Adam Voldán

Itôův a Stratonovičův stochastický integrál

Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové Práce: Doc. RNDr. Daniel Hlubinka Ph.D.

Studijní program: Matematika
Studijní obor: Teorie pravděpodobnosti a náhodné procesy

2008
Na tomto místě bych rád poděkoval Doc. RNDr. Danielu Hlubinkovi Ph.D. za jeho trpělivost, rady a pomoc při vedení mé práce. Dále bych rád poděkoval své rodině za jejich neutuchající podporu.

Prohlašuji, že jsem svou diplomovou práci napsal samostatně a výhradně s použitím citovaných pramenů. Souhlasím se zapůjčováním práce a jejím zveřejňováním.

V Berouně dne

Adam Voldán
# Contents

1 Construction of stochastic integrals 7  
1.1 Basics of probability theory 7  
1.2 Itô stochastic integral 13  
1.3 Properties of stochastic integrals 20  

2 Numerical stochastic analysis 27  
2.1 Stochastic differential equations 27  
2.2 Stochastic Taylor expansion 34  
2.3 Multiple Itô and Stratonovich integrals 43  

3 Numerical implementation 47  
3.1 Applied algorithms 47  
3.2 Computer examples 52  

4 Stochastic modeling in finance 68  
4.1 Introduction to finance mathematics 68  
4.2 European lookback call option 77  

Bibliography 84

Klíčová slova: Itôův stochastický integrál, Stratonovičův stochastický integrál, Itôova formule, SDE, Itô-Taylorův rozvoj, Stratonovich-Taylor rozvoj, vícnásobný stochastický integrál, aproximace v diskrétním čase, konvergenční kritérium, Eulerovo schema, Mielsteinovo schema, evropská opce, Black-Scholes-Mertonova formule, evropská lookback call opce.

Title: Spatial modelling
Author: Adam Voldán
Department: Departement of Probability and Mathematical Statistics
Supervisor: Doc. RNDr. Daniel Hlubinka Ph.D.
Supervisor’s e-mail address: Daniel.Hlubinka@mff.cuni.cz

Abstract: In this thesis the Itô stochastic integral and the Stratonovich stochastic integrals are studied. Their basic and some special properties are shown. Further the theory of the numerical solution of stochastic differential equations (SDE) is introduced. Using simple examples the properties of chosen numerical schemes are presented. Finally the Black-Scholes-Merton formula for pricing of European call option is sketched, and similar problems are numerically solved using the above presented algorithms.

Keywords: Itô stochastic integral, Stratonovich stochastic integral, Itô formula, SDE, Itô-Taylor expansion, Stratonovich-Taylor expansion, multiple stochastic integrals, discrete time approximations, convergence criterion, Euler scheme, Mielstein scheme, European option, Black-Scholes-Merton formula, European lookback call option.
Introduction

Stochastic analysis [2], [3] and stochastic differential equations [3], [4], [5] are a widely applicable tool appears in many areas. We can met its applications in engineering, genetics, biomechanics and before all in financial mathematics.

The Itô stochastic integral [2], [3], [4] became a cornerstone of the stochastic calculus. It is preferred for many practical reasons especially its martingale property. However, the Itô stochastic differential given by Itô formula [2], [3], [4] caused a lot of additional work and disabled mathematicians to use stronger numerical methods except the Euler scheme [4], [5].

In 1978 Wagner and Platen introduced the Itô-Taylor [4] resp. Stratonovich-Taylor expansion [4], which was the first step in developing more effective numerical schemes for solving of SDEs. The second step was done by Kloeden and Platen by introducing how to approximate multiple Stratonovich integrals [4]. Using this two results we are able to approximate any SDE by a finite sum of stochastic integrals which can be solved or at least approximately evaluated.

The aim of this thesis is to introduce the Itô and Stratonovich stochastic integrals [2], [3], [4] and their basic properties and differences. Then by means of simple numerical examples show the advantages of using both integral approaches and possible applications in the financial mathematics.

This thesis consists of four chapters. In the first chapter the basic definitions and results from the probability theory [1] are mentioned. Then the construction of an Itô integral is presented [2], [3] and the most important properties and relations between both Itô and Stratonovich stochastic integral are explained [2], [3], [4].

In the second chapter the basic results of numerical solutions of stochastic differential equations are introduced. At first the theorem of existence and uniqueness of the solution of SDE [3] is stated, then the theory of stochastic Itô-Taylor and Stratonovich-Taylor expansion is build up [4]. At last the
most common algorithm for evaluating of multiple Stratonovich stochastic integrals from Kloeden-Platen [4] is described and studied using the numerical examples [5].

The third chapter introduces the strong discrete time approximations for solution of stochastic differential equations [4], [5]. By numerical examples the main properties are shown for both one dimensional and multi dimensional cases.

The fourth chapter presents the application of numerical solution methods in the field of financial mathematics. The Black-Scholes-Merton formula for option pricing is sketched and numerically estimated. Finally the European lookback call option [6] is studied by means of numerical methods.
Chapter 1

Construction of stochastic integrals

In this chapter we shall start with the elements of probability theory and then we shall see the steps of building up the Itô, resp. Stratonovich stochastic integrals. Finally the main properties and the relation is shown.

1.1 Basics of probability theory

In this section we will mention some basic definitions and results from probability theory and clarify the notation which will be used later on.

We shall start with the stochastic process definition, where we consider a nonzero set \([0, T] \subset \mathbb{R}\), as a set of time indexes.

Definition 1.1.1 Let \(\{X(t), t \in [0, T]\}\) be a family of real random variables on a probability space \((\Omega, \mathcal{F}, P)\), then we call it stochastic process.

The second elementary definition is the filtration, here especially the definition of completeness and the right continuity of filtration are worth stressing because they will be crucial for further work.

Definition 1.1.2 Suppose we have a measurable space \((\Omega, \mathcal{F})\), then a system of \(\sigma\)-algebras \(\{\mathcal{F}_t, t \geq 0\}\), satisfying

\[
\mathcal{F}_t \subset \mathcal{F}, \text{ for all } t \geq 0 \quad \text{and} \quad \mathcal{F}_s \subset \mathcal{F}_t, \text{ whenever } s \leq t, \forall s, t \geq 0
\]
is called a filtration of the measurable space \((\Omega, \mathcal{F})\), or simply a filtration. Let \((\Omega, \mathcal{F}, P)\) be a complete probability space, then a filtration which includes all \(P\)-zero-measure sets is called complete, i.e.

\[
\mathcal{N}_p \subset \mathcal{F}_0, \text{ where } \mathcal{N}_p := \{N : \exists F \in \mathcal{F}, N \subset F, P(N) = 0\}.
\]

And a filtration is called right continuous if \(\mathcal{F}_{t+} = \mathcal{F}_t\) for all \(t \geq 0\), where \(\mathcal{F}_{t+} := \cap_{h>0} \mathcal{F}_{t+h}\).

In addition, suppose we have a stochastic process \(\{X(t), t \geq 0\}\) on \((\Omega, \mathcal{F}, P)\), such that \(X(t)\) is \(\mathcal{F}_t\)-measurable for all \(t \geq 0\) (i.e. \(\sigma(X(s), s \leq t) \subseteq \mathcal{F}_t\)), then \(X(t)\) is called \(\mathcal{F}_t\)-adapted process. Finally we denote \(\mathcal{F}_\infty := \sigma(\bigcup_{t \geq 0} \mathcal{F}_t)\).

Because in literature the notation of convergence of random values varies, we will introduce the following notation.

**Definition 1.1.3** Let \(X_n, X\) be random values on \((\Omega, \mathcal{F}_t, P)\) we denote convergence almost sure with respect to measure \(P\)

\[
X_n \xrightarrow{\text{P-a.s.}} X \text{ if } \exists A \in \mathcal{F} : \lim_{n \to \infty} X_n(\omega) = X(\omega), \forall \omega \in A, P(A) = 1
\]

and denote convergence in probability

\[
X_n \xrightarrow{\text{P}} X \text{ if } \forall \epsilon \lim_{n \to \infty} P[|X_k - X| \geq \epsilon] = 0.
\]

**Definition 1.1.4** Stochastic process \(X(t), t \in [0, T]\) is a collection of random variables hence for each \(t \in [0, T]\) there is a measurable map

\[
X(t, \cdot) : (\Omega, \mathcal{F}_t) \to (\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m)), \tag{1.1}
\]

where \(\mathcal{B}(\cdot)\) denote the Borel \(\sigma\)-algebra on the concerning set. On the other hand for fixed \(\omega \in \Omega\) we get a function

\[
X(\cdot, \omega) : [0, T] \to \mathbb{R}^m. \tag{1.2}
\]

Such a function is called trajectory of a stochastic process. When put together we get a map

\[
X(t, \omega) : ([0, T] \times \Omega) \to \mathbb{R}^m.
\]
This map is called progressive-measurable if it is $(\mathcal{B}([0,T]) \times \mathcal{F}_t)$-measurable. The process is called $(P-a.s.)$-continuous if its trajectory is a continuous function $P-a.s.$ i.e.

$$X(t) \xrightarrow{P-a.s.} X(s) \quad \text{if } t \to s, \ t, s \in [0,T]$$

respectively is $P-$continuous if its trajectory is a continuous function in probability.

The trajectory of a process is an important connection between processes and functions. We can always easily switch from processes to functions just by fixing the $\omega$ and on the other hand by supposing a real valued function to be $\omega$ dependent and measurable (i.e. $\mathcal{A}-$measurable if $f : (\Omega, \mathcal{A}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$) we obtain the features of the real valued process definition. Hence we will write $X(t, \omega) = X(t)(\omega) = X(t)$ or also for functions $f(t) = f(t)(\omega) = f(t, \omega)$. If we should use an deterministic function it will be explicitly stated.

Next definition tell us how much the process differs in time.

**Definition 1.1.5** The processes $X(t), t \geq 0$ and $Y(t), t \geq 0$ are said to have finite covariation if there exist a process $\langle X, Y \rangle(t), t \geq 0$ such that

$$S_n(t) \xrightarrow{P} \langle X, Y \rangle(t) \quad \text{for all } |\Delta_n([0,t])| \to 0 \quad \text{for all } t \geq 0,$$

where

$$S_n(t) = \sum_{j=1}^{n} (X(t_j) - X(t_{j-1})) \cdot (Y(t_j) - Y(t_{j-1})), \quad (1.3)$$

and

$$|\Delta_n([0,t])| := \max_{1 \leq j \leq n} |t_j - t_{j-1}|$$

for a partition

$$\Delta_n([0,t]) := \{0 = t_0 < t_1 < \ldots < t_n\};$$

where $t_n \to \infty$ as $n \to \infty$.

The process $\langle X, X \rangle(t)$ is called quadratic variation and we write $\langle X \rangle(t)$.

In addition, function $f : \mathbb{R}^+ \to \mathbb{R}$ is of finite variation if holds

$$f^v(t) := \sup_\mathcal{D} \sum_{i=0}^{n-1} \left| f(t_{i+1}) - f(t_i) \right| < +\infty,$$

where $\mathcal{D}$ is a set of all partitions $\Delta_n([0,t]) := \{0 = t_0 < t_1 < \ldots < t_n\}$, for any $t_n \to \infty$ as $n \to \infty$. The map $t \mapsto f^v(t)$ will be called variation of $f$. 
Remark 1.1.1 Suppose processes $X(t), Y(t), Z(t)$ to be $p$-continuous and to have finite quadratic variation and $a, b$ to be constants, then
\[
\langle (aX + Z), b(Y + Z) \rangle(t) = ab(X, Y)(t) + ab(X, Z)(t) + b(Z, Y)(t) + b(Z)(t).
\]

Further $f$ be function of finite variation, then
\[
\langle X, f \rangle(t) = \langle f, X \rangle(t) = \langle f, f \rangle(t) = 0.
\]

Proof: First equation is obtained from (1.3) because the $p$-limit saves the linearity. Denote $\Delta f_j = (f(t_1) - f(t_{j-1}))$, $\Delta X_j = (X(t_j) - X(t_{j-1}))$ for all $j = 1, \ldots, n$. Then the second statement is obtained as from Cauchy-Schwartz inequality
\[
S_n(t) \leq \sum_{j=1}^{n} (\Delta f_j)^2 \cdot (\Delta X_j)^2 \leq 0 \cdot C.
\]

Definition 1.1.6 A process $W(t), t \geq 0$ is called Wiener process if

(i) $W(0) = 0$,

(ii) $\mathcal{L}(W(t) - W(s)) = N(0, t - s)$ for $t > s$,

(iii) $P[W(t_1) - W(t_0) < w_1, \ldots, W(t_n) - W(t_{n-1}) < w_n] = \otimes_{i=1}^{n} P[W(t_i) - W(t_{i-1}) < w_i],$

for $n \in \mathbb{N}$, $w_i \in \mathbb{R}$ and all finite sequences $t_0 < t_1 < \ldots < t_n < \infty$. Furthermore, if $W(t), t \geq 0$ is $\mathcal{F}_t$-adapted and
\[
0 \leq s < t \leq \infty \Rightarrow \left( W(t) - W(s) \right) \text{ and } \mathcal{F}_s \text{ are independent},
\]

then $W(t)$ is called $\mathcal{F}_t$-Wiener process.

A process $W(t), t \geq 0$ such as $W(t) = (W^1(t), \ldots, W^d(t))$, where $W^i(t)$ and $W^j(t), i \neq j$ are independent $\mathcal{F}_t$-Wiener processes, is called a $d$-dimensional $\mathcal{F}_t$-Wiener process.

In this work the $\mathcal{F}_t$-adapted $W(t), t \geq 0$ is used.
Theorem 1.1.1 Let $W(t), t \geq 0$ be a $\mathcal{F}_t-$adapted Wiener process, suppose a partition $\Delta_n([s, t]) := \{s = t_1 \leq t_2 \leq \ldots \leq t_n = t\}$ and $0 \leq s \leq t < \infty$. Denote again $|\Delta_n| := \max_i |t_{i+1} - t_i|$. Then

\begin{enumerate}[(i)]
\item $|\Delta_n| \to 0 \implies S_n := \sum_{i=0}^{n-1} \left| (W(t_{i+1}) - W(t_i)) \right|^2 \to (t - s) \quad L^2,$
\item $\sum_{n=1}^{\infty} |\Delta_n| < \infty \implies S_n \to (t - s) \quad P-a.s.$
\end{enumerate}

Proof: The theorem is proved in [?] as theorem number VII. 1.4.

Now we need to recall some basic definitions and theorems to construct the space where the stochastic integrals will be built.

Definition 1.1.7 Suppose a probability space $(\Omega, \mathcal{F}, P)$ and $X(t), t \geq 0$ to be a $\mathcal{F}_t$-adapted process $X(t) \in L^1$ such as $\mathbb{E}[X(t)|\mathcal{F}_s] \overset{P-a.s.}{=} X(s)$, for $s \leq t$ we call a $\mathcal{F}_t$-martingale.

$\mathbb{E}[X(t)|\mathcal{F}_s] \overset{P-a.s.}{\geq} X(s)$, for $s \leq t$ we call a $\mathcal{F}_t$-submartingale.

$\mathbb{E}[X(t)|\mathcal{F}_s] \overset{P-a.s.}{\leq} X(s)$, for $s \leq t$ we call a $\mathcal{F}_t$-supermartingale.

Definition 1.1.8 A random variable $\tau : \Omega \to [0, \infty]$ such as

$\{\tau(\omega) \leq t\} \in \mathcal{F}_t$, for all $t \geq 0$

for a filtration $\{\mathcal{F}_t\}_{t \geq 0}$ is called $\mathcal{F}_t-$stopping time. Furthermore, we expand the filtration $\{\mathcal{F}_t\}_{t \geq 0}$ introducing

$\mathcal{F}_\tau = \{F \in \mathcal{F}_\infty : F \cup [\tau(\omega) \leq t] \in \mathcal{F}_t \ \forall t \in \mathbb{R}^+\}$.

Theorem 1.1.2 Let $X(t)$ be a right-continuous $\mathcal{F}_t-$martingale resp. $\mathcal{F}_t-$submartingale and $\nu, \tau$ be a $\mathcal{F}_t-$stopping times such as $\nu \leq \tau \leq T \leq \infty$ holds $P-a.s.$ Then

$X(\nu), X(\tau) \in L^1$, and $\mathbb{E}[X(\tau)|\mathcal{F}_\nu] = X(\nu)$, resp. $\mathbb{E}[X(\tau)|\mathcal{F}_\nu] \leq X(\nu)$.

holds $P-a.s.$
Proof: It is proofed in [2] as theorem III. 1.3.10.

Definition 1.1.9 Let \( X(t) \) be stochastic process and \( \tau(\omega) \) be a stopping time then the
\[
X^\tau(t) := X(\tau \wedge t) := \begin{cases} 
X(t) & \text{for } t < \tau(\omega), \\
X(\tau(\omega)) & \text{for } t \geq \tau(\omega).
\end{cases}
\]

Theorem 1.1.3 Let \( X(t), t \geq 0 \) be a right-continuous \( \mathcal{F}_t \)-martingale and \( \tau \) be a \( \mathcal{F}_t \)-stopping time then \( X^\tau \) is a right-continuous \( \mathcal{F}_t \)-martingale. And let \( \nu, \tau \) be \( \mathcal{F}_t \)-stopping times such as \( \nu \leq \tau \) holds \( P \)-a.s. Then
\[
\mathbb{E} \left[ X^\tau(t) \mid \mathcal{F}_\nu \right] = X^\nu(t), \quad \forall t \geq 0.
\]
Moreover is \( X(t) \) is a bounded process and \( \nu \leq \tau \) \( P \)-a.s. finite stopping times, then \( \mathbb{E} \left[ X(\tau) \mid \mathcal{F}_\nu \right] = X(\nu) \).

Proof: It is proofed in [2] as theorem III. 1.3.11.

Definition 1.1.10 Let \( L(t) \) be \( \mathcal{F}_t \)-adapted and \( P \)-a.s. continuous stochastic process on \((\Omega, \mathcal{F}, P)\). If there exist a sequence of \( \mathcal{F}_t \)-stopping times \( \tau_n(\omega) \) such as \( \tau_n(\omega) \nearrow \infty \) \( P \)-a.s. and for which holds
\[
L(\tau_n \wedge t) - L(0) = (L(t) - L(0))^{\tau_n} \text{ is } \mathcal{F}_t \text{-martingale for all } n \in \mathbb{N},
\]
then we call \( L(t) \) to be a local martingale and the sequence \( \tau_n(\omega) \) is called localization sequence.

Now we are able to introduce space on which we will build up the stochastic integral.

Definition 1.1.11 Let \( CM_2(\mathcal{F}, [0, T]) \) denote the space of all \( \mathcal{F}_t \)-adapted and \( P \)-a.s. continuous martingales \( M(t) \) with \( t \in [0, T] \), such as
\[
\mathbb{E} \left[ (M(t))^2 \right] < \infty \quad \text{for all } M(t) \in CM_2(\mathcal{F}, [0, T]). \tag{1.4}
\]
Finally denote \( CP_2(\mathcal{F}, [0, T]) \) as a set of all \( P \)-a.s. continuous, \( \mathcal{F}_t \)-adapted processes such as
\[
\mathbb{E} \left[ \max_{s \leq t} \{X^2(s)\} \right] < \infty \quad \text{for all } t \geq 0.
\]
Then for any $M(t), N(t) \in CP_2$ we define a pseudometric $m(\cdot, \cdot)$ by prescription

$$m(M(t), N(t)) := \sum_{t=1}^{\infty} \sqrt{\mathbb{E}[\max_{s \leq t} \{ |M(s) - N(s)|^2 \}] / 2^t} \wedge 1.$$ 

Finally the last theorem of this section provides us with the crucial result for the construction of Itô stochastic integral. Here in this theorem the complete filtration is required.

**Theorem 1.1.4** Suppose a fixed complete filtration $\{\mathcal{F}_t\}_{t \geq 0}$ the pseudometric space $CP_2(\mathcal{F}, [0, T])$ is complete and $CM_2(\mathcal{F}, [0, T]) \subset CP_2(\mathcal{F}, [0, t])$ is a closed subspace.

**Proof:** The theorem is proved in [2] as theorem number III. 1.5.2.

\[\square\]

### 1.2 Itô stochastic integral

Now we are coming to the "scope" of the work - the comparison of the two most used definitions of a stochastic integral. Suppose we want to compute

$$\mathbb{E} \left[ \int_0^T \phi(t, \omega) dW(t) \right],$$

for some suitable function $\phi(t, \omega)$ and any $T \in \mathbb{R}^+$. Then, as we shall see later, we will be looking for approximation like

$$\int_0^T \phi(t, \omega) dW(t) := \sum_{\Delta(T)} e_j(\omega)[W(t_{j+1}) - W(t_j)],$$

for some partition $\Delta_n([0, T]) = \{0 = t_0 < \cdots < t_n = T\}$ of the time interval $[0, T]$. Here the definition of $e_j(\omega)$ will be crucial. As we can see, $e_j(\omega)$ should approximate $\phi(u)(\omega)$ in some valuation point $u \in [t_j, t_{j+1}]$. So there are many possibilities which definition to take, but in general the different choice leads to a different random variable, as is shown in following example.
Example 1.2.1 Let $\phi(t) = W(t)$ and choose
\[
\begin{align*}
\text{i}) \quad e_j(\omega) &:= \phi(t_j)(\omega) \\
\end{align*}
\]
we will obtain
\[
\begin{align*}
\mathbb{E}\left[ \int_0^T \phi(t)dW(t) \right] &= \mathbb{E}\left[ \sum_{\Delta(T)} W(t_j)[W(t_{j+1}) - W(t_j)] \right] \\
&= \sum_{\Delta(T)} \mathbb{E}\left[ \mathbb{E}\left( W(t_j)[W(t_{j+1}) - W(t_j)]|\mathcal{F}_{t_j} \right) \right] \\
&= \sum_{\Delta(T)} \mathbb{E}[W(t_j)] \mathbb{E}[W(t_{j+1}) - W(t_j)] \\
&= \sum_{\Delta(T)} \left( \mathbb{E}[W(t_j)] \cdot 0 \right) \\
&= 0.
\end{align*}
\]
And, on the other hand, choose
\[
\begin{align*}
\text{ii}) \quad e_j(\omega) &:= \phi(t_{j+1})(\omega) \\
\end{align*}
\]
to get
\[
\begin{align*}
\mathbb{E}\left[ \int_0^T \phi(t)dW(t) \right] &= \mathbb{E}\left[ \sum_{\Delta(T)} W(t_{j+1})[W(t_{j+1}) - W(t_j)] \right] \\
&= \sum_{\Delta(T)} \mathbb{E}\left[ W(t_{j+1}) - W(t_j) \right]^2 + \mathbb{E}\left[ W(t_j)[W(t_{j+1}) - W(t_j)] \right] \\
&= \sum_{\Delta(T)} (t_{j+1} - t_j) + 0 \\
&= T.
\end{align*}
\]
However, both definitions seem to be reasonable. The results are very different. That's why the good choice of definition is so important. The most used definitions are the Itô’s and the Stratonovich’s definition.
Before we start building of Itô’s stochastic integral, we prepare following metric spaces
Definition 1.2.1 We call $K(t), t \geq 0$ an $\mathcal{F}_t-$adapted simple process if success

$$K(t) = K_0 \cdot I_0(t) + \sum_{i=0}^{\infty} K_i I_{(t_i, t_{i+1})}(t),$$

for a locally finite partition

$$\Delta_n([0, T]) = \{0 = t_0 < t_1 < t_2 < \ldots\},$$

where $K_i$ are bounded $\mathcal{F}_{t_i} -$adapted random values for all $i \in \mathbb{N}$.

A set of all $\mathcal{F}_t-$adapted simple processes $K(t), t \in [0, T]$ will be denoted $J(\mathcal{F}, [0, T])$.

Definition 1.2.2 (Integrable processes) Let $L^2(\mathcal{F}, [0, T])$ be a metric space of processes $X(t), t \in [0, T]$ such as

(i) $X(t)$ is $\mathcal{F}_t -$progressive,

(ii) $\mathbb{E}[\int_0^t X(s)^2 ds] < \infty$ for all $t \in [0, T]$.

with metric $l(G, H)$ for $G(t), H(t) \in L^2(\mathcal{F}, [0, T])$

$$l(G(t), H(t)) := \sum_{t=1}^{\infty} \min \left\{ 1; \sqrt{\mathbb{E}[\int_0^t (G(s) - H(s))^2 ds]} \right\} 2^t.$$
Here we should notice that the metric $m(\cdot, \cdot)$ is defined with respect to probability measure $P$, and in fact it is just pseudometric. Thus we see that the assumption of a right continuous and complete filtration $\mathcal{F}_t$ is crucial. Because under these assumptions $P - a.s.$ pseudometric is satisfying.

**Definition 1.2.3 (Itô simple stochastic integral)** Let $W(t), t \geq 0$ be a Wiener process, $K(t) \in J(\mathcal{F}, [0, T])$ and for any $T \in \mathbb{R}^+$ be

$$\Delta_n([0, T]) = \{0 = t_0 < \cdots < t_n = T\}$$

a partition of $[0, T]$. Then we define the Itô simple stochastic integral as follows

$$\left(\mathcal{I}^T[K]_0^T\right)(\omega) := \int_0^T K(s, \omega) dW(s, \omega) := \sum_{i=0}^{n-1} K_i(\omega) (W(t_{i+1}, \omega) - W(t_i, \omega)).$$

(1.5)

The completeness of $(CM_2(\mathcal{F}, [0, T]), m)$ is provided by theorem 1.1.4. Now we will sketch the proof of the density theorem.

**Theorem 1.2.1 (density theorem)** Let $(\mathcal{F}, [0, T])$ be an complete filtration, then $J(\mathcal{F}, [0, T])$ is a dense set in $(L_2(\mathcal{F}, [0, T]), l)$.

**Proof:** Suppose first that $G(t) \in L_2(\mathcal{F}, [0, T])$ is bounded and continuous for each $\omega$, then we can define a $K(t) \in J(\mathcal{F}, [0, T])$ with prescription

$$K_n(t)(\omega) := G(0)I_{[0]} + \sum_j G(t_j, \omega) \cdot I_{(t_j, t_{j+1}]}$$

for some division $\Delta_n([0, T])$. Then $K_n(t)$ is also bounded and we became

$$\int_0^T (G(t) - K_n(t))^2 dt \to 0 \text{ as } n \to \infty, \forall \omega$$

Thus also $l(X(t), K_n(t)) \to 0$ as $n \to \infty$ by bounded convergence.

Now we prove that all bounded processes $H(t) \in L_2(\mathcal{F}, [0, T])$ can be approximated by bounded processes $G_n(t) \in L_2(\mathcal{F}, [0, T])$. We will define

$$G_n(t) := n \cdot \int_{(t-\frac{1}{n})^+}^{t} H(t) dt, \ t \geq 0, n \in \mathbb{N},$$

16
Because \(|H(t)| \leq C < \infty\) also \(|G_n(t)| < C\). It is not crucial to prove that \(\int H(s)ds\) is a continuous and \((\mathcal{F}, [0, T])\)-adapted process everywhere on \(\Omega\). If we rewrite definition of \(G_n(t)\), by fundamental theorem of calculus we see that

\[
G_n(t) = \int_0^t H(s)ds - \int_0^{(t-\frac{1}{n})^+} H(s)ds \quad \text{as} \quad n \to \infty
\]

almost everywhere on \(\mathbb{R}^+\) with respect to Lebesque measure \(\lambda\). Both processes are bounded on \([0, T]\), hence we have convergence in \(l\) by the dominated convergence.

Finally, we prove that bounded processes \(H_n(t) \in L_2(\mathcal{F}, [0, T])\) can approximate all processes \(X(t) \in L_2(\mathcal{F}, [0, T])\). We define the appropriate sequence as \(H_n := X(t) \cdot I_{[-n,n]}(X(t))\). As supposed \(X(t) \in L_2(\mathcal{F}, [0, T])\) so \(E[X^2(t)] < \infty\) thus

\[
E[|X \cdot I_{[X>\cdot]}]|] \to 0.
\]

In other words \(E[H_n - X] \to 0 \ (\Leftrightarrow l(H_n, X) \to 0)\) for almost all \(\omega\) with respect to probability measure \(P\). Because we supposed a complete filtration, all such a zero sets are included.

\[\square\]

**Theorem 1.2.2** Let \(K(t) \in J(\mathcal{F}, [0, T])\) then

\[
E\left[\left(\int_0^T K(t)dW(t)\right)^2\right] = E\left[\int_0^T K^2(t)dt\right].
\]  

**Proof:** Put \(\Delta W_j = W(t)_{j+1} - W(t)_j\), since \(\Delta W_j\) is independent from \(\mathcal{F}_{t_j}\) hence by \(\mathcal{F}_t\)-adaptivity also independent from \(K_j(\omega)\) we have

\[
E[K_iK_j\Delta W_i\Delta W_j] = \begin{cases} E[K_iK_j\Delta W_i]E[\Delta W_j] = 0 & \text{if } i < j \\ E[K_iK_j\Delta W_j]E[\Delta W_i] = 0 & \text{if } j < i \\ E[K_i^2]E[(\Delta W_i)^2] & \text{if } j = i \end{cases}
\]

Thus

\[
E\left[\left(\int_0^T K(t)dW(t)\right)^2\right] = \sum_{i,j} E[K_iK_j\Delta W_i\Delta W_j] =
\]

\[
= \sum_i E[K_i^2]E[(\Delta W_i)^2] = \sum_i E[K_i^2](t_{i+1} - t_i) = E\left[\int_0^T K^2(t)dt\right].
\]

\[\square\]
Remark 1.2.1 Let $K(t) \in J(F, [0, T])$, then $T[K]_{0}^{T} \in CM_{2}(F, [0, T])$.

Proof: Simple process is bounded from the definition. We can use Itô isometry 1.6 to see that
\[
\mathbb{E}\left[\left( T[K]_{0}^{T} \right)^{2} \right] = \mathbb{E}\left[ \left( \int_{0}^{T} K(t) dW(t) \right)^{2} \right] = \mathbb{E}\left[ \int_{0}^{T} K^{2}(t) dt \right] \leq \max_{0 \leq j \leq n} K_{j}^{2} \int_{0}^{T} dt < \infty.
\]

Next we prove that $T[K]_{0}^{T}$ is $\mathcal{F}_{T}$-martingale. Simple integral is a countable sum of $\mathcal{F}_{T}$-adapted processes, hence it is $\mathcal{F}_{T}$-adapted.

Without loss of generality we can expect $s = t_{k}$ and $T = t_{n}$ where $s < T$, $k < n$, for any $T \in \mathbb{R}^{+}$, if not we will take more precise partition. Denote $\Delta W_{j} = (W(t_{j+1}) - W(t_{j}))$ for all $j = 0, \ldots, n - 1$ then
\[
\mathbb{E}\left[ T[K]_{0}^{T} \mid \mathcal{F}_{s} \right] = \mathbb{E}\left[ \sum_{j=0}^{n-1} K_{j} \Delta W_{j} \mid \mathcal{F}_{t_{k}} \right] = \sum_{j=0}^{k-1} K_{j} \left( W(t_{j+1}) - W(t_{j}) \right) + \sum_{j=k}^{n-1} \mathbb{E}\left[ \Delta W_{j} \mid \mathcal{F}_{t_{j}} \right] \mathbb{E}\left[ K_{j} \mid \mathcal{F}_{t_{j}} \right] \mid \mathcal{F}_{t_{k}} \right] = T_{0}^{t_{k}}(K) + \mathbb{E}\left[ \sum_{j=k}^{n-1} K_{j} \cdot 0 \right]
\]
\[
= T_{0}^{t_{k}}(K) + 0.
\]

So we proved the martingale property.

Finally we prove that $T_{0}^{T}$ is $P$-a.s.-continuous. For any $h$ we can make an equidistant partition $|\Delta([0, T])| = h$ then
\[
\left| (T[K]_{0}^{T}) (\omega) - (T[K]_{0}^{T-h}) (\omega) \right| = \left| \left( \sum_{j=0}^{n-1} K_{j} \Delta W_{j} \right) (\omega) - \left( \sum_{j=0}^{n-2} K_{j} \Delta W_{j} \right) (\omega) \right|
\]

since the Wiener process $W(t)$ is $P$-a.s.-continuous and $K_{j}$ is a bounded random value for any $j = 1, \ldots, n$ we obtain for $h \to 0$
\[
|K_{n-1}(\omega)| \cdot |\Delta W_{n-1}(\omega)| \to 0, \quad P - \text{almost all } \omega.
\]

$\diamond$
Now any \( X(t, \omega) \in L_2(\mathcal{F}, [0, T]) \) can be approximated by \( K(t) \in J(\mathcal{F}, [0, T]) \) such as
\[
\mathbb{E} \left[ \int_0^T \left( K_n(t) - X(t) \right)^2 dt \right] \to 0,
\]
because of density theorem 1.2.1. Then define
\[
\mathcal{I}_T^0(X(t))(\omega) := \int_0^T X(t, \omega) \, dW(t, \omega) := \lim_{n \to \infty} \int_0^T K_n(t)(\omega) \, dW(t)(\omega).
\]
Since \( \int_0^T (K_n(t))^2 dt \) is a Cauchy sequence in \( (L_2(\mathcal{F}, [0, T]), l) \) it is also Cauchy sequence in \( (CM_2(\mathcal{F}, [0, T]), m) \) by Itô isometry 1.6. Because \( (L_2(\mathcal{F}, [0, T]), l) \) is complete there exist one and only one limit so
\[
\int_0^T (K_n(t)) \, dW(t) \overset{m(\cdot, \cdot)}{\to} \int_0^T (X(t)) \, dW(t).
\]

**Definition 1.2.4 (Itô stochastic integral)** Let for any \( T \in \mathbb{R}^+ \) be \( X(t, \omega) \in L_2(\mathcal{F}, [0, T]) \). Then the Itô integral of \( X(t) \) is defined by
\[
\left( \mathcal{I}_T^0(X(t)) \right)(\omega) := \int_0^T X(t, \omega) \, dW(t, \omega) = \lim_{n \to \infty} \int_0^T K_n(\omega) \, dW(t)(\omega), \quad (1.7)
\]
limit in \( m(\cdot, \cdot) \), where \( K_n(t) \in J(\mathcal{F}, [0, T]) \) such that
\[
\mathbb{E} \left[ \int_0^T \left( K_n(t, \omega) - X(t, \omega) \right)^2 dt \right] \to 0, \text{ as } n \to \infty. \quad (1.8)
\]

We should remind, that such defined integrals \( \int_0^T X(s) \, dW(s) \) are unique \( P-a.s. \) on a complete probability space \( (\Omega, \mathcal{F}, P) \), and only the \( \mathcal{F}_t \)-Wiener process, which has its properties with respect to the probability measure \( P \), can be used. Furthermore its filtration \( \{\mathcal{F}_t\}_{t \geq 0} \) is right-continuous and complete.

As we saw in the example 1.2.1 we can choose a different evaluation point for summing up the simple integral. We have introduced the Itô choice. The second very used and sometimes advantageous choice is the one from Stratonovich i.e.

**Definition 1.2.5** For any \( T \in \mathbb{R}^+ \) and any function \( h \in L^2(\mathcal{F}, ([0, T])) \) we define the Stratonovich stochastic integral
\[
\mathcal{J}[h]_T^0(\omega) := \int_0^T h(u, \omega) \circ dW(u, \omega),
\]

as a mean-square limit of

\[ J_n[h]_T^0(\omega) = \sum_{j=1}^{n} h\left( \frac{1}{2}[t_j + t_{j+1}], \omega \right)[W(t_{j+1}) - W(t_j)](\omega) \quad (1.9) \]

for any partition

\[ \Delta_n([0,T]) = \{ 0 = t_0 < t_1 < \ldots < t_n = T \}, \]

and \( |\Delta_n| = \max_{1 \leq j \leq n} \{ t_{j+1} - t_j \} \to 0 \text{ as } n \to \infty \), where

\[ |\Delta_n([0,T])| = \max\{|t_{i+1} - t_i|, i = 1, \ldots, n\}. \]

### 1.3 Properties of stochastic integrals

In this section the basic features of both Itô and Stratonovich integrals are compared.

Firstly we should notice a very important corollary of the construction of Itô integral.

**Corollary 1.3.1** For all \( X(t) \in L_2(\mathcal{F}, \mathbb{R}^+) \) and \( 0 \leq s \leq t \leq T \) holds

\[ \mathbb{E}\left[ \left( \int_s^t X(u)dW(u) \right)^2 \right] = \mathbb{E}\left[ \int_s^t (X(u))^2 du \right] \quad (1.10) \]

**Proof:** The construction of the integral is enabled via this property, it is proved for simple integrals 1.6, and because of completeness 1.2.1 of the space there is one and only the one extension of this isometry which holds for all \( X(t) \in L_2(\mathcal{F}, \mathbb{R}^+) \).

\[ \square \]

**Theorem 1.3.1** For any \( X(t) \in L^2(\mathcal{F}, [0,T]) \) the Itô stochastic integral \( \mathcal{I}[H]_T^0 \in CM_2(\mathcal{F}, [0,T]) \).

**Proof:** Because Itô’s integral is defined as a \( m \)-limit of a simple Itô’s integral the martingale property is saved, as well as the \( \mathbb{E}[(\mathcal{I}_{[0,T]})^2] < \infty \), by remark 1.2.1. The only resolving problem is the time continuity in probability.

From 1.2.1 Let denote \( \mathcal{I}_n(t, \omega) \) simple integral with partition \( \Delta_n([0,T]) \),
(resp. \( \mathcal{I}_m'(t, \omega) \) simple integral with partition \( \Delta_m([0, T]) \)). Then we have by Doob’s inequality

\[
P \left[ \sup_{0 \leq t \leq T} |\mathcal{I}_n'(t, \omega) - \mathcal{I}_m'(t, \omega)| > \epsilon \right] \leq \frac{1}{\epsilon^2} \cdot \mathbb{E} \left[ |\mathcal{I}_n'(t, \omega) - \mathcal{I}_m'(t, \omega)|^2 \right] = \frac{1}{\epsilon^2} \cdot \mathbb{E} \left[ \int_0^T (K_n(t) - K_m(t))^2 dt \right] \to 0 \text{ as } m, n \to \infty.
\]

Hence we can choose a subsequence \( n_k \to \infty \) s.t.

\[
P \left[ \sup_{0 \leq t \leq T} |\mathcal{I}_{n_{k+1}}' - \mathcal{I}_{n_k}'| > 2^{-k} \right] < 2^{-k}.
\]

Then by the Borel-Cantelli lemma

\[
P \left[ \sup_{0 \leq t \leq T} |\mathcal{I}_{n_{k+1}}' - \mathcal{I}_{n_k}'| > 2^{-k} \text{ for infinitely many } k \right] = 0.
\]

Thus, for almost all \( \omega \) there exists \( k_1(\omega) \) such that

\[
|\mathcal{I}_{n_{k+1}}' - \mathcal{I}_{n_k}'| \leq 2^{-k} \text{ for } k \geq k_1(\omega).
\]

Therefore, \( I_{k_1}(t, \omega) \) is uniformly convergent for \( t \in [0, T] \), for almost all \( \omega \) and hence is \( P - a.s. \) time continuous for \( t \in [0, T] \). The limit in \( m \) i.e. the integral \( \mathcal{I}(t, \omega) \) is the same except the zero-sets. And zero-sets are included in \( F_i \) since it is complete as we expected.

\[ \square \]

Now, using this theorem 1.3.1, stopping times and stopping 1.1.3 theorem we can weaken the assumption

\[
\mathbb{E} \left[ \int_0^T (X(u))^2 du \right] < \infty, \ \forall T \in \mathbb{R}^+
\]

to weaker one

\[
P \left[ \int_0^T (X(u))^2 du < \infty \right] = 1, \ \forall T \in \mathbb{R}^+.
\]

Hence we can extend the set of integrable processes \( L_2(\mathcal{F}, [0, T]) \) to

**Definition 1.3.1 (Extended integrable processes)** Let \( L_2^*(\mathcal{F}, [0, T]) \) be a space of processes \( X(t) \) such as
(i) $X(t)$ is $\mathcal{F}_t-$progressive,

(ii) $P[\int_0^T X(t)^2 dt < \infty] = 1$ for all $t \in \mathbb{R}^+$. 

It shall be done as follows, suppose $X(t)$ to be from space of extended integrable functions $L^2_2(\mathcal{F}, [0, T])$, then we can build a sequence $\tau_n(\omega) \nearrow \infty$ as $n \to \infty$ such as

$$\tau_n(\omega) := n \land \inf \left\{ 0 \leq t < \infty; \int_0^t X(s, \omega) ds \geq n \right\}$$

obviously sequence $\{\tau_n\}_{n \in \mathbb{N}}$ is nondecreasing and holds

$$X(t)I_{\{t \leq \tau_n(\omega)\}} \in L_2(\mathcal{F}, [0, t \land \tau_n(\omega)])$$

Thus using 1.3.1 and 1.1.3

$$\left( \int_0^t X(s)I_{\{s \leq \tau_n\}}(s)dW(s) \right)(\omega) = \left\{ \begin{array}{ll}
\int_0^t X(s)dW(s)(\omega) & \text{for } t < \tau(\omega), \\
\int_0^{\tau_n} X(s)dW(s)(\omega) & \text{for } t \geq \tau(\omega).
\end{array} \right.$$ 

is stopped process $\in CM_2(\mathcal{F}, [0, \tau_n])$ for all $n \in \mathbb{N}$. We can denote

$$(I[X]_0^t)^{\tau_n} = \int_0^{t \land \tau_n} X(s)dW(s) = \int_0^t X(s)I_{\{s \leq \tau_n\}}(s)dW(s).$$

Therefore $\{\tau_n\}_{n \in \mathbb{N}}$ is a localization sequence and letting $\tau_n(\omega) \nearrow \infty$ as $n \to \infty$ we obtain a $P-a.s.-$continuous $\mathcal{F}_t-$local martingale. Let state the definition.

**Definition 1.3.2** Let $X(t) \in L^2_2(\mathcal{F}, [0, \infty))$ and $W(t), t \geq 0$ be $\mathcal{F}_t-$Wiener process. Then we define the Itô stochastic integral $\mathcal{I}[X]_0^t$ as a process

$$\mathcal{I}[X]_0^t := \int_0^t X(u)dW(u) \in CM_2(\mathcal{F}, [0, \infty)).$$

Given by its localization sequence $\tau_n, n \in \mathbb{N}$ with prescription

$$\left( \mathcal{I}[X]_0^t \right)^{\tau_n} := \int_0^{t \land \tau_n} X(u)dW(u) := \int_0^t X(s)I_{\{s \leq \tau_n\}}(s)dW(s), \quad t \in [0, \tau_n],$$

where

$$\left( \mathcal{I}[X]_0^t \right)^{\tau_n} \in CM_2(\mathcal{F}, [0, \tau_n]) \quad \text{for all } n \in \mathbb{N}.$$
Theorem 1.3.2 Let $G(t), H(t) \in L^2_2(\mathcal{F}, \mathbb{R}^+)$, let $0 \leq S \leq U \leq T$ and $c \in \mathbb{R}$. Then

(i) $\int_T^S G(t) dW(t) = \int_S^U G(t) dW(t) + \int_U^T G(t) dW(t) \quad P - a.s.,$

(ii) $\int_S^T (c \cdot G(t) + H(t)) dW(t) = c \int_S^T G(t) dW(t) + \int_S^T H(t) dW(t),$

(iii) $\int_S^T G(t) \circ dW(t) = \int_S^U G(t) \circ dW(t) + \int_U^T G(t) dW(t) \quad P - a.s.,$

(iv) $\int_S^T (c \cdot G(t) + H(t)) \circ dW(t) = c \int_S^T G(t) \circ dW(t) + \int_S^T H(t) \circ dW(t).$

Proof: This properties clearly holds for simple integral, so by taking a limits we obtain this for all $G(t), H(t) \in L^2_2(\mathcal{F}, \mathbb{R}^+)$. □

Theorem 1.3.3 Let $H(t, \omega) \in L^2_2(\mathcal{F}, \mathbb{R}^+)$ and $\mathcal{F}_t$ be complete filtration then $I[H]_0^t$ and $\mathcal{J}[H]_0^t$ are $\mathcal{F}, [0, T] -$ adapted processes.

Proof: Simple integrals $I_0^t$ are $\mathcal{F}, [0, T] -$ adapted as stated in remark 1.2.1, $I_0^t$ are mean-square limits so they can differ just on 0-sets. As we supposed the $\mathcal{F}_t$ filtration to be complete it completes our proof. Similarly for $J_0^t$. □

Definition 1.3.3 Let $I(\mathcal{F}, [0, \infty), W)$ denote a group of processes such as

$$X(t) := X(0) + \int_0^t a(s, \omega) ds + \int_0^t \sigma(s, \omega) dW(s), \quad t \geq 0,$$

where $X(0)$ is $\mathcal{F}_0 -$ adapted $\mathbb{E}[X^2(0)] < \infty$ and functions $a(t, \omega)$ and $b(t, \omega)$ are $\mathcal{F}_t -$ progressive satisfying $\mathbb{P}[\int_0^t |a(s, \omega)| ds < \infty] = 1$ for all $t \geq 0$ and finally function $\mathbb{P}[\int_0^t b^2(s, \omega) ds < \infty] = 1$ for all $t \geq 0$. Such processes are called Itô processes.

Now we shall introduce a Itô formula for Itô processes. This is one of the most common formula in stochastic calculus.
Theorem 1.3.4 Let $X(t)$ be an Itô process given by
\[ dX(t) = u(t, \omega) dt + v(t, \omega) dW(t). \]
Let $f(t, x) \in \mathcal{C}^2([0, \infty) \times \mathbb{R})$ (i.e. twice continuously differentiable on $(\mathbb{R}^+ \times \mathbb{R})$). Then
\[ Y(t) := f(t, X(t)) \]
is again an Itô process, and
\[ dY(t)(t, \omega) = \frac{\partial f}{\partial t}(t, X(t)) dt + \frac{\partial f}{\partial x}(t, X(t)) dX(t) + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(t, X(t))(dX(t))^2, \tag{1.12} \]
where $(dX(t))^2 = dX(t) \cdot dX(t)$ is computed according to the rules
\[ dt \cdot dt = dt \cdot W(t) = W(t) \cdot dt = 0, \quad dW(t) \cdot dW(t) = dt. \]
Proof: The theorem is proved in [3] as Theorem 4.1.2.

Theorem 1.3.5 Suppose $f(t, \omega) = f(W(t, \omega)) \in L_2(\mathcal{F}, \mathbb{R}^+)$ and $f(x) \in \mathcal{C}(\mathbb{R})$ then
\[ \int_0^T f(W(t)) \circ dW(t) = \int_0^T f(W(t))dW(t) + \frac{1}{2} \int_0^T \frac{\partial f}{\partial x}(W(t)) dt \tag{1.13} \]
Proof: By Taylor formula we get
\[ f(W(t_{j+1})) = f(W(t_j)) + \frac{\partial f}{\partial x}(W(t_j))(W(t_{j+1}) - W(t_j)) + \text{higher order terms}, \]
for some $0 \leq t_j \leq t_{j+1} \leq T$ and denote $W(t_{j+1}) - W(t_j) = \Delta W_{j+1}$, thus the sum 1.9 will appear like so
\[ \bar{S}_n = \sum_{j=1}^n \frac{1}{2} \left( f(W(t_j)) + f(W(t_{j+1})) \right) \Delta W_{j+1} \]
\[ = \sum_{j=1}^n \frac{1}{2} f(W(t_j)) \Delta W_{j+1} + \frac{1}{2} \left( f(W(t_j)) + \frac{\partial f}{\partial x}(W(t_j)) \Delta W_{j+1} \right) \Delta W_{j+1} \]
\[ = \sum_{j=1}^n f(W(t_j)) \Delta W_{j+1} + \frac{1}{2} \frac{\partial f}{\partial x}(W(t_j))(\Delta W_{j+1})^2 + \text{higher order terms}, \]
\[ \rightarrow \int_0^T f(W(t))dW(t) + \frac{1}{2} \int_0^T \frac{\partial f}{\partial x}(W(t)) dt \]
in the mean-square sense.
Theorem 1.3.6 Suppose \( f(t, \omega) = f(W(t, \omega)) \in L_2(\mathcal{F}, \mathbb{R}^+) \), and \( F \) be prime function to \( f \) i.e. \( \frac{\partial F(x)}{\partial x} = f(x) \) and \( f(x) \in C(\mathbb{R}) \) then

\[
\int_0^T f(W(t)) \circ dW(t) = F(W(T)) - F(W(0)). \tag{1.14}
\]

Proof: Applying Itô formula (1.12) to \( Y(t) = F(W(t)) \) we receive

\[
F(W(T)) = F(W(0)) + 0 \cdot dt + \int_0^T f(W(t))dW(t) + \frac{1}{2} \int_0^T \frac{\partial f}{\partial x}(W(t))dt
\]

now we use (1.13) and obtain the statement of the theorem. 

Thus, the rules of classical calculus are saved, which is the biggest advantage of using Stratonovich integrals.

Remark 1.3.1 Let \( X(t) \in L_2(\mathcal{F}, [0, T]) \) then

i) \( \int_s^t X(u)dW(u) = 0 \),

ii) \( \int_s^t X(u) \circ dW(u) = \mathbb{E}\left[\frac{1}{2} \int_s^t \frac{\partial f}{\partial x}(W(t))dt\right] \).

Proof: The first statement clearly holds for simple Itô integral, let \( \Delta u ([0, T]) \) be some partition denote \( \Delta W_i = W(t_{i+1}) - W(t_i) \) then

\[
\mathbb{E}\left[\sum_i f(t_i)\Delta W_i\right] = \sum_i \mathbb{E}[f(t_i)]\mathbb{E}[\Delta W_i] = 0
\]

since for \( \mathcal{F}_t \)-Wiener process \( \Delta W_i \) is independent of \( \mathcal{F}_t \). Thus the \( m(\cdot, \cdot) \)-limit has no influence, since we assume a complete filtration.

From the relation formula between Itô and Stratonovich integrals 1.13, we obtain the second statement.

Remark 1.3.2 Let \( X(t) \in L_2(\mathcal{F}, [0, T]) \) then

\[
\int_s^t X(u) \circ dW(u) \text{ is not a martingale.}
\]
Proof: Suppose $0 \leq s \leq r \leq t \leq T$ and $\mathcal{J}[H]^t_s = \int_s^t X(u) \circ dW(u)$ is a martingale, then

$$0 = \mathbb{E}[\mathcal{J}[H]^t_s - \mathcal{J}[H]^r_s | \mathcal{F}_r] = \mathbb{E}[\mathcal{J}[H]^t_r | \mathcal{F}_r],$$

since $\mathbb{E}[\mathbb{E}[\mathcal{J}[H]^t_r | \mathcal{F}_r]] = \mathbb{E}[\mathcal{J}[H]^t_r] \neq 0$ we obtain a contradiction with remark 1.3.1.
Chapter 2

Numerical stochastic analysis

The cornerstones of numerical analysis will be explained in this chapter. We shall start with the theorem of existence and uniqueness for the stochastic differential equation both Itô and Stratonovich. Further we will introduce a stochastic counterpart of the Taylor-expansion, which is the most important tool in deriving the numerical methods. Finally, chosen algorithms for evaluating of multiple stochastic integrals will be shown.

2.1 Stochastic differential equations

We will distinguish between Itô

\[ dX(t) = a(t, X(t))dt + b(t, X(t))dW(t), \quad X(t_0) = Z \quad (2.1) \]

and Stratonovich stochastic differential equation (SDE)

\[ X(t) = g(t, X(t))dt + b(t, X(t)) \circ dW(t), \quad X(t_0) = Z, \quad (2.2) \]

which differs in the stochastic integrals in the last term. With different choices of stochastic integral we would in general obtain a different solutions, even though they have the same coefficients.

In previous sections we defined the Stratonovich integral only for integrands of the form \( f(t, \omega) \) and \( h(W(t, \omega)) \). In this section we will extend this for integrands of the form \( h(t, X(t, \omega)) \) for a function \( h = h(t, x) \) and a diffusion process \( X(t) \). For our purposes we will restrict just on the solutions of stochastic differential equations. Before we redefine the Stratonovich integral we will give an explanation of such an enlargement.
Firstly, we will state the theorem about existence and uniqueness of an Itô stochastic differential equation. We are looking for a $m-$dimensional Itô process and having $d-$dimensional Wiener process $W(t) = (W^1(t),\ldots,W^d(t)).$

**Theorem 2.1.1** Let $T > 0$ and $a(\cdot, \cdot) : [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}^m,$ $b(\cdot, \cdot) : [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}^{m \times d}$ be jointly $(\mathcal{B}([0, T]) \times \mathcal{B}(\mathbb{R}^m)) -$measurable functions satisfying the linearity growth condition

$$||a(t, x)|| + ||b(t, x)|| \leq C(1 + ||x||); \quad x \in \mathbb{R}^m, t \in [0, T] \tag{2.3}$$

for some constant $C,$ where $||\sigma||^2 = \sum |\sigma_{ij}|^2$ and the Lipschitz condition

$$||a(t, x) - a(t, y)|| + ||b(t, x) - b(t, y)|| \leq D||x - y||; \quad x, y \in \mathbb{R}^m, t \in [0, T] \tag{2.4}$$

where $D$ is again some constant.

Further let $Z$ be a random variable which is independent of the $\sigma-$algebra $\mathcal{F}_\infty$ generated by $W(s, \cdot), s \geq 0$ and such that

$$\mathbb{E}[|Z|^2] < \infty.$$

Then the stochastic differential equation

$$dX(t) = a(t, X(t))dt + b(t, X(t))dW(t), \quad 0 \leq t \leq T \tag{2.5}$$

with the initial condition

$$X_0 = Z \tag{2.6}$$

has unique $P -$a.s. time continuous solution $X(t)(\omega)$ with the property that $X(t)(\omega)$ is adapted to the filtration $\mathcal{F}(t)^Z$ generated by $Z$ and $W(s, \cdot), s \geq 0$ and

$$\mathbb{E}\left[\int_0^T |X(t)|^2 dt\right] < \infty.$$

**Proof:** The theorem is proved in [3] as Theorem 5.2.1. 

Hence we are also getting the existence and uniqueness of the corresponding Stratonovich equation.

**Remark 2.1.1** We should stress that the conditions of theorem 2.1.1 above ensure that the both integrals and are meaningful.

Denote functions $e(t, \omega) = \sqrt{a(t, X(t))}, f(t, \omega) = b(t, X(t))$ then under the notation from previous chapter $e, f \in L^2(\mathcal{F}, [0, T]),$ i.e. $e(t, \omega)$ and $f(t, \omega)$ satisfy
\((i.)\) \(e(t, \omega)\) and \(f(t, \omega)\) must be \((B([0, T]) \times \mathcal{F}_t)\)-measurability

\((ii.)\) \(\mathbb{E}[^t_0 f^2(s)ds] < \infty\) for all \(t \in [0, T]\).

Proof: From assumptions of 2.1.1 \(X(t, \omega)\) is dependent only on \(X(0) \in \mathcal{F}_0\) and on \(\mathcal{F}_t\)-Wiener process, hence

\[
X(t, \cdot) : (\Omega, \mathcal{F}) \to (\mathbb{R}^m, B(\mathbb{R}^m)),
\]

and next assumption of 2.1.1 gives

\[
b(t, x) : \left([0, T]) \times \mathbb{R}^m, (B([0, T]) \times B(\mathbb{R}^m))\right) \to (\mathbb{R}^m, B(\mathbb{R}^m)).
\]

Since

\[
f(t, \omega) := b(t, X(t))
\]

is compounded map of the measurable maps, then the map \(f(t, \omega)\) is \((B([0, T]) \times \mathcal{F}_t)\)-measurable. Similarly for \(e(t, \omega)\).

The second property is satisfied from the linearity condition as follows

\[
\int_0^t b^2(t, x)ds \leq C(1 + 2||x|| + ||X||^2) \int_0^t 1ds \leq K \cdot T < \infty,
\]

holds \(P-a.s.\) for all \(t \in [0, T]\) and hence \(\mathbb{E}[\int_0^t f^2(s)ds] < \infty\) for all \(t \in [0, T]\).

\(\square\)

Second we state that the solutions are diffusion processes.

**Theorem 2.1.2** Let \(W(t)\) be a \(\mathcal{F}_t\)-Wiener process, \(a(t, \omega)\) and \(b(t, \omega)\) are continuous functions satisfying the linearity growth and Lipschitz condition and \(Z \in \mathcal{F}_0\) such that \(\mathbb{E}[|Z|^2] < \infty\) then the solution of (2.1) with initial condition \(X_0 = Z\) is a diffusion process on \([0, T]\) with drift \(a(t, x)\) and diffusion coefficient \(b(t, x)\).

Proof: The theorem is proved in [4] as Theorem 4.6.1

\(\square\)

Now we are ready to redefine the Stratonovich stochastic integral for a solution of stochastic differential equation in the sense of 2.1.1.
Definition 2.1.1 Suppose \( f(t, \omega) := h(t, X(t, \omega)) \in L_2(F_{[0,T]}), \) for any \( h(t, x) \in C^1([0,T] \times \mathbb{R}^m) \) and diffusion process \( X(t) \) then

\[
\int_s^t h(u, X(u)) \circ dW(u)
\]
is the mean-square limit of the sums

\[
S_n(\omega) = \sum_{j=1}^n h(t_j, \frac{1}{2}(X(t_j) + X(t_{j+1}))) [W(t_{j+1}) - W(t_j)] \tag{2.7}
\]

for some partition

\[
\Delta_n([s,t]) = \{ s = t_1 < t_2 < \ldots < t_n = t \},
\]
with \( |\Delta_n| \to 0 \) as \( n \to \infty \).

Under these conditions we can rewrite the relation between Itô and Stratonovich integrals also for diffusion processes.

Theorem 2.1.3 Let \( X(t) \) be a diffusion process such as

\[
dX(t) = a(t, \omega)dt + b(t, \omega)dW(t),
\]
then suppose the function \( h(t, x) \in C^1([0,T] \times \mathbb{R}^m) \), and \( f(t, \omega) := h(t, X(t, \omega)) \in L_2(F_{[0,T]}), \) then

\[
\int_s^t h(u, X(u)) \circ dW(u) = \int_s^t h(u, X(u))dW(u) + \frac{1}{2} b(t, X(t)) \frac{\partial h}{\partial x}(t, X(t))dt. \tag{2.8}
\]


This relation 2.8 has important corollaries for solving of SDE. Because it is says how the Itô and Stratonovich SDEs corresponds.

Corollary 2.1.1 Let assume we have some Itô SDE of the form (2.1) with initial condition \( X(t_0) = Z \) satisfying conditions of the theorem 2.1.1 Then \( X(t) \) is also a solution of a corresponding Stratonovich SDE of the form (2.2) with a modified drift \( a \) defined by

\[
a(t, x) = a(t, x) - \frac{1}{2} b(t, x) \frac{\partial b}{\partial x}(t, x).
\]
Proof: It follows immediately from the equation (2.8) just by putting $h(t, x) = b(t, x)$. □

The second important corollary is the Stratonovich analogy to the Itô formula.

**Corollary 2.1.2** Let $Y(t) = u(t, X(t))$, where $X(t)$ is a solution of the Stratonovich SDE (2.2), (i.e. by corollary 2.1.1 $X(t)$ is also solution of corresponding Itô SDE (2.1)) and suppose $u(t, x) \in C^2(\mathbb{R}^+ \times \mathbb{R}^m)$, then

$$dY(t) = \left( \frac{\partial u}{\partial t} + a \frac{\partial u}{\partial x} \right) dt + b \frac{\partial u}{\partial x} \diamond dW(t)$$  \hspace{2cm} (2.9)

**Proof:** Denote $\frac{\partial u}{\partial t} = u_t$, $\frac{\partial u}{\partial x} = u_x$, resp. $\frac{\partial^2 u}{\partial x^2}(t, x) = u_{xx}$ etc. and using Itô formula on $Y(t)$ we obtain

$$dY = u_t dt + u_x dX + \frac{1}{2} u_{xx} \langle X(t) \rangle$$

$$= \left( u_t + au_x \right) dt + bu_x dW(t) + \frac{1}{2} b^2 u_{xx} dt,$$

where all functions are evaluated at point $(t, X(t))$. Now let $h = bu_x$ then using (2.8) for $h$ we will proof the statement of the remark

$$dY = \left( u(t) + au_x \right) dt + \left( h \circ dW(t) - \frac{1}{2} b h_x dt \right) + \frac{1}{2} b^2 u_{xx} dt$$

$$= \left( u(t) + au_x - \frac{1}{2} b h_x u_x \right) dt + h \circ dW(t).$$

□

The last remark shows that Stratonovich integral saves rules of classical calculi also for diffusion processes $X(t)$ obtained from solution of some SDE.

**Remark 2.1.2** Let $h \in C^1[0, T]$ a and $H$ be a prime function to $h$ i.e. $H' = h$ and $X(t)$ be a solution of Stratonovich SDE (2.2) then for any $0 \leq s \leq t \leq T$ holds

$$H(X(t)) - H(X(s)) = \int_s^t h(X(u)) \circ dX(u).$$  \hspace{2cm} (2.10)
Proof: Using the Stratonovich counterpart of Itô formula (2.9) we see that
\[ dH(X(t)) = h(X(u))\left( a(u, X(u))dt + b(u, X(u)) \circ dW(u) \right), \]
rewritten into the integral form we obtain
\[ H(X(t)) = H(X(s)) + \int_s^t h(X(u)) \circ dX(u). \]

Thus we are provided with derivation of \( Y(t) \) both in Itô and in Stratonovich sense for any \( Y(t) := f(t, X(t)) \), where \( f \in L_2(\mathcal{F}_{[0,T]}) \) and \( X(t) \) is Itô process. Furthermore, we can always switch between Itô SDE (2.1) and Stratonovich (2.2) and use their strength either martingale property for Itô 1.3.1 or normal calculi rules for Stratonovich 2.1.2.

Following examples demonstrates usage of this properties and will be also studied by means of numerical analysis in the next chapter.

Example 2.1.1

We should solve Itô SDE:
\[ dX(t) = \frac{1}{3}X^{\frac{1}{3}}(t)dt + X^{\frac{2}{3}}(t)dW(t) \]

It is not so convenient to solve it in this form, but if we split into the corresponding Stratonovich SDE by corollary 2.1.1, we obtain
\[ dX(t) = \frac{1}{3}X^{\frac{1}{3}}(t)dt + X^{\frac{2}{3}}(t) \circ dW(t) - \frac{1}{2}X^{\frac{2}{3}}(t)\frac{2}{3}X^{-\frac{1}{3}}(t)dt = X^{\frac{2}{3}}(t) \circ dW(t) \]

Now we can separate variables and use the normal calculus for \( \int f(W(t)) \circ dW(t) \) by 1.14 thus
\[ \left( 3X(t)^{\frac{1}{3}} \right)' = \int 1 \circ dW(t), \]
and hence
\[ X(t) = \left( \frac{W(t)}{3} + C \right)^3, \]
where \( C = X^{\frac{1}{3}}(t_0) \).

Which is solution for both Stratonovich SDE and for corresponding Itô SDE. In addition, the theorem for the solution of linear stochastic differential equation is stated.
Theorem 2.1.4 Let $W(t)$ be a $d$-dimensional Wiener process and $X(t)$ be a solution of a $m$-dimensional linear stochastic differential equation of the form

$$dX(t) = \left( A(t)X(t) + a(t) \right) dt + \sum_{l=1}^{d} \left( B^l(t)X(t) + b^l(t) \right) dW^l(t), \quad (2.11)$$

where $A(t), B^1(t), \ldots, B^d(t)$ are $m \times m$ matrix functions and $a(t), b^1(t), \ldots, b^d(t)$ are $m$-dimensional vector functions. Then

$$X(t) = \Phi_{t,t_0}(X(t_0) + \int_{t_0}^{t} \Phi_{s,t_0}^{-1} \left( a(s) - \sum_{l=1}^{d} B^l(s)b^l(s) \right) ds + \sum_{l=1}^{d} \int_{t_0}^{t} \Phi_{s,t_0}^{-1} b^l(s) dW^l(s)), \quad (2.12)$$

where $\Phi_{t,t_0}$ is the $m \times m$ fundamental matrix satisfying $\Phi_{t_0,t_0} = I$ and the homogenous matrix stochastic differential equation

$$d\Phi_{t,t_0} = A(t)\Phi_{t,t_0} dt + \sum_{l=1}^{m} B^l(t)\Phi_{t,t_0} dW^l(t), \quad (2.13)$$

which we interpret column vector by column vector as vector stochastic differential equation. We cannot generally solve the (2.12) explicitly for its fundamental solution, even when all of the matrices are constant matrices. In a case that $A, B^1, \ldots, B^d$ are constants and commute i.e.

$$AB^l = B^lA \quad \text{and} \quad B^kB^l = B^lB^k,$$

for all $k, l = 1, \ldots, d$ then we obtain the following explicit expression for the fundamental matrix solution

$$\Phi_{t,t_0} = \exp \left\{ \left( A - \frac{1}{2} \sum_{l=1}^{d} (B^l)^2 \right) (t - t_0) + \sum_{l=1}^{d} B^l \left( W^l(t) - W^l(t_0) \right) \right\}. \quad (2.14)$$

2.2 Stochastic Taylor expansion

In this section the Itô-Taylor and Stratonovich-Taylor expansion and their properties will be described.

As in standard numerical analysis there are many valuation algorithms and each is more or less appropriate for some class of differential equations. But all methods are based on some approximation of the studied function. In the deterministic case the most used approximation is here the Taylor expansion. It is similar in stochastic numerical analysis. But here the stochastic counterpart of the Taylor expansion depends on the stochastic integral we use. Hence, we have Itô-Taylor and Stratonovich-Taylor expansion. These are not always the same, however we will see that in some special set they have the same order of convergence. We will see it later in this section.

For purposes of this chapter we will rewrite the $m$-dimensional Itô stochastic differential equation (2.1) per complement

$$dX^i(t) = a^i(t, X(t))dt + \sum_{j=1}^d b^{i,j}(t, X(t))dW^j(t), \quad i = 1, \ldots, m; \quad (2.15)$$

where $a^i : \mathbb{R}^+ \times \mathbb{R}^m \to \mathbb{R}$ and $b^{i,j} : \mathbb{R}^+ \times \mathbb{R}^m \to \mathbb{R}$ for each $i = 1, \ldots, m$ and each $j = 1, \ldots, d$ Borel measurable functions satisfying the linear growth and Lipschitz condition from theorem 2.1.1 and $W^j(t)$ are components of $d$-dimensional Wiener process. Before we could state the general formula of Itô-Taylor, resp. Stratonovich-Taylor expansion, we need to build up a notation structure.

**Definition 2.2.1** Suppose $X(t)$ to be a solution of $m$-dimensional Itô or Stratonovich SDE with $d$-dimensional Wiener process.

Then a multi-index is a row vector

$$\alpha = (j_1, \ldots, j_k), \text{ where } j_i \in \{0, 1, \ldots, d\} \text{ for all } i.$$  

A length of multi-index

$$l(\alpha) := k \in \{1, 2, \ldots\},$$

we denote by $(v)$ the case if $\alpha = \emptyset$, obviously $l(\emptyset) = 0$. Further define $n(\alpha)$ number of components of $\alpha$ which are equal to 0. And finally the set of all multi-indices

$$\mathcal{M} = \left\{(j_1, \ldots, j_k) : j_i \in \{0, 1, \ldots, d\}, i \in \{1, \ldots, k\}, \text{ for } k \in \mathbb{N}\right\} \cup \{v\}.$$
For any $\alpha \in \mathcal{M}$ such as $l(\alpha) \geq 1$ we write $-\alpha$ and $\alpha-$ for the multi-index in $\mathcal{M}$ obtained by deleting first resp. last component of $\alpha$.

**Definition 2.2.2** Suppose we have a $m-$dimensional Itô SDE of the form from equation (2.15) and $d-$dimensional $W(t)$. Then for any twice differentiable function $f : \mathbb{R}^+ \times \mathbb{R}^m \to \mathbb{R}$ we denote the Itô drift operator as

$$L^0 f = \frac{\partial f}{\partial t} + \sum_{k=1}^{m} a^k \frac{\partial f}{\partial x^k} + \frac{1}{2} \sum_{k,l=1}^{m} \sum_{j=1}^{d} b^{k,j} b^{l,j} \frac{\partial^2 f}{\partial x^k \partial x^l}$$

and for $j \in \{1, \ldots, d\}$ we denote the Itô diffusion operator as

$$L^j f = \sum_{k=1}^{m} b^{k,j} \frac{\partial f}{\partial x^k},$$

where functions $a^k, b^{k,j}, j = 1, \ldots, d, k = 1, \ldots, m$ were introduced in (2.15). Finally for each $\alpha = (j_1, \ldots, j_k) \in \mathcal{M}$ and function $f \in C^h(\mathbb{R}^+ \times \mathbb{R}^m \to \mathbb{R})$ with $h = l(\alpha) + n(\alpha)$ we define recursively following notation

$$f_{\alpha} = \begin{cases} f & \text{if } k = 0, \\ L^k f_{\alpha-} & \text{if } k \geq 1. \end{cases}$$

Similarly for Stratonovich SDE

**Definition 2.2.3** For any function $f \in C^1(\mathbb{R}^+ \times \mathbb{R}^m, \mathbb{R})$ we denote the Stratonovich drift operator

$$L^0 f = \frac{\partial f}{\partial t} + \sum_{k=1}^{m} a^k \frac{\partial f}{\partial x^k}$$

and for $j \in \{1, \ldots, d\}$ we denote the Stratonovich diffusion operator

$$L^j f = \sum_{k=1}^{m} b^{k,j} \frac{\partial f}{\partial x^k},$$

where

$$a^k = a^k - \frac{1}{2} \sum_{k=1}^{d} L^k b^{i,k}, \quad k = 1, \ldots, m;$$
and $a^k, b^{k,i}, k = 1, \ldots, m, j = 1, \ldots, d$ are introduced in (2.15). Finally for each $\alpha = (j_1, \ldots, j_k)$ and function $f \in C^h(\mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R})$ with $h = l(\alpha)$ we define recursively following notation

\[
 f_\alpha = \begin{cases} 
  f & \text{if } k = 0, \\
  L^{j_k}f_{\alpha} & \text{if } k \geq 1.
\end{cases}
\]

**Definition 2.2.4** Let $\mathcal{F}_t$ be $\sigma$–algebra generated by $d$–dimensional Wiener process $W(t) = (W^1(t), \ldots, W^d(t))$. Then the set of all $\mathcal{F}_t$–adapted right continuous process with left hand limits $f = \{f(t, \omega), t \geq 0\}$, satisfying

\[
 |f(t, \omega)| < \infty \quad P\text{–a.s. for each } t \geq 0,
\]

is denoted by $\mathcal{H}_v$, 

\[
 \int_{\rho}^{t} |f(s, \omega)| ds < \infty \quad P\text{–a.s. for each } t \geq 0,
\]

is denoted by $\mathcal{H}_0$, 

\[
 \int_{\rho}^{t} |f(s, \omega)|^2 ds < \infty \quad P\text{–a.s. for each } t \geq 0,
\]

is denoted by $\mathcal{H}_j$, $j = 1, \ldots, d$

Similarly for Stratonovich integral.

**Definition 2.2.5** Let $\{X(t), t \geq 0\}$ be a $m$–dimensional Itô process, satisfying the Stratonovich (2.2). Then the set of all functions $g : \mathbb{R}^+ \times \mathbb{R}^m \to \mathbb{R}$ for which $g(\cdot, X(\cdot)) \in \mathcal{H}_v$, resp. $g(\cdot, X(\cdot)) \in \mathcal{H}_0$ is denoted by $\mathcal{H}_v$, resp. $\mathcal{H}_1$. In addition the set of all functions $g : \mathbb{R}^+ \times \mathbb{R}^m \to \mathbb{R}$ for which $g(\cdot, X(\cdot)) \in \mathcal{H}_j$, $j = 1, \ldots, d$ such as

\[
 \int_{\rho}^{t} |L^j f(s, \omega)| ds < \infty \quad P\text{–a.s. for all } t \geq 0
\]

are denoted by $\mathcal{H}_j$.

**Definition 2.2.6** Let $\rho, \tau$ be two stopping times with

\[
 0 \leq \rho(\omega) \leq \tau(\omega) \leq T, \quad P\text{–a.s.}
\]

Then for a multi-index $\alpha = (j_1, \ldots, j_k) \in \mathcal{M}$ and $f \in \mathcal{H}_\alpha$ we define the multiple Itô integral $\mathcal{I}_\alpha[f(\cdot)]_{\rho, \tau}$ recursively by

\[
 \mathcal{I}_\alpha[f(\cdot)]_{\rho, \tau} = \begin{cases} 
  f(\tau) & \text{if } k = 0, \\
  \int_{\rho}^{\tau} \mathcal{I}_{\alpha-[f(\cdot)]_{\rho, s}} ds & \text{if } k \geq 1 \text{ and } j_k = 0, \\
  \int_{\rho}^{\tau} \mathcal{I}_{\alpha-[f(\cdot)]_{\rho, s}} dW^j & \text{if } k \geq 1 \text{ and } j_k > 0.
\end{cases}
\]
Where for $\alpha = (j_1, \ldots, j_k)$ with $k \geq 2$ we define recursively the set $\mathcal{H}_\alpha$ to be the totality of adapted right continuous processes $f = \{f(t), t \geq 0\}$ with left hand limits such that the integral process $\{I_{\alpha-}[f(\cdot)]_{\rho,t}, t \geq 0\}$ considered as function of $t$ satisfies

$$I_{\alpha-}[f(\cdot)]_{\rho,\cdot} \in \mathcal{H}_{j_k}.$$ 

At last denote $I_{\alpha}[1]_{0,t} = I_{\alpha}^{0,t}$ resp. if the borders are obvious we write just $I_{\alpha}$.

**Definition 2.2.7** Let $\rho, \tau$ be two stopping times with

$$0 \leq \rho(\omega) \leq \tau(\omega) \leq T, \quad P - a.s.$$ 

Then for a multi-index $\alpha = (j_1, \ldots, j_k) \in \mathcal{M}$ and $g \in \mathcal{H}_\alpha$, we define the multiple Stratonovich integral $\mathcal{J}_\alpha[g(\cdot, X(\cdot))]_{\rho,\tau}$ recursively by

$$\mathcal{J}_\alpha[g(\cdot, X(\cdot))]_{\rho,\tau} = \begin{cases} g(\tau, X(\tau)) & \text{if } k = 0 \\ \int_{\rho}^{\tau} \mathcal{J}_\alpha[g(\cdot, X(\cdot))]_{\rho,s} ds & \text{if } k \geq 1 \text{ and } j_k = 0 \\ \int_{\rho}^{\tau} \mathcal{J}_\alpha[g(\cdot, X(\cdot))]_{\rho,s} \circ dW_{j_k} & \text{if } k \geq 1 \text{ and } j_k > 0. \end{cases}$$

The set $\mathcal{H}_\alpha$ is also defined recursively as set of all functions $g : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$ for which

$$\mathcal{J}_\alpha[g(\cdot, X(\cdot))]_{\rho,\cdot} \in \mathcal{H}_{j_k}$$

and

$$\mathcal{J}_{(\alpha-)}[I_{\{j_k=j_k-1 \neq 0\}} g(\cdot, X(\cdot))]_{\rho,\cdot} \in \mathcal{H}_{j_k}.$$ 

Again denote $I_{\alpha}[1]_{0,t} = I_{\alpha}^{0,t}$ resp. if the borders are obvious we write just $\mathcal{J}_\alpha$.

**Example 2.2.1**

$$I_{(1,0,2,1)}[f(\cdot)]_{\rho,\tau} = I_{(1,0,2,1)}[f(\cdot)]_{\rho,\cdot} = \int_{\rho}^{\tau} \int_{\rho}^{u} \int_{\rho}^{s} f(t) dW_t^1 ds \, dW_u^2,$$

$$\mathcal{J}_{(1,0,2,1)}[f(\cdot)]_{\rho,\tau} = \mathcal{J}_{(1,0,2,1)}[f(\cdot)]_{\rho,\cdot} = \int_{\rho}^{\tau} \int_{\rho}^{u} \int_{\rho}^{s} g(s, X(s)) ds \, dW_u^2 \circ dW_r^1$$

$$\mathcal{H}_{(1,2)} = \{ f(t, \omega), t \geq 0 : \lim_{t \to s^+} f(t, \omega) = f(s, \omega); \lim_{t \to s^-} f(t, \omega); f \in \mathcal{H}_1, I_{(1)}[f(\cdot)]_{\rho,\cdot} \in \mathcal{H}_2 \}$$

Let introduce additional notation, which will simplify the writing.
Definition 2.2.8 We call a hierarchical set any $A \subset M$ such as

(i.) $A \neq \emptyset$,

(ii.) $\sup_{\alpha \in A} l(\alpha) < \infty$,

(iii.) $-\alpha \in A$ each $\alpha \in A \setminus \{v\}$.

In addition for any hierarchical set $A$ we define the remainder set as

$B(A) := \{\alpha \in M \setminus A : -\alpha \in A\}$.

Now we are ready to state the Itô-Taylor expansion formula.

Theorem 2.2.1 Let $X(t)$ be an Itô process as in (2.15) and $\rho, \tau$ be two stopping times with

$0 \leq \rho(\omega) \leq \tau(\omega) \leq T, \quad P-a.s.,$

let $A \subset M$ be a hierarchical set, and let $f : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$. Then the Itô-Taylor expansion is

$$f(\tau, X_\tau) = \sum_{\alpha \in A} I_\alpha[f_\alpha(\rho, X_\rho)]_{\rho,\tau} + \sum_{\alpha \in B(A)} I_\alpha[f_\alpha(\cdot, X_\cdot)]_{\rho,\tau}, \quad (2.16)$$

holds, provided all of the derivatives of $f, a$ and $b$ and all multiple Itô integrals in (2.16) exist.

Proof: The theorem is proved in [4] as Theorem 5.5.1.

Similarly we introduce the Stratonovich-Taylor formula.

Theorem 2.2.2 Let $X(t)$ be an Itô process as in (2.15) and $\rho, \tau$ be two stopping times with

$0 \leq \rho(\omega) \leq \tau(\omega) \leq T < \infty, \quad P-a.s.$

Let $A \subset M$ be a hierarchical set, and let $f : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$. Then the Stratonovich-Taylor expansion is

$$f(\tau, X_\tau) = \sum_{\alpha \in A} J_\alpha[f_\alpha(\rho, X_\rho)]_{\rho,\tau} + \sum_{\alpha \in B(A)} J_\alpha[f_\alpha(\cdot, X_\cdot)]_{\rho,\tau}, \quad (2.17)$$

holds, provided all of the derivatives of $f, a$ and $b$ and all multiple Stratonovich integrals in (2.17) exist.
Proof: The theorem is proved in [4] as Theorem 5.6.1. Hereby we obtained a very powerful tool which enables us to approximate any sufficiently smooth function in neighbourhood of a given point to any desired accuracy. The expansion contains a sum of multiple Itô resp. Stratonovich integrals with constant integrands with remainder which contains finite sum of multiple integrals depending on time. The next example shall show the difference between both expansions.

Example 2.2.2

Suppose a sufficient smooth function \( f : \mathbb{R} \rightarrow \mathbb{R} \) and a hierarchical set

\[ \mathcal{A} = \{(v), (1)\}, \text{ then } \mathcal{B}((A)) = \{(0), (0,1), (1,1)\} \]

is the remainder set and the Itô-Taylor expansion is

\[
f(t, X(t)) = f(\tau, X(\tau)) + b(\tau, X(\tau)) f'(\tau, X(\tau)) I_{(1)} + \]
\[
+ \mathcal{I}_{(0)} \left[ a(\cdot) f'(\cdot) + \frac{1}{2} b^2(\cdot) f''(\cdot) \right]_{\rho,\tau} + \mathcal{I}_{(0,1)} \left[ b(\cdot) \left( a(\cdot) f'(\cdot) + \frac{1}{2} b^2(\cdot) f''(\cdot) \right) \right]_{\rho,\tau},
\]

where for simplifying of the notation \( (\cdot) = (\cdot, X(\cdot)) \). Hence the function \( f \) must be from \( C^3(\mathbb{R}^+ \times \mathbb{R} \rightarrow \mathbb{R}) \) to determine the remainder of the approximation. And on the other hand the Stratonovich-Taylor expansion is

\[
f(t, X(t)) = f(\tau, X(\tau)) + b(\tau, X(\tau)) f(\tau, X(\tau)) J_{(1)} + \]
\[
+ \mathcal{J}_{(1)} \left[ a(\cdot) f'(\cdot) - \frac{1}{2} b(\cdot) b'(\cdot) f''(\cdot) \right]_{\rho,\tau} + \mathcal{J}_{(0,1)} \left[ b(\cdot) \left( a(\cdot) f'(\cdot) + \frac{1}{2} b^2(\cdot) f''(\cdot) \right) \right]_{\rho,\tau} + \mathcal{J}_{(1,1)} \left[ b(\cdot) \left( b(\cdot) f'(\cdot) \right) \right]_{\rho,\tau},
\]

where again \( (\cdot) = (\cdot, X(\cdot)) \). Thus the function \( f \) needs to be just from \( C^2(\mathbb{R}^+ \times \mathbb{R} \rightarrow \mathbb{R}) \).

\[\diamond\]

For both expansions the mean-square convergence criterion is proved in [4] (Proposition 5.9.1 for Itô-Taylor resp. Proposition 5.10.1 for Stratonovich-Taylor expansion).
For our purposes we just show that both expansions are equivalent but just on the same hierarchical set

$$\Lambda_k = \{ \alpha \in \mathcal{M} : l(\alpha) + n(\alpha) \leq k \}.$$ 

It will be demonstrated later by example.

Before we state the relation between multiple integrals.

**Theorem 2.2.3** Let $\rho$ and $\tau$ be two stopping times with

$$0 \leq \rho(\omega) \leq \tau(\omega) \leq T \quad \text{P-a.s.}$$

Then for any $\alpha = (j_1, \ldots, j_k) \in \mathcal{M}$ and a function $g \in \mathcal{H}_\alpha$ holds for $l(\alpha) = 0, \; l(\alpha) = 1$

$$\mathcal{J}_\nu[\mathcal{I}(\cdot, X(\cdot))]_{\rho,\tau}^{\rho,\nu} = \mathcal{I}(\cdot, X(\cdot))^{\rho,\nu}_{\rho,\tau} = g(\tau, X(\tau))$$

$$\mathcal{J}_{(j_1)}[\mathcal{I}(\cdot, X(\cdot))]_{\rho,\tau}^{\rho,\nu} = \mathcal{I}_{(j_1)}[\mathcal{I}(\cdot, X(\cdot))]_{\rho,\tau}^{\rho,\nu} + I_{\{j_1 \neq 0\}}[\frac{1}{2}\mathcal{J}_{j_1}(\cdot, X(\cdot))]_{\rho,\tau}^{\rho,\nu},$$

and for $l(\alpha) \geq 2$

$$\mathcal{J}_\alpha[\mathcal{I}(\cdot, X(\cdot))]_{\rho,\tau}^{\rho,\nu} = \mathcal{I}_{(j_k)}\left[\mathcal{J}_{\alpha-}[\mathcal{I}(\cdot, X(\cdot))]_{\rho,\tau}^{\rho,\nu}\right] +$$

$$+ I_{\{j_k = j_{k-1} \neq 0\}}[\mathcal{I}(\cdot, X(\cdot))_\rho_{\tau}] + I_{\{j_k - j_{k-1} \neq 0\}}[\frac{1}{2}\mathcal{J}_{(\alpha-)}[\mathcal{I}(\cdot, X(\cdot))]_{\rho,\tau}^{\rho,\nu}].$$

**Proof:** The theorem is proved in [4] as equation 2.34 in chapter 5 section 2.

Thus we can rewrite the most used relations.

**Remark 2.2.1**

$$\mathcal{J}_\alpha = \mathcal{I}_{\alpha}, \quad \text{where } l(\alpha) \in \{0, 1\},$$

$$\mathcal{J}_{\alpha} = \mathcal{I}_{\alpha} + \frac{1}{2}I_{\{j_1 = j_2 \neq 0\}}[\mathcal{I}(\cdot, X(\cdot))]_{\rho,\tau}^{\rho,\nu}, \quad \text{for } l(\alpha) = 2,$$

$$\mathcal{J}_{\alpha} = \mathcal{I}_{\alpha} + \frac{1}{2}\left(I_{\{j_1 = j_2 \neq 0\}}[\mathcal{I}(\cdot, X(\cdot))]_{\rho,\tau}^{\rho,\nu} + I_{\{j_2 = j_3 \neq 0\}}[\mathcal{I}(\cdot, X(\cdot))]_{\rho,\tau}^{\rho,\nu}\right) \quad \text{for } l(\alpha) \geq 2.$$ 

Hence, we can easily switch between multiple integrals and therefore compare the convergence on some hierarchical set.
Example 2.2.3

Let $W(t), t \geq 0$ be 1-dimensional Wiener process and $X(t), t \in [0, T]$ be an 1-dimensional Itô process such as $dX(t) = a(t, X(t)) dt + b(t, X(t)) dW(t)$, where functions $a : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$, and $b(\cdot, \cdot) : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ satisfy the conditions from theorem 2.1.1. Further let

$$X_k(t) = \sum_{\alpha \in \Lambda_k} I_{\alpha}[f_{\alpha}(t_0, X(t_0))]_{t_0, t}$$

be approximation of $X(t)$ as resulting from Itô-Taylor expansion, and

$$Z_k(t) = \sum_{\alpha \in \Lambda_k} J_{\alpha}[f_{\alpha}(t_0, X(t_0))]_{t_0, t}$$

be approximation resulting from Stratonovich-Taylor expansion. Put $f \equiv x$ and the hierarchical set

$$\Lambda_k = \{ \alpha \in \mathcal{M} : l(\alpha) + n(\alpha) \leq 3 \}.$$

Then for $k = 1$ we obtain

$$\Lambda_1 = \{(v), (1)\}$$

hence

$$X_1 = x + bI_{(1)}$$
$$Z_1 = x + bJ_{(1)},$$

for $k = 2$

$$\Lambda_2 = \{(v), (1), (0), (1, 1)\}$$
$$X_2 = x + bI_{(1)} + aI_{(0)} + bb'I_{(1, 1)}$$
$$Z_2 = x + bJ_{(1)} + aJ_{(0)} + bb'J_{(1, 1)}$$

and for $k = 3$

$$\Lambda_3 = \{v, (1), (0), (1, 1), (0, 1), (1, 0), (1, 1, 1)\},$$

$$X_3 = x + bI_{(1)} + aI_{(0)} + bb'I_{(1, 1)} +$$
$$+ ba'I_{(1, 0)} + (ab' + \frac{1}{2} b^2 b'')I_{(0, 1)} + b(b'^2 + bb'')I_{(1, 1, 1)}$$
$$Z_3 = x + bJ_{(1)} + aJ_{(0)} + bb'J_{(1, 1)} + ab'J_{(0, 1)} + ba'J_{(1, 0)} + b(b'^2 + bb'')J_{(1, 1, 1)}.$$
Using relations from remark 2.2.1 and writing $a = a - \frac{1}{2} bb'$ we obtain for $k = 1$

$$X_1 = x + bI_{(1)} = x + bJ_{(1)} = Z_1,$$

for $k = 2$

$$X_2 - Z_2 = aI_{(0)} - (a - \frac{1}{2} bb')J_{(0)} + bb'I_{(1,1)} - bb'J_{(1,1)}$$

$$= \frac{1}{2}bb'I_{(0)} + bb'I_{(1,1)} - bb'(I_{(1,1)} + \frac{1}{2}J_{(0)}) = 0$$

and for $k = 3$

$$X_3 - Z_3 = ba'I_{(1,0)} - ba'J_{(1,0)} + (ab' + \frac{1}{2}b^2b'')I_{(0,1)} - ab'J_{(0,1)} +$$

$$+ b(b^2 + bb'')(I_{(1,1,1)} - J_{(1,1,1)}),$$

where

$$ba'I_{(1,0)} - ba'J_{(1,0)} = ba'I_{(1,0)} - b(a - \frac{1}{2} bb')I_{(1,0)} = I_{(1,0)} - \frac{1}{2}(bb'' + (b')^2)$$

$$(ab' + \frac{1}{2}b^2b'')I_{(0,1)} - ab'J_{(0,1)} = (ab' + \frac{1}{2}b^2b'')I_{(0,1)} - (a - \frac{1}{2} bb')b'I_{(0,1)} =$$

$$= \frac{1}{2}(b^2b'' + b(b')^2)I_{(0,1)}$$

$$b(b^2 + bb'')(I_{(1,1,1)} - J_{(1,1,1)}) = b(b^2 + bb'')I_{(1,1,1)} -$$

$$+ \frac{1}{2}(I_{(0,1)} + I_{(1,0)} - I_{(1,0)}) = b(b^2 + bb'')\left[I_{(0,1)} + I_{(0,1)} - (I_{(0,1)} + I_{(1,0)})\right]$$

hence, we find that

$$X_3 - Z_3 = \frac{1}{2}(b^2b'' + b(b')^2)\left[I_{(0,1)} + I_{(0,1)} - (I_{(0,1)} + I_{(1,0)})\right] = 0.$$

One the other hand, it is easy to find a hierarchial set, where Itô-Taylor and Stratonovich-Taylor expansion do not merge, e.g.

$$\Gamma_k := \{\alpha \in \mathcal{M} : l(\alpha) \leq k\}$$

and put

$$X_{\Gamma_k} = \sum_{\alpha \in \Gamma_k} I_\alpha[f_{\alpha}(t_0, X(t_0))]_{t_0,t}$$
for the Itô-Taylor expansion on the hierarchial set $\Gamma_k$ and

$$Z_{\Gamma_k} = \sum_{\alpha \in \Gamma_k} \mathcal{J}_\alpha[f_{\alpha}(t_0, X(t_0))]_{t_0, t}$$

for the Stratonovich-Taylor expansion. Then for $k = 1$ we obtain $\Gamma_k = \{(v), (0), (1)\}$ and thus

$$X_{\Gamma_1} = x + aI_{(0)} + aI_{(1)} \neq x + aJ_{(0)} + aJ_{(1)} = Z_{\Gamma_1}.$$  

The hierarchial set $\Gamma_k$ is important for the Stratonovich-Taylor expansion. Which is in this case obviously similar to the deterministic Taylor expansion. Because of analogy between Stratonovich-Taylor expansion and deterministic Taylor expansion and because the Stratonovich-Taylor need weaker assumptions in determining the remainder as was showed in example 2.2.2. The Stratonovich-Taylor expansion is much more convenient to work with by developing the numerical schemes. However, we can always switch between both expansions.

We should always think about it and choose the suitable form of the stochastic differential equation by using the relation (2.8).

### 2.3 Multiple Itô and Stratonovich integrals

In the previous section we introduced the stochastic counterparts of the deterministic Taylor expansion. Hence, know how to approximate any function, supposed we are able to evaluate the multiple integrals both Itô and Stratonovich. But this is not a trivial because multiple stochastic integrals cannot be generally expressed in terms of simpler stochastic integrals, especially when the Wiener process is multidimensional.

In this section we show the standard method for evaluating the multiple Stratonovich integral proposed by Kloeden and Platen in [4]. The multiple Itô integral is then derived using the theorem 2.2.1.

The presented method is based on a Kuhunen-Loève series expansion of the Wiener process. The starting point is the Brownian bridge process

$$\{W(t) - \frac{t}{\Delta} W(\Delta), 0 \leq t \leq \Delta\}$$
formed from the given $d$–dimensional Wiener process $W(t) = (W^1(t), \ldots, W^d(t))$ on time interval $[0, \Delta]$. The componentwise Fourier expansion of this process

$$W^j(t) - \frac{t}{\Delta} W^j(\Delta) = \frac{1}{2} a_{j,0} + \sum_{r=1}^{\infty} \left( a_{j,r} \cos \left( \frac{2r \pi t}{\Delta} \right) + b_{j,r} \sin \left( \frac{2r \pi t}{\Delta} \right) \right),$$  

(2.18)

$j = 1, \ldots, d$ has random coefficients

$$a_{j,r} = \frac{2}{\Delta} \int_0^\Delta \left( W^j(s) - \frac{s}{\Delta} W^j(\Delta) \right) \cos \left( \frac{2r \pi t}{\Delta} \right) ds,$$

$$b_{j,r} = \frac{2}{\Delta} \int_0^\Delta \left( W^j(s) - \frac{s}{\Delta} W^j(\Delta) \right) \sin \left( \frac{2r \pi t}{\Delta} \right) ds,$$

where $a_{j,r}, b_{j,r}$ are i.i.d. $N(0, \frac{\Delta}{2 \pi^2 r^2})$ for all $r = 1, 2, \ldots$ and $j = 1, \ldots, d$. Now we will approximate the $j$–th component of $W(t)$ including just first $p$ terms of expansion (2.18) hence

$$W^{j,p}(t) = \frac{t}{\Delta} W^j(\Delta) + \frac{1}{2} a_{j,0} + \sum_{r=1}^{p} \left( a_{j,r} \cos \left( \frac{2r \pi t}{\Delta} \right) + b_{j,r} \sin \left( \frac{2r \pi t}{\Delta} \right) \right).$$  

(2.19)

Since Riemann-Stieltjes integrals with respect to $W^{j,p}(t)$ converge to the Stratonovich stochastic integrals the approximation of multiple Stratonovich equation can be defined using such integrals.

**Definition 2.3.1** The approximation of multiple Stratonovich integral $J^{p[1]}_{\alpha}[0,t]$ for any $t \in [0, \Delta]$ and any $\alpha = (j_1, \ldots, j_k) \in \mathcal{M}$ is the Riemann-Stieltjes integral

$$J^{p[1]}_{\alpha}[0,t] := \int_0^t \int_0^{s_1} \cdots \int_0^{s_k} \int_0^1 dW^{j_1,p}(s_1) dW^{j_2,p}(s_2) \cdots dW^{j_k,p}(s_k)$$  

(2.20)

with respect to the smooth functions defined by (2.19).

After some computations [4] chapter 5. section 8. we obtain for $t = \Delta$ following random values, which are the building blocks for generation of multiple Stratonovich integrals. For each $j = 1, \ldots, m$ and $r = 1, \ldots, p$ with $p = 1, 2, \ldots$ we have $\xi_j, \zeta_{j,r}, \eta_{j,r}, \mu_{j,p}$ and $\phi_{j,p}$ i.i.d. $N(0,1)$ distributed random values and defined by

$$\xi_j = \frac{1}{\sqrt{\Delta}} W^j(\Delta), \quad \zeta_{j,r} = \sqrt{\frac{2}{\Delta} \pi^r a_{j,r}}, \quad \eta_{j,r} = \sqrt{\frac{2}{\Delta} \pi^r b_{j,r}},$$  

(2.21)

$$\mu_{j,p} = \frac{1}{\sqrt{\Delta} \cdot \rho_p} \sum_{r=p+1}^{\infty} a_{j,r}, \quad \phi_{j,p} = \frac{1}{\sqrt{\Delta} \cdot \alpha_p} \sum_{r=p+1}^{\infty} b_{j,r},$$  

(2.22)
where
\[ \rho_p = \frac{1}{12} - \frac{1}{2\pi^2} \sum_{r=1}^{p} \frac{1}{r^2}, \quad \alpha_p = \frac{\pi}{180} - \frac{1}{2\pi^2} \sum_{r=1}^{p} \frac{1}{r^2}. \]

Approximation for the most used multiple integrals is mentioned below.

\[ J_p(0) = \Delta, \quad J_p(j) = \Delta(W^j) = \sqrt{\Delta} \xi_j, \quad J_p(0, 0) = \frac{1}{2} \Delta^2 \]
\[ J_p(j, 0) = \frac{1}{2} \Delta(\sqrt{\Delta} \xi_j + a_{j,0}), \quad J_p(0, 1) = \frac{1}{2} \Delta(\sqrt{\Delta} \xi_j - a_{j,0}), \]

where \( \Delta(W^j) = W^j(\Delta) - W^j(0) \) and
\[ a_{j,0} = -\frac{1}{\pi} \sqrt{2\Delta} \sum_{r=1}^{p} \frac{1}{r} \zeta_{j,r} \]
and
\[ J_p^p_{(j, j_2)} = \frac{1}{2} \Delta \xi_1 \xi_2 - \frac{1}{2} \sqrt{\Delta} (a_{j_2,0} \xi_{j_1} - a_{j_1,0} \xi_{j_2}) + \Delta A_{j_1, j_2}^p, \]
where
\[ A_{j_1, j_2}^p = \frac{1}{2\pi} \sum_{r=1}^{p} \frac{1}{r} (\zeta_{j_1,r} \eta_{j_2,r} - \zeta_{j_2,r} \eta_{j_1,r}). \]

Remark 2.3.1 It can be proved, that
\[ \mathbb{E} \left[ \left| J_{j_1, j_2} - J_{j_1, j_2}^p \right|^2 \right] \leq \Delta^2 \rho_p \]


□

The Itô multiple integrals are then obtained by 2.2.1 and by using following theorem

Theorem 2.3.1 Let \( j_1, j_k \in \{0, 1, \ldots, d\} \) and \( \alpha = (j_1, \ldots, j_k) \in M \) where \( k = 1, 2, \ldots \). Then
\[ W_i^j I_{\alpha}[1]_{0,t} = \sum_{i=0}^{k} I_{(j_1, j_i, \ldots, j_k)}[1]_{0,t} + \sum_{i=1}^{k} I_{\{j_i \neq 0\}} I_{(j_1, \ldots, j_i, 0, \ldots, j_k)}[1]_{0,t} \]
for all \( t \geq 0 \).
Proof: The theorem is proved in [4] as Proposition 5.2.3.

□

Similarly the Stratonovich multiple integrals are often evaluated by using other recently known multiple integrals.

**Theorem 2.3.2** Let \( j_1, \ldots, j_k \in \{0, 1, \ldots, d\} \) and \( \alpha = (j_1, \ldots, j_k) \in \mathcal{M} \) where \( k = 1, 2, \ldots \) Then

\[
W_t^j \mathcal{J}_{\alpha}[1]_{0,t} = \sum_{i=0}^{k} \mathcal{J}_{(j_1, j_2, \ldots, j_{i-1}, j_{i+1}, \ldots, j_k)}[1]_{0,t}
\]

for all \( t \geq 0 \).

Proof: The theorem is proved in [4] as Proposition 5.2.10.

□
Chapter 3

Numerical implementation

In this chapter the implemented algorithms shall be described and using of simple computer examples their characteristics should be computed statistically.

3.1 Applied algorithms

As in standard numerical analysis there are many evaluation algorithms and every more or less appropriate for some class of differential equations. Moreover in stochastic numerical analysis we can always decide whether we are looking for a pathwise approximation or just for a probability distribution or other functionals of the desired Itô process resulting from the given stochastic differential equation.

Here we restrict to the pathwise discrete time algorithms, especially to the class of strong Taylor schemes. In this chapter we will always suppose $X(t), t \in [0, T]$ to be solution of $m$—dimensional Itô SDE rewritten in complement form

$$dX^i(t) = a^i(t, X(t))dt + \sum_{j=1}^{d} b^{i,j}(t, X(t))dW^j(t), \quad i = 1, \ldots, m; \quad (3.1)$$

where $a^i : \mathbb{R}^+ \times \mathbb{R}^m \to \mathbb{R}$ and $b^{i,j} : \mathbb{R}^+ \times \mathbb{R}^m \to \mathbb{R}$ for each $i = 1, \ldots, m$ and each $j = 1, \ldots, d$ are Borel measurable functions satisfying the linear growth and Lipschitz condition from theorem 2.1.1 and $W^j(t)$ are components of $d$—dimensional Wiener process.

Then we will want to generate a pathwise discrete time approximation $Y^{(n)}_i$
of the Itô process \( X(t) \). We denote the \( m \)-dimensional discrete time approximation \( Y^{(n)}_i \) as

\[
Y^{(n)}_i = (Y^{(n),1}_i, \ldots, Y^{(n),m}_i) = (Y^{(n),1}(t_i), \ldots, Y^{(n),m}(t_i)), \tag{3.2}
\]

where \( Y^{(n),j}_i \) is the \( j \)-th component for \( j = 1, \ldots, m \), and \( i = 1, \ldots, n \), on some partition

\[
\Delta_n([0, T]) = \{0 = t_0 < \cdots < t_n = T\},
\]

with equidistant time steps \(|\Delta_n([0, T])| = t_{i+1} - t_i \) for all \( i = 0, \ldots, n - 1 \).

For convenience, we will write \( Y_i \) instead of \( Y^{(n)}_i \) and simply \( \Delta = |\Delta_n([0, T])| \) if the partition \( \Delta_n([0, T]) \) is obvious.

Finally, as before in the text we denote \( \Delta W^j_i = W^j(t_{i+1}) - W^j(t_i) \). The discrete time algorithms evaluate approximate values of the Itô process \( X(t), t \in [0, T] \) just on some partition \( \Delta_n([0, T]) \). Always by such algorithms a discrete time approximations are considered to be a time continuous stochastic processes on \([0, T] \).

For judging the correctness of the scheme and for the presentation of the results we need to know both the theoretical

\[
\epsilon = \mathbb{E}[|X(T) - Y_n|],
\]

and the estimated mean error of the approximation

\[
\hat{\epsilon}_n = \sum_{i=1}^{N} |X(T)(\omega_i) - Y_n(\omega_i)|, \tag{3.3}
\]

where \( N \) is a count of independent sampled paths and \( n \) determines the partition \( \Delta_n([0,1]) \).

The estimated mean error is a random value i.e. it can take two different values in two different runs. Thus the \((1 - \alpha)\)-confidence interval is the more suitable estimate of the error.

The following \((1 - \alpha)\)-confidence interval is constructed in [5] page 117-118.

**Remark 3.1.1** Let \( N \) be count of independent trajectories in a batch and \( M \) count of batches. Denote

\[
\hat{\epsilon}_j = \frac{1}{N} \sum_{i=1}^{N} |X(T)(\omega_{j,i}) - Y_n(\omega_{j,i})|, \quad j = 1, \ldots, M;
\]

48
and

\[ \sigma^2_\epsilon = \frac{1}{M-1} \sum_{j=1}^{M} (\hat{\epsilon}_j - \bar{\epsilon})^2, \]  (3.4)

where \( \hat{\epsilon} = \sum_{j=1}^{M} \hat{\epsilon}_j \). Then for \( N \geq 15 \) we can construct the \((1-\alpha)\)–confidence interval for \( \epsilon \) as

\[ (\hat{\epsilon} - \Delta \hat{\epsilon}, \hat{\epsilon} + \Delta \hat{\epsilon}), \]  (3.5)

with

\[ \Delta \hat{\epsilon} = t_{1-\frac{\alpha}{2}, M-1} \sqrt{\frac{\hat{\sigma}^2}{M}}, \]  (3.6)

where \( t_{1-\alpha/2, M-1} \) is a \((1-\alpha/2)\)–quantile of student-t distribution with \( M - 1 \) degrees of freedom.

For a better results \( N = 100 \) is usually set up. Furthermore, the equation (3.6) gives us the count of batches \( M \) needed to obtain enough accuracy of the confidence interval.

For judging the information about \( \hat{\epsilon}_n \) across various SDEs, also the \( \mathbb{E}[X(T)] \) must be included, i.e \( \hat{\epsilon}_n = 1 \) is worse result if \( \mathbb{E}[X(T)] = 1 \) than if \( \mathbb{E}[X(T)] = 10^3 \). Hence also the ratio \( \hat{\rho}_n \) will be observed.

\[ \hat{\rho}_n = \frac{1}{N} \frac{\hat{\epsilon}_n}{\sum_{i=1}^{N} X(T, \omega_i)} \]  (3.7)

Finally the following convergence criterion compares the effectiveness of numerical schemes.

**Definition 3.1.1** Let \( X(t), t \in [0, T] \) be a solution of SDE on probability space \((\Omega, \mathcal{F}_t, P)\), we say that its discrete time approximation \( Y(t) \) converges strongly with order \( \gamma > 0 \) at time \( t \in [0, T] \) if there exists a constant \( C > 0 \), which does not depend on \( \delta \) and \( \delta_0 > 0 \) such that

\[ \epsilon(\delta) = \mathbb{E}[|X(t) - Y(t)|] \leq C\delta^\gamma, \]

for all \( \delta \in (0, \delta_0) \).

The order of convergence can be increased by involving more terms from the Itô-Taylor resp. Stratonovich-Taylor expansion.

Now we shall introduce the algorithms. Let start with the most intuitive scheme of this class with the Euler scheme.
Algorithm 3.1.1 Suppose SDE 3.1 and notation as above, the \( k \)-th component of the Euler scheme has the form
\[
Y^k_{i+1} = Y^k_i + a^k(t_i, Y_i)\Delta + \sum_{j=1}^{d} b^{k,j}(t_i, Y_i)\Delta W^j_i. \tag{3.8}
\]
For Euler scheme it can be proved [4] Theorem 10.2.2, that the order of convergence is \( \gamma = 0.5 \). Euler scheme gives good results when drift and diffusion coefficient functions are nearly constants. However in general the use of higher order schemes are recommended.

Second less intuitive but stronger algorithm is the Mielstein scheme.

Algorithm 3.1.2 Consider again SDE (3.1) as above, then the \( k \)-th component of the Mielstein scheme has the form
\[
Y^k_{i+1} = Y^k_i + a^k(t_i, Y_i)\Delta + \sum_{j=1}^{d} b^{k,j}(t_i, Y_i)\Delta W^j_i + \sum_{j_1,j_2=1}^{d} L^{j_1} b^{k,j_2} I_{j_1,j_2}. \tag{3.9}
\]

The Mielstein scheme corresponds to the Itô-Taylor expansion on the hierarchial set \( \Lambda_2 = \{(\nu), (0), (1), (1,1)\} \). Hence the form of Mielstein scheme is the same using either Itô-Taylor or Stratonovich-Taylor expansion, see example 2.2.3. The Mielstein scheme provides us with convergence order \( \gamma = 1.0 \), proved in [4]. The additional term
\[
\sum_{j_1,j_2=1}^{d} L^{j_1} b^{k,j_2} I_{j_1,j_2}
\]
in Mielstein scheme marks the point of divergence of stochastic numerical analysis from the deterministic, because it gives us an addition information about the Wiener process.

This is generally very important for solving SDE, but also quite computationally demanding, because for \( d > 1 \) we have to evaluate the multiple stochastic integrals either Itô or Stratonovich. We should always seek for some simplifications, how to avoid the evaluation of multiple integrals.

The most common simplification of the Mielstein scheme is the case of commutativity.

Remark 3.1.2 Suppose a SDE of the form (3.1) such as
\[
L^{j_1} b^{k,j_2} = L^{j_2} b^{k,j_1}, \quad \forall j_1, j_2 = 1, 2, \ldots, d \text{ and } \forall k = 1, \ldots, m. \tag{3.10}
\]
Then we can use a Mielstein scheme of the form

$$Y_{i+1}^k = Y_i^k + a^k(t_i, Y_i)\Delta + \sum_{j=1}^{d} b^{k,j}(t_i, Y_i)\Delta W_i^j + \sum_{j_1, j_2=1}^{d} L^{j_1} b^{k,j_2} \frac{1}{2} \Delta W_i^j \Delta W_i^{k_j}. \tag{3.11}$$

**Proof:** For any $k = 1, \ldots, m$ we have

$$\sum_{j_1, j_2=1}^{d} L^{j_1} b^{k,j_2} I_{(j_1, j_2)} = \sum_{j_1=1}^{d} L^{j_1} b^{k,j_1} I_{(j_1, j_1)} + \sum_{j_1=1}^{d-1} \sum_{j_2=j_1+1}^{d} L^{j_1} b^{k,j_2} I_{(j_1, j_2)} + \sum_{j_2=1}^{d} \sum_{j_1=j_2+1}^{d} L^{j_2} b^{k,j_1} I_{(j_2, j_1)}.$$

Then

$$\sum_{j_1=1}^{d} L^{j_1} b^{k,j_1} \frac{1}{2} (\Delta W_i^{j_1})^2 = \sum_{j_1=1}^{d} L^{j_1} b^{k,j_1} \frac{1}{2} (\Delta W_i^{j_1})^2.$$

Further, the theorem 2.3.2 provides us with

$$J_{(j_1, j_2)}[1]_{0,\Delta} + J_{(j_2, j_1)}[1]_{0,\Delta} = W^{j_1}(\Delta)W^{j_2}(\Delta) = \Delta W_i^{j_1} \Delta W_i^{j_2}.$$

hence using the assumption (3.10) we obtain

$$\sum_{j_1=1}^{d-1} \sum_{j_2=1, j_2=2}^{d} L^{j_1} b^{k,j_2} (I_{(j_1, j_2)} + I_{(j_2, j_1)}) = \sum_{j_1=1}^{d-1} \sum_{j_2=1}^{d} L^{j_1} b^{k,j_2} (\Delta W_i^{j_1} \Delta W_i^{k_j}).$$

Using the assumption (3.10) backwards we obtain the statement.

□

As stated in previous chapter we build higher order schemes just by involving more particles from the Itô-Taylor, or Stratonovich-Taylor expansion respectively. The last chosen scheme is the strong 1.5 order, which corresponds to the Itô-Taylor scheme on the hierarchial set $\Gamma_3$, hence has not the same form for Stratonovich-Taylor expansion.
Algorithm 3.1.3 Suppose the SDE (3.1) from beginning of this chapter, then the \( k \)-th component of the order 1.5 strong Taylor scheme is equal to

\[
Y_{i+1}^k = Y_i^k + a^k(t_i, Y_i)\Delta + \frac{1}{2}L^0 a^k(t_i, Y_i)\Delta^2 + \\
+ \sum_{j=1}^{d} b^{kj}(t_i, Y_i)\Delta W_i^j + L^0 b^{kj}\mathcal{I}_{(0,j)} + L^j a^k(t_i, Y_i)\mathcal{I}_{(j,0)} + \\
+ \sum_{j_1,j_2=1}^{d} L^{j_1} b^{kj_2} I_{(j_1,j_2)} + \sum_{j_1,j_2,j_3=1}^{d} L^{j_1} L^{j_2} b^{kj_3} I_{(j_1,j_2,j_3)}, \tag{3.12}
\]

Order of convergence of this scheme is \( \gamma = 1.5 \) [4]. As will be demonstrated on PC examples.

### 3.2 Computer examples

In the following text the simple examples with known theoretical solutions are studied. Each considered stochastic differential equation is solved by any of numerical schemes 3.1.1-3.1.3. Then the numerical solution is compared with corresponding theoretical solution. The convergence and computational time will be evaluated and its consequences will be described. Examples are numbered as PC-example \( Z \), each example has its PC implementation on the attached CD under the name ”ex-Z.r”, where \( Z \) is number of PCexample.

The following settings are the same for each example:

(i.) The evaluation interval \([0, T]\) is set commonly as \([0, T] = [0, 1]\).

(ii.) Equidistant partition \( \Delta_n([0, 1]) \) is taken, where \( |\Delta_n| = t_{i+1} - t_i = 2^{-n} \), \( n \in \{3, 4, 5, 6, 7, 8, 9\} \), for all \( i = 1, \ldots, n \).

(iii.) The \( d \)-dimensional Wiener process \( W(t) = (W^1(t), \ldots, W^d(t)) \) is generated componentwise at the beginning of the each simulation then various schemes are evaluated using the same path of Wiener process.

(iv.) Each \( W^i(t), i = 1, \ldots, d \) is generated as a random walk with \( N(0, \delta) \) distributed steps, where \( \delta = 2^{-10} \). It is supposed to be enough accuracy for our purposes. This approach provides us with the same
sample paths for all schemes and simplifies the plotting of the theoretical solution.

We will distinguish between SDEs, where only $1$–dimensional Wiener process is involved, and SDEs where $d$–dimensional Wiener process with $d \geq 2$ takes part. We shall start with the case $d = 1$, where the multiple stochastic integrals do not need to be evaluated since from the theorem 2.3.2 we can see that

$$J_\alpha[1]_{0,\Delta} = \frac{\mathcal{J}_{(j_1)}^k}{k!} \frac{(\Delta W_{j_1})^k}{k!} \text{ for } \alpha = (j_1, \ldots, j_k) \in \mathcal{M} \text{ and } j_1 = \ldots = j_k.$$

**The Wiener process $W(t)$ is $1$–dimensional:** At the beginning the stochastic differential equations used in the $1$–dimensional examples and their solutions will be introduced.

$$dX(t) = aX(t)dt + bX(t)dW(t),$$

$$X(t) = X(t_0) \exp\{ (a - \frac{1}{2}b^2)(t - t_0) + b(W(t) - W(t_0)) \}.$$  

The solution is given by theorem 2.1.4. The next SDE was solved as a example 2.1.1.

$$dX(t) = \frac{1}{3}X(t)\frac{2}{3}dt + X(t)\frac{2}{3}dW(t),$$

$$X(t) = \left(X(0)\frac{1}{3} + \frac{W(t)}{3}\right)^3.$$  

Firstly, the motivation examples will be showed.

**PC-example 3.2.1** Using the algorithms 3.1.1-3.1.3 evaluate the solution for $1$–dimensional $X(t)$ Itô process given by (3.13) with $a = 1.5$, $b = 2$ and plot results for $|\Delta_n| = 2^{-n}$, where $n = 3, 6$ against the corresponding theoretical solution.

**PC-example 3.2.2** Repeat the PC-example 3.2.1 for the SDE of the form (3.14).
Hence we can easily see that in the case of (3.13) all of schemes work well especially for $n = 6$, however for the equation (3.14) the results of the Euler scheme are a bit worse as we shall survey in next examples.

Next example shows the convergence of used algorithms. It is an important criterion in deciding whether the algorithm is well chosen and well implemented, because lower discrete time step $|\Delta_n|$ always results into a lower theoretical mean error $\epsilon$. And hence if the scheme is well implemented then

![Figure 3.1: PC-example 3.2.2 $n = 3, n = 6$](image1)

![Figure 3.2: PC-example 3.2.2 $n = 3, n = 6$](image2)
lower discrete time step results into the lower estimated error $\hat{e}$. Moreover, for the discrete time step $|\Delta_n| = 2^{-n}$ we can obtain from (3.1.1)

$$\mathbb{E}[|X(T) - Y_n|] < K \cdot (2^{-n})^\gamma$$

$$\log_2 (\mathbb{E}[|X(T) - Y_n|]) < -n\gamma + \log_2(K),$$

hence also the estimated error decreases linearly with trend $-\gamma$ since $\hat{e}_n \to \epsilon(|\Delta_n|)$ as $N \to \infty$.

**PC-example 3.2.3** Using the algorithms 3.1.1-3.1.3 find the solution for 1-dimensional $X(t)$ Itô process given by (3.13) with $a = 1.5$, $b = 2$. For $M = 30$ batches and $N = 20$ independent trajectories in a batch evaluate the estimated error $\hat{e}_n$ defined by (3.3), its estimated variance (3.4), its confidence interval (3.6) and its ratio $\hat{\rho}_n$ for each discrete time step $|\Delta_n| = 2^{-n}$, where $n = 3, 4, 5, 6, 7, 8, 9$. Plot $\log_2(\hat{e}_n)$, $\hat{e}_n$, $\hat{\rho}_n$ against $\log(|\Delta_n|)$ and plot $\log_2(\hat{\epsilon}_n)$, $\hat{\epsilon}_n$ with its (0.9)-confidence intervals against $\log(|\Delta_n|)$.

![Figure 3.3: PC-example 3.2.3 $\hat{e}_n$ with 90%-conf. int., $\log_2(\hat{e}_n)$](image)

**Notes:** We can see from figure 3.3 that the convergence orders are in compliance with the theory, thus all the schemes are well implemented. Moreover, from the plot it can seen that for Euler scheme the confidence interval do not shorten. What shows an insufficient result and we should avoid the Euler scheme in this case.

◊
**PC-example 3.2.4** Repeat the PC-example 3.2.3 for the equation (3.13) with \( a = 1.5, \ b = 0.1 \).

![Graph showing error and log(error) against power]

**Figure 3.4:** PC-example 3.2.4 \( \hat{\epsilon}_n \) with 90%-conf. int., \( \log_2(\hat{\epsilon}_n) \)

**Notes:** Here we can observe an opposite, since the random noise given by \( b = 0.1 \) is too weak the SDE behaves almost like deterministic, and the Euler scheme proves order of convergence in \( \gamma = 1 \) in shorter the the other schemes, hence could be recommend for solution of such a SDE.  

---

**PC-example 3.2.5** Repeat the PC-example 3.2.3 for the equation (3.14).

**Notes:** The last equation was perfectly fitted by the 1.5 order scheme which is proven by the plot of error. On the other hand it could seem to be badly implemented since the \( \log(\hat{\epsilon}_n) \) rises, but the \( \hat{\epsilon}_n \sim 10^{-16} \) what causes probably in a truncation error of the programme R.

We should note that the construction \( |\Delta_n| = 2^{-n} \) is a very suitable one, since it corresponds very good with the definition of the convergence criterion and
with a natural question how to decrease $|\Delta_n|$ to obtain half error bound than $K \cdot (|\Delta_{n-1}|)^{\gamma}$. Then from (3.1.1) we have

$$\frac{1}{2} K \cdot (|\Delta_{n-1}|)^{\gamma} = K \cdot (|\Delta_n|)^{\gamma}$$

$$\left(\frac{1}{2}\right)^{\frac{1}{2}} |\Delta_{n-1}| = |\Delta_n|.$$

Therefore, the equidistant partition with $|\Delta_n| = 2^{-n}$ corresponds to the sequence of halved error bounds for a scheme with order $\gamma = 1$. Since for equidistant partition holds $|\Delta_n| = (\text{card}\{\Delta_n([0,1]) - 1\})^{-1}$, it can be seen that

$$\text{card}\{\Delta_n^{(1)}([0,1]) - 1\} = 2^{\frac{1}{\gamma}} \text{card}\{\Delta_{n-1}^{(1)}([0,1]) - 1\}.$$

Thus if we want to halve the error bound then the computational time of the scheme $\tau_n$ increases exponentially. This shall be demonstrated by following example.

**PC-example 3.2.6** Evaluate the (3.13) with the algorithms 3.1.1-3.1.3. For $|\Delta_n| = 2^{-n}$, $n = 3, 4, 5, 6, 7, 8, 9$ plot $\tau_n$ against $\log_2(|\Delta_n|)$ and $\log_2(\tau_n)$ against the $\log_2(|\Delta_n|)$.

This numerical example has an essential corollary. Since the time demand of any scheme grows exponentially, every scheme has limiting accuracy which
can be improved just by improving the order of convergence. The importance of convergence order criterion should be demonstrated on following example.

**PC-example 3.2.7** Using the algorithms 3.1.1-3.1.3 enumerate the solution for 1-dimensional $X(t)$ Itô process given by (3.13) with $a = 1.5$, $b = 2$. For $|\Delta_n| = 2^{-n}$, where $n = 3, 4, 5, 6, 7, 9$ evaluate the time spent by evaluation $\tau_n$ and estimated error $\hat{\epsilon}_n$. Plot $\tau_n$ against $\hat{\epsilon}_n$ and $\log(\tau_n)$ against $\log(\hat{\epsilon}_n)$.

**PC-example 3.2.8** Repeat the PC-example 3.2.7 for the equation (3.13) with $a = 1.5$, $b = 0.1$.

*Notes:* Because time of the simulation of one trajectory is not always measurably long, the time is observed for $k$ same trajectories. And the estimated error is then summed up over $M = 800$ independent trajectories. Which gives quite reasonable data, but we should always pay attention about randomness of such an estimated error.

The plots on the left hand side shows that as the time grows exponentially, it is needed to obtain desired accuracy with relatively long $|\Delta_n|$ i.e. for $|\Delta_n| = 2^{-n}$ with $n < 9$. If not, the additional improvement costs too much time and it would more effective to use a scheme with higher order of convergence.
The plots on the right hand side conclude the observations from 3.3 and 3.6. The estimated error $\hat{\epsilon}_n \sim \tau_n$ and $\log(\hat{\epsilon}_n) \sim \tau_n$.
We can conclude that this plot gives us the information which scheme is the most advantageous to use for appropriate SDE.

From examples above we can conclude that Euler scheme is a suitable method for solving equations with low influence of random noise. For other cases the Mielstein scheme should be supposed as the first choice, because of its satisfying order of converge with relatively weak assumptions, however always the properties of given SDE are determining.

**Problems with \( d \)-dimensional Wiener process:**

Using simple examples the importance of the convergence order was showed for \( d = 1 \). Then as we shall see the convergence order of any scheme is even more important if \( d, m \geq 2 \).

To show the insufficient results for Euler scheme we will evaluate a \( 2 \)-dimensional Itô process with \( 2 \)-dimensional Wiener processes, and compare it with its theoretical solution. As we shall see this process is of the commutative case and therefore the evaluation of multiple integrals can be avoided.

For the noncommutative case we are no more provided with the \( P-a.s.- \) convergence but just with the mean-square convergence by theorem 2.3.1. Hence just the convergence in distribution will be showed using boxplots and tested by Kolmogorov-Smirnov test.

Our first stochastic differential equations is of the commutative case. It is of the form

\[
dX(t) = B^1 X(t) dW^1(t) + B^2 X(t) dW^2(t)
\]

with

\[
B^1 = \begin{pmatrix} 0 & -1 \\ 1 & -2 \end{pmatrix}, \quad B^2 = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}
\]

From theorem 2.1.4 we can obtain a theoretical solution

\[
X(t) = \Phi(t) X(0), \quad \text{where} \quad \Phi^1 = \begin{pmatrix} \Phi_{11}(t) & \Phi_{12}(t) \\ \Phi_{21}(t) & \Phi_{22}(t) \end{pmatrix}
\]

\[
\Phi_{11}(t) = \left( 1 + W^1(t) + W^2(t) \right) \exp \left\{ -\frac{t}{n} - W^1(t) + W^2(t) \right\}
\]

\[
\Phi_{12}(t) = -\left( W^1(t) + W^2(t) \right) \exp \left\{ -\frac{t}{n} - W^1(t) + W^2(t) \right\}
\]

\[
\Phi_{21}(t) = \left( W^1(t) + W^2(t) \right) \exp \left\{ -\frac{t}{n} - W^1(t) + W^2(t) \right\}
\]

\[
\Phi_{22}(t) = -\left( 1 + W^1(t) + W^2(t) \right) \exp \left\{ -\frac{t}{n} - W^1(t) + W^2(t) \right\}.
\]
The same properties shall be observed using the same examples as was done for $d = 1$. Firstly, again the motivation example is showed.

**PC-example 3.2.9** Repeat the PCexample 3.2.1 for the SDE of the form (3.15).

![Figure 3.9: PC-example 3.2.9: n=3](image1)

![Figure 3.10: PC-example 3.2.9: n=6](image2)
Notes: We can see that the trajectory of "Mielstein" (Mielstein scheme with multiple Stratonovich integral evaluation) and "Mielstein II" (using the $\Delta W^1 \Delta W^2$ instead) are identical, which is in compliance with theory.

Next the estimated error, variance of estimated error, confidence interval of estimated error, ratio of estimated error and the order of convergence shall be evaluated for both dimensions separately.

**PC-example 3.2.10** Repeat the PC-example 3.2.3 for the equation (3.15) moreover, evaluate $\frac{1}{N} \sum_{i=1}^{N} X(t, \omega_i)$.

Notes: From the figure 3.12 we can see that all methods are well implemented and their estimated order of convergence is in compliance with the theoretical one.

But before all the Euler scheme proved very insufficient results. As can be seen in the table 3.1 especially the error of 2nd dimension is around 33% with base 0.9308 for $|\Delta_n| = 2^{-9}$, and from the observations for $d = 1$ it can be concluded that any other improvement would cost too much time.

At last it should be noticed that, all results for "Mielstein" and "Mielstein II" method are the same, which is again in compliance with theory, since the equation (3.15) is of commutative case.

As next step again the time demand for evaluating the scheme will be observed.

**PC-example 3.2.11** Repeat the PC-example 3.2.6 for the equation (3.15).

Notes: Here the time demand is aggregated per both dimensions as the dimensions depend on each other and therefore the evaluation cannot run separately.

The results are similar to the results of problems with 1-dimensional Wiener process $W(t)$. Again the time demand of any scheme grows exponentially as its $|\Delta_n|$ decreases exponentially, thus every scheme has an limiting accuracy.

Finally, the accuracy versus time needed to reach such an accuracy shall be observed.

**PC-example 3.2.12** Repeat the PC-example 3.2.7 for the equation (3.15).
Figure 3.11: PC-example 3.2.10: $\hat{\epsilon}_n$ and $\log(\hat{\epsilon}_n)$

Figure 3.12: PC-example 3.2.10: $\hat{\epsilon}_n$ and $\log(\hat{\epsilon}_n)$
The plot is again in compliance with the previous figures 3.12 and 3.15. Denote \( \hat{\epsilon}_n^{Euler} \) resp. \( \hat{\epsilon}_n^{Mielstein} \) an estimated error of Euler resp. Mielstein scheme. Then \( \hat{\epsilon}_n^{Euler} \sim \hat{\epsilon}_n^{Mielstein} \) and the estimated errors of any scheme \( \hat{\epsilon}_n \) decreases similarly (it is another random realization) as in figures 3.12. Finally the curves are again shifted at start time in compliance with the figure 3.15.

As in examples with 1-dimensional Wiener process \( W(t) \), it can concluded that this plot gives us the information which scheme is the most advantageous to use for appropriate SDE. For case of commutative multidimensional SDE it is obviously the Mielstein scheme without evaluation of multiple integrals.

Finally the example of the non-commutative SDE with \( d \)-dimensional Wiener process, with \( d \geq 1 \) is showed. The evaluated equation is

\[
dX(t) = W^1(t)dW^2(t), \quad X(0) = 0.
\] (3.18)

Because the convergence can no more be compared per trajectory, as there is just mean-square convergence instead of \( P- \)a.s. Just the distribution in various times can be tested. The distribution is observed using boxplots and tested with Kolmogorov-Smirnov test.
Scheme | $n=3$ | $n=5$ | $n=7$ | $n=9$ | $\frac{1}{N} \sum_{i=1}^{N} X(t, \omega_i)$
--- | --- | --- | --- | --- | ---
**dim 1** | Euler | 1.6459 | 0.8995 | 0.2759 | 0.1115 | 2.812 |
| Mielstein | 0.3951 | 0.1236 | 0.0222 | 0.0057 | 2.812 |
| 1.5 order | 0.3951 | 0.1236 | 0.0222 | 0.0057 | 2.812 |
**dim 2** | Euler | -2.0941 | -0.8004 | -0.6312 | -0.3361 | -0.9308 |
| Mielstein | -0.5040 | -0.1144 | -0.0477 | -0.0155 | -0.9308 |
| 1.5 order | -0.5040 | -0.1144 | -0.0477 | -0.0155 | -0.9308 |

Table 3.1: PC-examples 3.2.10: Ratio $\hat{\rho}_n$

![Figure 3.14: PC-example 3.2.11: $\hat{\epsilon}_n \sim \tau_n$ for both dimensions](image-url)

Figure 3.14: PC-example 3.2.11: $\hat{\epsilon}_n \sim \tau_n$ for both dimensions
**Figure 3.15:** PC-example 3.2.11: $\log(\epsilon_n) \sim \tau_n$ for both dimensions

**PC-example 3.2.13** For the SDE given by 3.18 evaluate $M = 800$ multiple Statonovich integrals $J_{(1,2)}[1]_{0,\Delta_n}$ using $\frac{1}{2} \Delta W^1 \Delta W^2$ and the approximation $J_n^{(1,2)}[1]_{0,\Delta_n}$. Compare the distribution in time $t = 2^{-n}, n = 0, \ldots, 8$ using boxplots and Kolmogorov-Smirnov tests, table and plot the results.

<table>
<thead>
<tr>
<th>scheme</th>
<th>$t = 1$</th>
<th>$t = 0.25$</th>
<th>$t = 2^{-4}$</th>
<th>$t = 2^{-6}$</th>
<th>$t = 2^{-8}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$J_{(1,2)}^{(1,2)}$</td>
<td>0.1981</td>
<td>0.4282</td>
<td>0.8643</td>
<td>0.5853</td>
<td>0.1981</td>
</tr>
<tr>
<td>$(\Delta W^1 \Delta W^2)/2$</td>
<td>e-05</td>
<td>e-05</td>
<td>e-06</td>
<td>0.0002</td>
<td>0.0012</td>
</tr>
</tbody>
</table>

Table 3.2: PC-examples 3.2.10: p-value of Kolmogorov-Smirnov test

**Notes:** From the plots and also from the p-value of Kolmogorov-Smirnov test is obvious that the it cannot be excluded, that the distribution of $J_{(1,2)}^{(1,2)}[1]_{0,\Delta_n}$ converge with $M \rightarrow \infty$ to the distribution of $J_{(1,2)}[1]_{0,\Delta_n}$. On the other hand, it can be observed (table 3.2) that the Kolmogorov-Smirnov test rejected that the distribution of $\frac{1}{2} \Delta W^1 \Delta W^2$ is similar to distribution of $J_{(1,2)}[1]_{0,\Delta_n}$ at any tested time. This result is very important for the solving of non-commutative SDEs with multidimensional Wiener process. Using the results from example 3.2.10 and this result it can be concluded, that only Mielstein scheme or any other higher order scheme with the evaluation of multiple stochastic integrals is suitable for solving the SDEs with multiple Wiener process.

66
Figure 3.16: PC-example 3.2.13

67
Chapter 4

Stochastic modeling in finance

4.1 Introduction to finance mathematics

In this section the fundamental theorems of financial mathematics will be presented. Our aim will be to sketch the proof of Black-Scholes-Merton theorem for pricing of an European option. It gives the theoretical background to the PC-example 4.1.1. Moreover, other financial models can be derived using similarities to Black-Scholes-Merton.

We will build up the market model in several steps. We expect that the only source of randomness is the 1-dimensional Wiener process $W(t), t \geq 0$ and the initial conditions of given assets. Hence the supposed probability space is $(\Omega, \mathcal{F}, P)$ with $\mathcal{F}_t$—Wiener process and holds $\mathcal{F} = \mathcal{F}_\infty$, and $X(0) \in \mathcal{F}_0$ for any initial condition.

Furthermore, we shall work with the interest rate model $B(t)$. It is supposed that if we invest an amount $B(t)$ into the money account with the variable interest rate $r(t)$, such as $r(t) = r$, in some interval $t \in [t, t + \Delta]$, our wealth at time $t + \Delta$ will be $B(t) + r(t)B(t)\Delta$ for each time increment $\Delta$. Hence putting $\Delta \to 0$ we obtain differential equation which defines a process.

**Definition 4.1.1** Let $r(t)$ be a random interest rate such as $\langle r(t) \rangle = 0$ and $P[\int_0^t |r(s)|ds < \infty] = 1$ for all $t \in [0, T]$ then we define the process $B(t)$ as solution of

$$dB(t) = r(t)B(t)dt, \quad B(0) = 1,$$

hence

$$B(t) = \exp\left\{ \int_0^t r(s)ds \right\}, \quad t \in [0, T].$$
The discount rate process $D(t)$ is defined as opposite to the interest rate model $B(t)$. It says which amount is needed to be invested into a money account at time $t = 0$ to receive the payment 1 at time $t$. Hence

$$D(t)B(t) = 1 \quad \text{i.e. } D(t) = \exp\left\{-\int_0^t r(s)ds\right\}. \quad (4.1)$$

Next, we model the asset price $S(t)$. There are many possibilities how to set up model for $S(t)$. One of the most common is the generalized geometric Brownian motion with drift i.e.

**Definition 4.1.2** Let $a : \mathbb{R}^+ \times \mathbb{R}^m \to \mathbb{R}^m$ and $\sigma : \mathbb{R}^+ \times \mathbb{R}^m \to \mathbb{R}^{m \times d}$ be Borel functions satisfying the conditions from theorem 2.1.1. Then the Itô process $S(t)$ such as

$$dS(t) = \alpha(t, S(t))S(t)dt + \sigma(t, S(t))S(t)dW(t), \ t \in [0, T],$$

with initial condition $S(0) = x$, is called geometric Brownian motion with drift.

To hedge the value of any derivative financial instrument we can build a self-financing (hedging) portfolio process $X(t)$. For the construction of a portfolio we assume that we can either buy a $\delta(t)$ units of an asset $S(t)$ or invest into a money account for interest rate $r(t)$ at each time $t \in [0, T]$.

Thus we define the self-financing portfolio as follows.

**Definition 4.1.3** Let Itô process $S(t)$ be a model of an asset price and $r(t)$ be an interest rate then the value of self-financing portfolio $X(t)$ is given by

$$dX(t) = \delta(t)dS(t) + r(t)(X(t) - \delta(t)S(t))dt, \ t \in [0, T],$$

with the terminal condition $X(T) = x$ and an $\mathcal{F}_t$-adapted process $\delta(t), 0 \leq t \leq T$. The process $\delta(t)$ is called delta hedging strategy.

For the financial mathematic models the no-arbitrage market is the key assumption. This assumption has a good sense, it is the expectation that nobody can make profit without any risk.

The arbitrage can be defined as follows.

**Definition 4.1.4** Suppose a probability space $(\Omega, \mathcal{F}, P)$ and a $\mathcal{F}_t$-adapted portfolio value process $X(t), t \geq 0$ with $X(0) = 0$ such as for some time $T > 0$ holds

$$P[X(T) \geq 0] = 1, \quad \text{and} \quad P[X(T) > 0] > 0,$$

Then we call $X(t)$ an arbitrage.
Next definition includes the assumption of no-arbitrage market.

**Definition 4.1.5** Let \((\Omega, \mathcal{F}, P)\) be a probability space and let \(Q\) be probability measure on measurable space \((\Omega, \mathcal{F})\), then \(Q\) is said to be risk neutral if

(i.) \(P\) and \(Q\) are equivalent, i.e \(P[A] = 0 \Leftrightarrow Q[A] = 0\) for each \(A \in \mathcal{F}\),

(ii.) the process \(D(t) X(t)\), \(t \in [0, T]\) is a \(\mathcal{F}_t\)-martingale with respect to (w.r.t) probability measure \(Q\) for each portfolio value process \(X(t)\) from \((\Omega, \mathcal{F}, P)\) and the discount process \(D(t)\).

First fundamental theorem of asset pricing gives us the relation between the no-arbitrage assumption and the risk neutral measure definition.

**Theorem 4.1.1** If a market model has a risk neutral measure \(Q\) it does not admit arbitrage.

**Proof:** Let \(X(t)\) be a self-financing portfolio and \(Q\) a risk neutral measure. Then \(\mathbb{E}[D(T)X(T)] = X(0)\) with respect to (w.r.t.) probability measure \(Q\) for any \(X(0) \in \mathbb{R}\). Suppose that \(X(0) = 0\) and \(P[X(T) < 0] = 0\). Since \(Q\) and \(P\) are equivalent measures from definition, we obtain \(Q[X(T) < 0] = 0\). Moreover, since \(D(t) > 0, t \leq 0\) also \(Q[D(T)X(T) < 0] = 0\) and from martingale property also \(Q[D(T)X(T) > 0] = 0\), backwards \(Q[X(T) > 0] = 0\) and hence \(P[X(T) > 0] = 0\), which excludes the arbitrage.

Next theorem provided by Girsanov gives the answer how to find a risk-neutral measure \(Q\).

**Theorem 4.1.2 (Girsanov theorem)** Suppose \(W(t)\) to be a \(d\)-dimensional \(\mathcal{F}_t\)-Wiener process on \((\Omega, \mathcal{F}, P)\), and \(\tilde{W}(t)\) be a \(d\)-dimensional Itô process on \((\Omega, \mathcal{F}, P)\) of the form

\[
d\tilde{W}(t) := \Theta(t, \omega)dt + dW(t), \quad 0 \leq t \leq T. \tag{4.4}
\]

Put

\[
M(t) := \exp \left( -\int_0^t \Theta(s, \omega)dW - \frac{1}{2} \int_0^t \Theta^2(s, \omega)ds \right), \quad 0 \leq t \leq T. \tag{4.5}
\]
Assume that \(M(t)\) is \(\mathcal{F}_t\)-martingale w.r.t. probability measure \(P\). Define the measure \(Q\) on \(\mathcal{F}_t\) by
\[
dQ(\omega) := M(T, \omega)dP(\omega).
\]
(4.6)

Then \(Q\) is a probability measure on \(\mathcal{F}_t\) and \(\tilde{W}(t)\) is a \(d\)-dimensional \(\mathcal{F}_t\)-Wiener process w.r.t. \(Q\) for \(t \in [0, T]\).

**Proof:** Is proved in [3] as theorem 8.6.4.

\[\square\]

Now we can apply the Girsanov theorem 4.1.2 to the price of an asset \(S(t)\) to obtain its form under the risk-neutral measure \(Q\).

**Corollary 4.1.1** Let the price model of an asset \(S(t)\) be an Itô process as in 4.1.2 on \((\Omega, \mathcal{F}, P)\) and \(W(t)\) be \(\mathcal{F}_t\)-Wiener process. Then
\[
ds(t) = r(t)S(t) + \sigma(t, S(t))d\tilde{W}(t), \quad t \in [0, T],
\]
(4.7)
under the risk neutral measure \(Q\) defined in (4.6), where
\[
\Theta(t, \omega) = \frac{a(t, S(t)) - r(t)}{\sigma(t, S(t))},
\]
(4.8)
supposed \(M(t)\) defined by (4.5) with \(\Theta(t, \omega)\) is a \(\mathcal{F}_t\)-martingale w.r.t \(P\).

Furthermore the discounted asset price model \(D(t)S(t)\) is \(\mathcal{F}_t\)-martingale w.r.t. the risk neutral measure \(Q\).

**Proof:** Firstly, we change the measure and Wiener process \(W(t)\)
\[
ds(t) = a(t, S(t))S(t)dt + \sigma(t, S(t))dW(t) + r(t)S(t)dt - r(t)S(t)dt
\]
\[
= S(t)\sigma(t, S(t)) \left[ r(t)S(t)dt + d\tilde{W}(t) \right] + r(t)S(t)dt
\]
\[
= S(t)\sigma(t, S(t))d\tilde{W}(t) + r(t)S(t)dt,
\]
hence from Itô formula we obtain
\[
d\left(D(t)S(t)\right) = -r(t)D(t)S(t)dt + D(t) \left[ S(t)\sigma(t, S(t))d\tilde{W}(t) + rX(t)dt \right] =
\]
\[
= D(t)S(t)\sigma(t, S(t))d\tilde{W}(t).
\]
Then by theorem 1.3.1 we know it is a martingale w.r.t. measure \(Q\).

\[\square\]
Similarly we obtain the value of self-financing portfolio $X(t)$ under the risk-neutral measure $Q$.

**Corollary 4.1.2** Suppose the value of a self-financing portfolio $X(t)$ to be an Itô process on $(\Omega, \mathcal{F}, P)$ and $W(t)$ be a $\mathcal{F}_t-$Wiener process. Then

$$dX(t) = \delta(t)S(t)\sigma(t, S(t))d\tilde{W}(t) + r(t)X(t)dt, \quad t \in [0, T],$$

(4.9)

under the risk neutral measure $Q$ same as in corollary 4.1.1. Furthermore the discounted value of self-financing portfolio $D(t)X(t)$ is $\mathcal{F}_t-$martingale under the measure $Q$.

**Proof:** Again we change the measure and Brownian motion

$$dX(t) = \delta(t)dS(t) + r(t)(X(t) - \delta(t)S(t))dt$$

$$= \delta(t)a(t, S(t))S(t)dt + \delta(t)\sigma(t, S(t))S(t)dW(t) +$$

$$+ r(t)X(t)dt - r(t)\delta(t)S(t)dt$$

$$= \delta(t)S(t)\sigma(t, S(t))\left[\frac{a(t, S(t)) - r(t)}{\sigma(t, S(t))}dt + dW(t)\right] + r(t)X(t)dt$$

$$= \delta(t)S(t)\sigma(t, S(t))d\tilde{W}(t) + r(t)X(t)dt,$$

hence from Itô formula we obtain

$$dD(t)X(t) = -r(t)D(t)X(t) + D(t)\left[\delta(t)S(t)\sigma(t, S(t))d\tilde{W}(t) + rX(t)dt\right]$$

$$= D(t)\delta(T)S(t)\sigma(t, S(t))d\tilde{W}(t).$$

Then again by theorem 1.3.1 we know it is a martingale.

$\square$

Now we are prepared to sketch the proof of the Black-Sholes-Merton formula for European call option.

A European call option is financial derivative instrument, which gives a right to its owner to buy an asset $S$ for a strike price $K$ at time $T$, where the price asset $S(t)$ is an $\mathcal{F}_t$-adapted random process. Thus the European call option pays to its owner payment

$$V(T) = (S(T) - K)^+ \quad \text{at time } T.$$

72
Any investor can sell the European call option for a price $X(t_0)$ where $0 \leq t_0 \leq T$ and immediately construct the self-financing portfolio $X(t)$, with initial value $X(t_0)$. At time $T$ he will have to pay $V(T) = (S(T) - K)^+$ to the buyer of the call option. Under the assumption of no-arbitrage market neither the seller nor the buyer of the option will make a profit in the mean value thus

$$
E^Q[V(T)] = E^Q[X(T)].
$$

Moreover, this must hold for any time $t$, where $0 \leq t \leq T$. Since $D(t)X(t)$ is a martingale w.r.t. $Q$ i.e.

$$
D(t)X(t) = E^Q[D(T)X(T)|\mathcal{F}_t]
$$

we obtain

$$
D(t)V(t) = E^Q[D(T)\left(S(T) - K\right)^+|\mathcal{F}_t]. \tag{4.1.10}
$$

Now we will add a simplifying assumptions $a(t, S(t)) \equiv a \equiv r$ and $\sigma(t, S(t)) = \sigma$, for all $t \in [0, T]$. Then we obtain constant $\Theta(t, \omega) \equiv \frac{r - \sigma \sqrt{\tau}}{\sigma}$, for which the assumption of Girsanov theorem 4.4 clearly holds. Thus from (4.1.1) we have $S(t)$ under $Q$

$$
dS(t) = S(t)\sigma d\tilde{W}(t) + rS(t)dt,
$$

hence from 2.1.4

$$
S(T) = S(t)\exp\left\{ (r - \frac{1}{2}\sigma^2)(T - t) + \sigma(\tilde{W}(T) - \tilde{W}(t)) \right\}.
$$

from definition of $\mathcal{F}_t$—Wiener process $(\tilde{W}(T) - \tilde{W}(t)) \sim N(0, T - t) \perp \mathcal{F}_t$, thus $S(T - t)$ can be rewritten as follows, denote $\tau = T - t$

$$
S(T) = S(t)\exp\left\{ (r - \frac{1}{2}\sigma^2)\tau - \sigma\sqrt{\tau}Y \right\},
$$

where $Y \sim N(0, 1)$ distributed random value. Finally let $S(t) = x$ then a price of an European call option is a function

$$
c(t, x) = \mathbb{E}^Q\left[ e^{-rt} \left( x \exp\left\{ (r - \frac{1}{2}\sigma^2)\tau - \sigma\sqrt{\tau}Y \right\} - K \right)^+ \right]
$$

$$
= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-rt} \left( x \exp\left\{ (r - \frac{1}{2}\sigma^2)\tau - \sigma\sqrt{\tau}y \right\} - K \right)^+ e^{-\frac{1}{2}y^2}. \tag{4.11}
$$
Thus we obtain the statement of the Black-Scholes-Merton formula. Denote \( N(\cdot) \) the distribution function of the standardized normal distribution. Then the price of an European call option at time \( t \in [0, T] \), with price at time \( t \) \( S(t) \), market interest rate \( r \), strike price \( K \) and \( T \) date of expiry is given by

\[
c(\tau, x) = xN(d_+(\tau, x)) - e^{-r\tau}KN(d_-(\tau, x)),
\]

where \( \tau = T - t \), \( x = S(t) \) and

\[
d_\pm = \frac{1}{\sigma \sqrt{T}} \left[ \log \left( \frac{x}{K} \right) + \left( r \pm \frac{\sigma^2}{2} \right) \tau \right].
\]

From the sketch of proof of the Black-Scholes-Merton formula we obtain a clue how to make a numerical estimates of the value of option \( V(t) \).

Let simulate \( k \) pseudo-independent paths of \( S(t)(\omega), t \in [0, T] \), given by (4.1.1). Then we obtain a random sample

\[
V_i(T) := \left( S(T, \omega_i) - K \right)^+, \quad i = 1, \ldots, k.
\]

and then it is easy to construct an estimate of expected value of \( S(T) \) w.r.t. \( Q \)

\[
\hat{V}_k(t) = \frac{1}{k} \sum_{i=1}^{k} \frac{D(T)}{D(t)} V_i(T).
\] (4.11)

This estimate shall be used for pricing of an European call option at time \( t \).

We introduce characteristics, which will be observed in following examples.

Suppose \( M \) is count of random samples, \( k \) is number of independent trajectories in each sample and \( n \) is the accuracy of partition for evaluating the trajectory of \( S(t) \), again \( |\Delta_n| = 2^{-n} \). Then we denote error of estimated price as

\[
\varepsilon_{n,k} := \hat{V}_k(t) - V(t),
\]

estimated error by

\[
\hat{\varepsilon}_{n,k,M} := \sum_{j=1}^{M} \frac{1}{M} |\hat{V}_k(t) - V(t)|,
\]

the ratio of error of estimated price by

\[
\vartheta = \frac{\varepsilon_{n,k}}{V(t)}
\]
and finally the estimated ratio of error by

$$\hat{\rho}_{n,k,M} = \frac{1}{M} \sum_{i=1}^{M} |q_{n,k}^{(i)} - 1|.$$ 

Now we are ready to estimate the price of European call option numerically.

**PC-example 4.1.1** Using algorithm 3.1.1 evaluate the price of European call option at time $\tau = 0.5$, for the diffusion function $\sigma(t, S(t)) \equiv 1$, interest rate $r(t) \equiv 0.1$, strike price $K = 3$ and the initial value of the asset $S(\tau) = 3$. Plot histograms of errors $\varepsilon_{n,k,M}$ and ratios $\rho_{n,k,M}$, and table the estimated errors (4.11) for $M = 10$, $k = 500, 1000, 1500, 2000, 5000$ and discrete time step $|\Delta_n| = 2^{-n}$, $n = 3, 5, 7$.

![Histograms of errors and ratios](image)

Figure 4.1: PC-example 4.1.1, histograms of $\varepsilon_{n,k}, \rho_{n,k}$ for $n=3,7, k=2000$

**PC-example 4.1.2** Consider same problem as in example 4.1.1, evaluate it using algorithm 3.1.2, for $k = 500, 1000, 1500, 2000$ and discrete time step $|\Delta_n| = 2^{-n}$, $n = 3, 5, 7$.

**Notes:** From the table 4.1 it can be seen that both algorithms gives good results, but the convergence is strongly depending on $k$ not on $n$. It is obvious, since the evaluation is depending just on distribution $S(T)$ and therefore the exact trajectory is not such important as enough large random...
Figure 4.2: PC-example 4.1.2, histograms of $\varepsilon_{n,k}$, $\varrho_{n,k}$ for $n=3,7$, $k=2000$

<table>
<thead>
<tr>
<th>Scheme</th>
<th>$n$</th>
<th>$k=500$</th>
<th>$k=1000$</th>
<th>$k=1500$</th>
<th>$k=2000$</th>
<th>$k=5000$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Euler</td>
<td>3</td>
<td>0.0203</td>
<td>0.0184</td>
<td>0.0142</td>
<td>0.0146</td>
<td>0.01098</td>
</tr>
<tr>
<td>Euler</td>
<td>7</td>
<td>0.0201</td>
<td>0.0188</td>
<td>0.0155</td>
<td>0.0139</td>
<td>0.0083</td>
</tr>
<tr>
<td>Mielstein</td>
<td>3</td>
<td>0.0214</td>
<td>0.0184</td>
<td>0.0159</td>
<td>0.0139</td>
<td>0.0082</td>
</tr>
<tr>
<td>Mielstein</td>
<td>7</td>
<td>0.0212</td>
<td>0.0189</td>
<td>0.0160</td>
<td>0.0140</td>
<td>0.0084</td>
</tr>
</tbody>
</table>

Table 4.1: $\hat{\varepsilon}_{n,k,M}$, PC-example 4.1.1

<table>
<thead>
<tr>
<th>Scheme</th>
<th>$n$</th>
<th>$k=500$</th>
<th>$k=1000$</th>
<th>$k=1500$</th>
<th>$k=2000$</th>
<th>$k=5000$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Euler</td>
<td>3</td>
<td>0.0782</td>
<td>0.0708</td>
<td>0.0548</td>
<td>0.0561</td>
<td>0.0422</td>
</tr>
<tr>
<td>Euler</td>
<td>7</td>
<td>0.0773</td>
<td>0.0722</td>
<td>0.0595</td>
<td>0.0535</td>
<td>0.0320</td>
</tr>
<tr>
<td>Mielstein</td>
<td>3</td>
<td>0.0822</td>
<td>0.0706</td>
<td>0.0611</td>
<td>0.0536</td>
<td>0.0314</td>
</tr>
<tr>
<td>Mielstein</td>
<td>7</td>
<td>0.0815</td>
<td>0.0725</td>
<td>0.0615</td>
<td>0.0538</td>
<td>0.0324</td>
</tr>
</tbody>
</table>

Table 4.2: $\hat{\varrho}_{n,k,M}$, PC-example 4.1.1

76
sample, which is given by $k$, therefore also the algorithms 3.1.1 and 3.1.2 gives almost similar results.
At last it can be denoted that, histograms of $\epsilon_{n,k,M}$ and $\rho_{n,k,M}$ give similar information, only $\rho_{n,k,M}$ is better measurable since it says how was the error no matter if the theoretical price of the option is high or low.

\[\begin{align*}
\text{4.2 European lookback call option} \\
\text{To be able to compare the numerical results, we stay with the financial derivatives whose price have a theoretical solution.} \\
\text{We will focus on the European lookback call option. It is a representative of the path dependent or exotic options.} \\
\text{Financial derivatives with path dependency are suitable class of problems for our observation, because they have to be evaluated by means of path dependent approximations as shall be demonstrated in following examples.} \\
\text{A European lookback option is a derivative payment, which pays to its holder a payment} \\
LC(T) := (S(T) - \min_{u\leq T} S(u)), \\
\text{where $T$ is the maturity time.} \\
\text{Theorem 4.2.1} \quad \text{The price at time $t \in [0, T]$ of a European lookback option call option equals} \\
\begin{align*}
LC(t) &= sN(\tilde{d}) - me^{-r\tau}N(\tilde{d} - \sigma\sqrt{\tau}) - \frac{s\sigma^2}{2r}N(-\tilde{d}) + e^{-r\tau} \frac{s^2 \sigma^2}{2r} \left(\frac{m}{s}\right)^{2 \sigma^{-2}} N(-\tilde{d} + 2r \sigma^{-1} \tau), \\
\text{where} \quad s &= S(t), \quad \tau = T - t \quad \text{and} \\
\tilde{d} &= \log \left(\frac{m}{s}\right) + r + \frac{1}{2} \sigma^2.
\end{align*}
\text{Proof:} \quad \text{The theorem is proved in [6] as a Proposition 9.7.1.}\n\]
\]
In evaluating such a derivative numerically the same approach is used as for the pricing of a standard European call option. We suppose that $a(t, S(t)) \equiv a$, thus the assumption of Girsanov theorem 4.4 is satisfied and the measure can be changed. Then the hedging portfolio such as $X(T) = LC(T)$ will be constructed and under the risk neutral measure $Q$, both $D(t)X(t)$ and $D(t)LC(t)$ are again $\mathcal{F}_t$-martingales. The form of $S(t)$ w.r.t. $Q$ is again given by corollary 4.1.1. Finally the price of European lookback option at time $t \in [0, T]$ is given by

$$LC(t) = \mathbb{E}\left[\frac{D(T)}{D(t)} LC(T) \bigg| \mathcal{F}_t\right] \text{ w.r.t } Q.$$  

Since the trajectory of $S(t)$ is known using the pathwise strong approximations. It is simple to evaluate the

$$LC_i(T) := \left( S(T, \omega_i) - \min_{u \leq T} S(u, \omega_i) \right), \ i = 1, \ldots, k$$

and then construct the estimate of price of European lookback option

$$\widehat{LC}(t) = \frac{1}{k} \sum_{i=1}^{k} \frac{D(T)}{D(t)} LC_i(T). \quad (4.14)$$

At last we denote the error of estimated price by

$$\varepsilon_{n,k} := \widehat{LC}_k(t) - LC(t),$$

estimated error by

$$\hat{\varepsilon}_{n,k,M} := \sum_{j=1}^{M} \frac{1}{M} |\widehat{LC}_k(t) - LC(t)|,$$

the ratio of error of estimated price by

$$\varrho = \frac{\varepsilon_{n,k}}{LC(t)}$$

and finally the estimated ratio of error by

$$\hat{\varrho}_{n,k,M} = \sum_{i=1}^{M} \frac{1}{M} |\varrho^{(i)}_{n,k} - 1|.$$  

In next examples $\widehat{LC}(t)$ will be evaluated using the Euler and Mielstein schemes.
PC-example 4.2.1 Using algorithm 3.1.1 evaluate the price of European lookback call option at time $\tau = 0.3$, for the diffusion function $\sigma(t, S(t)) \equiv 1$, interest rate $r(t) \equiv 0.1$, strike price $K = 3$ and the initial value of the asset $S(\tau) = 3$. Table the estimated errors $\hat{\epsilon}_{n,k,M}$ and estimated ratios of errors $\hat{\rho}_{n,k,M}$ for $M = 10$, $k = 500, 1500, 2500, 5000$ and discrete time step $|\Delta n| = 2^{-n}$, $n = 3, 5, 7$.

PC-example 4.2.2 Repeat the example 4.2.1 using the algorithm 3.1.2.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
Scheme & n & $k=500$ & $k=1500$ & $k=2000$ & $k=5000$
\hline
Euler & 4 & 0.1000 & 0.0868 & 0.0912 & 0.0911
Euler & 8 & 0.0657 & 0.0378 & 0.0346 & 0.0271
Mielstein & 4 & 0.1167 & 0.1099 & 0.1167 & 0.1154
Mielstein & 8 & 0.0681 & 0.0394 & 0.0363 & 0.0291
\hline
\end{tabular}
\caption{$\hat{\rho}_{n,k,M}$, PC-example 4.2.1}
\end{table}

Notes: From the table 4.3 it can be seen that the estimate converge even more with $n$ and than with $k$, moreover, from histograms we can observe that the estimates are biased. Both Euler and Mielstein schemes show that $E[\hat{LC}(t)] < LC(t)$. And finally there is no difference in using the Euler or the Mielstein scheme. The bias and the stronger converge with $n$ than with $k$ are caused by the properties of the trajectory of $S(t)$ and will be explained by following example using the plots of trajectories.

The absence of difference between both schemes is again in large sample $k$, which is needed to obtain sufficient information about distribution of $LC(T)$. But on the other hand the mean error disappears for both schemes due to the law of large numbers i.e.

$$\sum_{i=1}^{k} \frac{1}{k} \hat{\epsilon}_{n,i} \to 0.$$ 

Thus there is no significant difference.

\begin{center}
\begin{tikzpicture}
\end{tikzpicture}
\end{center}

This bias shall be cleared by following example.
Figure 4.3: PC-example 4.2.1, histograms of $\varepsilon_{n,k}$ and $\varrho_{n,k}$ for $n=4$, $k=1500$ and $n=8$, $k=1500$

Figure 4.4: PC-example 4.2.2, histograms of $\varepsilon_{n,k}$ and $\varrho_{n,k}$ for $n=4$, $k=1500$ and $n=8$, $k=1500$
**PC-example 4.2.3** Using algorithm 3.1.2 evaluate the price of European lookback option from example 4.2.1. Moreover, plot the trajectory of $S(t)$, $\min\{S(u), u \leq t\}$ and $\hat{LC}(t)$ for $t \in [0, T]$ and discrete time step $|\Delta_n| = 2^{-n}, n = 3, 5, 7, 9$.

![Diagram of S(t) and min(S(u), u <= t)](image1)

Figure 4.5: PC-example 4.2.3 mostly increasing trajectory of $S(t)$

![Diagram of V(t) and (x_mres - min_St)](image2)

Figure 4.6: PC-example 4.2.3 mostly decreasing trajectory of $S(t)$

**Notes:** From the plots on the right hand side it can be seen that the price $V(t)(\omega)$ is always higher with higher $n$ both in the case of mostly increasing
$S(t)$, figure 4.5 and in the case of mostly decreasing $S(t)$, figure 4.6. Only in the case of strictly increasing or strictly decreasing $S(t)$ it would be the same for all $n$. Hence our estimate $\hat{LC}(t, \omega) < LC(t, \omega)$ in most of cases.

The reason is obvious from plots on the left hand side, $V(t)$ is higher because $\min\{S(u), u \leq t\}$ is lower for higher $n$. It is since for two different equidistant partitions $\Delta_j([0,1])$ and $\Delta_i([0,1])$, such as $|\Delta_j| = 2^{-j}$ and $|\Delta_i| = 2^{-i}$ holds

$$ \Delta_i([0,1]) \subset \Delta_j([0,1]) \quad \text{if} \quad i \leq j $$

and therefore

$$ \min_{t_l \in \Delta_i([0,1])} f(t_l) \geq \min_{t_l \in \Delta_i([0,1])} f(t_l). $$

The last example shows us how important can be the a good understanding of trajectory of any process.
Appendix

Manual to evaluation schemes in R

Each example has its PC implementation on the attached CD under the name "ex-Z.r", where Z is number of given PCexample. All the plots and tables presented in the thesis were evaluated under the recent settings with the same random seed and can be rebuilt just by running the whole script. For a simpler approach to the procedures each implementation is split into a logical parts as follows

- **Parameters input** Here all the constant as number of batches, count of independent trajectories in each batch or discrete time step can be chosen, depending on the parameters of the PC-example.

- **Function input** In this section the drift and diffusion functions of the given SDE can be set up.

- **Theoretical solution** Here the theoretical solution of SDE must be given if the PC-example compares the results from the numerical solution with the theoretical one.

- **Definition part** In this part all the schemes are declared as a function giving vector or matrix respectively depend on the dimension of the solved SDE. In addition vectors for saving the evaluated characteristics are declared.

- **Evaluation part** Here the evaluation usually runs over various discrete time steps and finally the results are printed.
Bibliography


