sigma(x(0))$$

uniformly on compacts of $C[-r,0]$ for continuous functions b and σ with domain $\mathbb{R}$. For $h>0$ the corresponding time series $(X^{(h)}_{mh})_{m\geq -R^{(h)}}$ takes the form

$$\begin{cases} X^{(h)}_{mh} &= \xi^{(h)}_i, \quad 0\leq i\leq R^{(h)}, \quad \xi^{(h)}_i\in\mathbb{R} \\ X^{(h)}_{(m+1)h} &= X^{(h)}_{mh}+b^{(h)}(X^{(h)}_{mh},X^{(h)}_{(m-1)h},\dots,X^{(h)}_{(m-R^{(h)})h})h \\ &+ \sigma^{(h)}(X^{(h)}_{mh},X^{(h)}_{(m-1)h},\dots,X^{(h)}_{(m-R^{(h)})h})\sqrt{h}\epsilon_{m+1}, \quad m\in\mathbb{N}_0. \end{cases}$$

Then the processes $X^{(h)}$, derived from the time series $(X^{(h)}_{mh})_{m\geq -R^{(h)}}$ by linear interpolation, converge weakly to X, where X is assumed to be the unique weak solution of

$$\begin{cases} X(0) &= \xi \\ dX(t) &= b(X(t))\,dt+\sigma(X(t))\,dB(t), \quad t\geq 0, \end{cases}$$

if only for the initial conditions $\xi^{(h)}_i \xrightarrow[h\to 0]{} \xi$. If $b(x)=ax$ and $\sigma(x)=1$, then we recover the one-dimensional Ornstein-Uhlenbeck process.

2.5.6 Counterexamples

The aim of this section is to discuss the assumptions of our convergence theorems. In especially we are interested in what may happen if they are not obeyed.

Typically our convergence theorems had the structure: If for coefficients

$$b^{(h)} \xrightarrow[h\to 0]{} b, \quad \sigma^{(h)} \xrightarrow[h\to 0]{} \sigma$$

in a certain sense, and the martingale problem for b and σ is well-posed, then $\{X^{(h)}: h>0\}$ converges weakly. We shall show now that the by-clause "the martingale problem is well-posed" may not be omitted. The following example gives a sequence of coefficients $\sigma^{(h)}$ which tend to a coefficient σ, but the corresponding processes $X^{(h)}$ do not converge weakly. Define for each $h_n=1/n, n\in\mathbb{N}$ and $x\in\mathbb{R}$

$$\sigma^{(h_n)}(x) := 1(x\neq h_n)\cdot 1(\frac{1}{h_n}\in 2k\mathbb{N})+1(x\neq 0)\cdot 1(\frac{1}{h_n}\in 2k\mathbb{N}+1).$$

It is clear that

$$\sigma^{(h_n)}(x_n) \xrightarrow[n\to\infty]{} \sigma(x) := 1(x\neq 0)$$

for every sequence $\{x_n\} \subset \mathbb{R}$ approximating $x \in \mathbb{R}$ with exception of the discontinuity point $x = 0$. Take a sequence of independent, standard normally distributed random variables $(\epsilon_i)_{i\in\mathbb{N}}$ and define the following time series in terms of $\sigma^{(h_n)}$

$$X^{(h_n)}_{(m+1)h_n} = X^{(h_n)}_{mh_n} + \sigma^{(h_n)}(X^{(h_n)}_{mh_n})\sqrt{h_n}\epsilon_{m+1}, \quad m \in \mathbb{N}_0, \quad X^{(h)}_0 = 0.$$

Then one could think that the sequence $\{X^{(h_n)} : n \in \mathbb{N}\}$ converges weakly. But it does not. One checks easily by definition of $\sigma^{(h_n)}$ that for each $m \in \mathbb{N}_0$

$$X^{(h_n)}_{mh_n} = \left(\sum_{i=0}^{m-1} \sqrt{h_n}\epsilon_{i+1}\right) 1(\frac{1}{h_n} \in 2k\mathbb{N}).$$

This yields

$$X^{(h_{n'})} \xrightarrow[n'\to\infty]{d} G, \quad X^{(h_{n''})} \xrightarrow[n''\to\infty]{d} 0,$$

where G denotes a Brownian motion, n' the sequence of even numbers and n'' the sequence of uneven numbers. There are two different limit points, the whole sequence $\{X^{(h_n)} : n \in \mathbb{N}\}$ does not converge weakly. What goes wrong? For the stochastic differential equation

$$\begin{cases} X(0) &= 0 \\ dX(t) &= \sigma(X(t))\, dB(t) = 1(X(t) \neq 0)\, dB(t), \quad t \geq 0 \end{cases} \tag{2.5.15}$$

the martingale problem is not well-posed. There are at least two different weak solutions $X_1 \equiv G$ and $X_2 \equiv 0$. X_1 is the weak limit of $X^{(h_{n'})}$ and X_2 the weak limit of $X^{(h_{n''})}$. There are even infinitely many strong solutions of equation (2.5.15). Given a Brownian motion B, denote by τ a time where $B(\tau) = 0$. Then the continuous process X defined by

$$X(t) := B(t), \quad 0 \leq t \leq \tau, \quad X(t) := 0, \quad t \geq \tau$$

solves equation (2.5.15). There are infinitely many such times τ in the one-dimensional case.

We turn to another type of assumptions for discontinuous coefficients. We pointed out the requirement for the asymptotic behavior of processes $X^{(h)}$

$$\varlimsup_{h\to 0} P(X^{(h)}(u) \in B_\delta(x)) \xrightarrow[\delta\to 0]{} 0 \quad \forall x \in \mathbb{R}.$$

In the next example we shall construct a sequence of processes $X^{(h)}$ where this condition is violated. Define for each $h > 0$

$$\begin{aligned} b^{(h)}(x) := 1(x = 0) &\xrightarrow[h\to 0]{} 1(x = 0) =: b(x), \quad x \in \mathbb{R} \\ \sigma^{(h)}(x) := 0 &\xrightarrow[h\to 0]{} 0 =: \sigma(x), \quad x \in \mathbb{R}. \end{aligned}$$

The corresponding deterministic time series $X^{(h)}$ takes the form

$$X^{(h)}_{(m+1)h} = X^{(h)}_{mh} + 1(X^{(h)}_{mh} = 0), \quad m \in \mathbb{N}_0, \quad X^{(h)}_0 := 0.$$

Then it holds for each $m \in \mathbb{N}$ that $X^{(h)}_{mh} = h$. Therefore $\{X^{(h)} : h > 0\}$ converges weakly to the process $X \equiv 0$. This suggests that X is the weak solution of the "stochastic" equation

$$\begin{cases} X(0) &=& 0 \\ dX(t) &=& b(X(t))\,dt + \sigma(X(t))\,dB(t), \quad t \geq 0 \end{cases}$$

for $b(x) = 1(x = 0)$ and $\sigma = 0$. In other words, it should hold for $X \equiv 0$ that

$$X(t) = X(0) + \int_0^t 1(X(u) = 0)\,du.$$

But this is not the case since the integrand is 1, $X \equiv 0$ does not solve the "martingale problem" for b and σ which are the limit coefficients of $b^{(h)}$ and $\sigma^{(h)}$. What goes wrong? For $y = 0$ and $h < \delta$ it holds by construction that

$$P(X^{(h)}(u) \in B_\delta(y)) = 1,$$

and hence

$$\lim_{h \to 0} P(X^{(h)}(u) \in B_\delta(y)) \underset{\delta \to 0}{\longrightarrow} 1, \quad y = 0.$$

This tells us that the requirement that the lower limit of such type should tend to zero may not be omitted in the case of discontinuous coefficients. More exactly speaking, if b is measurable and bounded, then the sequence of the linearly interpolated deterministic processes, derived from the series

$$x^{(h)}_{(m+1)h} = x^{(h)}_{mh} + b(x^{(h)}_{mh})h, \quad m \in \mathbb{N}_0, \quad x^{(h)}_0 = 0$$

is "tight". There exists $x \in C[0,T]$ such that the sequence $\{x^{(h)} : h > 0\}$ converges to x uniformly on $[0,T]$. If

$$\int_0^T 1(x(u) \in D_b)\,du = 0, \quad D_b := \text{set of discontinuity points of } b, \tag{2.5.16}$$

then x solves the equation $dx(t) = b(x(t))\,dt$. If this is violated, as in the above example, then the limit function x need not solve $dx(t) = b(x(t))\,dt$. Note that (2.5.16) is the deterministic analogue to

$$Q\left(\int_0^T 1(X^\circ(u) \in D_b)\,du = 0\right) = 1,$$

which we used in the formulation of Theorem 2.4.4.

2.6 Solutions of Stochastic Delay Diff. Equations as Semimartingales

We assume in the sequel that the reader is familiar with semimartingale theory. For simplicity we restrict to the one-dimensional case. Suppose that the following stochastic delay differential equation

$$\begin{cases} Y_0 &= \xi \\ dY(t) &= b(Y_t)\,dt + \sigma(Y_t)\,dW(t), \quad t \geq 0, \end{cases} \tag{2.6.1}$$

driven by a Brownian motion W, has a unique weak solution. Then the solution Y of (2.6.1) is a semimartingale with respect to its natural filtration $\mathcal{F}^Y$. The characteristics (B, C, ν) of Y are for $t \geq 0$

$$B(t) = \int_0^t b(Y_u)\,du, \quad C(t) = \int_0^t \sigma^2(Y_u)\,du, \quad \nu \equiv 0.$$

This comes from the fact that $M(t) := \int_0^t \sigma(Y_u)\,dW(u)$ is a local martingale, and the function $t \mapsto \int_0^t b(X_u)\,du$ has bounded variation on compact intervals. Finally, ν is identically zero since the process Y is continuous. It is also necessary to consider the modified characteristic $\tilde{C}$ which is defined by

$$\tilde{C} := C + h^2 * \nu - \sum_{s \leq \cdot} |\triangle B(s)|^2$$

for a fixed truncation function h. Note that for the solution Y of (2.6.1) the characteristics $\tilde{C}$ and C coincide since the solution process Y is continuous, its jump measure and predictable compensator ν are zero, and the characteristic B is continuous. Now consider for each $h > 0$ an adapted series $\{X_{mh}^{(h)} : m \in \mathbb{N}_0\}$ with increments $U_{mh}^{(h)} := \triangle X_{mh}^{(h)} = X_{mh}^{(h)} - X_{(m-1)h}^{(h)}$ (and $U_0^{(h)} := X_0^{(h)}$) in discrete time on the stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_{mh}^{(h)}), P)$. Then the continuous-time process $Y^{(h)}$ defined by

$$Y_t^{(h)} := \sum_{0 \leq k \leq [\frac{t}{h}]} U_{kh}^{(h)} = X_{[\frac{t}{h}]h}^{(h)}, \quad t \geq 0$$

is a semimartingale with respect to its natural filtration $\mathcal{F}^{Y^{(h)}}$ with start in $X_0^{(h)}$. Note that, in contrast to the preceding sections, $Y^{(h)}$ is a random variable on $D[0, \infty)$, the space of right-continuous functions with left-hand limits. We assume that $D[0, \infty)$ is endowed with the Skorochod topology. Fix a truncation function ϕ. Then, referring to Theorem II.3.11 in Jacod and Shiryaev [10], the characteristics $(B^{(h)}, C^{(h)}, \nu^{(h)})$ of

$Y^{(h)}$ have for $t \geq 0$ the form

$$\begin{cases} B_t^{(h)} &= \sum_{0\leq k\leq[\frac{t}{h}]-1} E(\phi(U_{(k+1)h}^{(h)})|\mathcal{F}_{kh}^{(h)}) \\ C_t^{(h)} &= 0 \\ \tilde{C}_t^{(h)} &= \sum_{0\leq k\leq[\frac{t}{h}]-1} E(\phi^2(U_{(k+1)h}^{(h)})|\mathcal{F}_{kh}^{(h)}) - \sum_{0\leq k\leq[\frac{t}{h}]-1} \left|E(\phi(U_{(k+1)h}^{(h)})|\mathcal{F}_{kh}^{(h)})\right|^2 \\ \nu([0,t]\times g) &= \sum_{0\leq k\leq[\frac{t}{h}]-1} E(g(U_{(k+1)h}^{(h)})1_{\{U_{(k+1)h}^{(h)}\neq 0\}}|\mathcal{F}_{kh}^{(h)}), \quad \text{for } g\geq 0 \quad \text{Borel-mb.} \end{cases}$$

We extend $Y^{(h)}$ from a process on $D[0,\infty)$ to a process on $D[-r,\infty)$ by demanding that it starts in the initial function $\xi \in C[-r,0]$. Jacod and Shiryaev [10] treat the topic of convergence of a sequence of semimartingales. The aim of this chapter is to give a second proof of Theorem 2.5.1 by means of semimartingale theory if the coefficients b and σ of the SDDE in (2.6.1) are continuous and bounded. We introduce the notations

$$l_{kh}^{(h)}Y^{(h)} := l^{(h)}(Y_{kh}^{(h)},\dots,Y_{(k-r^{(h)})h}^{(h)}), \quad l_tY^{(h)} := (Y^{(h)}(t+u))_{-r\leq u\leq 0}.$$

We shall work with the same truncation function $\phi(x) = x1_{\{|x|\leq 1\}}$ as in the preceding sections. According to the setting of the previous sections assume that for $h > 0$ there are measurable functions $b^{(h)}, \sigma^{(h)}, \triangle_\delta^{(h)}$, all defined on $C[-r,0]$, such that for all $k \in \mathbb{N}_0$ a.s.

$$\begin{aligned} b^{(h)}(l_{kh}^{(h)}Y^{(h)}) &= \frac{1}{h}E(\phi(U_{(k+1)h}^{(h)})|\mathcal{F}_{kh}^{(h)}) \\ a^{(h)}(l_{kh}^{(h)}Y^{(h)}) &= \frac{1}{h}E(\phi^2(U_{(k+1)h}^{(h)})|\mathcal{F}_{kh}^{(h)}) \\ \triangle_\epsilon^{(h)}(l_{kh}^{(h)}Y^{(h)}) &= \frac{1}{h}P(|U_{(k+1)h}^{(h)})| > \epsilon|\mathcal{F}_{kh}^{(h)}) \end{aligned}$$

for the filtration

$$\mathcal{F}_{kh}^{(h)}) = \sigma(Y_0^{(h)},\dots,Y_{kh}^{(h)}), \quad k\in\mathbb{N}_0.$$

Now we are able to formulate and prove a theorem which is the semimartingale analogue to Theorem 2.4.2. An important modification is that the limit coefficients have domain $D[-r,0]$.

2.6.1 Theorem. *Let there be continuous bounded functions b and a such that for every compact set K of $C[-r,0]$*

$$\begin{aligned} \sup_{x\in K}|b^{(h)}(x)-b(x)| &\xrightarrow[h\to 0]{} 0 \\ \sup_{x\in K}|a^{(h)}(x)-a(x)| &\xrightarrow[h\to 0]{} 0. \end{aligned}$$

Suppose that the limit functions b and a are also defined on $D[-r,0]$. We demand that they are bounded on $D[-r,0]$, and that uniformly on compacts of $C[-r,0]$

$$b(\bar{x}^{(h)}) \xrightarrow[h\to 0]{} b(x), \quad a(\bar{x}^{(h)}) \xrightarrow[h\to 0]{} a(x),$$

where $\bar{x}^{(h)}$ is defined by $\bar{x}^{(h)}(u) := x([\frac{u}{h}]h)$. Assume that in addition it holds

$$\sup_{h>0} \sup_{x\in C[-r,0]} |b^{(h)}(x)| + |a^{(h)}(x)| < \infty$$

and for each $\epsilon > 0$

$$\sup_{x\in C[-r,0]} \triangle_\epsilon^{(h)}(x) \xrightarrow[h\to 0]{} 0.$$

Then the sequence $\{Y^{(h)} : h > 0\}$ is tight, and every limit point is a semimartingale on $D[-r,\infty)$ with start in ξ and with characteristics

$$\nu \equiv 0, \quad B(t)(m) = \int_0^t b(m_u)\,du, \quad C(t)(m) = \int_0^t a(m_u)\,du, \quad m \in D[-r,\infty)$$

for $t \geq 0$.

Proof. Our aim is to apply Theorem IX.2.11 in Jacod and Shiryaev [10] for tight sequences. To this end we have to show first of all that the sequence $\{Y^{(h)} : h > 0\}$ is tight on $D[0,\infty)$. This we shall do by the tightness criterion Theorem VI.4.18 in Jacod and Shiryaev [10]. We start with condition (iii) of that theorem and show that the sequence of characteristics $B^{(h)}$ is C-tight. By assumption we have that

$$\begin{aligned} B_t^{(h)} &= \sum_{0\le k\le [\frac{t}{h}]-1} E(\phi(U_{(k+1)h}^{(h)})|\mathcal{F}_{kh}^{(h)}) \\ &= \sum_{0\le k\le [\frac{t}{h}]-1} b^{(h)}(l_{kh}^{(h)}Y^{(h)})h = \int_0^{[\frac{t}{h}]h} b^{(h)}(l_{[\frac{u}{h}]h}^{(h)}Y^{(h)})\,du. \end{aligned}$$

One sees from Theorem 15.5 in Billingsley [2] that a sequence of processes of the form

$$I_t^{(h)} = \int_0^{[\frac{t}{h}]h} g^{(h)}\,du, \quad |g^{(h)}| \le K, \quad K = \text{constant}$$

is C-tight, since it holds for the modulus of continuity $w_{I^{(h)}}(\delta)$ that $w_{I^{(h)}}(\delta) \le K\delta$. Since the functions $b^{(h)}$ were assumed to be uniformly bounded, the sequence $\{B^{(h)} : h > 0\}$ is C-tight. For the component $\tilde{C}_t^{(h)}$ we may write

$$\tilde{C}_t^{(h)} = \sum_{0\le k\le [\frac{t}{h}]-1} E(\phi^2(U_{(k+1)h}^{(h)})|\mathcal{F}_{kh}^{(h)}) - \sum_{0\le k\le [\frac{t}{h}]-1} \left|E(\phi(U_{(k+1)h}^{(h)})|\mathcal{F}_{(kh)}^{(h)})\right|^2. \tag{2.6.2}$$

The second summand tends uniformly to the zero function on any compact interval $[0,T]$ in a.s. and in L^1 since

$$\begin{aligned}\sup_{0\le t\le T}\sum_{0\le k\le[\frac{t}{h}]-1}\left|E(\phi(U^{(h)}_{(k+1)h})|\mathcal{F}^{(h)}_{kh})\right|^2 &= \sup_{0\le t\le T}\sum_{0\le k\le[\frac{t}{h}]-1}|b^{(h)}(l^{(h)}_{[\frac{u}{h}]h}Y^{(h)})|^2h^2\\ &\le \sup_{h>0}\|b^{(h)}\|^2_\infty[\frac{T}{h}]h^2 \xrightarrow[h\to 0]{} 0.\end{aligned}$$

Therefore it suffices to show that the sequence of processes in the first summand in (2.6.2) is C-tight. But this sum can also be represented in an integral form, since we have by assumption that

$$\sum_{0\le k\le[\frac{t}{h}]-1}E(\phi^2(U^{(h)}_{(k+1)h})|\mathcal{F}^{(h)}_{kh}) = \int_0^{[\frac{t}{h}]h} a^{(h)}(l^{(h)}_{[\frac{u}{h}]h}Y^{(h)})\,du.$$

Since the functions $a^{(h)}$ are uniformly bounded by assumption, we have shown that the sequence $\tilde{C}^{(h)}$ is C-tight. Consider for $x\in\mathbb{R}$ the function $g_p(x):=(p|x|-1)^+\wedge 1$ for $p\in\mathbb{N}$. Clearly, it holds that

$$g_p(x)=0,\quad x\in B_{\frac{1}{p}}(0),\quad \sup_{x\in\mathbb{R}}|g(x)|\le 1.$$

Then we see that

$$E(g_p(U^{(h)}_{(k+1)h})1_{\{U^{(h)}_{(k+1)h}\neq 0\}}|\mathcal{F}^{(h)}_{kh}) \le E(1_{\{|U^{(h)}_{(k+1)h}|>\frac{1}{p}\}}|\mathcal{F}^{(h)}_{kh}) = \triangle^{(h)}_{\frac{1}{p}}(l^{(h)}_{kh}Y^{(h)})h.$$

This allows us to estimate

$$\begin{aligned}\sup_{0\le t\le T}(g_p*\nu^{(h)}_t) &= \sup_{0\le t\le T}(\nu^{(h)}_t([0,t]\times g_p)\le \sup_{0\le t\le T}\left(\sum_{0\le k\le[\frac{t}{h}]-1}E(1_{\{|U^{(h)}_{(k+1)h}|>\frac{1}{p}\}}|\mathcal{F}^{(h)}_{kh})\right)\\ &= \sup_{0\le t\le T}\left(\sum_{0\le k\le[\frac{t}{h}]-1}\triangle^{(h)}_{\frac{1}{p}}(l^{(h)}_{kh}Y^{(h)})h\right)\\ &\le [\frac{T}{h}]h\sup_{x\in C[-r,0]}\triangle^{(h)}_{\frac{1}{p}}(x)\xrightarrow[h\to 0]{} 0,\end{aligned}$$

where the convergence is also in L^1. Therefore the sequence $(g_p*\nu^{(h)})$ is also C-tight. Condition (ii) in the tightness criterion reads as follows

$$\lim_{a\uparrow\infty}\overline{\lim_{h\to 0}}P(\nu^{(h)}([0,N]\times\{x:|x|>a\})>\epsilon)=0.$$

But we have just seen that the sequence of random variables $\nu^{(h)}([0,N]\times 1_{\{|x|>a\}})$ for each $a\in\mathbb{R}$ tends to zero in L^1, thus also in probability. Finally, condition (i) is met since $Y_0^{(h)} = \xi(0)$ for all $h > 0$. Thereby, invoking Theorem VI.4.18 in Jacod and Shiryaev [10], the sequence $\{Y^{(h)} : h > 0\}$ is tight on $D[0,\infty)$. We shall now come to the conditions of Theorem IX.2.11 in Jacod and Shiryaev [10]. According to requirement (i) of that theorem we have to check the following conditions

$$\begin{cases} [\beta_7 - \mathbb{R}_+] & B_t^{(h)} - B(t) \circ Y^{(h)} \xrightarrow[h\to 0]{P} 0 \quad t \geq 0 \\ [\gamma_7 - \mathbb{R}_+] & \tilde{C}_t^{(h)} - \tilde{C}(t) \circ Y^{(h)} \xrightarrow[h\to 0]{P} 0 \quad t \geq 0 \\ [\delta_{7,1} - \mathbb{R}_+] & g * \nu^{(h)} \xrightarrow[h\to 0]{P} 0 \quad t \geq 0, \quad g \in C_1(\mathbb{R}). \end{cases}$$

We shall start with condition $[\beta_7 - \mathbb{R}_+]$. Note that the random variable $B_t^{(h)}$ and the (well-defined) random variable $B(t) \circ Y^{(h)}$ live on the same probability space $(\Omega, \mathcal{F}, P)$. We have to show that their difference tends to zero in probability. By assumption we have that

$$\begin{aligned} B_t^{(h)} = \sum_{0 \leq k \leq [\frac{t}{h}]-1} E(\phi(U_{(k+1)h}^{(h)}) | \mathcal{F}_{kh}^{(h)}) &= \int_0^{[\frac{t}{h}]h} b^{(h)}(l_{[\frac{u}{h}]h}^{(h)} Y^{(h)})\, du \\ B(t) \circ Y^{(h)} &= \int_0^t b(l_u Y^{(h)})\, du. \end{aligned}$$

Now we introduce a continuous auxiliary process $Z^{(h)}$ which equals $Y^{(h)}$ at discrete time points and which is linearly interpolated. Since

$$l_{kh}^{(h)} Y^{(h)} = l^{(h)}(Y_{kh}^{(h)}, \ldots, Y_{(k-r^{(h)})h}^{(h)}) = l^{(h)}(Z_{kh}^{(h)}, \ldots, Z_{(k-r^{(h)})h}^{(h)}) = l_{kh}^{(h)} Z^{(h)},$$

we have the following property

$$B_t^{(h)} = \int_0^{[\frac{t}{h}]h} b^{(h)}(l_{[\frac{u}{h}]h}^{(h)} Y^{(h)})\, du = \int_0^{[\frac{t}{h}]h} b^{(h)}(l_{[\frac{u}{h}]h}^{(h)} Z^{(h)})\, du.$$

Furthermore it holds that $Y^{(h)} = \overline{Z^{(h)}}^{(h)}$ if $\overline{m}^{(h)}(u) := m([\frac{u}{h}]h)$. Thereby we can rewrite $B(t) \circ Y^{(h)}$ in terms of $Z^{(h)}$ in the following way

$$B(t) \circ Y^{(h)} = \int_0^t b(l_u Z^{(h)})\, du + \int_0^t [b(l_u \overline{Z^{(h)}}^{(h)}) - b(l_u Z^{(h)})]\, du.$$

If K is a compact set on $C[-r,\infty)$, then

$$A := \overline{\bigcup_{0 \leq u \leq t} \{l_{[\frac{u}{h}]h}^{(h)} m : m \in K, h > 0\} \cup \{m_u : m \in K\}}$$

is a compact subset of $C[-r, 0]$. Therefore we obtain

$$\sup_{m \in K} \sup_{0 \leq u \leq t} |b^{(h)}(l_{[\frac{u}{h}]h}^{(h)} m) - b(l_{[\frac{u}{h}]h}^{(h)} m)| \leq \sup_{x \in A} |b^{(h)}(x) - b(x)| \xrightarrow[h \to 0]{} 0$$

and

$$\sup_{m\in K}\sup_{0\le u\le t}|b(l^{(h)}_{[\frac{u}{h}]h}m)-b(l_u m)|\le \sup_{m\in K}\sup_{x,x'\in A,\|x-x'\|_\infty\le w_m(h)}|b(x)-b(x')|\xrightarrow[h\to 0]{}0,$$

where $w_m(h)$ denotes the modulus of continuity on $C[-r,0]$. Thus we have shown that

$$\sup_{m\in K}\left|\int_0^{[\frac{t}{h}]h} b^{(h)}(l^{(h)}_{[\frac{u}{h}]h}m)\,du-\int_0^t b(l_u m)\,du\right|\xrightarrow[h\to 0]{}0,$$

having used uniform boundedness of $b^{(h)}$ and boundedness and uniform continuity of b on compact sets of $C[-r,0]$. Furthermore we have that

$$\sup_{m\in K}\sup_{0\le u\le t}|b(l_u\overline{m}^{(h)})-b(l_u m)|\le\sup_{x\in A}|b(\bar{x}^{(h)})-b(x)|\xrightarrow[h\to 0]{}0,$$

where we used the assumption of the theorem for the last convergence. Therefore we can deduce that

$$|B^{(h)}_t-B(t)\circ Y^{(h)}|1_{\{Z^{(h)}\in K\}}\xrightarrow[h\to 0]{}0$$

a.s. and by dominated convergence (all expressions are bounded) also in L^1. Combining Theorem 15.2, relation (14.9) and Theorem 8.2 in Billingsley [2] we see that the sequence $\{Z^{(h)}:h>0\}$ is tight on $C[0,\infty)$, if only the sequence $\{Y^{(h)}:h>0\}$ is tight on $D[0,\infty)$. Now we split up in the following way

$$\begin{aligned}E|B^{(h)}_t-B(t)\circ Y^{(h)}| &= E(|B^{(h)}_t-B(t)\circ Y^{(h)}|1_{\{Z^{(h)}\in K\}})\\ &+ E(|B^{(h)}_t-B(t)\circ Y^{(h)}|1_{\{Z^{(h)}\in K^C\}}).\end{aligned}$$

In view of tightness of the sequence $\{Z^{(h)}:h>0\}$ (which we already established) there exists a compact set K such that for given $\epsilon>0$

$$E(|B^{(h)}_t-B(t)\circ Y^{(h)}|1_{\{Z^{(h)}\in K^C\}})\le tM\epsilon,$$

where the constant M depends on the uniform upper bounds for $b^{(h)}$ and b. Then we established that the expectation

$$E(|B^{(h)}_t-B(t)\circ Y^{(h)}|1_{\{Z^{(h)}\in K\}})$$

tends to zero as h to zero. Therefore $B^{(h)}_t$ tends to $B(t)\circ Y^{(h)}$ in L^1 and in probability as desired. Next we shall check condition $[\gamma_7-\mathbb{R}_+]$. We recall that the characteristic $\tilde{C}^{(h)}$ has the form

$$\tilde{C}^{(h)}_t = \sum_{0\le k\le[\frac{t}{h}]-1}E(\phi^2(U^{(h)}_{(k+1)h})|\mathcal{F}^{(h)}_{kh})-\sum_{0\le k\le[\frac{t}{h}]-1}\left|E(\phi(U^{(h)}_{(k+1)h})|\mathcal{F}^{(h)}_{kh})\right|^2,$$

and that the second sum tends to zero in L^1. For the first summand it holds that

$$\left| \int_0^{[\frac{t}{h}]h} a^{(h)}(l^{(h)}_{[\frac{u}{h}]h} Y^{(h)})\, du - \int_0^t a(l_u Y^{(h)})\, du \right| \xrightarrow[h\to 0]{P} 0,$$

repeating the arguments for $a^{(h)}$ and a instead of $b^{(h)}$ and b. Therefore condition $[\gamma_7 - \mathbb{R}_+]$ holds. It remains to check condition $[\delta_{7,1} - \mathbb{R}_+]$. We shall do this for $[\delta_{7,2} - \mathbb{R}_+]$ with the underlying function class $C_2(\mathbb{R})$, see Jacod and Shiryaev [10]. It is the set of all continuous bounded functions which are zero in a neighborhood of zero and have a limit at infinity. Let $g_\epsilon \in C_2(\mathbb{R})$ be given with the properties

$$g_\epsilon(x) = 0, \quad x \in B_\epsilon(0), \quad \sup_{x\in\mathbb{R}} |g_\epsilon(x)| \le K < \infty.$$

We already established that $(g_p * \nu^{(h)}_t)$, therefore $(g_p * \nu^{(h)}_t)K$ tends to zero in L^1 for each $p \in \mathbb{N}$. Hence does $(g_\epsilon * \nu^{(h)}_t)$ for each $\epsilon > 0$. Therefore condition $[\delta_{7,2} - \mathbb{R}_+]$ holds for $\nu \equiv 0$, where ν is understood to be defined on $D[0,\infty)$. This means that the sequence $\{Y^{(h)} : h > 0\}$ itself is C-tight. This means, every limit point Q of the laws of $\{Y^{(h)} : h > 0\}$ is concentrated on $C[-r,\infty)$. We remark that condition (ii) of Theorem IX.2.11 in Jacod and Shiryaev [10]

$$\sup_{m\in D[-r,\infty)} |\tilde{C}(t)(m)| < \infty, \qquad \sup_{m\in D[-r,\infty)} |g * \nu_t(m)| < \infty$$

is met by assumption for $\tilde{C}$. To check condition (iii) we invoke that we already established that $\nu \equiv 0$. The functions

$$m \mapsto B(t)(m) = \int_0^t b(m_u)\, du, \quad m \mapsto \tilde{C}(t)(m) = \int_0^t a(m_u)\, du$$

are continuous for all $m \in C[-r,\infty)$ and Q-almost surely for all $m \in D[-r,\infty)$. We have proved the theorem. □

If every random variable $Y^{(h)}_{kh}$ is $(2+\delta)$-integrable for some $\delta > 0$, one needs not to truncate. Assume that for $h > 0$ there are measurable functions $b^{(h)*}, a^{(h)*}, \triangle^{(h)*}_\delta$, all defined on $C[-r,0]$, such that for all $k \in \mathbb{N}_0$ a.s.

$$\begin{aligned} b^{(h)*}(l^{(h)}_{kh} Y^{(h)}) &= \frac{1}{h} E(U^{(h)}_{(k+1)h} | \mathcal{F}^{(h)}_{kh}) \\ a^{(h)*}(l^{(h)}_{kh} Y^{(h)}) &= \frac{1}{h} E(|U^{(h)}_{(k+1)h}|^2 | \mathcal{F}^{(h)}_{kh}) \\ \triangle^{(h)*}_\delta(l^{(h)}_{kh} Y^{(h)}) &= \frac{1}{h} E(|U^{(h)}_{(k+1)h})|^{2+\delta} | \mathcal{F}^{(h)}_{kh}). \end{aligned}$$

Exactly as in the proof of Theorem 2.4.16 we have the following implications

$$\begin{aligned} \forall \delta > 0 \quad \sup_{x\in C[-r,0]} \triangle^{(h)*}_\delta \xrightarrow[h\to 0]{} 0 &\implies \sup_{x\in C[-r,0]} \triangle^{(h)}_\epsilon \xrightarrow[h\to 0]{} 0 \quad \forall \epsilon > 0 \\ \sup_{x\in K} |b^{(h)*}(x) - b(x)| \xrightarrow[h\to 0]{} 0 &\implies \sup_{x\in K} |b^{(h)}(x) - b(x)| \xrightarrow[h\to 0]{} 0 \\ \sup_{x\in K} |a^{(h)*}(x) - a(x)| \xrightarrow[h\to 0]{} 0 &\implies \sup_{x\in K} |a^{(h)}(x) - a(x)| \xrightarrow[h\to 0]{} 0, \end{aligned}$$

and Theorem 2.6.1 is applicable for the not truncated expectations $b^{(h)*}$ and $a^{(h)*}$.

In the previous sections the approximating process $X^{(h)}$ was linearly interpolated. Is $X^{(h)}$ also a semimartingale? The answer to this question is positive since a linearly interpolated process obviously has bounded variation. Therefore $X^{(h)}$ is a semimartingale with characteristics $(B^{(h)}, C^{(h)}, \nu^{(h)}) = (X^{(h)}, 0, 0)$ with respect to its natural filtration $\mathcal{F}^{X^{(h)}}$. This shows us the following: If a sequence of linearly interpolated processes $X^{(h)}$ converges weakly to a continuous process Y of unbounded variation, then this cannot be verified with the tools of semimartingale theory since the conditions

$$\begin{cases} [\beta_7 - \mathbb{R}_+] & B_t^{(h)} - B(t) \circ Z^{(h)} \xrightarrow[h \to 0]{P} 0 \quad t \geq 0 \\ [\gamma_7 - \mathbb{R}_+] & \tilde{C}_t^{(h)} - \tilde{C}(t) \circ Z^{(h)} \xrightarrow[h \to 0]{P} 0 \quad t \geq 0 \end{cases}$$

are not fulfilled for the semimartingale components $(B, \tilde{C})$ of Y. It gives us also the following warning about weak convergence of semimartingales: If a sequence of semimartingales $X^{(h)}$ with characteristics $(B^{(h)}, \tilde{C}^{(h)}, \nu^{(h)})$ converges weakly to another semimartingale Y with characteristics $(B, \tilde{C}, \nu)$, then it need not hold that the sequence of characteristics of $X^{(h)}$ converges weakly to the characteristics of Y for any choice of all characteristics involved. Theorem VII.3.4. in Jacod and Shiryaev [10] gives conditions for convergence of the characteristics if all underlying processes have independent increments. We close this section with the following remark.

2.6.2 Remark. *The strongest aspect in the theory of semimartingales in Jacod and Shiryaev [10] is that also processes with jumps are included. But this thesis is restricted to continuous processes.*

2.7 Comparison to Literature

This section is devoted to citing already known results on approximation of stochastic delay differential equations. There are two types of approximation: weak and strong approximation.

We shall begin with weak approximation. A sequence of probability measures $\{P^{(h)} : h > 0\}$ converges weakly to another probability measure Q if $\int f \, dP^{(h)} \xrightarrow[h \to 0]{} \int f \, dQ$ for all bounded, continuous functions f. In this chapter, with exception of the preceding section, the domain of $P^{(h)}$ and Q is the space $C[-r, \infty)$ endowed with the Skorochod topology. But it is also to study the laws for a fixed time points T. If

$$|E(f(X_T^{(h)})) - E(f(X(T)))| \xrightarrow[h \to 0]{} 0 \tag{2.7.1}$$

holds for each $T > 0$, then the sequence $\{X_T^{(h)} : h > 0\}$ converges weakly to $X(T)$. In this case one can also say that the processes $X^{(h)}$ converge weakly to the process X, since the random variables $X_T^{(h)}$ converge weakly to $X(T)$ for all $T > 0$. Usually one goes one step further and investigates how fast the convergence in (2.7.1) is obtained. A special class of continuous functions ϕ is picked out and the following property is studied

$$|E(\phi(X_T^{(h)})) - E(\phi(X(T)))| \leq Kh^\beta, \quad \beta > 0,$$

where the constant K depends on the function ϕ, the initial data and the time point T, but not on the step length h. In this case the sequence of approximations $X^{(h)}$ converges weakly to the process X with order β for the class of test functions ϕ. In Küchler and Platen [15] the d-dimensional delay equation of the form

$$dX(t) = a(X(t), X(t-r))\,dt + \sum_{j=1}^{d} b^j(X(t), X(t-r))\,dB^j(t)$$

is approximated weakly in the described sense with order 1 and with order 2. In Buckwar and Shardlow [3] the d-dimensional delay equation

$$dX(t) = \left(\int_{-r}^{0} X(t+u)\,da(u) + f(X(t))\right) dt + b(X(t))\,dB(t)$$

is approximated weakly with order 1 in the case of continuous delay

$$\left|\int_{-r}^{0} g(u)\,da(u)\right| \leq \int_{-r}^{0} \bar{a}(u)|g(u)|\,du.$$

Kloeden and Platen [13] treat weak approximation of higher orders for stochastic ordinary differential equations. The arguments are based on an Itô-Taylor formula and use the Markov property of the solution process.

In comparison to above cited authors, this thesis does not contain any estimations of orders. An estimation of orders on the space $C[-r,\infty)$ has not been done yet. To tackle this aspect, one has to investigate the inequality

$$|E(\phi(X^{(h)})) - E(\phi(X))| \leq Kh^\beta, \quad \beta > 0,$$

where ϕ belongs to a class of continuous test functions with domain $C[-r,\infty)$. One could also introduce an appropriate distance d between two probability measures on $C[-r,\infty)$ and study the relation

$$d(P^{(h)}, Q) \leq Kh^\beta,$$

where $P^{(h)}$ is the law of the process $X^{(h)}$ and Q the law of the solution process X. Those inequalities are surely not easy to establish if X is the solution of a stochastic equation, even in the ordinary case.

There is also weak approximation in another sense if it is known in advance that the process X has a density $p(T,x)$ at time T. In this case one studies the inequality

$$\sup_{x\in\mathbb{R}} |E(f_T^h - x) - p(T,x)| \leq Kh^\beta, \tag{2.7.2}$$

where the functions f_T^h are certain algorithms depending on the approximations $X^{(h)}$. In Kohatsu-Higa [14] the diffusion with boundary conditions

$$\begin{cases} dX(t) &= b(X(t))\,dt + \sigma(X(t)) \circ dB(t), \quad 0 \le t \le 1 \\ h_0 &= F_0 X(0) + F_1 X(1) \end{cases}$$

is studied. The stochastic integral is the Stratonovich integral. Note that due to the boundary condition the process X is anticipative. Therefore the proof in Kohatsu-Higa [14] contains Malliavin calculus. Concerning ordinary equations it is a success of the Malliavin calculus to prove existence of a density if Hörmander conditions are fulfilled. In the one-dimensional case they reduce to the condition $\sigma(\xi) \neq 0$ or $\sigma^n(\xi)b(\xi) \neq 0$ for some $n \in \mathbb{N}$, if for the solution X it holds that $X(0) = \xi$. The existence of a density for delay equations is proven in Bell and Mohammed [1] for the system

$$dX(t) = H(t, X) + g(t, X(t-r))\,dB(t)$$

for a non-anticipating functional H. In Hu, Mohammed, and Yan [8] it is shown in the case of point delay that $X(T) \in D^{1,\infty}$ for the solution X. An inequality of the type (2.7.2) has not been established yet for the approximation of densities of SDDE's. A proof of this inequality would use Malliavin calculus. This technique is not present in this thesis.

We shall now cite results on strong approximations. Assume that the solution process X and all approximating processes $X^{(h)}$ are defined on a common probability space. Then one can measure the L^1-distance between X and $X^{(h)}$. The approximations $X^{(h)}$ converge strongly to X with order γ if

$$E|X_T^{(h)} - X(T)| \le K h^\gamma$$

for a constant K not depending on h. Usually the approximation $X^{(h)}$ is constructed in terms of the driving force B of the process X. Strong approximations of SDDE's are treated in Hu, Mohammed, and Yan [8]. Equations of the form

$$dX(t) = h(t, \Pi_2(X_t), Q_2(X_t)) + g(t, \Pi_1(X_t), Q_1(X_t))\,dB(t)$$

for projections Π_i of discrete type and Q_i of continuous type for $i = 1, 2$ are approximated by Euler schemes (discretization in state but not in time) with order 0.5. The proof is based on Gronwall techniques alone. In Theorem 5.3 in Hu, Mohammed, and Yan [8] an Itô formula is presented for $\phi(t, \Pi(X_t))$ for discrete projections Π of the function segment X_t. An SDDE with point delay is approximated by Milstein schemes $X^{(h)}$ which converge strongly to the solution X with order 1 which is proven by means of this Itô formula.

If the coefficients f and g of the following SDDE

$$dX(t) = f(X_t)\,dt + g(X_t)\,dB(t)$$

satisfy a global Lipschitz condition, then the uniform estimation

$$E(\sup_{0\leq t\leq T}|X_t^{(h)}-X(t)|)\leq K\sqrt{h}$$

holds for Euler schemes as is pointed out in Mao [18].

It is also possible to firstly approximate the Brownian motion B by processes $B^{(h)}$ and secondly construct processes $X^{(h)}$ in terms of $B^{(h)}$. This direction was at first followed in Wong and Zakai [30]. They consider processes $X^{(h)}$ of the kind

$$dX^{(h)}(t)=m(X^{(h)}(t),t)\,dt+\sigma(X^{(h)}(t),t)\,dB^{(h)}(t),$$

where $B^{(h)}$ with bounded variation for each $h>0$ is an approximation of the Brownian motion B. Then one has the effect that the sequence $\{X^{(h)}:h>0\}$ converges a.s. to a process X uniformly on any compact interval, where the process X is the Stratonovich solution of

$$dX(t)=m(X(t),t)\,dt+\sigma(X(t),t)\circ\,dB(t),$$

or equivalently the Itô solution of

$$dX(t)=m(X(t),t)\,dt+\sigma(X(t),t)\,dB(t)+\frac{1}{2}\sigma(X(t),t)\frac{\partial}{\partial y}\sigma(X(t),t)\,dt.$$

The additional term

$$\frac{1}{2}\sigma(X(t),t)\frac{\partial}{\partial x}\sigma(X(t),t)\,dt$$

for X is explained by the fact that each $B^{(h)}$ has bounded variation, and $B(t)$ has quadratic variation t. How does the additional term for delay equations look like? An answer is given in Twardowska [29]. The processes $X^{(h)}$ given by

$$dX^{(h)}(t)=b(X_t^{(h)})\,dt+\sigma(X_t^{(h)})\,dB^{(h)}(t)$$

converge in mean square to the strong solution of the SDDE

$$dX(t)=b(X_t)\,dt+\sigma(X_t)\,dB(t)+\frac{1}{2}\tilde{D}\sigma(X_t)\sigma(X_t)\,dt.$$

The term $\tilde{D}\sigma$ requires explanation. The Fréchet derivative $D\sigma$ at the point g may be represented as an integral with respect to a measure μ_g

$$(D\sigma)(g)(\triangle)=\int_{-r}^{0}\triangle(v)\,d\mu_g(v).$$

Then $\tilde{D}\sigma$ is defined as $(\tilde{D}\sigma)(g):=\mu_g(\{0\})$. In the case of point delay with

$$\sigma(x)=\bar{\sigma}(x(u_0),\ldots,x(u_n)),\quad -r=u_n<\ldots<u_0=0,\quad x\in C[-r,0]$$

the Fréchet derivative and the quantity $\mu_g(\{0\})$ are given by

$$(D\sigma)(g)(\triangle) = \sum_{i=0}^{n} \frac{\partial \bar{\sigma}}{\partial x_i}(g(u_i))\triangle(u_i), \quad \mu_g(\{0\}) = \frac{\partial \bar{\sigma}}{\partial x_0}(g(0)).$$

For delay of the kind

$$\sigma(x) = h\left(\int_{-r}^{0} \psi(x(v))a(v)\,dv\right), \quad x \in C[-r,0]$$

those quantities are

$$(D\sigma)(g)(\triangle) = \int_{-r}^{0} h'\left(\int_{-r}^{0} \psi(g(v))a(v)\,dv\right)\psi'(g(v))a(v)\triangle(v)\,dv, \quad \mu_g(\{0\}) = 0.$$

This shows the following result of Twardowska [29]: The processes $X^{(h)}$ converge strongly to the Itô solution, if the diffusion coefficient does not depend on the actual state in the case of point delay, or if the diffusion coefficient has the form $\sigma(x) = h\left(\int_{-r}^{0} \psi(x(v))a(v)\,dv\right)$.

Chapter 3

Weak Limits of ARMA-Series

3.1 Introduction

In this chapter we shall deal with strictly stationary processes. The distribution of the initial condition is determined by the requirement of stationarity of the underlying process.

3.1.1 Definition. *Let $I \subset (-\infty, +\infty)$ be an index set. A stochastic process $(X_t)_{t\in I}$ is strictly stationary if*

$$(X_{t_1}, \dots, X_{t_n}) \stackrel{d}{=} (X_{t_1+s}, \dots, X_{t_n+s})$$

for all $0 \le t_1 < \dots < t_n, n \in \mathbb{N}, t_j \in I, t_j + s \in I$, where $\stackrel{d}{=}$ stands for equality in distribution.

There exist other notions of stationarity than strictly stationary. Since we regard only this type of stationarity in this section, we will omit the qualifier strictly in the sequel.

We shall now introduce an important class of stationary processes in discrete time.

3.1.2 Definition (ARMA(p,q))**.** *$(Y_n)_{n\in\mathbb{Z}}$ is an ARMA(p,q)-process, if it is stationary and if for every $n \in \mathbb{Z}$*

$$Y_n + b_1 Y_{n-1} + \dots + b_q Y_{n-p} = a_0\epsilon_n + a_1\epsilon_{n-1} + \dots + a_q\epsilon_{n-q}, \quad b_p \neq 0, \quad a_q \neq 0, \tag{3.1.1}$$

where $(\epsilon_n)_{n\in\mathbb{Z}}$ is a sequence of independent, $N(0,1)$-distributed random variables.

It is shown in Shiryaev [27] that, if the polynomial

$$P(z) := 1 + b_1 z + \dots + b_p z^p, \quad z \in \mathbb{C}$$

does not vanish in the closed unit circle, then there exists a unique stationary solution of (3.1.1). This solution has the spectral density

$$f(\lambda) = \frac{1}{2\pi}\frac{|Q(\mathrm{e}^{-\mathrm{i}\lambda})|^2}{|P(\mathrm{e}^{-\mathrm{i}\lambda})|^2}, \quad \lambda \in \mathbb{R}, \quad E(Y_n Y_0) = \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}\lambda n} f(\lambda)\, d\lambda, \quad n \in \mathbb{Z},$$

where $Q(z) := a_0 + a_1 z + \ldots + a_q z^q$ for $z \in \mathbb{C}$. We assume that $p^{(h)} := p/h \in \mathbb{N}_0$ and $q^{(h)} := q/h \in \mathbb{N}_0$ for some nonnegative real numbers p and q. In the sequel we shall write "for $h > 0$" meaning "for all $h > 0$ such that $p^{(h)} = p/h \in \mathbb{N}_0$ and $q^{(h)} = q/h \in \mathbb{N}_0$". For $h > 0$ define the following scheme

$$\begin{cases} Y^{(h)}_{(m+1)h} &= Y^{(h)}_{mh} + \sum_{j=0}^{p^{(h)}} a_j^{(h)} Y^{(h)}_{(m-j)h} h + \sum_{i=0}^{q^{(h)}} \sigma_i^{(h)} \sqrt{h}\epsilon_{m+1-i}, \quad m \in \mathbb{Z} \\ Y^{(h)}_t &= Y^{(h)}_{[\frac{t}{h}]h}, \quad t \in \mathbb{R} \end{cases} \tag{3.1.2}$$

The cases $p = 0$ and $q = 0$ are included. If p or q are greater than zero, then the numbers $p^{(h)}$ or $q^{(h)}$ tend to infinity as h tends to zero. The process $Y^{(h)}$ is right-continuous stochastic process on the whole real line. The series $(Y^{(h)}_{mh})_{m\in\mathbb{Z}}$ is an ARMA$(p^{(h)}+1, q^{(h)})$-process in discrete time. It is shown in Shiryaev [27] that it admits a stationary solution if the polynomial

$$P^{(h)}(z) := 1 - z - h\sum_{j=0}^{p^{(h)}} a_j^{(h)} z^{j+1} = 1 - z - zh\int_{-p}^{0} z^{-\frac{u}{h}}\, da^{(h)}(u), \quad z \in \mathbb{C}$$

does not vanish in the closed unit circle. Here $a^{(h)}$ denotes the discrete measure on $[-p, 0]$ obtained from the coefficients $a_j^{(h)}$ by

$$a^{(h)}(\{-jh\}) = a_j^{(h)}, \quad 0 \le j \le p^{(h)}.$$

It is shown in Reiß [23] that, if the sequence $a^{(h)}$ of discrete signed measures converges weakly to a signed measure a on $[-p, 0]$ and $a \in M^-[-p, 0]$, then $P^{(h)}(z)$ does not vanish in the closed unit circle for all sufficiently small h. A signed measure a belongs to $M^-[-p, 0]$ if

$$v_0(a) := \sup\{\mathrm{Re}(\lambda) : \chi_a(\lambda) = 0\} < 0, \quad \chi_a(\lambda) := \lambda - \int_{-p}^{0} \mathrm{e}^{\lambda u}\, da(u), \quad \lambda \in \mathbb{C}.$$

The spectral density is given by

$$f^{(h)}(\lambda) = \frac{1}{2\pi}\frac{|Q^{(h)}(\mathrm{e}^{-\mathrm{i}\lambda})|^2}{|P^{(h)}(\mathrm{e}^{-\mathrm{i}\lambda})|^2}, \quad \lambda \in \mathbb{R},$$

where the polynomial $Q^{(h)}$ is defined by

$$Q^{(h)}(z) := \sum_{i=0}^{q^{(h)}} \sigma_i^{(h)} \sqrt{h} z^i = \sqrt{h}\int_{-q}^{0} z^{-\frac{u}{h}}\, d\sigma^{(h)}(u), \quad z \in \mathbb{C}.$$

Here $\sigma^{(h)}$ denotes the discrete measure on $[-q, 0]$ obtained from the coefficients $\sigma_i^{(h)}$ by

$$\sigma^{(h)}(\{-ih\}) = \sigma_i^{(h)}, \quad 0 \le i \le q^{(h)}.$$

Note that no assumption on the measure $\sigma^{(h)}$ is needed for the existence of a stationary solution. We obtain for the covariance function of the series $(Y^{(h)}_{mh})_{m\in\mathbb{Z}}$ in (3.1.2)

$$\begin{aligned} E(Y^{(h)}_{mh}Y^{(h)}_0) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}\lambda m}\frac{|Q^{(h)}(\mathrm{e}^{-\mathrm{i}\lambda})|^2}{|P^{(h)}(\mathrm{e}^{-\mathrm{i}\lambda})|^2}\,d\lambda \\ &= \frac{1}{2\pi}\int_{-\pi/h}^{\pi/h} \mathrm{e}^{\mathrm{i}\lambda mh}\frac{|Q^{(h)}(\mathrm{e}^{-\mathrm{i}\lambda h})|^2}{|P^{(h)}(\mathrm{e}^{-\mathrm{i}\lambda h})|^2}h\,d\lambda \\ &= \frac{1}{2\pi}\int_{-\pi/h}^{\pi/h} \mathrm{e}^{\mathrm{i}\lambda mh}\frac{\left|\int_{-q}^{0}\mathrm{e}^{\mathrm{i}\lambda u}\,d\sigma^{(h)}(u)\right|^2}{|P^{(h)}(\mathrm{e}^{-\mathrm{i}\lambda h})|^2}h^2\,d\lambda \\ &= \frac{1}{2\pi}\int_{-\pi/h}^{\pi/h} \mathrm{e}^{\mathrm{i}\lambda mh}\frac{\left|\int_{-q}^{0}\mathrm{e}^{\mathrm{i}\lambda u}\,d\sigma^{(h)}(u)\right|^2}{|\chi^{(h)}(\mathrm{i}\lambda)|^2}\,d\lambda, \quad m\in\mathbb{Z}, \end{aligned}$$

where we used the notation

$$\chi^{(h)}(\mathrm{i}\lambda) := \frac{P^{(h)}(\mathrm{e}^{-\mathrm{i}\lambda h})}{h} = \frac{1-\mathrm{e}^{-\mathrm{i}\lambda h}}{h} - \mathrm{e}^{-\mathrm{i}\lambda h}\int_{-r}^{0}\mathrm{e}^{\mathrm{i}\lambda u}\,da^{(h)}(u), \quad \lambda\in\mathbb{R}.$$

Our aim is to establish a weak limit of the sequence of ARMA-processes $Y^{(h)}$ as h tends to zero.

3.2 Establishing the Limit

We assume that the discrete measures $a^{(h)}$ and $\sigma^{(h)}$ converge weakly: There are signed measures a on $[-p,0]$ and σ on $[-q,0]$ such that $a^{(h)} \Longrightarrow a$ and $\sigma^{(h)} \Longrightarrow \sigma$. In this case the following asymptotic behavior of the covariance function of the series $(Y^{(h)}_{mh})_{m\in\mathbb{Z}}$ is suggested

$$\begin{aligned} E(Y^{(h)}_{[\frac{t}{h}]h}Y^{(h)}_0) &= \frac{1}{2\pi}\int_{-\pi/h}^{\pi/h} \mathrm{e}^{\mathrm{i}\lambda[\frac{t}{h}]h}\frac{\left|\int_{-q}^{0}\mathrm{e}^{\mathrm{i}\lambda u}\,d\sigma^{(h)}(u)\right|^2}{|\chi^{(h)}(\mathrm{i}\lambda)|^2}\,d\lambda \\ &\underset{h\to 0}{\longrightarrow} \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{e}^{\mathrm{i}\lambda t}\frac{\left|\int_{-q}^{0}\mathrm{e}^{\mathrm{i}\lambda u}\,d\sigma(u)\right|^2}{|\chi_a(\mathrm{i}\lambda)|^2}\,d\lambda =: q_{a,\sigma}(t), \quad t\in\mathbb{R}. \end{aligned}$$

The following theorem reinforces the suggestion.

3.2.1 Theorem. *Assume that $a^{(h)} \Longrightarrow a$ such that $a\in M^-[-p,0]$ and*

$$\left|\mathrm{e}^{-\mathrm{i}\lambda h}\int_{-p}^{0}\mathrm{e}^{\mathrm{i}\lambda u}\,da^{(h)}(u) - \int_{-p}^{0}\mathrm{e}^{\mathrm{i}\lambda u}\,da(u)\right| \le K(1+\lambda^2)h \qquad (3.2.1)$$

for some constant K and all sufficiently small h. Furthermore assume that $\sigma^{(h)} \Longrightarrow \sigma$ such that

$$\left| \int_{-q}^{0} \mathrm{e}^{\mathrm{i}\lambda u}\, d\sigma^{(h)}(u) - \int_{-q}^{0} \mathrm{e}^{\mathrm{i}\lambda u}\, d\sigma(u) \right| \leq K(1+|\lambda|)h \tag{3.2.2}$$

for some constant K and all sufficiently small h. Then it holds

$$\sup_{m\in\mathbb{Z}} |E(Y_{mh}^{(h)} Y_0^{(h)}) - q_{a,\sigma}(mh)| \xrightarrow[h\to 0]{} 0.$$

Proof. First we shall get rid of the uniformity in m and of the tails

$$\sup_{m\in\mathbb{Z}} |E(Y_{mh}^{(h)} Y_0^{(h)}) - q_{a,\sigma}(mh)| \leq \int_{|\lambda|\geq \pi/h} \frac{\left|\int_{-q}^{0} \mathrm{e}^{\mathrm{i}\lambda u}\, d\sigma(u)\right|^2}{|\chi_a(\mathrm{i}\lambda)|^2}\, d\lambda$$
$$+\frac{1}{2\pi}\int_{-\pi/h}^{\pi/h} \left| \frac{\left|\int_{-q}^{0} \mathrm{e}^{\mathrm{i}\lambda u}\, d\sigma^{(h)}(u)\right|^2}{|\chi^{(h)}(\mathrm{i}\lambda)|^2} - \frac{\left|\int_{-q}^{0} \mathrm{e}^{\mathrm{i}\lambda u}\, d\sigma(u)\right|^2}{|\chi_a(\mathrm{i}\lambda)|^2} \right| d\lambda.$$

The first integral is lower or equal than

$$\int_{|\lambda|\geq \pi/h} \frac{\|\sigma\|_{\mathrm{TV}}^2}{|\chi_a(\mathrm{i}\lambda)|^2}\, d\lambda,$$

which tends to zero as h to zero since for each $a \in M^-[-p,0]$

$$\frac{1}{|\chi_a(\mathrm{i}\lambda)|^2} \sim \frac{1}{1+\lambda^2}. \tag{3.2.3}$$

The notation $A(\lambda) \sim B(\lambda)$ means that there are constants $C_1 \leq C_2$ such that $0 < C_1 \leq \frac{A(\lambda)}{B(\lambda)} \leq C_2 < \infty$ for all $\lambda \in \mathbb{R}$. Using the property $a^{(h)} \Longrightarrow a$, one can easily show that also

$$\frac{1}{|\chi^{(h)}(\mathrm{i}\lambda)|^2} \sim \frac{1}{1+\lambda^2} \tag{3.2.4}$$

uniformly for all sufficiently small h. For simplification of writing we introduce the following notations for $\lambda \in \mathbb{R}$

$$w^{(h)}(\lambda) := \int_{-q}^{0} \mathrm{e}^{\mathrm{i}\lambda u}\, d\sigma^{(h)}(u), \qquad w(\lambda) := \int_{-q}^{0} \mathrm{e}^{\mathrm{i}\lambda u}\, d\sigma(u)$$
$$z^{(h)}(\lambda) := \chi^{(h)}(\mathrm{i}\lambda), \qquad z(\lambda) := \chi_a(\mathrm{i}\lambda).$$

For each $\lambda \in \mathbb{R}$ we have the pointwise estimation

$$\left| \frac{|w^{(h)}|^2}{|z^{(h)}|^2} - \frac{|w|^2}{|z|^2} \right| \leq |w^{(h)}|^2 \frac{(|z|+|z^{(h)}|)|z-z^{(h)}|}{|z^{(h)}|^2|z|^2} + \frac{(|w|+|w^{(h)}|)|w-w^{(h)}|}{|z|^2}.$$

Now we estimate every single term starting with w and $w^{(h)}$

$$|w| + |w^{(h)}| \leq \|\sigma\|_{\mathrm{TV}} + \sup_{h>0} \|\sigma^{(h)}\|_{\mathrm{TV}} < \infty.$$

By assumption (3.2.1) we have that

$$\begin{aligned} |z - z^{(h)}| &\leq \left| \frac{1 - \mathrm{e}^{-\mathrm{i}\lambda h}}{h} - \mathrm{i}\lambda \right| + K(1+\lambda^2)h \\ &\leq \lambda^2 h \sum_{j=2}^{\infty} \frac{|\lambda h|^{j-2}}{j!} + K(1+\lambda^2)h \\ &\leq C_1(1+\lambda^2)h, \quad |\lambda| \leq \frac{\pi}{h}, \quad C_1 = \mathrm{const.} \end{aligned}$$

Using the behavior of χ_a and $\chi^{(h)}$ in (3.2.3) and (3.2.4) we obtain

$$\left(|w^{(h)}|^2 \frac{(|z| + |z^{(h)}|)|z - z^{(h)}|}{|z^{(h)}|^2 |z|^2} \right)(\lambda) \leq C_2(1+\lambda^2)^{-1/2}h, \quad C_2 = \mathrm{const.}$$

As for the other summand it suffices to combine assumption (3.2.2)

$$|w - w^{(h)}|(\lambda) \leq K(1+|\lambda|)h \leq C_3(1+\lambda^2)^{1/2}h, \quad C_3 = \mathrm{const}$$

with the behavior of χ_a in (3.2.3) to obtain

$$\left(\frac{(|w| + |w^{(h)}|)|w - w^{(h)}|}{|z|^2} \right)(\lambda) \leq C_4(1+\lambda^2)^{-1/2}h, \quad C_4 = \mathrm{const.}$$

Now it holds that

$$h \int_{|\lambda| \leq \frac{\pi}{h}} (1+\lambda^2)^{-1/2}\, d\lambda = 2h \log(\pi/h + (1+\pi^2/h^2)^{-1/2}) \underset{h\to 0}{\longrightarrow} 0,$$

which completes the proof of the theorem. □

3.2.2 Remark. *If for a given measure a on $[-p, 0]$ one chooses approximating measures $a^{(h)}$ by*

$$a^{(h)}_{p^{(h)}} = a^{(h)}(\{-p\}) := a(\{-p\}), \quad a^{(h)}_j = a^{(h)}(\{-jh\}) := a(-(j+1)h, -jh],$$

then

$$a^{(h)} \Longrightarrow a$$

(respectively for σ), and one can show by partial integration that the technical assumptions of Theorem 3.2.1 are satisfied both for $a^{(h)}$ and $\sigma^{(h)}$.

3.2.3 Remark. *Since $q_{a,\sigma}(t)$ is continuous, we have shown by Theorem 3.2.1 that for all $T > 0$ it holds that*

$$\sup_{0 \le s_0 \le s_1 \le T} |E(Y_{s_1}^{(h)} Y_{s_0}^{(h)}) - q_{a,\sigma}(s_1 - s_0)| \xrightarrow[h \to 0]{} 0.$$

For $h > 0$ we defined in (3.1.2) a stationary sequence $(Y_{mh}^{(h)})_{m \in \mathbb{Z}}$ with covariance function $E(Y_{[\frac{t}{h}]h}^{(h)} Y_0^{(h)})$ depending on h. So far we established that those covariance functions converge as h to zero uniformly to a function $q_{a,\sigma}(t)$. Now one can ask the following question. Does there exist a stationary process $(Y(t))_{t \in \mathbb{R}}$ on the whole real line such that

$$E(Y(t)Y(0)) = q_{a,\sigma}(t), \quad t \in \mathbb{R}?$$

The answer to this question is positive. Set

$$Y_{a,\sigma}(t) = Y(t) := \int_{-q}^{0} \left(\int_{-\infty}^{t} x_a(t-s)\, dB(s+u) \right) d\sigma(u), \quad t \in \mathbb{R}. \tag{3.2.5}$$

Here B is a Brownian motion on the whole real line, and x_a denotes the fundamental solution for the measure a which is defined by

$$\begin{cases} x_a(s) &= 0, \quad -p \le s < 0 \\ x_a(0) &= 1 \\ x_a(t) &= x_a(0) + \int_0^t \int_{-p}^0 x_a(s+u)\, da(u)\, ds, \quad t \ge 0. \end{cases}$$

By stochastic Fubini we obtain also the following representation for $Y_{a,\sigma}$

$$Y_{a,\sigma}(t) = \int_{-\infty}^{t} \left(\int_{-q}^{0} x_a(t+u-s)\, d\sigma(u) \right) dB(s), \quad t \in \mathbb{R}.$$

The function

$$f_{a,\sigma}(t) = f(t) := \int_{-q}^{0} x_a(t+u)\, d\sigma(u), \quad t \in \mathbb{R}$$

is called kernel function. Then the variable $Y(t) = \int_{-\infty}^{t} f(t-s)\, dB(s)$ is a moving average over differentials of a Brownian motion weighted by the kernel function. Since the kernel function vanishes for negative values one may also write $Y(t) = \int_{-\infty}^{+\infty} f(t-s)\, dB(s)$. An interesting special case is if σ is the Dirac-measure $\delta_{\{0\}}$ in zero. Then we obtain

$$Y_{a,\delta_{\{0\}}}(t) =: X(t) = \int_{-\infty}^{t} x_a(t-s)\, dB(s), \quad t \in \mathbb{R}.$$

Then X is the well-known Ornstein-Uhlenbeck process. This is a continuous Gaussian process and the stationary solution of

$$dX(t) = \left(\int_{-p}^{0} X(t+u)\, da(u) \right) dt + dB(t).$$

By the substitution $s' = s + u$ for fixed u we obtain the following representation of Y in terms of X

$$\begin{aligned} Y(t) &= \int_{-q}^{0} \left(\int_{-\infty}^{t} x_a(t-s)\, dB(s+u) \right) d\sigma(u) \\ &= \int_{-q}^{0} \left(\int_{-\infty}^{t+u} x_a(t+u-s')\, dB(s') \right) d\sigma(u) = \int_{-q}^{0} X(t+u)\, d\sigma(u), \quad t \in \mathbb{R}. \end{aligned}$$

We see that Y is a linear mixture of an Ornstein-Uhlenbeck process X with mixing signed measure σ. Hence Y is also a continuous Gaussian process. Furthermore, it is strictly stationary since by stationarity of X it follows that

$$\begin{aligned} E(Y(t)Y(w)) &= \int_{-q}^{0} \int_{-q}^{0} E(X(t+u)X(w+u'))\, d\sigma(u)\, d\sigma(u') \\ &= \int_{-q}^{0} \int_{-q}^{0} E(X(t-w+u)X(u'))\, d\sigma(u)\, d\sigma(u') \\ &= E(Y(t-w)Y(0)), \quad t, w \in \mathbb{R}. \end{aligned}$$

Our aim is now to show that

$$E(Y(t)Y(0)) = q_{a,\sigma}(t), \quad t \in \mathbb{R}.$$

First we shall establish this relation for the process X. A simple calculation shows that

$$E(X(t)X(0)) = \int_{0}^{+\infty} x_a(s)x_a(s+t)\, ds = \int_{0}^{+\infty} x_a(s)x_a(s+|t|)\, ds, \quad t \in \mathbb{R}.$$

Now we need the Laplace transform of the fundamental solution x_a. It has the form

$$\int_{0}^{+\infty} \mathrm{e}^{-\lambda t} x_a(t)\, dt = \frac{1}{\chi_a(\lambda)}, \quad \mathrm{Re}(\lambda) > v_0(a),$$

as can be found in Reiß [23]. By repeated use of Fubini we obtain

$$\begin{aligned} \frac{1}{|\chi_a(\mathrm{i}\xi)|^2} &= \left(\int_{0}^{+\infty} \mathrm{e}^{-\mathrm{i}\xi t} x_a(t)\, dt \right) \left(\int_{0}^{+\infty} \mathrm{e}^{\mathrm{i}\xi t} x_a(t)\, dt \right) \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} 1_{\{s \geq 0\}} 1_{\{t \geq 0\}} x_a(s) x_a(t) \mathrm{e}^{\mathrm{i}\xi s} \mathrm{e}^{-\mathrm{i}\xi t}\, dt\, ds \\ &= \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} x_a(s) 1_{\{s \geq 0\}} x_a(s+t) 1_{\{s+t \geq 0\}}\, ds \right) \mathrm{e}^{-\mathrm{i}\xi t}\, dt \\ &= \int_{-\infty}^{+\infty} \left(\int_{0}^{+\infty} x_a(s) x_a(s+t)\, ds \right) \mathrm{e}^{-\mathrm{i}\xi t}\, dt \\ &= \int_{-\infty}^{+\infty} E(X(t)X(0)) \mathrm{e}^{-\mathrm{i}\xi t}\, dt, \quad \xi \in \mathbb{R}, \end{aligned}$$

which by transforming back implies

$$E(X(t)X(0)) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{e}^{\mathrm{i}\lambda t}\frac{1}{|\chi_a(\mathrm{i}\lambda)|^2}\, d\lambda, \quad t \in \mathbb{R}.$$

Therefore we see by repeated use of Fubini that

$$\begin{aligned}
E(Y(t)Y(0)) &= \int_{-q}^{0}\int_{-q}^{0} E(X(t+u-u')X(0))\, d\sigma(u)\, d\sigma(u') \\
&= \int_{-q}^{0}\int_{-q}^{0} \frac{1}{2\pi}\left(\int_{-\infty}^{+\infty} \mathrm{e}^{\mathrm{i}\lambda(t+u-u')}\frac{1}{|\chi_a(\mathrm{i}\xi)|^2}\, d\lambda\right) d\sigma(u)\, d\sigma(u') \\
&= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \frac{\mathrm{e}^{\mathrm{i}\lambda t}}{|\chi_a(\mathrm{i}\lambda)|^2}\left(\int_{-q}^{0}\int_{-q}^{0} \mathrm{e}^{\mathrm{i}\lambda(u-u')}\, d\sigma(u)\, d\sigma(u')\right) d\lambda \\
&= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{e}^{\mathrm{i}\lambda t}\frac{\left|\int_{-q}^{0} \mathrm{e}^{\mathrm{i}\lambda u}\, d\sigma(u)\right|^2}{|\chi_a(\mathrm{i}\lambda)|^2}\, d\lambda = q_{a,\sigma}(t), \quad t \in \mathbb{R}.
\end{aligned}$$

In view of Remark 3.2.3 we have obtained so far the following: The finite-dimensional distributions of $Y^{(h)}$ in (3.1.2) converge to the finite-dimensional distributions of Y, where Y is given by (3.2.5). To give a proof for tightness of the sequence $\{Y^{(h)} : h > 0\}$ we need some more information on the limit Y. We start with the fact that Y solves a certain stochastic equation.

3.2.4 Theorem. *Let Y be the process defined by*

$$Y(t) := \int_{-q}^{0}\left(\int_{-\infty}^{t} x_a(t-s)\, dB(s+u)\right) d\sigma(u), \quad t \in \mathbb{R}.$$

Then Y is the stationary solution of the following stochastic equation

$$dY(t) = \left(\int_{-p}^{0} Y(t+u)\, da(u)\right) dt + \int_{-q}^{0} dB(t+u)\, d\sigma(u), \quad t \geq 0 \tag{3.2.6}$$

which means in integrated form

$$Y(t) = Y(0) + \int_{0}^{t}\int_{-p}^{0} Y(s+u)\, da(u)\, ds + \int_{-q}^{0} [B(t+u) - B(u)]\, d\sigma(u), \quad t \geq 0.$$

Proof. We already established that Y is stationary and continuous. It is also clear that Y is adapted to the filtration of the Brownian motion B. It remains to show that Y solves (3.2.6). Here we use the property that Y is a mixture of an Ornstein-Uhlenbeck process X

$$Y(t) = \int_{-q}^{0} X(t+u)\, d\sigma(u), \quad t \in \mathbb{R}.$$

The process X satisfies for all real time points $t_1 \leq t_2$

$$X(t_2) = X(t_1) + \int_{t_1}^{t_2} \int_{-p}^{0} X(s+w)\, da(w)\, ds + B(t_2) - B(t_1).$$

By inserting we see that

$$\begin{aligned}
Y(t) - Y(0) &= \int_{-q}^{0} [X(t+u) - X(u)]\, d\sigma(u) \\
&= \int_{-q}^{0} \big[B(t+u) - B(u) + \int_{u}^{t+u} \int_{-p}^{0} X(s+w)\, da(w)\, ds\big]\, d\sigma(u) \\
&= \int_{-q}^{0} B(t+u) - B(u)\, d\sigma(u) \\
&+ \int_{0}^{t} \int_{-p}^{0} \int_{-q}^{0} X(s+w+u)\, d\sigma(u)\, da(w)\, ds \\
&= \int_{-q}^{0} B(t+u) - B(u)\, d\sigma(u) + \int_{0}^{t} \int_{-p}^{0} Y(s+w)\, da(w)\, ds, \quad t \geq 0,
\end{aligned}$$

which is our desired result. □

Before we discuss the structure of the stochastic equation (3.2.6) we proceed with a moment estimation for stationary solutions of this equation.

3.2.5 Lemma. *Let Y be a stationary solution of*

$$dY(t) = \left(\int_{-p}^{0} Y(t+u)\, da(u) \right) dt + \int_{-q}^{0} dB(t+u)\, d\sigma(u), \quad t \geq 0$$

with $E(|Y(0)|^2) < \infty$. Then for each $T > 0$ there exist a constant K such that

$$E(|Y(t) - Y(0)|^2) \leq Kt, \quad 0 \leq t \leq T.$$

Proof. We have by definition that

$$Y(t) = Y(0) + \int_{0}^{t} \int_{-p}^{0} Y(s+u)\, da(u)\, ds + \int_{-q}^{0} [B(t+u) - B(u)]\, d\sigma(u), \quad t \geq 0,$$

and hence

$$|Y(t) - Y(0)|^2 \leq 2\Big(\int_{0}^{t} \int_{-p}^{0} Y(s+u)\, da(u)\, ds\Big)^2 + 2\Big(\int_{-q}^{0} [B(t+u) - B(u)]\, d\sigma(u)\Big)^2.$$

Firstly, we have by Fubini and Cauchy-Schwarz that

$$
\begin{aligned}
&E\Big(\Big(\int_{-q}^{0}[B(t+u)-B(u)]\,d\sigma(u)\Big)^2\Big)\\
&=\int_{-q}^{0}\int_{-q}^{0}E([B(t+u)-B(u)][B(t+u')-B(u')])\,d\sigma(u)\,d\sigma(u')\leq\\
&\int_{-q}^{0}\int_{-q}^{0}\sqrt{E(|B(t+u)-B(u)|^2)E(|B(t+u')-B(u')|^2)}\,d|\sigma|(u)\,d|\sigma|(u')\\
&=\int_{-q}^{0}\int_{-q}^{0}t^{1/2}t^{1/2}\,d|\sigma|(u)\,d|\sigma|(u')=t\|\sigma\|_{\mathrm{TV}}^2,\quad t\geq 0.
\end{aligned}
$$

Secondly, we have the pointwise estimation

$$
\begin{aligned}
\Big(\int_0^t\int_{-p}^{0}Y(s+u)\,da(u)\,ds\Big)^2 &\leq t\int_0^t\Big(\int_{-p}^{0}Y(s+u)\,da(u)\Big)^2\,ds\\
&\leq t\|a\|_{\mathrm{TV}}\int_0^t\int_{-p}^{0}|Y(s+u)|^2\,d|a|(u)\,ds,\quad t\geq 0,
\end{aligned}
$$

and hence

$$
\begin{aligned}
E\Big(\int_0^t\int_{-p}^{0}Y(s+u)\,da(u)\,ds\Big)^2 &\leq t\|a\|_{\mathrm{TV}}\int_0^t\int_{-p}^{0}E(|Y(s+u)|^2)\,d|a|(u)\,ds\\
&= t\|a\|_{\mathrm{TV}}\int_0^t\int_{-p}^{0}E(|Y(0)|^2)\,d|a|(u)\,ds\\
&\leq t\|a\|_{\mathrm{TV}}^2E(|Y(0)|^2)T,\quad t\geq 0.
\end{aligned}
$$

Therefore the lemma holds for the constant

$$K=2\|\sigma\|_{\mathrm{TV}}^2+2\|a\|_{\mathrm{TV}}^2E(|Y(0)|^2)T.$$

□

Now we are able to prove the following theorem.

3.2.6 Theorem. *Let for $h>0$ the process $Y^{(h)}$ be defined by*

$$
\begin{cases}
Y_{(m+1)h}^{(h)} = Y_{mh}^{(h)}+\sum_{j=0}^{p^{(h)}}a_j^{(h)}Y_{(m-j)h}^{(h)}h+\sum_{i=0}^{q^{(h)}}\sigma_i^{(h)}\sqrt{h}\epsilon_{m-i}, & m\in\mathbb{Z}\\
Y_{mh}^{(h)} = Y_{[\frac{t}{h}]h}^{(h)}, & t\in\mathbb{R}.
\end{cases}
$$

Let the assumptions of Theorem 3.2.1 be satisfied. Then the sequence $\{Y^{(h)}:h>0\}$ converges weakly to Y, where Y is given by

$$Y(t)=\int_{-q}^{0}\Big(\int_{-\infty}^{t}x_a(t-s)\,dB(s+u)\Big)\,d\sigma(u),\quad t\in\mathbb{R}.$$

Proof. It remains to establish the tightness of the sequence $\{Y^{(h)} : h > 0\}$. By Theorem 15.6 in Billingsley [2] tightness will follow, if for all $T > 0$

$$e^{(h)}_{t_0,t_1,t_2} := \overline{\lim_{h\to 0}} E(|Y^{(h)}_{t_1} - Y^{(h)}_{t_0}|^2 |Y^{(h)}_{t_2} - Y^{(h)}_{t_1}|^2) \leq C(t_2 - t_0)^2$$

for all $0 \leq t_0 \leq t_1 \leq t_2 \leq T$ and for a constant C. Since $Y^{(h)}$ and Y are stationary, and Y has covariance function $q_{a,\sigma}$, it follows from Remark 3.2.3 that

$$\sup_{0\leq s_0\leq s_1\leq T} |E(|Y^{(h)}_{s_1} - Y^{(h)}_{s_0}|^2) - E(|Y(s_1) - Y(s_0)|^2)| \xrightarrow[h\to 0]{} 0.$$

For Gaussian systems (A, B) with $E(A) = E(B) = 0$ it holds that

$$E(A^2B^2) = E(A^2)E(B^2) + 2(E(AB))^2 \leq 3E(A^2)E(B^2).$$

Now we can complete the proof in the following manner

$$\begin{aligned} e^{(h)}_{t_0,t_1,t_2} &\leq 3\overline{\lim_{h\to 0}} E(|Y^{(h)}_{t_1} - Y^{(h)}_{t_0}|^2)E(|Y^{(h)}_{t_2} - Y^{(h)}_{t_1}|^2) \\ &= 3E(|Y(t_1) - Y(t_0)|^2)E(|Y(t_2) - Y(t_1)|^2) \\ &= 3E(|Y(t_1 - t_0) - Y(0)|^2)E(|Y(t_2 - t_1) - Y(0)|^2) \\ &\leq 3K(t_1 - t_0)(t_2 - t_1) \leq 3K(t_2 - t_0)^2. \end{aligned}$$

For the last line we used Lemma 3.2.5. The theorem has been shown. □

3.3 Discussion of the Limit

For theoretical interest we will investigate the system

$$\begin{cases} Y_0 &= \rho \\ dY(t) &= \left(\int_{-p}^0 Y(t+u)\,da(u)\right)\,dt + \int_{-q}^0 dB(t+u)\,d\sigma(u), \quad t \geq 0 \end{cases} \tag{3.3.1}$$

in detail. Here ρ is an arbitrary random initial condition on $[-p, 0]$ and B a Brownian motion starting at time $-q$. The case of the stationary solution for system (3.3.1) is included. Especially we are interested in the structure of the driving force. Its differentiated form is

$$\int_{-q}^0 dB(t+u)\,d\sigma(u). \tag{3.3.2}$$

This stochastic differential cannot be written in the form

$$g(Y_t, t, \omega)\,dB(t),$$

even for a coefficient g explicitly depending on ω. The differential in (3.3.2) is not only a differential of the Brownian motion at time point t, but a mixture of differentials for time points in the interval $[t-q,t]$. Hence existence and uniqueness for equations driven by (3.3.2) are not covered by existence theorems with random coefficients in Mohammed [20]. If we formally write down the integrated form and formally use Fubini, then we obtain

$$\begin{aligned} Z_1(t) \;&:=\; \int_0^t \int_{-q}^0 dB(s+u)\,d\sigma(u) = \int_{-q}^0 \int_0^t dB(s+u)\,d\sigma(u) \\ &=\; \int_{-q}^0 \int_0^t dB(s+u)\,d\sigma(u) = \int_{-q}^0 [B(t+u)-B(0)])\,d\sigma(u). \end{aligned}$$

In the preceding section we showed that Z_1 is the driving force of a weak limit of ARMA-processes. We see that the random variable

$$I(0) := \int_{-q}^0 -B(0)\,d\sigma(u)$$

is $\mathcal{F}_0$-measurable, if $\mathcal{F}$ denotes the natural filtration of the Brownian motion B. It is common to collect $\mathcal{F}_0$-measurable variables at the initial conditions. Then the stochastic equation takes the form

$$\left\{ \begin{array}{rcl} Y_0 &=& \rho \\ Y(t) &=& \rho(0) + I(0) + \int_0^t \int_{-p}^0 Y(s+u)\,da(u)\,dt + Z(t), \quad t \geq 0 \end{array} \right. \tag{3.3.3}$$

with

$$Z(t) := \int_{-q}^0 B(t)\,d\sigma(u), \quad t \geq 0.$$

One can interpret equation (3.3.3) such that for each ω one solves a deterministic delay differential equation for $Z(\omega)$. We would like to study the process Z precisely. It is easily seen that it is a centered Gaussian process. Its covariance is given by

$$\begin{aligned} E(Z(t)Z(s)) \;&=\; \int_{-q}^0 \int_{-q}^0 E(B(t+u)B(s+u'))\,d\sigma(u)\,d\sigma(u') \\ &=\; \int_{-q}^0 \int_{-q}^0 (t+u) \wedge (s+u')\,d\sigma(u)\,d\sigma(u'), \quad t,s \geq 0. \end{aligned}$$

A next important characteristic is the quadratic variation. Therefore we need a lemma.

3.3.1 Lemma. *It holds for the quadratic covariation of the Brownian motion with itself that a.s.*

$$[B(\cdot+u), B(\cdot+u')](t) = t \cdot 1_{\{u=u'\}}, \quad -q \leq u, u' \leq 0, \quad t \geq 0.$$

Proof. For $u = u'$ it is known that a.s.

$$\sum_{0 \le t_i^n \le t} [B(t_{i+1}^n + u) - B(t_i^n + u)]^2 \xrightarrow[n\to\infty]{} t, \quad t \ge 0.$$

Let $u \neq u'$. We consider the dyadic partition of $[0, t]$. For

$$t_{i+1}^n - t_i^n \le |u - u'|$$

it holds that

$$E\left(\sum_{0 \le t_i^n \le t} [B(t_{i+1}^n + u) - B(t_i^n + u)][B(t_{i+1}^n + u') - B(t_i^n + u')]\right)^2$$
$$= \sum_{0 \le t_i^n \le t} E[B(t_{i+1}^n + u) - B(t_i^n + u)]^2 E[B(t_{i+1}^n + u') - B(t_i^n + u')]^2$$
$$= \sum_{0 \le t_i^n \le t} (t_{i+1}^n - t_i^n)^2 \le t/2^n \xrightarrow[n\to\infty]{} 0, \quad t \ge 0.$$

Since

$$\sum_{n=1}^{\infty} t/2^n < \infty,$$

the convergence is a.s. The lemma has been shown. □

Now we are able to calculate the quadratic variation of Z.

3.3.2 Theorem. *For*

$$Z(t) = \int_{-q}^{0} B(t+u)\, d\sigma(u), \quad t \ge 0$$

it holds that a.s.

$$[Z](t) = t \cdot \int_{-q}^{0}\int_{-q}^{0} 1_{\{u=u'\}}\, d\sigma(u)\, d\sigma(u'), \quad t \ge 0.$$

Proof. We have to establish the limit of the sum

$$\sum_{0 \le t_i^n \le t} [Z(t_{i+1}^n) - Z(t_i^n)]^2, \quad t \ge 0.$$

By Fubini and linearity of the integral we see that

$$\sum_{0 \le t_i^n \le t} [Z(t_{i+1}^n) - Z(t_i^n)]^2 = \int_{-q}^{0}\int_{-q}^{0} g^n(u, u')\, d\sigma(u)\, d\sigma(u'), \quad t \ge 0$$

with

$$g^n(u,u') := \sum_{0\le t_i^n\le t} [B(t_{i+1}^n+u)-B(t_i^n+u)][B(t_{i+1}^n+u')-B(t_i^n+u')], \quad -q\le u,u'\le 0.$$

According to Lemma 3.3.1 it holds that a.s.

$$g^n(u,u') \xrightarrow[n\to\infty]{} t\cdot 1_{\{u=u'\}}, \quad -q\le u,u'\le 0.$$

Since

$$|a||b| \le \frac{1}{2}(a^2+b^2),$$

and hence a.s.

$$\begin{aligned} |g^n(u,u')|(\omega) &\le \frac{1}{2}\sum_{0\le t_i^n\le t}[B(t_{i+1}^n+u,\omega)-B(t_i^n+u,\omega)]^2 \\ &+ \frac{1}{2}\sum_{0\le t_i^n\le t}[B(t_{i+1}^n+u',\omega)-B(t_i^n+u',\omega)]^2 \le K(\omega)<\infty, \end{aligned}$$

we obtain a.s. by dominated convergence

$$\int_{-q}^0\int_{-q}^0 g^n(u,u')\,d\sigma(u)\,d\sigma(u') \xrightarrow[n\to\infty]{} \int_{-q}^0\int_{-q}^0 t\cdot 1_{\{u=u'\}}\,d\sigma(u)\,d\sigma(u'), \quad t\ge 0.$$

□

3.3.3 Example. *If the measure σ is absolutely continuous with respect to the Lebesgue-measure, then the quadratic variation of Z vanishes. For discrete measures*

$$\sigma = \sum_{i=0}^k \sigma_i\delta_{\{u_i\}}, \quad \sigma_i\in\mathbb{R}/\{0\}, \quad -q\le u_n<\ldots<u_1<u_0\le 0$$

the quadratic variation is strictly positive

$$[Z](t) = \left(\sum_{i=0}^k \sigma_i^2\right)\cdot t, \quad t\ge 0.$$

Now we assume that σ is absolutely continuous with respect to the Lebesgue-measure $d\sigma(u)=f(u)\,du$. Then

$$Z(t) = \int_{-q}^0 B(t+u)\,d\sigma(u) = \int_{-q}^0 B(t+u)f(u)\,du, \quad t\ge 0$$

is a centered Gaussian process with vanishing quadratic variation. This resembles so far a fractional Brownian motion. However we will see in the next theorem that Z is differentiable in time, hence has locally bounded total variation. Therefore Z cannot be a fractional Brownian motion.

3.3.4 Theorem. *For $f \in C^1$ the process*

$$Z(t) = \int_{-q}^{0} B(t+u) f(u)\, du, \quad t \geq 0$$

is differentiable, and its derivative is given by

$$Z'(t) = \int_{-q}^{0} f(u)\, dB(t+u), \quad t \geq 0.$$

Proof. For $h > 0$ we calculate explicitly

$$\begin{aligned} \frac{Z(t+h) - Z(t)}{h} &= \frac{1}{h}\left(\int_{h-q}^{h} B(t+u) f(u-h)\, du - \int_{-q}^{0} B(t+u) f(u)\, du\right) \\ &= \frac{1}{h}\left(\int_{h-q}^{h} B(t+u) f(u)\, du - \int_{-q}^{0} B(t+u) f(u)\, du\right) \\ &+ \frac{1}{h}\int_{h-q}^{h} B(t+u)[f(u-h) - f(u)]\, du, \quad t \geq 0. \end{aligned}$$

The summand in the second line equals

$$\frac{1}{h}\left(\int_{0}^{h} B(t+u) f(u)\, du - \int_{-q}^{h-q} B(t+u) f(u)\, du\right)$$

and tends to

$$B(t)f(0) - B(t-q)f(-q)$$

as h tends to zero. Since $f \in C^1$, the summand in the last line converges to

$$-\int_{-q}^{0} B(t+u) f'(u)\, du.$$

Hence we obtain by partial integration

$$\begin{aligned} Z'(t) &= B(t)f(0) - B(t-q)f(-q) - \int_{-q}^{0} B(t+u) f'(u)\, du \\ &= \int_{-q}^{0} f(u)\, dB(t+u), \quad t \geq 0 \end{aligned}$$

as required. □

3.3.5 Remark. *Since Z has differentiable paths, it is a semimartingale with vanishing martingale part.*

We shall now turn to the case that σ is a discrete measure on $[-q,0]$

$$Z(t)=\sum_{i=0}^{k}\sigma_i B(t+u_i),\quad \sigma_i\in\mathbb{R}/\{0\},\quad -q\le u_n<\ldots<u_1<u_0\le 0.$$

If $k=0$, then Z/σ_0 is a Brownian motion martingale, and Z is a semimartingale. Its paths are not differentiable. The question is if Z still is a semimartingale for $k\in\mathbb{N}$ with respect to the natural filtration of Z. Consider $A=Z-M$ for an arbitrary $\mathcal{F}^Z$-martingale M. Then A has the form

$$A(t)=\sum_{i=1}^{k}\sigma_i B(t+u_i)+\sigma_0 B(t)-M(t).$$

Since the process $\hat{A}(t):=\sum_{i=1}^{k}\sigma_i B(t+u_i)$ neither has bounded variation nor is an $\mathcal{F}^Z$-martingale, we deduce from this representation that A has unbounded variation. This means, Z is no semimartingale for $k\in\mathbb{N}$. Neither is the solution process Y driven by Z in this case.

3.3.6 Remark. *We see that the class of semimartingales is not closed under weak convergence. In Theorem 3.2.6 we established that the ARMA-processes $Y^{(h)}$, which are semimartingales since they are piecewise constant, converge weakly to the process Y which is not necessarily a semimartingale.*

We return to the stochastic equation with arbitrary random initial conditions

$$\left\{\begin{array}{rcl}(Y_\rho)_0&=&\rho\\ dY_\rho(t)&=&\left(\int_{-p}^{0}Y_\rho(t+u)\,da(u)\right)dt+\int_{-q}^{0}dB(t+u)\,d\sigma(u),\quad t\ge 0.\end{array}\right.\tag{3.3.4}$$

Next we shall investigate the following question. If X_ρ is the solution of

$$\left\{\begin{array}{rcl}(X_\rho)_0&=&\rho\\ dX_\rho(t)&=&\left(\int_{-p}^{0}X_\rho(t+u)\,da(u)\right)dt+\,dB(t),\quad t\ge 0,\end{array}\right.$$

may then Y_ρ be expressed in terms of X_ρ as in the stationary case? We need to introduce one more notation. Let x_ρ be the solution of the corresponding homogeneous system

$$\left\{\begin{array}{rcl}(x_\rho)_0&=&\rho\\ dx_\rho(t)&=&\left(\int_{-p}^{0}x_\rho(t+u)\,da(u)\right)dt,\quad t\ge 0.\end{array}\right.\tag{3.3.5}$$

The next theorem shows that strong uniqueness and strong existence hold for the system (3.3.4).

3.3.7 Theorem. *The system (3.3.4) has one unique strong solution. It is given by*

$$Y_\rho(t)=x_\rho(t)+\int_{-q}^{0}\int_{0}^{t}x_a(t-s)\,dB(s+u)\,d\sigma(u),\quad t\ge 0,\quad (Y_\rho)_0=\rho.$$

Proof. Let Y^1 and Y^2 be two solutions of the system (3.3.4). Then it holds for $\Upsilon := Y^1 - Y^2$

$$\begin{cases} \Upsilon(u) &= 0, \quad u < 0 \\ \Upsilon(t) &= \int_0^t \int_{-p}^0 \Upsilon(t+s)\, da(u)\, ds, \quad t \geq 0. \end{cases}$$

Hence it follows from Gronwall's lemma that $\Upsilon \equiv 0$ which proves strong uniqueness. That

$$Y_\rho(t) = x_\rho(t) + \int_{-q}^0 \int_0^t x_a(t-s)\, dB(s+u)\, d\sigma(u), \quad t \geq 0$$

solves system (3.3.4) is verified by inserting

$$\int_0^t \int_{-p}^0 Y_\rho(s+u)\, da(u)\, ds = \int_0^t \int_{-p}^0 x_\rho(s+u)\, da(u)\, ds$$

$$+ \int_0^t \int_{-p} \left(\int_{-q}^0 \int_0^{s+u} x_a(s+u-v)\, dB(v+z)\, d\sigma(z) \right) da(u)\, ds = x_\rho(t)$$

$$-x_\rho(0) + \int_{-q}^0 \left(\int_0^t \int_{-p} \int_0^{s+u} x_a(s+u-v)\, dB(v+z)\, da(u)\, ds \right) d\sigma(z), \quad t \geq 0.$$

For the last line we used ordinary Fubini and that the fundamental solution vanishes for negative values. It is verified with stochastic Fubini that for each fixed z it holds that a.s.

$$\int_0^t \int_{-p} \int_0^{s+u} x_a(s+u-v)\, dB(v+z)\, da(u)\, ds = \int_0^t (x_a(t-v) - 1)\, dB(v+z).$$

Hence it follows that a.s.

$$\int_0^t \int_{-p}^0 Y_\rho(s+u)\, da(u)\, ds = Y_\rho(t) - Y_\rho(0) - \int_{-q}^0 B(t+z) - B(z)\, d\sigma(z), \quad t \geq 0.$$

The theorem has been shown. □

As a special case we receive for the Ornstein-Uhlenbeck process X_ρ

$$X_\rho(t) = x_\rho(t) + \int_0^t x_a(t-s)\, dB(s), \quad t \geq 0, \quad (X_\rho)_0 = \rho.$$

Now we are able to express Y_ρ in terms of X_ρ

$$\begin{aligned} Y_\rho(t) &= x_\rho(t) + \int_{-q}^0 \int_0^t x_a(t-s)\, dB(s+u)\, d\sigma(u) \\ &= x_\rho(t) + \int_{-q}^0 \int_u^{t+u} x_a(t+u-v)\, dB(v)\, d\sigma(u) \\ &= x_\rho(t) + \int_{-q}^0 \int_u^0 x_a(t+u-v)\, dB(v)\, d\sigma(u) \\ &+ \int_{-q}^0 1_{\{t+u \geq 0\}} \int_0^{t+u} x_a(t+u-v)\, dB(v)\, d\sigma(u), \quad t \geq 0. \end{aligned}$$

In the last line we can replace the integrand by

$$1_{\{t+u\geq 0\}}(X_\rho(t+u)-x_\rho(t+u)).$$

Since for $-p \leq t+u < 0$ it holds that

$$X_\rho(t+u)=x_\rho(t+u),$$

we can omit the indicator function and obtain finally

$$\begin{aligned} Y_\rho(t)=x_\rho(t) \quad &+ \int_{-q}^{0}\int_{u}^{0} x_a(t+u-v)\,dB(v)\,d\sigma(u) \\ &+ \int_{(-q\vee -p)}^{0} X_\rho(t+u)\,d\sigma(u)-\int_{(-q\vee -p)}^{0} X_\rho(t+u)\,d\sigma(u), \quad t\geq 0. \end{aligned} \tag{3.3.6}$$

This is the desired representation of Y_ρ in terms of the Ohrstein-Uhlenbeck process X_ρ with the same initial condition. The initial segment of the Brownian motion on $[-q,0]$ enters as well as the values of the solution x_ρ of the homogeneous system. Properties as stability, mixing are studied for the Ohrstein-Uhlenbeck process $X_\rho(t)$ as t tends to infinity. The representation in (3.3.6) shows that then Y_ρ has the same asymptotic behavior as X_ρ, if the deterministic homogeneous solution $x_\rho(t)$ tends to zero for $t\to\infty$. In Theorem 3.2.4 we established that, if one chooses the initial condition

$$\rho(t)=\int_{-q}^{0}\left(\int_{-\infty}^{t} x_a(t-s)\,dB(s+u)\right)d\sigma(u), \quad -p\leq t\leq 0,$$

then Y_ρ is a stationary solution of the system (3.3.4). We do not know yet if this the only stationary solution. More exactly speaking, are there two different distributions of the initial conditions ρ_1,ρ_2 such that Y_{ρ_1} and Y_{ρ_2} are stationary solutions of the system (3.3.4)? To tackle this question we start with a moment estimation for solutions of the homogeneous system x_ρ.

3.3.8 Lemma. *The system (3.3.5) has exactly one strong solution. It is given by*

$$x_\rho(t)=\rho(0)x_a(t)+\int_{-p}^{0}\int_{-u}^{0} x_a(t-s+u)\rho(s)\,ds\,da(u), \quad t\geq 0, \tag{3.3.7}$$

where x_a denotes the fundamental solution. Moreover, if

$$E(\int_{-p}^{0}|\rho(s)|^2\,ds)<\infty,$$

then it holds that

$$E(\sup_{t-p\leq v\leq t}|x_\rho(v)|^2)\leq C\mathrm{e}^{-2\delta t}, \quad t\geq 0$$

for $0<\delta<-v_0(a)$.

Proof. That the solution is unique follows from Gronwall's lemma. That x_ρ in (3.3.7) is a solution is verified by inserting. In view of the representation for x_ρ in (3.3.7) we have by repeated use of the Cauchy-Schwartz inequality the following pointwise estimation

$$|x_\rho(t)|^2 \le 2|\rho(0)|^2|x_a(t)|^2 + 2\|a\|_{\mathrm{TV}}^2 p \int_{-p}^{0}\int_{-p}^{0} |x_a(t-s+u)|^2|\rho(s)|^2\, ds\, d|a|(u), \quad t \ge 0.$$

Hence it follows that

$$\begin{aligned} \sup_{t-p\le v\le t} |x_\rho(v)|^2 &\le 2|\rho(0)|^2 \sup_{t-p\le v\le t} |x_a(v)|^2 \\ &+ 2\|a\|_{\mathrm{TV}}^2 p \int_{-p}^{0}\int_{-p}^{0} \sup_{t-p\le v\le t} |x_a(v-s+u)|^2|\rho(s)|^2\, ds\, d|a|(u). \end{aligned}$$

Taking expected values on both sides and using the estimation

$$|x_a(t)| \le K\mathrm{e}^{-\delta t}, \quad t \ge 0$$

for the fundamental solution we obtain

$$\begin{aligned} E(\sup_{t-p\le v\le t} |x_\rho(v)|^2) &\le 2E(|\rho(0)|^2)K^2\mathrm{e}^{2\delta p}\mathrm{e}^{-2\delta t} \\ &+ 2\|a\|_{\mathrm{TV}}^3 p^2 K^2 E\left(\int_{-p}^{0} |\rho(s)|^2\, ds\right)\mathrm{e}^{4\delta p}\mathrm{e}^{-2\delta t}, \quad t \ge 0, \end{aligned}$$

which completes the proof. □

The next lemma estimates the L^2-distance of the stationary solution Y_ρ to any other solution of (3.3.4).

3.3.9 Lemma. *Let the process Y_ρ be given by*

$$Y_\rho(t) = \int_{-q}^{0}\left(\int_{-\infty}^{t} x_a(t-s)\, dB(s+u)\right) d\sigma(u), \quad t \in \mathbb{R}.$$

Assume that $\bar{Y}$ is another solution of (3.3.4) with a square-integrable initial condition. Then it holds that

$$E(\|(Y_\rho)_t - \bar{Y}_t\|_\infty^2) \le C\mathrm{e}^{-2\delta t}, \quad t \ge 0,$$

where $(Y_\rho)_t$ and $\bar{Y}_t$ denote segments on the space $C[-p,0]$.

Proof. We have in view of Theorem 3.3.7 that

$$Y_\rho(t) - \bar{Y}(t) = x_\rho(t) - \int_{-q}^{0}\int_{-\infty}^{0} x_a(t-s)\, dB(s+u)\, d\sigma(u), \quad t \ge 0,$$

and hence

$$|Y_\rho(t) - \bar{Y}(t)|^2 \le 2|x_\rho(t)|^2 + 2\left(\int_{-q}^0 \int_{-\infty}^0 x_a(t-s)\, dB(s+u)\, d\sigma(u)\right)^2, \quad t \ge 0.$$

By partial integration the last expression is a.s. lower or equal than

$$\begin{aligned} 2\left|\int_{-q}^0 B(u) x_a(t)\, d\sigma(u)\right|^2 \;&+\; 2\left|\int_{-q}^0 \int_{-\infty}^0 \dot{x}_a(t-s) B(s+u)\, ds\, d\sigma(u)\right|^2 \\ &\le\; 2|x_a(t)|^2 \left|\int_{-q}^0 |B(u)|\, d|\sigma|(u)\right|^2 \\ &+\; \left|\int_{-q}^0 \int_{-\infty}^0 |\dot{x}_a(t-s)||B(s+u)|\, ds\, d|\sigma|(u)\right|^2, \quad t \ge 0. \end{aligned}$$

The last expression in square brackets may be bounded from above by

$$\begin{aligned} &\|\sigma\|_{\mathrm{TV}} \int_{-q}^0 \left(\int_{-\infty}^0 K \mathrm{e}^{-\delta(t-s)} |B(s+u)|\, ds\right)^2 d|\sigma|(u) \\ \le\; &\|\sigma\|_{\mathrm{TV}} \mathrm{e}^{-2\delta t} \int_{-q}^0 \frac{1}{\delta} \int_{-\infty}^0 K^2 \mathrm{e}^{\delta s} |B(s+u)|^2\, ds\, d|\sigma|(u), \quad t \ge 0, \end{aligned}$$

where we have used the inequality

$$\left(\int_{-\infty}^0 f(s) \mathrm{e}^{\delta s}\, ds\right)^2 \le \frac{1}{\delta} \int_{-\infty}^0 f^2(s) \mathrm{e}^{\delta s}\, ds$$

for all functions $f \ge 0$ such that the integrals exist. This yields for $t \ge 0$

$$\begin{aligned} E(\|(Y_\rho)_t - \bar{Y}_t\|_\infty^2) \quad &\le\quad 2E(\sup_{t-p \le v \le t} |x_\rho(v)|^2) \\ &+\quad 2E \sup_{t-p \le w \le t} \left(\int_{-q}^0 \int_{-\infty}^0 x_a(w-s)\, dB(s+u)\, d\sigma(u)\right)^2 \\ &\le\quad 2K\mathrm{e}^{-2\delta t} + 4\mathrm{e}^{2\delta p} \mathrm{e}^{-2\delta t} E \left|\int_{-q}^0 |B(u)|\, d|\sigma|(u)\right|^2 + \\ 4\|\sigma\|_{\mathrm{TV}} \quad &\mathrm{e}^{2\delta p} \mathrm{e}^{-2\delta t} E(\int_{-q}^0 \frac{1}{\delta} \int_{-\infty}^0 K^2 \mathrm{e}^{\delta s} |B(s+u)|^2\, d, s\, d|\sigma|(u)), \end{aligned}$$

where we used Lemma 3.3.8 for the estimation of the moments of x_ρ. Since all expectations involved are finite, the lemma has been shown. □

Now we are able to answer the question of the uniqueness of a stationary solution.

3.3.10 Theorem. *There exists exactly one stationary solution of (3.3.4) in the class of stochastic processes with square-integrable initial conditions. This solution is given by*

$$Y(t) = \int_{-q}^{0} \left(\int_{-\infty}^{t} x_a(t-s)\, dB(s+u) \right) d\sigma(u), \quad t \in \mathbb{R}.$$

Proof. This follows from Lemma 3.3.9 by a standard argument. □

We saw that the Ornstein-Uhlenbeck process

$$X(t) = \int_{-\infty}^{t} x_a(t-s)\, dB(s), \quad t \in \mathbb{R}$$

is an integral over a differentials of a Brownian motion weighted by the fundamental solution. For the process Y we find that

$$\begin{aligned} Y(t) &= \int_{-q}^{0} X(t+u)\, d\sigma(u) = \int_{-q}^{0} \int_{-\infty}^{t+u} x_a(t+u-s')\, dB(s')\, d\sigma(u) \\ &= \int_{-q}^{0} \left(x_a(0)B(t+u) - \int_{-\infty}^{t+u} B(s')\, dx_a(t+u-s') \right) d\sigma(u) \\ &= \int_{-q}^{0} \left(x_a(0)B(t+u) - \int_{-\infty}^{t} B(u+s)\, dx_a(t-s) \right) d\sigma(u) \\ &= x_a(0)Z(t) - \int_{-\infty}^{t} Z(s)\, dx_a(t-s) =: \int_{-\infty}^{t} x_a(t-s)\, dZ(s), \quad t \in \mathbb{R}. \end{aligned}$$

The integral with respect to the process Z, which is not necessarily a semimartingale, is defined by the expression on the left side following partial integration. Thereby we have expressed Y as an integral weighted by the fundamental solution in analogy to an Ornstein-Uhlenbeck process.

The last intention in this section is to approximate the following stochastic equation with arbitrary random square-integrable initial condition ρ

$$\left\{ \begin{aligned} Y_0 &= \rho \\ dY(t) &= \left(\textstyle\int_{-p}^{0} Y(t+u)\, da(u) \right) dt + \textstyle\int_{-q}^{0} dB(t+u)\, d\sigma(u), \quad t \geq 0. \end{aligned} \right. \tag{3.3.8}$$

We already established that there exists a strong solution Y_ρ. For two different initial conditions standard estimations and a use of Gronwall's lemma give that

$$E\big(\sup_{0 \leq s \leq T} |Y_{\rho_1}(s) - Y_{\rho_2}(s)|^2\big) \leq K_T \|\rho_1 - \rho_2\|^2_{L^2(\Omega, C[-p,0])}, \quad T \geq 0.$$

Hence strong uniqueness holds for the system (3.3.8), and we have continuity in the initial condition. Hence it suffices in view of Theorem 2.4.17 to consider the case of deterministic initial conditions ξ. For $h > 0$ define a process $Y^{(h)}$ with initial condition $\xi^{(h)}$ by

$$\begin{cases} Y^{(h)}_{(m+1)h} &= Y^{(h)}_{mh} + \sum_{j=0}^{p^{(h)}} a^{(h)}_j Y^{(h)}_{(m-j)h} h + \sum_{i=0}^{q^{(h)}} \sigma^{(h)}_i \sqrt{h} \epsilon_{m+1-i}, \quad m \in \mathbb{N}_0 \\ Y^{(h)}_{ih} &= \xi^{(h)}(ih), \quad -p^{(h)} \le i \le 0, \quad Y^{(h)}_t = Y^{(h)}_{[\frac{t}{h}]h}, \quad t \ge -p, \end{cases}$$

where $\{\epsilon_i : i \in \mathbb{Z}\}$ is a sequence of independent, standard Gaussian random variables. Denote the corresponding discrete measures on $[-p, 0]$ by $a^{(h)}$ and on $[-q, 0]$ by $\sigma^{(h)}$. Our goal is to show the following: If for the initial conditions $\xi^{(h)} \xrightarrow[h \to 0]{} \xi$ on the space $C[-p, 0]$ and

$$a^{(h)} \Longrightarrow a, \quad \sigma^{(h)} \Longrightarrow \sigma$$

as weak convergence of measures, then $\{Y^{(h)} : h > 0\}$ converges weakly to the unique strong solution of

$$\begin{cases} Y_0 &= \xi \\ dY(t) &= \left(\int_{-p}^0 Y(t+u)\, da(u)\right) dt + \int_{-q}^0 dB(t+u)\, d\sigma(u), \quad t \ge 0, \end{cases}$$

where B is a Brownian motion on $[-q, \infty)$. The time series $(Y^{(h)}_{mh})_{m \ge -p^{(h)}}$ is not stationary. It is not possible to analyze spectral densities as we did in the stationary case. Furthermore, the weak limit Y is in general no semimartingale. Neither the semimartingale theory can be used to obtain a convergence result. Therefore we have to follow another strategy. We start with a lemma for processes $S^{(h)}$ with vanishing measures $a^{(h)}$.

3.3.11 Lemma. *Let for $h > 0$ a process $S^{(h)}$ be given by*

$$\begin{cases} S^{(h)}_{(m+1)h} &= S^{(h)}_{mh} + \sum_{i=0}^{q^{(h)}} \sigma^{(h)}_i \sqrt{h} \epsilon_{m+1-i}, \quad m \in \mathbb{N}_0 \\ S^{(h)}_0 &= 0, \quad S^{(h)}_t = S^{(h)}_{[\frac{t}{h}]h}, \quad t \ge 0, \end{cases}$$

where $\{\epsilon_i : i \in \mathbb{Z}\}$ is a sequence of independent, standard Gaussian random variables. If $\sigma^{(h)} \Longrightarrow \sigma$ as weak convergence of measures, then $\{S^{(h)} : h > 0\}$ converges weakly to S, where S is given by

$$S(t) = \int_{-q}^0 [B(t+u) - B(u)]\, d\sigma(u), \quad t \ge 0$$

for a Brownian motion B on $[-q, \infty)$.

Proof. We shall prove the tightness of the sequence $\{S^{(h)} : h > 0\}$ by the criterion in Theorem 15.6 in Billingsley [2]. Recall that for a Gaussian system (A, B) with $E(A) = E(B) = 0$ it holds that

$$E(A^2 B^2) = E(A^2)E(B^2) + 2(E(AB))^2 \le 3E(A^2)E(B^2).$$

Since $\{S_t^{(h)} : t \geq 0\}$ is a Gaussian system we obtain for time points $0 \leq t_0 \leq t_1 \leq t_2$

$$E(|S_{t_2}^{(h)} - S_{t_1}^{(h)}|^2 |S_{t_1}^{(h)} - S_{t_0}^{(h)}|^2) \leq 3E(|S_{t_2}^{(h)} - S_{t_1}^{(h)}|^2) E(|S_{t_2}^{(h)} - S_{t_1}^{(h)}|^2).$$

We calculate explicitly

$$\begin{aligned} E(|S_{t_2}^{(h)} - S_{t_1}^{(h)}|^2 &= E\left(\sum_{j=[\frac{t_1}{h}]}^{[\frac{t_2}{h}]-1} \sum_{i=0}^{q^{(h)}} \sigma_i^{(h)} \sqrt{h} \epsilon_{j+1-i}\right)^2 \\ &= E\left(\sum_{j=[\frac{t_1}{h}]}^{[\frac{t_2}{h}]-1} \sum_{i=0}^{q^{(h)}} \sigma_i^{(h)} \sqrt{h} \epsilon_{j+1-i} \sum_{j'=[\frac{t_1}{h}]}^{[\frac{t_2}{h}]-1} \sum_{i'=0}^{q^{(h)}} \sigma_{i'}^{(h)} \sqrt{h} \epsilon_{j'+1-i'}\right) \\ &= \int_{-q}^{0} \int_{-q}^{0} h \sum_{j=[\frac{t_1}{h}]}^{[\frac{t_2}{h}]-1} \sum_{j'=[\frac{t_1}{h}]}^{[\frac{t_2}{h}]-1} E(\epsilon_{j+1-[\frac{u}{h}]} \epsilon_{j'+1-[\frac{u'}{h}]})\, d\sigma^{(h)}(u)\, d\sigma^{(h)}(u'). \end{aligned}$$

By independence, vanishing mean and standard variance of the sequence ϵ the integrand as function of (u, u') equals

$$h \,\sharp \left\{ [\frac{t_1}{h}] \leq j, j' \leq [\frac{t_2}{h}] - 1 : j - j' = [\frac{u}{h}] - [\frac{u'}{h}] \right\} \leq h([\frac{t_2}{h}] - [\frac{t_1}{h}]),$$

where $\sharp$ denotes the number of elements of a finite set (here the indices j and j'). Since by assumption the measures $\sigma^{(h)}$ converge weakly, it holds that

$$\sup_{h>0} \int_{-q}^{0} \int_{-q}^{0} 1\, d|\sigma^{(h)}|(u)\, d|\sigma^{(h)}|(u') \leq C < \infty.$$

This enables us to prove tightness. It holds that

$$\begin{aligned} \overline{\lim_{h\to 0}} E(|S_{t_2}^{(h)} - S_{t_1}^{(h)}|^2 |S_{t_1}^{(h)} - S_{t_0}^{(h)}|^2) &\leq 3 \overline{\lim_{h\to 0}} E(|S_{t_2}^{(h)} - S_{t_1}^{(h)}|^2) E(|S_{t_2}^{(h)} - S_{t_1}^{(h)}|^2) \\ &\leq 3(t_2 - t_1) C (t_1 - t_0) C \leq 3C^2 (t_1 - t_0)^2. \end{aligned}$$

Therefore the sequence $\{S^{(h)} : h > 0\}$ is tight. For the convergence of finite dimensional distributions it suffices to prove convergence of the underlying covariance functions. The same computation as above yields for $0 \leq t_1 \leq t_2$

$$E(S_{t_2}^{(h)} S_{t_1}^{(h)}) = \int_{-q}^{0} \int_{-q}^{0} f_{t_1,t_2}^{(h)}(u, u')\, d\sigma^{(h)}(u)\, d\sigma^{(h)}(u')$$

with

$$\begin{aligned}
f^{(h)}_{t_1,t_2}(u,u') &:= h \sum_{j=0}^{[\frac{t_2}{h}]-1} \sum_{j'=0}^{[\frac{t_2}{h}]-1} E(\epsilon_{j+1-[\frac{u}{h}]}\epsilon_{j'+1-[\frac{u'}{h}]})\, d\sigma^{(h)}(u)\, d\sigma^{(h)}(u') \\
&= h\, \sharp \Big\{ 0 \le j \le [\frac{t_2}{h}]-1, 0 \le j' \le [\frac{t_1}{h}]-1 : j-j' = [\frac{u}{h}] - [\frac{u'}{h}] \Big\} \\
&= h \begin{cases} \Big(([\frac{t_2}{h}] - ([\frac{u}{h}]-[\frac{u'}{h}]) \vee 0\big) \wedge [\frac{t_1}{h}] \Big), & |u'| \ge |u| \\ \Big(([\frac{t_1}{h}] - ([\frac{u'}{h}]-[\frac{u}{h}]) \vee 0\big) \wedge [\frac{t_2}{h}] \Big), & |u'| < |u| \end{cases}.
\end{aligned}$$

We see that $f^{(h)}_{t_1,t_2}(u,u')$ converges uniformly in (u,u') to the function $f_{t_1,t_2}(u,u')$ which is defined by

$$f_{t_1,t_2}(u,u') := \begin{cases} \Big((t_2 - (u-u') \vee 0) \wedge t_1 \Big), & |u'| \ge |u| \\ \Big((t_1 - (u'-u) \vee 0) \wedge t_2 \Big), & |u'| < |u|. \end{cases}$$

It is not hard to see that $f_{t_1,t_2}(u,u')$ equals

$$\begin{cases} (t_1+u-u') \vee t_2 - (u-u') \vee t_2, & |u'| \ge |u| \\ (t_2+u'-u) \vee t_1 - (u'-u) \vee t_1, & |u'| < |u| \end{cases} = E(B_{t_1+u} - B_u)(B_{t_2+u'} - B_{u'}).$$

Now we can finish the proof. We see that

$$\begin{aligned}
E(S^{(h)}_{t_2} S^{(h)}_{t_1}) &= \int_{[-q,0]^2} f^{(h)}_{t_1,t_2}(u,u')\, d(\sigma^{(h)} \otimes \sigma^{(h)})(u,u') \\
&\xrightarrow[h\to 0]{} \int_{[-q,0]^2} f_{t_1,t_2}(u,u')\, d(\sigma \otimes \sigma)(u,u') = E(S(t_2)S(t_1)),
\end{aligned}$$

where we used uniform convergence of the functions $f^{(h)}_{t_1,t_2}(u,u')$ to $f_{t_1,t_2}(u,u')$ on compacts of (u,u'). The lemma has been shown. □

Next we need a purely analytical lemma which is proven in Lorenz [17].

3.3.12 Lemma. *Let for $h > 0$ a deterministic function $y^{(h)}$ be given by*

$$\begin{cases} y^{(h)}_{ih} = \xi^{(h)}_{ih}, & -p^{(h)} \le i \le 0 \\ y^{(h)}_{(m+1)h} = y^{(h)}_{mh} + \sum_{j=0}^{p^{(h)}} a^{(h)}_j y^{(h)}_{(m-j)h} + (s^{(h)}_{(m+1)h} - s^{(h)}_{mh}), & m \in \mathbb{N}_0 \\ y^{(h)}_t = y^{(h)}_{[\frac{t}{h}]h}, & t \ge -p \end{cases}$$

for an arbitrary deterministic function $s^{(h)}$ with value zero at time point zero. If for functions $s^{(h)} \xrightarrow[h\to 0]{} s$ uniformly on $[0,T]$, for the initial conditions $\xi^{(h)} \xrightarrow[h\to 0]{} \xi$

uniformly on $[-p, 0]$ *and for the measures* $a^{(h)} \Longrightarrow a$*, then the functions* $y^{(h)}$ *converge to the unique solution of the system*

$$\begin{cases} y_0 &= \xi \\ y(t) &= y(0) + \int_0^t \int_{-p}^0 y(s+u)\, da(u)\, ds + s(t), \quad t \geq 0 \end{cases}$$

uniformly on $[-p, T]$.

Now we are able to prove the following theorem.

3.3.13 Theorem. *Let for* $h > 0$ *a process* $Y^{(h)}$ *be given by*

$$\begin{cases} Y_{ih}^{(h)} &= \xi_{ih}^{(h)}, \quad -p^{(h)} \leq i \leq 0 \\ Y_{(m+1)h}^{(h)} &= Y_{mh}^{(h)} + \sum\limits_{j=0}^{p^{(h)}} a_j^{(h)} Y_{(m-j)h}^{(h)} h + \sum\limits_{i=0}^{q^{(h)}} \sigma_i^{(h)} \sqrt{h} \epsilon_{m+1-i}, \quad m \in \mathbb{N}_0 \\ Y_t^{(h)} &= Y_{[\frac{t}{h}]h}^{(h)}, \quad t \geq -p, \end{cases}$$

where $\{\epsilon_i : i \in \mathbb{Z}\}$ *is a sequence of independent, standard Gaussian random variables. If for the initial conditions* $\xi^{(h)} \xrightarrow[h \to 0]{} \xi$ *uniformly on* $[-p, 0]$ *and for the measures* $a^{(h)} \Longrightarrow a$*,* $\sigma^{(h)} \Longrightarrow \sigma$*, then the sequence of processes* $Y^{(h)}$ *converges weakly to* Y*, where* Y *is the unique strong solution of the system*

$$\begin{cases} Y_0 &= \xi \\ Y(t) &= Y(0) + \int_0^t \int_{-p}^0 Y(s+u)\, da(u)\, ds + \int_{-q}^0 [B(t+u) - B(u)]\, d\sigma(u), \quad t \geq 0 \end{cases}$$

with B *a Brownian motion on* $[-q, \infty)$.

Proof. By Lemma 3.3.11 the sequence of processes $S^{(h)}$ defined by

$$\begin{cases} S_{(m+1)h}^{(h)} &= S_{mh}^{(h)} + \sum\limits_{i=0}^{q^{(h)}} \sigma_i^{(h)} \sqrt{h} \epsilon_{m+1-i}, \quad m \in \mathbb{N}_0 \\ S_0^{(h)} &= 0, \quad S_t^{(h)} = S_{[\frac{t}{h}]h}^{(h)}, \quad t \geq -p \end{cases}$$

converges weakly to the process

$$S(t) = \int_{-q}^0 [B(t+u) - B(u)]\, d\sigma(u), \quad t \geq 0.$$

By the almost sure representation theorem there exists a probability space $(\bar{\Omega}, \bar{\mathfrak{F}}, \bar{P})$ and a sequence of processes $\bar{S}^{(h)}$ and $\bar{S}$ with values in $C[0, \infty)$, all defined on $\bar{\Omega}$, such that

$$\text{Law}(\bar{S}^{(h)}) = \text{Law}(S^{(h)}) \quad \forall h > 0, \quad \text{Law}(\bar{S}) = \text{Law}(S), \quad \bar{S}^{(h)} \xrightarrow[h \to 0]{} \bar{S} \quad \text{a.s.}$$

Then the process $\bar{Y}^{(h)}$ defined on $\bar{\Omega}$ by

$$\begin{cases} \bar{Y}^{(h)}_{ih} &= \xi^{(h)}_{ih}, \quad -p^{(h)} \leq i \leq 0 \\ \bar{Y}^{(h)}_{(m+1)h} &= \bar{Y}^{(h)}_{mh} + \sum_{j=0}^{p^{(h)}} a^{(h)}_j \bar{Y}^{(h)}_{(m-j)h} h + (\bar{S}^{(h)}_{(m+1)h} - \bar{S}^{(h)}_{mh}), \quad m \in \mathbb{N}_0 \\ \bar{Y}^{(h)}_t &= \bar{Y}^{(h)}_{[\frac{t}{h}]h}, \quad t \geq -p \end{cases}$$

has the same distribution as $Y^{(h)}$. By Lemma 3.3.12 $\{\bar{Y}^{(h)} : h > 0\}$ converges a.s. to the solution of

$$\begin{cases} \bar{Y}_0 &= \xi \\ \bar{Y}(t) &= \bar{Y}(0) + \int_0^t \int_{-p}^0 \bar{Y}(s+u)\, da(u)\, ds + \bar{S}(t), \quad t \geq 0. \end{cases}$$

Since $\bar{Y}$ has the same distribution as Y given by

$$\begin{cases} Y_0 &= \xi \\ Y(t) &= Y(0) + \int_0^t \int_{-p}^0 Y(s+u)\, da(u)\, ds + S(t), \quad t \geq 0, \end{cases}$$

the proof is finished. □

We shall now return to GARCH(p,q)-models in financial mathematics. Let us recall the general GARCH(p, q)-scheme

$$\begin{cases} X^{(h)}_{(m+1)h} &= X^{(h)}_{mh} \quad + \rho^{(h)}_{mh} \sqrt{h} \epsilon_{m+1} \\ \rho^{(h)2}_{(m+1)h} &= v^{(h)}_0 \quad + \beta^{(h)} \rho^{(h)2}_{mh} + \sum_{j=1}^{p^{(h)}} \beta^{(h)}_j \rho^{(h)2}_{(m-j)h} \\ & \qquad + \alpha^{(h)} \rho^{(h)2}_{mh} h \epsilon^2_{m+1} + \sum_{i=1}^{q^{(h)}} \alpha^{(h)}_i \rho^{(h)2}_{(m-i)h} h \epsilon^2_{m+1-i}, \quad m \in \mathbb{N}_0. \end{cases}$$

The last line of this scheme may be rewritten as

$$\begin{aligned} \rho^{(h)2}_{(m+1)h} - \rho^{(h)2}_{mh} &= \frac{v^{(h)}_0}{h} h + \left(\frac{\beta^{(h)} - 1}{h} + \alpha^{(h)} \right) h \rho^{(h)2}_{mh} + \sqrt{h} \alpha^{(h)} \rho^{(h)2}_{mh} \sqrt{h} (\epsilon^2_m - 1) \\ &+ \sum_{k=1}^{p^{(h)} \vee q^{(h)}} \left(\frac{\beta^{(h)}_k}{h} + \alpha^{(h)}_k \right) h \rho^{(h)2}_{(m-k)h} + \sqrt{h} \alpha^{(h)}_k \rho^{(h)2}_{(m-k)h} \sqrt{h} (\epsilon^2_{m+1-k} - 1) \end{aligned}$$

with the convention that $\beta^{(h)}_k = 0$ for $k > p^{(h)}$ and $\alpha^{(h)}_k = 0$ for $k > q^{(h)}$. Then this is an ARMA-scheme for $\rho^{(h)2}$ with three modifications:

1. An additional constant term v_0 occurs.

2. The distribution of the driving force $\eta_m := \epsilon^2_m - 1$ is not Gaussian.

3. The most important difference is that the factors $(\alpha_k^{(h)}\rho_{(m-k)h}^{(h)2}h)$ at η_{m+1-k} are random.

Associate discrete measures $\kappa^{(h)}$ on $[-(p\vee q),0]$ with the set of coefficients $\kappa_k^{(h)} := \frac{\beta_k^{(h)}}{h}+\alpha_k^{(h)}$ and $\phi^{(h)}$ on $[-q,0]$ with the set of coefficients $\phi_k^{(h)} := \alpha_k^{(h)}$. Now assume the following asymptotic behavior for the real numbers

$$\frac{v_0^{(h)}}{h}\xrightarrow[h\to 0]{} v_0,\quad -\left(\frac{\beta^{(h)}-1}{h}+\alpha^{(h)}\right)\xrightarrow[h\to 0]{}\lambda,\quad \sqrt{h}\alpha^{(h)}\xrightarrow[h\to 0]{}\alpha$$

and for the discrete measures

$$\kappa^{(h)}\Longrightarrow\kappa \quad\text{on}\quad [-(p\vee q),0],\quad \sqrt{h}\phi^{(h)}\Longrightarrow\phi\quad\text{on}\quad[-q,0].$$

Under those assumptions it is suggested that $(X^{(h)},\rho^{(h)2})$ converges weakly to (X,ρ^2), where (X,ρ^2) is the solution of

$$\begin{cases} dX(t) &=& \rho(t)\,dW(t)\\ d\rho^2(t) &=& \left[v_0-\lambda\rho^2(t)+\int_{-p}^0\rho^2(t+u)\,d\kappa(u)\right]\,dt+c\alpha\rho^2(t)\,dB(t)\\ && +\int_{-q}^0\rho^2(t+u)c\,dB(t+u)\,d\phi(u),\quad t\geq 0\end{cases}$$

for a two-dimensional Brownian motion (W,B) and $c=\sqrt{E|\epsilon_1|^4-1}$. The integrated form of the last differential is

$$\int_{-q}^0\int_u^{t+u}\rho^2(s)c\,dB(s)\,d\phi(u).$$

A proof of this heuristic result is unknown. Note that ρ^2 is in general no semimartingale. Hence no semimartingale characterization is possible. On the other hand, it is the weak limit of modified ARMA-processes, where we listed three modifications which do not allow us to deduce convergence directly from the established result of convergence of ARMA-processes. If especially $p=q$ and $\alpha_i^{(h)}=\alpha(-ih)h$ for a continuous function α on $[-q,0]$, then $\phi\equiv 0$ and $d\kappa(u)=d\gamma(u)+\alpha(u)\,d(u)$. In this case we recover system (2.5.14) of the previous chapter.

3.4 Comparison to Literature

In this chapter, in contrary to the previous one, reference to literature is little. The weak convergence of processes $Y^{(h)}$ defined by

$$\begin{cases} Y_{(m+1)h}^{(h)} &=& Y_{mh}^{(h)}+\sum_{j=0}^{p^{(h)}}a_j^{(h)}Y_{(m-j)h}^{(h)}h+\sum_{i=0}^{q^{(h)}}\sigma_i^{(h)}\sqrt{h}\epsilon_{m+1-i},\quad m\in\mathbb{Z}\\ Y_t^{(h)} &=& Y_{[\frac{t}{h}]h}^{(h)},\quad t\in\mathbb{R}\end{cases}$$

is found in Reiß [23] in the case $q^{(h)} = 0$ for all $h > 0$. We established that the weak limit of $\{Y^{(h)} : h > 0\}$ is the unique stationary strong solution of

$$Y(t) = Y(0) + \int_0^t \int_{-p}^0 Y(s+u)\,da(u)\,ds + \int_{-q}^0 [B(t+u) - B(u)]\,d\sigma(u). \tag{3.4.1}$$

If $\sigma = \delta_{\{0\}}$, then the last integral in (3.4.1) reduces to $B(t)$. Gushchin and Küchler [7] treat the question of existence of a stationary solution of (3.4.1) if the driving force B is replaced by a Lévy process.

The process Y can be interpreted as continuous-time analogue of ARMA-processes Y_n in discrete time given by

$$a_p Y_n + a_{p-1} Y_{n-1} + \ldots + a_1 Y_{n-p+1} + Y_{n-p} = b_0 \epsilon_n + b_1 \epsilon_{n-1} + \ldots + b_q \epsilon_{n-q},$$

where the number of coefficients increases unboundedly as n tends to infinity. However, if p and q remain fixed for all $n \in \mathbb{N}$, one obtains another continuous-time analogue of an ARMA-process called CARMA-process. It occurs in Fasen [5] in Example 1.1.11. and is defined as follows. If $q < p$ and polynomials a and b are defined by

$$a(z) := a_p + \ldots + a_1 z^{p-1} + z^p, \quad b(z) := b_0 + b_1 z + \ldots + b_q z^q, \quad z \in \mathbb{C},$$

then one can formally write down a process Z in continuous time given by the stochastic differential equation

$$a(D)Z(t) = b(D)DB(t). \tag{3.4.2}$$

Here D denotes the differential operator with respect to t. The solution of (3.4.2) is called CARMA(p, q)-process. The first letter "C" stands for "continuous". An explicit form for CARMA(p, q)-processes can be constructed in the following way. At first a $p-$dimensional process X is defined by

$$\begin{cases} DX_i(t) &= X_{i+1}(t), \quad i = 1, \ldots, p-1 \\ DX_p(t) &= -a_p X_1(t) - a_{p-1} X_2(t) - \ldots - a_1 X_p(t) + DB(t), \quad t \geq 0. \end{cases}$$

The compact form of this equation on $\mathbb{R}^p$ is

$$dX(t) = AX(t)\,dt + e\,dB(t), \quad e = (0, \ldots, 0, 1).$$

This tells us that X is an Ornstein-Uhlenbeck process on $\mathbb{R}^p$ with the following stationary representation

$$X(t) = \int_{-\infty}^{+\infty} \mathrm{e}^{(t-s)A} e\,dB(s), \quad t \in \mathbb{R}$$

if all eigenvalues of A, which equal the zeroes of the polynomial a, have negative real part. Then the process

$$Z(t) := b_0 X_1(t) + b_1 X_2(t) + \ldots + b_q X_{q+1}(t), \quad t \in \mathbb{R}$$

is a stationary solution of equation (3.4.2). From this representation we deduce that in general the CARMA-process Z in (3.4.2) is a different processes than the process Y in (3.4.1). The CARMA-process is the sum of components of a p-dimensional Markov process, whereas the process Y is the solution of a stochastic delay equation. Also the kernel functions f and spectral densities g have a different structure

$$f_{\text{CARMA}}(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{e}^{\mathrm{i}t\omega}\frac{b(\mathrm{i}\omega)}{a(\mathrm{i}\omega)}\,d\omega, \qquad f_Y(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{e}^{\mathrm{i}\omega t}\frac{\int_{-q}^{0}\mathrm{e}^{\mathrm{i}\omega u}\,d\sigma(u)}{\chi_a(\mathrm{i}\omega)}\,d\omega$$

$$g_{\text{CARMA}}(\lambda) = \frac{1}{2\pi}\left|\frac{b(\mathrm{i}\lambda)}{a(\mathrm{i}\lambda)}\right|^2, \qquad g_Y(\lambda) = \frac{1}{2\pi}\frac{\left|\int_{-q}^{0}\mathrm{e}^{\mathrm{i}\lambda u}\,d\sigma(u)\right|^2}{|\chi_a(\mathrm{i}\lambda)|^2}.$$

Only in the case $\sigma = \delta_{\{0\}}$ and $x_a(t) = \mathrm{e}^{-\lambda t}$ for $\lambda > 0$ the kernel functions are the same. In this case Z is a CARMA(1,0)-process

$$Z(t) = \int_{-\infty}^{+\infty} 1_{\{t\geq s\}}\mathrm{e}^{-\lambda(t-s)}\,dB(s), \quad t \in \mathbb{R},$$

also known as one-dimensional Ornstein-Uhlenbeck process. As already mentioned CARMA-processes can be viewed as weak limits of ARMA-processes where the number of coefficients is constant and the memory length shrinks to zero. We will illustrate this for the CARMA(2,0)-process. Let the polynomials a of degree 2 and b of degree 0 be defined by

$$a(z) := a_2 + a_1 z + z^2, \quad b(z) :\equiv 1, \quad z \in \mathbb{C},$$

where the zeroes of a are assumed to have negative real part. Then define for $h > 0$ a two-dimensional scheme $(X^{1,(h)}, X^{2,(h)})$ by

$$\begin{cases} X^{1,(h)}_{(m+1)h} - X^{1,(h)}_{mh} &= X^{2,(h)}_{mh}h \\ X^{2,(h)}_{(m+1)h} - X^{2,(h)}_{mh} &= -a_2 X^{1,(h)}_{mh}h - a_1 X^{2,(h)}_{mh}h + \sqrt{h}\epsilon_{m+1}, \quad m \in \mathbb{Z}. \end{cases}$$

Then one sees after some computations that for $Y^{(h)} := X^{1,(h)}$ it holds that

$$Y^{(h)}_{(m+2)h} + (a_1 h - 2)Y^{(h)}_{(m+1)h} + (a_2 h^2 - a_1 h + 1)Y^{(h)}_{mh} = h\sqrt{h}\epsilon_{m+1}, \quad m \in \mathbb{Z}.$$

We know from the beginning of this chapter that the covariance of $(Y^{(h)}_{mh})_{m\in\mathbb{Z}}$ has the following representation

$$E(Y^{(h)}_{[\frac{t}{h}]h}Y^{(h)}_0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}\lambda[\frac{t}{h}]}\frac{|Q^{(h)}(\mathrm{e}^{-\mathrm{i}\lambda})|^2}{|P^{(h)}(\mathrm{e}^{-\mathrm{i}\lambda})|^2}\,d\lambda = \frac{1}{2\pi}\int_{-\pi/h}^{\pi/h} \mathrm{e}^{\mathrm{i}\lambda[\frac{t}{h}]h}\frac{|Q^{(h)}(\mathrm{e}^{-\mathrm{i}\lambda h})|^2}{|P^{(h)}(\mathrm{e}^{-\mathrm{i}\lambda h})|^2}h\,d\lambda$$

with

$$P^{(h)}(z) := 1 + (a_1 h - 2)z + (a_2 h^2 - a_1 h + 1)z^2, \quad Q^{(h)}(z) := h\sqrt{h}z, \quad z \in \mathbb{C}.$$

By differentiating the nominator and denominator of the fraction $\sqrt{h}Q^{(h)}/P^{(h)}$ two times with respect to h it follows for the asymptotic behavior of the spectral densities that

$$\lim_{h\to 0}\frac{1}{2\pi}\left|\frac{\sqrt{h}Q^{(h)}(\mathrm{e}^{-\mathrm{i}\lambda h})}{P^{(h)}(\mathrm{e}^{-\mathrm{i}\lambda h})}\right|^2 = \frac{1}{2\pi}\left|\frac{1}{a_2 + a_1(\mathrm{i}\lambda) + (\mathrm{i}\lambda)^2}\right|^2 = \frac{1}{2\pi}\left|\frac{1}{a(\mathrm{i}\lambda)}\right|^2 = g_{\mathrm{CARMA}}(\lambda).$$

Bibliography

[1] D. R. Bell and S.-E. Mohammed. Smooth densities for degenerate stochastic delay equations with hereditary drift. *The Annals of Probability*, 23(4):1875–1894, 1995.

[2] P. Billingsley. *Convergence of probability measures.* Wiley, New York, 1968.

[3] E. Buckwar and T. Shardlow. Weak approximation of stochastic differential delay equations. *IMA Journal of Numerical Analysis*, 25(1):57–86, 2005.

[4] E. Buckwar and R. Winkler. Multi-step methods for SDDEs and their applications to problems with small noise. Preprint, Humboldt University Berlin, 2003.

[5] V. Fasen. *Extremes of Lévy driven moving average processes with applications in finance.* PhD thesis, TU München, 2004.

[6] I. I. Gichman and A. V. Skorochod. *Vvedenie v teoriju slucajnych processov.* Nauka, Moskva, 1977.

[7] A. A. Gushchin and U. Küchler. On stationary solutions of delay differential equations driven by a Lévy process. *Stochastic Processes and their Applications*, 88(2):195–211, 2000.

[8] Y. Hu, S.-E. Mohammed, and F. Yan. Numerical solutions of stochastic differential systems with memory. Preprint, Southern Illinois University, Carbondale, 2001.

[9] K. Itô and M. Nisio. On stationary solutions of a stochastic differential equation. *Journal of Mathematics of Kyoto University*, 4:1–75, 1964.

[10] J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes.* Springer, Berlin, 1987.

[11] T. Jeantheau. A link between complete models with stochastic volatility and ARCH models. *Finance and Stochastics*, 8(1):111–131, 2004.

[12] I. Karatzas and S. Shreve. *Brownian motion and stochastic calculus. 2nd ed.* Springer, New York, 1991.

[13] P. E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations.* Springer, Berlin, 1992.

[14] A. Kohatsu-Higa. Weak approximations. A Malliavin calculus approach. *Mathematics of Computation*, 70(233):135–172, 2001.

[15] U. Küchler and E. Platen. Weak discrete time approximation of stochastic differential equations with time delay. *Mathematics and Computers in Simulation*, 59(6):497–507, 2002.

[16] T. Kurtz and P. Protter. Weak limit theorems for stochastic integrals and stochastic differential equations. *The Annals of Probability*, 19:1035–1070, 1991.

[17] R. Lorenz. Weak approximation of stochastic differential equations by discrete time series. Master's thesis, Humboldt University Berlin, 2002.

[18] X. Mao. Numerical solutions of stochastic functional differential equations. *LMS Journal of Computation and Mathematics*, 6:297–313, 2003.

[19] X. Mao. *Stochastic differential equations and their applications.* Horwood Publ., Chichester, 1997.

[20] S.-E. Mohammed. Stochastic differential systems with memory: theory, examples and applications. In L. Decreusefond, J. Gjerde, B. Øksendal, and A.-S. Ustunel, editors, *Stochastic Analysis and Related Topics VI. The Geilo Workshop, 1996.* Birkhäuser 1-77, 1998.

[21] S.-E. Mohammed. *Stochastic functional differential equations.* Pitman, Boston, 1984.

[22] J. Pedersen. A class of continuous-time garch-type processes. *Stochastics and Stochastics Report*, 65(3-4):229–254, 1999.

[23] M. Reiß. *Nonparametric estimation for stochastic delay differential equations.* PhD thesis, Humboldt University Berlin, 2002.

[24] D. Revuz and M. Yor. *Continuous martingales and Brownian motion.* Springer, Berlin, 1991.

[25] M. Riedle. *Stochastische Differentialgleichungen mit unendlichem Gedächtnis.* PhD thesis, Humboldt University Berlin, 2003.

[26] M. Riedle and X. Mao. Mean square stability of stochastic Volterra integro-differential equations. Preprint, University of Strathclyde, Glasgow, 2003.

[27] A. Shiryaev. *Wahrscheinlichkeit.* Deutscher Verlag der Wissenschaften, Berlin, 1987.

[28] D. W. Stroock and S. R. Varadhan. *Multidimensional diffussion processes.* Springer, Berlin, 1979.

[29] K. Twardowska. On the approximation theorem of the Wong-Zakai type for the functional stochastic differential equations. *Probability and Mathematical Statistics*, 12(2):319–334, 1991.

[30] E. Wong and M. Zakai. On the convergence of ordinary integrals to stochastic integrals. *The Annals of Mathematical Statistics*, 36:1560–1564, 1965.

[31] L. Yan. The Euler scheme with irregular coefficients. *The Annals of Probability*, 30:1172–1194, 2002.

Appendix A

Acknowledgement

The process of writing a thesis is impossible without the help of many people.

It was very helpful for me to join the working team of Prof. Küchler. The collaboration with my colleagues Dr. Markus Riedle, Dr. Markus Reiss, Dr. Evelyn Buckwar and Stefan Tappe at the Humboldt University Berlin gave new ideas and inspirations.

Also my table tennis friends helped me unconsciously. Regularly they showed interest in how far I have proceeded with my thesis and wished me all the best.

My parents gave me love and emotional support as well. Sometimes I told my father - who also studied mathematics - about my results. It was very interesting for me to formulate mathematical results in simple words for someone who has not been dealing with mathematics for 30 years.

I gratefully thank all of you who supported me, consciously or unconsciously.

I gratefully acknowledge that the research has been financed by the Deutsche Forschungsgemeinschaft (DFG), Graduiertenkolleg 251 "Stochastische Prozesse und probabilistische Analysis".

Selbständigkeitserklärung

Hiermit erkläre ich, die vorliegende Arbeit selbständig ohne fremde Hilfe verfasst und nur die angegebene Literatur und die angegebenen Hilfsmittel verwendet zu haben.

RobertLorenz
25. Oktober 2005