Alma Mater Studiorum - Università di Bologna

DOTTORATO DI RICERCA IN

Scienze Statistiche

Ciclo 32

Settore Concorsuale:01/A3

Settore Scientifico Disciplinare:MAT/06

On SDE systems with non-Lipschitz diffusion coefficients

Presentata da:
Vinayak Chuni

Supervisore
Prof. Enrico Bernardi

Coordinatore Dottorato
Prof. Alessandra Luati

Esame finale anno 2020

"To most outsiders, modern mathematics is unknown territory. its borders are protected by dense thickets of technical terms; its landscapes are a mass of indecipherable equations and incomprehensible concepts.Few realize that the world of modern mathematics is rich with vivid images and provocative ideas."

Ivars Peterson

Statutory declaration

I hereby declare that I have developed and written the following PhD Thesis completely by myself. Wherever contributions of others are involved, every effort has been made to indicate this clearly through references to the Bibliography and acknowledgements.
Bologna, February 17th, 2020

Vinayak Chuni

Contents

Abstract

This thesis is a compilation of two papers. In the first paper we investigate a class of two dimensional stochastic differential equations related to susceptible-infected-susceptible epidemic models with demographic stochasticity. While preserving the key features of the model considered in [1], where an *ad hoc* approach has been utilized to prove existence, uniqueness and non explosivity of the solution, we consider an encompassing family of models described by a stochastic differential equation with random and Hölder continuous coefficients. We prove the existence of a unique strong solution by means of a Cauchy-Euler-Peano approximation scheme which is shown to converge in the proper topologies to the unique solution.

In the second paper we link a general method for modeling random phenomena using systems of stochastic differential equations to the class of affine stochastic differential equations. This general construction emphasizes the central role of the Duffie-Kan system [2] as a model for first order approximations of a wide class of nonlinear systems perturbed by noise. We also specialize to a two dimensional framework and propose a direct proof of the Duffie-Kan theorem which does not pass through the comparison with an auxiliary process. Our proof produces a scheme to obtain an explicit representation of the solution once the one dimensional square root process is assigned.

Key words and phrases: stochastic differential equations, square root process, Feller condition, two dimensional susceptible-infected-susceptible epidemic model, Brownian motion

AMS 2000 classification: 60H10, 60H30, 92D30

Acknowledgements

First and foremost I want to thank my advisor Prof. Enrico Bernardi. It has been an honor to be his Ph.D. student. I am deeply grateful for his guidance and the time he invested in supervising me and making my PhD experience very stimulating and productive. This dissertation would not have been possible without his support and help.

Secondly I would like to thank Prof Alberto Lanconelli for his valuable time, support and guidance. I have learned a lot from him and our discussions have been instrumental in writing my dissertation.

Finally, I must express my very profound gratitude to my parents specially my dad who encouraged me to go and study abroad and supported me financially. If it were not for him, I would have never taken up the challenge of undertaking doctoral studies. I am deeply indebted to my mother for supporting me through thick and thin and instilling in me the resilience to never give up on my dreams. I would also like to thank my sister and my girlfriend for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them.

Chapter 1

Introduction

The term Stochastic differential equation was introduced by S. Bernstein(see [3],[4]) in the limiting study of a sequence of Markov chains arising in a stochastic differential scheme. He was only interested in the distribution of limiting processes and showed that the latter had a density satisfying the Kolmogorov equations. However according to Gihman and Skorohod(see [5]) it would be an exaggeration to consider Bernstein the founder of this theory. Independently of Itô's work, I.I. Gihman(see [6],[7] and [8]) developed a theory of stochastic differential equations complete with results on existence, uniqueness, smooth dependence on initial conditions and Kolmogorov's equations for the transition density.

Since the early work of Itô and Gihman, the interest in the methodology and the mathematical theory of Stochastic differential equations has enjoyed remarkable success. The constructive and intuitive nature of the concept as well as the strong physical appeal, has been responsible for its popularity among applied scientists. Stochastic differential equations are now one of the most popular tools to model real world phenomenon. They have many applications in domains such epidemiology, financial modeling (interest rate modeling, option pricing etc), target tracking and medical technology methodologies such as filtering, smoothing, parameter estimation, and machine learning.There are also a wide range of examples of applications of SDEs arising in physics and electrical engineering. In order to simulate and model real world phenomenon using stochastic differential equations and draw conclusions from the solutions, it is imperative to know the existence and uniqueness of solutions. Moreover the theory of of existence and uniqueness of solutions of stochastic differential equations is quite deep and challenging particularly when the coefficients of the SDE are non-regular.

The first result on strong existence and uniqueness of SDE's was due to Itô ([9]) where he assumed that both the drift and the diffusion coefficients b and σ respectively in equation (1.0.1) were uniformly Lipschitz.

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t \tag{1.0.1}$$

In practice, one often needs stochastic differential equations with non-Lipschitz coefficients to model real-world systems. It is often the case that the volatility of the process is the square root of the solution. In other words the dispersion coefficient is Hölder-continuous in the space variable. From a mathematical point of view the analysis of existence and uniqueness for strong solutions of SDEs with Hölder-continuous coefficients is quite challenging. In the one dimensional case, resorting to the famous Yamada-Watanabe principle (i.e. weak existence plus pathwise uniqueness implies strong existence) one can prove the existence of a unique strong solution for SDEs where the drift coefficient is locally Lipshiptz-continuous while the diffusion coefficient is of the type $\sigma(x) = |x|^\alpha$ for $\alpha \in [1/2, 1]$. The hard part of this proof is the pathwise uniqueness which heavily relies on an *ad hoc* technique introduced by Yamada and Watanabe [10] (see also the books Ikeda and Watanabe [11] and Karatzas and Shreve [12] for comparison theorems obtained with a similar approach).

Multi-dimensional linear SDE's are used to model many real world phenomenon for example in stochastic demographic models(see Mao [1]) and in interest rate modeling(see Duffie and Kan [2] and Cairns [13]) and have a rich theory when the system of SDEs is linear, but the moment we start working with systems of SDE's with non-Lipschitz diffusion and dispersion coefficients, the analysis of existence and uniqueness becomes quite intractable. In this thesis we attempt to investigate the existence and uniqueness of two dimensional stochastic differential equations with non-regular diffusion coefficients.

The thesis is organized as follows. In the first chapter we will give a high-level introduction to Stochastic differential equations. In the second chapter we will give a very detailed introduction to the tools needed to study and investigate the existence and uniqueness of solutions of stochastic differential equations. In particular we will provide an extensive introduction to stochastic analysis and stochastic integration using results from the book of Karatzas and Shreve (see [12]) and the book of Ikeda and Watanabe(see [11]). This is followed by a short section which contains some major strong existence and uniqueness results due to Itô and Yamada and Watanabe.

The last two chapters contain two papers "On a general model system related to affine stochastic differential equations" ([14]) and "On a class of stochastic differential equations with random and Hölder continuous coefficients arising in biological modeling"(see [15]). These papers are joint work with my PhD supervisor Prof Enrico Bernardi and Prof Alberto Lanconelli. In these papers we prove existence and uniqueness results for systems of stochastic differential equations with non Lipschitz diffusion coefficients.

I have also included for the sake of completeness a short appendix containing some important results on weak convergence, tightness, convergence of finite-dimensional distributions and the invariance principle. I spent a considerable amount of time on these topics during my Phd studies and these subjects are closely related to the study of Brownian motion which is the process driving the stochastic differential equations under study in this thesis.

The first paper we investigate a class of two dimensional stochastic differential equations related to susceptible-infected-susceptible epidemic models with demographic stochasticity. While preserving the key features of the model considered in [1], where an *ad hoc* approach has been utilized to prove existence, uniqueness and non explosivity of the solution, we consider an encompassing family of models described by a stochastic differential equation with random and Hölder continuous coefficients. We prove the existence of a unique strong solution by means of a Cauchy-Euler-Peano approximation scheme which is shown to converge in the proper topologies to the unique solution.

In the second paper we link a general method for modeling random phenomena using systems of stochastic differential equations to the class of affine stochastic differential equations. This general construction emphasizes the central role of the Duffie-Kan system [2] as a model for first order approximations of a wide class of nonlinear systems perturbed by noise. We also specialize to a two dimensional framework and propose a direct proof of the Duffie-Kan theorem which does not pass through the comparison with an auxiliary process. Our proof produces a scheme to obtain an explicit representation of the solution once the one dimensional square root process is assigned.

Chapter 2

Preliminaries

In this chapter we will introduce some of the important tools and machinery that will be subsequently used in the following sections and chapters. The first section on Stopping times gives a number of definitions and results (without proof) which will be used repeatedly in the text. The second section will outline some important results from stochastic analysis, in particular the construction of stochastic integrals with respect to local martingales. In the second section we precisely define what it means for a stochastic differential equation to have a strong solution. We present the most important results from literature on strong existence and uniqueness of SDEs and a comparison result which will play a very important role later on. In the third section we introduce two important SDEs-the square root process and the mean reverting square root process which are used extensively in interest rate modeling and play a central role in the thesis.

2.1 Stopping Times

In this section we provide some definitions and preliminary results on stopping times which will be used later. We skip the proofs in the section for the sake of brevity.

Definition 2.1.1. *Let us consider a measurable space* $(\Omega, \mathcal{F})$ *equipped with the filtration* $\{\mathcal{F}_t\}$*. The random time* T *is a stopping time of the filtration, if the event* $\{T \leq t\}$ *belongs to the sigma-field* $\{\mathcal{F}_t\}$*, for every* $t \geq 0$*. A random time* T *is an optional time of the filtration , if* $\{T < t\} \in \mathcal{F}_t$*, for every* $t \geq 0$*.*

Lemma 2.1.2. *Let* X *be a stochastic process and* T *a stopping time of* $\{\mathcal{F}_t^X\}$*. Suppose that for the pair* $\omega, \omega' \in \Omega$*, we have* $X_t(\omega) = X_t(\omega')$ *for all* $t \in [0, T(\omega)] \cap [0, \infty)$*. Then* $T(\omega) = T(\Omega')$

Proposition 2.1.3. *Every random time equal to a nonnegative constant is a stopping time. Every stopping time is optional, and the two concepts coincide if the filtration is right continuous.*

Corollary 2.1.4. *T is an optional time of the filtration $\{\mathcal{F}_t\}$ if and only if it is the stopping time of the right continuous filtration $\{\mathcal{F}_{t+}\}$*

Definition 2.1.5. *Let X be a stochastic process with right-continuous paths, which is adapted to the filtration $\{\mathcal{F}_t\}$ and consider a subset $\Gamma \in \mathcal{B}(\mathbb{R}^d)$ of the state space of the process then we define the hitting time as*

$$H_\Gamma(\omega) = \inf\{t \geq 0 : X_t(\omega) \in \Gamma\}$$

Theorem 2.1.6. *H_Γ defined in Definition 2.1.5 is a stopping time.*

Lemma 2.1.7. *If T is optional and θ is a positive constant , then $T + \theta$ is a stopping time.*

Lemma 2.1.8. *If T, S are stopping time, then so are $T \wedge S, T \vee S, T + S$*

Lemma 2.1.9. *Let T, S be optional times, then $T + S$ is optional.It is a stopping time if one of the following condition holds*

1. *$T > 0, S > 0$;*
2. *$T > 0$, T is a stopping time.*

Lemma 2.1.10. *Let $\{T_n\}_{n=1}^\infty$ be a sequence of optional times; then the random times*

$$\sup_{n\geq 1} T_n, \inf_{n\geq 1} T_n, \limsup_{n\to\infty} T_n, \liminf_{n\to\infty} T_n$$

are all optional. Furthermore if the T_n's are stopping times then so are $\sup_{n\geq 1} T_n$.

Definition 2.1.11. *Let T be a stopping time of the filtration $\{\mathcal{F}_t\}$. The sigma-field $\mathcal{F}_T$ of events determined prior to the stopping time T consist of those events $A \in \mathcal{F}$ for which $A \cap \{T \leq t\} \in \mathcal{F}_t$ for every $t \geq 0$.*

Lemma 2.1.12. *$\mathcal{F}_T$ is a sigma-field and T is $\mathcal{F}_T$-measurable. Moreover if $T(\omega) = t$ for some constant $t \geq 0$ and every $\omega \in \Omega$, then $\mathcal{F}_T = \mathcal{F}_t$*

Lemma 2.1.13. *Let T be a stopping time and S a random time such that $S \geq T$ on Ω. If S is $\mathcal{F}_T$-measurable, then it is also a stopping time.*

Lemma 2.1.14. *For any two stopping times T and S, and for any $A \in \mathcal{F}_S$, we have $A \cap \{S \leq T\} \in \mathcal{F}_T$. In particular if $S \leq T$ on Ω, we have $\mathcal{F}_S \subseteq \mathcal{F}_T$*

Lemma 2.1.15. *Let T and S be stopping times. Then $\mathcal{F}_{T\wedge S} = \mathcal{F}_T \cap \mathcal{F}_S$, and each of the events*

$$\{T < S\}, \{S < T\}, \{T \leq S\}, \{S \leq T\}, \{T = S\}$$

belongs to $\mathcal{F} \cap \mathcal{F}_S$.

Lemma 2.1.16. *Let T, S be stopping times and Z an integrable random variable. We have*

1. $E\left[Z \mid \mathcal{F}_T\right] = E\left[Z \mid \mathcal{F}_{S\wedge T}\right]$, *$P$-a.s on $\{T \leq S\}$*

2. $E\left[E\left[Z \mid \mathcal{F}_T\right] \mid \mathcal{F}_S\right] = E\left[Z \mid \mathcal{F}_{S\wedge T}\right]$ *P-a.s.*

Proposition 2.1.17. *Let $X = \{X_t, \mathcal{F}_t; 0 \leq t < \infty\}$ be a progressively measurable process and let T be a stopping time of the filtration $\{\mathcal{F}_t\}$. Then the random variable X_T defined on the set $\{T < \infty\} \in \mathcal{F}_T$ by $X_T(\omega) := X_{T(\omega)}(\omega)$ is $\mathcal{F}_T$-measurable random variable., and the "stopped process" $\{X_{T\wedge t}, \mathcal{F}_t; 0 \leq t < \infty\}$ is progressively measurable.*

Lemma 2.1.18. *Under the same assumptions as in Proposition 2.1.17, and with $f(t,x) : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}$ a bounded, $\mathcal{B}\left([0,\infty)\right) \otimes \mathcal{B}(\mathbb{R}^d)$-measurable function, then the process $Y_t := \int_0^t f(s, X_s)ds; t \geq 0$ is progressively measurable with respect to $\mathcal{F}_t$ and Y_T is an $\{\mathcal{F}_T\}$-measurable random variable.*

Definition 2.1.19. *Let T be an optional time of the filtration $\{\mathcal{F}_t\}$. The sigma field $\mathcal{F}_{T+}$ of events determined immediately after the optional time T consist of those events $A \in \mathcal{F}$ for which $A \cap \{T \leq t\} \in \mathcal{F}_{t+}$ for every $t \geq 0$*

Lemma 2.1.20. *The class of sets $\mathcal{F}_{T+}$, is indeed a sigma-field with respect to which T is measurable and it coincides with $\{A \in \mathcal{F}; A \cap \{T < t\} \in \mathcal{F}_t, \forall t \geq 0\}$ and that if T is a stopping time (so both $\mathcal{F}_T$ and $\mathcal{F}_{T+}$ are defined), then $\mathcal{F}_T \subseteq \mathcal{F}_{T+}$*

Lemma 2.1.21. *The analogues of Lemmas 2.1.14 and 2.1.15 hold if T and S are assumed to be optional and $\mathcal{F}_T, \mathcal{F}_S$ and $\mathcal{F}_{T\wedge S}$ are replaced by $\mathcal{F}_{T+}, \mathcal{F}_{S+}$ and $\mathcal{F}_{(T\wedge S)+}$ respectively. Moreover if S is an optional time and T is a positive stopping time with $S \leq T$, and $S < T$ on $\{S < \infty\}$, then $\mathcal{F}_{S+} \subseteq \mathcal{F}_T$*

Lemma 2.1.22. *If $\{T_n\}_{n=1}^{\infty}$ is a sequence of optional times and $T = \inf_{n\geq 1} T_n$ then $\mathcal{F}_{T+} = \bigcap_{n=1}^{\infty} \mathcal{F}_{T_n+}$. Besides if each T_n is a positive stopping time and $T < T_n$ on $\{T < \infty\}$, then we have $\mathcal{F}_{T+} = \bigcap_{n=1}^{\infty} \mathcal{F}_{T_n}$*

Lemma 2.1.23. *Given an optional time T of the filtration $\{\mathcal{F}_t\}$, consider a sequence $\{T_n\}_{n=1}^{\infty}$ of random times given by*

$$T_n(\omega) = \begin{cases} T(\omega) & \text{on } \{\omega : T(\omega) = +\infty\} \\ \frac{k}{2^n} & \text{on} \left\{\omega : \frac{k-1}{2^n} \leq T(\omega) < \frac{k}{2^n}\right\} \end{cases}$$

for $n \geq 1, k \geq 1$. Obviously $T_n \geq T_{n+1} \geq T$, for every $n \geq 1$. Moreover T_n is a stopping time for every n and that $\lim_{n\to\infty} T_n = T$, and that for every $A \in \mathcal{F}_{T+}$, we have $A \cap \{T_n = (k/2^n)\} \in \mathcal{F}_{k/2^n}; n \geq 1, k \geq 1$.

Definition 2.1.24. *A filtration $\{\mathcal{F}_t\}$ is said to satisfy the usual conditions if it is right continuous and $\mathcal{F}_0$ contains all P-negligible events in $\mathcal{F}$.*

2.2 Introduction to Stochastic Integration

Let us consider a continuous square-integrable martingale $M = \{M_t, \mathcal{F}_t; 0 \leq t < \infty\}$ on a probability space $(\Omega, \mathcal{F}, P)$ equipped with the filtration $\{\mathcal{F}_t\}$ which is assumed throughout this chapter to satisfy the usual conditions i.e the filtration is complete and right continuous. We assume that $M_0 = 0$ a.s P. Such a process $M \in \mathcal{M}_2^c$(the space of square integrable continuous martingales) is of unbounded variation on any finite interval $[0, T]$ and consequently the integrals of the form

$$I_T(X) = \int_0^T X_t(\omega) dM_t(\omega) \tag{2.2.1}$$

cannot be defined pathwise(i.e for each ω separately) as ordinary Lebesgue-Stieltges integrals.Nevertheless , the martingale M has a finite second (quadratic) variation given by the continuous increasing process $\langle M \rangle$. It is precisely this fact which allows one to proceed in a highly non-trivial yet straightforward manner with the construction of stochastic integral (2.2.1) with respect to a continuous square-integrable martingale M for an appropriate class of integrands X. The construction is due to Itô (see [16] and [17]) for the special case that M is a Brownian motion and to Kunita and Watanabe(see [18]) for the general case.

2.2.1 Simple Processes and Approximation

In this section we will first define a class of stochastic processes(called simple processes) for which we will define the stochastic integral. These simple processes are chosen such that they are dense in L^2 and subsequently the stochastic integral will be defined for all processes in L^2 as limiting operation.

Definition 2.2.1. *Let $\mathcal{L}$ denote the set of all equivalence classes of all measurable $\{\mathcal{F}_t\}$-adapted processes X for which $[X]_T < \infty$ for all $T > 0$ where*

$$[X]_T^2 = E \int_0^T X_t^2 d\langle M \rangle_t \tag{2.2.2}$$

We define an metric on $\mathcal{L}$ by $[X-Y]$, where

$$[X] := \sum_{n=1}^{\infty} 2^{-n}(1 \wedge [X]_n) \tag{2.2.3}$$

Definition 2.2.2. *Let $\mathcal{L}^*$ denote the set of equivalence classes of progressively measurable processes satisfying $[X]_T < \infty$ for all $T > 0$, and the metric on $\mathcal{L}^*$ the same way as in Definition 2.2.1*

Definition 2.2.3. *Simple Process: A process X is called simple if there exists an increasing sequence of real numbers $\{t_n\}_{n=0}^{\infty}$ with $t_0 = 0$ and $\lim_{n\to\infty} t_n = \infty$, as well as a sequence of random variables $\{\xi_n\}_{n=1}^{\infty}$ and a non-random constant $C < \infty$ such that $\sup_{n\geq 0}|\xi_n(\omega)| \leq C$ for every $\omega \in \Omega$ such that ξ_n is $\mathcal{F}_{t_n}$-measurable for every $n \geq 0$ and*

$$X_t(\omega) = \xi_0(\omega)1_{\{0\}}(t) + \sum_{i=1}^{\infty} \xi_i(\omega)1_{(t_i,t_{i+1}]}(t),\ 0 \leq t < \infty, \omega \in \Omega$$

The class of all simple processes will be denoted by $\mathcal{L}_0$.Note that because the members of $\mathcal{L}_0$ are progressively measurable and bounded we have $\mathcal{L}_0 \subseteq \mathcal{L}^*(M) \subseteq \mathcal{L}(M)$

The stochastic integral of $X \in \mathcal{L}_0$ wr.t to the martingale $M \in \mathcal{M}_c^2$ can be defined as a martingale transform.

$$\begin{aligned} I_t(X) &:= \sum_{i=0}^{n-1} \xi_i(M_{t_i+1} - M_{t_i}) + \xi_n(M_t - M_{t_n}) \\ &= \sum_{i=1}^{\infty} \xi_i(M_{t\wedge t_{i+1}} - M_{t\wedge t_i}),\ 0 \leq t < \infty \end{aligned} \tag{2.2.4}$$

where $n \geq 0$ is the unique integer for which $t_n \leq t < t_{n+1}$. The definition is then extended to integrands $X \in \mathcal{L}^*$ and $X \in \mathcal{L}$, thanks to the crucial results which show that the elements of $\mathcal{L}$ and $\mathcal{L}^*$ can be approximated in a suitable sense by simple processes. Before proceeding to the next lemma we define $\mathcal{L}_T^*$ to be the class of processes X in $\mathcal{L}^*$ for which $X_t(\omega) = 0, \forall t > T, \omega \in \Omega$. For $T = \infty$, $\mathcal{L}_T^*$ is defined as the class of processes $X \in \mathcal{L}^*$ for which $E[\int_0^t X_t^2 d\langle M\rangle_t] < \infty$ (a condition we already have for $T < \infty$ by virtue of its membership of $\mathcal{L}^*$). Note that a process $X \in \mathcal{L}_T^*$ can only be identified with one defined for $(t,\omega) \in [0,T] \times \Omega$.

Lemma 2.2.4. *(Lemma 3.2.2 Shreve) For $0 < T \leq \infty$, $\mathcal{L}_T^*$ is a closed subspace of $\mathcal{H}_T$. In particular $\mathcal{L}_T^*$ is complete under the norm $[X]_T$. $\mathcal{H}_T$ is defined as*

$$\mathcal{H}_T = L^2\left([0,T] \times \Omega, \mathcal{B}([0,T]) \otimes \mathcal{F}_T, \mu_M\right)$$

where μ_M *is a measure on* $([0,\infty)\times\Omega, \mathcal{B}([0,\infty))\otimes\mathcal{F})$ *given by*

$$\mu_M(A) = E\int_0^\infty 1_A(t,\omega)d(\langle M\rangle_t)(\omega)$$

Proof. Let $\{X^{(n)}\}_{n=1}^\infty$ be a convergent sequence in $\mathcal{L}_T^*$ with the limit $X \in \mathcal{H}_T$(Since $X \in \mathcal{H}_T \supseteq \mathcal{L}_T^*$ and $\mathcal{H}_T$ is a Hilbert space and every convergent sequence has a limit). The limit is with respect to the norm $L^2([0,T]\times\Omega, \mathcal{B}([0,T])\otimes\mathcal{F}_T, \mu_M)$. So $\{X^{(n)}\}_{n=1}^\infty$ converges in probability(wr.t the measure μ_M) and therefore there exists a sub-sequence which converges almost surely i.e

$$\mu_M\{(t,\omega)\in[0,T]\times\Omega; \lim_{n\to\infty} X_t^{(n)}(\omega) \neq X_t(\omega)\} = 0$$

In order to show that $X \in \mathcal{L}_T^*$, we need to show that its progressively measurable. By the virtue of its membership in $\mathcal{H}_T$, it is $\mathcal{B}([0,T])\otimes\mathcal{F}$-measurable Now setting

$$A = \{(t,\omega)\in[0,T]\times\Omega; \lim_{n\to\infty} X_t^{(n)}(\omega) \text{ exists in } \mathbb{R}\}$$

Since the limit of the sequence $\{X^{(n)}\}_{n=1}^\infty$ is in $\mathcal{H}_T$, we have that the L^2 norm of X is finite and therefore $X < \infty$ μ_M a.s and hence the measure of the set A is one. Now the process $Y_t(\omega) = 1_A \lim_{n\to\infty} X_t^{(n)}(\omega) + 01_{A^c}$ inherits the progressive measurability since its the product of the indicator of a progressive set A and the limit of progressively measurable processes $X^{(n)}$. □

Lemma 2.2.5. *(Lemma 3.2.4 Shreve) Let X be a bounded , measurable ,* $\{\mathcal{F}_t\}$*- adapted process. Then there exists a sequence* $\{X^{(m)}\}_{m=1}^\infty$ *of simple processes such that*

$$\sup_{T>0}\lim_{m\to\infty} E\int_0^T |X_t^{(m)} - X_t|^2 dt = 0 \tag{2.2.5}$$

Proof. We shall show how to construct for each fixed $T > 0$, a sequence $\{X^{(n,T)}\}_{n=1}^\infty$ of simple processes so that

$$\lim_{n\to\infty} E\int_0^T |X_t^{(n,T)} - X_t|^2 dt = 0$$

Thus for each positive integer m, there is another integer n_m such that

$$E\int_0^m |X_t^{(n_m,m)} - X_t|^2 dt \leq \frac{1}{m}$$

and hence the sequence $\{X^{(n_m,m)}\}_{m=1}^\infty$ has the desired properties since the integrand in the last expression is positive and therefore

$$0 \leq \sup_{T>0}\lim_{m\to\infty} E\int_0^T |X_t^{(m)} - X_t|^2 dt \leq \lim_{m\to\infty} E\int_0^m |X_t^{(n_m,m)} - X_t|^2 dt = 0$$

Henceforth , T is a fixed positive number . We proceed in three steps.

1. Suppose that X is continuous, then the sequence of simple processes

$$X_t^{(n)}(\omega) = X_0(\omega)1_{\{0\}}(t) + \sum_{k=0}^{2^n-1} X_{kT/2^n}(\omega)1_{(kT/2^n,(k+1)T/2^n]}(t)$$

It is obvious by the definition of $X_t^{(n)}$(it takes the same value as that of X_t on intervals of the form $(kT/2^n, (k+1)T/2^n]$ which become smaller and smaller and hence $X_t^{(n)}$ approximates X_t) that

$$\lim_{n\to\infty} X_t^{(n)} = X_t \text{ a.s}$$

And since X is bounded(by assumption) so is $X^{(n)}$ by construction and hence $|X_t^n - X_t|^2 \leq C$ and therefore the bounded convergence theorem and almost sure convergence yields.

$$\lim_{n\to\infty} E \int_0^T |X_t^{(n)} - X_t|^2 dt = E \int_0^T \lim_{n\to\infty} |X_t^{(n)} - X_t|^2 dt = 0$$

2. Now suppose that X is progressively measurable; we consider the continuous progressively measurable processes

$$F_t(\omega) := \int_0^{t\wedge T} X_s(\omega)ds; \tilde{X}_t^{(m)} := m[F_t - F_{(t-1/m)\wedge 0}]; m \geq 1 \qquad (2.2.6)$$

for $t \geq 0, \omega \in \Omega$. Since X is bounded and progressively measurable(hence it is jointly measurable and hence for a fixed ω, it is Lebesgue measurable) , the Lebesgue integral is well defined and $F_t(\omega)$ is absolutely continuous and by the Fundamental theorem of Lebesgue integral calculus , differentiable almost everywhere(with respect to the variable t) with the derivative being equal to $X_t(\omega)$. Since F_t is absolutely continuous, $\tilde{X}_t^{(m)}$ is absolutely continuous and hence continuous, and therefore by virtue of step 1) we can conclude the existence of a sequence of simple processes $\{\tilde{X}^{(m,n)}\}_{n=1}^{\infty}$ such that $\lim_{m\to\infty} E \int_0^T |\tilde{X}_t^{(m,n)} - \tilde{X}_t^{(m)}|^2 dt = 0$

Consider the set

$$A := \{(t,\omega) \in [0,T] \times \Omega; \lim_{m\to\infty} \tilde{X}_t^{(m)}(\omega) = X_t(\omega)\}^c$$

Clearly A is in $\mathcal{B}([0,T])\otimes\mathcal{F}_T$. Indeed X_t is progressive and therefore $\mathcal{B}([0,T])\otimes\mathcal{F}_T$-measurable and $\tilde{X}_t^{(m)}$ is continuous and adapted and hence progressive and so is it limit and therefore their difference is $\mathcal{B}([0,T]) \otimes \mathcal{F}_T$- measurable and hence

$$A^c = (\lim_{m\to\infty} (\tilde{X}_t^{(m)}(\omega) - X_t(\omega))^{-1}(0) \in \mathcal{B}([0,T]) \otimes \mathcal{F}_T$$

The bounded convergence theorem implies

$$\lim_{m\to\infty} E\int_0^T |\tilde{X}_t^{(m)} - X_t|^2 dt = E\int_0^T \lim_{m\to\infty} |\tilde{X}_t^{(m)} - X_t|^2 dt = 0$$

Indeed we have that

$$\begin{aligned} |\tilde{X}_t^{(m)} - X_t| &= m\left|\int_{t-1/m}^t (X_s - X_t)ds\right| \\ \le m\int_{t-1/m}^t |X_s - X_t| ds &\le m\int_{t-1/m}^t 2C ds \le 2C \end{aligned} \tag{2.2.7}$$

where C is such that $X_t(\omega) \le C\ \forall t \in [0,T]$ and $\omega \in \Omega$

We can approximate the continuous process $\tilde{X}^m$ by a sequence $\{\tilde{X}^{m,n}\}_{n=1}^\infty$ w.r.t to the L^2 norm which we call for the sake of convenience $\|.\|$. Similarly we can approximate X by $\tilde{X}^{(m)}$ w.r.t to the same norm and therefore we can conclude that there exists a sequence of bounded simple processes such that

$$\lim_{m\to\infty} E\int_0^T |\tilde{X}_t^{(m,n_m)} - \tilde{X}_t|^2 dt := \left\|\tilde{X}_t^{(m,n_m)} - \tilde{X}_t\right\| = 0$$

Indeed we have that given $\epsilon > 0, \forall n \ge N(\epsilon)$ for some $N(\epsilon) \in \mathbb{N}$ we have

$$\left\|\tilde{X}^{(n)} - X\right\| < \epsilon/2$$

. Similarly we have that for all $n \in \mathbb{N}$ there exists $m_n \in \mathbb{N}$ such that

$$\left\|\tilde{X}^{(n,m_n)} - X^{(n)}\right\| < \epsilon/2$$

. And hence by the triangular inequality we have that

$$\left\|\tilde{X}^{(n,m_n)} - X\right\| \le \left\|\tilde{X}^{(n,m_n)} - \tilde{X}^{(n)}\right\| + \left\|\tilde{X}^{(n)} - X\right\| < \epsilon$$

for all $n \ge N(\epsilon)$

Note that the assumption of progressive measurability is necessary in this step to claim the existence of a sequence of approximating simple processes since progressivity of X implies the adaptedness of the process F and hence $\tilde{X}^m$ which is essential in order to use part 1 of the lemma to show the existence of approximating simple processes which are adapted(a requirement for a process to be simple).

3. Finally let X be measurable and adapted. We cannot guarantee immediately that the continuous process $F = \{F_t; 0 \le t < \infty\}$ is progressive.We do however know that any measurable and adapted process has a progressively measurable modification Y(Proposition 1.1.12 Karatzas and Shreve's Brownian Motion and Stochastic Calculus). We now show that progressive measurable process (proof just like before) $\{G_t := \int_0^{t\wedge T} Y_s ds; 0 \le t \le T\}$ is a modification of F. For the measurable process $\eta_t(\omega) = 1_{\{X_t(\omega)\neq Y_t(\omega)\}}, 0 \le t \le T, \omega \in \Omega$, we have from Fubini: $E\int_0^T \eta_t(\omega)dt = \int_0^T E\eta_t(\omega)dt = \int_0^T P(X_t(\omega) \neq Y_t(\omega)) = \int_0^T 0dt = 0$ where the second last equality follows from the fact that X is a modification of Y. Now this implies that $\int_0^T \eta_t(\omega)dt = 0$ P-a.e $\omega \in \Omega$.

   This implies that the event $\{\omega \in \Omega : \int_0^T \eta_t(\omega)dt > 0\}$ is a measure zero set which contains $\{F_t \neq G_t\}$ .Indeed we have that $\{F_t \neq G_t\} = \{\int_0^{t\wedge T} |X_s - Y_s| 1_{\{X_s(\omega)\neq Y_s(\omega)\}} > 0\} \subseteq \{\int_0^T 1_{\{X_s(\omega)\neq Y_s(\omega)\}} ds > 0\}$

   Now since G_t is $\mathcal{F}_t$- measurable and $\mathcal{F}_t$ contains all P-null sets we have that F_t is also $\mathcal{F}_t$- measurable(since we can add and subtract subsets of the null set $\{F_t \neq G_t\}$ from G_t to get F_t).Now adaptivity and continuity of F implies progressivity and we can now repeat the argument in step 2).

□

Lemma 2.2.6. *(Proposition 3.2.6 Shreve) If the function $t \mapsto \langle M\rangle_t(\omega)$ is absolutely continuous with respect to Lebesgue measure for $P-$a.e $\omega \in \Omega$, then $\mathcal{L}_0$ is dense in $\mathcal{L}$ with respect to the metric*

$$d(x, y) = \sum_{n=1}^{\infty} 2^{-n} (1 \wedge [x - y]_n)$$

where

$$[X]_n^2 = E \int_0^n X_u^2 d\langle M\rangle_u$$

Proof. (a) If $X \in \mathcal{L}$ is bounded then Lemma 2.2.5 guarantees the existence of a bounded sequence $\{X^{(m)}\}$ of simple processes satisfying

$$\sup_{T>0} \lim_{m\to\infty} E \int_0^T |X_t^{(m)} - X_t|^2 dt = 0$$

From it we can extract a subsequence $\{X^{(m_k)}\}$ such that the set

$$\{(t, \omega) \in [0, \infty) \times \Omega : \lim_{k\to\infty} X_t^{(m_k)}(\omega) = X_t(\omega)\}^c$$

has product measure zero. Now the absolute continuity of $t \mapsto \langle M\rangle_t(\omega)$ with respect to the lebesgue measure implies the existence of a density function which is defined almost everywhere such that $d\langle M\rangle_t(\omega) = \langle M\rangle'(\omega)dt$. Now the bounded convergence theorem implies that we can take the limit inside the integral.

$$\lim_{k\to\infty}[X^{(m_k)} - X] = \sum_{n=1}^{\infty} 2^{-n}\left(1 \wedge \lim_{k\to\infty}\left(E\int_0^n (X_t^{(m_k)} - X_t)^2 \langle M\rangle'(\omega)dt\right)\right) = 0 \tag{2.2.8}$$

(b) If $X \in \mathcal{L}$ is not necessarily bounded we define

$$X_t^{(n)}(\omega) := X_t(\omega)1_{\{|X_t(\omega)|\leq n\}};\ 0 \leq t < \infty, \omega \in \Omega$$

and thereby obtain a sequence of bounded processes in $\mathcal{L}$. The dominated convergence theorem implies

$$[X^{(n)} - X]_T^2 = E\int_0^T X_t^2 1_{\{|X_t|>n\}} d\langle M\rangle_t \to 0 \text{ as } n \to \infty$$

for every $T > 0$ whence $\lim_{n\to\infty}[X^{(n)} - X] = 0$. Each $X^{(n)}$ can be approximated by bounded simple processes , so X can be as well. Indeed its enough to prove this for sequences in $\mathbb{R}$ since the exact same argument follows for sequences in the given norm.

So let us assume that $(x_n)_{n\in\mathbb{N}}$ be a real sequence such that $x_n \to x$ as $n \to \infty$ and $\forall n \geq 1$ be a sequence $\{x_n^m\}_{m\in\mathbb{N}}$ such that $x_n^m \to x_n$ as $m \to \infty$ for all $n \geq 1$. Then $\forall \epsilon > 0, \exists n_\epsilon \in \mathbb{N}$ such that $\forall n \geq n_\epsilon$ we have $|x_n - x| < \epsilon/2$. We also have that $\forall n \geq 1, \exists m_n \in \mathbb{N}$ such that $|x_n^{m_n} - x_n| < \frac{1}{n}$ due to the second assumption. Then $\forall \epsilon > 0, \exists \hat{n}_\epsilon \in \mathbb{N}$ such that $\forall n \geq \hat{n}_\epsilon$ we have that

$$|x_n^{m_n} - x| \leq |x_n^{m_n} - x_n| + |x_n - x| < \epsilon/2 + 1/n < \epsilon$$

if we choose $\hat{n}_\epsilon = \max(n_\epsilon, \lfloor \frac{2}{\epsilon} \rfloor)$

□

Definition 2.2.7. *An adapted process A is called increasing if for P-a.e. $\omega \in \Omega$ we have*

1. $A_0(\omega) = 0$

2. *$t \mapsto A_t(\omega)$ is a nondecreasing , right continuous function, and $E(A_t) < \infty$ holds for every $t \in [0, \infty)$. An increasing process is called integrable if $E(A_\infty) < \infty$, where $A_\infty = \lim_{t\to\infty} A_t$*

Lemma 2.2.8. *(Lemma 3.2.7 Shreve)*

Let $\{A_t; 0 \leq t < \infty\}$ *be a continuous increasing process(usual definition) adapted to the filtration of the martingale* $M = \{M_t, \mathcal{F}_t, 0 \leq t < \infty\}$*. If* $X = \{X_t, \mathcal{F}_t, 0 \leq t < \infty\}$ *is a progressively measurable process satisfying*

$$E \int_0^T X_t^2 dA_t < \infty$$

for each $T > 0$*, then there exists a sequence* $\{X^{(n)}\}_{n=1}^{\infty}$ *of simple processes such that*

$$\sup_{T>0} \lim_{n\to\infty} E \int_0^T |X_t^{(n)} - X_t|^2 dA_t = 0$$

Proof. We may assume without loss of generality that X is bounded (if not we use the same argumentation as in part (b) of Lemma 2.2.6) i.e

$$|X_t(\omega)| \leq C; \;\; \forall t \geq 0, \omega \in \Omega \tag{2.2.9}$$

As in the proof of Lemma 2.2.5, it suffices to show how to construct , for each fixed $T > 0$, a sequence $\{X^{(n)}\}_{n=1}^{\infty}$ of simple processes for which

$$\lim_{n\to\infty} E \int_0^T |X_t^{(n)} - X_t|^2 dA_t = 0 \tag{2.2.10}$$

Henceforth $T > 0$ is fixed, and we assume without loss of generality(as the integral above doesn't change) that

$$X_t(\omega) = 0; \;\; \forall t > T, \omega \in \Omega. \tag{2.2.11}$$

Now we describe the time change. Since $A_t(\omega) + t$ is strictly increasing in $t \geq 0$ for P-a.e ω, there exists a continuous strictly increasing inverse function $T_s(\omega)$, defined for $s \geq 0$ such that

$$A_{T_s(\omega)}(\omega) + T_s(\omega) = s; \;\; \forall s \geq 0 \tag{2.2.12}$$

In particular we have $T_s \leq s$ since from the equation just above $T_s(\omega) = s - A_{T_s(\omega)}(\omega)$ and A_t is an increasing process(and hence by Definition 2.2.7 always non-negative). Its not very hard to see that

$$\{T_s \leq t\} = \{A_t + t \geq s\} \in \mathcal{F}_t \tag{2.2.13}$$

Indeed we have that $\{T_s \leq t\} \subseteq \{A_t + t \geq s\}$ since $A_t + t$ is strictly increasing and hence $\omega \in \{T_s(\omega) \leq t\} \subseteq \{A_{T_s(\omega)}(\omega) + T_s(\omega) \leq A_t + t\} = \{s \leq A_t + t\}$ where the last equality is a consequence of equation (2.2.12)

On the other hand $\{A_t + t \geq s\} \subseteq \{T_s \leq t\}$ is obvious. Thus for each $s \geq 0$, T_s is a bounded stopping time for $\{\mathcal{F}_t\}$. Taking s as our new time variable we define the filtration $\{\mathcal{G}_s\}$ by

$$\mathcal{G}_s = \mathcal{F}_{T_s};\ \ s \geq 0$$

and introduce the time changed process

$$Y_s(\omega) = X_{T_s(\omega)}(\omega);\ \ s \geq 0, \omega \in \Omega$$

which is adapted to $\mathcal{G}_s$ because the progressive measurability of X(Lemma 1.2.18 in [12]). On the other hand Lemma 2.2.5 implies that, given any $\epsilon > 0$ and $R > 0$, there is a simple process $\{Y_s^\epsilon, \mathcal{G}_s, 0 \leq s < \infty\}$ for which

$$E\int_0^R |Y_s^\epsilon - Y_s|^2 ds \leq \epsilon/2 \tag{2.2.14}$$

But from equation (2.2.9) and (2.2.10) it follows that

$$\begin{aligned} E\int_0^\infty Y_s^2 ds = E\int_0^\infty 1_{\{T_s \leq T\}} X_{T_s}^2 ds = E\int_0^{A_T+T} X_{T_s}^2 ds \\ \leq C^2(EA_T + T) < \infty \end{aligned} \tag{2.2.15}$$

where the first equality follows from the definition of of Y_s and the fact that we assume $X_t(\omega) = 0$ for all $t > T$ and the second inequality follows from equation (2.2.13). So now by choosing R to be sufficiently large and setting $Y_s^\epsilon = 0$ for $s > R$ we get

$$\begin{aligned} E\int_0^\infty |Y_s^\epsilon - Y_s|^2 ds = E\int_0^R |Y_s^\epsilon - Y_s|^2 ds + E\int_R^\infty |Y_s^\epsilon - Y_s|^2 ds \\ = E\int_0^R |Y_s^\epsilon - Y_s|^2 ds + E\int_R^\infty |Y_s|^2 ds < \epsilon/2 + \epsilon/2 = \epsilon \end{aligned} \tag{2.2.16}$$

where the the inequality follows from equation (2.2.15) and (2.2.14).(First we choose an R large enough and we already know that equation (2.2.15)) is true for all $R > 0$

Now since Y_s^ϵ is simple and because it vanishes for $s > R$, there is a finite partition $0 = s_0 < s_1 < \cdots s_n \leq R$ with

$$Y_s^\epsilon = \xi_0(\omega)1_{\{0\}}(s) + \sum_{j=1}^n \xi_{s_{j-1}}(\omega)1_{(s_{j-1},s_j]}(s),\ \ 0 \leq s < \infty$$

where each ξ_{s_j} is measurable with respect to $\mathcal{G}_{s_j} = \mathcal{F}_{T_{s_j}}$ and bounded in absolute value by a constant K. Now reverting to the original clock we observe that

$$X_t^\epsilon := Y_{t+A_t}^\epsilon = \xi_0(\omega)1_{\{0\}}(t + A_t) + \sum_{j=1}^n \xi_{s_{j-1}}(\omega)1_{(s_{j-1},s_j]}(t + A_t),\ \ 0 \leq t < \infty$$

Also note that using the same argumentation as before(using (2.2.13)) we an show that $s_{j-1} < t + A_t \le s_j$ is the same as $T_{s_{j-1}} < t \le T_{s_j}$ and hence rewrite X_t^ϵ as follows

$$X_t^\epsilon = \xi_0(\omega)1_{\{0\}}(t) + \sum_{j=1}^{n} \xi_{s_{j-1}}(\omega)1_{(T_{s_{j-1}},T_{s_j}]}(t),\ 0 \le t < \infty$$

Now recalling that a random variable Z is $\mathcal{F}_T$ measurable if and only if $Z1_{\{T\le t\}}$ is a $\mathcal{F}_t$ measurable random variable for all $t \ge 0$ Now since ξ_{s_j} is $\mathcal{F}_{T_{s_j}}$ we can conclude that $\xi_{s_j}1_{\{T_{s_j}\le t\}}$ is $\mathcal{F}_t$ measurable and therefore we have that $\xi_{s_{j-1}}1_{\{T_{s_{j-1}}\le t\}}1_{\{T_{s_{j-1}}\ne t\}}1_{\{T_{s_{j-1}}\ge t\}} = \xi_{s_{j-1}}1_{(T_{s_{j-1}},T_{s_j}]} \in \mathcal{F}_t$ as the L.H.S is a product of $\mathcal{F}_t$ measurable functions since T_{s_j} is a $\mathcal{F}_t$ stopping time. We have

$$\begin{aligned} E\int_0^T |X_t^\epsilon - X_t|^2 dA_t &\le E\int_0^T |X_t^\epsilon - X_t|^2 d(A_t + t) \\ = E\int_0^T |Y_{A_t+t}^\epsilon - Y_{A_t+t}|^2 d(A_t+t) &= E\int_0^T |Y_s^\epsilon - Y_s|^2 ds \\ &\le E\int_0^\infty |Y_s^\epsilon - Y_s|^2 ds < \epsilon \end{aligned}$$

In order to complete the proof we need to show that X^ϵ is a simple process. For this we refer the reader to [12].

□

Proposition 2.2.9. *(Proposition 3.2.8 Shreve) The set $\mathcal{L}_0$ of simple processes is dense in $\mathcal{L}^*$ with respect to the metric of Definition* 2.2.1

Proof. Take $A = \langle M \rangle$ in Lemma 2.2.8 □

In the next section we will give the most important properties of the Stochastic integral, many of which are used in the theory of stochastic differential equations. The presentation style and the results are from [12].

2.2.2 Construction and Elementary Properties of the Stochastic Integral

We have already defined the stochastic integral of a simple process $X \in \mathcal{L}_0$. Let us list certain properties of the integral : for $X, Y \in \mathcal{L}_0$ and $0 \le s < t < \infty$ we have

$$I_0(X) = 0,\ \text{a.s. } P \tag{2.2.17}$$

$$E\left[I_t(X)|\mathcal{F}_s\right] = I_s(X),\ \text{a.s. } P \tag{2.2.18}$$

$$E\left(I_t(X)\right)^2 = E\int_0^t X_u^2 d\langle M\rangle_u \tag{2.2.19}$$

$$\|I(X)\| = [X] \tag{2.2.20}$$

$$E\left[\left(I_t(X) - I_s(X)\right)^2 |\mathcal{F}_s\right] = E\left[\int_s^t X_u^2 d\langle M\rangle_u |\mathcal{F}_s\right] \tag{2.2.21}$$

$$I(\alpha X + \beta Y) = \alpha I(X) + \beta I(Y);\ \alpha, \beta \in \mathbb{R} \tag{2.2.22}$$

Properties (2.2.17) and (2.2.22) are obvious for simple integrands and follows directly from the definition of stochastic integral of simple integrands. Property 2.2.18 follows from the fact that for any $0 \leq s < t < \infty$ and any integer $i \geq 1$, we have in the notation of (2.2.4),

$$E\left[\xi_i(M_{t\wedge t_{i+1}} - M_{t\wedge t_i})|\mathcal{F}_s\right] = \xi_i(M_{s\wedge t_{i+1}} - M_{s\wedge t_i}),\ \text{a.s } P$$

which can be verified easily using the properties of conditional expectation for each of the the three cases: $s \leq t_i$, $t_i < s \leq t_{i+1}$ and $t_{i+1} < s$

For example when $s < t_i$ we have

$$\begin{aligned} E\left[E\left[\xi_i(M_{t\wedge t_{i+1}} - M_{t\wedge t_i})|\mathcal{F}_{t_i}\right]|\mathcal{F}_s\right] &= E\left[\xi_i E\left[(M_{t\wedge t_{i+1}} - M_{t\wedge t_i})|\mathcal{F}_{t_i}\right]|\mathcal{F}_s\right] \\ &= E\left[\xi_i E\left[(M_{t_{i+1}} - M_{t_i})|\mathcal{F}_{t_i}\right]|\mathcal{F}_s\right] = 0 = \xi_i(M_{s\wedge t_{i+1}} - M_{s\wedge t_i}) \end{aligned}$$

The other two cases can be proved similarly.

Moreover it follows from construction of the stochastic integral of a simple process as a martingale transform that it is continuous and hence $I(X) = \{I_t(X), \mathcal{F}_t, 0 \leq t < \infty\}$ is a continuous martingale. Now with $0 \leq s < t < \infty$ and m and n chosen such that $t_{m-1} \leq s < t_m$ and $t_n \leq t < t_{n+1}$, we have the following

$$\begin{aligned} &E\left[\left(I_t(X) - I_s(X)\right)^2 \mid \mathcal{F}_s\right] \\ &= E\left[\left(\xi_{m-1}(M_{t_m} - M_s) + \sum_{i=m}^{n-1}\xi_i(M_{t_{i+1}} - M_{t_i}) + \xi_n(M_t - M_{t_n})\right)^2 |\mathcal{F}_s\right] \\ &= E\left[\xi_{m-1}^2(M_{t_m} - M_s)^2 + \sum_{i=m}^{n-1}\xi_i^2(M_{t_{i+1}} - M_{t_i})^2 + \xi_n^2(M_t - M_{t_n})^2|\mathcal{F}_s\right] \\ &= E\left[\xi_{m-1}^2(M_{t_m}^2 - M_s^2) + \sum_{i=m}^{n-1}\xi_i^2(M_{t_{i+1}}^2 - M_{t_i}^2) + \xi_n^2(M_t^2 - M_{t_n}^2)|\mathcal{F}_s\right] \\ &= E\left[\xi_{m-1}^2\left(\langle M\rangle_{t_m} - \langle M\rangle_s\right) + \sum_{i=m}^{n-1}\xi_i^2\left(\langle M\rangle_{t_{i+1}} - \langle M\rangle_{t_i}\right) + \xi_n^2\left(\langle M\rangle_t - \langle M\rangle_{t_n}\right)|\mathcal{F}_s\right] \\ &= E\left[\int_s^t X_u^2 d\langle M\rangle_u|\mathcal{F}_s\right] \end{aligned}$$

Where we have used the following properties of square integrable martingales for $0 \leq t \leq u < v$

1. $E\left[(M_u - M_v)^2|\mathcal{F}_t\right] = E\left[M_u^2 - M_v^2|\mathcal{F}_t\right]$
2. $E\left[M_u^2 - M_v^2|\mathcal{F}_t\right] = E\left[\langle M\rangle_u - \langle M\rangle_v|\mathcal{F}_t\right]$

This proves (2.2.21) and establishes the fact that the continuous martingale is square integrable (by putting $s = 0$ in (2.2.21) and then taking expectations and recalling that the quadratic variation is a continuous process and hence bounded and therefore integrable on all compact sets)

And its quadratic variation is given by

$$\langle I(X)\rangle_t = \int_0^t X_u^2 d\langle M\rangle_u$$

because quadratic variation of $I(X)$ is the unique (up to indistinguishability) stochastic process $\langle I(X)\rangle$ such that

$$\left(I(X)_t^2 - \langle I(X)\rangle_t\right)_{t\geq 0} \text{ is a martingale}$$

and since

$E\left[I(X)_t^2 - I(X)_s^2 - \int_s^t X_u^2 d\langle M\rangle_u|\mathcal{F}_s\right] = 0$, the result follows.

Lemma 2.2.10. *Let $W = \{W_t, \mathcal{F}_t; 0 \leq t < \infty\}$ be a standard , one-dimensional Brownian Motion , and let T be a stopping time of $\{\mathcal{F}_t\}$ with $ET < \infty$. Then the following Wald's identities hold*

$$E(W_T) = 0 \text{ and } E(W_T^2) = E[T]$$

Proof. Let T be a stopping time with respect to the filtration $\{\mathcal{F}_t\}$. For a fixed $0 \leq t < \infty$,$t \wedge T$ is also a $\mathcal{F}_t$-stopping time(elementary fact). Since W is progressively measurable(sample paths of Brownian motion are continuous almost surely and adapted), $W_{t\wedge T}$ is $\mathcal{F}_{t\wedge T}$-msb and hence $\mathcal{F}_t$-msb.

Since $(W_t^2 - t)_{t\geq 0}$ is a martingale, it follows from the optional stopping theorem that

$$\mathbb{E}(W_{T\wedge n}^2) = \mathbb{E}(T \wedge n).$$

This implies

$$\mathbb{E}((W_{T\wedge n} - W_{T\wedge m})^2) = \mathbb{E}(T \wedge n) - \mathbb{E}(T \wedge m) \xrightarrow{m,n\to\infty} 0.$$

This shows that $(W_{T\wedge n})_{n\geq 1}$ is an L^2-Cauchy sequence and so $W_{T\wedge n} \to W_T$ in L^2. Hence, in particular, $W_{T\wedge n} \to W_T$ in L^1 and so

$$\mathbb{E}(W_{T\wedge n}) \xrightarrow{n\to\infty} \mathbb{E}(W_T).$$

Obviously taking $n \to \infty$ we get $\lim_{n\to\infty} E\left[W^2_{T\wedge n}\right] = \lim_{n\to\infty} E\left[T\wedge n\right] = E\left[T\right]$ where the last inequality follows from the monotone convergence theorem. Note that even thought we show this for $n \in \mathbb{N}$, the same proof works if we chose a sequence $(t_n)_{n\in\mathbb{N}}$ such that $t_n \to \infty$ as $n \to \infty$ And therefore we can conclude

$$E\left[W_T^2\right] = E[T]$$

Also by the optional sampling theorem $E\left[W_{t\wedge T}\right] = E\left[W_0\right] = 0$ and L^2 convergence implies convergence in L_1 we have $\lim_{t\to\infty} E\left[W_{t\wedge T}\right] = E\left[W_T\right]$ which yields

$$E\left[W_T\right] = 0$$

□

Lemma 2.2.11. *Let $W = \{W_t, \mathcal{F}_t; 0 \leq t < \infty\}$ be a standard , one-dimensional Brownian Motion, let b be a real number and let T_b be the first passage time to b. Show that for $b \neq 0$ we have that $E[T_b] = \infty$*

Proof. (Proof by contradiction) Recall that the Brownian passage time T_b is defined as

$$T_b = \inf\{t \geq 0 : W_t = b\}$$

If $E\left[T_b\right] < \infty$ then by Lemma 2.2.10(Wald's identity) $E\left[W_{T_b}\right] = 0$ but by the definition of T_b we have that $E\left[W_{T_b}\right] = E[b] > 0$(Contradiction!!!). And hence we have to have that $E\left[T_b\right] = \infty$ □

2.2.3 Characterization of the Integral

Suppose that $M = \{M_t, \mathcal{F}_t, 0 \leq t < \infty\}$ and $N = \{N_t, \mathcal{F}_t, 0 \leq t < \infty\}$ are in $\mathcal{M}_2^c$, and take $X \in \mathcal{L}^*(M), Y \in \mathcal{L}^*(N)$. Then we will show that $I_t^M(X) := \int_0^t X_s dM_s, I_t^N(Y) := \int_0^t Y_s dN_s$ are also in $\mathcal{M}_2^c$. We have already seen that

$$\langle I^M(X)\rangle_t = \int_0^t X_u^2 d\langle M\rangle_u,\ \langle I^N(Y)\rangle_t = \int_0^t Y_u^2 d\langle N\rangle_u$$

We now propose to now establish the cross variational formula

$$\langle I^M(X), I^N(Y)\rangle_t = \int_0^t X_u Y_u d\langle M, N\rangle_u;\ t \geq 0, P \text{ a.s} \tag{2.2.23}$$

If X and Y is simple its is straightforward to show that that for $0 \leq s < t < \infty$ we have

$$E\left[\left(I_t^M(X) - I_s^M(X)\right)\left(I_t^N(Y) - I_t^N(Y)\right)|\mathcal{F}_s\right] = E\left[\int_s^t X_u Y_u d\langle M, N\rangle_u | \mathcal{F}_s\right]\ P \text{ a.s} \tag{2.2.24}$$

(2.2.24) is equivalent too (2.2.23) . We first show that (2.2.24) $\implies$ (2.2.23).

$$
\begin{aligned}
&E\left[\left(I_t^M(X) - I_s^M(X)\right)\left(I_t^N(Y) - I_s^N(Y)\right)|\mathcal{F}_s\right] \\
&= E\left[I_t^M(X)I_t^N(Y) - I_t^M(X)I_s^N(Y) - I_s^M(X)I_t^N(Y) + I_s^N(Y)I_s^M(X)|\mathcal{F}_s\right] \\
&= E\left[I_t^M(X)I_t^N(Y) - I_t^M(X)I_s^N(Y) - I_s^M(X)I_t^N(Y) + I_s^N(Y)I_s^M(X)|\mathcal{F}_s\right] = \\
&E\left[I_t^M(X)I_t^N(Y)|\mathcal{F}_s\right] - I_s^N(Y)E\left[I_t^M(X)|\mathcal{F}_s\right] - I_s^M(X)E\left[I_t^N(Y)|\mathcal{F}_s\right] + I_s^N(Y)I_s^M(X) = \\
&E\left[I_t^M(X)I_t^N(Y)|\mathcal{F}_s\right] - I_s^N(Y)I_s^M(X) - I_s^M(X)I_t^N(Y) + I_s^N(Y)I_s^M(X) \\
&= E\left[I_t^M(X)I_t^N(Y) - I_s^M(X)I_s^N(Y)|\mathcal{F}_s\right]
\end{aligned}
\tag{2.2.25}
$$

Therefore from (2.2.24) we have

$$E\left[I_t^M(X)I_t^N(Y) - I_s^M(X)I_s^N(Y) - \int_s^t X_u Y_u d\langle M,N\rangle_u|\mathcal{F}_s\right] = 0$$

Now if we assume that

$$\langle I^M(X), I^N(Y)\rangle_t = \int_0^t X_u Y_u d\langle M,N\rangle_u;\ t \geq 0, P \text{ a.s}$$

then recalling that the quadratic variation $\langle I^M(X), I^N(Y)\rangle_t$ is the unique (upto indistinguishability) process which makes $(I_t^M(X), I_t^N(Y) - \langle I^M(X), I^N(Y)\rangle_t)_{t\geq 0}$ a martingale, we can conclude that (2.2.24) is equivalent too (2.2.23)

Now it remains to extend the result to the case when $X \in \mathcal{L}^*(M), Y \in \mathcal{L}^*(N)$. In order to do it we will need the following propositions. The following result is due to Kunita and Watanabe(see [18])

Proposition 2.2.12. *(Proposition 3.2.14 Shreve) If $M, N \in \mathcal{M}_2^c, X \in \mathcal{L}^*(M), Y \in \mathcal{L}^*(N)$ then a.s*

$$\int_0^t |X_s Y_s| d\xi_s \leq \left(\int_0^t X_s^2 d\langle M\rangle_s\right)^{1/2}\left(\int_0^t Y_s^2 d\langle N\rangle_s\right)^{1/2} \quad 0 \leq t < \infty$$

where $\hat{\xi}_s$ denotes the total variation of the process $\xi = \langle M,N\rangle$ on $[0,s]$

Proof. According to problem 1.5.7(iv), on page 31-32 in [12], $\hat{\xi}(\omega)$ is absolutely continuous with respect to $\psi(\omega) = \frac{1}{2}\left[\langle M\rangle + \langle N\rangle\right](\omega)$ for every $\omega \in \hat{\Omega}$ with $P(\hat{\Omega}) = 1$ and for every such ω, the Radon-Nikodym theorem implies the existence of functions $f_i(.,\omega) : [0,\infty) \to \mathbb{R}, i = 1,2,3$, such that

$$\langle M\rangle_t(\omega) = \int_0^t f_1(s,\omega)d\psi_s(\omega)\ ,\ \langle N\rangle_t(\omega) = \int_0^t f_2(s,\omega)d\psi_s(\omega)\ ,$$

$$\xi_t(\omega) = \langle M,N\rangle_t(\omega) = \int_0^t f_3(s,\omega)d\psi_s(\omega)\ ,\ 0 \leq t < \infty$$

Note that $f_1, f_2 \geq 0$ but f_3 is not necessarily positive. In order to see this recall that if V is a finite variation process then the total variation of V given by S^V can be written as the sum of two increasing functions as

$$S_t^V = A_t^1 + A_t^2$$

where as V can be written as

$$V_t = A_t^1 - A_t^2$$

where A^1 and A^2 are defined as

$$A_t^1 = \frac{S_t^V + V_t}{2} \text{ and } A_t^2 = \frac{S_t^V - V_t}{2}$$

In the context of this equalilty since we can write $\hat{\xi}_t(\omega) = A_t^1(\omega) + A_t^2(\omega)$ where both $A_t^1(\omega)$ and $A_t^2(\omega)$ are both non-negative and increasing. Now since $\hat{\xi}_t(\omega)$ is absolutely continuous with respect to $\psi_t(\omega)$ so is $A_t^1(\omega)$ and $A_t^2(\omega)$(since when $\psi_t(\omega) = 0$ then so is $\hat{\xi}_t(\omega)$ and hence $A_t^1(\omega)$ and $A_t^1(\omega)$). Now Radon Nikodym theorem implies the existence of a densities $f_4(\omega)$ and $f_5(\omega)$ such that

$$A_t^1(\omega) = \int_0^t f_4(s,\omega)d\psi_s(\omega) \text{ and } A_t^2(\omega) = \int_0^t f_5(s,\omega)d\psi_s(\omega)$$

and moreover for the total variation process $\hat{\xi}_t(\omega)$ we have

$$\hat{\xi}_t(\omega) = \int_0^t f_6(s,\omega)d\psi_s(\omega)$$

Hence we can conclude that

$$\begin{aligned} V_t = \xi_t(\omega) = A_t^1 - A_t^2 &= \int_0^t f_4(s,\omega)d\psi_s(\omega) - \int_0^t f_5(s,\omega)d\psi_s(\omega) \\ &= \int_0^t \left(f_4(s,\omega) - f_5(s,\omega)\right) d\psi_s(\omega) \end{aligned} \tag{2.2.26}$$

Now this implies that the $f_3(s,\omega)$ defined above is given by the difference of the densities $f_4(s,\omega)$ and $f_5(s,\omega)$, that is

$$f_3(s,\omega) = f_4(s,\omega) - f_5(s,\omega)$$

Similarly $\hat{\xi}_t(\omega) = A_t^1(\omega) + A_t^2(\omega)$ for all t,ω implies

$$f_6(s,\omega) = f_4(s,\omega) + f_5(s,\omega)$$

Consequently for $\alpha, \beta \in \mathbb{R}$ and $\omega \in \tilde{\Omega}_{\alpha\beta} \subseteq \hat{\Omega}$ satisfying $P(\tilde{\Omega}_{\alpha\beta}) = 1$, we have

$$0 \leq \langle \alpha M + \beta N \rangle_t(\omega) - \langle \alpha M + \beta N \rangle_u(\omega)$$
$$= \int_u^t \left(\alpha^2 f_1(s,\omega) + 2\alpha\beta f_3(s,\omega) + \beta^2 f_2(s,\omega)\right) d\psi_s(\omega); \quad 0 \leq u < t < \infty$$

Now obviously this can happen only if , for every $\omega \in \tilde{\Omega}_{\alpha\beta}$, there exists a set $T_{\alpha\beta}(\omega) \in \mathcal{B}([0,\infty))$ with $\int_{T_{\alpha\beta}(\omega)} d\psi_t(\omega) = 0$ and such that

$$\alpha^2 f_1(t,\omega) + 2\alpha\beta f_3(t,\omega) + \beta^2 f_2(t,\omega) \geq 0 \tag{2.2.27}$$

holds for every $t \notin T_{\alpha\beta}(\omega)$.Now let $\tilde{\Omega} := \bigcap_{\alpha,\beta\in\mathbb{Q}} \tilde{\Omega}_{\alpha\beta}$ and $T(\omega) = \bigcap_{\alpha,\beta\in\mathbb{Q}} T_{\alpha\beta}(\omega)$ so that $P(\tilde{\Omega}) = 1, \int_{T(\omega)} d\psi_t(\omega) = 0; \forall\omega \in \tilde{\Omega}$. Fix $\omega \in \tilde{\Omega}$ then (2.2.27) holds for every $t \notin T(\omega)$ and every pair (α, β) of rational numbers and thus also for every $t \notin T(\omega), (\alpha,\beta) \in \mathbb{R}^2$. In particular,

$$\alpha^2|X_t(\omega)|^2 f_1(t,\omega) + 2\alpha|X_t(\omega)Y_t(\omega)||f_3(t,\omega)| + |Y_t(\omega)|^2 f_2(t,\omega) \geq 0; \forall t \notin T(\omega)$$

Integrating with respect to $d\psi_t(\omega)$ we obtain

$$\alpha^2 \int_0^t |X_s|^2 d\langle M\rangle_s + 2\alpha \int_0^t |X_s Y_s||f_4(s,\omega) - f_5(s,\omega)| d\psi_s + \int_0^t |Y_s|^2 d\langle N\rangle_s \geq 0;\ 0 \leq t < \infty$$

almost surely. And hence we have

$$\alpha^2 \int_0^t |X_s|^2 d\langle M\rangle_s + 2\alpha \int_0^t |X_s Y_s| \left(f_4(s,\omega) + f_5(s,\omega)\right) d\psi_s + \int_0^t |Y_s|^2 d\langle N\rangle_s \geq 0;\ 0 \leq t < \infty$$

and therefore

$$\alpha^2 \int_0^t |X_s|^2 d\langle M\rangle_s + 2\alpha \int_0^t |X_s Y_s| f_6(s,\omega) d\psi_s + \int_0^t |Y_s|^2 d\langle N\rangle_s \geq 0;\ 0 \leq t < \infty$$

and hence

$$\alpha^2 \int_0^t |X_s|^2 d\langle M\rangle_s + 2\alpha \int_0^t |X_s Y_s| \hat{\xi}_s(\omega) + \int_0^t |Y_s|^2 d\langle N\rangle_s \geq 0;\ 0 \leq t < \infty$$

Now noting that the equation above is a quadratic equation in the variable α,in order to ensure that its aways positive , the discriminant has to be less than zero and hence we have

$$\int_0^t |X_s Y_s| d\hat{\xi}_s \leq \left(\int_0^t |X_s|^2 d\langle M\rangle_s\right)^{1/2} \left(\int_0^t |Y_s|^2 d\langle N\rangle_s\right)^{1/2}$$

□

Lemma 2.2.13. *(Lemma 3.2.15 Shreve) If $M, N \in \mathcal{M}_2^c$, $X \in \mathcal{L}^*(M)$ and $\{X^{(n)}\}_{n=1}^{\infty} \subseteq \mathcal{L}^*(M)$ is such that for some $T > 0$*

$$\lim_{n\to\infty} \int_0^T |X_u^{(n)} - X_u|^2 d\langle M\rangle_u \ ;\ a.s\ P$$

then

$$\lim_{n\to\infty} \langle I(X^{(n)}), N\rangle_t = \langle I(X), N\rangle_t$$

Proof. Problem 1.5.7(iii) on page 31-32 in [12] implies for $0 \leq t \leq T$

$$|\langle I(X^{(n)}) - I(X), N\rangle_t|^2 \leq \langle I(X^{(n)} - X)\rangle_t \langle N\rangle_t \leq \int_0^T |X_u^{(n)} - X_u|^2 d\langle M\rangle_u \cdot \langle N\rangle_T$$

But when we take limits , the last term on the right is zero and hence the desired result follows immediately by the sandwich theorem □

Lemma 2.2.14. *(Lemma 3.2.16 Shreve) If $M, N \in \mathcal{M}_2^c$ and $X \in \mathcal{L}^*(M)$ then*

$$\langle I^M(X), N\rangle_t = \int_0^t X_u d\langle M, N\rangle_u;\ 0 \leq t < \infty\ a.s. \tag{2.2.28}$$

Proof. According to Lemma 2.2.8, there exists a sequence $\{X^{(n)}\}_{n=1}^{\infty}$ of simple processes such that

$$\sup_{T>0} \lim_{n\to\infty} E\int_0^T |X_u^{(n)} - X_u|^2 d\langle M\rangle_u = 0$$

and hence consequently for each $T > 0$, a subsequence $\{\tilde{X}^{(n)}\}$ can be extracted for which

$$\lim_{n\to\infty} \int_0^T |\tilde{X}_u^{(n)} - X_u|^2 d\langle M\rangle_u = 0$$

But since (2.2.23) holds for simple processes, so we have

$$\langle I^M(\tilde{X}^{(n)}), N\rangle_t = \int_0^t \tilde{X}_u^{(n)} d\langle M, N\rangle_u;\ 0 \leq t \leq T \text{ a.s.}$$

Now letting $n \to \infty$ in the equation above Lemma 2.2.13 makes the L.H.S equal to $\langle I(X), N\rangle_t$

Now in order to show that the R.H.S converges to the right limit it is sufficient to show that

$$\lim_{n\to\infty} \left| \int_0^t \tilde{X}_u^{(n)} d\langle M, N\rangle_u - \int_0^t X_u d\langle M, N\rangle_u \right| = 0$$

We have from the triangular inequality, the Kunita-Watanabe Inequality and the assumption in the lemma that

$$\lim_{n\to\infty}\left|\int_0^t \left(\tilde{X}_u^{(n)} - X_u\right) d\langle M,N\rangle_u\right| \le \lim_{n\to\infty}\int_0^t \left|\tilde{X}_u^{(n)} - X_u\right| d\langle M,N\rangle_u$$

$$\le \lim_{n\to\infty}\int_0^t \left|\tilde{X}_u^{(n)} - X_u\right| d\hat{\xi}_u \le \lim_{n\to\infty}\left(\int_0^t \left|\tilde{X}_u^{(n)} - X_u\right|^2 d\langle M\rangle_u\right)^{1/2}\langle N\rangle_t$$

$$\le \lim_{n\to\infty}\left(\int_0^T \left|\tilde{X}_u^{(n)} - X_u\right|^2 d\langle M\rangle_u\right)^{1/2}\langle N\rangle_T = 0$$

where the second inequality follows from the fact that if $\langle M,N\rangle_t = A_t^1 - A_t^2$ then $\hat{\xi}_t = A_t^1 + A_t^2$ where A^1 and A^2 are non -decreasing processes starting at 0. And hence we have the right side which concludes the proof. □

Proposition 2.2.15. *(Proposition 3.2.17 Shreve) Let $M,N \in \mathcal{M}_2^c, X \in \mathcal{L}^*(M)$, and $Y \in \mathcal{L}^*(N)$, then the equivalent formulas (2.2.23) and (2.2.22) hold.*

Proof. Lemma 2.2.14 states that $d\langle M, I^N(Y)\rangle_u = Y_u d\langle M,N\rangle$. Replacing N in (2.2.28) by $I^N(Y)$, we have

$$\langle I^M(X), I^N(Y)\rangle_t = \int_0^t X_u d\langle M, I^N(Y)\rangle_u = \int_0^t X_u Y_u d\langle M,N\rangle_u;\ t\ge 0 \quad P \text{ a.s}$$

□

Proposition 2.2.16. *(Proposition 3.2.19 Shreve) Consider a martingale $M \in \mathcal{M}_2^c$ and a process $X \in \mathcal{L}^*(M)$. The stochastic integral $I^M(X)$ is the unique martingale which satisfies*

$$\langle \phi, N\rangle_t = \int_0^t X_u d\langle M,N\rangle_u;\ 0\le t<\infty, \qquad a.s.P \tag{2.2.29}$$

for every $N \in \mathcal{M}_2^c$

Proof. We already know from (2.2.28) that $\phi = I^M(X)$ satisfies (2.2.29) for every $N \in \mathcal{M}_2^c$. Subtracting (2.2.28) from (2.2.29) we have

$$\langle \phi - I^M(X), N\rangle_t = 0;\ 0\le t<\infty$$

Since this is true for all $N \in \mathcal{M}_2^c$, by setting $N = \phi - I^M(X)$, we see that the continuous martingale $\phi - I^M(X)$ has quadratic variation zero and hence $\phi = I^M(X)$.This is a direct consequence of Lemma 1.5.12 in [12]. □

Corollary 2.2.17. *(Corollary 3.2.20 Shreve) Suppose $M \in \mathcal{M}_2^c$, $X \in \mathcal{L}^*(M)$, and $N = I^M(X)$. Suppose further that $Y \in \mathcal{L}^*(N)$. Then $XY \in \mathcal{L}^*(M)$ and $I^N(Y) = I^M(XY)$*

Proof. Because $\langle N \rangle = \int_0^t X_s^2 d\langle M \rangle_s$, we have

$$E \int_0^T X_t^2 Y_t^2 d\langle M \rangle_t = E \int_0^T X_t^2 d\langle N \rangle_t < \infty$$

for all $T > 0$, so $XY \in \mathcal{L}^*(M)$. For any $\tilde{N} \in \mathcal{M}_2^c$, (2.2.23) gives (we take $Y = 1$) $d\langle \tilde{N}, N \rangle = X_s d\langle M, \tilde{N} \rangle_s$, and thus

$$\langle I^M(XY) \rangle = \int_0^t X_s Y_s d\langle M, \tilde{N} \rangle_s = \int_0^t Y_s d\langle N, \tilde{N} \rangle_s = \langle N, \tilde{N} \rangle_t$$

Now uniqueness of the representation of stochastic integral in Proposition 2.2.16 implies $I^M(XY) = I^N(Y)$ which completes the proof. □

Corollary 2.2.18. *(Corollary 3.2.21 Shreve) Suppose $M, \tilde{M} \in \mathcal{M}_2^c, X \in \mathcal{L}^*(M)$ and $\tilde{X} \in \mathcal{L}^*(\tilde{M})$ and there exists a stopping time T of the common filtration for these processes, such that for P-almost everywhere ω*

$$X_{t\wedge T(\omega)}(\omega) = \tilde{X}_{t\wedge T(\omega)}(\omega), ,\ M_{t\wedge T(\omega)}(\omega) = \tilde{M}_{t\wedge T(\omega)}(\omega);\ 0 \le t < \infty$$

Then

$$I^M_{t\wedge T(\omega)}(X)(\omega) = I^{\tilde{M}}_{t\wedge T(\omega)}(\tilde{X})(\omega);\ \ 0 \le t < \infty,\ \textit{for } P\textit{-a.e. } \omega$$

Proof. See Corollary 3.2.20 [12]. □

2.2.4 Integration with respect to continuous semi-martingales

Corollary 2.2.18, shows that stochastic integrals are determined locally by the local values of the integrator and the integrand. This fact allows us to broaden the classes of both integrands and integrators , a project we now undertake.

We begin by defining a continuous local martingale

Definition 2.2.19. *Let $X = \{X_t, \mathcal{F}_t; 0 < t \le \infty\}$ be a (continuous) process with $X_0 = 0$ a.s.. If there exists a non-decreasing sequence $\{T_n\}_{n=1}^{\infty}$ of stopping times of $\{\mathcal{F}_t\}$ such that $\{X_t^{(n)} := X_{t\wedge T_n}, \mathcal{F}_t; 0 \le t < \infty\}$ is a martingale for each $n \ge 1$ and $P[\lim_{n\to\infty} T_n = \infty] = 1$, then we say X is a (continuous) local martingale and write $X \in \mathcal{M}^{loc}$ (respectively , $X \in \mathcal{M}^{c,loc}$ if X is continuous).*

Let $M = \{M_t, \mathcal{F}_t; 0 \le t < \infty\}$ be a continuous local martingale on a probability space $(\Omega, \mathcal{F}, P)$ with $M_0 = 0$ a.s., i.e $M \in \mathcal{M}^{c,\text{loc}}$. Note that $\{\mathcal{F}_t\}$ satisfies the usual assumptions. Now we define an equivalence relation on the set of measurable, $\{\mathcal{F}_t\}$-adapted processes.

Definition 2.2.20. *We denote by $\mathcal{P}$, the collection of equivalence classes of all measurable, adapted processes $X = \{X_t, \mathcal{F}_t, 0 \le t < \infty\}$ satisfying*

$$P\left[\int_0^T X_t^2 d\langle M\rangle_t < \infty\right] \text{ for every } T \in [0,\infty) \tag{2.2.30}$$

We denote by $\mathcal{P}^*$ the collection of all progressively measurable processes satisfying the condition in equation (2.2.30)

We shall continue our development only for integrands in $\mathcal{P}^*$. If almost every path $t \mapsto \langle M\rangle_t$ of the quadratic variation process $\langle M\rangle$ is absolutely continuous, we can choose integrands from the wider class $\mathcal{P}$. Because M is in $\mathcal{M}^{c,\text{loc}}$, there exists a localizing sequence $\{\tau_n\}_{n=1}^{\infty}$ such that $\tau_n \uparrow \infty$ a.s. and for every $n \in \mathbb{N}$ we have that $(M_{t\wedge\tau_n})_{t\ge 0}$ is a continuous martingale. One can then define a sequence of stopping times σ_n by

$$\sigma_n = \inf\{t \ge 0 : M_t \ge n\}.$$

Almost sure continuity of $(M_t)_{t\ge 0}$ implies that $\sigma_n \uparrow \infty$ a.s. Hence we can conclude that there exists a non-decreasing sequence of stopping times $\{S_n\}_{n=1}^{\infty}$ given by $S_n = \sigma_n \wedge \tau_n$ such that $S_n \uparrow \infty$ a.s. P such that $\{M_{t\wedge S_n,}\mathcal{F}_t, 0 \le t < \infty\}$ is in $\mathcal{M}_2^c$.For $X \in \mathcal{P}^*$ one constructs a sequence of bounded stopping times by setting

$$R_n(\omega) = n \wedge \inf\left\{0 \le t < \infty; \int_0^t X_s^2(\omega) d\langle M\rangle_s(\omega) \ge n\right\}$$

It is not hard to see that R_n is a non-decreasing sequence in n since both the maps $n \mapsto n$ and $n \mapsto \inf\left\{0 \le t < \infty; \int_0^t X_s^2(\omega) d\langle M\rangle_s(\omega) \ge n\right\}$ are non-decreasing. The latter is non-decreasing because for $m \ge n$

$$\left\{0 \le t < \infty : \int_0^t X_s(\omega)^2 d\langle M\rangle(\omega) \ge m\right\} \subseteq \left\{0 \le t < \infty : \int_0^t X_s(\omega)^2 d\langle M\rangle(\omega) \ge n\right\}$$

implies

$$\inf\left\{0 \le t < \infty : \int_0^t X_s(\omega)^2 d\langle M\rangle(\omega) \ge m\right\} \ge \inf\left\{0 \le t < \infty : \int_0^t X_s(\omega)^2 d\langle M\rangle(\omega) \ge n\right\}$$

It is also not so hard to see that $R_n \uparrow \infty$ a.s since the map $n \mapsto n$ increases to ∞ as $n \to \infty$. On the other hand we know know that the map

$$n \mapsto \inf\left\{0 \le t < \infty; \int_0^t X_s^2(\omega) d\langle M\rangle_s(\omega) \ge n\right\}$$

is non decreasing and we have the following two possibilities

1. Either $t \mapsto \int_0^t X_s^2(\omega)d\langle M\rangle_s(\omega)$ is bounded by some finite M.In this case we have that for $n \geq M, R_n = \infty$ since the infimum of the empty set is infinity.

2. If there is no such finite M, right continuity of the map $t \mapsto \int_0^t X_s^2(\omega)d\langle M\rangle_s(\omega)$ implies that for all $n \in \mathbb{N}$ there exists an increasing sequence of times $(t_n)_{n\in\mathbb{N}}$ where $t_n < \infty$ such that $\int_0^{t_n} X_s^2(\omega)d\langle M\rangle_s(\omega) \geq n$ and $\int_0^{t_{n-1}} X_s^2(\omega)d\langle M\rangle_s(\omega) < n$.

By equation (2.2.30) it follows that $t_n \to \infty$ and hence

$$\inf\left\{0 \leq t < \infty : \int_0^t X_s(\omega)^2 d\langle M\rangle(\omega) \geq n\right\} \to \infty$$

as n goes to infinity which implies that $R_n \uparrow \infty$ a.s P. For $n \geq 1, \omega \in \Omega$, set

$$T_n(\omega) = R_n(\omega) \wedge S_n(\omega), \tag{2.2.31}$$

$$M_t^{(n)}(\omega) := M_{t\wedge T_n}(\omega),\ \ X_t^{(n)}(\omega) := X_t(\omega)1_{\{T_n(\omega)\geq t\}}; 0 \leq t < \infty \tag{2.2.32}$$

Then $M^{(n)} \in \mathcal{M}_2^c$ since stopped square integrable continuous martingale is again a square integrable continuous martingale by the Doob's optional sampling theorem (martingale variant of Problem 1.3.24 (i) in [12]).

Also $X^{(n)} \in \mathcal{L}^*(M^{(n)})$ since

$$E\int_0^T X_t^2 1_{\{T_n(\omega)\geq t\}} d\langle M^{(n)}\rangle_t = E\int_0^T X_t^2 1_{\{T_n(\omega)\geq t\}} d\langle M\rangle_{t\wedge T_n} = E\int_0^{t\wedge T_n} X_t^2 d\langle M\rangle_t \leq n$$

because $T_n \leq R_n$ and by the definition of R_n. Obviously the integral $I^{M^{(n)}}(X^{(n)})$ is well defined as a result of the construction above. Corollary 2.2.18 implies that for $1 \leq n \leq m$,

$$I_t^{M^{(m)}}(X^{(m)}) = I_t^{M^{(n)}}(X^{(n)}) \text{ for } 0 \leq t \leq T_n \tag{2.2.33}$$

so we may define the stochastic integral as

$$I_t(X) := I_t^{M^{(n)}}(X^{(n)}) \text{ on } \{0 \leq t \leq T_n\} \tag{2.2.34}$$

This definition is consistent, independent of the choice of $\{S_n\}_{n=1}^{\infty}$ and determines a continuous process(this part is obvious) which is a local martingale.

Proposition 2.2.21. *(Proposition 3.2.24 Shreve) Consider a local martingale $M \in \mathcal{M}^{c,loc}$ and a process $X \in \mathcal{P}^*(M)$. The stochastic integral $I^M(X)$ is the unique local martingale $\phi \in \mathcal{M}^{c,loc}$ which satisfies equation (2.2.35) for every $N \in \mathcal{M}_2^c$(or equivalently for every $N \in \mathcal{M}^{c,loc}$)*

$$\langle \phi, N\rangle_t = \int_0^t X_u d\langle M, N\rangle_u;\ \ 0 \leq t < \infty,\ \ a.s.\ P \tag{2.2.35}$$

Proof. In order to prove this proposition we refer the reader to the construction of the stochastic integral with respect to a continuous local martingale for processes $X \in \mathcal{P}^*(M)$. Using the same notation we have that $X^{(n)} \in \mathcal{L}^*(M)$ and $M^{(n)} \in \mathcal{M}_2^c$ for each $n \in \mathbb{N}$ and hence by equation (2.2.29) it follows that for all $N \in \mathcal{M}_2^c$ we have

$$\begin{aligned}\langle I^{M^{(n)}}(X^{(n)}), N\rangle_t &= \int_0^t X_u^{(n)} d\langle M^{(n)}, N\rangle_u = \int_0^t X_u 1_{\{T_n \geq u\}} d\langle M, N\rangle_{T_n \wedge u} \\ &= \int_0^t X_u 1_{\{T_n \geq u\}} d\langle M, N\rangle_u = \int_0^{t\wedge T_n} X_u d\langle M, N\rangle_u\end{aligned} \tag{2.2.36}$$

Now from the construction of the stochastic process, in particular equation (2.2.34) we have that

$$I_t^M(X) = I_t^{M^{(n)}}(X^{(n)}) \text{ for } 0 \leq t \leq T_n$$

and hence by together with (2.2.36) we have the following for $0 \leq t \leq T_n$

$$\langle I^M(X), N\rangle_t = \langle I^{M^{(n)}}(X^{(n)}), N\rangle_t = \int_0^{t\wedge T_n} X_u d\langle M, N\rangle_u$$

and since by Proposition 2.2.16 the uniqueness of $I^{M^{(n)}}(X^{(n})$ to be the only martingale to satisfy equation (2.2.29) for every $n \in \mathbb{N}$ implies that $I^M(X)$ is the unique local martingale which satisfies (2.2.29) for every $t \in \mathbb{R}_+$ upto T_n but as $T_n \uparrow \infty$ a.s. P ,we have the result for all $t \in \mathbb{R}_+$. □

Proposition 2.2.22. *(Problem 3.2.25 Shreve) Suppose $M, N \in \mathcal{M}^{c,loc}$ and $X \in \mathcal{P}^*(M) \cap \mathcal{P}^*(N)$. Show that for every pair (α, β) of real numbers we have*

$$I^{\alpha M+\beta N}(X) = \alpha I^M(X) + \beta I^N(X)$$

Proof. Since $M, N \in \mathcal{M}^{c,\text{loc}}$, it follows that $\alpha M + \beta N \in \mathcal{M}^{c,\text{loc}}$ and hence it follows by Proposition 2.2.21 that $I^{\alpha M+\beta N}(X)$ is the unique stochastic integral which satisfies $\langle I^{\alpha M+\beta N}(X), Z\rangle_t = \int_0^t X_u d\langle \alpha M + \beta N, Z\rangle_u$; $0 \leq t < \infty$, a.s. P for every $Z \in \mathcal{M}^{c,\text{loc}}$ Now using the bi-linearity of the quadratic co-variation process and the linearity of the riemann-stieltges integral

$$\begin{aligned}\langle I^{\alpha M+\beta N}(X), Z\rangle_t &= \int_0^t X_u d\langle \alpha M + \beta N, Z\rangle_u = \int_0^t X_u d\langle \alpha M, Z\rangle_u + \int_0^t X_u d\langle \beta N, Z\rangle_u \\ &= \alpha \int_0^t X_u d\langle M, Z\rangle_u + \beta \int_0^t X_u d\langle N, Z\rangle_u = \alpha\langle I^M(X), Z\rangle_t + \beta\langle I^N(X), Z\rangle_t \\ &= \langle \alpha I^M(X) + \beta I^N(X), Z\rangle\end{aligned} \tag{2.2.37}$$

and therefore we conclude that for all $Z \in \mathcal{M}^{c,\text{loc}}$

$$\langle I^{\alpha M+\beta N}(X), Z\rangle_t = \langle \alpha I^M(X) + \beta I^N(X), Z\rangle$$

Proposition 2.2.21 immediately implies that $I^{\alpha M+\beta N}(X) = \alpha I^M(X) + \beta I^N(X)$ and hence the proof is complete.

□

Lemma 2.2.23. *(Proposition 3.2.26 Shreve) Let $M \in \mathcal{M}^{c,loc}$, $\{X^{(n)}\}_{n=1}^{\infty} \subseteq \mathcal{P}^*(M)$ and suppose that for stopping time T of $\{\mathcal{F}_t\}$ we have $\lim_{n\to\infty} \int_0^T |X_t^{(n)} - X_t|^2 d\langle M\rangle_t = 0$ in probability. Then*

$$\sup_{0\le t\le T} \left| \int_0^t X_s^{(n)} dM_s - \int_0^t X_s dM_s \right| \to 0$$

in probability as $n \to \infty$.

Proof. See [12]. □

Lemma 2.2.24. *(Problem 3.2.27 Shreve) Let $M \in \mathcal{M}^{c,loc}$ and choose $X \in \mathcal{P}^*$. Show that there exists a sequence of simple processes $\{X^{(n)}\}_{n=1}^{\infty}$ such that for every $T > 0$*

$$\lim_{n\to\infty} \int_0^T |X_t^{(n)} - X_t|^2 d\langle M\rangle_t = 0$$

and

$$\lim_{n\to\infty} \sup_{0\le t\le T} |I_t(X^{(n)}) - I_t(X)| = 0$$

holds a.s P.If M is a one dimensional standard Brownian Motion, then the preceding holds with $X \in \mathcal{P}$

Proof. The proof is due to S. Dayanik. With $X \in \mathcal{P}^*(M)$, we construct a sequence of bounded stopping times $\{T_n\}_{n=1}^{\infty}$(see equation (2.2.31) in the section discussing the construction of stochastic integral with respect to continuous local martingales) such that each $X^{(n)} \in \mathcal{L}^*(M^{(n)})$ and therefore can be approximated by a sequence of simple processes $\{X^{n,k}\}_{k=1}^{\infty} \subseteq \mathcal{L}_0$ in the sense

$$\lim_{k\to\infty} E \int_0^T |X_t^{(n,k)} - X_t^{(n)}|^2 d\langle M^{(n)}\rangle_t = 0 \ \forall T < \infty$$

by Proposition 2.2.9. Let us now fix a positive $T < \infty$. By the equation just above we can find some m_n such that

$$E \int_0^T |X_t^{(n,m_n)} - X_t^{(n)}|^2 d\langle M^{(n)}\rangle_t < \frac{1}{n}$$

We claim that

$$\int_0^T |X_t^{(n,m_n)} - X_t|^2 d\langle M^{(n)}\rangle_t \xrightarrow{P} 0 \text{ as } n \to \infty$$

To show this we first observe that for every n, $X_t^{(n)} = X_t$ and $\langle M^{(n)}\rangle_t = \langle M\rangle_t$ for $0 \le t \le T$ on $\{T \le T_n\}$. Therefore for every fixed $\epsilon > 0$, we have

$$\begin{aligned} &P\left[\int_0^T |X_t^{(n,m_n)} - X_t|^2 d\langle M\rangle_t > \epsilon\right] \\ &\le P\left[\left\{\int_0^T |X_t^{(n,m_n)} - X_t|^2 d\langle M\rangle_t > \epsilon\right\} \cap \{T \le T_n\}\right] + P\left[T > T_n\right] \\ &= P\left[\left\{\int_0^T |X_t^{(n,m_n)} - X_t^{(n)}|^2 d\langle M^{(n)}\rangle_t > \epsilon\right\} \cap \{T \le T_n\}\right] + P\left[T > T_n\right] \\ &\le P\left[\left\{\int_0^T |X_t^{(n,m_n)} - X_t^{(n)}|^2 d\langle M^{(n)}\rangle_t > \epsilon\right\}\right] + P\left[T > T_n\right] \\ &\le \frac{1}{\epsilon} E\int_0^T |X_t^{(n,m_n)} - X_t^{(n)}|^2 d\langle M^{(n)}\rangle_t + P\left[T > T_n\right] \\ &\le \frac{1}{n\epsilon} + P\left[T > T_n\right] \end{aligned}$$

for every n. Since $\lim_{n\to\infty} P[T_n < T] = 0$(because $T_n \uparrow \infty$ a.s) and the inequality above is true for every ϵ we conclude

$$\int_0^T |X_t^{(n,m_n)} - X_t^{(n)}|^2 d\langle M^{(n)}\rangle_t \xrightarrow{P} 0 \text{ as } n \to \infty$$

We denote the simple process $X^{(n,m_n)}$ by $Y^{(T,n)}$ to emphasize its dependence on T. Now the equation just above and Proposition 2.2.23 together imply, that the following sequences of random variables

$$\int_0^T |Y_t^{(T,n)} - X_t^{(n)}|^2 d\langle M\rangle_t, \quad \sup_{0\le t\le T} |I_t(Y^{(T,n)}) - I_t(X)|$$

converge to zero in probability and hence there exists a subsequence for which the convergence takes place almost surely. Having done this construction for a fixed T, we use a diagonalization argument , as in the first paragraph of the Proof of Lemma 2.2.5, to obtain a sequence which works for all T. In case that M is a Brownian Motion we use Proposition 2.2.6, rather than Proposition 2.2.9 in the construction. □

Lemma 2.2.25. *(Problem 3.2.28 Shreve) Let $M{=}W$ be a standard Brownian Motion and $X \in \mathcal{P}$. We define for $0 \le s < t < \infty$*

$$\zeta_t^s(X) := \int_s^t X_u dW_u - \frac{1}{2}\int_s^t X_u^2 du; \ \zeta_t(X) := \zeta_t^0(X) \tag{2.2.38}$$

the process $\{\exp\zeta_t(X), \mathcal{F}_t, 0 \le t < \infty\}$ *is a supermartingale, it is a martingale if* $X \in \mathcal{L}_0$

Proof. We first show that for $X \in \mathcal{L}_0$,$\{\exp\zeta_t(X), \mathcal{F}_t, 0 \le t < \infty\}$ is a martingale. Let $\{\xi_n\}_{n=1}^{\infty}$ be a sequence of random variables such that $\sup_{n\ge 1}|\xi_n(\omega)|$. We only show the martingale property since adaptedness is obvious and integrability follows from the uniform boundedness of $\{\xi_n\}_{n=1}^{\infty}$ and integrability of the exponential of brownian increments.

Hence we need to show that

$$E\left[\zeta_t(X)|\mathcal{F}_s\right] = \zeta_s(X)$$

And since

$$E\left[\zeta_t(X)|\mathcal{F}_s\right] = \exp\zeta_s(X)E\left[\exp\zeta_t^s(x)|\mathcal{F}_s\right]$$

it is sufficient to show that

$$E\left[\exp\zeta_t^s(x)|\mathcal{F}_s\right] = 1$$

Now we define $V_i = \xi_i(M_{t_{i+1}} - M_{t_i}) - \frac{1}{2}\xi_i^2(t_{i+1} - t_i)$,$V_s = \xi_{m-1}(M_{t_m} - M_s) - \frac{1}{2}\xi_{m-1}^2(t_m - s)$ and $V_t = \xi_n(M_t - M_{t_n}) - \frac{1}{2}\xi_n^2(t_n - t)$ where $0 \le s < t < \infty$, m and n are chosen so that $t_{m-1} \le s < t_m$ and $t_n \le t < t_{n+1}$. Using the definition of $\zeta_t^s(X)$ and the definition of the stochastic integral of a process $X \in \mathcal{L}_0$ w.r.t to an integrand in $\mathcal{M}_2^c$ we have

$$E\left[\exp\left(V_s\right)\exp\left(\sum_{i=m}^{n-1} V_i\right)\exp\left(V_t\right)\middle|\mathcal{F}_s\right]$$
$$= E\left[E\left[E\left[\exp\left(V_s\right)|\mathcal{F}_s\right]\right]\left(\prod_{i=m}^{n-1} E\left[E\left[\exp\left(V_i\right)|\mathcal{F}_{t_i}\right]\right]\right)E\left[E\left[\exp\left(V_s\right)|\mathcal{F}_{t_n}\right]\right]\middle|\mathcal{F}_s\right]$$

We conclude by showing that all the conditional expectations above are 1. We have that

$$E\left[\exp\left(V_s\right)|\mathcal{F}_s\right] = E\left[\exp\left(\xi_{m-1}(M_{t_m} - M_s) - \frac{1}{2}\xi_{m-1}^2(t_m - s)\right)|\mathcal{F}_s\right]$$

where ξ_{m-1} are $\mathcal{F}_s$ measurable and $M_{t_m} - M_s$ is independent of $\mathcal{F}_s$ since $M = W$ is a standard Brownian motion.

To prove $E\left[\exp\left(V_s\right)|\mathcal{F}_s\right] = 1$, consider a sigma-algebra $\mathcal{G}$, a random variable U measurable with respect to $\mathcal{G}$ (such that U^2 has finite exponential moments) and a random variable V independent of $\mathcal{G}$, centered normal with variance s, then the goal is to show that

$$E\left[M \mid \mathcal{G}\right] = 1, \qquad M = \mathrm{e}^{UV - \frac{1}{2}U^2 s}.$$

That is, since U is measurable with respect to $\mathcal{G}$, $E\left[M \mid \mathcal{G}\right] = A(U)$, where

$$A(u) = E\left[\mathrm{e}^{uV - \frac{1}{2}u^2 s}|\mathcal{G}\right].$$

One sees that

$$A(u) = \mathrm{e}^{-\frac{1}{2}u^2 s} E\left[\mathrm{e}^{uV} | \mathcal{G}\right] = \mathrm{e}^{-\frac{1}{2}u^2 s} E\left[\mathrm{e}^{uV}\right],$$

where the last equality stems from the independence of V and $\mathcal{G}$. Finally, if V is standard normal with variance s, then $E\left[\mathrm{e}^{uV}\right] = \mathrm{e}^{\frac{1}{2}u^2 s}$, hence $A(u) = 1$ for every u.

In our context $U = \xi_{m-1}$, $V = M_{t_m} - M_s$ and $\mathcal{G} = \mathcal{F}_s$ and the variance of V is $t_m - s$. The exact same argument works for the conditional expectations of V_i with respect to the filtration $\mathcal{F}_{t_i}$and V_t w.r.t the filtration $\mathcal{F}_{t_n}$. Now and Fatou's lemma implies that $\{\zeta_t, \mathcal{F}_t, 0 \le t < \infty\}$ is a supermartingale. Indeed we have

$$\begin{aligned}
& E\left[\zeta_t(X) | \mathcal{F}_s\right] \\
&= E\left[\exp\left(\int_0^t X_u dW_u - \frac{1}{2}\int_0^t X_u^2 du\right) \middle| \mathcal{F}_s\right] \\
&= E\left[\exp\left(\lim_{n\to\infty}\int_0^t X_u^{(n)} dW_u - \frac{1}{2}\int_0^t (X_u^{(n)})^2 du\right) \middle| \mathcal{F}_s\right] \\
&= E\left[\lim_{n\to\infty}\exp\left(\int_0^t X_u^{(n)} dW_u - \frac{1}{2}\int_0^t (X_u^{(n)})^2 du\right) \middle| \mathcal{F}_s\right] \\
&\le \lim_{n\to\infty} E\left[\exp\left(\int_0^t X_u^{(n)} dW_u - \frac{1}{2}\int_0^t (X_u^{(n)})^2 du\right) \middle| \mathcal{F}_s\right] \\
&= \lim_{n\to\infty}\exp\left(\int_0^s X_u^{(n)} dW_u - \frac{1}{2}\int_0^s (X_u^{(n)})^2 du\right) \\
&= \exp\left(\lim_{n\to\infty}\int_0^s X_u^{(n)} dW_u - \frac{1}{2}\int_0^s (X_u^{(n)})^2 du\right) \\
&= \exp\left(\int_0^s X_u dW_u - \frac{1}{2}\int_0^s X_u^2 du\right) = \zeta_s(X)
\end{aligned} \tag{2.2.39}$$

This completes the proof □

Lemma 2.2.26. *(Exercise 3.2.30 Karatzas and Shreve) For $M \in \mathcal{M}^{c,loc}, X \in \mathcal{P}^*$, and Z an $\mathcal{F}_s$-measurable random variable , show that*

$$\int_s^t Z X_u dM_u = Z \int_s^t X_u dM_u;\ \ s \le t < \infty,\ \ a.s. \tag{2.2.40}$$

Proof. First note that for the stochastic integral $\int_s^t Z X_u dM_u$ to be well-defined we need that $ZX \in \mathcal{P}^*$ i.e $P\left[\int_0^T Z^2 X_t^2 d\langle M\rangle_t < \infty\right] = 1$ for every $T \in [0, \infty)$. Suppose that we have proved (2.2.40) for bounded Z then we have that for $Z_k = Z 1_{\{Z \le k\}}$

$$\int_s^t Z_k X_u dM_u = Z_k \int_s^t X_u dM_u;\ \ s \le t < \infty,\ \ \text{a.s.} \tag{2.2.41}$$

Now in order to compute $\lim_{k\to\infty}\int_s^t Z_k X_u dM_u$ we use the fact that $Z_k^2 X_u^2 \leq Z^2 X_u^2$, for all $k \in \mathbb{N}$ with $\int_0^T Z^2 X_u^2 d\langle M\rangle_u < \infty$ we can apply the Lebesgue dominate convergence theorem to conclude that

$$\lim_{n\to\infty}\int_0^T |Z_k X_t - ZX_t|^2 = \int_0^T \lim_{n\to\infty} |Z_k X_t - ZX_t|^2 = 0 \text{ a.s.}$$

And hence the convergence also holds in probability and therefore we can apply Lemma 2.2.23 to conclude that

$$\sup_{0\leq t\leq T}\left|\int_0^t Z_k X_s dM_s - \int_0^t ZX_s dM_s\right| \to 0 \text{ as } k \to \infty$$

in probability for the deterministic stopping time $T > 0$. And therefore there exists a subsequence $(k_n)_{n\in\mathbb{N}}$ such that

$$\sup_{0\leq t\leq T}\left|\int_0^t Z_{k_n} X_s dM_s - \int_0^t ZX_s dM_s\right| \to 0 \text{ a.s. as } n \to \infty$$

Now taking limits in equation (2.2.41) as $n \to \infty$

$$\lim_{n\to\infty}\int_s^t Z_{k_n} X_u dM_u = \lim_{n\to\infty} Z_{k_n}\int_s^t X_u dM_u$$

we get

$$\int_s^t ZX_u dM_u = Z\int_s^t X_u dM_u$$

Now in order to complete the proof we need to show that the result holds for bounded $\mathcal{F}_s$-measurable random variable Z. For the sake of notational simplicity we prove it for $s = 0$. Proposition 2.2.21 tells us that for any continuous local martingale N we have the following

$$\langle I^M(ZX), N\rangle_t = \int_0^t ZX_u d\langle M, N\rangle_u = Z\int_0^t X_u d\langle M, N\rangle_u = Z\langle I^M(X), N\rangle_t \qquad (2.2.42)$$

The second equality follows from the fact that the integral is computed ω-wise, we can pull $Z(\omega)$ out of the integral. Since by definition

$$\langle I^M(X), N\rangle_t = \left\langle \int_0^t X_u dM_u, N\right\rangle_t = \langle \Phi, N\rangle_t$$

where we assume $\Phi_t = \int_0^t X_u dM_u$ Using the local martingale analogue of Problem 1.5.14 in Karatzas and Shreve [12] we have that for $X, Y \in \mathcal{M}^{c,\text{loc}}$ and a partition $\Pi = \{t_0, t_1, \cdots, t_m\}$ of $[0, t]$ we have that

$$\lim_{\|\Pi\|\to 0}\sum_{k=1}^m \left(X_{t_k} - X_{t_{k-1}}\right)\left(Y_{t_k} - Y_{t_{k-1}}\right) = \langle X, Y\rangle_t \text{ in probability}$$

Hence it follows by the definition of quadratic covariation that

$$\lim_{\|\Pi\|\to 0} \sum_{k=1}^{m} \left(ZX_{t_k} - ZX_{t_{k-1}}\right)\left(Y_{t_k} - Y_{t_{k-1}}\right) = \langle ZX, Y\rangle_t \text{ in probability .}$$

Moreover we have that if a sequence of random variables $G_n \to G$ in probability and F is another random variable on the same probability space then $FG_n \to FG$ in probability. In the context of this problem we have that $Z\sum_{k=1}^{m}\left(X_{t_k} - X_{t_{k-1}}\right)\left(Y_{t_k} - Y_{t_{k-1}}\right) = \sum_{k=1}^{m}\left(ZX_{t_k} - ZX_{t_{k-1}}\right)\left(Y_{t_k} - Y_{t_{k-1}}\right)$ such that $Z\sum_{k=1}^{m}\left(X_{t_k} - X_{t_{k-1}}\right)\left(Y_{t_k} - Y_{t_{k-1}}\right)$ converges to $Z\langle X, Y\rangle_t$ in probability whereas $\sum_{k=1}^{m}\left(ZX_{t_k} - ZX_{t_{k-1}}\right)\left(Y_{t_k} - Y_{t_{k-1}}\right)$ converges to $\langle ZX, Y\rangle_t$ in probability and since limit in probability is almost surely unique we get

$$Z\langle X, Y\rangle_t = \langle ZX, Y\rangle_t \text{ a.s } P \quad \forall 0 \le t < \infty$$

□

2.2.5 The Change of Variables Formula

One of the most important tools in the study of stochastic processes of the martingale type is the *change-of-variable formula* or *Itô's* rule as it is better known. It provides an integral-differential calculus for the sample paths of such processes. Let us consider a basic probability space $(\Omega, \mathcal{F}, P)$ with an associated filtration $\{\mathcal{F}_t\}$ which we always assume to satisfy the usual conditions.

Definition 2.2.27. *A continuous semi-martingale $X = \{X_t, \mathcal{F}_t; 0 \le t < \infty\}$ is an adapted process which has the decomposition P a.s.,*

$$X_t = X_0 + M_t + B_t; \quad 0 \le t < \infty, \tag{2.2.43}$$

where $M = \{M_t, \mathcal{F}_t; 0 \le t < \infty\}$ is a continuous local martingale and $B = \{B_t, \mathcal{F}_t; 0 \le t < \infty\}$ is the difference of continuous, nondecreasing adapted proceses $\{A_t^{\pm}, \mathcal{F}_t; 0 \le t < \infty\}$:

$$B_t = A_t^{+} - A_t^{-}; \quad 0 \le t < \infty \tag{2.2.44}$$

with $A_0^{\pm} = 0, P$ a.s.

Itô's rule states that a "smooth function" of a continuous semi-martingale is a continuous semi-martingale and provides its decomposition. We state the theorem without proof

Theorem 2.2.28. *Let $\{M_t =: (M_t^{(1)}, \dots, M_t^{(d)}), \mathcal{F}_t; 0 \le t < \infty\}$ be a vector of continuous local martingales, $\{B_t =: (B_t^{(1)}, \dots, B_t^{(d)}), \mathcal{F}_t; 0 \le t < \infty\}$ a vector of adapted* *processes of bounded variation with $B_0 = 0$, and set $X_t = X_0 + M_t + B_t; 0 \leq t < \infty$, where X_0 is a $\mathcal{F}_0$ measurable random vector in $\mathbb{R}^d$. Let $f(t,x) : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}$ be of class $C^{1,2}$. Then , a.s P,*

$$\begin{aligned} f(t,X_t) = f(0,X_0) + \int_0^t \frac{\partial}{\partial t} f(s,X_s)ds + \sum_{i=1}^d \int_0^t \frac{\partial}{\partial x_i} f(s,X_s)dB_s^{(i)} \\ + \sum_{i=1}^d \int_0^t \frac{\partial}{\partial x_i} f(s,X_s)dM_s^{(i)} \qquad (2.2.45) \\ + \frac{1}{2}\sum_{i=1}^d \sum_{j=1}^d \int_0^t \frac{\partial^2}{\partial x_i \partial x_j} f(s,X_s)d\langle M^{(i)}, M^{(j)}\rangle_s, \quad 0 \leq t < \infty \end{aligned}$$

Lemma 2.2.29. *(Problem 3.3.12 Karatzas and Shreve) Suppose we have two continuous semi-martingales*

$$X_t = X_0 + M_t + B_t,\ Y_t = Y_0 + N_t + C_t;\ ; 0 \leq t < \infty$$

where $M, N \in \mathcal{M}^{c,loc}$ and B and C are adapted continuous processes of bounded variation with $B_0 = C_0 = 0$ a.s. Prove the integration by parts formula

$$\int_0^t X_s dY_s = X_t Y_t = X_0 Y_0 - \int_0^t Y_s dX_s - \langle M, N\rangle_t. \qquad (2.2.46)$$

Proof. Using the linearity of the integrator in the stochastic integral we get

$$d(XY) = \frac{1}{4} d((X+Y)^2 - (X-Y)^2) = \frac{1}{4}\left(d(X+Y)^2 - d(X-Y)^2\right) \qquad (2.2.47)$$

A simple application of Itô's lemma gives us

$$d(X+Y)^2 = 2(X+Y)d(X+Y) + d\langle X+Y\rangle$$

and

$$d(X-Y)^2 = 2(X-Y)d(X-Y) + d\langle X-Y\rangle$$

Now from (2.2.47) we get

$$d(XY) = \frac{1}{4}\left(2(X+Y)d(X+Y) + d\langle X-Y\rangle + 2(X-Y)d(X-Y) - d\langle X-Y\rangle\right)$$

which simplifies to

$$\begin{aligned} d(XY) &= \frac{1}{4}(4XdY + 4YdX) + \frac{1}{4}(\langle X+Y\rangle - \langle X-Y\rangle) \\ &= XdY + YdX + d\langle X,Y\rangle = XdY + YdX + d\langle M,N\rangle \end{aligned}$$

Note that $\langle X, Y\rangle = \langle X_0 + M + B, Y_0 + N + C\rangle = \langle M, N\rangle$ using the bi-linearity of the quadratic variation , the fact that the quadratic co-variation of a finite variation process and a martingale is zero and the quadratic covariation of two finite variation processes is zero Now integrating the L.H.S and the R.H.S from 0 to t we get

$$X_t Y_t = X_0 Y_0 + \int_0^t Y_s dX_s + \int_0^t X_s dY_s + \langle M, N\rangle_t$$

□

Lemma 2.2.30. *(3.3.25 Exercise Shreve) With $W = \{W_t, \mathcal{F}_t; 0 \le t < \infty\}$ is a standard one dimensional Brownian Motion and X is measurable, adapted process satisfying*

$$E\int_0^T |X_t|^{2m} dt < \infty \tag{2.2.48}$$

for some real numbers $T > 0$ and $m \ge 1$, then

$$E\left|\int_0^T X_t dW_t\right|^{2m} \le (m(2m-1))^m T^{m-1} E\int_0^T |X_t|^{2m} dt \tag{2.2.49}$$

Proof. Applying Itô's lemma to the continuous semi-martingale (submartingale) $|M|^{2m}$ we get

$$M_t^{2m} = 2m\int_0^t M_s^{2m-1} dM_s + m(2m-1)\int_0^t M_s^{2m-2} d\langle M\rangle_s$$

Taking expectations and reapplying Itô's lemma to $M^{2m-2}, M^{2m-4}, \cdots, M^2$, using Fubini's theorem to interchange the integral and the expectation and using the fact the expectation of the stochastic integral(which is a martingale starting at zero) is 0 we get

$$\begin{aligned}
E\left[M_t^{2m}\right] &= 2mE\left[\int_0^t M_s^{2m-1} dM_s\right] + m(2m-1)E\left[\int_0^t M_s^{2m-2} d\langle M\rangle_s\right] \\
&= m(2m-1)\int_0^t E\left[M_s^{2m-2}\right] d\langle M\rangle_s \\
&= \frac{1}{2}2m(2m-1)\int_0^t \frac{1}{2}(2m-2)(2m-3)E\left[\int_0^s M_u^{2m-4} d\langle M\rangle_u\right] \\
= \cdots &= \left(\frac{1}{2}\right)^m (2m).(2m-1).(2m-2)\ldots 1E\int_0^t\int_0^s\cdots\int_0^u \langle M\rangle_v d\langle M\rangle_v \\
&\le (m(2m-1))^m E\left[\langle M\rangle_t^m\right] = (m(2m-1))^m E\left[\left(\int_0^t X_u^2 du\right)^m\right] \\
&\overset{\text{Holders's}}{\le} (m(2m-1))^m E\left[\left(\int_0^t X_u^{2m} du\right)^{m/m}\left(\int_0^t du\right)^{m(1-\frac{1}{m})}\right] \\
&= (m(2m-1))^m t^{m-1} E\left[\left(\int_0^t X_u^{2m} du\right)\right]
\end{aligned}$$

And hence we have the result □

2.3 Introduction to Stochastic Differential Equations

In this section we will introduce the concept of strong solutions of stochastic differential equations with respect to a Brownian motion. We will follow the presentation in the book of (Karatzas and Shreve [12]) and provide some important results on existence and uniqueness of SDE's due to (Itô [9] and Yamada and Watanabe [10])

Let us start with Borel-measurable functions $b_i(t,x), \sigma_{ij}(t,x); 1 \leq i \leq d, 1 \leq j \leq r$, from $[0,\infty) \times \mathbb{R}^d$ into $\mathbb{R}$, and define the $(d \times 1)$ drift vector $b(t,x) = \{b_i(t,x)\}_{1\leq i\leq d}$ and the $(d \times r)$ dispersion matrix $\sigma(t,x) = \{\sigma_{ij}(t,x)\}_{\substack{1\leq i\leq d \\ 1\leq j\leq r}}$. The intent is to assign a meaning to a stochastic differential equation

$$dX_t = b(t,X_t)dt + \sigma(t,X_t)dW_t \tag{2.3.1}$$

written component wise as

$$dX_t^i = b_i(t,X_t)dt + \sum_{j=1}^{r} \sigma_{ij}(t,X_t)dW_t^j; \quad 1 \leq i \leq d \tag{2.3.2}$$

where $W = \{W_t; 0 \leq t < \infty\}$ is an r-dimensional Brownian motion and $X = \{X_t; 0 \leq t < \infty\}$ is as suitable stochastic process with continuous sample paths and values in $\mathbb{R}^d$, the **solution** of the the equation. The drift vector $b(t,x)$ and the dispersion matrix $\sigma(t,x)$ are the coefficients of the equation; the $(d \times d)$ matrix $a(t,x) := \sigma(t,x)\sigma^T(t,x)$ with elements

$$a_{ij}(t,x) := \sum_{j=1}^{r} \sigma_{ij}(t,x)\sigma_{kj}(t,x); \quad 1 \leq i,k \leq d \tag{2.3.3}$$

will be called the diffusion matrix.

In order to develop the concept of strong solution, we choose a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ as well as a r-dimensional Brownian motion $W = \{W_t, \mathcal{F}_t^W; 0 \leq t < \infty\}$ on it. We assume also that this space is rich enough to accommodate a random vector ξ taking values in $\mathbb{R}^d$, independent of $\mathcal{F}_\infty^W$, and with the given distribution

$$\mu(T) = P[\xi \in \Gamma]; \quad \Gamma \in \mathcal{B}(\mathbb{R}^d)$$

We consider the filtration

$$\mathcal{G}_t := \sigma(\xi) \vee \mathcal{F}_t^W = \sigma(\xi, W_s, 0 \leq s \leq t); \quad 0 \leq t < \infty$$

as well as the collection of null sets

$$\mathcal{N} = \{N \subseteq \Omega, \exists G \in \mathcal{G}_\infty \text{ with } N \subseteq G \text{ and } P(G) = 0\}$$

and create the augmented filtration

$$\mathcal{F}_t := \sigma(\mathcal{G}_t \cup \mathcal{N}), \quad 0 \le t < \infty; \quad \mathcal{F}_\infty = \sigma\left(\bigcup_{t \ge 0} \mathcal{F}_t\right) \tag{2.3.4}$$

Obviously $\{W_t, \mathcal{G}_t; 0 \le t \le \infty\}$ is an r-dimensional Brownian motion, and then so is $\{W_t, \mathcal{F}_t; 0 \le t \le \infty\}$(see Theorem 2.7.9 in [12]). Note that the filtration thus satisfies the usual conditions(see Theorem 2.7.7 in [12]).

2.3.1 Strong Solutions of SDE

Definition 2.3.1. *A strong solution of the stochastic differential equation (2.3.1), on a given probability space $(\Omega, \mathcal{F}, P)$ and with respect to the Brownian motion W and initial condition ξ, is a process $X = \{X_t; 0 \le t < \infty\}$ with continuous sample paths and with the following properties:*

(i) X *is adapted to the filtration* $\mathcal{F}_t$ *of (2.3.4),*

(ii) $P[X_0 = \xi] = 1$,

(iii) $P\left[\int_0^t \{|b_i(s, X_s)| + \sigma_{ij}^2(s, X_s)\} ds < \infty\right]$

(iv) the integral version of (2.3.1)

$$X_t = X_0 + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s; \quad 0 \le t < \infty \tag{2.3.5}$$

or equation

$$X_t^{(i)} = X_0^{(i)} + \int_0^t b_i(s, X_s) ds + \sum_{j=1}^{r} \int_0^t \sigma_{ij}(s, X_s) dW_s^{(j)}; \quad 0 \le t < \infty, 1 \le i \le d \tag{2.3.6}$$

Remark 2.3.2. *Note that the crucial requirement of this definition is captured in condition (i); it corresponds to our intuitive understanding of X as the **output** of a dynamical system described by a pair of coefficients (b, σ) whose input is W and which is fed by the initial datum ξ. The principal of causality for dynamical systems requires that the output X_t at time t depend only on ξ and the values of the input $\{W_s; 0 \le s \le t\}$ up to that time. This principal finds it mathematical expression in (i)*

Definition 2.3.3. *Let the drift vector $b(t, x)$ and dispersion matrix $\sigma(t, x)$ be given. Suppose that, whenever W is an r-dimensional Brownian motion on some $(\Omega, \mathcal{F}, P), \xi$ is an independent, d-dimensional random vector, $\{\mathcal{F}_t\}$ is given by (2.3.4), and $X, \tilde{X}$ are* *two strong solutions of (2.3.1) relative to* W *with initial condition* ξ *then* $P[X_t = \tilde{X}_t; 0 \leq t < \infty] = 1$. *Under these conditions we say that strong uniqueness holds for the pair* (b, σ).

In the early 1940's K. Itô(see [16] and [9]) proved a series of results on the existence and uniqueness of strong solutions to stochastic differential equations with Lipschitz drift and dispersion coefficients. I begin this program with a short help lemma:

Lemma 2.3.4. *Suppose that the continuous function*

$$0 \leq g(t) \leq \alpha(t) + \beta \int_0^t g(s)ds; \quad 0 \leq t \leq T \tag{2.3.7}$$

with $\beta \geq 0$ *and* $\alpha : [0, T] \mapsto R$ *integrable. Then*

$$g(t) \leq \alpha(t) + \beta \int_0^t \alpha(s)e^{\beta(t-s)}ds \quad 0 \leq t \leq T \tag{2.3.8}$$

Proof. It follows from (2.3.7)

$$\frac{d}{dt}\left(e^{-\beta t}\int_0^t g(s)ds\right) = \left(g(t) - \beta\int_0^t g(s)ds\right)e^{-\beta t} \leq \alpha(t)e^{-\beta t}$$

Integrating with respect to the variable t we get

$$e^{-\beta t}\int_0^t g(s)ds \leq \int_0^t \alpha(s)e^{-\beta s}ds$$

and hence we can conclude

$$\int_0^t g(s)ds \leq e^{\beta t}\int_0^t \alpha(s)e^{-\beta s}ds$$

and Gronwall's inequality follows from (2.3.7)

□

Theorem 2.3.5. *Suppose that the coefficients* $b(t, x), \sigma(t, x)$ *are locally Lipschitz-continuous in the space variable; i.e.,for every integer* $n \geq 1$ *there exists a constant* $K_n > 0$ *such that for every* $t \geq 0, \|x\| \leq n, \|y\| \leq n$:

$$\|b(t, x) - b(t, y)\| + \|\sigma(t, x) - \sigma(t, y)\| \leq K_n \|x - y\|. \tag{2.3.9}$$

Then strong uniqueness holds for equation (2.3.1)

Proof. Let us suppose that X and $\tilde{X}$ are both strong solutions defined for all $t \geq 0$, of (2.3.1) relative to the same Brownian motion W and the same initial condition ξ, on some $(\Omega, \mathcal{F}, \mathcal{P})$. We define the stopping times $\tau_n = \inf\{t \geq 0; \|X_t\| \geq n\}$ for $n \geq 1$, as well as their tilded counterparts and set $S_n := \tau_n \wedge \tilde{\tau}_n$. The almost sure continuity of the the stochastic processes $X, \tilde{X}$(as a consequence of the assumption that they are strong solutions) implies that $\lim_{n\to\infty} \tau_n = \infty$ a.s P.and $\lim_{n\to\infty} \tilde{\tau}_n = \infty$ a.s. P. As a consequence we have $\lim_{n\to\infty} S_n = \infty$ a.s. P.

Since X and $\tilde{X}$ are solutions to the SDE (2.3.1) they satisfy equation (2.3.5) and hence we get

$$X_{t\wedge S_n} - \tilde{X}_{t\wedge S_n} = \int_0^{t\wedge S_n} \{b(u, X_u) - b(u, \tilde{X}_u)\}du + \int_0^{t\wedge S_n} \{\sigma(u, X_u) - \sigma(u, \tilde{X}_u)\}dW_u$$

Using the vector inequality $\|v_1 + \cdots + v_k\|^2 \leq k^2(\|v_1\|^2 + \cdots + \|v_k\|^2)$, the triangular inequality ,the Hölder inequality for Lebesgue integrals, the basic property of stochastic integrals(2.2.24) and equation (2.3.9) we may write for $0 \leq t \leq T$:

$$\begin{aligned} E\left\|X_{t\wedge S_n} - \tilde{X}_{t\wedge S_n}\right\|^2 &\leq 2^2 E\left[\int_0^{t\wedge S_n} \left\|b(u, X_u) - b(u, \tilde{X}_u)\right\| du\right]^2 \\ &+2^2 E\sum_{i=1}^{d}\left[\sum_{j=1}^{r}\int_0^{t\wedge S_n} \sigma_{ij}(u, X_u) - \sigma_{ij}(u, \tilde{X}_u)dW_u^{(j)}\right]^2 \\ \leq 4tE\int_0^{t\wedge S_n} \left\|b(u, X_u) - b(u, \tilde{X}_u)\right\|^2 du &+ 4E\int_0^{t\wedge S_n} \left\|\sigma(u, X_u) - \sigma(u, \tilde{X}_u)\right\|^2 du \\ &\leq 4(T+1)K_n^2 \int_0^t E\left\|X_{u\wedge S_n} - \tilde{X}_{u\wedge S_n}\right\|^2 du \end{aligned}$$

Now a simple application of Lemma 2.3.4 with $g(t) := E\left\|X_{t\wedge S_n} - \tilde{X}_{t\wedge S_n}\right\|^2$ allows us to conclude that $\{X_{t\wedge S_n}; 0 \leq t < \infty\}$ and $\{\tilde{X}_{t\wedge S_n}; 0 \leq t < \infty\}$ are modifications of each other and hence indistinguishable (Since $X, \tilde{X}$ are strong solutions to the SDE in (2.3.1), they are continuous almost surely) Now letting $n \to \infty$ we get that $\{X_t; 0 \leq t < \infty\}$ and $\{\tilde{X}_t; 0 \leq t < \infty\}$ are indistinguishable and hence we have strong uniqueness.

□

Theorem 2.3.6. *Suppose that the coefficients $b(t, x), \sigma(t, x)$ satisfy the global Lipschitz and linear growth conditions*

$$\|b(t, x) - b(t, y)\| + \|\sigma(t, x) - \sigma(t, y)\| \leq K\|x - y\|, \tag{2.3.10}$$

$$\|b((t, x)\|^2 + \|\sigma(t, x)\|^2 \leq K^2(1 + \|x\|^2), \tag{2.3.11}$$

for every $0 \le t < \infty, x \in \mathbb{R}^d, y \in \mathbb{R}^d$, *where* K *is a positive constant. On some probability space* $(\Omega, \mathcal{F}, P)$, *let* ξ *be an* $\mathbb{R}^d$*-values random vector, independent of the* r*-dimensional Brownian motion* $W = \{W_t, \mathcal{F}_t^W; 0 \le t < \infty\}$, *with a finite second moment*

$$E\left\|\xi\right\|^2 < \infty \tag{2.3.12}$$

Let $\mathcal{F}_t$ *be as in (2.3.4). Then there exists a continuous adapted process* $X = \{X_t, \mathcal{F}_t; 0 \le t < \infty\}$ *which is a strong solution of equation (2.3.1) relative to* W *with the initial condition* ξ. *Moreover, this process is square-integrable: for every* $T > 0$, *there exists a constant* C, *depending only on* K *and* T *such that*

$$E\left\|X_t\right\|^2 \le C(1 + E\left\|\xi\right\|^2)e^{Ct}; \quad 0 \le t \le T \tag{2.3.13}$$

Proof. The idea of the proof is to mimic the deterministic situation and construct recursively a sequence with successive approximations by setting $X_t^{(0)} = \xi$ and

$$X_t^{(k+1)} := \xi + \int_0^t b(s, X_s^{(k)})ds + \int_0^t \sigma(s, X_s^{(k)})dW_s; \quad 0 \le t < \infty \tag{2.3.14}$$

The processes $\{X^{(k)}\}_{k=1}^{\infty}$ are obviously continuous and adapted to the filtration $\{\mathcal{F}_t\}$. The hope is that the sequence $\{X^{(k)}\}_{k=1}^{\infty}$ will converge to solution of the equation (2.3.1)

Before continuing let us first establish that for every $T > 0$, there exists a positive constant C depending only on K and T such that for the iterations in (2.3.14) we have

$$E\left\|X_t^{(k)}\right\|^2 \le C(1 + E\left\|\xi\right\|^2)e^{Ct}; \quad 0 \le t \le T, k \ge 0 \tag{2.3.15}$$

We first check that each $X_t^{(k)}$ is well defined for all $t \ge 0$. In particular we must show that for all $k \ge 0$,

$$\int_0^t \left(\left\|b(s, X_s^{(k)})\right\| + \left\|\sigma(s, X_s^{k})\right\|^2\right) ds < \infty; \quad 0 \le T < \infty$$

In light of (2.3.11) this will follow immediately if one demonstrates the following

$$\sup_{0 \le t \le T} E\left\|X_t^{(k)}\right\|^2 < \infty; \quad 0 \le T < \infty \tag{2.3.16}$$

Equation (2.3.16) can be proved using induction. For $k = 0$, it is a simple consequence of (2.3.12). Now assume that (2.3.16) holds for some value of k. Proceeding similarly to the proof of Theorem 2.3.5, we obtain the following bound for $0 \le t \le T$:

$$E\left\|X_t^{(k+1)}\right\|^2 \le 9E\left\|\xi\right\|^2 + 9(T+1)K^2 \int_0^t \left(1 + E\left\|X_s^{(k)}\right\|\right) \tag{2.3.17}$$

which gives us (2.3.16) for $k+1$. From (2.3.17) it follows

$$E\left\|X_t^{(k+1)}\right\|^2 \le C\left(1+E\|\xi\|^2\right)+C\int_0^t E\left\|X_s^{(k)}\right\|^2 ds; \quad 0\le t\le T$$

where C depends only on K and T. Iteration of this inequality gives

$$\begin{aligned} E\left\|X_t^{(k+1)}\right\|^2 &\le C\left(1+E\|\xi\|^2\right)\left[1+Ct+\frac{(Ct)^2}{2!}+\cdots+\frac{(Ct)^{k+1}}{(k+1)!}\right] \\ &\le C\left(1+E\|\xi\|^2\right)e^{Ct} \end{aligned}$$

and thus we have shown (2.3.15) holds.

From (2.3.14) we have that $X_t^{(k+1)}-X_t^{(k)}=B_t+M_t$, where

$$B_t:=\int_0^t \left\{b(s,X_s^{(k)})-b(s,X_s^{(k-1)})\right\}ds, \quad M_t:=\int_0^t \left\{\sigma(s,X_s^{(k)})-\sigma(s,X_s^{(k-1)})\right\}dW_s$$

Thanks to the inequalities (2.3.11), (2.3.15), the process $\{M_t=(M_t^{(1),\dots,M_t^{(d)}}),\mathcal{F}_t;0\le t<\infty\}$,is seen to be a vector of square-integrable martingales. Using a variant of the Burkholder-Davis-Gundy-Inequality(see Problem 3.3.29 and Remark 3.3.30 on page 166 in [12]) to we get

$$\begin{aligned} E\left[\max_{0\le s\le t}\|M_s\|^2\right] &\le \Lambda_1 E\int_0^t \left\|\sigma(s,X_s^{(k)})-\sigma(s,X_s^{(k-1)})\right\|^2 ds \\ &\le \Lambda_1 K^2 E\int_0^t \left\|X_s^{(k)}-X_s^{(k-1)}\right\|^2 ds \end{aligned}$$

where the last inequality is a consequence of (2.3.10). Again using (2.3.10) we get

$$E\|B_t\|^2 \le K^2\int_0^t E\left\|X_s^{(k)}-X_s^{(k-1)}\right\|^2 ds$$

and therefore using $|a+b|^2\le 2(|a|^2+|b|^2)$ we get

$$E\left[\max_{0\le s\le t}\left\|X_s^{(k+1)}-X_s^{(k)}\right\|^2\right]\le L\int_0^t E\left\|X_s^{(k)}-X_s^{(k-1)}\right\|^2 ds; \quad 0\le t\le T \tag{2.3.18}$$

Iterating the inequality (2.3.18) we get to yield successive upper bounds we get

$$E\left[\max_{0\le s\le t}\left\|X_s^{(k+1)}-X_s^{(k)}\right\|^2\right]\le C^*\frac{(Lt)^k}{k!}; \quad 0\le t\le T \tag{2.3.19}$$

where $C^*=\max_{0\le t\le T}E\left\|X_t^{(1)}-\xi\right\|^2$ a finite quantity because of (2.3.15) and (2.3.12) . Relation (2.3.19) and the Chebyshev inequality now give

$$P\left[\max_{0\le s\le t}\left\|X_s^{(k+1)}-X_s^{(k)}\right\|^2>\frac{1}{2^{k+1}}\right]\le 4C^*\frac{(4Lt)^k}{k!}; \quad k=1.2,\dots, \tag{2.3.20}$$

and this upper bound is a general term in a convergent series. From the Borel-Cantelli lemma, we conclude that there exists an event $\Omega^* \in \mathcal{F}$ with $P(\Omega^*) = 1$ and an integer-valued random variable $N(\omega)$ such that for every $\omega \in \Omega^*$: $\max_{0\leq t\leq T} \left\| X_t^{(k+1)}(\omega) - X_t^{(k)}(\omega) \right\| < 2^{-(k+1)}, \forall k \geq N(\omega)$.Consequently

$$\max_{0\leq t\leq T} \left\| X_t^{(k+m)}(\omega) - X_t^{(k)}(\omega) \right\| < 2^{-k}, \forall m \geq 1, k \geq N(\omega) \tag{2.3.21}$$

We see that the sequence of sample paths $\{X_t^{(k)}(\omega); 0 \leq t \leq T\}_{k=1}^{\infty}$ is convergent in the supremum norm on continuous functions, from which follows the existence of a continuous limit $\{X_t(\omega); 0 \leq t \leq T\}$ for all $\omega \in \Omega^*$. Since T is arbitrary, we have the existence of a continuous process $X = \{X_t; 0 \leq t < \infty\}$ with the property that for P-a.e. ω , the sample paths $X = \{X^{(k)}(\omega)\}_{k=1}^{\infty}$ converge to $X(\omega)$, uniformly on compact subsets of $[0, \infty)$. Inequality (2.3.13) is an immediate consequence of (2.3.15) and Fatou's lemma. From (2.3.13) and (2.3.11) we have condition $iii)$ of Definition 2.3.1. Conditions $i)$ and $ii)$ are clearly satisfied by X. For condition $iv)$ of Definition 2.3.1 refer to Problem 5.2.11 on page 290 in [12]. □

In the one-dimensional case the Lipschitz condition was relaxed considerably by Yamada and Watanabe in 1971(see [10]). They proved pathwise uniqueness of solutions which implies the existence of a unique strong solution via weak existence(see [19] and [20]).

Theorem 2.3.7. *(Yamada and Watanabe 1971) Let us suppose that the coefficients of the one-dimensional equation(d=r=1)*

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t$$

satisfy the conditions

$$|b(t, x) - b(t, y)| \leq K(|x - y|) \tag{2.3.22}$$

$$|\sigma(t, x) - \sigma(t, y)| \leq h(|x - y|) \tag{2.3.23}$$

for every $0 \leq t < \infty$ and $x \in \mathbb{R}, y \in \mathbb{R}$, where K is a positive constant and we assume that $h : [0, \infty] \mapsto [0, \infty)$ is strictly increasing and concave with $h(0) = 0$ and $\int_{(0,\epsilon)} (du/h^2(u)) = \infty$ for every $\epsilon > 0$. Then strong uniqueness holds for equation (2.3.1)

Example 2.3.8. *One can take the function h in this proposition to be $h(u) = u^{\alpha}$ for $\alpha \geq (1/2)$*

Proof. Because of the conditions imposed on the function h, there exists a decreasing sequence $\{a_n\}_{n=1}^{\infty} \subseteq (0, 1]$ with $a_0 = 1$, $\lim_{n\to\infty} a_n = 0$ and $\int_{a_n}^{a_{n-1}} h^{-2}(u)du = n$ for every

$n \geq 1$. For each $n \geq 1$, there exists a continuous function ρ_n on $\mathbb{R}$ with support in (a_n, a_{n-1}) so that $0 \leq \rho_n(x) \leq (2/nh^2(x))$ holds for every $x > 0$ and $\int_{a_n}^{a_{n-1}} \rho_n(x)dx = 1$. Then the function

$$\psi_n(x) = \int_0^{|x|}\int_0^y \rho_n(u)dudy; x \in \mathbb{R} \tag{2.3.24}$$

is even and twice continuously differentiable with $|\psi_n'(x)| \leq 1$ and $\lim_{n\to\infty} \psi_n(x) = |x|$. Furthermore the sequence $\{\psi_n\}_{n=1}^{\infty}$ is non-decreasing. Now let us suppose that there are two strong solutions $X^{(1)}$ and $X^{(2)}$ of (2.3.1) with $X_0^{(1)} = X_0^{(2)}$ a.s.

It suffices to prove the indistinguishability of $X^{(1)}$ and $X^{(2)}$ under the assumption

$$E\int_0^t |\sigma(s, X_s^{(i)})|^2 ds < \infty; 0 \leq t < \infty, i = 1, 2, \tag{2.3.25}$$

otherwise, we may use condition (iii) of Definition 2.3.1 and a localization argument to reduce the situation to one in which (2.3.25) holds. We have

$$\Delta_t := X_t^{(1)} - X_t^{(2)} = \int_0^t \{b_1(s, X_s^{(1)}) - b_2(s, X_s^{(2)})\}ds + \int_0^t \{\sigma(s, X_s^{(1)}) - \sigma(s, X_s^{(2)})\}dW_s \tag{2.3.26}$$

$$\begin{aligned}\psi_n(\Delta_t) = \int_0^t \psi_n'(\Delta_s)[b_1(s, X_s^{(1)}) - b_2(s, X_s^{(2)})]ds + \int_0^t \psi_n'(\Delta_s)[\sigma(s, X_s^{(1)}) - \sigma(s, X_s^{(2)})]dW_s \\ + \frac{1}{2}\int_0^t \psi_n''(\Delta_s)[\sigma(s, X_s^{(1)}) - \sigma(s, X_s^{(2)})]]^2 ds\end{aligned} \tag{2.3.27}$$

Now taking expectation and recalling that the expectation of the stochastic integral is zero and as a consequence of assumption 2) the third integral is bounded above by $E\int \psi_n''(\Delta_s)h^2(|\Delta_s|)ds \leq 2t/n$. We can therefore conclude that

$$\begin{aligned}E\psi_n(\Delta_t) &\leq E\int_0^t \psi_n'(\Delta_s)[b_1(s, X_s^{(1)}) - b_2(s, X_s^{(2)})]ds + t/n \\ &\leq KE|\Delta_s| + t/n \quad t \geq 0, n \geq 0\end{aligned} \tag{2.3.28}$$

A passage to the limit as $n \to \infty$ yields $E|\Delta_t| \leq E\int_0^t |\Delta_s|ds; \quad t \geq 0$ and the conclusion follows from Gronwall's inequality and sample path continuity. □

Example 2.3.9. *(Girsanov 1962) From what we have just proved, it follows that strong uniqueness holds for the one-dimensional stochastic equation*

$$X_t = \int_0^t |X_s|^\alpha dW_s; \quad 0 \leq t < \infty \tag{2.3.29}$$

as long as $\alpha \geq (1/2)$ *and it is obvious that the unique solution is the trivial one* $X_t \equiv 0$. *This is also the solution when* $0 < \alpha < (1/2)$ *but is no longer the only solution.*

Proposition 2.3.10. *Suppose that on a certain probability space* $(\Omega, \mathcal{F}, P)$ *equipped with a filtration* $\{\mathcal{F}_t\}$ *which satisfies the usual conditions, we have standard, one-dimensional Brownian motion* $\{W_t; \mathcal{F}_t; 0 \leq t < \infty\}$ *and two continuous, adapted processes* $X^{(j)}, j = 1, 2$ *such that for* $\tau_0^{X^{(1)}} = \inf\{t \geq 0 : X^{(1)}(t) = 0\}$

$$X_t^{(j)} = X_0^{(j)} + \int_0^t b_j(X_s^j)ds + \int_0^t \sigma(X_s^j)dW_s;\ \ 0 \leq t < \tau_0^{X^{(1)}} \tag{2.3.30}$$

holds a.s. for $j = 1, 2$. *We assume that*

1. *the coefficients* $\sigma(x), b_j(x)$ *are continuous, real-valued functions on* $\mathbb{R}$
2. *the dispersion matrix* $\sigma(x)$ *satisfies the condition*

$$|\sigma(x) - \sigma(y)| \leq h(|x - y|)$$

where $h : [0, \infty) \to [0, \infty)$ *is a strictly increasing function with* $h(0) = 0$ *and it satisfies the following condition*

$$\int_{(0,\epsilon)} h^{-2}(u)du = \infty, \forall \epsilon > 0$$

3. $X_0^{(1)} = X_0^{(2)}$ *a.s.,*
4. $b_1(x) \leq b_2(x), x \in \mathbb{R}$
5. *either* $b_1(x)$ *or* $b_2(x)$ *satisfies the following condition*

$$|b_i(x) - b_i(y)| \leq K|x - y| \text{ for } i = 1, 2$$

Then

$$P[X_t^{(1)} \leq X_t^{(2)}, \forall 0 \leq t < \tau_0^{X^{(1)}}] = 1$$

Proof. For concreteness let us assume that $b_1(x)$ satisfies condition 5). We assume

$$E\int_0^t |\sigma(X_s^{(i)})|^2 ds < \infty; 0 \leq t < \infty, i = 1, 2, \tag{2.3.31}$$

otherwise we may use a localization argument to reduce the situation to the one in (2.3.31). We have for $0 \leq t < \tau_0^{X^{(1)}}$

$$\Delta_t := X_t^{(1)} - X_t^{(2)} = \int_0^t \{b_1(X_s^{(1)}) - b_2(X_s^{(2)})\}ds + \int_0^t \{\sigma(X_s^{(1)}) - \sigma(X_s^{(2)})\}dW_s \quad (2.3.32)$$

Because of the conditions imposed on the function h, there exists a decreasing sequence $\{a_n\}_{n=1}^{\infty} \subseteq (0,1]$ with $a_0 = 1$, $\lim_{n\to\infty} a_n = 0$ and $\int_{a_n}^{a_{n-1}} h^{-2}(u)du = n$ for every $n \geq 1$. For each $n \geq 1$, there exists a continuous function ρ_n on $\mathbb{R}$ with support in (a_n, a_{n-1}) so that $0 \leq \rho_n(x) \leq (2/nh^2(x))$ holds for every $x > 0$ and $\int_{a_n}^{a_{n-1}} \rho_n(x)dx = 1$. Then the function

$$\psi_n(x) = \int_0^{|x|} \int_0^y \rho_n(u)dudy; x \in \mathbb{R} \quad (2.3.33)$$

is even and twice continuously differentiable with $|\psi_n'(x)| \leq 1$ and $\lim_{n\to\infty} \psi_n(x) = |x|$. Furthermore the sequence $\{\psi_n\}_{n=1}^{\infty}$ is non-decreasing. Now we create a new sequence of auxillary functions $\varphi_n(x) = \psi_n(x)1_{(0,\infty)}(x)$. From a simple application of Itô's rule we get for $0 \leq t < \tau_0^{X^{(1)}}$

$$\varphi_n(\Delta_t) = \int_0^t \varphi_n'(\Delta_s)[b_1(X_s^{(1)}) - b_2(X_s^{(2)})]ds + \int_0^t \psi_n'(\Delta_s)[\sigma(X_s^{(1)}) - \sigma(X_s^{(2)})]dW_s$$
$$+\frac{1}{2}\int_0^t \varphi_n''(\Delta_s)[\sigma(X_s^{(1)}) - \sigma(X_s^{(2)})]]^2 ds \quad (2.3.34)$$

Now taking expectation and recalling that the expectation of the stochastic integral is zero and as a consequence of assumption 2) the third integral is bounded above by $E\int \varphi_n''(\Delta_s)h^2(|\Delta_s|)ds \leq 2t/n$. We can therefore conclude that

$$E[\varphi_n(\Delta_t)] - \frac{t}{n} \leq E\left[\int_0^t \varphi_n'(\Delta_s)[b_1(X_s^{(1)}) - b_2(X_s^{(2)})]ds\right]$$
$$= E\left[\int_0^t \varphi_n'(\Delta_s)[b_1(X_s^{(1)}) - b_1(X_s^{(2)})]ds\right] + E\left[\int_0^t \varphi_n'(\Delta_s)[b_1(X_s^{(2)}) - b_2(X_s^{(2)})]ds\right]$$

Now using the fact that $\{\psi_n\}_{n=1}^{\infty}$ is non- decreasing we get that $\{\varphi_n\}_{n=1}^{\infty}$ is non-decreasing and hence φ_n' is non-negative. We can therefore conclude that for all $0 \leq s < \tau_0^{X^{(1)}}$ we have $E\left[\int_0^t \varphi_n'(\Delta_s)[b_1(X_s^{(2)}) - b_2(X_s^{(2)})]ds\right] \leq 0$ and hence we can conclude using assumption (5) and $\varphi_n(x) = \psi_n(x)1_{(0,\infty)}(x)$ that for $0 \leq t < \tau_0^{X^{(1)}}$

$$E[\varphi_n(\Delta_t)] - \frac{t}{n} \leq E\left[\int_0^t \varphi_n'(\Delta_s)[b_1(X_s^{(1)}) - b_1(X_s^{(2)})]ds\right]$$
$$= E\left[\int_0^t \varphi_n'(X_s^{(1)} - X_s^{(2)})[b_1(X_s^{(1)}) - b_1(X_s^{(2)})]ds\right] \leq K\int_0^t E(\Delta_s^+)ds$$

since $\varphi_n'(x) = 0$ for $x \leq 0$ Now taking $n \to \infty$ we get

$$\lim_{n\to\infty} E[\varphi_n(\Delta_t)] - \frac{t}{n} \leq K \int_0^t E(\Delta_s^+)ds$$

and a simple application of Gronwall's inequality implies $E[\Delta_t^+] = 0$ for all $0 \leq t < \tau_0^{X^{(1)}}$ and hence $X_t^{(1)} \leq X_t^{(2)}$ for $0 \leq t < \tau_0^{X^{(1)}}$. □

2.4 Some important one dimensional SDEs

In this section I will introduce some important one dimensional SDE's in literature which are going to later play an important role in the subsequent chapters. The reference for the material of this section is the book of Mao and Cairns (see [21]) and [13] respectively)

2.4.1 The square root process

Anl close to the geometric Brownian motion is the square root process:

$$dr(t) = \mu r(t) + \sigma\sqrt{r(t)}dW(t) \tag{2.4.1}$$

Here the mean is made to follow an exponential trend while the standard deviation is made a function of the square root of $r(t)$. This makes the "variance" of the error term proportional to $r(t)$. Hence, if we are modeling asset prices using the SDE in (2.4.1), if asset price volatility does not increase "too much" when $r(t)$ increases (greater than 1, of course), this model may be more appropriate. For equation (2.4.1), one may ask whether $r(t)$ will become negative. If so, $r(t)$ would become a complex number and this would not make sense in most practical modeling situations. This is impossible and a simple proof can be found on page 307-308 of in the book of Mao([21]). A discussion about positivity of solutions of the SDE (2.4.1) is quite meaningless without ascertaining if the solutions actually exists. In the case of a square root one has the existence and uniqueness of a strong solution due to Theorem 2.3.7(path wise uniqueness) and weak existence(for more details see Yamada and Watanabe [10])

2.4.2 Mean Reverting Square Root Process

Combining the square root idea with the mean reverting one gives us the model of the mean reverting square root process:

$$dr(t) = \alpha(\mu - r(t)) + \sigma\sqrt{r(t)}dW(t) \tag{2.4.2}$$

This process has a unique strong solution because its coefficients satisfy the same properties as the coefficients of the square root process. This SDE is used in modeling the evolution of interest rate and is popularly known as the CIR SDE(Cox-Ross-Ingersoll see [22]). The parameter α corresponds to the speed of adjustment, μ , to the mean and σ to volatility. The drift factor, $\alpha(\mu - r(t))$ in the SDE (2.4.2) which is the same as in the Vasicek SDE(see [23]) ensures mean reversion of the interest rate towards the long run value μ, with speed of adjustment governed by the strictly positive parameter α. Just as in the case of the square root SDE, one can show that the solution of the mean reverting square root is almost surely non-negative. We conclude this section with a very important result due to Feller(see [24]). The detailed proof has been taken from Cairns (see [13]).

Theorem 2.4.1. *Given the one-dimensional square root mean reverting SDE*

$$dr_t = \alpha(\mu - r(t))dt + \sigma\sqrt{r(t)}dW(t)$$

Assume that $r(0) = r > 0$*, let* $U = \inf\{t : r(t) \leq 0\}$*(* *where* $\inf \phi = \infty$*)* *Then* $2\mu\alpha \geq \sigma^2 \implies Q(U = \infty) = 1$ *and* $2\mu\alpha < \sigma^2 \implies Q(U < \infty) = 1$ *where* Q *is the probability measure under which* W *is a Brownian motion.*

Proof. The key steps of the proof will be stated first followed by a detailed development filling out these initial statements

(i) Define the function $s(r) = \int_1^r e^{2\alpha v/\sigma^2} v^{-2\alpha\mu/\sigma^2} dv$ for $0 < r < \infty$. Then

$$\alpha(\mu - r)\frac{\partial s}{\partial r} + \frac{1}{2}\sigma^2 r \frac{\partial^2 s}{\partial r^2} = 0$$

(ii) For each t, and given $r(0) = r$,

$$s(r(t)) = s(r(0)) + \int_0^t \frac{ds}{dr}(r(u))\sigma\sqrt{r(u)}dW(u)$$

In particular $s(r(t))$ is a local a local martingale under Q.

(iii) Define $\tau_x = \inf\{t \geq 0 : r(t) = x\}$, and $p \wedge q = \inf\{p, q\}$. Let ϵ, M be such that $0 < \epsilon < r(0) < M < \infty$. Then we exploit the local-martingale properties of $s(r(t))$ and the boundedness of $\frac{ds}{dr}$, where $\epsilon < r < M$ to demonstrate that $Pr_Q(\tau_\epsilon \wedge \tau_M < \infty) = 1$

(iv) The martingale property then implies that

$$s(r(0)) = s(\epsilon)Pr_Q(\tau_\epsilon < \tau_M) + s(M)Pr_Q(\tau_\epsilon > \tau_M)$$

(v) If $2\alpha\mu \geq \sigma^2$ then $s(\epsilon) \to -\infty$ as $\epsilon \to 0$. This implies $Pr_Q(\tau_0 < \tau_M) = 0$ for all $0 < r(0) < M < \infty$. Hence $Pr_Q(\tau_0 < \infty) = 0$

(vi) If $0 < 2\alpha\mu < \sigma^2$ with $\alpha, \mu, \sigma^2 > 0$, then $-\infty < \lim_{\epsilon\to 0} s(\epsilon) < 0$. Hence $s(r(0)) = s(0)Pr_Q(\tau_0 < \tau_M) + s(M)Pr_Q(\tau_M < \tau_0)$, where $s(0)$ is defined as $\lim_{\epsilon\to 0} s(\epsilon)$. Since in addition, $s(M) \to +\infty$ as $M \to +\infty$, $Pr_Q(\tau_0 < \infty) = 1$.

Now we work through the steps more rigorously Consider the twice continuously differentiable function $s(r)$. By a simple application of Itô's lemma we get

$$\begin{aligned} ds(r(t)) &= \frac{\partial s}{\partial r} dr(t) + \frac{1}{2}\frac{\partial^2 s}{\partial r^2} d\langle r\rangle_t \\ &= \frac{\partial s}{\partial r}(\alpha(\mu - r(t))dt + \sigma\sqrt{r(t)}dW(t)) + \frac{1}{2}\frac{\partial^2 s}{\partial r^2}\sigma^2 r(t)dt \\ &= \frac{\partial s}{\partial r}(\alpha(\mu - r(t))dt + \frac{1}{2}\frac{\partial^2 s}{\partial r^2}\sigma^2 r(t)dt + \sigma\sqrt{r(t)}dW(t) \end{aligned}$$

Now since $r(t)$ is a continuous semi-martingale , Itô's lemma implies that $s(r(t))$ is a continuous semi-martingale and it follows from the definition of a continuous semi-martingale that $s(r(t))$ is a continuous local martingale iff the drift term is zero i.e

$$\frac{\partial s}{\partial r}(\alpha(\mu - r(t)) + \frac{1}{2}\frac{\partial^2 s}{\partial r^2}\sigma^2 r(t) = 0$$

Now with drift equal to zero, the new SDE becomes

$$ds(r(t)) = s'(r(t))\sigma\sqrt{r(t)}dW(t)$$

which can be reformulated in the integral form as

$$s(r(t)) = s(r(0)) + \int_0^t s'(r(t))\sigma\sqrt{r(t)}dW(u)$$

where $s'(r(t)) = \frac{ds}{dr} = \frac{d}{dr}\int_1^r e^{2\alpha v/\sigma^2} v^{-2\alpha\mu/\sigma^2} dv = e^{2\alpha r/\sigma^2} r^{-2\alpha\mu/\sigma^2}$ and hence in this form $s(r(t))$ is a continuous local martingale. For $0 < \epsilon < r(0) < M < \infty$ and since $s'(r) = e^{2\alpha r/\sigma^2} r^{-2\alpha\mu,/\sigma^2}$ is positive for all $r > 0$, $s(r)$ is non decreasing and hence it follows that $s(\epsilon) \leq s(r) \leq s(M)$ for $\epsilon \leq r \leq M$.

It not so hard to see that for $0 < \epsilon < r(0) < M$, $s'(r)$ is bounded below by $\delta = M^{-2\alpha\mu/\sigma^2}$ where $\delta > 0$. Indeed for for $r \geq 0, \alpha, \mu, \sigma^2 > 0$ we have $1 \leq e^{2\alpha r/\sigma^2}$ and $M^{-2\alpha\mu/\sigma^2} \leq r^{-2\alpha\mu/\sigma^2}$ which together yield
$M^{-2\alpha\mu/\sigma^2} \leq e^{2\alpha r/\sigma^2} r^{-2\alpha\mu/\sigma^2}$

Let us now consider the stopped process

$$s(r(t \wedge \tau_\epsilon \wedge \tau_M)) = s(r(0)) + \int_0^t I(u)s'(r(u))\sigma\sqrt{r(u)}dW(u)$$

where $\tau_x = \inf\{t \geq 0 : r(t) = x\}, p \wedge q = \inf\{p, q\}$ and

$$I(u) = \begin{cases} 1 & u < \tau_\epsilon \wedge \tau_M \\ 0 & u \geq \tau_\epsilon \wedge \tau_M \end{cases}$$

Now $s(r(t \wedge \tau_\epsilon \wedge \tau_M))$ is not just a local martingale but a martingale since $I(u)s'(r(u))\sigma\sqrt{r(u)}$ is bounded (we have already established the boundedness of $s'(r)$) since $0 \leq I(u) \leq 1$ ans $\sigma\sqrt{\epsilon} \leq \sigma\sqrt{r(u)} \leq \sigma\sqrt{M}$.

And since $\int_0^t I(u)s'(r(u))\sigma\sqrt{r(u)}dW(u)$ is a martingale starting at 0, taking expectations we get

$$E_Q\left(s(r(t \wedge \tau_\epsilon \wedge \tau_M))\right) = E_Q(s(r(0))) + E_Q\left(\int_0^t I(u)s'(r(u))\sigma\sqrt{r(u)}dW(u)\right)$$

which finally yields

$$s(r(0)) = E_Q\left(s(r(t \wedge \tau_\epsilon \wedge \tau_M))\right)$$

$$\begin{aligned} Var_Q(s(r(t \wedge \tau_\epsilon \wedge \tau_M)) = E_Q\left[(s(r(t \wedge \tau_\epsilon \wedge \tau_M) - E_Q[s(r(t \wedge \tau_\epsilon \wedge \tau_M)])^2\right] = \\ E_Q\left[\left(\int_0^t I(u)s'(r(u))\sigma\sqrt{r(u)}dW(u)^2\right)\right] = E_Q\left[\int_0^{t \wedge \tau_\epsilon \wedge \tau_M} s'(r(u))^2\sigma^2 r(u)du\right] \\ \geq \delta^2\sigma^2\epsilon E_Q\left[\int_0^{t \wedge \tau_\epsilon \wedge \tau_M} du\right] = \delta^2\sigma^2\epsilon E_Q[t \wedge \tau_\epsilon \wedge \tau_M] \end{aligned} \tag{2.4.3}$$

where the third equality is a consequence of Itô's Isometry. But we also have $Var_Q(s(r(t \wedge \tau_\epsilon \wedge \tau_M)) \leq (s(M) - s(\epsilon))^2 < \infty$ since the random variable takes values in $[s(\epsilon), s(M)]$. Hence we have that $\delta^2\sigma^2\epsilon E_Q[t \wedge \tau_\epsilon \wedge \tau_M] \leq (s(M) - s(\epsilon))^2 < \infty$ for all $t \geq 0$, implying that

$$E_Q[t \wedge \tau_\epsilon \wedge \tau_M] < \infty \text{ for all } t \geq 0$$

and monotone convergence theorem implies $E[\tau_\epsilon \wedge \tau_M] < \infty$ and, therefore , that $Pr_Q[\tau_\epsilon \wedge \tau_M < \infty] = 1$. Now

$$\begin{aligned} s(r(0)) = E_Q[s(r(t \wedge \tau_\epsilon \wedge \tau_M))] \\ = E_Q[s(r(t \wedge \tau_\epsilon \wedge \tau_M))(1_{\tau_\epsilon \leq t \wedge \tau_M} + 1_{\tau_M \leq t \wedge \tau_\epsilon} + 1_{t < \tau_\epsilon \wedge \tau_M})] \\ = s(\epsilon)Pr_Q(\tau_\epsilon \leq t \wedge \tau_M) + \\ s(M)Pr_Q(\tau_M \leq t \wedge \tau_\epsilon) + E_Q[s(r(t \wedge \tau_\epsilon \wedge \tau_M))1_{t < \tau_\epsilon \wedge \tau_M}] \\ = s(\epsilon)Pr_Q(\tau_\epsilon \leq t \wedge \tau_M) + \\ s(M)Pr_Q(\tau_M \leq t \wedge \tau_\epsilon) + \\ E_Q[s(r(t \wedge \tau_\epsilon \wedge \tau_M) \mid t < \tau_\epsilon \wedge \tau_M]Pr_Q(t < \tau_\epsilon \wedge \tau_M)] \end{aligned} \tag{2.4.4}$$

Now in the limit $t \to \infty$, $E_Q[s(r(t)) \mid t < \tau_\epsilon \wedge \tau_M]$ is bounded below and above by $s(\epsilon)$ and $s(M)$ respectively, while $Pr_Q(t < \tau_\epsilon \wedge \tau_M) \to 0$ since $E[\tau_\epsilon \wedge \tau_M] < \infty$, $Pr_Q(\tau_\epsilon < t \wedge \tau_M) \to Pr_Q(\tau_\epsilon < \tau_M)$ and $Pr_Q(\tau_M < t \wedge \tau_\epsilon) \to Pr_Q(\tau_M < \tau_\epsilon)$ because probability measures are continuous from below and above.

Hence we have

$$s(r(0)) = s(\epsilon)Pr_Q(\tau_\epsilon < \tau_M) + s(M)Pr_Q(\tau_M < \tau_\epsilon)$$

Now suppose that $2\alpha\mu \geq \sigma^2$ and $0 < \epsilon < 1$. Then

$$-s(\epsilon) = \int_\epsilon^1 e^{2\alpha v/\sigma^2} v^{-2\alpha\mu/\sigma^2} dv \geq \int_\epsilon^1 e^{2\alpha v/\sigma^2} \frac{1}{v} dv \geq \int_\epsilon^1 \frac{1}{v} dv \to \infty \text{ as } \epsilon \to 0$$

since for $v \in (\epsilon, 1)$ with $\epsilon > 0$ and $2\alpha\mu \geq \sigma^2$ we have

$$\frac{1}{v} \leq \frac{1}{v} e^{2\alpha v/\sigma^2} \leq v^{-2\alpha\mu/\sigma^2} e^{2\alpha v/\sigma^2}$$

So we have the two results $s(r(0)) = s(\epsilon)Pr_Q(\tau_\epsilon < \tau_M) + s(M)Pr_Q(\tau_M < \tau_\epsilon)$ and $s(\epsilon) \to -\infty$ as $\epsilon \to 0$. Hence it follows that for a fixed M, as $\epsilon \to 0$ we must have $Pr_Q(\tau_\epsilon < \tau_M) \to 0$. And since $\{\tau_{\epsilon_1} < \tau_M\} \subseteq \{\tau_{\epsilon_2} < \tau_M\}$ for $\epsilon_2 < \epsilon_1$ and continuity from above implies $Pr_Q(\tau_0 < \tau_M) = 0$ for all M such that $0 < r(0) < M < \infty$.Now consider the event that $r(t)$ hits zero in finite time:

$$\Omega_0 = \left\{ \omega : \tau_0(\omega) < \infty, \sup_{0<t<\tau_0(\omega)} r(t)(\omega) < \infty \right\}$$

Note that we have excluded from Ω_0 the sample path , $r(t)(\omega)$, which explode before $\tau_0(\omega)$; that is

$$\Omega_e = \left\{ \omega : \tau_0(\omega) < \infty, \sup_{0<t<\tau_0(\omega)} r(t)(\omega) = \infty \right\}$$

Theorem 2.4 on page 177 of Ikeda and Watanabe's Stochastic differential equations and diffusion processes(second edition) implies that the stochastic differential equation $dr_t = \alpha(\mu - r(t))dt + \sigma\sqrt{r(t)}dW(t)$ does not explode with probability 1. And thus $Pr_Q(\Omega_e) = 0$.

Now let , for integers n,

$$\Omega_n = \left\{ \omega : \tau_0(\omega) < \tau_n(\omega), \sup_{0<t<\tau_0(\omega)} r(t)(\omega) < \infty \right\}$$

Clearly $\tau_{r(0)} \in (\tau_0, \tau_n)$ and hence $\tau_n \leq \tau_{n+1}$ which allows us to immediately conclude that for all $n \in \mathbb{N}$ we have $\{\tau_0 < \tau_n\} \subseteq \{\tau_0 < \tau_{n+1}\}$ and hence $\Omega_n \subseteq \Omega_{n+1}$.

For each $\omega \in \Omega_0$ there exists $n_0(\omega)$ such that $\tau_n(\omega) > \tau_0(\omega)$ for all $n \geq n_0(\omega)$ where $n_0(\omega) := \sup_{0<t<\tau_0(\omega)} r(t)(\omega)$. This statement can be mathematically expressed as

$$\Omega_0 = \bigcup_{n=1}^{\infty} \Omega_n$$

But since $Pr_Q(\tau_0 < \tau_M) = 0$ for all $\infty > M > r(0) > 0$

$$Pr_Q(\Omega_n) \leq Pr_Q(\{\omega : \tau_0(\omega) < \tau_n(\omega)\}) = 0$$

Countable sub-additivity implies

$$Pr(\Omega_0) \leq \sum_{n=}^{\infty} Pr_Q(\Omega_n) = 0$$

and finally $Pr_Q(\tau_0 < \infty) = Pr_Q(\Omega_e) + Pr_Q(\Omega_0) = 0$

Now we study the case where $0 < 2\alpha\mu < \sigma^2$. Now for $0 < \epsilon < 1$

$$0 > s(\epsilon) = -\int_{\epsilon}^{1} e^{2\alpha v/\sigma^2} v^{-2\alpha\mu/\sigma^2} dv > -\int_{\epsilon}^{1} e^{2\alpha/\sigma^2} v^{-2\alpha\mu/\sigma^2} dv$$

because for $v \in (\epsilon, 1)$ we have $e^{2\alpha/\sigma^2} > e^{2\alpha v/\sigma^2}$

But the limit as $\epsilon \to 0$ of $\int_{\epsilon}^{1} v^{-2\alpha\mu\sigma^2} dv$ is finite since $2\alpha\mu/\sigma^2 < 1$. Thus , the limit as ϵ tends to zero of $s(\epsilon)$ lies strictly between $-\infty$ and 0. Define

$$s(0) = \lim_{\epsilon \to 0} s(\epsilon)$$

We have already shown earlier that for $0 < \epsilon < r(0) < M < \infty$

$$s(r(0)) = s(\epsilon) Pr_Q(\tau_\epsilon < \tau_M) + s(M) Pr_Q(\tau_M < \tau_\epsilon)$$

We now modify out arguments above to show that this is true for $s(\epsilon)$ replaced by $s(0)$(which we just defined above) when $\frac{1}{2} \leq 2\alpha\mu/\sigma^2 < 1$

Recall that from equation (2.4.3)

$$Var_Q(s(r(t \wedge \tau_\epsilon \wedge \tau_M)) = E_Q\left[\int_0^{t\wedge\tau_\epsilon\wedge\tau_M} s'(r(u))^2 \sigma^2 r(u) du\right]$$

Now recalling that $s'(r) = e^{2\alpha r/\sigma^2} r^{-2\alpha\mu/\sigma^2}$ we get

$$Var_Q(s(r(t \wedge \tau_\epsilon \wedge \tau_M)) = E_Q\left[\int_0^{t\wedge\tau_\epsilon\wedge\tau_M} e^{4\alpha r(u)/\sigma^2} r(u)^{-4\alpha\mu/\sigma^2+1} \sigma^2 du\right]$$

Let

$$f(r) = s'(r)^2 r = e^{4\alpha r/\sigma^2} r^{1-d} = e^{dr/\mu} r^{1-d}$$

where $d = 4\alpha\mu/\sigma^2$. We have already specified that $\frac{1}{2} \leq 2\alpha\mu/\sigma^2 < 1$ which is the same as $1 \leq d < 2$. Now if $1 < d < 2$, then $f(r)$ is minimized(the minimum of this continuously differentiable function can be computed by setting the derivative equal to zero) in the range $0 \leq r < \infty$ at $\hat{r} = (d-1)\mu/d$ with

$$f(\hat{r}) = e^{d-1} \left(\frac{(d-1)\mu}{d} \right)^{1-d} > 0$$

If $d = 1$, then $f(r)$ is minimized in the range $0 \leq r < \infty$ at $\hat{r} = 0$ with $f(0) = 1$ Let δ be the minimum value of $f(\hat{r})$ in either case. Hence

$$\sigma^2 s'(r(u))^2 r(u) \geq \sigma^2 \delta \text{ for all } 0 < u < t \wedge \tau_0 \wedge \tau_M$$

which yields the following inequality

$$Var_Q(s(r(t \wedge \tau_\epsilon \wedge \tau_M)) = E_Q \left[\int_0^{t \wedge \tau_\epsilon \wedge \tau_M} s'(r(u))^2 \sigma^2 r(u) du \right] \geq \sigma^2 \delta E_Q[t \wedge \tau_0 \wedge \tau_M]$$

But as we saw before $Var_Q(s(r(t \wedge \tau_\epsilon \wedge \tau_M)) \leq (s(M) - s(\epsilon))^2$, so

$$E_Q[t \wedge \tau_0 \wedge \tau_M] \leq \frac{(s(M) - s(\epsilon))^2}{\sigma^2 \delta} < \infty \text{ for all } t,$$

And hence the monotone convergence theorem implies

$$E_Q[\tau_0 \wedge \tau_M] < \infty$$

and therefore we can conclude

$$Pr_Q(\tau_0 \wedge \tau_M < \infty) = 1$$

And hence taking the limits as $t \to \infty$ in

$$s(r(0)) = E_Q[s(r(t \wedge \tau_0 \wedge \tau_M))] = E_Q[s(r(t \wedge \tau_0 \wedge \tau_M))(1_{\tau_0 \leq t \wedge \tau_M} + 1_{\tau_M \leq t \wedge \tau_0} + 1_{t < \tau_0 \wedge \tau_M})]$$

just as in (2.4.4) we get

$$s(r(0)) = s(0) Pr_Q(\tau_0 < \tau_M) + s(M) Pr_Q(\tau_M < \tau_0)$$

Moreover since $e^{2\alpha v/\sigma^2} v^{-2\alpha\mu/\sigma^2} \to \infty$ as $v \to \infty$, and therefore so does $s(M) = \int_1^M e^{2\alpha v/\sigma^2} v^{-2\alpha\mu}$ ∞ as $M \to \infty$. As $s(r(0))$ is finite , we must have $Pr_Q[\tau_M < \tau_0] \to 0$ and hence $Pr_Q[\tau_0 < \tau_M] \to 1$ as $M \to \infty$. Thus we get $Pr_Q(\tau_0 < \infty) = 1$.

Finally, suppose that $0 < 2\alpha\mu/\sigma^2 < \frac{1}{2}$(or equivalently $0 < d < 1$). Let $X(t) = \sqrt{r(t)}$. After a simple application of Itô's Lemma and substituting $X(t) = \sqrt{r(t)}$ we get

$$dX_t = \frac{1}{2\sqrt{r(t)}}dr(t) + \frac{1}{2}\frac{-1}{4r(t)^{3/2}}d\langle r\rangle(t)$$
$$= \frac{1}{2X(t)}(\alpha(\mu - X(t)^2))dt + \sigma X(t)dW(t)) - \frac{1}{8X(t)^3}\sigma^2 X(t)^2$$
$$= \left[\frac{\sigma^2}{4X(t)}\left(\frac{2\alpha\mu}{\sigma^2} - \frac{1}{2}\right) - \frac{\alpha X(t)}{2}\right]dt + \frac{\sigma}{2}dW(t)$$

Now let $Y(t)$ be the Ornstein-Uhlenbeck process governed by the SDE :

$$dY(t) = -\frac{1}{2}\alpha Y(t)dt + \frac{1}{2}\sigma dW(t), Y(0) = X(0) \tag{2.4.5}$$

Define

$$\tau_0^X = \inf\{t \geq 0 : X(t) = 0\} \text{ and } \tau_0^Y = \inf\{t \geq 0 : Y(t) = 0\}$$

For each outcome, ω, for all $0 < t < \tau_0^X(\omega)$,

$$\frac{\sigma^2}{8X(t)(\omega)}(d-1) - \frac{\alpha X(t)(\omega)}{2} < -\frac{\alpha X(t)(\omega)}{2}$$

This is because $\frac{\sigma^2}{8X(t)(\omega)}(d-1)$ is strictly negative for $0 < d < 1$ for all $0 < t < \tau_0^X(\omega)$ since $X(t), \sigma^2$ is strictly positive for these t and all outcomes ω. Now it follows from Proposition 5.2.18 in Brownian Motion and Stochastic calculus by Karatzas and Shreve(modified with $0 < t < \tau_0^X$) that $X(t) \leq Y(t)$ for $0 < t < \tau_0^X$. A detailed proof is given in Proposition 2.3.10

This implies that if we can show that $\tau_0^Y < \infty$ then we can immediately conclude $\tau_0^X < \infty$ and therefore $\tau_0 < \infty$ a.s. However we know from the basic properties of the Ornstein-Uhlenbeck process that $Pr_Q(\tau_0^Y < \infty) = 1$. Hence this implies $Pr_Q(\tau_0^X < \infty) = 1$:, that is the same as saying $r(t)$ will hit zero with probability 1 under Q. □

Before I end this section I would like to mention that the square-root type SDEs dealt above and in the following two chapters are the subject of recent research on SDEs with non-Lipschitz coefficients. A possible avenue for further study could be to establish a relationship between the existence results in the following two chapters to the many existence and approximation results by [25],[26], [27] and [28].

Chapter 3

On a class of stochastic differential equations with random and Hölder continuous coefficients arising in biological modeling

3.1 Introduction

Susceptible-infected-susceptible (SIS) epidemic model is one of the most popular models for how diseases spread in a population. In such a model an individual starts off being susceptible to a disease and at some point of time gets infected and then recovers after some time becoming susceptible again. The literature of such mathematical models is very rich: for probabilistic/stochastic models one may look for instance at Allen [29], Allen and Burgin [30], A. Gray et al. [31], Hethcote and van den Driessche [32], Kryscio and Lefvre [33], McCormack and Allen [34] and Nasell [35]. We also refer the reader to the detailed account presented in Greenhalgh et al. [1] for an overview on both deterministic and stochastic models.

The focus of the present paper is on the model presented in [1]. One of its distinguishing features is the nature of the births and deaths that are regarded as stochastic processes with per capita disease contact rate depending on the population size. Contrary to many other previously proposed models, this stochasticity produces a variable population size which turns out to be a reasonable assumption for slowly spreading diseases.

From a mathematical point of view, the SIS model proposed in [1] amounts at the following two dimensional stochastic differential equation for the vector (S_t, I_t) where S_t

and I_t stand for the number of susceptible and infected individuals at time t, respectively:

$$\begin{cases} dS &= \left[-\frac{\lambda(N)SI}{N} + (\mu+\gamma)I\right]dt + \sqrt{\frac{\lambda(N)SI}{N} + (\mu+\gamma)I + 2\mu S}dW_3 \\ dI &= \left[\frac{\lambda(N)SI}{N} - (\mu+\gamma)I\right]dt + \sqrt{\frac{\lambda(N)SI}{N} + (\mu+\gamma)I}dW_4. \end{cases} \tag{3.1.1}$$

Here, $N := S + I$ denotes the total population size while μ, γ and $\lambda : [0, +\infty[\to [0, +\infty[$ are suitably chosen parameters. The system (3.1.1) is driven by the two dimensional correlated Brownian motion (W_3, W_4) resulting from a certain application of the martingale representation theorem (see Section 3.2.1 below for technical details). The system (3.1.1) is then shown to be equivalent to the triangular system

$$\begin{cases} dI &= \left[\frac{\lambda(N)}{N}(N-I)I - (\mu+\gamma)I\right]dt + \sqrt{\frac{\lambda(N)}{N}(N-I)I + (\mu+\gamma)I}dW_4 \\ dN &= \sqrt{2\mu N}dW_5 \end{cases} \tag{3.1.2}$$

where now the second equation, the so-called square root process (see for instance the book by Mao [21] for the properties of this process), is independent of the first one. To prove the existence of a solution to the first equation in (3.1.2) the authors resort to Theorem 2.2 in Chapter IV of Ikeda and Watanabe [11] while for the uniqueness they need to construct a localized version of Theorem 3.2, Chapter IV in [11]. The equation for I in (3.1.2) exhibits random (for the dependence on the process N) and Hölder continuous (for the presence of the square root in the diffusion term) coefficients resulting in a stochastic differential equation for which the issue of the existence of a unique solution has not been addressed in the literature yet.

Our aim in the present paper is to propose a more general approach allowing for the investigation of a richer family of models characterized by the same distinguishing features of the model analyzed in [1].

The paper is articulated as follows: In Section 2 we present a general review using the exposition in the book by Allen (see [36]) of a two-state dynamics leading to a Fokker-Planck partial differential equation and its associated stochastic system. This is followed by Section 2.1 where we consider the more specific situation of a bio-demographic model like the one presented in [1]. Our idea is to embed the rather special system of SDE's of the model in a slightly more encompassing class, like the one in (3.3.9) below, in order to establish a general proof of strong existence and uniqueness. Our technique relies on the construction of an explicit approximating sequence of stochastic processes (inspired by the work of Zubchenko [37]) in such a way that all the relevant features of the solution appear to be directly constructed from scratch. In Section 3 we give a detailed proof of existence and uniqueness of the SDE (3.3.9). We would like to point out that systems of SDE's with non-Lipschitz or Hölder coefficients exhibit non-standard difficulties as far as general results for existence and uniqueness are concerned. This model conforms to the aforementioned difficulties and that is what has motivated us in approaching the problem. Our idea has been to how we could encase the model proposed in [1] within a more general framework , thus bypassing some of the computations done there, and hopefully allowing for larger class of models to be treated.

3.2 A general two-state system

In this section we review the construction of a general two-state system presented in the book by Allen ([36]). The model will then be made concrete through the assumptions contained in the paper by Greenhalgh et al. ([1]) and this will lead to the class of stochastic differential equations investigated in the present manuscript.

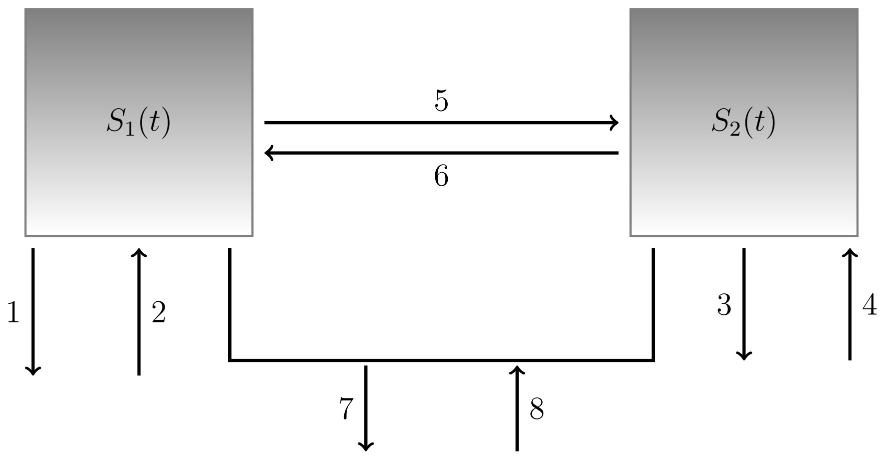

Figure 3.1: A two-state dynamical process

We begin by considering a representative two-state dynamical process which is illustrated in Figure 3.1. Let $S_1(t)$ and $S_2(t)$ represent the values of the two states of the system at time t. It is assumed that in a small time interval Δt, state S_1 can change by $-\lambda_1$, 0 or λ_1 and state S_2 can change by $-\lambda_2$, 0 or λ_2, where $\lambda_1, \lambda_2 \geq 0$. Let $\Delta S := [\Delta S_1, \Delta S_2]^T$ be the change in a small time interval Δt. As illustrated in Figure 3.1 , there are eight possible changes for the two states in the time interval Δt not including the case where there is no change in the time interval. The possible changes and the probabilities of these changes are given in Table 3.1. It is assumed that the probabilities are given to $O((\Delta t)^2)$. For example, change 1 represents a loss of λ_1 in S_1 with probability $d_1 \Delta t$, change 5 represents a transfer of λ_1 out of state S_1 with a corresponding transfer of λ_2 into state S_2 with probability $m_{12}\Delta t$ and change 7 represents a simultaneous reduction in both states S_1 and S_2. As indicated in the table, all probabilities may depend on $S_1(t)$, $S_2(t)$ and the time t. Also notice that it is assumed that the probabilities for the changes are proportional to the time interval Δt.

Table 3.1: Possible changes in the representative two-state system with the corresponding probabilities

Change	Probability
$\Delta\mathbf{S}^{(1)} = [-1, 0]^T$	$p_1 = d_1(t, S_1, S_2)\Delta t$
$\Delta\mathbf{S}^{(2)} = [1, 0]^T$	$p_2 = b_1(t, S_1, S_2)\Delta t$
$\Delta\mathbf{S}^{(3)} = [0, -1]^T$	$p_3 = d_2(t, S_1, S_2)\Delta t$
$\Delta\mathbf{S}^{(4)} = [0, 1]^T$	$p_4 = b_2(t, S_1, S_2)\Delta t$
$\Delta\mathbf{S}^{(5)} = [-1, 1]^T$	$p_5 = m_{12}(t, S_1, S_2)\Delta t$
$\Delta\mathbf{S}^{(6)} = [1, -1]^T$	$p_6 = m_{21}(t, S_1, S_2)\Delta t$
$\Delta\mathbf{S}^{(7)} = [-1, -1]^T$	$p_7 = m_{11}(t, S_1, S_2)\Delta t$
$\Delta\mathbf{S}^{(8)} = [1, 1]^T$	$p_8 = m_{22}(t, S_1, S_2)\Delta t$
$\Delta\mathbf{S}^{(9)} = [0, 0]^T$	$p_9 = 1 - \sum_{j=1}^{8} p_j$

It is useful to calculate the mean vector and covariance matrix for the change $\Delta S = [\Delta S_1, \Delta S_2]^T$ fixing the value of S at time t. Using the table below,

$$E[\Delta S] = \sum_{j=1}^{9} p_j \Delta S^{(j)} = \begin{bmatrix} (-d_1 + b_1 - m_{12} + m_{21} + m_{22} - m_{11})\lambda_1 \\ (-d_2 + b_2 + m_{12} - m_{21} + m_{22} - m_{11})\lambda_2 \end{bmatrix} \Delta t$$

$$\begin{aligned} E[\Delta S(\Delta S)^T] &= \sum_{j=1}^{9} p_j (\Delta S^{(j)})(\Delta S^{(j)})^T \\ &= \begin{bmatrix} (d_1 + b_1 + m_a)\lambda_1^2 & (-m_{12} - m_{21} + m_{22} + m_{11})\lambda_1\lambda_2 \\ (-m_{12} - m_{21} + m_{22} + m_{11})\lambda_1\lambda_2 & (d_2 + b_2 + m_a)\lambda_2^2 \end{bmatrix} \Delta t \end{aligned}$$

where we set $m_a := m_{12} + m_{21} + m_{11} + m_{22}$. Notice that the covariance matrix is set equal to $E(\Delta S(\Delta S)^T)/\Delta t$ because $E(\Delta S)(E(\Delta S))^T = O((\Delta t)^2)$. We now define

$$\mu(t, S_1, S_2) = E[\Delta S]/\Delta t \quad \text{and} \quad V(t, S_1, S_2) = E[\Delta S(\Delta S)^T]/\Delta t \tag{3.2.1}$$

and we denote by $B(t, S_1, S_2)$ the symmetric square root matrix of V. A forward Kolmogorov equation can be determined for the probability distribution at time $t + \Delta t$ in terms of the distribution at time t. If we write $p(t, x_1, x_2)$ for the probability that $S_1(t) = x_1$ and $S_2(t) = x_2$, then referring to Table 3.1 we get

$$p(t + \Delta t, x_1, x_2) = p(t, x_1, x_2) + \Delta t \sum_{i=1}^{10} T_i \tag{3.2.2}$$

where

$$\begin{aligned}
T_1 &= p(t,x_1,x_2)(-d_1(t,x_1,x_2) - b_1(t,x_1,x_2) - d_2(t,x_1,x_2) - b_2(t,x_1,x_2)) \\
T_2 &= p(t,x_1,x_2)(-m_a(t,x_1,x_2)) \\
T_3 &= p(t,x_1+\lambda_1,x_2)d_1(t,x_1+\lambda_1,x_2) \\
T_4 &= p(t,x_1-\lambda_1,x_2)b_1(t,x_1-\lambda_1,x_2) \\
T_5 &= p(t,x_1,x_2-\lambda_2)b_2(t,x_1,x_2-\lambda_2) \\
T_6 &= p(t,x_1,x_2+\lambda_2)d_2(t,x_1,x_2+\lambda_2) \\
T_7 &= p(t,x_1+\lambda_1,x_2-\lambda_2)m_{12}(t,x_1+\lambda_1,x_2-\lambda_2) \\
T_8 &= p(t,x_1-\lambda_1,x_2+\lambda_2)m_{21}(t,x_1-\lambda_1,x_2+\lambda_2) \\
T_9 &= p(t,x_1+\lambda_1,x_2+\lambda_2)m_{11}(t,x_1+\lambda_1,x_2+\lambda_2) \\
T_{10} &= p(t,x_1-\lambda_1,x_2-\lambda_2)m_{22}(t,x_1-\lambda_1,x_2-\lambda_2).
\end{aligned}$$

Now, expanding out the terms T_3 through T_{10} in second order Taylor polynomials around the point (t,x_1,x_2), it follows that

$$\begin{aligned}
T_3 &\approx pd_1 + \partial_{x_1}(pd_1)\lambda_1 + \frac{1}{2}\frac{\partial^2(pd_1)}{\partial x_1^2}\lambda_1^2 \\
T_4 &\approx pb_1 - \frac{\partial(pb_1)}{\partial x_1}\lambda_1 + \frac{1}{2}\frac{\partial^2(pb_1)}{\partial x_1^2}\lambda_1^2 \\
T_5 &\approx pb_2 - \frac{\partial(pb_2)}{\partial x_2}\lambda_2 + \frac{1}{2}\frac{\partial^2(pb_2)}{\partial x_2^2}\lambda_2^2 \\
T_6 &\approx pd_2 + \frac{\partial(pd_2)}{\partial x_2}\lambda_2 + \frac{1}{2}\frac{\partial^2(pd_2)}{\partial x_2^2}\lambda_2^2 \\
T_7 &\approx pm_{12} + \frac{\partial(pm_{12})}{\partial x_1}\lambda_1 - \frac{\partial(pm_{12})}{\partial x_2}\lambda_2 + \frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2}(-1)^{i+j}\frac{\partial^2(pm_{12})}{\partial x_i \partial x_j}\lambda_i\lambda_j \\
T_8 &\approx pm_{21} - \frac{\partial(pm_{21})}{\partial x_1}\lambda_1 + \frac{\partial(pm_{21})}{\partial x_2}\lambda_2 + \frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2}(-1)^{i+j}\frac{\partial^2(pm_{21})}{\partial x_i \partial x_j}\lambda_i\lambda_j \\
T_9 &\approx pm_{11} + \frac{\partial(pm_{11})}{\partial x_1}\lambda_1 + \frac{\partial(pm_{11})}{\partial x_2}\lambda_2 + \frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2}(-1)^{i+j}\frac{\partial^2(pm_{11})}{\partial x_i \partial x_j}\lambda_i\lambda_j \\
T_{10} &\approx pm_{22} - \frac{\partial(pm_{22})}{\partial x_1}\lambda_1 - \frac{\partial(pm_{22})}{\partial x_2}\lambda_2 + \frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2}(-1)^{i+j}\frac{\partial^2(pm_{22})}{\partial x_i \partial x_j}\lambda_i\lambda_j
\end{aligned}$$

Substituting these expressions into (3.2.2) and assuming that Δt, λ_1 and λ_2 are small, then it is seen that $p(t,x_1,x_2)$ approximately solves the Fokker-Planck equation

$$\begin{aligned}\frac{\partial p(t,x_1,x_2)}{\partial t} &= -\sum_{i=1}^{2}\frac{\partial}{\partial x_i}\left[\mu_i(t,x_1,x_2)p(t,x_1,x_2)\right] \\ &+\frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2}\frac{\partial}{\partial x_i \partial x_j}\left[\sum_{k=1}^{2} b_{ik}(t,x_1,x_2)b_{jk}(t,x_1,x_2)p(t,x_1,x_2)\right]\end{aligned} \tag{3.2.3}$$

where $\mu = (\mu_1,\mu_2)$ and $B = \{b_{ij}\}_{1\le i,j\le 2}$. On the other hand, it is well known that the probability distribution $p(t,x_1,x_2)$ that solves equation (3.2.3) coincides with the distribution of the solution at time t to the following system of stochastic differential equations

$$dS = \mu(t,S)dt + B(t,S)dW(t), \quad S(0) = S_0 \tag{3.2.4}$$

where W is a two-dimensional standard Brownian motion and S_0 is a given deterministic initial condition. The stochastic differential equation (3.2.4) describes the random evolution of the two-state system S related to the changes described in Table 3.1.

3.2.1 The *Greenhalgh et al.* [1] model

We now specialize the general model introduced in the previous section to the case investigated in Greenhalgh et al. [1] (where the process (S_1,S_2) is denoted as (S,I)). The values of the parameters in Table 3.1 are chosen as follows:

Table 3.2: Probabilities in *Greenhalgh et al.*'s paper

Change	Probability
$\Delta\mathbf{S}^{(1)} = [-1,0]^T$	$\mu S_1 \Delta t$
$\Delta\mathbf{S}^{(2)} = [1,0]^T$	$\mu N \Delta t$
$\Delta\mathbf{S}^{(3)} = [0,-1]^T$	$\mu S_2 \Delta t$
$\Delta\mathbf{S}^{(4)} = [0,1]^T$	0
$\Delta\mathbf{S}^{(5)} = [-1,1]^T$	$\frac{\lambda(N)S_1S_2}{N}\Delta t$
$\Delta\mathbf{S}^{(6)} = [1,-1]^T$	$\gamma S_2 \Delta t$
$\Delta\mathbf{S}^{(7)} = [-1,-1]^T$	0
$\Delta\mathbf{S}^{(8)} = [1,1]^T$	0
$\Delta\mathbf{S}^{(9)} = [0,0]^T$	$1-\sum_{j=1}^{8} p_j$

where $N := S_1+S_2$, $\lambda : [0,+\infty[\to [0,+\infty[$ is a continuous monotone increasing function and μ and γ are positive constants. We refer to the paper [1] for the biological interpretation of these quantities. Now, according to Table 3.2 the vector μ and matrix V in

(3.2.1) read

$$\mu(t, S_1, S_2) = \begin{bmatrix} -\frac{\lambda(N)S_1S_2}{N} + (\mu+\gamma)S_2 \\ \frac{\lambda(N)S_1S_2}{N} - (\mu+\gamma)S_2 \end{bmatrix}$$

and

$$V(t, S_1, S_2) = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

where to ease the notation we set

$$\begin{aligned} a &:= \frac{\lambda(N)S_1S_2}{N} + (\mu+\gamma)S_2 + 2\mu S_1 \\ b &:= -\frac{\lambda(N)S_1S_2}{N} - \gamma S_2 \\ c &:= \frac{\lambda(N)S_1S_2}{N} + (\mu+\gamma)S_2. \end{aligned}$$

Therefore,

$$B(t, S_1, S_2) = V(t, S_1, S_2)^{\frac{1}{2}} = \frac{1}{d}\begin{bmatrix} a+w & b \\ b & c+w \end{bmatrix}$$

with

$$w := \sqrt{ac - b^2} \quad \text{and} \quad d := \sqrt{a + c + 2w}.$$

We are then lead to study the following two dimensional system of stochastic differential equations

$$\begin{cases} dS_1 &= \left[-\frac{\lambda(N)S_1S_2}{N} + (\mu+\gamma)S_2\right]dt + \frac{a+w}{d}dW_1 + \frac{b}{d}dW_2 \\ dS_2 &= \left[\frac{\lambda(N)S_1S_2}{N} - (\mu+\gamma)S_2\right]dt + \frac{b}{d}dW_1 + \frac{c+w}{d}dW_2 \end{cases} \tag{3.2.5}$$

where $W = (W_1, W_2)$ is a standard two dimensional Brownian motion. We observe that by construction

$$\left(\frac{a+w}{d}\right)^2 + \left(\frac{b}{d}\right)^2 = a.$$

Therefore, by the martingale representation theorem (see for instance Theorem 3.9 Chapter V in [38]) there exists a Brownian motion W_3 such that the first equation in (3.2.5) can be rewritten as

$$dS_1 = \left[-\frac{\lambda(N)S_1S_2}{N} + (\mu+\gamma)S_2\right]dt + \sqrt{\frac{\lambda(N)S_1S_2}{N} + (\mu+\gamma)S_2 + 2\mu S_1}dW_3$$

Similarly, since

$$\left(\frac{b}{d}\right)^2 + \left(\frac{c+w}{d}\right)^2 = c$$

by the martingale representation theorem there exists a Brownian motion W_4 such that the second equation in (3.2.5) can be rewritten as

$$dS_2 = \left[\frac{\lambda(N)S_1S_2}{N} - (\mu+\gamma)S_2\right]dt + \sqrt{\frac{\lambda(N)S_1S_2}{N} + (\mu+\gamma)S_2}dW_4.$$

This implies that the system (3.2.5) is equivalent to

$$\begin{cases} dS_1 &= \left[-\frac{\lambda(N)S_1S_2}{N} + (\mu+\gamma)S_2\right]dt + \sqrt{\frac{\lambda(N)S_1S_2}{N} + (\mu+\gamma)S_2 + 2\mu S_1}dW_3 \\ dS_2 &= \left[\frac{\lambda(N)S_1S_2}{N} - (\mu+\gamma)S_2\right]dt + \sqrt{\frac{\lambda(N)S_1S_2}{N} + (\mu+\gamma)S_2}dW_4. \end{cases} \tag{3.2.6}$$

We remark that by construction the Brownian motions W_3 and W_4 are now correlated. Moreover, if we notice that the drift of the first equation in (3.2.5) is the opposite of the one in the second equation in (3.2.5), recalling that $N = S_1 + S_2$ we may write

$$dN = \frac{a+b+w}{d}dW_1 + \frac{b+c+w}{d}dW_2$$

and, exploiting the definitions of a, b, c, d and w, we conclude as before that there exists a Brownian motion W_5 such that

$$dN = \sqrt{2\mu N}dW_5. \tag{3.2.7}$$

Hence, instead of studying the system (3.2.5), the authors in [1] study the equivalent system

$$\begin{cases} dS_2 &= \left[\frac{\lambda(N)}{N}(N-S_2)S_2 - (\mu+\gamma)S_2\right]dt + \sqrt{\frac{\lambda(N)}{N}(N-S_2)S_2 + (\mu+\gamma)S_2}dW_4 \\ dN &= \sqrt{2\mu N}dW_5 \end{cases} \tag{3.2.8}$$

where the Brownian motions W_4 and W_5 are correlated. In the system (3.2.8) the equation for N does not depend on S_2 and it belongs to the family of the square root processes ([21]). Once the equation for N is solved, the equation for S_2 contains random (for the presence of N) Hölder continuous coefficients. Moreover, due to the presence of the square root in the diffusion coefficient of S_2, the authors of [1] consider a modified version of the first equation in (3.2.8) to make the coefficients defined on the whole real line. They consider

$$dS_2(t) = \bar{a}(t, N(t), S_2(t))dt + \bar{g}(t, N(t), S_2(t))dW_4(t) \tag{3.2.9}$$

where

$$\bar{a}(t,y,x) = \begin{cases} 0 & \text{for } x < 0 \\ \frac{\lambda(y)x}{y}(y-x) - (\mu+\gamma)x & \text{for } 0 \le x \le y\left(1 + \frac{\mu+\gamma}{\lambda(y)}\right) \\ \bar{a}\left(t, y, y\left(1 + \frac{\mu+\gamma}{\lambda(y)}\right)\right) & \text{for } x > y\left(1 + \frac{\mu+\gamma}{\lambda(y)}\right) \end{cases}$$

and

$$\bar{g}(t,y,x) = \begin{cases} 0 & \text{for } x < 0 \\ \sqrt{\frac{\lambda(y)x}{y}(y-x) + (\mu+\gamma)x} & \text{for } 0 \le x \le y\left(1 + \frac{\mu+\gamma}{\lambda(y)}\right) \\ 0 & \text{for } x > y\left(1 + \frac{\mu+\gamma}{\lambda(y)}\right) \end{cases}$$

The existence of a unique non explosive strong solution to equation (3.2.9) is obtained through a localization argument in terms of stopping times and comparison inequalities to control the non explosivity of the solution. In the next section we will consider a class of stochastic differential equations, which includes equation (3.2.9), allowing for more general models where the existence of a unique non explosive strong solution is proved via a standard Caychy-Euler-Peano approximation method.

3.3 Main theorem

Motivated by the discussion in the previous sections, we are now ready to state and prove the main result of our manuscript. We begin by specifying the class of coefficients involved in the stochastic differential equations under investigation.
Let $g : [0, +\infty[\times\mathbb{R}\times\mathbb{R} \to \mathbb{R}$ be a function of the form

$$g(t,y,x) = \sqrt{-x^2 + \alpha(t,y)x + \beta(t,y)} \tag{3.3.1}$$

where $\alpha, \beta : [0, +\infty[\times\mathbb{R} \to \mathbb{R}$ are measurable functions satisfying the condition

$$\alpha(t,y)^2 + 4\beta(t,y) \ge 0 \quad \text{for all } (t,y) \in [0,+\infty[\times\mathbb{R}. \tag{3.3.2}$$

We observe that condition (3.3.2) implies that

$$-x^2 + \alpha(t,y)x + \beta(t,y) \ge 0 \quad \text{if and only if} \quad r_1(t,y) \le x \le r_2(t,y)$$

where we set

$$r_1(t,y) := \frac{\alpha(t,y) - \sqrt{\alpha(t,y)^2 + 4\beta(t,y)}}{2}$$

and

$$r_2(t,y) := \frac{\alpha(t,y) + \sqrt{\alpha(t,y)^2 + 4\beta(t,y)}}{2}.$$

Now, we define

$$\bar{g}(t,y,x) := \begin{cases} 0 & \text{if} \quad x < r_1(t,y) \\ g(t,y,x) & \text{if} \quad r_1(t,y) \le x \le r_2(t,y) \\ 0 & \text{if} \quad x > r_2(t,y) \end{cases} \tag{3.3.3}$$

The function $\bar{g}$ will be the diffusion coefficient of our stochastic differential equation.

Assumption 3.3.1. *There exist a positive constant M such that*

$$|\alpha(t,y)| \le M(1+|y|) \quad \text{and} \quad |\beta(t,y)| \le M(1+|y|) \tag{3.3.4}$$

for all $(t,y) \in [0,\infty[\times\mathbb{R}$. Moreover, there exists a positive constant H such that

$$|\bar{g}(t,y_1,x_1) - \bar{g}(t,y_2,x_2)| \le H(\sqrt{|y_1-y_2|} + \sqrt{|x_1-x_2|}) \tag{3.3.5}$$

for all $t \in [0,\infty[$ and $y_1, y_2, x_1, x_2 \in \mathbb{R}$.

We observe that assumption (3.3.4) implies the bound

$$\begin{aligned} |\bar{g}(t,y,x)| &\le \max_{x\in\mathbb{R}} |\bar{g}(t,y,x)| \\ &= \sqrt{\frac{\alpha(t,y)^2}{4} + \beta(t,y)} \\ &\le M(1+|y|) \end{aligned}$$

for all $t \in [0,\infty[$ and $y \in \mathbb{R}$. Here the constant M may differ from the one appearing in (3.3.4); we will adopt this convention for the rest of the paper. We also remark that by construction inequality (3.3.5) for $y_1 = y_2$ is satisfied with a constant $H = \sqrt{|\alpha(t,y_1)|}$.

We now introduce the drift coefficient of our SDE. We start with a measurable function $a : [0,+\infty[\times\mathbb{R}\times\mathbb{R} \to \mathbb{R}$ with the following property.

Assumption 3.3.2. *There exists a positive constant M such that*

$$|a(t,y,x)| \le M(1+|y|+|x|) \tag{3.3.6}$$

for all $t \in [0,\infty[$ and $x, y \in \mathbb{R}$. Moreover, there exists a positive constant L such that

$$|a(t,y_1,x_1) - a(t,y_2,x_2)| \le L(|y_1-y_2|+|x_1-x_2|) \tag{3.3.7}$$

for all $t \in [0,\infty[$ and $y_1, y_2, x_1, x_2 \in \mathbb{R}$.

Then, we set

$$\bar{a}(t,y,x) := \begin{cases} a(t,y,r_1(t,y)) & \text{if} \quad x < r_1(t,y) \\ a(t,y,x) & \text{if} \quad r_1(t,y) \leq x \leq r_2(t,y) \\ a(t,y,r_2(t,y)) & \text{if} \quad x > r_2(t,y) \end{cases} \tag{3.3.8}$$

Observe that by construction also the function $\bar{a}$ satisfies Assumption 3.3.2.

We now consider the following one dimensional stochastic differential equation

$$dX_t = \bar{a}(t, Y_t, X_t)dt + \bar{g}(t, Y_t, X_t)dW_t^2, \quad X_0 = x \in \mathbb{R} \tag{3.3.9}$$

where $\{Y_t\}_{t\geq 0}$ is the unique strong solution of the stochastic differential equation

$$dY_t = m(t, Y_t)dt + \sigma(t, Y_t)dW_t^1, \quad Y_0 = y \in \mathbb{R}. \tag{3.3.10}$$

Here $\{(W_t^1, W_t^2)\}_{t\geq 0}$ is a two dimensional correlated Brownian motion defined on a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t\geq 0})$ where the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ is generated by the process $\{(W_t^1, W_t^2)\}_{t\geq 0}$. Strong solutions are meant to be $\{\mathcal{F}_t\}_{t\geq 0}$-adapted. Regarding equation (3.3.10), the coefficients m and σ are assumed to entail existence and uniqueness of a strong solution $\{Y_t\}_{t\geq 0}$ such that

$$E\Big[\sup_{t\in[0,T]} |Y_t|^2\Big] \text{ is finite for all } T > 0.$$

Equations (3.3.9) and (3.3.10) describe a class of equations which includes equations (3.2.9) and (3.2.7) as a particular case.

Remark 3.3.3. *If $r_1(t,y) = r_2(t,y)$ for all $(t,y) \in [0,\infty[\times\mathbb{R}$, which is equivalent to say that $\alpha(t,y)^2 + 4\beta(t,y) = 0$, then the diffusion coefficient $\bar{g}$ is identically zero and the drift coefficient becomes $\bar{a}(t,y,x) = a(t,y,\alpha(t,y)/2)$. Therefore, in this particular case the SDE (3.3.9) takes the form*

$$dX_t = a(t, Y_t, \alpha(t, Y_t)/2)dt, \quad X_0 = x$$

whose solution is explicitly given by the formula

$$X_t = x + \int_0^t a(s, Y_s, \alpha(s, Y_s)/2)ds.$$

Theorem 3.3.4 (Strong existence and uniqueness)**.** *Let Assumption 3.3.1 and Assumption 3.3.2 be fulfilled. Then, the stochastic differential equation (3.3.9) possesses a unique strong solution $\{X_t\}_{t\geq 0}$.*

Proof. To ease the notation we consider the time-homogeneous case and hence we drop the explicit dependence on t from all the coefficients.
We fix an arbitrary $T > 0$ and prove existence and uniqueness of a solution for the SDE

$$X_t = x + \int_0^t \bar{a}(Y_s, X_s)ds + \int_0^t \bar{g}(Y_s, X_s)dW_s^2, \quad X_0 = x. \tag{3.3.11}$$

on the time interval $t \in [0, T]$. The proof for the existence is rather long and proceeds as follows: using a Cauchy-Euler-Peano approximate solutions technique we define, associated to a partition Δ_n of $[0, T]$ a stochastic process X^n. We will, at the beginning, prove a convergence result for X^n in the space $L^1([0, T] \times \Omega)$, then we will prove a convergence result for X^n in the space $\mathcal{C}[0, T]$ with the norm of the uniform convergence and this will eventually yield the result.

Existence: We consider a sequence of partitions $\{\Delta_n\}_{n\geq 1}$ of the interval $[0, T]$ with $\Delta_n \subseteq \Delta_{n+1}$. Each partition Δ_n will consist of a set of $N_n + 1$ points $\{t_0^n, t_1^n, ..., t_{N_n}^n\}$ satisfying

$$0 = t_0^n < t_1^n < \cdots < t_{N_n}^n = T.$$

We denote by $\|\Delta_n\| := \max_{0\leq k\leq N_n-1}|t_{k+1}^n - t_k^n|$, the mesh of the partition Δ_n, and assume that $\lim_{n\to\infty}\|\Delta_n\| = 0$. In the sequel, we will write t_k instead of t_k^n when the membership to the partition Δ_n will be clear from the context.
For a given partition Δ_n we construct a continuous and $\{\mathcal{F}_t\}_{t\geq 0}$-adapted stochastic process $\{X_t^n\}_{t\in[0,T]}$ as follows: for $t = 0$ we set $X_t^n = x$ while for $t \in]t_k, t_{k+1}]$ we define

$$X_t^n := X_{t_k}^n + \bar{a}(Y_{t_k}, X_{t_k}^n)(t - t_k) + g(Y_{t_k}, X_{t_k}^n)(W_t - W_{t_k}). \tag{3.3.12}$$

It is useful to observe that, denoting $\eta_n(t) = t_k$ when $t \in]t_k, t_{k+1}]$, we may represent X_t^n in the compact form:

$$X_t^n = x + \int_0^t \bar{a}(Y_{\eta_n(s)}, X_{\eta_n(s)}^n)ds + \int_0^t \bar{g}(Y_{\eta_n(s)}, X_{\eta_n(s)}^n)dW_s^2. \tag{3.3.13}$$

Step one: $\mathbb{E}|X_{\eta_n(t)}^n|$ *is uniformly bounded with respect to n and t*

We begin with equation (3.3.12). Using the triangle inequality and upper bounds for $\bar{a}$

and $\bar{g}$ we get

$$\begin{aligned}
\mathbb{E}[|X^n_{t_{k+1}}|] &\leq \mathbb{E}[|X^n_{t_k}|] + \mathbb{E}[|\bar{a}(Y_{t_k}, X^n_{t_k})(t_{k+1} - t_k)|] \\
&\quad + \mathbb{E}[|\bar{g}(Y_{t_k}, X^n_{t_k})(W_{t_{k+1}} - W_{t_k})|] \\
&\leq \mathbb{E}[|X^n_{t_k}|] + M|t_{k+1} - t_k|\mathbb{E}\left[1 + |Y_{t_k}|\right] + M|t_{k+1} - t_k|\mathbb{E}\left[|X^n_{t_k}|\right] \\
&\quad + M\mathbb{E}\left[(1 + |Y_{t_k}|)|W_{t_{k+1}} - W_{t_k}|\right] \\
&\leq (1 + M\|\Delta_n\|)\mathbb{E}[|X^n_{t_k}|] + M|t_{k+1} - t_k|\mathbb{E}\left[1 + |Y_{t_k}|\right] \\
&\quad + \frac{M}{2}\left(\mathbb{E}\left[(1 + |Y_{t_k}|)^2\right] + \mathbb{E}\left[|W_{t_{k+1}} - W_{t_k}|^2\right]\right) \\
&\leq (1 + M\|\Delta_n\|)\mathbb{E}[|X^n_{t_k}|] + M\|\Delta_n\| \sup_{t\in[0,T]} \mathbb{E}\left[1 + |Y_t|\right] \\
&\quad + \frac{M}{2} \sup_{t\in[0,T]} \mathbb{E}\left[(1 + |Y_t|)^2\right] + \frac{M}{2}|t_{k-1} - t_k| \\
&\leq (1 + M\|\Delta_n\|)\mathbb{E}[|X^n_{t_k}|] + M\|\Delta_n\| \sup_{t\in[0,T]} \mathbb{E}\left[1 + |Y_t|\right] \\
&\quad + \frac{M}{2} \sup_{t\in[0,T]} \mathbb{E}\left[(1 + |Y_t|)^2\right] + \frac{M}{2}\|\Delta_n\| \\
&\leq (1 + M\|\Delta_n\|)\mathbb{E}[|X^n_{t_k}|] + \frac{M}{2} \sup_{t\in[0,T]} \mathbb{E}\left[(1 + |Y_t|)^2\right] + \varepsilon.
\end{aligned}$$

Here we used the fact that $\|\Delta_n\|$ tends to zero as n tends to infinity and that $\sup_{t\in[0,T]} \mathbb{E}\left[1 + |Y_t|\right]$ is finite: we can therefore choose n big enough to make

$$M\|\Delta_n\| \sup_{t\in[0,T]} \mathbb{E}\left[1 + |Y_t|\right] + \frac{M}{2}\|\Delta_n\|$$

smaller than a given positive ε. Comparing the first and last terms of the previous chain of inequalities we get for all $k \in \{0, ..., N_n - 1\}$

$$\mathbb{E}[|X^n_{t_{k+1}}|] \leq (1 + M\|\Delta_n\|)\mathbb{E}[|X^n_{t_k}|] + \frac{M}{2} \sup_{t\in[0,T]} \mathbb{E}\left[(1 + |Y_t|)^2\right] + \varepsilon$$

which by recursion implies

$$\begin{aligned}
\mathbb{E}[|X^n_{t_k}|] &\leq \gamma_1^k|x| + \frac{\gamma_1^k - 1}{\gamma_1 - 1}\gamma_2 \\
&\leq \gamma_1^{N_n}|x| + \frac{\gamma_1^{N_n} - 1}{\gamma_1 - 1}\gamma_2
\end{aligned}$$

where for notational convenience we set

$$\gamma_1 := 1 + M\|\Delta_n\| \quad \text{and} \quad \gamma_2 := \frac{M}{2} \sup_{t\in[0,T]} \mathbb{E}\left[(1 + |Y_{t_k}|)^2\right] + \varepsilon.$$

Since $\eta_n(t)$ is a step function in $[0,T]$ with values $\{t_0,t_1,...,t_{N_n}\}$, the previous estimate for $k\in\{0,...,N_n-1\}$ entails the boundedness of the function $[0,T]\ni t\to E[|X^n_{\eta_n(t)}|]$. We now obtain an estimate for $E[|X^n_{\eta_n(t)}|]$ which is also uniform with respect to n. Using the triangle inequality in (3.3.13) we can write

$$\begin{aligned}\mathbb{E}[|X^n_{\eta_n(t)}|] &\leq |x|+\mathbb{E}\left[\left|\int_0^{\eta_n(t)}\bar{a}(Y_{\eta_n(s)},X^n_{\eta_n(s)})ds\right|\right]\\ &\quad+\mathbb{E}\left[\left|\int_0^{\eta_n(t)}\bar{g}(Y_{\eta_n(s)},X^n_{\eta_n(s)})dW^2_s\right|\right]. \end{aligned}\tag{3.3.14}$$

For the first expected value on the right hand side above we employ the assumptions on $\bar{a}$:

$$\begin{aligned}\mathbb{E}\left[\left|\int_0^{\eta_n(t)}\bar{a}(Y_{\eta_n(s)},X^n_{\eta_n(s)})ds\right|\right] &\leq \mathbb{E}\left[\int_0^t|\bar{a}(Y_{\eta_n(s)},X^n_{\eta_n(s)})|ds\right]\\ &\leq M\mathbb{E}\left[\int_0^t(1+|X^n_{\eta_n(s)}|+|Y_{\eta_n(s)}|)ds\right]\\ &= M\int_0^t\mathbb{E}[|X^n_{\eta_n(s)}|]ds+M\int_0^t\mathbb{E}\left[1+|Y_{\eta_n(s)}|\right]ds\\ &\leq M\int_0^t\mathbb{E}[|X^n_{\eta_n(s)}|]ds+MT\sup_{t\in[0,T]}\mathbb{E}\left[1+|Y_t|\right].\end{aligned}$$

Using the Itô isometry and the assumptions on $\bar{g}$ we can treat the second expected value as follows:

$$\begin{aligned}\mathbb{E}\left[\left|\int_0^{\eta_n(t)}\bar{g}(Y_{\eta_n(s)},X^n_{\eta_n(s)})dW^2_s\right|\right] &\leq \left(\mathbb{E}\left[\left|\int_0^{\eta_n(t)}\bar{g}(Y_{\eta_n(s)},X^n_{\eta_n(s)})dW^2_s\right|^2\right]\right)^{\frac{1}{2}}\\ &\leq \left(\int_0^t\mathbb{E}[|\bar{g}(Y_{\eta_n(s)},X^n_{\eta_n(s)})|^2]ds\right)^{\frac{1}{2}}\\ &\leq M\left(\int_0^t\mathbb{E}[(1+|Y_{\eta_n(s)}|)^2]ds\right)^{\frac{1}{2}}\\ &\leq M\sqrt{T\sup_{t\in[0,T]}\mathbb{E}[(1+|Y_t|)^2]}.\end{aligned}$$

Plugging the last two estimates in (3.3.14) gives

$$
\begin{aligned}
\mathbb{E}[|X^n_{\eta_n(t)}|] &\leq |x|+\mathbb{E}\left[\left|\int_0^{\eta_n(t)} \bar{a}(Y_{\eta_n(s)}, X^n_{\eta_n(s)})ds\right|\right] \\
&\quad +\mathbb{E}\left[\left|\int_0^{\eta_n(t)} \bar{g}(Y_{\eta_n(s)}, X^n_{\eta_n(s)})dW^2_s\right|\right] \\
&\leq |x|+M\int_0^t \mathbb{E}[|X^n_{\eta_n(s)}|]ds + MT\sup_{t\in[0,T]}\mathbb{E}\left[1+|Y_t|\right] \\
&\quad +M\sqrt{T\sup_{t\in[0,T]}\mathbb{E}[(1+|Y_t|)^2]} \\
&= G+M\int_0^t \mathbb{E}[|X^n_{\eta_n(s)}|]ds
\end{aligned}
$$

where

$$
G := |x|+MT\sup_{t\in[0,T]}\mathbb{E}\left[1+|Y_t|\right]+M\sqrt{T\sup_{t\in[0,T]}\mathbb{E}[(1+|Y_t|)^2]}.
$$

By the Gronwall inequality (we proved before that $t\to\mathbb{E}[|X^n_{\eta_n(t)}|]$ is a non negative, bounded and measurable function) we conclude that

$$
\mathbb{E}[|X^n_{\eta_n(t)}|]\leq Ge^{Mt}\leq Ge^{MT} \tag{3.3.15}
$$

which provides the desired uniform bound (with respect to n and t) for $\mathbb{E}[|X^n_{\eta_n(t)}|]$.

Step two: $\mathbb{E}[|X^n_t - X^n_{\eta_n(t)}|]$ *tends to zero as* n *tends to infinity, uniformly with respect to* $t\in[0,T]$

We proceed as in step one. Recalling the identity (3.3.13) we can write

$$\begin{aligned}
\mathbb{E}[|X_t^n - X_{\eta_n(t)}^n|] &= \mathbb{E}\left[\left|\int_{\eta_n(t)}^t \bar{a}(Y_{\eta_n(s)}, X_{\eta_n(s)}^n)ds + \int_{\eta_n(t)}^t \bar{g}(Y_{\eta_n(s)}, X_{\eta_n(s)}^n)dW_s^2\right|\right] \\
&\leq \int_{\eta_n(t)}^t \mathbb{E}[|\bar{a}(Y_{\eta_n(s)}, X_{\eta_n(s)}^n)|]ds + \mathbb{E}\left[\left|\int_{\eta_n(t)}^t \bar{g}(Y_{\eta_n(s)}, X_{\eta_n(s)}^n)dW_s^2\right|\right] \\
&\leq M\int_{\eta_n(t)}^t \mathbb{E}[(1+|X_{\eta_n(s)}^n|+|Y_{\eta_n(s)}|)]ds \\
&\quad + \left(\mathbb{E}\left[\left|\int_{\eta_n(t)}^t \bar{g}(Y_{\eta_n(s)}, X_{\eta_n(s)}^n)dW_s^2\right|^2\right]\right)^{\frac{1}{2}} \\
&\leq M(t-\eta_n(t))\left(Ge^{MT} + \sup_{t\in[0,T]}\mathbb{E}[1+|Y_t|]\right) \\
&\quad + \left(\int_{\eta_n(t)}^t \mathbb{E}[|\bar{g}(Y_{\eta_n(s)}, X_{\eta_n(s)}^n)|^2]ds\right)^{\frac{1}{2}} \\
&\leq M(t-\eta_n(t))\left(Ge^{MT} + \sup_{t\in[0,T]}\mathbb{E}[1+|Y_t|]\right) \\
&\quad + M\sqrt{t-\eta_n(t)}\sqrt{\sup_{t\in[0,T]}\mathbb{E}[(1+|Y_t|)^2]} \\
&\leq M\sqrt{\|\Delta_n\|}\left(Ge^{MT} + \sup_{t\in[0,T]}\mathbb{E}[1+|Y_t|] + \sqrt{\sup_{t\in[0,T]}\mathbb{E}[(1+|Y_t|)^2]}\right).
\end{aligned}$$

Here, in the third equality, we utilized the uniform upper bound (3.3.15). We have therefore proved that

$$\begin{aligned}
\mathbb{E}[|X_t^n - X_{\eta_n(t)}^n|] &\leq M\sqrt{\|\Delta_n\|}\left(Ge^{MT} + \sup_{t\in[0,T]}\mathbb{E}[1+|Y_t|] + \sqrt{\sup_{t\in[0,T]}\mathbb{E}[(1+|Y_t|)^2]}\right) \\
&=: M_1\sqrt{\|\Delta_n\|}
\end{aligned}$$

This in turn implies that $\mathbb{E}[|X_t^n - X_{\eta_n(t)}^n|]$ tends to zero as n tends to infinity, uniformly with respect to $t \in [0,T]$.

Step three: *$\{X^n\}_{n\geq 1}$ is a Cauchy sequence in $L^1([0,T]\times\Omega)$.*

We need to prove that for any $\varepsilon > 0$ there exists $n_\varepsilon \in \mathbb{N}$ such that

$$\mathbb{E}\left[\int_0^T |X_t^n - X_t^m|dt\right] < \varepsilon \quad \text{for all } n, m \geq n_\varepsilon.$$

We have:

$$
\begin{aligned}
X_t^n - X_t^m &= \int_0^t \left[\bar{a}(Y_{\eta_n(s)}.X^n_{\eta_n(s)}) - \bar{a}(Y_{\eta_m(s)}, X^m_{\eta_m(s)})\right] ds \\
&\quad + \int_0^t \left[\bar{g}(Y_{\eta_n(s)}, X^n_{\eta_n(s)}) - \bar{g}(Y_{\eta_m(s)}, X^m_{\eta_m(s)})\right] dW_s^2
\end{aligned}
$$

We now aim to apply the Itô formula to the stochastic process $\{X_t^n - X_t^m\}_{t\in[0,T]}$ for a suitable smooth function that we now describe.
Consider the decreasing sequence of real numbers $\{a_h\}_{h\geq 0}$ defined by induction as follows:

$$
a_0 = 1 \quad \text{and for } h \geq 1, \int_{a_h}^{a_{h-1}} \frac{1}{u} du = h.
$$

It is easy to see that $a_h = e^{-\frac{h(h+1)}{2}}$ and therefore that $\lim_{h\to+\infty} a_h = 0$. Define the function $\Phi_h(u)$ for $u \in [0,\infty)$ such that $\Phi_h(0) = 0$, $\Phi_h(u) \in \mathcal{C}^2([0,\infty[)$ and

$$
\Phi_h''(u) = \begin{cases} 0, & 0 \leq u \leq a_h \\ \text{a value between } 0 \text{ and } \frac{2}{hu}, & a_h < u < a_{h-1} \\ 0, & u \geq a_{h-1} \end{cases} \tag{3.3.16}
$$

such that Φ_h'' is continuous and

$$
\int_{a_h}^{a_{h-1}} \Phi_h''(u) du = 1.
$$

Integrating Φ_h'' we get

$$
\Phi_h'(u) = \begin{cases} 0, & 0 \leq u \leq a_h \\ \text{a value between } 0 \text{ and } 1, & a_h < u < a_{h-1} \\ 1, & u \geq a_{h-1} \end{cases} \tag{3.3.17}
$$

Finally we choose $\theta_h(u) = \Phi_h(|u|)$. Then, we have:

$$
\begin{aligned}
\theta_h(X_t^n - X_t^m) &= \int_0^t \theta_h'(X_s^n - X_s^m) \left[\bar{a}(Y_{\eta_n(s)}.X^n_{\eta_n(s)}) - \bar{a}(Y_{\eta_m(s)}, X^m_{\eta_m(s)})\right] ds \\
&\quad + \int_0^t \theta_h'(X_s^n - X_s^m) \left[\bar{g}(Y_{\eta_n(s)}, X^n_{\eta_n(s)}) - \bar{g}(Y_{\eta_m(s)}, X^m_{\eta_m(s)})\right] dW_s^2 \\
&\quad + \frac{1}{2}\int_0^t \theta''(X_s^n - X_s^m) \left[\bar{g}(Y_{\eta_n(s)}, X^n_{\eta_n(s)}) - \bar{g}(Y_{\eta_m(s)}, X^m_{\eta_m(s)})\right]^2 ds \\
&=: I_1(\theta_h) + I_2(\theta_h) + I_3(\theta_h)
\end{aligned}
$$

Since for any $h \geq 0$ and $u \in \mathbb{R}$ we have by construction that $|u| - a_{h-1} \leq \theta_h(u)$, we can write

$$\begin{aligned}
\mathbb{E}[|X_t^n - X_t^m|] &\leq a_{h-1} + \mathbb{E}[\theta_h(X_t^n - X_t^m)] \\
&= a_{h-1} + \mathbb{E}[I_1(\theta_h) + I_2(\theta_h) + I_3(\theta_h)] \\
&= a_{h-1} + \mathbb{E}[I_1(\theta_h)] + \mathbb{E}[I_3(\theta_h)]. \qquad (3.3.18)
\end{aligned}$$

Let us now estimate $\mathbb{E}[|I_1(\theta_h)|]$:

$$\begin{aligned}
\mathbb{E}[|I_1(\theta_h)|] &= \mathbb{E}\left[\left|\int_0^t \theta_h'(X_s^n - X_s^m)\left[\bar{a}(Y_{\eta_n(s)}.X^n_{\eta_n(s)}) - \bar{a}(Y_{\eta_m(s)}, X^m_{\eta_m(s)})\right] ds\right|\right] \\
&\leq \mathbb{E}\left[\int_0^t |\theta_h'(X_s^n - X_s^m)| \cdot |\bar{a}(Y_{\eta_n(s)}.X^n_{\eta_n(s)}) - \bar{a}(Y_{\eta_m(s)}, X^m_{\eta_m(s)})| ds\right] \\
&\leq \mathbb{E}\left[\int_0^t |\bar{a}(Y_{\eta_n(s)}.X^n_{\eta_n(s)}) - \bar{a}(Y_{\eta_m(s)}, X^m_{\eta_m(s)})| ds\right] \\
&\leq L\int_0^t \mathbb{E}[|X^n_{\eta_n(s)} - X^m_{\eta_m(s)}|]ds + L\int_0^t \mathbb{E}[|Y_{\eta_n(s)} - Y_{\eta_m(s)}|]ds
\end{aligned}$$

In the second inequality we utilized the bound $|\theta_h'(u)| \leq 1$ which is valid for all $h \geq 0$ and $u \in \mathbb{R}$. By means of the estimate obtained in step two we can write

$$\begin{aligned}
\mathbb{E}[|X^n_{\eta_n(s)} - X^m_{\eta_m(s)}|] &\leq \mathbb{E}[|X^n_{\eta_n(s)} - X_s^n|] + \mathbb{E}[|X_s^n - X_s^m|] + \mathbb{E}[|X_s^m - X^m_{\eta_m(s)}|] \\
&\leq M_1(\sqrt{\|\Delta_n\|} + \sqrt{\|\Delta_m\|}) + \mathbb{E}[|X_s^n - X_s^m|].
\end{aligned}$$

Similarly we get

$$\begin{aligned}
\mathbb{E}[|Y_{\eta_n(s)} - Y_{\eta_m(s)}|] &\leq \mathbb{E}[|Y_{\eta_n(s)} - Y_s|] + \mathbb{E}[Y_s - Y_{\eta_m(s)}|] \\
&\leq C(\sqrt{\|\Delta_n\|} + \sqrt{\|\Delta_m\|})
\end{aligned}$$

where the last inequality is due to well known estimates for strong solutions of stochastic differential equations. Combining the last two bounds we conclude that

$$\begin{aligned}
\mathbb{E}[|I_1(\theta_h)|] &\leq L\int_0^t \mathbb{E}[|X^n_{\eta_n(s)} - X^m_{\eta_m(s)}|]ds + L\int_0^t \mathbb{E}[|Y_{\eta_n(s)} - Y_{\eta_m(s)}|]ds \\
&\leq TL(M_1 + C)(\sqrt{\|\Delta_n\|} + \sqrt{\|\Delta_m\|}) + L\int_0^t \mathbb{E}[|X_s^n - X_s^m|]ds \qquad (3.3.19)
\end{aligned}$$

We now treat $\mathbb{E}[I_3(\theta_h)]$; by the assumption (3.3.5) and properties of θ_h we get:

$$\begin{aligned}
\mathbb{E}[I_3(\theta_h)] &= \frac{1}{2}\mathbb{E}\left[\int_0^t \theta_h''(X_s^n - X_s^m)(\bar{g}(Y_{\eta_n(s)}, X_{\eta_n(s)}^n) - \bar{g}(Y_{\eta_m(s)}, X_{\eta_m(s)}^m))^2 ds\right] \\
&\leq \frac{H^2}{2}\mathbb{E}\left[\int_0^t \theta_h''(X_s^n - X_s^m)\left(\sqrt{|X_{\eta_n(s)}^n - X_{\eta_m(s)}^m|} + \sqrt{|Y_{\eta_n(s)} - Y_{\eta_m(s)}|}\right)^2 ds\right] \\
&\leq H^2\mathbb{E}\left[\int_0^t \theta_h''(X_s^n - X_s^m)\left(|X_{\eta_n(s)}^n - X_{\eta_m(s)}^m| + |Y_{\eta_n(s)} - Y_{\eta_m(s)}|\right) ds\right] \\
&\leq H^2\mathbb{E}\left[\int_0^t \frac{2}{h|X_s^n - X_s^m|}|X_s^n - X_s^m| ds\right] \\
&\quad + H^2\|\theta_h''\|\mathbb{E}\left[\int_0^t (|X_{\eta_n(s)}^n - X_s^n| + |X_{\eta_m(s)}^m - X_s^m|)ds\right] \\
&\quad + H^2\|\theta_h''\|\mathbb{E}\left[\int_0^t (|Y_{\eta_n(s)} - Y_s| + |Y_s - Y_{\eta_m(s)}|)ds\right] \\
&\leq \frac{2H^2T}{h} + \|\theta_h''\|TH^2(M_1 + C)(\sqrt{\|\Delta_n\|} + \sqrt{\|\Delta_m\|}). \qquad (3.3.20)
\end{aligned}$$

Here $\|\theta_h''\|$ denotes the supremum norm of θ_h'' while in the last inequality we used the same bound to obtain inequality (3.3.19). Now, let us fix $\varepsilon > 0$. For this ε let h be such that $0 < a_{h-1} < \varepsilon$ and $\frac{2H^2T}{h} < \varepsilon$. With this h being so chosen and fixed, $\|\theta_h''\|$ is bounded. Then, there exists $n_\varepsilon \in \mathbb{N}$ such that

$$(M_1 + C)(T + \|\theta_h''\|TH^2)(\sqrt{\|\Delta_n\|} + \sqrt{\|\Delta_m\|}) < \varepsilon$$

for all $n, m \geq n_\varepsilon$. We can now insert estimates (3.3.19) and (3.3.20) in (3.3.18) to obtain

$$\begin{aligned}
\mathbb{E}[|X_t^n - X_t^m|] &\leq a_{h-1} + \mathbb{E}[I_1(\theta_h)] + \mathbb{E}[I_3(\theta_h)] \\
&\leq a_{h-1} + TL(M_1 + C)(\sqrt{\|\Delta_n\|} + \sqrt{\|\Delta_m\|}) + L\int_0^t \mathbb{E}[|X_s^n - X_s^m|]ds \\
&\quad + \frac{2H^2T}{h} + \|\theta_h''\|TH^2(M_1 + C)(\sqrt{\|\Delta_n\|} + \sqrt{\|\Delta_m\|}) \\
&\leq 3\varepsilon + L\int_0^t \mathbb{E}[|X_s^n - X_s^m|]ds.
\end{aligned}$$

By Gronwall's inequality we conclude then that

$$\mathbb{E}[|X_t^n - X_t^m|] \leq 3e^{Lt}\varepsilon \leq 3e^{LT}\varepsilon,$$

for all $n, m \geq n_\varepsilon$ and all $t \in [0, T]$. Hence,

$$\begin{aligned}
\mathbb{E}\left[\int_0^T |X_t^n - X_t^m| dt\right] &= \int_0^T \mathbb{E}[|X_t^n - X_t^m|]dt \\
&\leq T \sup_{t \in [0,T]} \mathbb{E}[|X_t^n - X_t^m|] \\
&\leq 3Te^{LT}\varepsilon.
\end{aligned}$$

The claim of step three is proved.

Step four: $\{X^n\}_{n\geq 1}$ *is a Cauchy sequence in* $L^1(\Omega; C([0,T]))$.

We know that $\{X^n\}_{n\geq 1}$ is a Cauchy sequence in $L^1([0,T]\times\Omega)$ which is a complete space. We can therefore conclude that there exists a stochastic process $X \in L^1([0,T]\times\Omega)$ such that

$$\lim_{n\to\infty} \mathbb{E}\left[\int_0^T |X_t^n - X_t| dt\right] = 0.$$

From Step two we can also deduce that

$$\lim_{n\to\infty} \mathbb{E}\left[\int_0^T |X_{\eta_n(t)}^n - X_t| dt\right] = 0.$$

Hence, there exists a subsequence (we keep the same indexes though for easy notations) such that

$$\lim_{n\to\infty} X_t^n(\omega) = \lim_{n\to\infty} X_{\eta_n(t)}^n(\omega) = X_t(\omega) \quad dt \times d\mathbb{P}\text{-almost surely.}$$

Since the process $\{X_t^n\}_{t\in[0,T]}$ is $\{\mathcal{F}_t\}_{t\in[0,T]}$-adapted for any $n \in \mathbb{N}$ and almost sure convergence preserves measurability, we deduce that $\{X_t\}_{t\in[0,T]}$ is also $\{\mathcal{F}_t\}_{t\in[0,T]}$-adapted. To prove the continuity of $\{X_t\}_{t\in[0,T]}$ we need to check the convergence in the uniform topology, i.e. we need to estimate $\mathbb{E}\left[\sup_{t\in[0,T]}|X_t^n - X_t^m|\right]$.
As before we employ the representation (3.3.13):

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\in[0,T]} |X_t^n - X_t^m|\right] &\leq \mathbb{E}\left[\sup_{t\in[0,T]} \int_0^t |\bar{a}(Y_{\eta_n(s)}, X_{\eta_n(s)}^n) - \bar{a}(Y_{\eta_m(s)}, X_{\eta_m(s)}^m)| ds\right] \\
&\quad + \mathbb{E}\left[\sup_{t\in[0,T]} \left|\int_0^t (\bar{g}(Y_{\eta_n(s)}, X_{\eta_n(s)}^n) - \bar{g}(Y_{\eta_m(s)}, X_{\eta_m(s)}^m)) dW_s^2\right|\right] \\
&\leq \int_0^T \mathbb{E}[|\bar{a}(Y_{\eta_n(s)}, X_{\eta_n(s)}^n) - \bar{a}(Y_{\eta_m(s)}, X_{\eta_m(s)}^m)|] ds \\
&\quad + \mathbb{E}\left[\sup_{t\in[0,T]} \left|\int_0^t (\bar{g}(Y_{\eta_n(s)}, X_{\eta_n(s)}^n) - \bar{g}(Y_{\eta_m(s)}, X_{\eta_m(s)}^m)) dW_s^2\right|^2\right]^{\frac{1}{2}} \\
&=: J_1 + J_2
\end{aligned}$$

To treat J_1 we proceed as before; using inequality (3.3.19) we obtain

$$\begin{aligned} J_1 &= \int_0^T \mathbb{E}[|\bar{a}(Y_{\eta_n(s)}, X^n_{\eta_n(s)}) - \bar{a}(Y_{\eta_m(s)}, X^m_{\eta_m(s)})|]ds \\ &\leq L\int_0^T \mathbb{E}[|X^n_{\eta_n(s)} - X^m_{\eta_m(s)}|]ds + \int_0^T \mathbb{E}[|Y_{\eta_n(s)} - Y_{\eta_m(s)}|]ds \qquad (3.3.21) \\ &\leq TL(M_1 + C)(\sqrt{\|\Delta_n\|} + \sqrt{\|\Delta_m\|}) + L\int_0^T \mathbb{E}[|X^n_s - X^m_s|]ds. \end{aligned}$$

Since we proved in Step three that $\{X^n\}_{n\geq 1}$ is a Cauchy sequence in $L^1([0,T[\times\Omega)$ and by assumption $\|\Delta_n\|$ tends to zero as n tends to infinity, we can find n and m big enough to make the last row of the previous chain of inequalities smaller than any positive ε. We now evaluate J_2. Invoking the Doob maximal inequality, i.e.

$$E\left[\left(\sup_{t\in[a,b]} X_t\right)^2\right] \leq 2^2 E[X_b^2] \quad \text{where } \{X_t\}_{t\in[a,b]} \text{ is a non negative submartingale}$$

(see for instance Karatzas and Shreve [12] page 14) and Itô isometry we can write

$$\begin{aligned} J_2 &= \left(\mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^t (\bar{g}(Y_{\eta_n(s)}, X^n_{\eta_n(s)}) - \bar{g}(Y_{\eta_m(s)}, X^m_{\eta_m(s)}))dW^2_s\right|^2\right]\right)^{\frac{1}{2}} \\ &\leq 2\left(\mathbb{E}\left[\left|\int_0^T (\bar{g}(Y_{\eta_n(s)}, X^n_{\eta_n(s)}) - \bar{g}(Y_{\eta_m(s)}, X^m_{\eta_m(s)}))dW^2_s\right|^2\right]\right)^{\frac{1}{2}} \\ &= 2\left(\mathbb{E}\left[\int_0^T |\bar{g}(Y_{\eta_n(s)}, X^n_{\eta_n(s)}) - \bar{g}(Y_{\eta_m(s)}, X^m_{\eta_m(s)})|^2 ds\right]\right)^{\frac{1}{2}} \\ &\leq 2H\left(\mathbb{E}\left[\int_0^T \left(\sqrt{|X^n_{\eta_n(s)} - X^m_{\eta_m(s)}|} + \sqrt{|Y_{\eta_n(s)} - Y_{\eta_m(s)}|}\right)^2 ds\right]\right)^{\frac{1}{2}} \\ &\leq 2\sqrt{2}H\left(\mathbb{E}\left[\int_0^T |X^n_{\eta_n(s)} - X^m_{\eta_m(s)}| + |Y_{\eta_n(s)} - Y_{\eta_m(s)}|ds\right]\right)^{\frac{1}{2}} \\ &= 2\sqrt{2}H\left(\int_0^T \mathbb{E}[|X^n_{\eta_n(s)} - X^m_{\eta_m(s)}|] + \mathbb{E}[|Y_{\eta_n(s)} - Y_{\eta_m(s)}|]ds\right)^{\frac{1}{2}}. \end{aligned}$$

If we now observe that the last member above is equivalent to (3.3.21), we can proceed as before and conclude that for any $\varepsilon > 0$ there exists $n_\varepsilon \in \mathbb{N}$ such that

$$\mathbb{E}\left[\sup_{t\in[0,T]} |X^n_t - X^m_t|\right] < \varepsilon \quad \text{for all } n, m \geq n_\varepsilon.$$

This proves that $\{X_n\}_{n\geq 1}$ is a Cauchy sequence in $L^1(\Omega; C([0,T])$ and thus

$$\lim_{n\to\infty} \mathbb{E}\left[\sup_{t\in[0,T]} |X_t^n - X_t|\right] = 0$$

where $\{X_t\}_{t\in[0,T]}$ is the stochastic process obtained in Step three. Moreover, we can find a subsequence (we keep the same indexes though for easy notations) such that

$$\lim_{n\to\infty} \sup_{t\in[0,T]} |X_t^n(\omega) - X_t(\omega)| = 0 \quad d\mathbb{P}\text{-almost surely.}$$

Since the processes $\{X_t^n\}_{t\in[0,T]}$ are continuous by construction for each $n \in \mathbb{N}$, we deduce that the process $\{X_t\}_{t\in[0,T]}$ is also continuous being a uniform limit of continuous functions.

Step five: *The stochastic process $\{X_t\}_{t\in[0,T]}$ solves equation (3.3.9).*

Finally we show that

$$\mathbb{P}\left(X(t) = x + \int_0^t \bar{a}(Y_s, X_s)ds + \int_0^t \bar{g}(Y_s, X_s)dW_s^2 \quad \text{for all } t \in [0,T]\right) = 1.$$

This in turn will be proven by showing that

$$\mathbb{E}\left[\sup_{t\in[0,T]} \left|X_t - x - \int_0^t \bar{a}(Y_s, X_s)ds - \int_0^t \bar{g}(Y_s, X_s)dW_s^2\right|\right] = 0$$

In fact, the equality

$$\begin{aligned} & X_t - x - \int_0^t \bar{a}(Y_s, X_s)ds - \int_0^t \bar{g}(Y_s, X_s)dW_s^2 \\ = \; & X_t - X_{\eta_n(t)}^n + \int_0^t \bar{a}(Y_{\eta_n(s)}, X_{\eta_n(s)}^n) - \bar{a}(Y_s, X_s)ds \\ & + \int_0^t \bar{g}(Y_{\eta_n(s)}, X_{\eta_n(s)}^n) - \bar{g}(Y_s, X_s)dW_s^2 \end{aligned}$$

implies

$$\begin{aligned} & \sup_{t\in[0,T]} \left|X_t - x - \int_0^t \bar{a}(Y_s, X_s)ds - \int_0^t \bar{g}(Y_s, X_s)dW_s^2\right| \\ \leq \; & \sup_{t\in[0,T]} |X_t - X_{\eta_n(t)}^n| + \int_0^T |\bar{a}(Y_{\eta_n(s)}, X_{\eta_n(s)}^n) - \bar{a}(Y_s, X_s)|ds \\ & + \sup_{t\in[0,T]} \left|\int_0^t \bar{g}(Y_{\eta_n(s)}, X_{\eta_n(s)}^n) - \bar{g}(Y_s, X_s)dW_s^2\right|. \end{aligned}$$

If we take the expectation and use the technique utilized in Step four to bound the terms in the right hand side of the previous inequality we get

$$\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[0,T]}\left|X_t - x - \int_0^t \bar{a}(Y_s,X_s)ds - \int_0^t \bar{g}(Y_s,X_s)dW_s^2\right|\right]\\
= &\lim_{n\to\infty}\mathbb{E}\left[\sup_{t\in[0,T]}\left|X_t - x - \int_0^t \bar{a}(Y_s,X_s)ds - \int_0^t \bar{g}(Y_s,X_s)dW_s^2\right|\right]\\
\leq &\lim_{n\to\infty}\left(\mathbb{E}\left[\sup_{t\in[0,T]}|X_t - X^n_{\eta_n(t)}|\right] + \mathbb{E}\left[\int_0^T |\bar{a}(Y_{\eta_n(s)},X^n_{\eta_n(s)}) - \bar{a}(Y_s,X_s)|ds\right]\right)\\
&+\lim_{n\to\infty}\mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^t \bar{g}(Y_{\eta_n(s)},X^n_{\eta_n(s)}) - \bar{g}(Y_s,X_s)dW_s^2\right|\right]\\
= &\ 0.
\end{aligned}$$

Uniqueness: We use a standard approach. Let $\{X_t\}_{t\in[0,T]}$ and $\{Z_t\}_{t\in[0,T]}$ be two strong solutions of equation (3.3.9). Setting,

$$\delta_t := X_t - Z_t = \int_0^t [\bar{a}(Y_s,X_s) - \bar{a}(Y_s,Z_s)]ds + \int_0^t [\bar{g}(Y_s,X_s) - \bar{g}(Y_s,Z_s)]dW_s^2 \quad (3.3.22)$$

we get by the Itô formula

$$\begin{aligned}
\theta_h(\delta_t) = &\int_0^t \theta_h'(\delta_s)[\bar{a}(Y_s,X_s) - \bar{a}(Y_s,Z_s)]ds\\
&+\int_0^t \theta_h'(\delta_s)[\bar{g}(Y_s,X_s) - \bar{g}(Y_s,Z_s)]dW_s^2\\
&+\frac{1}{2}\int_0^t \theta_h''(\delta_s)[\bar{g}(Y_s,X_s) - \bar{g}(Y_s,Z_s)]^2 ds
\end{aligned}$$

where $\{\theta_h\}_{h\geq 0}$ is the collection of functions defined in Step three. Using the assumptions on $\bar{a}$ and $\bar{g}$ and the bounds $|\theta_h'(u)|\leq 1$ and $|\theta_h''(u)|\leq \frac{2}{hu}$ we get

$$\begin{aligned}
\mathbb{E}[\theta_h(\delta_t)] \leq &\ \mathbb{E}\left[\int_0^t \theta_h'(\delta_s)[\bar{a}(Y_s,X_s) - \bar{a}(Y_s,Z_s)]ds\right] + \frac{tH^2}{h}\\
\leq &\ L\int_0^t \mathbb{E}[|\delta_s|]ds + \frac{tH^2}{h}
\end{aligned}$$

If we let $h\to\infty$, the function θ_h approaches the absolute value function; hence, Gronwall's inequality and sample path continuity imply that $\{X_t\}_{t\in[0,T]}$ and $\{Z_t\}_{t\in[0,T]}$ are indistinguishable. □

Chapter 4

On a general model system related to affine stochastic differential equations

4.1 Introduction

Stochastic differential equations (SDEs, for short) with Hölder-continuous coefficients appear in the modeling of several evolutionary systems perturbed by noise. The most important instance is probably the so-called *square root process* defined to be the unique strong solution of the following one dimensional SDE

$$dX_t = (aX_t + b)dt + \sigma\sqrt{X_t}dW_t, \quad X_0 = x \tag{4.1.1}$$

where $a, b \in \mathbb{R}$, $\sigma, x \in]0, +\infty[$ and $\{W_t\}_{t \geq 0}$ denotes a standard one dimensional Brownian motion. This equation is very popular in interest rate modeling due to the properties of its solution. We refer the reader to the book Cairns [13] for a detailed analysis of this topic (see also Mao [21]). SDEs with Hölder-continuous coefficients appear in the description of certain epidemic models as well: in this case the solution process represents the number of susceptible individuals in a given population. We mention the papers Greenhalgh et al. [1] and Bernardi et al. [15] which consider models described by SDEs with random and Hölder-continuous coefficients.

From a mathematical point of view the analysis of existence and uniqueness for strong solutions of SDEs with Hölder-continuous coefficients is quite challenging. In the one dimensional case, resorting to the famous Yamada-Watanabe principle (i.e. weak existence plus pathwise uniqueness implies strong existence) one can prove the existence of a unique strong solution for SDEs where the drift coefficient is locally Lipshiptz-continuous while the diffusion coefficient is of the type $\sigma(x) = |x|^\alpha$ for $\alpha \in [1/2, 1]$. The hard part of this proof is the pathwise uniqueness which heavily relies on an *ad hoc* technique introduced by Yamada and Watanabe [10] (see also the books Ikeda and Watanabe [11] and Karatzas and Shreve [12] for comparison theorems obtained with a similar approach). When we move to systems of SDEs with Hölder-continuous coefficients, then only few particular cases can be found in the literature; in fact, the lack of a multidimensional version of the Yamada-Watanabe technique to prove pathwise uniqueness forced the authors of those papers to consider equations that can be investigated with a slight modification of the one dimensional approach. The most important paper in this stream of results is certainly Duffie and Kan [2] where the authors, motivated by financial applications, consider a multidimensional version of the square root process (4.1.1). They prove existence, uniqueness and positivity for the strong solution of an SDE where the components of the drift vector are affine functions of the solution and the diffusion matrix is a constant matrix times a diagonal matrix with entries being square roots of affine functions of the solution. Their proof is based on a suitable application of the comparison theorem mentioned above, which we recall is based on the one dimensional Yamada-Watanabe technique. We now mention a series of results where the Yamada-Watanabe approach has been utilised in some multidimensional problems: Graczyk and J. Malecki [39] and Kumar [40] consider SDEs where for $i \in \{1, ..., m\}$ the i-th row of the diffusion matrix depends only on the i-th component of the solution; Luo [41] investigates a nested system of SDEs where the i-th row of the diffusion matrix depends only on the first i components of the solution; Wand and Zhang [42] introduce an integrability condition involving the determinant of the diffusion matrix and an auxiliary function fulfilling certain requirements.

The aim of the present paper is to link the general method presented in the book Allen [36] for modeling random phenomena using SDEs to the multidimensional system studied in Duffie and Kan [2]. More precisely, in Allen [36] pages 138-139 it is shown how, assigning probabilities to the possible changes of a general two dimensional system, one can deduce a Fokker-Planck partial differential equation for the candidate density of the system and from that a suitable SDE describing the random motion of the system. Following this procedure we consider an m-dimensional system with some prescribed admissible (i.e. with positive probability) changes and we deduce after some simplifying assumptions an m-dimensional SDE with Hölder continuous coefficients. Then, Taylor-expanding up to the first order the coefficients of the SDE around the initial condition, we end up with the multidimensional SDE investigated in Duffie-Kan [2] for which the existence of a unique strong solution is guaranteed under proper restrictions (we also present a detailed proof of this result, elaborating some technical aspects missing in the original proof). Therefore, this general construction emphasises the central role of the Duffie-Kan SDE as a model for first order approximations of a wide class of nonlinear systems perturbed by noise. We also remark that the positivity property guaranteed by the Duffie-Kan theorem entails the consistency of our procedure: in fact, such property will ensure the positivity of the probabilities originally assigned to the m-dimensional system according to the Allen's method. We then specialise to the two dimensional case and we suggest a direct proof of the Duffie-Kan theorem which does not passes through the comparison with an auxiliary process. Our proof is based on the sole properties of the one dimensional square root process (4.1.1) and produces a scheme to obtain an explicit solution of the two dimensional system once the process in (4.1.1) is assigned.

The paper is organised as follows: in Section 2 we adapt the Allen's procedure to an m-dimensional system assigning probabilities of admissible changes and making some simplifying assumptions; Section 3 contains the description of the first order approximation, link to the Duffie-Kan SDE, statement and detailed proof of the Duffie-Kan theorem; lastly, in Section 4 we specialise to the two dimensional framework and propose a constructive alternative proof of the Duffie-Kan theorem.

4.2 A general m-dimensional system

Let us consider a model system with $m \in \mathbb{N}$ different states evolving in time according to some probabilistic rules specified below. We write

$$S_t = (S_t^1, S_t^2, ..., S_t^m)^T, \quad t \geq 0$$

to represent the values of the m states of the system at time t.

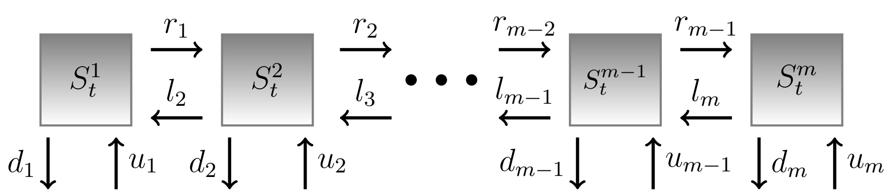

Figure 4.1: An m-state dynamical process

It is assumed that in a small time interval $[t, t+\Delta t]$ every state can change by -1, 0 or $+1$. This produces a total of 3^m possible different changes (the number of vectors of length m with components taking values in the set $\{-1, 0, 1\}$). We let $\Delta S_t := S_{t+\Delta t} - S_t$ be the global change of the system in the time interval $[t, t+\Delta t]$; for instance, $\Delta S_t = (-1, 0, 1, 0, ..., 0)^T$ means that in the time interval $[t, t+\Delta t]$ state S^1 has decreased of one unit, state S^3 has increased of one unit while all the other states remained unchanged. As illustrated in Figure 4.1, we denote

$$
\begin{aligned}
r_j(t,x) &:= \mathbb{P}(\Delta S_t = -e_j + e_{j+1} | S_t = x)/\Delta t, \quad j \in \{1, ..., m-1\} && (4.2.1)\\
l_j(t,x) &:= \mathbb{P}(\Delta S_t = -e_j + e_{j-1} | S_t = x)/\Delta t, \quad j \in \{2, ..., m\} && (4.2.2)\\
d_j(t,x) &:= \mathbb{P}(\Delta S_t = -e_j | S_t = x)/\Delta t, \quad j \in \{1, ..., m\} && (4.2.3)\\
u_j(t,x) &:= \mathbb{P}(\Delta S_t = e_j | S_t = x)/\Delta t, \quad j \in \{1, ..., m\} && (4.2.4)\\
p_0(t,x) &:= 1 - \Delta t \cdot \sum_{j=1}^{m} (r_j(t,x) + l_j(t,x) + d_j(t,x) + u_j(t,x)) && (4.2.5)
\end{aligned}
$$

where $\{e_1, ..., e_m\}$ denotes the canonical base of $\mathbb{R}^m$ and $r_m(t,x) = l_1(t,x) \equiv 0$. We remark that the probabilities associated to those changes not specified by (4.2.1)-(4.2.5) are identically zero. We also observe that $p_0(t,x)$ represents the probability of no changes during the interval $[t, t+\Delta t]$ given that $S_t = x$. According to Figure 4.1 the evolution of the states of the system is determined by interaction between the neighboring states(r_j's and l_j's) and exchanges with the outside world(u_j's and d_j's).

Given the probabilities (4.2.1)-(4.2.5) one can introduce, following Allen [36] pages 137-139, a Fokker-Planck equation solved by the density $p(t,x) := \mathbb{P}(S_t = x)$ of the system which in turn is related to the stochastic differential equation

$$
\begin{cases} dS_t = \mu(t, S_t)dt + B(t, S_t)dW_t \\ S_0 = s \end{cases} \tag{4.2.6}
$$

where $\{W_t\}_{t\geq 0}$ is an m-dimensional standard Brownian motion,

$$
\mu(t,x) := \mathbb{E}\left[\Delta S_t | S_t = x\right]/\Delta t
$$

is the mean vector and $B(t,x)$ denotes the symmetric square root of the covariance matrix

$$
V(t,x) := \mathbb{E}\left[(\Delta S_t)(\Delta S_t)^T | S_t = x\right]/\Delta t.
$$

According to equations (4.2.1)-(4.2.4) we can write

$$
\mu(t,x) = \left(-r_1(t,x) + l_2(t,x) + u_1(t,x) - d_1(t,x)\right) e_1
$$

$$+\sum_{j=2}^{m-1}\left(r_{j-1}(t,x)-r_j(t,x)+l_{j+1}(t,x)-l_j(t,x)+u_j(t,x)-d_j(t,x)\right)e_j$$
$$+\left(r_{m-1}(t,x)-l_m(t,x)+u_m(t,x)-d_m(t,x)\right)e_m \tag{4.2.7}$$

and

$$V(t,x)=\sum_{j=1}^{m}(u_j(t,x)+d_j(t,x))e_j\otimes e_j+\sum_{j=1}^{m-1}(r_j(t,x)+l_{j+1}(t,x))M_j \tag{4.2.8}$$

where for $j\in\{1,...,m-1\}$ we set $M_j:=(e_j-e_{j+1})\otimes(e_j-e_{j+1})$. We remark that the previous general system has been proposed in Bernardi et al. [43] as a model to study risks aggregation in a Bonus-Malus migration system. To proceed in the analysis of the SDE (4.2.6) we need to find the symmetric square root of the matrix $V(t,x)$. To this aim we assume the following.

Assumption 4.2.1. *For any $i,j\in\{1,...,m\}$ we have*

$$u_i(t,x)+d_i(t,x)=u_j(t,x)+d_j(t,x)=:\gamma(t,x)$$

and for any $i,j\in\{1,...,m-1\}$ we have

$$r_i(t,x)+l_{i+1}(t,x)=r_j(t,x)+l_{j+1}(t,x)=:\theta(t,x).$$

Assumption 4.2.1 introduces some symmetries in the evolution of our system. More precisely, the first condition implies that each state has the same probability of an exchange with the outside, while the second condition means that the probability of exchanges between neighboring states does not depend on the specific states considered. As a result we can now rewrite equation (4.2.8) in the simplified form

$$V(t,x)=\gamma(t,x)I+\theta(t,x)M \tag{4.2.9}$$

where I is the $m\times m$ identity matrix while M is the $m\times m$ matrix defined as

$$M=\begin{bmatrix}1 & -1 & 0 & 0 & 0 & \cdots & 0\\ -1 & 2 & -1 & 0 & 0 & \cdots & 0\\ 0 & -1 & 2 & -1 & 0 & \cdots & 0\\ 0 & 0 & -1 & 2 & -1 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & 0 & \cdots & -1 & 1\end{bmatrix}$$

According to Theorem 4 page 73 in Yueh [44] (with $a = c = -1$, $\alpha = \beta = \sqrt{ac} = 1$ and $b = 2$) the matrix M has m distinct eigenvalues of the form

$$\lambda_k = 2 + 2\cos(k\pi/m), \quad k = 1, ..., m \tag{4.2.10}$$

and hence there exists an orthogonal matrix Σ such that

$$M = \Sigma \mathcal{M} \Sigma^T \quad \text{with} \quad \mathcal{M} = \text{diag}\,[\lambda_1, \ldots, \lambda_m]\,.$$

Therefore, setting $y(t,x) := \theta(t,x)/\gamma(t,x)$ from equation (4.2.9) we deduce that

$$\begin{aligned} V(t,x) &= \gamma(t,x)\cdot(I + y(t,x)M) \\ &= \gamma(t,x)\cdot(I + y(t,x)\Sigma\mathcal{M}\Sigma^T) \\ &= \gamma(t,x)\cdot\Sigma(I + y(t,x)\mathcal{M})\Sigma^T. \end{aligned}$$

Since

$$(I + y(t,x)\mathcal{M})^{1/2} = \text{diag}\left[\sqrt{1 + y(t,x)\lambda_1}, ..., \sqrt{1 + y(t,x)\lambda_m}\right]$$

we conclude that

$$\begin{aligned} B(t,x) &= \sqrt{V(t,x)} \\ &= \sqrt{\gamma(t,x)}\cdot\Sigma\ \text{diag}\left[\sqrt{1 + y(t,x)\lambda_1}, ..., \sqrt{1 + y(t,x)\lambda_m}\right]\Sigma^T \\ &= \Sigma\ \text{diag}\left[\sqrt{\gamma(t,s) + \theta(t,x)\lambda_1}, ..., \sqrt{\gamma(t,x) + \theta(t,x)\lambda_m}\right]\Sigma^T. \end{aligned} \tag{4.2.11}$$

To sum up, given the probabilities (4.2.1)-(4.2.5) together with Assumption 4.2.1 our model system evolves according to the stochastic differential equation

$$\begin{cases} dS_t = & \mu(t,S_t)dt \\ & +\Sigma\ \text{diag}\left[\sqrt{\gamma(t,S_t) + \theta(t,S_t)\lambda_1}, ..., \sqrt{\gamma(t,S_t) + \theta(t,S_t)\lambda_m}\right]\Sigma^T dW_t \\ S_0 = & s \end{cases}$$

or equivalently

$$\begin{cases} dS_t = & \mu(t,S_t)dt \\ & +\Sigma\ \text{diag}\left[\sqrt{\gamma(t,S_t) + \theta(t,S_t)\lambda_1}, ..., \sqrt{\gamma(t,S_t) + \theta(t,S_t)\lambda_m}\right] d\tilde{W}_t \\ S_0 = & s \end{cases} \tag{4.2.12}$$

where $\tilde{W}_t := \Sigma^T W_t$ is a new m-dimensional standard Brownian motion (recall that by construction Σ^T is orthogonal) while $\mu(t,S_t)$ and the λ_j's are defined in (4.2.7) and (4.2.10), respectively.

4.3 First order approximation and the Duffie-Kan's theorem

The aim of the present section is to prove the existence of a unique strong solution for an SDE of the type (4.2.12) under suitable regularity assumptions on the coefficients of the equation. First of all we observe that according to equation (4.2.7) and Assumption 4.2.1 the components of the drift coefficient μ and the scalar functions γ and β are linear combinations of the functions r_j's, l_j's, u_j's and d_j's defined in (4.2.1)-(4.2.4).
If we assume for simplicity that the functions r_j's, l_j's, u_j's and d_j's are time independent and we expand each of them into its first order Taylor polynomial around the point s (which is the initial condition of the SDE (4.2.12)), then we obtain a corresponding family of affine functions on $\mathbb{R}^m$. Linear combinations of these affine functions will result in new affine functions representing the components of the drift coefficient μ and the scalar functions γ and θ. More precisely, introducing the notation $f^\star$ to denote the first order Taylor polynomial around s of the smooth function $f : \mathbb{R}^m \to \mathbb{R}$, i.e

$$\begin{aligned} f^\star : \mathbb{R}^m &\to \mathbb{R} \\ x &\mapsto f^\star(x) := f(s) + \langle \nabla f(s), x - s \rangle, \end{aligned}$$

we approximate the functions r_j's, l_j's, u_j's and d_j's with $r_j^\star$'s, $l_j^\star$'s, $u_j^\star$'s and $d_j^\star$'s, respectively. This results in the first order approximation of μ, γ and θ transforming equation (4.2.12) into

$$\begin{cases} dS_t = & \mu^\star(S_t)dt \\ & +\Sigma \text{ diag}\left[\sqrt{\gamma^\star(S_t) + \theta^\star(S_t)\lambda_1}, ..., \sqrt{\gamma^\star(S_t) + \theta^\star(S_t)\lambda_m}\right] d\tilde{W}_t \\ S_0 = & s \end{cases} \tag{4.3.1}$$

The SDE (4.3.1) now falls into the class of *affine* stochastic differential equations which is a class of equations having a relevant role in the theory of interest rate models (see for instance Cairns [13]). Existence, uniqueness and positivity for affine SDEs have been investigated in the remarkable paper Duffie and Kan [2]. Here we recall their main theorem together with a detailed proof.

Theorem 4.3.1 (Duffie and Kan [2]). *Consider the m-dimensional stochastic differential equation*

$$dS_t = (aS_t + b)dt + \Sigma \text{ diag}\left(\sqrt{v_1(S_t)}, \sqrt{v_2(S_t)}, ..., \sqrt{v_m(S_t)}\right) dW_t \tag{4.3.2}$$

where $a, \Sigma \in M_{m\times m}$, $b \in \mathbb{R}^m$ and $v_i(x) := \alpha_i + \langle \beta_i, x\rangle$ with $\alpha_1, ..., \alpha_m \in \mathbb{R}$ and $\beta_1, ..., \beta_m \in \mathbb{R}^m$. Assume that

1. *If $x \in \mathbb{R}^m$ is such that $v_i(x) = 0$, then*

$$\langle \beta_i, ax + b \rangle > |\Sigma^T \beta_i|^2/2$$

2. *For all $j \in \{1, ..., m\}$ if $(\Sigma^T \beta_i)_j \neq 0$, then $v_i(x) = v_j(x)$ for all $x \in \mathbb{R}^m$.*

Then, for any initial condition $S_0 = s \in \mathbb{R}^m$ belonging to

$$D := \{x \in \mathbb{R}^m : v_i(x) > 0 \text{ for all } i \in \{1, , , .m\}\}$$

the SDE (4.3.2) admits a unique global strong solution. Moreover, such solution satisfies for all $i \in \{1, ..., m\}$ and $t \geq 0$

$$v_i(S_t) > 0 \text{ almost surely.}$$

Proof. We first consider the case in which

$$v_i(x) = v(x) = \alpha + \langle \beta, x \rangle \quad \text{ for all } i \in \{1, ..., m\}$$

making the second assumption trivially satisfied. In this case equation (4.3.2) reduces to

$$dS_t = (aS_t + b)dt + \sqrt{v(S_t)}\Sigma dW_t. \tag{4.3.3}$$

Let $\{\varepsilon_n\}_{n \geq 1}$ be a positive strictly decreasing sequence of numbers converging to zero. For each $n \geq 1$, let $\{S_t^{(n)}\}_{t \geq 0}$ be the unique solution of the stochastic differential equation defined by (4.3.3) for $t \leq \tau_n := \inf\{r \geq 0 : v(S_r^{(n)}) = \varepsilon_n\}$ and by $S_t^{(n)} = S_{\tau_n}^{(n)}$ for $t \geq \tau_n$. This is the process satisfying (4.3.3) that is absorbed at the boundary $\{x \in \mathbb{R}^m : v(x) = \varepsilon_n\}$. Since the coefficient functions defining (4.3.3) are uniformly Lipschitz on the domain $\{x \in \mathbb{R}^m : v(x) \geq \varepsilon_n\}$, the process $\{S_t^{(n)}\}_{t \geq 0}$ is well defined and is a strong Markov process by standard SDE results.

With $\tau_0 = 0$ we can now define a unique process $\{S_t\}_{t \geq 0}$ on the closed time interval $[0, +\infty]$ by $S_t = S_t^{(n)}$ for $\tau_{n-1} \leq t \leq \tau_n$ and by $S_t = s$ for $t \geq \tau := \lim_{n \to +\infty} \tau_n$. If $\tau = +\infty$ almost surely, then $\{S_t\}_{t \geq 0}$ uniquely solves (4.3.3) on $[0, +\infty[$, as desired, and is strong Markov. To prove that $\tau = +\infty$ almost surely we will construct an auxiliary positive process that lower bounds $v(S_t)$. We begin by considering the scalar process

$$V_t := v(S_t) = \alpha + \langle \beta, S_t \rangle, \quad t \geq 0$$

which clearly satisfies

$$dV_t = \langle \beta, aS_t + b \rangle dt + \sqrt{V_t} \cdot \langle \beta, \Sigma dW_t \rangle. \tag{4.3.4}$$

If we set

$$\hat{W}_t := \langle \Sigma^T \beta, W_t \rangle / |\Sigma^T \beta|, \quad t \geq 0$$

we see that $\{\hat{W}_t\}_{t\geq 0}$ is a one dimensional Brownian motion and equation (4.3.4) can be rewritten as

$$dV_t = \langle \beta, aS_t + b \rangle dt + |\Sigma^T \beta| \sqrt{V_t} d\hat{W}_t. \tag{4.3.5}$$

According to the first assumption the inequality

$$\langle \beta_i, ax + b \rangle - |\Sigma^T \beta|^2/2 > 0$$

holds on the hyper-plane $v(x) = 0$. Therefore, by continuity there exists $\varepsilon > 0$ such that the previous inequality is valid on the strip $\{x \in \mathbb{R}^m : 0 \leq v(x) \leq \varepsilon\}$. We can assume without loss of generality that such ε coincides with ε_1. In particular, we can find a $\delta > 0$ such that

$$\langle \beta_i, ax + b \rangle - |\Sigma^T \beta|^2/2 > \delta \tag{4.3.6}$$

holds for all x belonging to the aforementioned strip. Denoting by $\bar{\eta} := |\Sigma^T \beta|^2/2 + \delta$ we have that

$$\langle \beta_i, ax + b \rangle > \bar{\eta} > |\Sigma^T \beta|^2/2 \tag{4.3.7}$$

on the set $\{x \in \mathbb{R}^m : 0 \leq v(x) \leq \varepsilon_1\}$. We can also assume that $V_0 > \varepsilon_1$.
We now introduce the *excursions* of the process V from ε_2 to ε_1. We set $T^\star(0) = 0$ and for $k \geq 1$ we define

$$T(k) := \inf\{t \geq T^\star(k-1) : V_t = \varepsilon_2\} \quad \text{and} \quad T^\star(k) := \inf\{t \geq T(k) : V_t = \varepsilon_1\}.$$

These stopping times realize a partition of $[0, +\infty[$:

$$0 = T^\star(0) < T(1) < T^\star(1) < T(2) < T^\star(2) < \cdots$$

In addition, we consider the auxiliary process $\{\hat{V}_t\}_{t\geq 0}$ defined as follows:

$$\begin{aligned} \hat{V}_t &= \varepsilon_2 + \int_{T(k)}^T \bar{\eta} ds + \int_{T(k)}^T |\Sigma^T \beta| \sqrt{\hat{V}_s} d\hat{W}_s, \quad \text{if } t \in [T(k), T^\star(k)] \\ \hat{V}_t &= V_t, \quad \text{if } t \in]T^\star(k), T(k+1)[\end{aligned}$$

The process $\{\hat{V}_t\}_{t\geq 0}$ satisfies

$$0 < \hat{V}_t \leq V_t \quad \text{for all } t \in [0, +\infty[. \tag{4.3.8}$$

In fact, when $t \in]T^\star(k), T(k)[$ then $\hat{V}_t = V_t$ and by the construction of the stopping times $V_t \geq \varepsilon_2 > 0$ on that time interval. On the other hand, when $t \in [T(k), T^\star(k)]$ then $\hat{V}_t$ is a one dimensional square root process satisfying the Feller condition $\bar{\eta} > |\Sigma^T \beta|^2/2$ (compare with the second inequality in (4.3.7)). This gives the positivity of $\hat{V}_t$. Moreover, recalling the dynamic of the process $\{V_t\}_{t\geq 0}$ in (4.3.5), the first inequality in (4.3.7) together with Theorem 1.1 page 437 in Ikeda and Watanabe [11] implies $\hat{V}_t \leq V_t$.
We now consider the general case: let $\{\varepsilon_n\}_{n\geq 1}$ be a positive strictly decreasing sequence of numbers converging to zero and define as before for each $n \geq 1$ the process $\{S_t^{(n)}\}_{t\geq 0}$ to be the solution of the stochastic differential equation defined by (4.3.2) for $t \leq \tau_n := \inf\left\{r \geq 0 : \min_{i\in\{1,\dots,d\}} v_i(S_r^{(n)}) = \varepsilon_n\right\}$ and by $S_t^{(n)} = S_{\tau_n}^{(n)}$ for $t \geq \tau_n$. This is the process satisfying (4.3.2) that is absorbed at the boundary $\left\{x \in \mathbb{R}^m : \min_{i\in\{1,\dots,d\}} v_i(x) = \varepsilon_n\right\}$. Since the coefficient functions defining (4.3.2) are uniformly Lipschitz on the domain $\left\{x \in \mathbb{R}^m : \min_{i\in\{1,\dots,d\}} v_i(x) \geq \varepsilon_n\right\}$, the process $\{S_t^{(n)}\}_{t\geq 0}$ is uniquely well defined and is a strong Markov process by standard SDE results.
With $\tau_0 = 0$ we can now define a unique process $\{S_t\}_{t\geq 0}$ on the closed time interval $[0, +\infty]$ by $S_t = S_t^{(n)}$ for $\tau_{n-1} \leq t \leq \tau_n$ and by $S_t = s$ for $t \geq \tau := \lim_{n\to+\infty} \tau_n$. If $\tau = +\infty$ almost surely, then $\{S_t\}_{t\geq 0}$ uniquely solves (4.3.2) on $[0, +\infty[$. For $i \in \{1, ..., m\}$ let

$$V_t^i := v_i(S_t) = \alpha_i + \langle \beta_i, S_t \rangle, \quad t \geq 0$$

which clearly satisfies

$$\begin{aligned}
dV_t^i &= \langle \beta_i, aS_t + b\rangle dt + \left\langle \beta_i, \Sigma \text{ diag}\left(\sqrt{V_t^1}, \sqrt{V_t^2}, ..., \sqrt{V_t^d}\right) dW_t \right\rangle \\
&= \langle \beta_i, aS_t + b\rangle dt + \left\langle \Sigma^T\beta_i, \text{ diag}\left(\sqrt{V_t^1}, \sqrt{V_t^2}, ..., \sqrt{V_t^d}\right) dW_t \right\rangle \\
&= \langle \beta_i, aS_t + b\rangle dt + \sum_{j=1}^m (\Sigma^T\beta_i)_j \cdot \sqrt{V_t^j} dW_t^j \\
&= \langle \beta_i, aS_t + b\rangle dt + \sum_{j\in\mathcal{C}_i} (\Sigma^T\beta_i)_j \cdot \sqrt{V_t^j} dW_t^j
\end{aligned}$$

where

$$\mathcal{C}_i := \left\{ j \in \{1, ..., m\} : (\Sigma^T\beta_i)_j \neq 0 \right\}.$$

According to the second assumption of the theorem, we have that $V_t^j = V_t^i$ for all $j \in \mathcal{C}_i$ and $t \geq 0$. Therefore,

$$dV_t^i = \langle \beta_i, aS_t + b\rangle dt + \sum_{j\in\mathcal{C}_i} (\Sigma^T\beta_i)_j \cdot \sqrt{V_t^j} dW_t^j$$

$$
\begin{aligned}
&= \langle \beta_i, aS_t + b\rangle dt + \sqrt{V_t^i}\sum_{j\in\mathcal{C}_i}(\Sigma^T\beta_i)_j dW_t^j \\
&= \langle \beta_i, aS_t + b\rangle dt + \sqrt{V_t^i} d\hat{W}_t^i
\end{aligned}
$$

with

$$
\hat{W}_t^i := \sum_{j\in\mathcal{C}_i}(\Sigma^T\beta_i)_j W_t^j / |\Sigma^T\beta_i|
$$

being a one dimensional Brownian motion (observe that $|\Sigma^T\beta_i|^2 = \sum_{j\in\mathcal{C}_i}(\Sigma^T\beta_i)_j^2$ by the definition of $\mathcal{C}_i$). One can now proceed as before introducing m auxiliary process $\hat{V}^i$ which satisfy $0 < \hat{V}_t^i \leq V_t^i$ for all $i \in \{1, ..., m\}$ and $t \geq 0$. This completes the proof. □

By means of the previous theorem we can now set concrete assumptions on the probabilities (4.2.1)-(4.2.4) for the existence of a unique strong solution for the SDE (4.3.1). These assumptions will also guarantee the non negativity of the probabilities in our original model system making the whole construction consistent. Before stating the result we recall that by Assumption 4.2.1 we have

$$
\gamma(x) = u_j(x) + d_j(x) \quad \text{for all } j \in 1, ..., m.
$$

Corollary 4.3.2. *If $\theta^\star \equiv 0$, $\gamma(s) > 0$ and the inequality*

$$
\langle \nabla\gamma(s), \mu^\star(x)\rangle > |\nabla\gamma(s)|^2/2 \quad \text{holds true on the set } \{x \in \mathbb{R}^m : \gamma(x) = 0\} \tag{4.3.9}
$$

then equation (4.3.1) admits a unique strong solution $\{S_t\}_{t\geq 0}$ such that $\gamma^\star(S_t) > 0$ almost surely for all $t \geq 0$.

Proof. We have simply to verify that our assumptions imply those of Theorem 4.3.1. First of all, $\theta^\star \equiv 0$ is by Assumptions 4.2.1 equivalent to $r_j^\star + l_j^\star \equiv 0$ for all $j \in \{1, ..., m\}$ and hence $r_j^\star = l_j^\star \equiv 0$. With $\theta^\star \equiv 0$ the system (4.3.1) reduces to

$$
\begin{cases}
dS_t = & \mu^\star(S_t)dt + \sqrt{\gamma^\star(S_t)}dW_t \\
S_0 = & s
\end{cases} \tag{4.3.10}
$$

Observe that $\Sigma\tilde{W}_t = W_t$ by definition of $\tilde{W}_t$ and orthogonality of Σ. Equation (4.3.10) trivially satisfies the second assumption of Theorem 4.3.1 since, in the notation of that theorem, $v_1(x) = \cdots = v_m(x)$. We are left with the verification of the first assumption in Theorem 4.3.1. We note that

$$
\gamma^\star(x) = \gamma(s) + \langle \nabla\gamma(s), x - s\rangle = \alpha + \langle \beta, x\rangle
$$

if $\beta := \nabla\gamma(s)$ and $\alpha := \gamma(s) - \langle \nabla\gamma(s), s\rangle$. Since $\mu^\star(x)$ corresponds to $ax + b$ using the orthogonality of Σ we get that (4.3.9) is equivalent to the first assumption of Theorem 4.3.1. □

We observe that from the previous corollary we get the positivity of

$$\gamma^\star(S_t) = u_j^\star(S_t) + d_j^\star(S_t), \quad j \in \{1, ..., m\}$$

which is the aggregated probability of an increase and a decrease for each single state.

4.4 Two dimensional system

We now focus our attention on the two dimensional version of the general model system presented above. For the sake of clarity we schematise in Figure 4.2 below the dynamic investigated in the present section

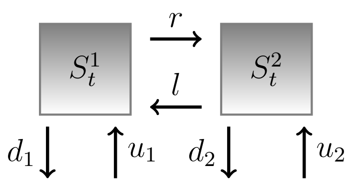

Figure 4.2: Two dimensional system

and we set

$$\begin{aligned}
r(t,x) &:= \mathbb{P}(\Delta S_t = (-1,1)|S_t = x)/\Delta t && (4.4.1)\\
l(t,x) &:= \mathbb{P}(\Delta S_t = (1,-1)|S_t = x)/\Delta t && (4.4.2)\\
d_1(t,x) &:= \mathbb{P}(\Delta S_t = (-1,0)|S_t = x)/\Delta t && (4.4.3)\\
u_1(t,x) &:= \mathbb{P}(\Delta S_t = (1,0)|S_t = x)/\Delta t && (4.4.4)\\
d_2(t,x) &:= \mathbb{P}(\Delta S_t = (0,-1)|S_t = x)/\Delta t && (4.4.5)\\
u_2(t,x) &:= \mathbb{P}(\Delta S_t = (0,1)|S_t = x)/\Delta t. && (4.4.6)
\end{aligned}$$

In addition, we denote

$$\begin{aligned}
p_0(t,x) &:= \mathbb{P}(\Delta S_t = (0,0)|S_t = x)\\
&= 1 - \Delta t \cdot (r(t,x) + l(t,x) + d_1(t,x) + u_1(t,x) + d_2(t,x) + u_2(t,x))
\end{aligned}$$

implying that

$$\mathbb{P}(\Delta S_t = (-1,-1)|S_t = x) = \mathbb{P}(\Delta S_t = (1,1)|S_t = x) = 0.$$

According to the scheme presented in the previous sections, if we employ the first order Taylor approximation of the functions defined above (which are assumed to be time independent), then the stochastic differential equation under investigation takes now the form

$$\begin{aligned} dS_t &= \mu^\star(S_t)dt + B^\star(S_t)dW_t \\ &= (aS_t + b)dt + \Sigma \begin{bmatrix} \sqrt{v_1(S_t)} & 0 \\ 0 & \sqrt{v_2(S_t)} \end{bmatrix} \Sigma^T dW_t \\ &= (aS_t + b)dt + \Sigma \begin{bmatrix} \sqrt{\alpha_1 + \langle \beta_1, S_t \rangle} & 0 \\ 0 & \sqrt{\alpha_2 + \langle \beta_2, S_t \rangle} \end{bmatrix} d\tilde{W}_t \end{aligned} \tag{4.4.7}$$

where for suitable choices of $a \in M_{2\times 2}$, $b, \beta_1, \beta_2 \in \mathbb{R}^2$ and $\alpha_1, \alpha_2 \in \mathbb{R}$ we find that

$$\begin{aligned} (aS_t + b)_1 &= u_1^\star(S_t) - d_1^\star(S_t) - r^\star(S_t) + l^\star(S_t) \\ (aS_t + b)_2 &= u_2^\star(S_t) - d_2^\star(S_t) + r^\star(S_t) - l^\star(S_t) \end{aligned}$$

(this follows from equation (4.2.7)) and

$$\alpha_1 + \langle \beta_1, S_t \rangle = d_1^\star(S_t) + u_1^\star(S_t) + 2(r^\star(S_t) + l^\star(S_t)) \tag{4.4.8}$$

$$\alpha_2 + \langle \beta_2, S_t \rangle = d_1^\star(S_t) + u_1^\star(S_t) \tag{4.4.9}$$

(which follows from equation (4.2.11)). We remark that in the present case

$$\lambda_1 = 2, \quad \lambda_2 = 0, \quad \gamma^\star(x) = d_1^\star(x) + u_1^\star(x) \quad \text{and} \quad \theta^\star(x) = r^\star(x) + l^\star(x).$$

and Assumption 4.2.1 reduces to

$$d_1^\star(x) + u_1^\star(x) = d_2^\star(x) + u_2^\star(x).$$

Moreover, we have

$$\Sigma = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

If we look through the proof of Theorem 4.3.1, we see that the second assumption in the statement of the theorem, namely

$$\textit{for all } j \in \{1, ..., m\} \textit{ if } (\Sigma^T \beta_i)_j \neq 0, \textit{ then } v_i(x) = v_j(x) \textit{ for all } x \in \mathbb{R}^m \tag{4.4.10}$$

serves to reduce the diffusion matrix

$$\operatorname{diag}\left(\sqrt{v_1(S_t)}, \sqrt{v_2(S_t)}, ..., \sqrt{v_m(S_t)}\right)$$

to one of the form $\sqrt{v_i(S_t)}I$ where I stands for $m \times m$ identity matrix. Therefore, there is no loss of generality in considering only the case

$$v_1(x) = v_2(x) = \cdots = v_m(x).$$

The next result is the two dimensional version of Theorem 4.3.1 for the case

$$v_1(x) = v_2(x) = \alpha + \langle \beta, x \rangle. \tag{4.4.11}$$

The proof is, however, different: it is based on a direct approach rather than the Yamada-Watanabe comparison method utilised in the proof of Theorem 4.3.1. This direct approach has the advantage of providing an explicit representation of the solution. Let us also point out that condition (4.4.11) together with (4.4.8) and (4.4.9) implies

$$r(x) = l(x) = 0.$$

With reference to Figure 4.2 this means that the interactions between the two states of the system take place in the probabilities u_1, u_2, d_1 and d_2 rather than from direct exchanges.

Theorem 4.4.1. *Consider the two dimensional stochastic differential equation*

$$dS_t = (aS_t + b)dt + \sqrt{\alpha + \langle \beta, S_t \rangle} dW_t, \quad S_0 = s \in \mathbb{R}^2 \tag{4.4.12}$$

where $a \in M_{2\times 2}$, $b, \beta \in \mathbb{R}^2$ and $\alpha \in \mathbb{R}$. If the inequality

$$\langle \beta, ax + b \rangle \geq |\beta|^2/2 \quad \textit{holds true on the set } \{x \in \mathbb{R}^2 : \alpha + \langle \beta, x \rangle = 0\} \tag{4.4.13}$$

then for any initial condition s satisfying $\alpha + \langle \beta, s \rangle > 0$ the SDE (4.4.12) admits a unique strong solution $\{S_t\}_{t \geq 0}$ with the property that $\alpha + \langle \beta, S_t \rangle > 0$ almost surely for all $t \geq 0$.

Proof. The idea of the proof is to reduce via an orthogonal transformation the system (4.4.12) to a system where the equation describing the first component is independent of the second. The first component will turn out to be a one dimensional square root process while the equation for the second component will be explicitly solvable once the first is known.

We may assume without loss of generality that $\beta \neq 0$ (if $\beta = 0$ then equation (4.4.12) admits a unique strong solution for any $\alpha \geq 0$). Let $K \in M_{2\times 2}$ be the unique orthogonal matrix such that $K\beta = |\beta| e_1$ and define the stochastic process $Y_t := KS_t$, $t \geq 0$. Then, by the linearity of the Itô differential we can write

$$dY_t \;=\; (KaS_t + Kb)dt + \sqrt{\alpha + \langle \beta, S_t \rangle} dKW_t$$

$$
\begin{aligned}
&= (KaK^{-1}KS_t + Kb)dt + \sqrt{\alpha + \langle \beta, K^{-1}KS_t\rangle}d\tilde{W}_t \\
&= (\tilde{a}Y_t + \tilde{b})dt + \sqrt{\alpha + \langle K\beta, Y_t\rangle}d\tilde{W}_t \\
&= (\tilde{a}Y_t + \tilde{b})dt + \sqrt{\alpha + |\beta|Y_t^1}d\tilde{W}_t \qquad (4.4.14)
\end{aligned}
$$

where $\tilde{a} := KaK^{-1}$, $\tilde{b} := Kb$ and $\tilde{W}_t := KW_t$ being a new two-dimensional standard Brownian motion. The initial condition is $Y_0 = KS_0 = Ks =: \tilde{s}$. We observe that condition (4.4.13) corresponds to

$$\tilde{a}_{11}y_1 + \tilde{a}_{12}y_2 + \tilde{b}_1 > |\beta|/2 \quad \text{holds true on the set } \{y \in \mathbb{R}^2 : \alpha + |\beta|y_1 = 0\} \quad (4.4.15)$$

Indeed,

$$
\begin{aligned}
\alpha + \langle \beta, x\rangle &= \alpha + \langle \beta, K^{-1}Kx\rangle \\
&= \alpha + \langle K\beta, Kx\rangle \\
&= \alpha + |\beta|y_1 \qquad (4.4.16)
\end{aligned}
$$

and

$$
\begin{aligned}
\langle \beta, ax + b\rangle &= \langle K^TK\beta, ax + b\rangle \\
&= |\beta|\langle e_1, Kax + Kb\rangle \\
&= |\beta|\langle e_1, KaK^{-1}y + \tilde{b}\rangle \\
&= |\beta|\langle e_1, \tilde{a}y + \tilde{b}\rangle \\
&= |\beta|\Big(\tilde{a}_{11}y_1 + \tilde{a}_{12}y_2 + \tilde{b}_1\Big)\,.
\end{aligned}
$$

Since the set $\{y \in \mathbb{R}^2 : \alpha + |\beta|y_1 = 0\}$ in (4.4.15) coincides with $\{y \in \mathbb{R}^2 : y_1 = -\alpha/|\beta|\}$, a substitution of the last condition in the inequality of (4.4.15) gives

$$\tilde{a}_{12}y_2 + \tilde{b}_1 > |\beta|/2 + (\alpha\tilde{a}_{11})/|\beta|.$$

The last inequality has to be true for all $y_2 \in \mathbb{R}$; hence, we get that $\tilde{a}_{12} = 0$ and

$$\tilde{b}_1 > |\beta|/2 + (\alpha\tilde{a}_{11})/|\beta|. \qquad (4.4.17)$$

Therefore, we can write equation (4.4.14) as

$$
\begin{cases}
dY_t^1 = (\tilde{a}_{11}Y_t^1 + \tilde{b}_1)dt + \sqrt{\alpha + |\beta|Y_t^1}d\tilde{W}_t^1, & Y_0^1 = \tilde{s}_1 \\
dY_t^2 = (\tilde{a}_{21}Y_t^1 + \tilde{a}_{22}Y_t^2 + \tilde{b}_2)dt + \sqrt{\alpha + |\beta|Y_t^1}d\tilde{W}_t^2 & Y_0^2 = \tilde{s}_2
\end{cases}
\qquad (4.4.18)
$$

Let us study the first equation in (4.4.18). Setting $\mathcal{Y}_t := |\beta|Y_t^1 + \alpha$ and applying the Itô formula we get

$$d\mathcal{Y}_t = \Big(\tilde{a}_{11}\mathcal{Y}_t + \tilde{b}_1|\beta| - \alpha\tilde{a}_{11}\Big)\,dt + |\beta|\sqrt{\mathcal{Y}_t}d\tilde{W}_t^1, \quad \mathcal{Y}_0 = |\beta|\tilde{s}_1 + \alpha.$$

The previous SDE has a unique positive solution (see e.g. Cairns [13]) if

$$\tilde{b}_1|\beta|-\alpha\tilde{a}_{11} \geq |\beta|^2/2$$

which corresponds to (4.4.17). The positivity of $\mathcal{Y}_t$ is equivalent to the positivity of $|\beta|Y_t^{(1)} + \alpha$ which in turn is equivalent by (4.4.16) to the positivity of $\alpha + \langle S_t, \beta\rangle$. We can now solve the equation for Y_t^2 in (4.4.18), namely

$$\begin{aligned} dY_t^2 &= \left(\tilde{a}_{21}Y_t^1 + \tilde{a}_{22}Y_t^2 + \tilde{b}_2\right) dt + \sqrt{\alpha + |\beta|Y_t^1} d\tilde{W}_t^2 \\ &= \tilde{a}_{22}Y_t^2 dt + \left[\left(\tilde{a}_{21}Y_t^1 + \tilde{b}_2\right) dt + \sqrt{\alpha + |\beta|Y_t^1} d\tilde{W}_t^2\right]. \end{aligned}$$

Its solution is give by the formula

$$Y_t^2 = e^{\tilde{a}_{22}t}\tilde{s}_2 + \int_0^T e^{\tilde{a}_{22}(t-s)} \left[\left(\tilde{a}_{21}Y_s^1 + \tilde{b}_2\right) ds + \sqrt{\alpha + |\beta|Y_s^1} d\tilde{W}_s^2\right].$$

Setting $S_t = K^{-1}Y_t$ we obtain the solution of the original system completing the proof. □

We now summarise the construction of the solution of the system (4.4.12) suggested in the previous proof:

- define the orthogonal matrix K imposing that $K\beta = |\beta|e_1$ and set $\tilde{a} := KaK^{-1}$, $\tilde{b} := Kb$, $\tilde{s} := Ks$ and $\tilde{W}_t := KW_t$
- let $\{\mathcal{Y}_t\}_{t\geq 0}$ to be unique positive strong solution of the (one dimensional) square root SDE

  $$d\mathcal{Y}_t = \left(\tilde{a}_{11}\mathcal{Y}_t + \tilde{b}_1|\beta|-\alpha\tilde{a}_{11}\right) dt + |\beta|\sqrt{\mathcal{Y}_t} d\tilde{W}_t^1, \quad \mathcal{Y}_0 = |\beta|\tilde{s}_1 + \alpha$$

  (note that the driving noise is $\tilde{W}_t^1$)
- set $Y_t^1 := (\mathcal{Y}_t - \alpha)/|\beta|$ and

  $$Y_t^2 := e^{\tilde{a}_{22}t}\tilde{s}_2 + \int_0^t e^{\tilde{a}_{22}(t-s)} \left[\left(\tilde{a}_{21}Y_s^1 + \tilde{b}_2\right) ds + \sqrt{\alpha + |\beta|Y_s^1} d\tilde{W}_s^2\right]$$

  (note that the driving noise is $\tilde{W}_t^2$)
- the process $S_t := K^{-1}Y_t$ solves (4.4.12).

In the following example we show that Theorem 4.3.1 without its second assumption no longer holds in general.

Example 4.4.2. *We consider the system*

$$\begin{cases} dX_t^1 = 2\sqrt{X_t^2 - 1}dW_t^1, & X_0^1 = x_1 \\ dX_t^2 = 3dt + 2\sqrt{X_t^2}dW_t^2, & X_0^2 = x_2. \end{cases} \tag{4.4.19}$$

In the notation of Theorem 4.3.1 it corresponds to

$$m = 2 \quad a = 0 \quad b = (0,3)^T \quad \Sigma = 2I \quad \alpha = (-1,0)^T \quad \beta_1 = (0,1)^T \quad \beta_2 = (0,1)^T.$$

Recalling that $v_i(x) = \alpha_i + \langle \beta_i, x \rangle$ *for* $i = 1,2$ *we get*

$$v_1(x) = -1 + x_2 \quad \text{and} \quad v_2(x) = x_2.$$

Since the second component of β_1 *is not zero and* $v_1 \neq v_2$*, the second condition of Theorem 4.3.1 does not hold. However, since* $a = 0$ *the first condition reduces to*

$$\langle \beta_i, b \rangle > |\beta_i|^2/2, \quad i = 1,2$$

which is clearly true. The positivity region D *is now given by* $D = \{x \in \mathbb{R}^2 : x_2 > 1\}$*. If the result of Theorem 4.3.1 were true we should be able to get a unique strong solution of (4.4.19) lying in* D *for all* $t \geq 0$ *almost surely.*
We observe that the process X^2 *in* (4.4.19) *falls in the class of the* squared Bessel processes, *i.e. processes that are strong solutions of SDEs of the form*

$$Z_t = z + 2\int_0^t \sqrt{Z_s}dB_s + \delta t$$

where $z, \delta \geq 0$ *(see Revuz and Yor [38] for a deep analysis of this family of processes). The parameters* δ *and* $\nu := \frac{\delta}{2} - 1$ *are called* dimension *and* index *of* Z*, respectively. It is well known that the transition density of* Z *is given by the formula*

$$f_t^\delta(z,y) = \frac{1}{2t}\left(\frac{y}{z}\right)^{\frac{\nu}{2}} e^{-\frac{z+y}{2t}} I_\nu\left(\frac{\sqrt{zy}}{t}\right) \mathbb{1}_{\{y>0\}}$$

where $I_\nu(z)$ *stands for the modified Bessel function of the first kind of order* ν*, i.e.*

$$I_\nu(z) = \sum_{n=0}^\infty \frac{(\frac{z}{2})^{\nu+2n}}{n!\,\Gamma(n+\nu+1)}, \quad \nu, z \in \mathbb{C}.$$

From this we see that $\mathbb{P}(0 < X_t^2 < 1) > 0$*, even starting with* $x_2 > 1$*. For instance, taking* $x_2 = 2$ *and* $t = 1$ *we have*

$$\mathbb{P}(0 < X_1^2 < 1) = \int_0^1 f_1^3(2,y)dy \approx 0.08.$$

This violates the positivity condition defined by $D = \{x \in \mathbb{R}^2 : x_2 > 1\}$ *which ensures* $\sqrt{X_t^2 - 1}$ *to be well defined.*

Appendices

Appendix A

The space $C[0,\infty)$, Weak Convergence and the Wiener measure

The "canonical" space for Brownian motion, is $C[0,\infty)$, the space of all continuous real-valued functions on $[0,\infty)$ with the metric

$$\rho(\omega_1,\omega_2) := \sum_{n=1}^{\infty} \frac{1}{2^n} \max_{0\le t\le n} \left(|\omega_1(t)-\omega_2(t)|\wedge 1\right). \tag{A.0.1}$$

In this appendix we show how to construct a measure, called the Wiener measure on this space so that the coordinate mapping process is Brownian Motion. This construction is given as the Donsker's invariance principle(also known as the functional central limit theorem) and involves the notion of weak convergence of random walks to brownian motion.

A.1 Weak Convergence

Definition A.1.1. *Let (S,ρ) be a metric space with a Borel sigma-field $\mathcal{B}(S)$. Let $\{P_n\}_{n=1}^{\infty}$ be a sequence of probability measures on $(S,\mathcal{B}(S))$, and let P be another probability measure on this space. We say that $\{P_n\}_{n=1}^{\infty}$* ***converges weakly*** *to P and $P_n \stackrel{w}{\to} P$ if and only if*

$$\lim_{n\to\infty} \int_S f(s)dP_n(s) = \int_S f(s)dP(s)$$

for every bounded, continuous, real-valued function f on S.

One can show that in particular from the definition above that the weak limit P is a probability measure and is unique.

Definition A.1.2. *Let $\{(\Omega_n, \mathcal{F}_n, P_n)\}_{n=1}^{\infty}$ be a sequence of probability spaces and on each of them consider a random variable X_n with values i the metric space (S, ρ). Let $(\Omega, \mathcal{F}, P)$ be another probability space on which a random variable X with values in (S, ρ) is given. We say that $\{X_n\}_{n=1}^{\infty}$ converges to X in distribution, and write $X_n \xrightarrow{\mathcal{D}} X$, if the sequence of measures $\{P_n X_n^{-1}\}_{n=1}^{\infty}$ converges weakly to the measure PX^{-1}.*

Equivalently $X_n \xrightarrow{\mathcal{D}} X$ if and only if

$$\lim_{n\to\infty} E_n(f(X_n)) = E(f(X))$$

for every bounded continuous function f on S, where E_n and E denote expectations with respect to P_n and P, respectively. Indeed since $\{P_n X_n^{-1}\}_{n=1}^{\infty}$ converges weakly to the measure PX^{-1} we have by definition that for all bounded continuous functions f on S that

$$\lim_{n\to\infty} \int_S f(s) dP_n X_n^{-1}(s) = \lim_{n\to\infty} \int_{\Omega_n} f(X_n(\omega)) dP_n(\omega) = \lim_{n\to\infty} E_n(f(X_n))$$

and

$$\int_S f(s) dPX^{-1}(s) = \int_{\Omega} f(X(\omega)) dP(\omega) = E(f(X))$$

which together imply convergence in distribution

$$\lim_{n\to\infty} E_n(f(X_n)) = E(f(X)).$$

The other direction is easily proven by a similar argument.

Recall that if S in Definition A.1.2 is $\mathbb{R}^d$, then $X_n \xrightarrow{\mathcal{D}} X$ if and only if the sequence of characteristic functions $\varphi_n(u) := E_n(\exp i(u, X_n))$ converges to $\varphi(u) := E(\exp i(u, X))$, for every $u \in \mathbb{R}^d$. This is called the Cramer Wold device and is a simple consequence of the celebrated Levy Continuity theorem.

The most important example of convergence in distribution is that provided by the central limit theorem. In the Lindeberg-Levy form used here, the theorem asserts that if $\{\psi_n\}_{n=1}^{\infty}$ is an i.i.d sequence of random variables with mean zero and variance σ^2, then $\{S_n\}$ defined by

$$S_n = \frac{1}{\sigma\sqrt{n}} \sum_{k=1}^{n} \psi_k$$

converges in distribution to a standard normal random variable. It is this fact that dictates that a properly normalized sequence of random walks will converge in distribution to a Brownian motion(Donsker's invariance principal).

Lemma A.1.3. *Suppose $\{X_n\}_{n=1}^{\infty}$ is a sequence of random variables taking values in a metric space (S_1, ρ_1) and converging in distribution to X. Suppose (S_2, ρ_2) is another metric space, and $\varphi : S_1 \to S_2$ is continuous. Show that $Y_n := \varphi(X_n)$ converges in distribution to $Y := \varphi(X)$.*

Proof. In order to show $Y_n \xrightarrow{\mathcal{D}} Y$ it is sufficient to show that(by definition of convergence in distribution) for all bounded continuous functions f we have that

$$\lim_{n\to\infty} E_n(f(\varphi(X_n))) = E(f(\varphi(X)))$$

Observe that the composition function $f\varphi$ is bounded and continuous since f is bounded and continuous and φ is continuous.The assumption that $X_n \xrightarrow{\mathcal{D}} X$ which implies that $\lim_{n\to\infty} E_n(g(X_n)) = E(g(X))$ for all bounded continuous g. In particular it is true for $g = f\varphi$. This completes the proof.

□

A.2 Tightness

Definition A.2.1. *Let (S, ρ) be a metric space and let Π be a family of probability measures on $(S, \mathcal{B}(S))$. We say that Π is relatively compact if every sequence of elements of Π contains a weakly convergent subsequence. We say that Π is tight if for every $\epsilon > 0$, there exists a compact set $K \subseteq S$ such that $P(K) \geq 1 - \epsilon$ for every $P \in \Pi$.*

If $\{X_\alpha\}_{\alpha\in A}$ is a family of random variables , each defined on a probability space $(\Omega_\alpha, \mathcal{F}_\alpha, P_\alpha)$ and taking values in S, we say that this family is relatively compact or tight if the family of induced measures $\{P_\alpha X_\alpha^{-1}\}_{\alpha\in A}$ has the appropriate property.

Theorem A.2.2. *Let Π be a family of probability measures on a complete separable metric space S. This family is relatively compact if and only if its tight*

Proof. For the proof refer to Convergence of Probability Measures by Billingsley(1968) pp.35-40 □

We are interested in the case $S = C[0, \infty)$. For this case we shall provide a characterization for tightness.To do so we will need the following definition

Definition A.2.3. *For each $\omega \in C[0, \infty), T > 0,$ and $\delta > 0$ the modulus of continuity on $[0, T]$ is defined as*

$$m^T(\omega, \delta) := \max_{\substack{|s-t|\leq\delta \\ 0\leq s;t\leq T}} |\omega(s) - \omega(t)| \tag{A.2.1}$$

Lemma A.2.4. *(Problem 2.4.8 Shreve) Show that $m^T(\omega, \delta)$ is continuous in $\omega \in C[0, \infty)$ under the metric ρ defined in (A.0.1), is non decreasing in δ and $\lim_{\delta\downarrow 0} m^T(\omega, \delta) = 0$ for each $\omega \in C[0, \infty)$.*

Proof. We first show that $\omega \mapsto m^T(\omega,\delta)$ is continuous with respect to the metric ρ i.e given a fixed $\delta > 0$ and $T > 0$ we can find a $\eta > 0$ such that

$$\text{whenever} \quad \rho(\omega_1,\omega_2) < \eta \text{ then} \quad |m^T(\omega_1,\delta) - m^T(\omega_2,\delta)| < \epsilon$$

.

$\rho(\omega_1,\omega_2) < \eta$ yields by the definition of ρ and choosing an $n^* \in \mathbb{N}$ such that $T+1 > n^* \geq T$

$$\max_{0\leq t\leq n^*} |\omega_1(t) - \omega_2(t)| < C_T\eta$$

and hence we can choose η such that

$$\max_{0\leq t\leq T} |\omega_1(t) - \omega_2(t)| < C_T\eta < \epsilon/3$$

Now we have by the triangular inequality that

$$\begin{aligned} |\omega_1(s) - \omega_1(t)| &= |\omega_1(s) - \omega_2(s) + \omega_2(s) - \omega_2(t) + \omega_2(t) - \omega_1(t)| \\ &\leq |\omega_1(s) - \omega_2(s)| + |\omega_2(s) - \omega_2(t)| + |\omega_2(t) - \omega_1(t)| \end{aligned} \tag{A.2.2}$$

And hence we have

$$\begin{aligned} |\omega_1(s) - \omega_1(t)| \leq |\omega_1(s) - \omega_2(s)| + |\omega_2(s) - \omega_2(t)| + |\omega_2(t) - \omega_1(t)| \\ \leq \epsilon/3 + |\omega_2(s) - \omega_2(t)| + \epsilon/3 = 2/3\epsilon + |\omega_2(s) - \omega_2(t)| \leq 2/3\epsilon + m^T(\omega_2,\delta) \end{aligned} \tag{A.2.3}$$

and therefore we have

$$|\omega_1(s) - \omega_1(t)| - m^T(\omega_2,\delta) \leq 2/3\epsilon$$

which yields after taking maximum on both the LHS and RHS

$$m^T(\omega_1,\delta) - m^T(\omega_2,\delta) \leq 2/3\epsilon < \epsilon$$

and a similar argument (by using triangular inequality on $|\omega_2(s) - \omega_2(t)|$) we can conclude that

$$|m^T(\omega_1,\delta) - m^T(\omega_2,\delta)| < \epsilon$$

In order to show that given a fixed $\omega \in C[0,\infty)$ and $T > 0$ the map $\delta \mapsto m^T(\omega,\delta)$ is non decreasing in δ we let $\delta_1 \leq \delta_2$ and then observe that

$$\max_{\substack{|s-t|\leq\delta_1 \\ 0\leq s,t\leq T}} |\omega(s) - \omega(t)| \leq \max_{\substack{|s-t|\leq\delta_2 \\ 0\leq s,t\leq T}} |\omega(s) - \omega(t)|$$

since

$$\left\{{}^{|s-t|\leq\delta_1}_{0\leq s,t\leq T}\right\} \subseteq \left\{{}^{|s-t|\leq\delta_2}_{0\leq s,t\leq T}\right\}$$

Now since $\omega \in C[0,\infty)$ it is uniformly continuous on $[0,T]$ and hence uniform continuity of ω implies

$$\lim_{\delta\downarrow 0} \max_{\substack{|s-t|\leq\delta \\ 0\leq s;t\leq T}} |\omega(s)-\omega(t)|=0$$

which completes the proof.

□

In the sequel we will need the following version of the Arzelà-Ascoli theorem

Theorem A.2.5. *(Theorem 2.4.9 Shreve) A set $A \subseteq C[0,\infty)$ ha s a compact closure if and only if the following conditions hold:*

$$\sup_{\omega\in A}|\omega(0)|<\infty \tag{A.2.4}$$

$$\lim_{\delta\downarrow 0}\sup_{\omega\in A} m^T(\omega,\delta)=0 \text{ for every } T>0. \tag{A.2.5}$$

Proof. Assume the closure of A is denoted by $\bar{A}$, is compact. $\bar{A}$ is contained in the union of open sets

$$G_n=\{\omega\in C[0,\infty):\omega(0)<n\}\ n=1,2,\dots$$

Indeed we have that $\omega\in\bar{A} \implies \omega\in C[0,\infty)$ and hence by continuity $\forall\omega\in\bar{A}\exists n\in\mathbb{N}$ such that $\omega(0)<n$ and hence $\{G_n\}_{n=1}^{\infty}$ covers $\bar{A}$. Each G_n is open because it is a cylinder set which are by definition open. By the definition of compactness of $\bar{A}$ we know that every open cover of $\bar{A}$ has a finite sub-cover and since $G_1\subseteq G_2\subseteq\cdots$, we have that $A\subseteq\cup_{i=1}^{n}G_i=G_n$ And hence this means $\omega\in A \implies \omega\in G_n$ and hence by the definition of G_n above we have $\sup_{\omega\in A}\omega(0)\leq\sup_{\omega\in G_n}\omega(0)<n<\infty$ and hence we have shown (A.2.4).

Now for each ϵ, let $K_\delta=\{\omega\in\bar{A}; m^T(\omega,\delta)\geq\epsilon\}$. Since $\omega\mapsto m^T(\omega,\delta)$ is continuous, the set $\{\omega\in C[0,\infty):m^T(\omega,\delta)\geq\epsilon\}$ is closed as the inverse image of a closed set by a continuous function. Since closed subset of compact sets are compact, K_δ is compact. Indeed $K_\delta=\bar{A}\cap\{\omega\in C[0,\infty):m^T(\omega,\delta)\geq\epsilon\}$.

Lemma A.2.4 implies $\cap_{\delta>0}K_\delta=\phi$ since for a fixed $\epsilon>0$ there exists a $\delta(\omega)>0$ for which $m^T(\omega,\delta)<\epsilon$ as $\lim_{\delta\downarrow 0}m^T(\omega,\delta)=0$ for all $\omega\in C[0,\infty)$. So for some $\delta(\epsilon)>0$, we must have $K_{\delta(\epsilon)}=\phi$. That is to say that for every $\epsilon>0$ we have that there exists a $\delta(\epsilon)>0$ such that $\sup_{\omega\in\bar{A}}m^T(\omega,\delta(\epsilon))<\epsilon$. And hence we have shown (*A.2.5*).

We now assume (*A.2.5*) and (*A.2.4*) and prove the compactness of $\bar{A}$. Since $C[0,\infty)$ is a metric space, it suffices to prove that every sequence $\{\omega_n\}_{n=1}^{\infty}\subseteq A$ has a weakly convergent subsequence. We fix $T > 0$ and note that some $\delta_1 > 0$, we have $m^T(\omega, \delta_1) \leq 1$ by lemma ($A.2.4$) for each $\omega \in A$, so for a fixed integer $m \geq 1$ and $t \in (0, T]$ with $(m-1)\delta_1 < t \leq m\delta_1 \wedge T$, we have

$$|\omega(t)| = |\omega(0) + \sum_{k=1}^{m-1} |\omega(k\delta_1) - \omega((k-1)\delta_1)| + |\omega(t) - \omega((m-1)\delta_1)|$$

$$\leq |\omega(0)| + \sum_{k=1}^{m-1} 1 + 1 = |\omega(0)| + m$$

where we have used the fact that

$$|\omega(k\delta_1) - \omega((k-1)\delta_1)| \leq m^T(\omega, \delta_1)| \leq 1 \quad \text{and} \quad |\omega(t) - \omega((m-1)\delta_1) \leq m^T(\omega, \delta_1) \leq 1$$

And hence by equation ($A.2.4$) it follows that

$$\sup_{\omega \in A} \omega(t) \leq \sup_{\omega \in A} \omega(0) + m.$$

So if we consider a sequence $\{\omega_n\}_{n=1}^{\infty} \subseteq A$, then for each $r \in \mathbb{Q}^+$, the set of non-negative rationals, $\{\omega_n(r)\}_{n=1}^{\infty}$ is bounded above by $\sup_{\omega \in A} \omega(0) + m$ as shown above. Now let $\{r_0, r_1, r_2, \ldots, \}$ be an enumeration of $\mathbb{Q}^+$. The Bolzano Weierstrass theorem implies that we can choose a subsequence $\{\omega_n^{(0)}\}_{n=1}^{\infty}$ of $\{\omega_n\}_{n=1}^{\infty}$ with $\omega_n^{(0)}(r_0)$ converging to a limit $\omega(r_0)$. Now again from $\{\omega_n^{(0)}\}_{n=1}^{\infty}$, we can choose a further subsequence $\{\omega_n^{(1)}\}_{n=1}^{\infty}$ such that $\omega_n^{(1)}(r_1)$ converging to a limit $\omega(r_1)$. We can continue this process ad-infinitum in a diagonal way i.e by letting $\{\tilde{\omega}_n\}_{n=1}^{\infty} = \{\omega_n^{(n)}\}_{n=1}^{\infty}$ to be the diagonal sequence. We have that $\tilde{\omega}_n(r) \to \omega(r)$ for each $r \in \mathbb{Q}^+$. It follows from equation ($A.2.5$) that for every $n \geq 1$ given an $\epsilon > 0$, $\exists$ a $\delta(\epsilon) > 0$ such that $|\tilde{\omega}_n(s) - \tilde{\omega}_n(t)| \leq \epsilon$ whenever $0 \leq s, t \leq T$ and $|s - t| \leq \delta(\epsilon)$. Indeed this a a consequence of the fact that the limit as $\delta \downarrow 0$ of $\sup_{\omega \in A} m^T(\omega, \delta)$ is zero, $\{\tilde{\omega}_n\}_{n=1}^{\infty} \subseteq A$ and the definition of modulus of continuity. The same inequality therefore holds for ω when we impose the additional condition that $s, t \in \mathbb{Q}^+$ since $\tilde{\omega}_n(k) \to \omega(k)$, for all $k \in \mathbb{Q}^+$ and we can take the limits as n goes to ∞ in $|\tilde{\omega}_n(s) - \tilde{\omega}_n(t)| \leq \epsilon$ to get $|\omega(s) - \omega(t)| \leq \epsilon$. It follows that ω is continuous and hence uniformly continuous on $[0, T] \cap \mathbb{Q}^+$ so has an extension to a continuous function called ω' on $[0, T]$ and furthermore $|\omega(s) - \omega(t)| \leq \epsilon$whenever $0 \leq s, t \leq T$ and $|s - t| \leq \delta(\epsilon)$. Indeed for $t \in [0, T] \cap \mathbb{Q}^+$ set $\omega(t) = \omega'(t)$ and for $t \in [0, T] \cap (\mathbb{Q}^+)^c$ there exists sequence $\{t_n\}_{n=1}^{\infty}$ such that $t_n \to t$ as $n \to \infty$. Uniform continuity implies that $\{\omega(t_n)\}_{n=1}^{\infty}$ is a cauchy sequence since whenever $|t_n - t_m| < \delta(\epsilon)$(this is always possible for a large enough $n, m \in \mathbb{N}$ since $t_n \to t$) then $|\omega(t_n) - \omega(t_m)| < \epsilon$

□

Theorem A.2.6. *(Theorem 2.4.10 Shreve) A sequence $\{P_n\}_{n=1}^{\infty}$ of probability measures on $(C[0,\infty), \mathcal{B}(C[0,\infty))$ is tight if and only if*

$$\lim_{\lambda\uparrow\infty} \sup_{n\geq 1} P_n[\omega; |\omega(0)| > \lambda] = 0 \tag{A.2.6}$$

$$\lim_{\delta\downarrow 0} \sup_{n\geq 1} P_n[\omega; m^T(\omega,\delta) > \epsilon] = 0; \text{ for all } T > 0, \epsilon > 0 \tag{A.2.7}$$

Proof. See Theorem 2.4.10 on page 63 in [12]

□

Lemma A.2.7. *Suppose $\{P_n\}_{n=1}^{\infty}$ is a sequence of probability measures on $(C[0,\infty), \mathcal{B}(C[0,\infty)))$ which converges weakly to a probability measure P. Suppose, in addition that $\{f_n\}_{n=1}^{\infty}$ is a uniformly bounded sequences of real-valued, continuous functions on $C[0,\infty)$ converging to a continuous function f, the convergence being uniform on compacts of $C[0,\infty]$. Then*

$$\lim_{n\to\infty} \int_{C[0,\infty)} f_n(\omega) dP_n(\omega) = \int_{C[0,\infty)} f(\omega) dP(\omega) \tag{A.2.8}$$

Proof. First note that since the sequence $\{f_n\}_{n=1}^{\infty}$ is uniformly bounded means that $\sup_{\omega\in C[0,\infty)} \sup_{n\geq 1} |f_n(\omega)| < K$ for some $K > 0$. We have that

$$\begin{aligned}\lim_{n\to\infty} \left| \int f_n dP_n - \int f dP \right| &= \lim_{n\to\infty} \left| \int (f_n - f) dP_n + \int f dP_n - \int f dP \right| \\ &\leq \lim_{n\to\infty} \left| \int (f_n - f) dP_n \right| + \lim_{n\to\infty} \left| \int f dP_n - \int f dP \right|\end{aligned}$$

Since the limits of uniformly bounded functions is bounded. Indeed $f_n(\omega) \to f(\omega)$ means that $\forall \omega \in C[0,\infty)$ we have that $|f_n(\omega) - f(\omega)| < 1$ and since since the sequence $\{f_n\}_{n=1}^{\infty}$ is uniformly bounded by K we have that $|f| < K + 1$. We already know f is continuous, and we just showed that its bounded and since by assumption $P_n \xrightarrow{\text{weak}} P$ we have that $\lim_{n\to\infty} |\int f dP_n - \int f dP| = 0$

Now since $\{P_n\}_{n=1}^{\infty}$ weakly convergent and hence compact and therefore tight by the Theorem A.2.2 i.e for every $\epsilon > 0$ there exists a compact set $K \subseteq C[0,\infty)$ such that for all $n \geq 1$ we have $P_n(K) \geq 1 - \epsilon \implies -P_n(K) \leq \epsilon - 1 \implies 1 - P_n(K) \leq \epsilon \implies P_n(K^c) \leq \epsilon$.

We have that

$$\begin{aligned}\lim_{n\to\infty} \left| \int_{C[0,\infty)} (f_n - f) dP_n \right| &\leq \lim_{n\to\infty} \int_{C[0,\infty)} |f_n - f| dP_n \\ &= \lim_{n\to\infty} \int_K |f_n - f| dP_n + \lim_{n\to\infty} \int_{K^c} |f_n - f| dP_n \\ \leq \lim_{n\to\infty} \sup_{\omega\in K} |f_n(\omega) - f(\omega)| P_n(K) &+ (2K+1) P_n(K^c) \leq (2K+1)\epsilon\end{aligned}$$

and since this is true for all $\epsilon > 0$, the result follows immediately.

□

A.3 Convergence of Finite-dimensional distributions

Suppose that X is a continuous process on some $(\Omega, \mathcal{F}, P)$. For each ω, the function $t \mapsto X_t(\omega)$ is a member of $C[0,\infty)$, which we denote by $X(\omega)$. Since $\mathcal{B}(C[0,\infty))$ is generated by 1-dimensional cylinder sets and $X_t(.)$ is $\mathcal{F}$ measurable for each t, the random variable $X : \Omega \to C[0,\infty)$ is $\mathcal{F}/\mathcal{B}(C[0,\infty))$-measurable. This is the consequence of Problem 2.4.2 in [12] and that in order to show that X is a $\mathcal{F}/\mathcal{B}(C[0,\infty))$-measurable random variable, it is sufficient to show(by Theorem 8.1 in Probability essentials by Jacod and Protter [45]) that X is a $\mathcal{F}/\mathcal{C}$-measurable random variable where $\mathcal{C}$ is the collection of finite-dimensional cylinder sets of the form

$$C = \{\omega \in C[0,\infty); (\omega(t_1), \dots, \omega(t_n)) \in A\}, \quad n \geq 1, A \in \mathcal{B}(\mathbb{R}^n)$$

where for all $i = 1, \dots, n, t_i \in [0,\infty)$.

Thus if $\{X_n\}_{n=1}^{\infty}$ is a sequence of continuous processes(with each $X^{(n)}$ defined on perhaps a different probability space $(\Omega_n, \mathcal{F}_n), P_n)$ we can ask whether $X^{(n)} \xrightarrow{\mathcal{D}} X$ in distribution in the sense of Definition A.1.2. We can also ask whether the finite-dimensional distributions of $\{X^{(n)}\}_{n=1}^{\infty}$ converge to those of X, i.e whether

$$(X_{t_1}^{(n)}, X_{t_2}^{(n)}, \dots, X_{t_d}^{(n)}) \xrightarrow{\mathcal{D}} (X_{t_1}, X_{t_2}, \dots, X_{t_d})$$

The latter question is considerably easy to answer than the former , since the convergence in distribution of finite-dimensional random vectors can be resolved by studying their characteristic functions.

For any finite subset $\{t_1, t_2, \dots, t_d\}$ of $[0,\infty)$, let us define the projection mapping $\pi_{t_1,\dots\dots,t_d} : C[0,\infty) \to \mathbb{R}^d$ as

$$\pi_{t_1,\dots\dots,t_d}(\omega) = (\omega(t_1), \dots, \omega(t_d))$$

If the function $f : \mathbb{R}^d \to \mathbb{R}$ is bounded and continuous, then the composite mapping $f \circ \pi_{t_1,\dots\dots,t_d} : C[0,\infty) \to \mathbb{R}$ enjoys the same properties; thus $X^{(n)} \xrightarrow{\mathcal{D}} X$ as $n \to \infty$ implies

$$\begin{aligned}\lim_{n\to\infty} E_n(f(X_{t_1}^{(n)}, \dots, X_{t_d}^{(n)})) &= \lim_{n\to\infty} E_n(f \circ \pi_{t_1,\dots\dots,t_d})(X^{(n)}) \\ &= E(f \circ \pi_{t_1,\dots\dots,t_d})(X) = Ef(X_{t_1}, \dots, X_{t_d})\end{aligned}$$

where the second equality is a consequence of the definition of weak-convergence. In other words this tells us that if the sequence of processes $\{X^{(n)}\}_{n=1}^{\infty}$ converges in distribution to the process X, then all the finite dimensional distributions converge as well.

The converse holds in the presence of tightness but not in general and this failure is illustrated by the following example.

Theorem A.3.1. *(Theorem2.4.15 Shreve) Let* $\{X^{(n)}\}_{n=1}^{\infty}$ *be a tight sequence of continuous processes with the property that, whenever* $0 \le t_1 < \cdots < t_d < \infty$*, then the sequence of random vectors* $\{(X_{t_1}^{(n)}, \ldots, X_{t_d}^{(n)})\}_{n=1}^{\infty}$ *converges in distribution. Let* P_n *be the measure induced on* $(C[0,\infty), \mathcal{B}(C[0,\infty)))$ *by* $X^{(n)}$*. Then* $\{P_n\}_{n=1}^{\infty}$ *converges weakly to a measure* P*, under which the coordinate mapping process* $W_t(\omega) := \omega(t)$ *on* $C[0.\infty)$ *satisfies*

$$(X_{t_1}^{(n)}, \ldots, X_{t_d}^{(n)}) \xrightarrow{\mathcal{D}} (W_{t_1}. \ldots, W_{t_d}),\ 0 \le t_1 < \cdots < t_d < \infty, d \ge 1$$

Proof. Every subsequence $\{\tilde{X}^{(n)}\}$ of $\{X^{(n)}\}$ is tight (by the definition of tightness of a sequence of random variables) and so has a further subsequence $\{\hat{X}^{(n)}\}$ such that the measures induced on $C[0,\infty)$ by $\{\hat{X}^{(n)}\}$ converge weakly to a probability measure P, by Prohorov's theorem A.2.2. Indeed since the measures $\{P_n\}$ is tight and hence has a weakly convergent subsequence $\{\tilde{P}_n\}$(say converges to the probability measure P) and hence the so does the further subsequence $P(\hat{X}^{(n)})^{-1}$ of probability measures corresponding to the sequence $\{X^{(n)}\}$ of random variables. Similarly if a different subsequence $\{\bar{X}^{(n)}\}$ of $\{X^{(n)}\}$ induces a measure on $C[0,\infty)$ converging to a probability measure Q, then P and Q must have the same finite-dimensional distribution. Indeed since $(X_{t_1}^{(n)}, \ldots, X_{t_d}^{(n)}) \xrightarrow{\mathcal{D}} (W_{t_1}. \ldots, W_{t_d})$, any subsequence $(\tilde{X}_{t_1}^{(n)}, \ldots, \tilde{X}_{t_d}^{(n)})$ will also converge in distribution to the same random vector . In other words any subsequence of $(X_{t_1}^{(n)}, \ldots, X_{t_d}^{(n)})$ induces a sequence of measures on $\mathcal{B}(\mathbb{R}^d)$ which converge to the same probability measure induced by the random vector on $(X_{t_1}^{(n)}, \ldots, X_{t_d}^{(n)})$ on $\mathcal{B}(\mathbb{R}^d)$.

$$P[\omega \in C[0,\infty); (\omega(t_1), \ldots, \omega(t_d)) \in A] = Q[\omega \in C[0,\infty); (\omega(t_1), \ldots, \omega(t_d)) \in A]$$

whenever $0 \le t_1 < t_2 < \cdots < t_d < \infty, A \in \mathcal{B}(\mathbb{R}^d), d \ge 1$

This means $P = Q$, since the probability measure induced by a continuous process is determined by its finite-dimensional distributions.

Now suppose the sequence of measures $\{P_n\}_{n=1}^{\infty}$ induced by $\{X^{(n)}\}_{n=1}^{\infty}$ did not converge weakly to P. Then there must be a bounded and continuous function $f : C[0,\infty) \to \mathbb{R}$ such that $\lim_{n\to\infty} \int f(\omega)dP_n(\omega)$ does not exist, or else this limit exists but is different from $\int f(\omega)dP(\omega)$ In either case we an choose a subsequence(since $\{P_n\}_{n=1}^{\infty}$ is tight and hence relatively compact) $\{\tilde{P}_n\}_{n=1}^{\infty}$ for which $\lim_{n\to\infty} \int f(\omega)d\tilde{P}_n(\omega)$ exists but is different from $\int f(\omega)dP(\omega)$. This is true because if $\int f(\omega)dP_n(\omega)$ does not converge to $\int f(\omega)dP(\omega)$, its subsequence $\int f(\omega)d\tilde{P}_n(\omega)$ must converge to the same value and hence cannot converge to $\int f(\omega)dP(\omega)$. This subsequence cannot cannot have a further subsequence $\{\hat{P}_n\}_{n=1}^{\infty}$ such that $\hat{P} \xrightarrow{\text{W}} P$, and this contradicts the conclusion of the previous paragraph which says that every sub subsequence converges weakly to the probability measure P.

□

.

Lemma A.3.2. *(Problem 2.4.16 Shreve) Let $\{X^{(n)}\}_{n=1}^{\infty}$, $\{Y^{(n)}\}_{n=1}^{\infty}$ and X be random variables with values in a separable metric space (S,ρ);we assume that for each $n \geq 1, X^{(n)}$ and $Y^{(n)}$ are defined on the same probability space. If $X^{(n)} \xrightarrow{\mathcal{D}} X$ and $\rho(X^{(n)}, Y^{(n)}) \to 0$ in probability, as $n \to \infty$, then $Y^{(n)} \xrightarrow{\mathcal{D}} X$ as $n \to \infty$*

Proof. Let $(\Omega_n, \mathcal{F}_n, P_n)$ denote the space on which X_n and Y_n are defined, and let E_n denote expectation with respect to P_n. Let X be defined on $(\Omega, \mathcal{F}, P)$. We are given that $\lim_{n\to\infty} E_n(f(X^{(n)})) = E(f(X))$ for every bounded continuous function $f : S \to \mathbb{R}$ and that $\lim_{n\to\infty} P_n(\rho(X^{(n)}, Y^{(n)}) \geq \epsilon) = 0$ for all $\epsilon > 0$.To prove $Y^{(n)} \xrightarrow{\mathcal{D}} X$, it suffices to show that

$$\lim_{n\to\infty} E[f(X^{(n)}) - f(Y^{(n)})] = 0$$

whenever f is bounded and continuous as $\lim_{n\to\infty} E_n(f(X^{(n)})) = E(f(X))$ and $Y^{(n)} \xrightarrow{\mathcal{D}} X$ is the same as showing $\lim_{n\to\infty} E_n(f(Y^{(n)})) = E(f(X))$ Let such an f be given and set $M = \sup_{x\in S}|f(x)| < \infty$. Since X_n converges to X in distribution and therefore its relatively compact(since the induced probabillity measures $P_n X_n^{-1}$ converge and hence have a convergent subsequence). By the Prohorov's theorem its tight and hence for all $\epsilon > 0$ there exists a compact set $K \subseteq S$ such that $P_n(X^{(n)} \in K) \geq 1 - \epsilon/6M$ for all $n \geq 1$. Since f is continuous and hence uniformly continuous on the compact set K we have that there exists a δ such that $0 < \delta < 1$ so that $|f(x) - f(y)| < \epsilon/3$ whenever $\rho(x, y) < \delta$ and $x \in K$. We also choose a positive integer N large enough so that $P_n(\rho(X^{(n)}, Y^{(n)}) \geq \delta) \leq \epsilon/6M$ for all $n \geq N$. We can do this since $\rho(X^{(n)}, Y^{(n)})$ converge in probability to 0.

Putting all the above together we have that

$$\left|\int_{\Omega_n} \left[f(X^{(n)}) - f(Y^{(n)})\right] dP_n\right| = \left|\int_{\Omega_n\cap\{X^{(n)}\in K,\rho(X^{(n)},Y^{(n)})<\delta\}} \left[f(X^{(n)}) - f(Y^{(n)})\right] dP_n\right|$$

$$+\left|\int_{\Omega_n\cap\{X^{(n)}\in K,\rho(X^{(n)},Y^{(n)})<\delta\}^c} \left[f(X^{(n)}) - f(Y^{(n)})\right] dP_n\right| \leq \epsilon/3 \int_{\Omega_n\cap\{X^{(n)}\in K,\rho(X^{(n)},Y^{(n)})<\delta\}} dP_n$$

$$+2M \int_{\Omega_n\cap\{X^{(n)}\in K,\rho(X^{(n)},Y^{(n)})<\delta\}^c} dP_n = \epsilon/3 P_n(X^{(n)} \in K, \rho(X^{(n)}, Y^{(n)}) < \delta)$$

$$+2MP_n(\{X^{(n)} \in K, \rho(X^{(n)}, Y^{(n)}) < \delta\}^c)$$

$$\leq \epsilon/3 + 2MP(X^{(n)} \in K^c) + 2MP(\rho(X^{(n)}, Y^{(n)}) \geq \delta) \leq \epsilon$$

And since this is true for any epsilon the proof is complete.

□

A.4 The invariance principal and the Wiener measure

Let us consider a sequence $\{\xi_j\}_{j=1}^{\infty}$ of independent identically distributed random variables with mean 0 and variance $\sigma^2, 0 < \sigma^2 < \infty$, as well as the sequence of partial sums $S_0 = 0, S_k = \sum_{j=1}^{k} \xi_j, k \geq 1$. A continuous time process $Y = \{Y_t, t \geq 0\}$ can be obtained from the sequence $\{S_k\}_{k=0}^{\infty}$ by linear interpolation; i.e,

$$Y_t = S_{\lfloor t \rfloor} + (t - \lfloor t \rfloor)\xi_{\lfloor t \rfloor + 1}, \ t \geq 0 \tag{A.4.1}$$

where $\lfloor t \rfloor$ denotes the greatest integer less or equal to t. Scaling appropriately both time and space , we obtain from Y a sequence of processes $\{X^{(n)}\}$:

$$X_t^{(n)} = \frac{1}{\sigma\sqrt{n}} Y_{nt}, \ t \geq 0 \tag{A.4.2}$$

Note that with $s = k/n$ and $t = (k+1)/n$, the increment $X_t^{(n)} - X_s^{(n)} = \frac{Y_{nt} - Y_{st}}{\sigma\sqrt{n}} = \frac{Y_{k+1} - Y_k}{\sigma\sqrt{n}} = \frac{S_{k+1} - S_k}{\sigma\sqrt{n}} = \frac{\xi_{k+1}}{\sigma\sqrt{n}}$, which is obviously independent of $\mathcal{F}_s^{X^{(n)}} = \sigma(\xi_1, \ldots, \xi_k)$. Furthermore, $X_t^{(n)} - X_s^{(n)}$ has zero mean and variance $t - s$. This suggests that $\{X^{(n)}; t \geq 0\}$ is approximately a Brownian motion. We now show that , even though , the random variables ξ_j are not necessarily normal , the central limit theorem dictates that the limiting distribution of the increments of $X^{(n)}$ are normal

Theorem A.4.1. *With $\{X^{(n)}\}$ defined by (A.4.2) and $0 \leq t_1 < \cdots < t_d < \infty$ we have*

$$(X_{t_1}^{(n)}, \ldots, X_{t_d}^{(n)}) \xrightarrow{\mathcal{D}} (B_{t_1}, \ldots, B_{t_d}) \ \textit{as } n \to \infty$$

where $\{B_t, \mathcal{F}_t^B \ t \geq 0\}$ is a standard one dimensional Brownian motion

Proof. We take the case $d = 2$; the other cases differ from this one only by being notationally more cumbersome. Set $s = t_1, t = t_2$. We wish to show that

$$(X_s^{(n)}, X_t^{(n)}) \xrightarrow{\mathcal{D}} (B_s, B_t)$$

We have that

$$\left| X_t^{(n)} - \frac{1}{\sigma\sqrt{n}} S_{\lfloor tn \rfloor} \right| = \left| \frac{1}{\sigma\sqrt{n}} Y_{nt} - \frac{1}{\sigma\sqrt{n}} S_{\lfloor tn \rfloor} \right| = \left| \frac{1}{\sigma\sqrt{n}} (nt - \lfloor nt \rfloor) \xi_{\lfloor nt \rfloor + 1} \right| \leq \left| \frac{1}{\sigma\sqrt{n}} \xi_{\lfloor nt \rfloor + 1} \right|$$

And hence Chebyshev's inequality implies

$$P\left[|X_t^{(n)}-\frac{1}{\sigma\sqrt{n}}S_{\lfloor tn\rfloor}|>\epsilon\right]\leq\frac{1}{\epsilon^2}E\left[|X_t^{(n)}-\frac{1}{\sigma\sqrt{n}}S_{\lfloor tn\rfloor}|^2\right]\leq\frac{1}{\sigma^2n\epsilon^2}E\left[|\xi_{\lfloor tn\rfloor+1}|^2\right]=\frac{1}{n\epsilon^2}$$

Now taking the limits we get

$$\lim_{n\to\infty}P\left[|X_t^{(n)}-\frac{1}{\sigma\sqrt{n}}S_{\lfloor tn\rfloor}|>\epsilon\right]=0$$

It is not so hard to see that $\left\|(X_s^{(n)},X_t^{(n)})-\frac{1}{\sigma\sqrt{n}}(S_{\lfloor sn\rfloor},S_{\lfloor tn\rfloor})\right\|\to 0$ in probability. Indeed

$$\begin{aligned}
&P\left[\left\|(X_s^{(n)},X_t^{(n)})-\frac{1}{\sigma\sqrt{n}}(S_{\lfloor sn\rfloor},S_{\lfloor tn\rfloor})\right\|>\epsilon\right]\\
&=P\left[\sqrt{\left(X_s^{(n)}-\frac{1}{\sigma\sqrt{n}}S_{\lfloor sn\rfloor}\right)^2+\left(X_t^{(n)}-\frac{1}{\sigma\sqrt{n}}S_{\lfloor tn\rfloor}\right)^2}>\epsilon\right]\\
&=P\left[\left(X_s^{(n)}-\frac{1}{\sigma\sqrt{n}}S_{\lfloor sn\rfloor}\right)^2+\left(X_t^{(n)}-\frac{1}{\sigma\sqrt{n}}S_{\lfloor tn\rfloor}\right)^2>\epsilon^2\right]\\
&\leq P\left[\left(X_s^{(n)}-\frac{1}{\sigma\sqrt{n}}S_{\lfloor sn\rfloor}\right)^2>\epsilon^2/2\right]+P\left[\left(X_t^{(n)}-\frac{1}{\sigma\sqrt{n}}S_{\lfloor tn\rfloor}\right)^2>\epsilon^2/2\right]\\
&=P\left[|X_s^{(n)}-\frac{1}{\sigma\sqrt{n}}S_{\lfloor sn\rfloor}|>\epsilon/\sqrt{2}\right]+P\left[|X_t^{(n)}-\frac{1}{\sigma\sqrt{n}}S_{\lfloor tn\rfloor}|>\epsilon/\sqrt{2}\right]
\end{aligned}$$

And now taking the limits we get the desired result. Now Lemma A.3.2 implies that in order to prove

$$(X_{t_1}^{(n)},X_{t_2}^{(n)})\xrightarrow{\mathcal{D}}(B_{t_1},B_{t_2})$$

it is sufficient to show that

$$\frac{1}{\sigma\sqrt{n}}\left(\sum_{j=1}^{\lfloor sn\rfloor}\xi_j,\sum_{j=\lfloor sn\rfloor+1}^{\lfloor tn\rfloor}\xi_j\right)\xrightarrow{\mathcal{D}}(B_s,B_t-B_s)$$

In order to show this we use the Levy-continuity theorem(which says that $X_n\xrightarrow{\mathcal{D}}X$ iff $\varphi_{X_n}(t)\to\varphi_X(t)$ for all $t\in\mathbb{R}^n$ if X_n,X are n-dimensional random vectors. In other words we need to show that

$$\lim_{n\to\infty}E\left[\exp\left(i\left\langle u,\frac{1}{\sigma\sqrt{n}}\left(\sum_{j=1}^{\lfloor sn\rfloor}\xi_j,\sum_{j=\lfloor sn\rfloor+1}^{\lfloor tn\rfloor}\xi_j\right)\right\rangle\right)\right]$$
$$=E\left[\exp\left(i\left\langle u,(B_s,B_t-B_s)\right\rangle\right)\right]$$

Expanding the inner product we need to show

$$\lim_{n\to\infty} E\left[\exp\left(iu_1\frac{1}{\sigma\sqrt{n}}\sum_{j=1}^{\lfloor sn\rfloor}\xi_j\right)\exp\left(iu_2\frac{1}{\sigma\sqrt{n}}\sum_{j=\lfloor sn\rfloor+1}^{\lfloor tn\rfloor}\xi_j\right)\right] = E\left[\exp\left(iu_1B_s\right)\exp\left(iu_2(B_t-B_s)\right)\right] \tag{A.4.3}$$

Now using the independence of B_s, B_t-B_s and $\sum_{j=1}^{\lfloor sn\rfloor}\xi_j, \sum_{j=\lfloor sn\rfloor+1}^{\lfloor tn\rfloor}\xi_j$, it suffices to show the following

$$\lim_{n\to\infty} E\left[\exp\left(iu_1\frac{1}{\sigma\sqrt{n}}\sum_{j=1}^{\lfloor sn\rfloor}\xi_j\right)\right]E\left[\exp\left(iu_2\frac{1}{\sigma\sqrt{n}}\sum_{j=\lfloor sn\rfloor+1}^{\lfloor tn\rfloor}\xi_j\right)\right] = E\left[\exp\left(iu_1B_s\right)\right]E\left[\exp\left(iu_2(B_t-B_s)\right)\right]$$

We compute the following limit

$$\lim_{n\to\infty} E\left[\exp\left(iu_1\frac{1}{\sigma\sqrt{n}}\sum_{j=1}^{\lfloor sn\rfloor}\xi_j\right)\right]$$

and

$$E\left[\exp\left(iu_2\frac{1}{\sigma\sqrt{n}}\sum_{j=\lfloor sn\rfloor+1}^{\lfloor tn\rfloor}\xi_j\right)\right]$$

can be computed similarly. It is not hard to see that

$$\lim_{n\to\infty} P\left[\left|\frac{1}{\sigma\sqrt{n}}\sum_{j=1}^{\lfloor sn\rfloor}\xi_j - \frac{\sqrt{s}}{\sigma\sqrt{\lfloor sn\rfloor}}\sum_{j=1}^{\lfloor sn\rfloor}\xi_j\right| > \epsilon\right] \le \lim_{n\to\infty}\frac{1}{\epsilon^2}E\left[\left|\frac{1}{\sigma\sqrt{n}}\sum_{j=1}^{\lfloor sn\rfloor}\xi_j - \frac{\sqrt{s}}{\sigma\sqrt{\lfloor sn\rfloor}}\sum_{j=1}^{\lfloor sn\rfloor}\xi_j\right|^2\right]$$

$$= \lim_{n\to\infty}\frac{1}{\epsilon^2}E\left[\left|\frac{1}{\sigma\sqrt{n}} - \frac{\sqrt{s}}{\sigma\sqrt{\lfloor sn\rfloor}}\right|^2\left|\sum_{j=1}^{\lfloor sn\rfloor}\xi_j\right|^2\right] = \lim_{n\to\infty}\frac{1}{\epsilon^2}\left|\frac{1}{\sigma\sqrt{n}} - \frac{\sqrt{s}}{\sigma\sqrt{\lfloor sn\rfloor}}\right|^2 E\left[\left|\sum_{j=1}^{\lfloor sn\rfloor}\xi_j\right|^2\right]$$

$$= \lim_{n\to\infty}\frac{1}{\epsilon^2}\left|\frac{\sqrt{\lfloor sn\rfloor}-\sqrt{sn}}{\sigma\sqrt{n}\sqrt{\lfloor sn\rfloor}}\right|^2 E\left[\left|\sum_{j=1}^{\lfloor sn\rfloor}\xi_j\right|^2\right] \le \lim_{n\to\infty}\frac{1}{\epsilon^2\sigma^2 n\lfloor sn\rfloor}E\left[\sum_{j=1}^{\lfloor sn\rfloor}\xi_j^2\right] = \lim_{n\to\infty}\frac{\sigma^2\lfloor sn\rfloor}{\epsilon^2\sigma^2 n\lfloor sn\rfloor} = 0$$

Now using Lemma A.3.2 with $X^{(n)} = \frac{\sqrt{s}}{\sigma\sqrt{\lfloor sn\rfloor}}\sum_{j=1}^{\lfloor sn\rfloor}\xi_j, Y^{(n)} = \frac{1}{\sigma\sqrt{n}}\sum_{j=1}^{\lfloor sn\rfloor}\xi_j$ and $X = \mathcal{N}(0,s)$. Now the Lindeberg-Levy version of the Central Limit Theorem implies $X^{(n)} \xrightarrow{\mathcal{D}} X$ and since we have already shown that $\rho(X^{(n)}, Y^{(n)}) \to 0$ in probability with $\rho(x,y) = |x-y|$ i.e the euclidean distance on $\mathbb{R}^1$ we get that $Y^{(n)} \xrightarrow{\mathcal{D}} X$ or in other words $\frac{1}{\sigma\sqrt{n}}\sum_{j=1}^{\lfloor sn\rfloor}\xi_j \xrightarrow{\mathcal{D}} \mathcal{N}(0,s)$. Therefore the Levy Cramer continuity theorem implies

$$\lim_{n\to\infty} E\left[\exp\left(iu_1\frac{1}{\sigma\sqrt{n}}\sum_{j=1}^{\lfloor sn\rfloor}\xi_j\right)\right] = E\left[\exp\left(iu_1\mathcal{N}(0,s)\right)\right] = \exp\left(-u_1^2 s/2\right)$$

Using a similar argument we can show that

$$\lim_{n\to\infty} E\left[\exp\left(iu_2\frac{1}{\sigma\sqrt{n}}\sum_{j=\lfloor sn\rfloor+1}^{\lfloor tn\rfloor}\xi_j\right)\right] = \exp\left(iu^2(t-s)/2\right)$$

and hence we have shown A.4.3 since using the properties of Brownian motion we have $E\left[\exp\left(iu_1B_s\right)\right] = \exp\left(-u_1^2s/2\right)$ and $E\left[\exp\left(iu_2(B_t-B_s)\right)\right] = \exp\left(iu^2(t-s)/2\right)$ □

Actually the sequence $\{X^{(n)}\}$ of linearly interpolated and normalized random walks given by equation (A.4.2) converges to Brownian Motion in distribution. For the tightness required to carry out such and extension we shall need the following to auxiliary results.

Lemma A.4.2. *(Lemma 2.4.18 Shreve) Set $S_k = \sum_{j=1}^k \xi_j$ where $\{\xi_j\}_{j=1}^\infty$ is a sequence of independent and identically distributed random variables , with mean zero and finite variance $\sigma^2 > 0$. Then for any $\epsilon > 0$,*

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\frac{1}{\delta}P\left[\max_{1\le j\le\lfloor n\delta\rfloor+1}|S_j|>\epsilon\sigma\sqrt{n}\right] = 0$$

Proof. By the central limit theorem, we have for each $\delta > 0$ that $\frac{1}{\sigma\sqrt{1+\lfloor n\delta\rfloor}}S_{\lfloor n\delta\rfloor+1}$ converges in distribution to a standard normal random variable Z. Using the exact same argument as in Theorem A.4.1 we can show that

$$\left|\frac{1}{\sigma\sqrt{1+\lfloor n\delta\rfloor}}S_{\lfloor n\delta\rfloor+1} - \frac{1}{\sigma\sqrt{n\delta}}S_{\lfloor n\delta\rfloor+1}\right| \to 0 \text{ in probability}$$

and therefore just as in Theorem A.4.1 we can apply Lemma A.3.2 to conclude that

$$\frac{1}{\sigma\sqrt{n\delta}}S_{\lfloor n\delta\rfloor+1} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$$

Now fix $\lambda > 0$ and let $\{\varphi_k\}_{k=1}^\infty$ be a sequence of bounded continuous functions on $\mathbb{R}$ with $\varphi_k \downarrow 1_{\{(-\infty,\lambda]\cup[\lambda,\infty)\}}$ We have for each k

$$\begin{aligned}\limsup_{n\to\infty} P\left[|S_{\lfloor n\delta\rfloor+1}|\ge\lambda\sigma\sqrt{n\delta}\right] &\le \lim_{n\to\infty} E\left[1_{\left\{\left|\frac{S_{\lfloor n\delta\rfloor+1}}{\sigma\sqrt{n\delta}}\right|\ge\lambda\right\}}\right]\\ &\le \lim_{n\to\infty} E\left[\varphi_k\left(\frac{S_{\lfloor n\delta\rfloor+1}}{\sigma\sqrt{n\delta}}\right)\right] = E\left[\varphi_k(Z)\right]\end{aligned}$$

where Z is a standard normal random variable. The second inequality is a consequence of the fact that $\varphi_k \downarrow 1_{\{(-\infty,\lambda]\cup[\lambda,\infty)\}}$ and the linearity of the expectation operator. The last equality is a consequence the definition of convergence in distribution (of $|S_{\lfloor n\delta\rfloor+1}|\xrightarrow{\mathcal{D}} Z$)

in Definition A.1.2 and the boundedness and continuity of φ_k. Now letting $k \to \infty$ and using the monotone convergence theorem for decreasing positive sequences and $\varphi_k \downarrow 1_{\{(-\infty,\lambda]\cup[\lambda,\infty)\}}$ and the Chebyshev's inequality we get

$$\begin{aligned}\limsup_{n\to\infty} P\left[|S_{\lfloor n\delta\rfloor+1}|\geq \lambda\sigma\sqrt{n\delta}\right] &\leq \lim_{k\to\infty} E\left[\varphi_k(Z)\right] \\ = E\left[\lim_{k\to\infty}\varphi_k(Z)\right] &= P\left[|Z|\geq \lambda\right] \leq \frac{1}{\lambda^3}E\left[|Z|^3\right]\end{aligned} \tag{A.4.4}$$

Now we define $\tau = \min\{j \geq 1; |S_j|> \epsilon\sigma\sqrt{n}\}$ Now with $0 < \delta < \epsilon^2/2$, we have the following

$$\begin{aligned}\left\{\max_{0\leq j\leq\lfloor n\delta\rfloor+1}|S_j|> \epsilon\sigma\sqrt{n}\right\} \subseteq \left\{|S_{\lfloor n\delta\rfloor+1}|\geq \sigma\sqrt{n}(\epsilon-\sqrt{2\delta})\right\}\bigcup \\ \bigcup_{j=1}^{\lfloor n\delta\rfloor+1}\left\{|S_{\lfloor n\delta\rfloor+1}|< \sigma\sqrt{n}(\epsilon-\sqrt{2\delta}), \tau=j\right\}\end{aligned}$$

Note that for $j = \lfloor n\delta\rfloor + 1$, $\left\{|S_{\lfloor n\delta\rfloor+1}|< \sigma\sqrt{n}(\epsilon-\sqrt{2\delta}), \tau=j\right\}$ is empty so the union above is only until $j = \lfloor n\delta\rfloor$ which is reflected in the sequel Now taking probabilities of these sets and using sub-additivity we get

$$\begin{aligned}P\left[\left\{\max_{0\leq j\leq\lfloor n\delta\rfloor+1}|S_j|> \epsilon\sigma\sqrt{n}\right\}\right] &\leq P\left[\left\{|S_{\lfloor n\delta\rfloor+1}|\geq \sigma\sqrt{n}(\epsilon-\sqrt{2\delta})\right\}\right] \\ &+ \sum_{j=1}^{\lfloor n\delta\rfloor+1} P\left[\left\{|S_{\lfloor n\delta\rfloor+1}|< \sigma\sqrt{n}(\epsilon-\sqrt{2\delta}), \tau=j\right\}\right]\end{aligned} \tag{A.4.5}$$

or equivalently using conditional probabilities we can write the inequality above as

$$\begin{aligned}P\left[\left\{\max_{0\leq j\leq\lfloor n\delta\rfloor+1}|S_j|> \epsilon\sigma\sqrt{n}\right\}\right] &\leq P\left[\left\{|S_{\lfloor n\delta\rfloor+1}|\geq \sigma\sqrt{n}(\epsilon-\sqrt{2\delta})\right\}\right] \\ &+ \sum_{j=1}^{\lfloor n\delta\rfloor+1} P\left[\left\{|S_{\lfloor n\delta\rfloor+1}|< \sigma\sqrt{n}(\epsilon-\sqrt{2\delta})|\tau=j\right\}\right]P\left[\tau=j\right]\end{aligned}$$

But if $\tau = j$ then $|S_{\lfloor n\delta\rfloor+1}|< \sigma\sqrt{n}(\epsilon-\sqrt{2\delta})$ implies(by simple observation) that $|S_j - S_{\lfloor n\delta\rfloor+1}|> \sigma\sqrt{2n\delta}$ or more precisely

$$\left\{|S_{\lfloor n\delta\rfloor+1}|< \sigma\sqrt{n}(\epsilon-\sqrt{2\delta}), \tau=j\right\} \subseteq \left\{|S_j - S_{\lfloor n\delta\rfloor+1}|> \sigma\sqrt{2n\delta}, \tau=j\right\}$$

Now monotonicity and Chebychev's inequality and independence of $\{\xi_j\}_{j=1}^{\infty}$ implies the following

$$P\left[\left\{|S_{\lfloor n\delta\rfloor+1}|< \sigma\sqrt{n}(\epsilon-\sqrt{2\delta}),\tau=j\right\}\right] \leq P\left[\left\{|S_j - S_{\lfloor n\delta\rfloor+1}|> \sigma\sqrt{2n\delta},\tau=j\right\}\right]$$
$$\leq \frac{1}{2n\delta\sigma^2}E\left[\left(S_j - S_{\lfloor n\delta\rfloor+1}\right)^2|\tau=j\right] = \frac{1}{2n\delta\sigma^2}E\left[\sum_{i=j+1}^{\lfloor n\delta\rfloor+1}\xi_i^2\right]$$
$$= \frac{1}{2n\delta\sigma^2}\sigma^2(\lfloor n\delta\rfloor - j) \leq \frac{1}{2n\delta\sigma^2}\sigma^2(\lfloor n\delta\rfloor) \leq \frac{1}{2}$$

And hence we have the following (in the last inequality below we use the definition of τ and monotonicity)

$$P\left[\left\{\max_{0\leq j\leq\lfloor n\delta\rfloor+1}|S_j|> \epsilon\sigma\sqrt{n}\right\}\right]$$
$$\leq P\left[\left\{|S_{\lfloor n\delta\rfloor+1}|\geq \sigma\sqrt{n}(\epsilon-\sqrt{2\delta})\right\}\right] + \sum_{j=1}^{\lfloor n\delta\rfloor}\frac{1}{2}P\left[\tau=j\right]$$
$$= P\left[\left\{|S_{\lfloor n\delta\rfloor+1}|\geq \sigma\sqrt{n}(\epsilon-\sqrt{2\delta})\right\}\right] + \frac{1}{2}P\left[\tau\leq\lfloor n\delta\rfloor\right]$$
$$\leq P\left[\left\{|S_{\lfloor n\delta\rfloor+1}|\geq \sigma\sqrt{n}(\epsilon-\sqrt{2\delta})\right\}\right] + \frac{1}{2}P\left[\left\{\max_{0\leq j\leq\lfloor n\delta\rfloor+1}|S_j|> \epsilon\sigma\sqrt{n}\right\}\right]$$

from which follows

$$P\left[\left\{\max_{0\leq j\leq\lfloor n\delta\rfloor+1}|S_j|> \epsilon\sigma\sqrt{n}\right\}\right] \leq 2P\left[\left\{|S_{\lfloor n\delta\rfloor+1}|\geq \sigma\sqrt{n}(\epsilon-\sqrt{2\delta})\right\}\right] \tag{A.4.6}$$

Now setting $\lambda = (\epsilon-\sqrt{2\delta})/\sqrt{\delta}$ in equation (A.4.4) and apply Chebyshev's inequality we get

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\frac{1}{\delta}P\left[\left\{\max_{0\leq j\leq\lfloor n\delta\rfloor+1}|S_j|> \epsilon\sigma\sqrt{n}\right\}\right] \leq \lim_{\delta\downarrow 0}\frac{2\sqrt{\delta}}{(\epsilon-\sqrt{2\delta})^3}E[|Z|^3] = 0$$

□

Lemma A.4.3. *(Lemma2.4.19 Shreve) Under the assumption of Lemma A.4.2 we have that for $T>0$,*

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty}P\left[\max_{\substack{1\leq j\leq\lfloor n\delta\rfloor+1\\0\leq k\leq\lfloor nT\rfloor+1}}|S_{j+k}-S_k|> \epsilon\sigma\sqrt{n}\right] = 0$$

Proof. For $0<\delta\leq T$,let $m=m(\delta)\geq 2$ be the unique integer satisfying $T/m<\delta\leq T/(m-1)$. Since

$$\lim_{n\to\infty}\frac{\lfloor nT\rfloor+1}{\lfloor n\delta\rfloor+1} = \frac{T}{\delta} < m$$

we have that $\lfloor nT \rfloor + 1 \leq (\lfloor n\delta \rfloor + 1)m$ for sufficiently large n. For such a large n, suppose $|S_{j+k} - S_k| > \epsilon\sigma\sqrt{n}$ for some $0 \leq k \leq \lfloor nT \rfloor + 1$ and some $j, 1 \leq j \leq \lfloor n\delta \rfloor + 1$. Then there exists a unique integer $p, 0 \leq p \leq m-1$, such that

$$(\lfloor n\delta \rfloor + 1)\, p \leq k < (\lfloor n\delta \rfloor + 1)\,(p+1)$$

Since k is an integer between 0 and $(\lfloor n\delta \rfloor + 1)m$ and $[0, (\lfloor n\delta \rfloor + 1)m] = \bigsqcup_{p=0}^{m-1} [(\lfloor n\delta \rfloor + 1)p, (\lfloor n\delta \rfloor + 1)(p+1)]$ Now clearly for a k given a p such that $0 \leq p \leq m-1$

$$(\lfloor n\delta \rfloor + 1)p \leq k + j \leq (\lfloor n\delta \rfloor + 1)(p+2)$$

And hence it follows that there are two possibilities for $k+j$. The first being

$$(\lfloor n\delta \rfloor + 1)\, p \leq k + j \leq (\lfloor n\delta \rfloor + 1)\,(p+1)$$

in which case either $|S_k - S_{(\lfloor n\delta \rfloor + 1)p}| > \frac{1}{3}\epsilon\sigma\sqrt{n}$, or else $|S_{k+j} - S_{(\lfloor n\delta \rfloor + 1)p}| > \frac{1}{3}\epsilon\sigma\sqrt{n}$. Indeed if both of the quantities were less than $\frac{1}{3}\epsilon\sigma\sqrt{n}$ we would have a contradiction in the following sense

$$\begin{aligned} \epsilon\sigma\sqrt{n} < |S_{j+k} - S_k| &= |S_k - S_{(\lfloor n\delta \rfloor + 1)p} + S_{(\lfloor n\delta \rfloor + 1)p} - S_{k+j}| \\ &\leq |S_k - S_{(\lfloor n\delta \rfloor + 1)p}| + |S_{(\lfloor n\delta \rfloor + 1)p} - S_{k+j}| \leq \frac{2}{3}\epsilon\sigma\sqrt{n} \end{aligned}$$

The second possibility is that

$$(\lfloor n\delta \rfloor + 1)\,(p+1) \leq k + j \leq (\lfloor n\delta \rfloor + 1)\,(p+2)$$

in which case either $|S_k - S_{(\lfloor n\delta \rfloor + 1)p}| > \frac{1}{3}\epsilon\sigma\sqrt{n}, |S_{(\lfloor n\delta \rfloor + 1)p} - S_{(\lfloor n\delta \rfloor + 1)(p+1)}| > \frac{1}{3}\epsilon\sigma\sqrt{n}, |S_{(\lfloor n\delta \rfloor + 1)(p+1)} - S_{k+j}| > \frac{1}{3}\epsilon\sigma\sqrt{n}$ which can again be proved as by contradiction using the triangular inequality just as before. In conclusion we see that

$$\left\{ \max_{\substack{1 \leq j \leq \lfloor n\delta \rfloor + 1 \\ 0 \leq k \leq \lfloor nT \rfloor + 1}} |S_{j+k} - S_k| > \epsilon\sigma\sqrt{n} \right\} \subseteq \bigcup_{p=0}^{m} \left\{ \max_{1 \leq j \leq \lfloor n\delta \rfloor + 1} |S_{j+p(\lfloor n\delta \rfloor + 1)} - S_{(\lfloor n\delta \rfloor + 1)p}| > \frac{1}{3}\epsilon\sigma\sqrt{n} \right\} \tag{A.4.7}$$

The set inequality above is seen to be true both the cases classified above i.e when $(\lfloor n\delta \rfloor + 1)\, p \leq k + j \leq (\lfloor n\delta \rfloor + 1)\,(p+1)$ or when $(\lfloor n\delta \rfloor + 1)\,(p+1) \leq k + j \leq (\lfloor n\delta \rfloor + 1)\,(p+2)$.

But independence of $\{\xi_j\}_{j=1}^{\infty}$ implies that

$$P\left[\max_{1 \leq j \leq \lfloor n\delta \rfloor + 1} |S_{j+p(\lfloor n\delta \rfloor + 1)} - S_{(\lfloor n\delta \rfloor + 1)p}| > \frac{1}{3}\epsilon\sigma\sqrt{n} \right] = P\left[\max_{1 \leq j \leq \lfloor n\delta \rfloor + 1} |S_j| > \frac{1}{3}\epsilon\sigma\sqrt{n} \right]$$

and since $m \leq (T/\delta)+1$ thus we have

$$P\left[\left\{\max_{\substack{1\leq j\leq \lfloor n\delta\rfloor+1\\ 0\leq k\leq \lfloor nT\rfloor+1}} |S_{j+k}-S_k| > \epsilon\sigma\sqrt{n}\right\}\right] \leq (m+1)P\left[\max_{1\leq j\leq \lfloor n\delta\rfloor+1} |S_j| > \frac{1}{3}\epsilon\sigma\sqrt{n}\right]$$

$$\leq ((T/\delta)+2)\, P\left[\max_{1\leq j\leq \lfloor n\delta\rfloor+1} |S_j| > \frac{1}{3}\epsilon\sigma\sqrt{n}\right]$$

Now on taking the limits and using the result of Lemma A.4.2 and equation A.4.6 we have

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty} P\left[\max_{\substack{1\leq j\leq \lfloor n\delta\rfloor+1\\ 0\leq k\leq \lfloor nT\rfloor+1}} |S_{j+k}-S_k| > \epsilon\sigma\sqrt{n}\right]$$
$$\leq T\lim_{\delta\downarrow 0}\limsup_{n\to\infty} \frac{1}{\delta} P\left[\max_{1\leq j\leq \lfloor n\delta\rfloor+1} |S_j| > \frac{1}{3}\epsilon\sigma\sqrt{n}\right]$$
$$+\lim_{\delta\downarrow 0}\limsup_{n\to\infty} 2P\left[\max_{1\leq j\leq \lfloor n\delta\rfloor+1} |S_j| > \frac{1}{3}\epsilon\sigma\sqrt{n}\right] = 0+0$$

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty} P\left[\max_{\substack{1\leq j\leq \lfloor n\delta\rfloor+1\\ 0\leq k\leq \lfloor nT\rfloor+1}} |S_{j+k}-S_k| > \epsilon\sigma\sqrt{n}\right] = 0$$

□

We are now in a position to establish the main result of this section , namely the convergence in distribution of the sequence of normalized random walks in equation (A.4.2) to Brownian motion. This result is known as the invariance principle.

Theorem A.4.4. *(Theorem 2.4.20 Shreve) Let $(\Omega,\mathcal{F},P)$ be a probability space on which a given sequence $\{\xi_j\}_{j=1}^{\infty}$ of independent, identically distributed random variables with mean zero and finite variance $\sigma^2 > 0$. Defined $X^{(n)} = \{X_t^{(n)}, t \geq 0\}$ by equation (A.4.2 and let $P^{(n)}$ be the measure induced by $X^{(n)}$ on $(C[0,\infty),\mathcal{B}(C[0,\infty)))$. Then $\{P_n\}_{n=1}^{\infty}$ converges weakly to a measure P_* under which the coordinate mapping process $W_t(\omega) := \omega(t)$ on $C[0,\infty)$ is a standard one dimensional Brownian Motion.*

Proof. In light of Theorem A.3.1 which says that convergence is finite dimensional distribution of a continuous stochastic process implies convergence in distribution of probability measures induced by these continuous process on $C[0,\infty)$ under tightness and Theorem A.4.1 which proves that the finite dimensional distributions of linearly interpolated normalized random walks converge to the finite dimensional distribution of Brownian motion, we just need to show the tightness of the sequence $\{X^{(n)}\}_{n=1}^{\infty}$ of linearly interpolated and normalized random walks defined in equation (A.4.2). In order to show tightness we use A.2.6. Since by definition $X_0^{(n)} = 0$ a.s for every n, equation (*A.2.6*) follows immediately since $P_n(\omega(0) = 0)$ for all $n \geq 1$. In order to show tightness of the sequence $\{X^{(n)}\}_{n=1}^{\infty}$ we need to establish equation (*A.2.7*) i.e for an arbitrary $\epsilon > 0$ and $T > 0$, the convergence

$$\lim_{\delta \downarrow 0} \sup_{n \geq 1} P\left[\max_{\substack{|s-t| \leq \delta \\ 0 \leq s,t \leq T}} |X_s^{(n)} - X_t^{(n)}| > \epsilon\right] = 0 \tag{A.4.8}$$

We may replace $\sup_{n \geq 1}$ in this expression by $\limsup_{n \to \infty}$,since for a finite number of integers n we can make the probability appearing in (*A.4.8*) as small as we choose by reducing δ. Indeed we have that

$\lim_{\delta \downarrow 0}\left[\sup_{n \geq 1} a_{n,\delta}\right] \implies \forall \epsilon > 0 \exists \delta_\epsilon > 0, \forall 0 < \delta < \delta_\epsilon$

$0 \leq \sup_{n \geq 1} a_{n,\delta} < \epsilon$

On the other hand $\lim_{\delta \downarrow 0} \limsup_{n \to \infty} a_{n,\delta} = 0$ implies that for all $\epsilon > 0$ there exists a δ_ϵ such that for all $0 < \delta < \delta_\epsilon$ we have $0 \leq \limsup_{n \to \infty} a_{n,\delta} < \epsilon$ or $0 \leq \inf_{n \geq 1} \sup_{k \geq n} a_{k,\delta} < \epsilon$. Now by the definition of infimum and the fact that its strictly less that ϵ we have that there exists a n_ϵ such that for all $n \geq n_\epsilon$ we have $\sup_{k \geq n} a_{k,\delta} < \epsilon$. This together with the fact that for a finite number of integers n we can make the probability appearing in (*A.4.8*) as small as we choose by reducing δ explains why we can replace the sup and the lim sup.

But by definition of $X^{(n)}$ in equation (*A.4.2*)

$$P\left[\max_{\substack{|s-t| \leq \delta \\ 0 \leq s,t \leq T}} |X_s^{(n)} - X_t^{(n)}| > \epsilon\right] = P\left[\max_{\substack{|s-t| \leq n\delta \\ 0 \leq s,t \leq nT}} |Y_s - Y_t| > \epsilon \sigma \sqrt{n}\right]$$

and

$$\max_{\substack{|s-t| \leq n\delta \\ 0 \leq s,t \leq nT}} |Y_s - Y_t| \leq \max_{\substack{|s-t| \leq \lfloor n\delta \rfloor + 1 \\ 0 \leq s,t \leq \lfloor nT \rfloor + 1}} |Y_s - Y_t| \leq \max_{\substack{|s-t| \leq \lfloor n\delta \rfloor + 1 \\ 0 \leq s,t \leq \lfloor nT \rfloor + 1}} |S_{k+j} - S_k|$$

where the first inequality is a consequence of the fact that the maximum is taken over a larger set and the last inequality follows from the fact that Y is piecewise linear constructed by interpolating the discrete process S and hence changes slope only at integral values of t. Now equation (A.4.8) follows from Lemma A.4.3 . Indeed

$$\lim_{\delta\downarrow 0}\sup_{n\geq 1} P\left[\max_{\substack{|s-t|\leq\delta \\ 0\leq s,t\leq T}} |X_s^{(n)} - X_t^{(n)}| > \epsilon\right] = \lim_{\delta\downarrow 0}\limsup_{n\to\infty} P\left[\max_{\substack{|s-t|\leq\delta \\ 0\leq s,t\leq T}} |X_s^{(n)} - X_t^{(n)}| > \epsilon\right]$$

$$= \lim_{\delta\downarrow 0}\limsup_{n\to\infty} P\left[\max_{\substack{|s-t|\leq n\delta \\ 0\leq s,t\leq nT}} |Y_s - Y_t| > \epsilon\sigma\sqrt{n}\right] \leq \lim_{\delta\downarrow 0}\limsup_{n\to\infty} P\left[\max_{\substack{|s-t|\leq\lfloor n\delta\rfloor+1 \\ 0\leq s,t\leq\lfloor nT\rfloor+1}} |Y_s - Y_t| > \epsilon\sigma\sqrt{n}\right]$$

$$\leq \lim_{\delta\downarrow 0}\limsup_{n\to\infty} P\left[\max_{\substack{1\leq j\leq\lfloor n\delta\rfloor+1 \\ 0\leq k\leq\lfloor nT\rfloor+1}} |S_{j+k} - S_k| > \epsilon\sigma\sqrt{n}\right] = 0$$

□

Index

Chapter 5

Bibliography

[1] D. Greenhalgh, Y. Liang, and X. Mao, "SDE SIS epidemic model with demographic stochasticity and varying population size," *Applied Mathematics and Computation , 276 (2016) 218-238*, 2016.

[2] D. Duffie and R. Kan, "A yield factor model of interest rates," *Mathematical Finance 6, 379-406*, 1996.

[3] S. Bernstein, "Principes de la théorie des équations différentielles stochastique," *Trudy Fiz.-Mat.,Stekov Inst., Akad. Nauk.5, 95-124*, 1934.

[4] S. Bernstein, "Équations différentielles stochastique," *Conf. Intern. Sci. Math. Univ Geneve .Act. Sci. Ind. 738,5-31*, 1938.

[5] I. Gihman and A. Skorohod, *Stochastic differential equations.* 1972.

[6] I. Gihman, "Certain differential equations with random functions (in russian)," *Uskr. Math. Z. 2 (3),45-69*, 1947.

[7] I. Gihman, "On the theory of differential equations of random processes (in russian)," *Uskr. Math. Z. 2 (4),37-63*, 1947.

[8] I. Gihman, "A method of constructing random processes (in russian)," *Akad. Nauk. SSSR 58,961-964*, 1947.

[9] K. Itô, "On a stochastic integral equation," *Proc. Imp. Acad. Tokyo,* ***22*** *32-35*, 1946.

[10] T. Yamada and S. Watanabe, "On the uniqueness of solutions of stochastic differential equations," *J. Math. Kyoto Univ* ***11*** *155-167*, 1971.

[11] N. Ikeda and S. Watanabe, *Stochastic differential equations and Diffusion Processes.* North-Holland Publishing Companny/Kodansha Ltd., 1981.

[12] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus, Second Edition.* Springer, 1996.

[13] A. Cairns, *Interest Rate Models: An Introduction.* Princeton University Press, Princeton and Oxford, 2004.

[14] E. Bernardi, V. Chuni, and A. Lanconelli, "On a general model system related to affine stochastic differential equations," *PrePrint*, 2019.

[15] E. Bernardi, V. Chuni, and A. Lanconelli, "On a class of stochastic differential equations with random and hölder continuous coefficients arising in biological modeling," *Journal of Nonlinear Science*, Jun 2019.

[16] K. Itô, "Differential equations determining markov processes(in japanese)," *Zenkoku Shijo Sugaku Danwakai **1077** 1352-1400*, 1942.

[17] K. Itô, "Stochastic integral," *Proc. Imp. Acad. Tokyo, **20** 519-524*, 1944.

[18] H. Kunita and S. Watanabe, "On square integrable martingales," *Nagoya Math J. **30** 209-245*, 1967.

[19] A. Skorohod, *Studies in the theory of Random Processes.* 1965.

[20] D. Stroock and S. Varadhan, "Diffusion processes with continuous coefficients," *Comm. Pure and Appl. Math. 22,345-400 and 479-530*, 1969.

[21] X. Mao, *Stochastic differential equations and Applications 2nd Edition.* Second edition, Horwood, Chichester, UK, 2008.

[22] J. Cox, J. Ingersoll, and S. Ross, "Ta theory of the term structure of interest rates," Econometrica *. 53: 385-407.*, 1985.

[23] O. Vasicek, "An equilibrium characterization of the term structure," Journal of Financial Economics *5 (2): 177-188.*, 1977.

[24] W. Feller, "The parabolic differential equations and the associated semi-groups of transformations," Ann. Math*. 55, 468-519*, 1952.

[25] G. Leobacher and M. Szlgyenyi, "A strong order 1/2 method for multidimensional sdes with discontinuous drift," *Ann. Appl. Probab.*, vol. 27, pp. 2383–2418, 08 2017.

[26] G. Leobacher, M. Szolgyenyi, and S. Thonhauser, "On the existence of solutions of a class of sdes with discontinuous drift and singular diffusion," 2015.

[27] A. Jentzen, B. Kuckuck, T. Mller-Gronbach, and L. Yaroslavtseva, "On the strong regularity of degenerate additive noise driven stochastic differential equations with respect to their initial values," 2019.

[28] T. Mller-Gronbach and L. Yaroslavtseva, "A note on strong approximation of sdes with smooth coefficients that have at most linearly growing derivatives," 2017.

[29] L. Allen, "An introduction to stochastic epidemic models in mathematical epidemiology," F. Brauer, P. van den Driessche, J. Wu (Eds.), Lecture Notes in Biomathematics, Mathematical Biosciences Subseries*, vol. 1945, Springer-Verlag, Berlin (2008) 81-130.*, 2008.

[30] L. Allen and A. Burgin, "Comparison of deterministic and stochastic sis and sir models in discrete time," Math. Biosci. *163 (2000) 1-33.*, 2000.

[31] A. Gray, D. Greenhalgh, L. Hu, X. Mao, and J. Pan, "Comparison of deterministic and stochastic sis and sir models in discrete time," Math. Biosci. *163 (2000) 1-33.*, 2000.

[32] H. Hethcote and P. van den Driessche, "An SIS epidemic model with variable population size and a delay," J. Math. Biol. *34, 177-194*, 1995.

[33] R. Kryscio and C. Lefévre, "On the extinction of the sis stochastic logistic epidemic," J. Appl. Probab. *26 (4) (1989) 685-694.*, 1989.

[34] R. McCormack and L. Allen, "Stochastic SIS and SIR multihost epidemic models," R.P. Agarwal, K. Perera (Eds.), Proceedings of the Conference on Differential and Difference Equations and Applications*, Hindawi, New York (2006) 775-786.*, 2006.

[35] I. Nasell, "The quasi-stationary distribution of the closed endemic SIS model," Adv. Appl. Probab. *28 ,895-932.*, 1996.

[36] E. Allen, *Modeling with Ito Stochastic Differential Equations.* Springer, 2011.

[37] V. Zubchenko, "Properties of solutions of stochastic differential equations with random coefficients NON-LIPSCHITZIAN DIFFUSION, AND POISSON MEASURES," *Theor. Probability and Math. Statist.*, 2011.

[38] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion Corrected 3rd edition.* Springer, 2004.

[39] P. Graczyk and J. Malecki, "Multidimensional Yamada-Watanabe theorem and its applications to particle systems," J. Math. Phys. ***54*** *021503*, 2013.

[40] K. Kumar, "A class of degenerate stochastic differential equations with non-lipschitz coefficients," *Proceedings-Mathematical Sciences 123 443-454*, 2013.

[41] D. Luo, "Pathwise uniqueness of multi-dimensional stochastic differential equations with hölder diffusion coefficients," Front. Math. China ***6*** *(2011) 129-136*, 2011.

[42] Z. Wang and X. Zhang, "Strong uniqueness of degenerate SDEs with hölder diffusion coefficients,," arXiv:1805.05526, 2018.

[43] E. Bernardi, A. Lanconelli, and S. Ramagnoli, "A distorted copula-based evolution model: Risks aggregation in a bonus-malus migration system," *Preprint*, 2019.

[44] Y. W.C., "Eigenvalues of several tridiagonal matrices," *Proceedings-Mathematical Sciences 123 443-454*, 2013.

[45] J. Jacod and P. E. Protter, *Probability Essentials, Second Edition.* Springer, 2004.