# Stochastic Calculus for Models in Finance

João Guerra

16/09/2013

# Contents

# Chapter 1

# Introduction

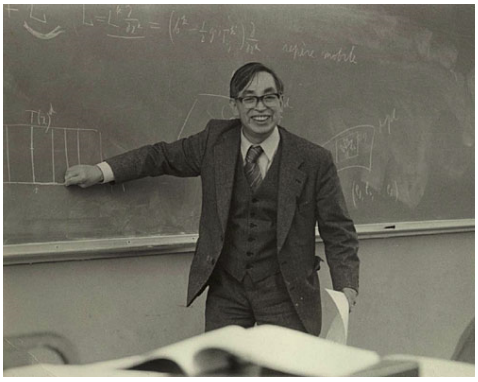
## 1.1 Outline

The aim of this course is to develop the necessary mathematical skills in order to understand and apply the mathematical methods of analytical, stochastic and numerical type, that play an important role in financial stochastic models either in discrete or continuous time. In particular, we are interested in models for the valuation of derivative securities. These skills are also important in order to communicate with other financial professionals and to critically evaluate modern financial theories. These lecture notes were prepared for the first part of the course "Models in Finance", of the Msc. degree in Actuarial Science in ISEG, Technical University of Lisbon, in the academic year 2012/2013. They cover the first five chapters of the programme that correspond to the theory of stochastic calculus, which is the core mathematical theory behind the models for the valuation of derivative securities. Therefore, it is necessary that the students understand the basic methods of stochastic calculus in order to be able to deduce the main properties of stochastic models for the valuation of derivative securities. In some parts of the text, we follow the references [9] and [10].

## 1.2 What is stochastic calculus?

What is stochastic calculus? Briefly, it is an integral (and differential) calculus with respect to certain stochastic processes (for example: Brownian motion). It allows to define integrals (and "derivatives") of stochastic processes where the "integrating function" is also a stochastic process. It allows to define and solve stochastic differential equations (where there is a random factor). The most important stochastic process for stochastic calculus and financial applications is the Brownian motion. The main financial applications of stochastic calculus are the pricing and hedging of financial derivatives, the study of the Black-Scholes model, interest rate models and credit risk modelling.

For a very interesting and modern account of the history of stochastic calculus, we refer to [4]. Some of the most important authors involved in the stochastic calculus development were Louis Bachelier, Albert Einstein, Norbert Wiener, Andrey Kolmogorov, Vincent Doeblin, Kiyosi Itô, Joseph Doob and Paul-André Meyer.



Kiyosi Itô

Andrey Kolmogorov

# Chapter 2

# The Brownian motion

## 2.1 Definition

Let us begin by presenting the definition of stochastic process.

**Definition 2.1** *A stochastic process is a family of random variables* $\{X_t, t \in T\}$ *defined on a probability space* $(\Omega, \mathcal{F}, P)$*, where* $T$ *is the set where the "time" parameter* $t$ *is defined. If* $T = \mathbb{N}$*, we say that the process is a discrete time process; if* $T = [a, b] \subset \mathbb{R}$ *or if* $T = \mathbb{R}$*, we say it is a continuous time process.*

A stochastic process depends on $t \in T$ and on $\omega \in \Omega$, i.e.

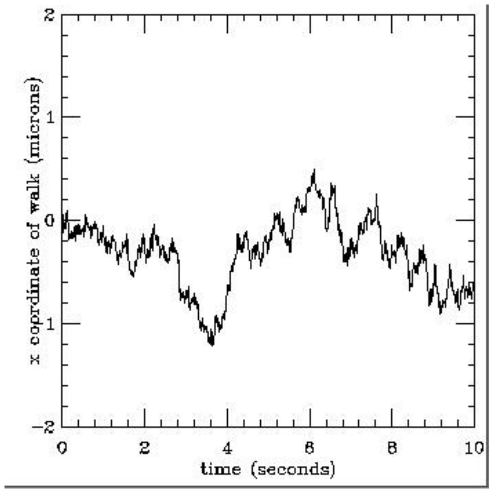
$$\{X_t, t \in T\} = \{X_t(\omega), \omega \in \Omega,\ t \in T\},$$

where $X_t$ is the state or position of the process at time $t$.

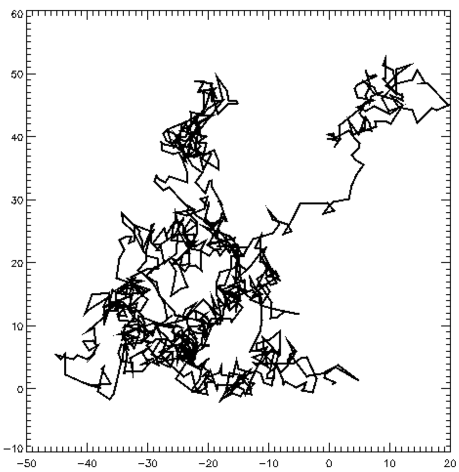
The space of the states $\mathbf{S}$ (space where the random variables take values) is usually $\mathbb{R}$ (continuous state space) or $\mathbb{N}$ (discrete state space).

For each fixed $\omega$ ($\omega \in \Omega$), the mapping $t \to X_t(\omega)$ or $X_.(\omega)$ is called a realization, trajectory or sample path of the process.

As an example of a trajectory, we present below some trajectories of Brownian motion




A trajectory of a one-dimensional
Brownian motion.




A trajectory of a bidimensional
Brownian motion.

The name of the process was given after Robert Brown, a 19th century botanist who first observed the physical motion of grains of polen suspended in water under a microscope. In 1900, Louis Bachelier, in is thesis "Théorie de la spéculation" used the Brownian motion to model financial assets evolution. Albert Einstein, in one of his 1905 papers, used Brownian motion as a tool to indirectly confirm the existence of atoms and molecules.

Now, we present a rigorous definition of the process.

**Definition 2.2** *The standard Brownian motion (also called Wiener Process) is a stochastic process $B = \{B_t; t \geq 0\}$ with state space $\mathbf{S} = \mathbb{R}$ and satisfying the following properties:*

1. $B_0 = 0$.
2. *$B$ has independent increments (i.e. $B_t - B_s$ is independent of $\{B_u, u \leq s\}$ whenever $s < t$).*
3. *$B$ has stationary increments (i.e., the distribution of $B_t - B_s$ depends only on $t - s$).*
4. *$B$ has Gaussian increments (i.e., the distribution of $B_t - B_s$ is the normal distribution $N(0, t-s)$.)*
5. *$B$ has continuous sample paths (i.e. for each fixed $\omega$ ($\omega \in \Omega$), the mapping $t \to X_t(\omega)$ is continuous).*

Some authors consider that the term Brownian motion refers to a process $W = \{W_t; t \geq 0\}$ which satisfies conditions 2,3 e 5 of the previous definition of standard Brownian motion and also condition 4': the distribution of $W_t - W_s$ is $N(\mu(t-s), \sigma^2(t-s))$, where $\mu$ is the drift coefficient and $\sigma$ is the diffusion coefficient. A Brownian motion $W$ with drift $\mu$ and diffusion coefficient $\sigma$ can be constructed from a standard Brownian motion $B$ by:

$$W_t = W_0 + \mu t + \sigma B_t.$$

**Exercise 2.3** *Prove the previous statement, i.e., prove that if $B$ is a standard Brownian motion, then $W_t = W_0 + \mu t + \sigma B_t$ is a Brownian motion with drift $\mu$ and diffusion coefficient $\sigma$.*

It can be difficult to prove that the conditions on the definition of standard Brownian motion are compatible. However, using the Kolmogorov extension theorem (see ) one can show that there exists a stochastic process which satisfies all the conditions of the definition of standard Brownian motion. Condition (4) or condition (5) can be dropped from the definition of standard Brownian motion or Brownian motion, since each of these properties can be shown to be a consequence of the other properties. The Brownian motion is the only process with stationary independent increments and continuous sample paths.

Let us consider a simple symmetric random walk, i. e., a discrete time process defined by

$$X_n = \sum_{i=1}^{n} Z_i, \text{ where } Z_i = \begin{cases} +1 & \text{with probability } \frac{1}{2}. \\ -1 & \text{with probability } \frac{1}{2}. \end{cases}$$

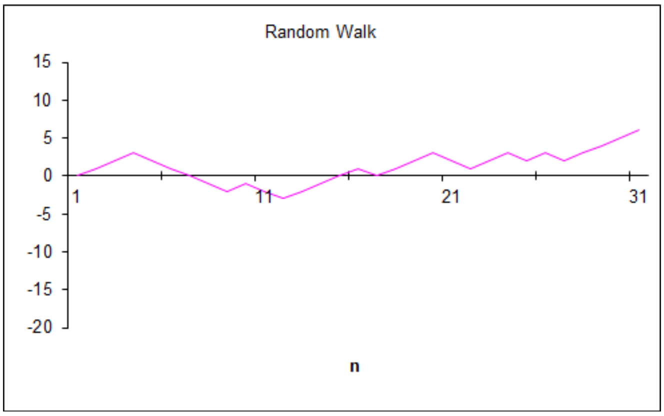



Figure 2.1: A random walk path

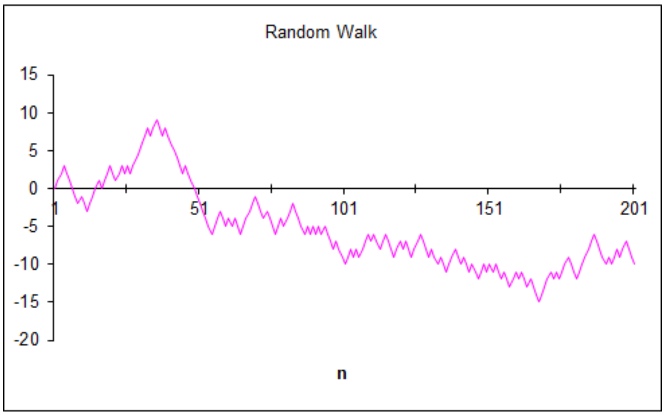



Figure 2.2: Another random walk path

If we reduce the step size progessively from 1 unit until it is infinitesimal (and rescale the values of $X$) then the simple symmetric random walk tends to a Brownian motion.

## 2.2 Main properties of the Brownian motion

In order to prove some properties of the Brownian motion, we can use the following decomposition (for $s < t$):

$$B_t = B_s + (B_t - B_s)\,, \tag{2.1}$$

where $B_s$ is known at time $s$ and $B_t - B_s$ is a random variable independent of the history of the process up until time $s$. In particular, $B_t - B_s$ is inde-

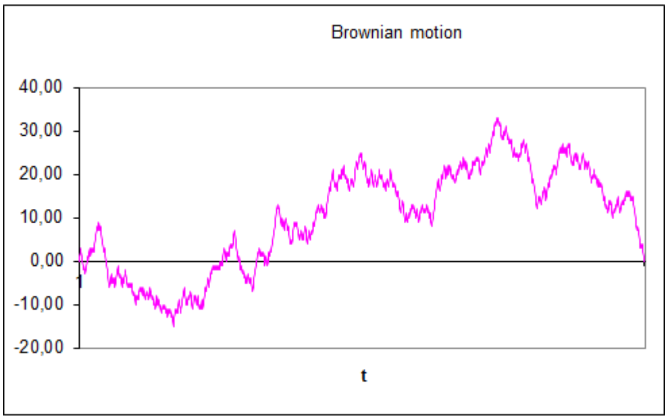



Figure 2.3: A Brownian motion path as a limit of a random walk path

pendent of $B_s$ (this is a consequence of the independent increments property of Brownian motion).

**Proposition 2.4** *Properties of the standard Brownian motion $B$:*

*(1) The standard Brownian motion or the Brownian motion are Gaussian processes.*

*(2) cov $(B_t, B_s) = \min\{t, s\}$.*

*(3) $B$ is a Markov process.*

*(4) $B$ is a martingale.*

*(5) $B$ returns infinitely often to 0 (or to any other level $a \in \mathbb{R}$).*

*(6) (scaling property of Brownian motion or self-similar property): If $B_1$ is a stochastic process defined by $B_1(t) := \frac{1}{\sqrt{c}} B_{ct}$, with $c > 0$, then $B_1$ is also a standard Brownian motion.*

*(7) (time inversion property): If If $B_2$ is a stochastic process defined by $B_2(t) := tB_{(1/t)}$ then $B_2$ is also a standard Brownian motion.*

A Gaussian process is essentially a process where its random variables are Gaussian random variables: this is clear for standard Brownian motion by condition 4. of the definition. A Gaussian stochastic process is completely determined by its expectation and covariance function (as a normal random variable is determined by its expectation and variance). If we know that a stochastic process has Gaussian increments and we know the first two moments of these increments, then we can determine all the statistical properties of the process. Therefore, in order to prove that a Gaussian process is a standard Brownian motion, we only need to compute the expectation and the covariance function of the process and prove that they are equal to zero and equal to the coveriance function given by property (2).

**Proof.** Proof of property (2): Let $s < t$. Then, using (2.1), we obtain

$$\begin{aligned} cov\left(B_t, B_s\right) &= cov\left[B_s + \left(B_t - B_s\right), B_s\right] \\ &= cov\left(B_s, B_s\right) + cov\left(B_t - B_s, B_s\right) \\ &= \mathbb{E}\left[B_s^2\right] + 0 = s. \end{aligned}$$

Proof of property (3): Recall that $X$ is a Markov process if the probability of obtaining a state at time $t$ depends only of the state of the process at the previous last observed instant $t_k$ and not from the previous history, i.e., if $t_1 < t_2 < \cdots < t_k < t$, then

$$\begin{aligned} &P\left[a < X_t < b | X_{t_1} = x_1, X_{t_2} = x_2, \ldots, X_{t_k} = x_k\right] \\ &= P\left[a < X_t < b | X_{t_k} = x_k\right]. \end{aligned}$$

A Markov process with discrete state space is a Markov chain. A Markov process with continuous state space and continuous time is a diffusion process. For the Brownian motion, we have

$$\begin{aligned} &P\left[a < B_t < b | B_{t_1} = x_1, B_{t_2} = x_2, \ldots, B_{t_k} = x_k\right] \\ &= P\left[a - x_k < B_t - B_{t_k} < b - x_k | B_{t_k} = x_k\right], \end{aligned}$$

by the independent increments property of standard Brownian motion (condition 2. of the definition).

Proof of property (6): Clearly, $B_1(0) = \frac{1}{\sqrt{c}} B_{ct} = \frac{1}{\sqrt{c}} B_0 = 0$ and $B_1(t) - B_1(s) = \frac{1}{\sqrt{c}}\left(B_{ct} - B_{cs}\right)$ is independent of $\{B_{cu}, u \leq s\}$ by the independent increments property of the standard Brownian motion. Therefore $B_1(t) - B_1(s)$ is independent of $\{B_1(u), u \leq s\}$ and $B_1$ has independent increments. The distribution of $B_1(t) - B_1(s) = \frac{1}{\sqrt{c}}\left(B_{ct} - B_{cs}\right) \sim \frac{1}{\sqrt{c}} N(0, ct - cs) \sim N(0, t - s)$ and threfore $B_1$ has stationary increments. Moreover, $B_1(t) = \frac{1}{\sqrt{c}} B_{ct}$ has continuous sample paths because $B_{ct}$ has continuous sample paths. ■

**Exercise 2.5** *Prove the time inversion property (property 7) by computing the expectation and the covariance function of* $B_2$.

Another important property of Brownian motion is the following one.

**Proposition 2.6** *Property of non-differentiability of sample paths: The sample paths of a Brownian motion are not differentiable anywhere a.s. (with probability 1).*

We will prove a weaker result: for any fixed time $t_0$, the probability that the sample path of a standard Brownian motion is differentiable at $t_0$ is 0.
**Proof.** Let us assume $t_0 = 0$ (the proof can be generalized to any $t_0$). If $B$ has derivative $a$ at 0 then:

$$a - \delta < \frac{B_t - B_0}{t} < a + \delta$$

for $t$ small enough. This means that (with variable change: $s = \frac{1}{t}$) we have $a - \delta < sB_{(1/s)} < a + \delta$ and by the time inversion property (7), $sB_{(1/s)}$ is a standard Brownian motion, so if we make $t \to 0$ then $s \to +\infty$ and the probability that a standard Brownian motion remains confined between $a - \delta$ and $a + \delta$, when $s \to +\infty$, is zero.

More details about the Brownian motion and its properties can be found in references [2], [5], [7], [9] [10], [11] and [12]. For reviews of probability theory and stochastic processes we refer to [2], [3] and [8]. ■

## 2.3 The geometric Brownian motion

The Brownian motion is not very useful for modeling market prices (at the long run) since it can take negative values and the Brownian motion model would imply that the sizes of price movements are independent of the level of the prices. A more useful and realistic model is the geometric Brownian motion:

$$S_t = e^{W_t},$$

where $W$ is a Brownian motion $W_t = W_0 + \mu t + \sigma B_t$.

$S_t$ is lognormally distributed with mean $W_0 + \mu t$ and variance $\sigma^2 t$, i.e, the $\log(S_t) \sim N(W_0 + \mu t, \sigma^2 t)$. It is also clear that $S_t \geq 0$ for all $t$ and it is easy to prove that

$$\mathbb{E}[S_t] = \exp\left[(W_0 + \mu t) + \frac{1}{2}\sigma^2 t\right],$$
$$\mathrm{var}[S_t] = [\mathbb{E}[S_t]]^2 \left\{\exp\left[\sigma^2 t\right] - 1\right\}.$$

In the Black-Scholes model, the underlying asset price follows geometric Brownian motion. The geometric Brownian motion properties are less helpful than those of Brownian motion: increments of $S$ are neither independent nor stationary.

In order to do some analysis of geometric Brownian motion $S$ one can proceed as follows:

1. Take the logarithm of the observations;

2. Use techniques for the Brownian motion.

As an example, let us consider the log-return of a time series under geometric Brownian motion:

$$\log \frac{S_t}{S_s} = \log \frac{e^{W_t}}{e^{W_s}} = W_t - W_s.$$

Therefore, the log-returns (and the returns themselves) are independent over disjoint time periods.

## 2.4 Martingales in discrete and in continuous time

A martingale is essentially a stochastic process for which its "current value" is the "optimal estimator" of its expected "future value", i.e., given the martingale $\{M_j, j \in \mathbb{N}\}$ and the information $\mathcal{F}_n$ at instant $n$, then $M_n$ is the best estimator for $M_{n+1}$. A martingale has "no drift" and its expected value remains constant in time.

Martingale theory is fundamental in modern financial theory. Indeed, the modern theory of pricing and hedging of financial derivatives is strongly based on martingale properties.

In order to define martingales, let us present the conditional expectation definition and properties. Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathcal{B} \subset \mathcal{F}$ be a $\sigma$-algebra.

**Definition 2.7** *The conditional expectation of the integrable random variable $X$ given $\mathcal{B}$ (or $E(X|\mathcal{B})$) is an integral random variable $Z$ such that*

1. *$Z$ is $\mathcal{B}$-measurable*

2. *For each $A \in \mathcal{B}$, we have*

$$E\left(Z\mathbf{1}_A\right) = E\left(X\mathbf{1}_A\right) \tag{2.2}$$

If $X$ is integrable (i.e., $E\left[|X|\right] < +\infty$) then $Z = E(X|\mathcal{B})$ exists and is unique (a.s.).

**Proposition 2.8** *(Properties of the conditional expectation). Let $X$, $Y$ and $Z$ be integrable random variables, and $a, b \in \mathbb{R}$. Then*

1. $$E(aX + bY|\mathcal{B}) = aE(X|\mathcal{B}) + bE(Y|\mathcal{B}). \tag{2.3}$$

2. $$E\left(E(X|\mathcal{B})\right) = E\left(X\right). \tag{2.4}$$

3. *If $X$ and the $\sigma$-algebra $\mathcal{B}$ are independent then:*
$$E(X|\mathcal{B}) = E\left(X\right) \tag{2.5}$$

4. *If $X$ is $\mathcal{B}$-measurable (or if $\sigma\left(X\right) \subset \mathcal{B}$) then:*
$$E(X|\mathcal{B}) = X. \tag{2.6}$$

5. *If $Y$ is $\mathcal{B}$-measurable (or if $\sigma\left(X\right) \subset \mathcal{B}$) then*
$$E(YX|\mathcal{B}) = YE(X|\mathcal{B}) \tag{2.7}$$

6. *Given two $\sigma$-algebras $\mathcal{C} \subset \mathcal{B}$ then*
$$E(E\left(X|\mathcal{B}\right)|\mathcal{C}) = E(E\left(X|\mathcal{C}\right)|\mathcal{B}) = E(X|\mathcal{C}) \tag{2.8}$$

7. *Consider that $Z$ is $\mathcal{B}$-measurable and $X$ is independent of $\mathcal{B}$. Let $h(x,z)$ be a measurable function such that $h(X,Z)$ is an integrable random variable. Then*
$$E\left(h\left(X,Z\right)|\mathcal{B}\right) = E\left(h\left(X,z\right)\right)|_{z=Z}.$$
*Note: First we compute $E\left(h\left(X,z\right)\right)$ for a $z$ fixed value of $Z$ and then we replace $z$ by $Z$.*

Given several random variables $Y_1, Y_2, \ldots, Y_n$, we can consider the conditional expectation

$$E\left[X|Y_1, Y_2, \ldots, Y_n\right] = E\left[X|\beta\right],$$

where $\beta$ is the $\sigma$-algebra generated by $Y_1, Y_2, \ldots, Y_n$. The $\sigma$-algebra generated by a random variable $X$ is given by sets of the form

$$\sigma\left(X\right) := \left\{X^{-1}(B) : B \in \mathcal{B}_{\mathbb{R}}\right\}.$$

Let $\underline{Y} = (Y_1, Y_2, \ldots, Y_n)$ (notation). Then

$$E\left[E\left[X|\underline{Y}\right]\right] = E\left[X\right].$$

A very important property (it is the reason why conditional expectation is so important) is that $E[X|\underline{Y}]$ is the optimal estimator of $X$ based on $\underline{Y}$ in the sense that for every function $h$, we have:

$$E\left\{(X - E[X|\underline{Y}])^2\right\} \leq E\left\{(X - h(\underline{Y}))^2\right\}. \tag{2.9}$$

**Definition 2.9** *Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\{\mathcal{F}_n, n \geq 0\}$ be a sequence of $\sigma$-algebras such that*

$$\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots \subset \mathcal{F}_n \subset \cdots \subset \mathcal{F} \tag{2.10}$$

*The sequence $\{\mathcal{F}_n, n \geq 0\}$ is called a filtration*

A filtration can be considered as an "information flow".

**Definition 2.10** *The process $M = \{M_n; n \geq 0\}$ (in discrete time) is a martingale with respect to the filtration $\{\mathcal{F}_n, n \geq 0\}$ if:*

1. *For each $n$, $M_n$ is a $\mathcal{F}_n$-measurable random variable (i.e., $M$ is a stochastic process adapted to the filtration $\{\mathcal{F}_n, n \geq 0\}$).*
2. *For each $n$, $E[|M_n|] < \infty$.*
3. *For each $n$, we have:*

$$E[M_{n+1}|\mathcal{F}_n] = M_n. \tag{2.11}$$

If we consider the filtration $\mathcal{F}_n = \sigma(M_0, M_1, \ldots, M_n)$ , then we say that $M = \{M_n; n \geq 0\}$ is a martingale (with respect to this filtration) if

1. For each $n$, $E[|M_n|] < \infty$.
2. For each $n$, we have:

$$E[M_{n+1}|\mathcal{F}_n] = M_n. \tag{2.12}$$

**Proposition 2.11** *It is easy to show that if $M = \{M_n; n \geq 0\}$ is a martingale then*

1. $E[M_n] = E[M_0]$ for all $n \geq 1$.
2. $E[M_n|\mathcal{F}_k] = M_k$ for all $n \geq k$.

**Exercise 2.12** *Prove properties 1. and 2. above.*

The "current value" $M_k$ of a martingale is the "optimal estimator" of its "future value" $M_n$.

The martingale concept allows us to define a risk neutral probability measure in the financial context. If the discounted price of a financial asset is a martingale when calculated using a particular probability distribution, then this probability distribution is called a "risk-neutral" probability measure (meaning that the asset price has no "drift").

**Example 2.13** *Assume that a share has a price process $S_t$ and a discounted price process*

$$\widetilde{S}_t = e^{-rt} S_t, \tag{2.13}$$

*where $r$ is the risk-free interest rate. If we assume that for a probability measure $Q$, the process $\widetilde{S}_t$ is a martingale then, under $Q$, we have that*

$$E_Q \left[ \widetilde{S}_{n+1} | \widetilde{S}_0, \widetilde{S}_1, \ldots, \widetilde{S}_n \right] = \widetilde{S}_n.$$

*Since $\widetilde{S}_n$ is known (it is measurable) with respect to $\sigma\left(\widetilde{S}_0, \widetilde{S}_1, \ldots, \widetilde{S}_n\right)$, then by property (2.6), we have that*

$$E_Q \left[ \frac{e^{-r(n+1)} S_{n+1}}{e^{-rn} S_n} | \widetilde{S}_0, \widetilde{S}_1, \ldots, \widetilde{S}_n \right] = 1$$
$$\Longleftrightarrow E_Q \left[ \frac{S_{n+1}}{S_n} | S_0, S_1, \ldots, S_n \right] = e^r.$$

*Therefore, the expected return in period from time $n$ to time $n+1$ is the risk-free rate: that is why the distribution $Q$ is called risk-neutral measure.*

**Definition 2.14** *Consider the probability space $(\Omega, \mathcal{F}, P)$ and the family of $\sigma$-algebras in continuous time $\{\mathcal{F}_t, t \geq 0\}$, such that*

$$\mathcal{F}_s \subset \mathcal{F}_t, \quad 0 \leq s \leq t. \tag{2.14}$$

*The family $\{\mathcal{F}_t, t \geq 0\}$ is called a filtration in continuous time.*

Let $\mathcal{F}_t^X$ be the $\sigma$-algebra generated by process $X$ on the interval $[0, t]$, i.e. $\mathcal{F}_t^X = \sigma(X_s, 0 \leq s \leq t)$. Then $\mathcal{F}_t^X$ is the "information generated by $X$ on interval $[0, t]$" or the "history of the process $X$ up until time $t$". $A \in \mathcal{F}_t^X$ means that is possible to decide if event $A$ occurred or not, based on the observation of the paths of the process $X$ on $[0, t]$.

**Example 2.15** *If* $A = \{\omega : X(5) > 1\}$ *then* $A \in \mathcal{F}_5^X$ *but* $A \notin \mathcal{F}_4^X$.

A stochastic process $Y$ is said to be adapted to the filtration $\{\mathcal{F}_t, t \geq 0\}$ if $Y_t$ is $\mathcal{F}_t$ measurable for all t. If $\mathcal{F}_t^X = \sigma(X_s, 0 \leq s \leq t)$ is the filtration generated by $X$, then any continuous function of $X_t$ is adapted to $\mathcal{F}_t^X$.

1. $E[X|\mathcal{F}_t]$ is the optimal estimator of $X$ among all $\mathcal{F}_t$-measurable random variables with finite expectation, or equivalently

$$E\{(X - E[X|\mathcal{F}_t])Y\} = 0 \tag{2.15}$$

   for all $\mathcal{F}_t$-measurable bounded random variables $Y$.

2. $E\{E[X|\mathcal{F}_t]\} = E[X]$.
3. If $X$ is $\mathcal{F}_t$-measurable then $E[X|\mathcal{F}_t] = X$.
4. If $Y$ is $\mathcal{F}_t$-measurable and bounded then $E[XY|\mathcal{F}_t] = YE[X|\mathcal{F}_t]$.
5. If $X$ is independent of $\mathcal{F}_t$ then $E[X|\mathcal{F}_t] = E[X]$.

We can now define what is a martingale in continuous time.

**Definition 2.16** *A stochastic process* $M = \{M_t; t \geq 0\}$ *is a martingale with respect to the filtration* $\{\mathcal{F}_t, t \geq 0\}$ *if*

1. *For each* $t \geq 0$, $M_t$ *is a* $\mathcal{F}_t$*-measurable random variable (i.e.,* $M$ *is adapted to* $\{\mathcal{F}_t, t \geq 0\}$*).*
2. *For each* $t \geq 0$, $E[|M_t|] < \infty$.
3. *For each* $s \leq t$,

$$E[M_t|\mathcal{F}_s] = M_s. \tag{2.16}$$

The condition (3) is equivalent to $E[M_t - M_s|\mathcal{F}_s] = 0$. If $t \in [0, T]$ then $M_t = E[M_T|\mathcal{F}_t]$. Moreover, condition (3) implies that $E[M_t] = E[M_0]$ for all $t$.

Consider a standard Brownian motion $B = \{B_t; t \geq 0\}$ defined on $(\Omega, \mathcal{F}, P)$ and

$$\mathcal{F}_t^B = \sigma\{B_s, s \leq t\}. \tag{2.17}$$

**Proposition 2.17** *The following processes are* $\mathcal{F}_t^B$*-martingales:*

1. $B_t$.

2. $B_t^2 - t$.

3. $\exp\left(aB_t - \frac{a^2 t}{2}\right)$.

**Proof.** 1. $B_t$ is $\mathcal{F}_t^B$-measurable and therefore it is adapted. $E[|B_t|] < \infty$. Moreover $B_t - B_s$ is independent of $\mathcal{F}_s^B$. Hence

$$E\left[B_t - B_s | \mathcal{F}_s^B\right] = E\left[B_t - B_s\right] = 0.$$

2. Clearly, $B_t^2 - t$ is $\mathcal{F}_t^B$-measurable and adapted. Moreover $E\left[|B_t^2 - t|\right] < \infty$. By the properties of the conditional expectation, we have

$$\begin{aligned} E\left[B_t^2 - t | \mathcal{F}_s^B\right] &= E\left[(B_t - B_s + B_s)^2 | \mathcal{F}_s^B\right] - t \\ &= E\left[(B_t - B_s)^2\right] + 2B_s E\left[B_t - B_s | \mathcal{F}_s^B\right] + B_s^2 - t \\ &= t - s + B_s^2 - t = B_s^2 - s. \end{aligned}$$

■

**Exercise 2.18** *Prove that the process* $X_t = \exp\left(aB_t - \frac{a^2 t}{2}\right)$ *is a* $\left\{\mathcal{F}_t^B, t \geq 0\right\}$*-martingale.*

# Chapter 3

# The Itô integral

## 3.1 Motivation

A stochastic differential equation (SDE) is a differential equation with "noise" of the type:

$$\frac{dX}{dt} = b(t, X_t) + \sigma(t, X_t) \text{"} \frac{dB_t}{dt} \text{"}.$$

The term "$\frac{dB_t}{dt}$" is a stochastic "noise". Does not exist in a classical sense since $B$ is not differentiable. The stochastic differential equation (SDE) in integral form can be written as

$$X_t = X_0 + \int_0^t b(s, X_s) ds + \text{"} \int_0^t \sigma(s, X_s) dB_s \text{"} \tag{3.1}$$

How to define the integral

$$\int_0^T u_s dB_s, \tag{3.2}$$

where $B$ is a Brownian motion and $u$ is an appropriate adapted process?

The SDE's that we deal with are the continuous time versions of the equations used to define time series (processes in discrete time).

**Example 3.1** *A zero-mean random walk can be defined by:*

$$X_t = X_{t-1} + \sigma Z_t,$$

*where $Z_t$ is a standard normal random variable (the $Z_i$ variables are called "white noise"). This equation is a stochastic difference equation and is equivalent to $\Delta X_t = \sigma Z_t$. Its solution is*

$$X_t = X_0 + \sigma \sum_{s=1}^{t} Z_s.$$

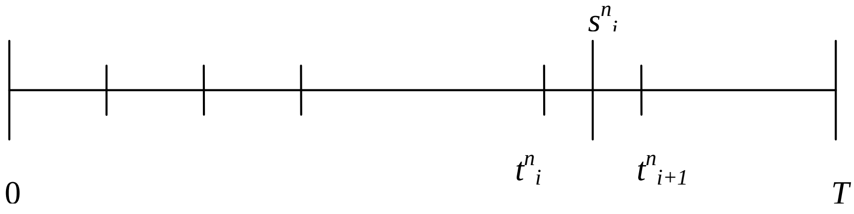



Figure 3.1: The partition

*In continuous time, the analog of a zero-mean random walk is a zero-mean Brownian motion* $W_t$.

In order to define the stochastic integral $\int_0^T u_s dB_s$, one could try to apply the following strategy:

- Consider a sequence of partitions of $[0, T]$ and a sequence of points:

$$\tau_n\colon\ 0 = t_0^n < t_1^n < t_2^n < \cdots < t_{k(n)}^n = T$$
$$s_n\colon\ t_i^n \leq s_i^n \leq t_{i+1}^n, \quad i = 0, \ldots, k(n) - 1,$$

  such that $\lim\limits_{n\to\infty} \sup\limits_i \left(t_{i+1}^n - t_i^n\right) = 0$.

- Consider Riemann-Stieltjes (R-S) integral, defined as the limit of Riemann sums:

$$\int_0^T f dg := \lim_{n\to\infty} \sum_{i=0}^{n-1} f\left(s_i^n\right) \Delta g_i,$$

  where $\Delta g_i := g(t_{i+1}^n) - g(t_i^n)$, if the limit exists and is independent of the sequences $\tau_n$ and $s_n$.

- If $g$ is a differentiable function and $f$ is continuous, the (R-S) integral is well defined and $\int_0^T f(t)\, dg(t) = \int_0^T f(t)\, g'(t)\, dt$.

- The (R-S) integral $\int_0^T f dg$ exists if $f$ is continuous and $g$ has bounded total variation, i.e.,

$$\sup_{\tau_n} \sum_i |\Delta g_i| < \infty.$$

- In the Brownian motion case $B$, it is clear that $B'(t)$ does not exist, so we cannot define the path integral:

$$\int_0^T u_t(\omega)\, dB_t(\omega) \overset{\times}{\neq} \int_0^T u_t(\omega)\, B_t'(\omega)\, dt$$

- Moreover, we can prove that the Brownian motion total variation is not bounded and therefore we cannot define the (R-S) integral $\int_0^T u_t(\omega)\, dB_t(\omega)$ in general.

If $u$ has sample paths of class $C^1$, then integrating by parts, the (R-S) integral exists and we can compute:

$$\int_0^T u_t(\omega)\, dB_t(\omega) = u_T(\omega)\, B_T(\omega) - \int_0^T u_t'(\omega)\, B_t(\omega)\, dt.$$

However, we have a big problem: for instance, the integral $\int_0^T B_t(\omega)\, dB_t(\omega)$ does not exist as a R-S integral. We need to consider processes $u$ with sample paths more irregular than $C^1$ trajectories. How to define the integral (3.2) for these processes? We need to consider a new strategy. We will construct the stochastic integral $\int_0^T u_t dB_t$ using a probabilistic approach.

## 3.2 The Itô integral for simple processes

**Definition 3.2** *Consider processes $u$ of class $L^2_{a,T}$, which is defined as the class of processes $u = \{u_t,\ t \in [0,T]\}$, such that:*

1. *$u$ is adapted and measurable (measurable in the sense that the mapping $(t,\omega) \rightarrow u_t(\omega)$ is measurable on the product space $[0,T] \times \Omega$, with respect to the $\sigma$-algebra $\beta([0,T]) \times \mathcal{F}$).*

2. $E\left[\int_0^T u_t^2 dt\right] < \infty$.

The condition 2. allows us to show that $u$ as a map of two variables $t$ and $\omega$ belongs to the space $L^2([0,T] \times \Omega)$ and that:

$$E\left[\int_0^T u_t^2 dt\right] = \int_0^T E\left[u_t^2\right] dt = \int_{[0,T]\times\Omega} u_t^2(\omega)\, dt P(d\omega).$$

In order to define $\int_0^T u_t dB_t$ for $u \in L^2_{a,T}$, we will consider the limit in mean-square (i.e., a limit in $L^2(\Omega)$) of integrals of simple processes.

**Definition 3.3** *The process $u \in \mathcal{S}$ (set of simple processes in $[0,T]$) is called a simple process if*

$$u_t = \sum_{j=1}^{n} \phi_j \mathbf{1}_{(t_{j-1},t_j]}(t), \tag{3.3}$$

*where $0 = t_0 < t_1 < \cdots < t_n = T$, the random variables $\phi_j$ are square-integrable ($E\left[\phi_j^2\right] < \infty$) and $\mathcal{F}_{t_{j-1}}$-measurable*

**Definition 3.4** *If $u$ is a simple process of form (3.3) ($u \in \mathcal{S}$), then the stochastic Itô integral of $u$ with respect to Brownian motion $B$ is*

$$\int_0^T u_t dB_t := \sum_{j=1}^n \phi_j \left( B_{t_j} - B_{t_{j-1}} \right).$$

**Example 3.5** *Consider the simple process*

$$u_t = \sum_{j=1}^n B_{t_{j-1}} \mathbf{1}_{(t_{j-1}, t_j]}(t).$$

*Then*

$$\int_0^T u_t dB_t = \sum_{j=1}^n B_{t_{j-1}} \left( B_{t_j} - B_{t_{j-1}} \right).$$

*Therefore*

$$\begin{aligned} E\left[ \int_0^T u_t dB_t \right] &= \sum_{j=1}^n E\left[ B_{t_{j-1}} \left( B_{t_j} - B_{t_{j-1}} \right) \right] \\ &= \sum_{j=1}^n E\left[ B_{t_{j-1}} \right] E\left[ B_{t_j} - B_{t_{j-1}} \right] = 0. \end{aligned}$$

**Proposition:** (Isometry property or norm preservation property). Let $u \in \mathcal{S}$. Then:

$$E\left[ \left( \int_0^T u_t dB_t \right)^2 \right] = E\left[ \int_0^T u_t^2 dt \right] = \int_0^T E\left[ u_t^2 \right] dt. \tag{3.4}$$

**Proof.** With $\Delta B_j := B_{t_j} - B_{t_{j-1}}$, we have

$$\begin{aligned} E\left[ \left( \int_0^T u_t dB_t \right)^2 \right] &= E\left[ \left( \sum_{j=1}^n \phi_j \Delta B_j \right)^2 \right] \\ &= \sum_{j=1}^n E\left[ \phi_j^2 \left( \Delta B_j \right)^2 \right] + 2 \sum_{i<j}^n E\left[ \phi_i \phi_j \Delta B_i \Delta B_j \right]. \end{aligned}$$

Note that since $\phi_i \phi_j \Delta B_i$ is $\mathcal{F}_{j-1}$-measurable and $\Delta B_j$ is independent of $\mathcal{F}_{j-1}$, then

$$\sum_{i<j}^n E\left[ \phi_i \phi_j \Delta B_i \Delta B_j \right] = \sum_{i<j}^n E\left[ \phi_i \phi_j \Delta B_i \right] E\left[ \Delta B_j \right] = 0.$$

On the other hand, since $\phi_j^2$ is $\mathcal{F}_{j-1}$-measurable and $\Delta B_j$ is independent of $\mathcal{F}_{j-1}$,

$$\begin{aligned}\sum_{j=1}^{n} E\left[\phi_j^2 (\Delta B_j)^2\right] &= \sum_{j=1}^{n} E\left[\phi_j^2\right] E\left[(\Delta B_j)^2\right] \\ &= \sum_{j=1}^{n} E\left[\phi_j^2\right] (t_j - t_{j-1}) = \\ &= E\left[\int_0^T u_t^2 dt\right].\end{aligned}$$

■

**Proposition 3.6** *Consider that $u$ and $v$ are simple processes. We have the following properties.*

1. *Linearity: If $a, b \in \mathbb{R}$,*

$$\int_0^T (au_t + bv_t)\, dB_t = a \int_0^T u_t dB_t + b \int_0^T v_t dB_t. \tag{3.5}$$

2. *Zero mean:*

$$E\left[\int_0^T u_t dB_t\right] = 0. \tag{3.6}$$

**Exercise 3.7** *Prove the property 2.*

**Exercise 3.8** *Compute $\int_0^5 f(s)\, dB_s$ with $f(s) = 1$ if $0 \leq s \leq 2$ and $f(s) = 4$ if $2 < s \leq 5$. What is the distribution of the resulting random variable?*

## 3.3 The Itô integral for adapted processes

A process $u \in L^2_{a,T}$ can be approximated by a sequence of simple processes, in the sense of the following lemma.

**Lemma 3.9** *If $u \in L^2_{a,T}$ then exists a sequence of simple processes $\left\{u^{(n)}\right\}$ such that*

$$\lim_{n\to\infty} E\left[\int_0^T \left|u_t - u_t^{(n)}\right|^2 dt\right] = 0. \tag{3.7}$$

For a proof of this important Lemma, we refer to [9] or [10].

**Definition 3.10** *The Itô stochastic integral of $u \in L^2_{a,T}$ is defined as the limit (in the $L^2(\Omega)$ sense)*

$$\int_0^T u_t dB_t = \lim_{n\to\infty} (L^2) \int_0^T u_t^{(n)} dB_t, \tag{3.8}$$

*where $\left\{u^{(n)}\right\}$ is a sequence of simple processes satisfying (3.7).*

**Proposition 3.11** *Properties of the Itô integral $\int_0^T u_t dB_t$ for $u \in L^2_{a,T}$.*

1. Isometry (or norm preservation):

$$E\left[\left(\int_0^T u_t dB_t\right)^2\right] = E\left[\int_0^T u_t^2 dt\right] = \int_0^T E\left[u_t^2\right] dt. \tag{3.9}$$

2. Zero mean:

$$E\left[\int_0^T u_t dB_t\right] = 0 \tag{3.10}$$

3. Linearity:

$$\int_0^T (au_t + bv_t)\, dB_t = a\int_0^T u_t dB_t + b\int_0^T v_t dB_t. \tag{3.11}$$

4. The process $\left\{\int_0^t u_s dB_s, t \geq 0\right\}$ is a martingale.

5. The sample paths of $\left\{\int_0^t u_s dB_s, t \geq 0\right\}$ are continuous.

**Example 3.12** *Let us show that*

$$\int_0^T B_t dB_t = \frac{1}{2}B_T^2 - \frac{1}{2}T.$$

*Since $u_t = B_t$, let us consider the sequence of simple processes*

$$u_t^n = \sum_{j=1}^n B_{t_{j-1}^n} \mathbf{1}_{\left(t_{j-1}^n, t_j^n\right]}(t),$$

*with* $t_j^n := \frac{j}{n}T$.

$$\begin{aligned}\int_0^T B_t dB_t &= \lim_{n\to\infty} (L^2) \int_0^T u_t^{(n)} dB_t = \\ &= \lim_{n\to\infty} (L^2) \sum_{j=1}^n B_{t_{j-1}^n}\left(B_{t_j^n} - B_{t_{j-1}^n}\right) \\ &= \lim_{n\to\infty} (L^2) \frac{1}{2}\sum_{j=1}^n \left[\left(B_{t_j^n}^2 - B_{t_{j-1}^n}^2\right) - \left(B_{t_j^n} - B_{t_{j-1}^n}\right)^2\right] \\ &= \frac{1}{2}\left(B_T^2 - T\right),\end{aligned}$$

*where we used:* $E\left[\left(\sum_{j=1}^n \left(\Delta B_{t_j^n}\right)^2 - T\right)^2\right] = 0$ *and* $\frac{1}{2}\sum_{j=1}^n \left(B_{t_j^n}^2 - B_{t_{j-1}^n}^2\right) = \frac{1}{2}B_T^2$.

*Let us prove that* $E\left[\left(\sum_{j=1}^n \left(\Delta B_{t_j^n}\right)^2 - T\right)^2\right] = 0$. *Using the independence of increments and* $E\left[\left(\Delta B_{t_j^n}\right)^2\right] = \Delta t_j^n$, *then*

$$\begin{aligned}E\left[\left(\sum_{j=1}^n \left(\Delta B_{t_j^n}\right)^2 - T\right)^2\right] &= E\left[\left(\sum_{j=1}^n \left[\left(\Delta B_{t_j^n}\right)^2 - \Delta t_j^n\right]\right)^2\right] \\ &= \sum_{j=1}^n E\left[\left(\Delta B_{t_j^n}\right)^2 - \Delta t_j^n\right]^2.\end{aligned}$$

*Using the fact that* $E\left[(B_t - B_s)^{2k}\right] = \frac{(2k)!}{2^k \cdot k!}(t-s)^k$, *then*

$$\begin{aligned}E\left[\left(\sum_{j=1}^n \left(\Delta B_{t_j^n}\right)^2 - T\right)^2\right] &= \sum_{j=1}^n \left[3\Delta t_j^n - 2\left(\Delta t_j^n\right)^2 + \left(\Delta t_j^n\right)^2\right] \\ &= 2\sum_{j=1}^n \left(\Delta t_j^n\right)^2 = 2T \sup_j \left|\Delta t_j^n\right| \underset{n\to\infty}{\to} 0.\end{aligned}$$

Let us note that, by using the formula $E\left[(B_t - B_s)^{2k}\right] = \frac{(2k)!}{2^k \cdot k!}(t-s)^k$, we have that

$$\begin{aligned}Var\left[(\Delta B)^2\right] &= E\left[(\Delta B)^4\right] - \left(E\left[(\Delta B)^2\right]\right)^2 \\ &= 3(\Delta t)^2 - (\Delta t)^2 = 2(\Delta t)^2.\end{aligned}$$

We also know that

$$E\left[(\Delta B)^2\right] = \Delta t.$$

Therefore, if $\Delta t$ is small, the variance of $(\Delta B)^2$ is very small when compared with its expected value. Therefore, when $\Delta t \to 0$ or "$\Delta t = dt$", we have

$$(dB_t)^2 \approx dt. \tag{3.12}$$

For an elementary introduction to stochastic integrals, see [1] and [7]. For detailed presentations of stochastic integration, please see [5], [9], [10] and [11].

# Chapter 4

# Itô's Formula

## 4.1 The One dimensional Itô formula

The Itô formula or Itô's lemma is simply a stochastic version of the classical chain rule. Suppose we have a function of a function $f(b_t)$ and we

consider that $f$ is a $C^2$ class function. In order to find $\frac{d}{dt}f(b_t)$, by Taylor's formula (second order expansion), we obtain

$$\delta f(b_t) = f'(b_t)\,\delta b_t + \frac{1}{2}f''(b_t)(\delta b_t)^2 + \cdots$$

Dividing by $\delta t$ and letting $\delta t \to 0$, we obtain the classical chain rule

$$\frac{d}{dt}f(b_t) = f'(b_t)\frac{db_t}{dt} + \frac{1}{2}f''(b_t)\frac{db_t}{dt}\lim_{\delta t\to 0}(\delta b_t) = f'(b_t)\frac{db_t}{dt},$$

or

$$df(b_t) = f'(b_t)\,db_t.$$

What if we replace the deterministic function $b_t$ by the standard Brownian motion $B_t$? Then, the second order term $\frac{1}{2}f''(B_t)(\delta B_t)^2$ cannot be ignored because $(\delta B_t)^2 \approx (dB_t)^2 \approx dt$ is not of the order $(dt)^2$, that is, we obtain the Itô formula

$$df(B_t) = f'(B_t)\,dB_t + \frac{1}{2}f''(B_t)\,dt. \tag{4.1}$$

**Example 4.1** *Consider the stochastic differential of $B_t^2$. In order to represent this process using a stochastic integral, let $B_t^2 = f(B_t)$ with $f(x) = x^2$. Therefore, by (4.1), we obtain*

$$\begin{aligned} d\left(B_t^2\right) &= 2B_t dB_t + \frac{1}{2}2(dB_t)^2 \\ &= 2B_t dB_t + dt, \end{aligned}$$

*which is the Taylor expansion of $B_t^2$ as a function of $B_t$, assuming that $(dB_t)^2 = dt$. Note that in integral form, the result is equivalent to*

$$\int_0^t B_s dB_s = \frac{1}{2}\left(B_t^2 - t\right).$$

If $f$ is a $C^2$ function, then we can write

$$\begin{aligned} f(B_t) &= \text{stochastic integral+process with differentiable paths} \\ &= \text{Itô process.} \end{aligned}$$

We can replace condition 2) $E\left[\int_0^T u_t^2 dt\right] < \infty$ in the definition of $L_{a,T}^2$ by the weaker condition

$$2') \; P\left[\int_0^T u_t^2 dt < \infty\right] = 1.$$

**Definition 4.2** *Let $L_{a,T}$ be the space of processes that satisfy condition 1 of the definition of $L_{a,T}^2$ and condition 2').*

The Itô integral can be defined for $u \in L_{a,T}$ but, in this case, the stochastic integral may fail to have zero expected value and the Itô isometry may fail to be verified.

**Definition 4.3** *Define $L_{a,T}^1$ as the space of processes $v$ such that:*

1. *$v$ is an adapted and measurable process ($v_t$ is $\{\mathcal{F}_t\}$-adapted, and the map $(t,\omega) \to u_t(\omega)$, defined on $[0,T] \times \Omega$ is measurable with respect to the $\sigma$-algebra $\mathcal{B}([0,T]) \times \mathcal{F}$).*

2. *$P\left[\int_0^T |v_t|\, dt < \infty\right] = 1$.*

**Definition 4.4** *An adapted and continuous process $X = \{X_t, 0 \leq t \leq T\}$ is called an Itô process if it satisfies the decomposition:*

$$X_t = X_0 + \int_0^t u_s dB_s + \int_0^t v_s ds, \tag{4.2}$$

*where $u \in L_{a,T}$ and $v \in L_{a,T}^1$.*

**Theorem 4.5** *(One-dimensional Itô's formula or Itô's lemma): Let* $X = \{X_t, 0 \le t \le T\}$ *be a Itô process of type (4.2). Let* $f(t,x)$ *be a* $C^{1,2}$ *function. Then* $Y_t = f(t, X_t)$ *is an Itô process and we have*

$$f(t,X_t) = f(0,X_0) + \int_0^t \frac{\partial f}{\partial t}(s,X_s)\,ds + \int_0^t \frac{\partial f}{\partial x}(s,X_s)\,u_s dB_s$$
$$+ \int_0^t \frac{\partial f}{\partial x}(s,X_s)\,v_s ds + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(s,X_s)\,u_s^2 ds.$$

In differential form, the Itô formula is given by

$$df(t,X_t) = \frac{\partial f}{\partial t}(t,X_t)\,dt + \frac{\partial f}{\partial x}(t,X_t)\,dX_t$$
$$+ \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t,X_t)\,(dX_t)^2.$$

where $(dX_t)^2$ can be computed using (4.2) and the table of products

| $\times$ | $dB_t$ | $dt$ |
|---|---|---|
| $dB_t$ | $dt$ | 0 |
| $dt$ | 0 | 0 |

The Itô formula for $f(t,x)$ and $X_t = B_t$, or $Y_t = f(t, B_t)$ is given by

$$f(t,B_t) = f(0,0) + \int_0^t \frac{\partial f}{\partial t}(s,B_s)\,ds + \int_0^t \frac{\partial f}{\partial x}(s,B_s)\,dB_s$$
$$+ \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(s,B_s)\,ds.$$

or (in differential form)

$$df(t,B_t) = \frac{\partial f}{\partial t}(t,B_t)\,dt + \frac{\partial f}{\partial x}(t,B_t)\,dB_t$$
$$+ \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t,B_t)\,dt.$$

The Itô formula for $f(x)$ and $X_t = B_t$, or $Y_t = f(B_t)$ is given by

$$df(B_t) = \frac{\partial f}{\partial x}(B_t)\,dB_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(B_t)\,dt.$$

## 4.2 The multidimensional Itô formula

Suposse that $B_t := (B_t^1, B_t^2, \ldots, B_t^m)$ is an $m$-dimensional standard Brownian motion, that is, components $B_t^k$, $k = 1, ..., m$ are one-dimensional independent standard Brownian motion. Consider a Itô process of dimension $n$, defined by

$$\begin{aligned} X_t^1 &= X_0^1 + \int_0^t u_s^{11} dB_s^1 + \cdots + \int_0^t u_s^{1m} dB_s^m + \int_0^t v_s^1 ds, \\ X_t^2 &= X_0^2 + \int_0^t u_s^{21} dB_s^1 + \cdots + \int_0^t u_s^{2m} dB_s^m + \int_0^t v_s^2 ds, \\ &\vdots \\ X_t^n &= X_0^n + \int_0^t u_s^{n1} dB_s^1 + \cdots + \int_0^t u_s^{nm} dB_s^m + \int_0^t v_s^n ds. \end{aligned}$$

In differential form, we can write

$$dX_t^i = \sum_{j=1}^m u_t^{ij} dB_t^j + v_t^i dt,$$

with $i = 1, 2, \ldots, n$. Or, in compact form:

$$dX_t = u_t dB_t + v_t dt,$$

where $v_t$ is $n$-dimensional and $u_t$ is a $n \times m$ matrix of processes. We assume that the components of $u$ belong to $L_{a,T}$ and the components of $v$ belong to $L^1_{a,T}$.

**Theorem 4.6** *(Multidimensional Itô formula or Itô's lemma): If $f : [0,T] \times \mathbb{R}^n \to \mathbb{R}^p$ is a $C^{1,2}$ function, then $Y_t = f(t, X_t)$ is a Itô process and we have the Itô formula:*

$$\begin{aligned} dY_t^k = {} & \frac{\partial f_k}{\partial t}(t, X_t)\, dt + \sum_{i=1}^n \frac{\partial f_k}{\partial x_i}(t, X_t)\, dX_t^i \\ & + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 f_k}{\partial x_i \partial x_j}(t, X_t)\, dX_t^i dX_t^j. \end{aligned}$$

The product of the differentials $dX_t^i dX_t^j$ is computed following the product rules

$$\begin{aligned} dB_t^i dB_t^j &= \begin{cases} 0 & \text{se } i \neq j \\ dt & \text{se } i = j \end{cases}, \\ dB_t^i dt &= 0, \\ (dt)^2 &= 0. \end{aligned}$$

If $B_t$ is a $n$-dimensional standard Brownian motion and $f : \mathbb{R}^n \to \mathbb{R}$ is a $C^2$ function with $Y_t = f(B_t)$, then

$$f(B_t) = f(B_0) + \sum_{i=1}^{n} \int_0^t \frac{\partial f}{\partial x_i}(B_t)\, dB_s^i + \frac{1}{2} \int_0^t \left( \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2}(B_t) \right) ds.$$

**Example 4.7** *(Integration by parts formula) If $X_t^1$ and $X_t^2$ are Itô processes and $Y_t = X_t^1 X_t^2$, then by Itô's formula applied to $f(x) = f(x_1, x_2) = x_1 x_2$, we get*

$$d\left(X_t^1 X_t^2\right) = X_t^2 dX_t^1 + X_t^1 dX_t^2 + dX_t^1 dX_t^2.$$

*That is:*

$$X_t^1 X_t^2 = X_0^1 X_0^2 + \int_0^t X_s^2 dX_s^1 + \int_0^t X_s^1 dX_s^2 + \int_0^t dX_s^1 dX_s^2.$$

**Example 4.8** *Consider the process*

$$Y_t = \left(B_t^1\right)^2 + \left(B_t^2\right)^2 + \cdots + \left(B_t^n\right)^2.$$

*Represent this process in terms of Itô stochastic integrals with respect to $n$-dimensional standard Brownian motion. By the multidimensional Itô formula applied to $f(x) = f(x_1, x_2, \ldots, x_n) = x_1^2 + \cdots + x_n^2$, we obtain*

$$\begin{aligned} dY_t &= 2B_t^1 dB_t^1 + \cdots + 2B_t^n dB_t^n \\ &\quad + ndt. \end{aligned}$$

*That is:*

$$Y_t = 2\int_0^t B_s^1 dB_s^1 + \cdots + 2\int_0^t B_s^n dB_s^n + nt.$$

**Exercise 4.9** *Let $B_t := (B_t^1, B_t^2)$ be a two dimensional Brownian motion. Represent the process*

$$Y_t = \left(B_t^1 t, \left(B_t^2\right)^2 - B_t^1 B_t^2\right)$$

*as an Itô process.*

**Solution 4.10** *By the multidimensional Itô's formula applied to $f(t, x) = f(t, x_1, x_2) = (x_1 t, x_2^2 - x_1 x_2)$, we obtain*

$$\begin{aligned} dY_t^1 &= B_t^1 dt + t dB_t^1, \\ dY_t^2 &= -B_t^2 dB_t^1 + \left(2B_t^2 - B_t^1\right) dB_t^2 + dt \end{aligned}$$

*that is*

$$Y_t^1 = \int_0^t B_s^1 ds + \int_0^t s dB_s^1,$$
$$Y_t^2 = -\int_0^t B_s^2 dB_s^1 + \int_0^t \left(2B_s^2 - B_s^1\right) dB_s^2 + t.$$

**Exercise 4.11** *Assume that a process $X_t$ satisfies the SDE*

$$dX_t = \sigma(X_t) dB_t + \mu(X_t) dt.$$

*Compute the stochastic differential of the process $Y_t = X_t^3$ and represent this process as an Itô process.*

We now present a sketch of the proof of the Itô formula. Consider the process

$$Y_t = f(0, X_0) + \int_0^t \frac{\partial f}{\partial t}(s, X_s) ds + \int_0^t \frac{\partial f}{\partial x}(s, X_s) u_s dB_s$$
$$+ \int_0^t \frac{\partial f}{\partial x}(s, X_s) v_s ds + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(s, X_s) u_s^2 ds.$$

This is an Itô process. We assume that $f$ and its partial derivatives are bounded (the general case can be proved approximating $f$ by bounded functions with bounded derivatives). The Itô stochastic integral can be approximated by a sequence of stochastic integrals of simple processes and so we can assume that $u$ and $v$ are simple processes.

Consider a partition of $[0, t]$ into $n$ equal sub-intervals:

$$f(t, X_t) = f(0, X_0) + \sum_{k=0}^{n-1} \left(f\left(t_{k+1}, X_{t_{k+1}}\right) - f(t_k, X_{t_k})\right).$$

By Taylor formula,

$$f\left(t_{k+1}, X_{t_{k+1}}\right) - f(t_k, X_{t_k}) = \frac{\partial f}{\partial t}(t_k, X_{t_k}) \Delta t + \frac{\partial f}{\partial x}(t_k, X_{t_k}) \Delta X_k$$
$$+ \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t_k, X_{t_k}) (\Delta X_k)^2 + Q_k,$$

where $Q_k$ is the remainder or error. We also have that

$$\Delta X_k = X_{t_{k+1}} - X_{t_k} = \int_{t_k}^{t_{k+1}} v_s ds + \int_{t_k}^{t_{k+1}} u_s dB_s$$
$$= v(t_k) \Delta t + u(t_k) \Delta B_k + S_k,$$

where $S_k$ is the remainder or error. Therefore:

$$\begin{aligned}(\Delta X_k)^2 &= (v(t_k))^2 (\Delta t)^2 + (u(t_k))^2 (\Delta B_k)^2 \\ &\quad + 2v(t_k)\, u(t_k)\, \Delta t \Delta B_k + P_k,\end{aligned}$$

where $P_k$ is the remainder or error term. If we replace all this terms, we obtain

$$f(t, X_t) - f(0, X_0) = I_1 + I_2 + I_3 + \frac{1}{2} I_4 + \frac{1}{2} K_1 + K_2 + R,$$

where

$$\begin{aligned}I_1 &= \sum_k \frac{\partial f}{\partial t}(t_k, X_{t_k})\, \Delta t, \\ I_2 &= \sum_k \frac{\partial f}{\partial t}(t_k, X_{t_k})\, v(t_k)\, \Delta t, \\ I_3 &= \sum_k \frac{\partial f}{\partial x}(t_k, X_{t_k})\, u(t_k)\, \Delta B_k, \\ I_4 &= \sum_k \frac{\partial^2 f}{\partial x^2}(t_k, X_{t_k})\, (u(t_k))^2 (\Delta B_k)^2 .\end{aligned}$$

$$\begin{aligned}K_1 &= \sum_k \frac{\partial^2 f}{\partial x^2}(t_k, X_{t_k})\, (v(t_k))^2 (\Delta t)^2 , \\ K_2 &= \sum_k \frac{\partial^2 f}{\partial x^2}(t_k, X_{t_k})\, v(t_k)\, u(t_k)\, \Delta t \Delta B_k, \\ R &= \sum_k (Q_k + S_k + P_k) .\end{aligned}$$

When $n \to \infty$, it is easy to show that

$$\begin{aligned}I_1 &\to \int_0^t \frac{\partial f}{\partial t}(s, X_s)\, ds, \\ I_2 &\to \int_0^t \frac{\partial f}{\partial x}(s, X_s)\, v_s ds, \\ I_3 &\to \int_0^t \frac{\partial f}{\partial x}(s, X_s)\, u_s dB_s.\end{aligned}$$

As we have seeen before (quadratic variation of standard Brownian motion), we have that

$$\sum_k (\Delta B_k)^2 \to t,$$

hence

$$I_4 \to \int_0^t \frac{\partial^2 f}{\partial x^2}(s, X_s)\, u_s^2 ds.$$

On the other hand, we also have

$$K_1 \to 0,$$
$$K_2 \to 0.$$

It is also possible to show (but more technical and hard) that

$$R \to 0.$$

Therefore, in the limit, when $n \to \infty$, we obtain the one-dimensional Itô's formula.

## 4.3 The martingale representation theorem

Let $u \in L^2_{a,T}$ ($u$ adapted, measurable and squared integrable) and let

$$M_t = \mathbb{E}[M_0] + \int_0^t u_s dB_s. \tag{4.3}$$

The process $M$ is a $\mathcal{F}_t$-martingale. We can also that any squared integrable martingale has the form (4.3).

**Theorem 4.12** *(Itô integral representation): Let $F \in L^2(\Omega, \mathcal{F}_T, P)$. Then, exists a unique process $u \in L^2_{a,T}$ such that*

$$F = \mathbb{E}[F] + \int_0^t u_s dB_s. \tag{4.4}$$

**Proof.**

1. Assume that

$$F = \exp\left(\int_0^T h(s)\, dB_s - \frac{1}{2}\int_0^T h(s)^2\, ds\right), \tag{4.5}$$

with $h$ a deterministic function and $\int_0^T h(s)^2\, ds < \infty$. Applying the Itô formula with $f(x) = e^x$, $X_t = \int_0^t h(s)\, dB_s - \frac{1}{2}\int_0^t h(s)^2\, ds$ and $Y_t = f(X_t)$, we obtain

$$\begin{aligned} dY_t &= Y_t\left(h(t)\, dB_t - \frac{1}{2}h(t)^2\, dt\right) + \frac{1}{2}Y_t\,(h(t)\, dB_t)^2 \\ &= Y_t h(t)\, dB_t. \end{aligned}$$

Hence,

$$Y_t = 1 + \int_0^t Y_s h(s)\, dB_s.$$

and

$$\begin{aligned} F = Y_T &= 1 + \int_0^T Y_s h(s)\, dB_s \\ &= \mathbb{E}[F] + \int_0^T Y_s h(s)\, dB_s \end{aligned}$$

Note that

$$\mathbb{E}\left[\int_0^T (Y_s h(s))^2\, ds\right] < \infty,$$

since $\mathbb{E}[Y_t^2] = \exp\left(\int_0^t h(u)^2\, du\right) < \infty$. Therefore

$$\begin{aligned} \mathbb{E}\left[\int_0^T (Y_s h(s))^2\, ds\right] &\leq \int_0^T \exp\left(\int_0^s h(u)^2\, du\right) h(s)^2\, ds \\ &\leq \exp\left(\int_0^T h(u)^2\, du\right) \int_0^T h(s)^2\, ds. \end{aligned}$$

2. The representation (4.4) also holds (by the linear property) for linear combinations of random variables of the form (4.5). In general, $F \in L^2(\Omega, \mathcal{F}_T, P)$ can be approximates (in the mean square sense) by a sequence $\{F_n\}$ of linear combinations of random variables of the type (4.5). Therefore

$$F_n = \mathbb{E}[F_n] + \int_0^t u_s^{(n)} dB_s.$$

By the Itô isometry, we have that

$$\begin{aligned} \mathbb{E}\left[(F_n - F_m)^2\right] &= (\mathbb{E}[F_n - F_m])^2 + \mathbb{E}\left[\int_0^t \left(u_s^{(n)} - u_s^{(m)}\right)^2 ds\right] \\ &\geq \mathbb{E}\left[\int_0^t \left(u_s^{(n)} - u_s^{(m)}\right)^2 ds\right]. \end{aligned}$$

$\{F_n\}$ is a Cauchy sequence in $L^2(\Omega, \mathcal{F}_T, P)$. Hence

$$\mathbb{E}\left[(F_n - F_m)^2\right] \longrightarrow 0 \text{ when. } n, m \to \infty.$$

Therefore

$$\mathbb{E}\left[\int_0^t \left(u_s^{(n)} - u_s^{(m)}\right)^2 ds\right] \longrightarrow 0 \text{ when } n, m \to \infty.$$

and $\left\{u^{(n)}\right\}$ is a Cauchy sequence in $L^2([0,T]\times\Omega)$. Since this is a complete space, $u^{(n)}\to u$ in $L^2([0,T]\times\Omega)$. The process $u$ is adapted because $u^{(n)}\in L^2_{a,T}$ and exists a subsequence $\left\{u^{(n)}(t,\omega)\right\}$ that converges to $u(t,\omega)$ for a.a. $(t,\omega)\in[0,T]\times\Omega$. Therefore, $u(t,\cdot)$ is $\mathcal{F}_t$-measurable for a.a. $t$. Modifying this process $u$ in a set of zero measure in the $t$ variable, we obtain an adapted process $u$.

We have that

$$\lim_{n\to\infty}\mathbb{E}\left[(F_n-F)^2\right]=\lim_{n\to\infty}\mathbb{E}\left(\mathbb{E}[F_n]+\int_0^T u_s^{(n)}dB_s-F\right)^2=0.$$

On the other hand, by Itô isometry, we obtain

$$\lim_{n\to\infty}\mathbb{E}\left(\mathbb{E}[F_n]-\mathbb{E}[F]\right)^2=0$$
$$\lim_{n\to\infty}\mathbb{E}\left(\int_0^T\left(u_s^{(n)}-u_s\right)dB_s\right)^2=\lim_{n\to\infty}\mathbb{E}\int_0^T\left(u_s^{(n)}-u_s\right)^2ds=0.$$

Therefore, $F=\mathbb{E}[F]+\int_0^T u_s dB_s$.

3. Uniqueness: suppose that $u^{(1)}$,$u^{(2)}\in L^2_{a,T}$ and

$$F=\mathbb{E}[F]+\int_0^T u_s^{(1)}dB_s=\mathbb{E}[F]+\int_0^T u_s^{(2)}dB_s.$$

By Itô isometry,

$$\mathbb{E}\left[\left(\int_0^T\left(u_s^{(1)}-u_s^{(2)}\right)dB_s\right)^2\right]=\mathbb{E}\left[\int_0^T\left(u_s^{(1)}-u_s^{(2)}\right)^2ds\right]=0.$$

Hence

$$u^{(1)}(t,\omega)=u(t,\omega)^{(2)}\quad\text{a.a. }(t,\omega)\in[0,T]\times\Omega.$$

■

**Theorem 4.13** *(Martingale representation theorem): Let $\{M_t,t\in[0,T]\}$ be a $\{\mathcal{F}_t\}$-martingale and $\mathbb{E}[M_T^2]<\infty$. Then exists a unique process $u\in L^2_{a,T}$ such that*

$$M_t=\mathbb{E}[M_0]+\int_0^t u_s dB_s\qquad\forall t\in[0,T].$$

**Proof.** By the Itô integral representation theorem applied to $F = M_T$, $\exists^1 u \in L^2_{a,T}$ such that

$$M_T = \mathbb{E}\left[M_T\right] + \int_0^T u_s dB_s.$$

Since $\{M_t, t \in [0,T]\}$ is a martingale, $\mathbb{E}\left[M_T\right] = \mathbb{E}\left[M_0\right]$ and

$$\begin{aligned} M_t &= \mathbb{E}\left[M_T|\mathcal{F}_t\right] = \mathbb{E}\left[\mathbb{E}\left[M_T\right]|\mathcal{F}_t\right] + \mathbb{E}\left[\int_0^T u_s dB_s|\mathcal{F}_t\right] \\ &= \mathbb{E}\left[M_0\right] + \int_0^t u_s dB_s. \end{aligned}$$

where we have used the martingale property for the stochastic integral. ■

**Example 4.14** *Let $F = B_T^3$. What is the Itô integral representation of this random variable? By the Itô formula with $f(x) = x^3$ and $B_T^3 = f(B_t)$, we obtain*

$$B_T^3 = \int_0^T 3B_t^2 dB_t + 3\int_0^T B_t dt.$$

*Integrating by parts,*

$$\int_0^T B_t dt = TB_T - \int_0^T t dB_t = \int_0^T (T-t)\, dB_t.$$

*Therefore,*

$$F = B_T^3 = \int_0^T 3\left[B_t^2 + (T-t)\right] dB_t. \tag{4.6}$$

*Clearly $E\left[B_T^3\right] = 0$ (since $B_T \sim N(0,T)$). Therefore, the integral representation is (4.6).*

**Exercise 4.15** *What is the process $u$ such that $\int_0^T tB_t^2 dt - \frac{T^2}{2}B_T^2 = -\frac{T^3}{6} + \int_0^T u_t dB_t$ ?*

**Solution 4.16** *Applying the Itô formula to $X_t = f(t, B_t) = t^2 B_t^2$, with $f(t,x) = t^2 x^2$, we obtain*

$$T^2 B_T^2 = \int_0^T 2tB_t^2 dt + \int_0^T 2t^2 B_t dB_t + \int_0^T t^2 dt.$$

*Hence*

$$\int_0^T tB_t^2 dt - \frac{T^2}{2}B_T^2 = -\frac{T^3}{6} - \int_0^T t^2 B_t dB_t$$

*and therefore*

$$u_t = -t^2 B_t.$$

*Note that* $\mathbb{E}\left[\int_0^T tB_t^2 dt - \frac{T^2}{2}B_T^2\right] = -\frac{T^3}{6}$.

In general, the integration by parts formula is the one stated in the following exercise.

**Exercise 4.17** *(Integration by parts): Assume that* $f(s)$ *is a deterministic function of class* $C^1$*. Prove that*

$$\int_0^t f(s)\, dB_s = f(t)\, B_t - \int_0^t f'(s)\, B_s ds.$$

**Solution 4.18** *This formula can be deduced by the Itô formula applied to* $g(t,x) = f(t)\, x$*, which results in*

$$f(t)\, B_t = \int_0^t f'(s)\, B_s ds + \int_0^t f(s)\, dB_s.$$

# Chapter 5

# Stochastic Differential Equations

## 5.1 Itô processes and diffusions

Consider a deterministic ordinary differential equation (ODE)

$$f\left(t, x\left(t\right), x'\left(t\right), x''\left(t\right), \ldots\right) = 0, \qquad 0 \leq t \leq T.$$

The first order ODE can be represented by

$$\frac{dx\left(t\right)}{dt} = \mu\left(t, x\left(t\right)\right),$$

or

$$dx\left(t\right) = \mu\left(t, x\left(t\right)\right) dt.$$

A discrete version of this equation is

$$\Delta x\left(t\right) = x\left(t + \Delta t\right) - x\left(t\right) \approx \mu\left(t, x\left(t\right)\right) \Delta t.$$

**Example 5.1** *The first order linear homogeneous ODE is*

$$\frac{dx\left(t\right)}{dt} = cx\left(t\right)$$

*and has solution*

$$x\left(t\right) = x\left(0\right) e^{ct}.$$

A stochastic differential equation (SDE) has the general form

$$\begin{aligned} dX_t &= \mu\left(t, X_t\right) dt + \sigma\left(t, X_t\right) dB_t, \qquad (5.1) \\ X_0 &= X_0, \end{aligned}$$

where $\mu(t, X_t)$ is the drift coefficient and $\sigma(t, X_t)$ is the diffusion coefficient. The SDE can also be written in the integral form

$$X_t = X_0 + \int_0^t \mu(s, X_s)\, ds + \int_0^t \sigma(s, X_s)\, dB_s. \tag{5.2}$$

A "naif" interpretation of SDE is that the increment $\Delta X_t \approx \mu(t, X_t)\Delta t + \sigma(t, X_t)\Delta B_t$. and the distribution of this increment can be approximated by $\Delta X_t \sim N\left(\mu(t, X_t)\Delta t, (\sigma(t, X_t))^2 \Delta t\right)$, when $\Delta t$ is very small.

**Definition 5.2** *A solution of the SDE (5.1) or (5.2) is a stochastic process $\{X_t\}$ which satisfies:*

1. *$\{X_t\}$ is an adapted process (to the standard Brownian motion) and has continuous sample paths.*
2. $\mathbb{E}\left[\int_0^T (\sigma(s, X_s))^2\, ds\right] < \infty$.
3. *$\{X_t\}$ satisfies the SDE (5.1) or (5.2)*

The solutions of stochastic differential equations are called diffusions or "diffusion processes".

**Definition 5.3** *The process $\{X_t, t \geq 0\}$ is said to be a time-homogeneous diffusion process if*

- 1. it is a Markov process.
  2. it has continuous sample paths.
  3. there exist functions $\mu(x)$ and $\sigma^2(x) > 0$ such that, as $h \to 0^+$,

$$\begin{aligned} E\left[X_{t+h} - X_t | X_t = x\right] &= h\mu(x) + o(h), \\ E\left[(X_{t+h} - X_t)^2 | X_t = x\right] &= h\sigma^2(x) + o(h), \\ E\left[(X_{t+h} - X_t)^3 | X_t = x\right] &= o(h). \end{aligned}$$

A diffusion is "locally" like a Brownian motion with drift, but with a variable drift coefficient $\mu(x)$ and diffusion coefficient $\sigma(x)$. Fitting a diffusion model involves estimating the drift function $\mu(x)$ and the diffusion function $\sigma(x)$. Estimating arbitrary drift and diffusion coefficients is virtually impossible unless a very large quantity of data is to hand. Usually, a parametric form is specified for the mean and the variance and the parameters are estimated from data.

## 5.2 The existence and uniqueness theorem

The following theorem gives sufficient conditions to ensure that a unique solution exists for a stochastic differential equation. For a proof of this theorem, we refer to [10].

**Theorem 5.4** *Let $T > 0$, $\mu(\cdot,\cdot) : [0,T] \times \mathbb{R} \to \mathbb{R}$ and $\sigma(\cdot,\cdot) : [0,T] \times \mathbb{R} \to \mathbb{R}$ be measurable functions such that*

1. $\mathbb{E}\left[|X_0|^2\right] < \infty$. *and $X_0$ is independent of the standard Brownian motion $B$.*

2. *(linear growth condition)*

$$|\mu(t,x)| + |\sigma(t,x)| \leq C(1+|x|), x \in \mathbb{R}, t \in [0,T].$$

3. *(Lipschitz condition)*

$$|\mu(t,x) - \mu(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leq D|x-y|, \ x,y \in \mathbb{R}, \ t \in [0,T].$$

*Then the SDE*

$$X_t = X_0 + \int_0^t \mu(s,X_s)\,ds + \int_0^t \sigma(s,X_s)\,dB_s \tag{5.3}$$

*has a unique solution. That is, exists a unique stochastic process $X = \{X_t, 0 \leq t \leq T\}$ continuous and adapted, which satisfies (5.3) and*

$$E\left[\int_0^T |X_s|^2\,ds\right] < \infty.$$

## 5.3 The geometric Brownian motion and the OU process

**Example 5.5** *The standard model for a risky asset price is the SDE*

$$dS_t = \alpha S_t dt + \sigma S_t dB_t \tag{5.4}$$

*or*

$$S_t = S_0 + \alpha \int_0^t S_s ds + \sigma \int_0^t S_s dB_s \tag{5.5}$$

*How to solve this SDE?*

*Assume that* $S_t = f(t, B_t)$ *with* $f \in C^{1,2}$*. By Itô formula:*

$$S_t = f(t, B_t) = S_0 + \int_0^t \left( \frac{\partial f}{\partial t}(s, B_s) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(s, B_s) \right) ds + \tag{5.6}$$

$$+ \int_0^t \frac{\partial f}{\partial x}(s, B_s)\, dB_s.$$

*Comparing (5.5) with (5.6) then, by the uniqueness of representation as an itô process, we have*

$$\frac{\partial f}{\partial s}(s, B_s) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(s, B_s) = \alpha f(s, B_s), \tag{5.7}$$

$$\frac{\partial f}{\partial x}(s, B_s) = \sigma f(s, B_s). \tag{5.8}$$

*Differentiating (5.8), we get*

$$\frac{\partial^2 f}{\partial x^2}(s, x) = \sigma \frac{\partial f}{\partial x}(s, x) = \sigma^2 f(s, x)$$

*and replacing in (5.7) we obtain that*

$$\left(\alpha - \frac{1}{2}\sigma^2\right) f(s, x) = \frac{\partial f}{\partial s}(s, x)$$

*Separating the variables:* $f(s, x) = g(s)\, h(x)$*, we get*

$$\frac{\partial f}{\partial s}(s, x) = g'(s)\, h(x)$$

*and*

$$g'(s) = \left(\alpha - \frac{1}{2}\sigma^2\right) g(s)$$

*which is a linear ODE, with solution:*

$$g(s) = g(0) \exp\left[\left(\alpha - \frac{1}{2}\sigma^2\right) s\right]$$

*Using (5.8),* $h'(x) = \sigma h(x)$ *and*

$$f(s, x) = f(0, 0) \exp\left[\left(\alpha - \frac{1}{2}\sigma^2\right) s + \sigma x\right].$$

*Hence*

$$S_t = f(t, B_t) = S_0 \exp\left[\left(\alpha - \frac{1}{2}\sigma^2\right) t + \sigma B_t\right] \tag{5.9}$$

*which is the geometric Brownian motion. Therefore,* $\frac{S_t}{S_0}$ *has lognormal distribution with parameters* $\left(\alpha - \frac{1}{2}\sigma^2\right) t$ *and* $\sigma^2 t$*. Moreover*

$$E\left[\frac{S_t}{S_0}\right] = e^{\alpha t},\ var\left[\frac{S_t}{S_0}\right] = e^{2\alpha t}\left(e^{\sigma^2 t} - 1\right).$$

*Note that the solution of the SDE was obtained by solving a deterministic PDE (partial differential equation).*

Let us verify that (5.9) satisfies SDE (5.4) or (5.5). Apllying the Itô formula to $S_t = f(t, B_t)$ with

$$f(t,x) = S_0 \exp\left[\left(\alpha - \frac{1}{2}\sigma^2\right) t + \sigma x\right],$$

we obtain

$$\begin{aligned} S_t &= S_0 + \int_0^t \left[\left(\alpha - \frac{1}{2}\sigma^2\right) S_s + \frac{1}{2}\sigma^2 S_s\right] ds + \int_0^t \sigma S_s dB_s \\ &= S_0 + \alpha \int_0^t S_s ds + \sigma \int_0^t S_s dB_s \end{aligned}$$

or:

$$dS_t = \alpha S_t dt + \sigma S_t dB_t.$$

**Example 5.6** *The Ornstein-Uhlenbeck process (or Langevin equation) is the solution of the SDE*

$$dX_t = \mu X_t dt + \sigma dB_t$$

*or*

$$X_t = X_0 + \mu \int_0^t X_s ds + \sigma \int_0^t dB_s.$$

*In discrete time, this SDE transforms into*

$$X_{t+1} = (1+\mu) X_t + \sigma (B_{t+1} - B_t)$$

*or*

$$X_{t+1} = \phi X_t + Z_t,$$

*with* $\phi = 1 + \mu$ *and* $Z_t \sim N(0, \sigma^2)$*. This is the equation for an autoregressive time series of order 1.*

*Let*

$$Y_t = e^{-\mu t} X_t$$

*or* $Y_t = f(t, X_t)$ *with* $f(t,x) = e^{-\mu t}x$. *By Itô formula,*

$$\begin{aligned} Y_t &= Y_0 + \int_0^t \left( -\mu e^{-\mu s} X_s + \mu e^{-\mu s} X_s + \frac{1}{2}\sigma^2 \times 0 \right) ds \\ &+ \int_0^t \sigma e^{-\mu s} dB_s. \end{aligned}$$

*Therefore,*

$$X_t = e^{\mu t} X_0 + e^{\mu t} \int_0^t \sigma e^{-\mu s} dB_s.$$

*If* $X_0$ *is constant, this process is called the Ornstein-Uhlenbeck process.*

**Example 5.7** *Consider the SDE for the geometric Brownian motion again:*

$$dS_t = \alpha S_t dt + \sigma S_t dB_t \tag{5.10}$$

*or*

$$S_t = S_0 + \alpha \int_0^t S_s ds + \sigma \int_0^t S_s dB_s. \tag{5.11}$$

*Assume that*

$$S_t = e^{Z_t}.$$

*or*

$$Z_t = \ln(S_t).$$

*By the Itô formula, with* $f(x) = \ln(x)$, *we have*

$$\begin{aligned} dZ_t &= \frac{1}{S_t} dS_t + \frac{1}{2}\left(\frac{-1}{S_t^2}\right)(dS_t)^2 \\ &= \left(\alpha - \frac{1}{2}\sigma^2\right) dt + \sigma dB_t. \end{aligned}$$

*That is* $Z_t = Z_0 + \left(\alpha - \frac{1}{2}\sigma^2\right) t + \sigma B_t$ *and*

$$S_t = S_0 \exp\left[\left(\alpha - \frac{1}{2}\sigma^2\right) t + \sigma B_t\right].$$

In general, the solution of the homogeneous linear SDE

$$dX_t = \mu(t) X_t dt + \sigma(t) X_t dB_t$$

is

$$X_t = X_0 \exp\left[\int_0^t \left(\mu(s) - \frac{1}{2}\sigma(s)^2\right) ds + \int_0^t \sigma(s) dB_s\right].$$

## 5.4 Mean reverting processes

**Example 5.8** *The Ornstein-Uhlenbeck process with mean reversion is the solution of the SDE*

$$\begin{aligned} dX_t &= a\,(m - X_t)\,dt + \sigma dB_t, \\ X_0 &= x. \end{aligned}$$

*where* $a, \sigma > 0$ *and* $m \in \mathbb{R}$. *The solution of the associated ODE* $dx_t = -ax_t dt$ *is* $x_t = xe^{-at}$. *Consider the variable change* $X_t = Y_t e^{-at}$ *or* $Y_t = X_t e^{at}$. *By the Itô formula applied to* $f(t,x) = xe^{at}$, *we have*

$$Y_t = x + m\left(e^{at} - 1\right) + \sigma \int_0^t e^{as} dB_s.$$

*Therefore*

$$X_t = m + (x - m)\,e^{-at} + \sigma e^{-at} \int_0^t e^{as} dB_s.$$

*This is a Gaussian process, since the random part is* $\int_0^t f(s)\,dB_s$, *where* $f$ *is deterministic. The expected value is*

$$E[X_t] = m + (x - m)\,e^{-at}$$

*and the covariance is (by Itô isometry)*

$$\begin{aligned} Cov[X_t, X_s] &= \sigma^2 e^{-a(t+s)} E\left(\int_0^t e^{ar} dB_r\right)\left(\int_0^s e^{ar} dB_r\right) \\ &= \sigma^2 e^{-a(t+s)} \int_0^{t\wedge s} e^{2ar} dr \\ &= \frac{\sigma^2}{2a}\left(e^{-a|t-s|} - e^{-a(t+s)}\right). \end{aligned}$$

*Note that*

$$X_t \sim N\left[m + (x - m)\,e^{-at}, \frac{\sigma^2}{2a}\left(1 - e^{-2at}\right)\right].$$

*When* $t \to \infty$, *the distribution of* $X_t$ *converges to*

$$\nu := N\left[m, \frac{\sigma^2}{2a}\right],$$

*which is the invariant or stationary distribution. If* $X_0$ *has distribution* $\nu$ *then the distribution of* $X_t$ *will be* $\nu$ *for all* $t$.

**Remark 5.9** *Some financial applications of the Ornstein-Uhlenbeck process with mean reversion are:*

- *The Vasicek model for the interest rate* $r_t$*:*

$$dr_t = a\left(b - r_t\right)dt + \sigma dB_t,$$

*with* $a, b, \sigma$ *real constants. The solution of the SDE is*

$$r_t = b + \left(r_0 - b\right)e^{-at} + \sigma e^{-at}\int_0^t e^{as}dB_s.$$

- *The Black-Scholes model with stochastic volatility: assume that the volatility* $\sigma\left(t\right) = f\left(Y_t\right)$ *is a function of an Ornstein-Uhlenbeck process with mean reversion*

$$dY_t = a\left(m - Y_t\right)dt + \beta dW_t,$$

*where* $\{W_t, 0 \leq t \leq T\}$ *is a standard Brownian motion. The SDE which models the asset price evolution is*

$$dS_t = \alpha S_t dt + f\left(Y_t\right)S_t dB_t$$

*where* $\{B_t, 0 \leq t \leq T\}$ *is a standard Brownian motion and the processes* $W_t$ *and* $B_t$ *may be correlated, i.e.,*

$$E\left[B_t W_s\right] = \rho\left(s \wedge t\right).$$

An important and useful theoretical result is the following one.

**Proposition 5.10** *Let* $f : [0, +\infty) \to \mathbb{R}$ *be a deterministic function. Then*

1. $M_t = \exp\left(\int_0^t f(s)dB_s - \frac{1}{2}\int_0^t \left(f(s)\right)^2 ds\right)$ is a martingale.

2. $\int_0^t f(s)dB_s$ has a normal distribution with mean 0 and variance $\int_0^t \left(f(s)\right)^2 ds$.

The property 1 is a simple generalization of the fact that $\exp\left(\lambda B_t - \frac{1}{2}\lambda^2 t\right)$ is a martingale. The property 2 follows from 1, because martingales have constant mean, $E\left[M_0\right] = 1$ and $E\left[\exp\left(\lambda\int_0^t f(s)dB_s\right)\right] = \exp\left(\frac{1}{2}\lambda^2\int_0^t \left(f(s)\right)^2 ds\right)$, which is the moment generating function of the $N\left(0, \int_0^t \left(f(s)\right)^2 ds\right)$ distribution.

**Remark 5.11** *The AR(1) process is related with the mean reverting OU process. Consider the AR(1) process*

$$X_t = \phi X_{t-1} + Z_t,$$

*with* $\phi = 1 + \mu$ *and* $Z_t \sim N\left(0, \sigma_e^2\right)$. *Then*

$$E\left[X_t\right] = \phi^n X_0,$$
$$Var\left[X_t\right] = \sigma_e^2 \frac{\left(1 - \alpha^{2n}\right)}{1 - \alpha^2}.$$

*These coincide with the mean and variance values of the mean-reverting Ornstein-Uhlenbeck with* $m = 0$, *if we put* $\alpha = e^{-a}$ *and* $\frac{\sigma_e^2}{1-\alpha^2} = \frac{\sigma^2}{2a}$. *Therefore, the mean-reverting Ornstein-Uhlenbeck process is the continuous time equivalent of a AR(1) process such as standard Brownian motion is the continuous time equivalent of a random walk.*

**Exercise 5.12** *(Exam style problem): A derivatives trader is modelling the volatility of an equity index using the following time-discrete model (model 1):*

$$\sigma_t = 0.12 + 0.4\sigma_{t-1} + 0.05\varepsilon_t, \quad t = 1, 2, 3, \ldots$$

*where* $\sigma_t$ *is the volatility at time t years and* $\varepsilon_1, \varepsilon_2, \ldots$ *are a sequence of i.i.d. random variables with standard normal distribution. The initial volatility is* $\sigma_0 = 0.15$ *(that is, 15%). The trader is developing a related continuous-time model for use in derivative pricing. The model is defined by the following SDE (model 2):*

$$d\sigma_t = -\alpha\left(\sigma_t - \mu\right) dt + \beta dB_t,$$

*where* $\sigma_t$ *is the volatility at time t years,* $B_t$ *is the standard Brownian motion and the parameters* $\alpha, \beta$ *and* $\mu$ *all take positive values.*

*(a) Determine the long-term distribution of* $\sigma_t$ *for model 1.*

*(b) Show that for model 2 (solve the SDE), we have that*

$$\sigma_t = \sigma_0 e^{-\alpha t} + \mu\left(1 - e^{-\alpha t}\right) + \int_0^t \beta e^{-\alpha(t-s)} dB_s.$$

*(c) Determine the numerical value of* $\mu$ *and a relationship between parameters* $\alpha$ *and* $\beta$ *if it is required that* $\sigma_t$ *has the same long-term mean and variance under each model (models 1 and 2)*

*(d) State another consistency property between the two models that could be used to determine precise numerical values for* $\alpha$ *and* $\beta$.

(e) *The derivative pricing formula used by the trader involves the squared volatility $V_t = \sigma_t^2$, which represents the variance of the returns on the index. Determine the SDE for $V_t$ in terms of the parameters $\alpha, \beta$ and $\mu$.*

For more details on the theory of stochastic differential equations, see [5], [10] or [11]. For numerical methods, see [6].

# Chapter 6

# The Girsanov Theorem

## 6.1 Basic idea

The Girsanov Theorem, in its simplest version, says that a Brownian motion with drift

$$\widetilde{B}_t = B_t + \lambda t$$

can be transformed into a standard Brownian motion if we transform the probability measure $P$, of our probability space $(\Omega, \mathcal{F}, P)$ into a new probability measure $Q$. In financial appliations, this new probability measure, is the so-called risk neutral measure or the equivalent martingale measure.

In more general terms, the Girsanov Theorem says that we can transform the drift coefficient of an Itô process in such a way that the law of the process does not change "too much". The law of the new Itô process will be absolutely continuous with respect to the law of the original process and we can calculate explicitly the Radon-Nikodym derivative associated to the measure change

## 6.2 Change of probability measures

Consider the space $L_{a,T}$, which is the space of adapted and measurable stochastic processes $u$ such that $P\left[\int_0^T u_t^2 dt < \infty\right] = 1$. Let us define define $L^1_{a,T}$ as the space of adapted and measurable processes $v$ such that $P\left[\int_0^T |v_t| \, dt < \infty\right] = 1$.

**Definition 6.1** *A continuous and adapted process* $X = \{X_t, 0 \le t \le T\}$ *is* *said to be an Itô process if it has the form*

$$X_t = X_0 + \int_0^t u_s dB_s + \int_0^t v_s ds, \tag{6.1}$$

*where* $u \in L_{a,T}$ *and* $v \in L^1_{a,T}$.

The drift of an Itô process is the integral term $\int_0^t v_s ds$.

Let $L \geq 0$ be a random variable defined on the probability space $(\Omega, \mathcal{F}, P)$. Then, we can define a new probability measure $Q$, by

$$Q(A) = E[\mathbf{1}_A L], \quad \text{for any } A \in \mathcal{F}.$$

It is clear that we must have

$$Q(\Omega) = E[L] = 1,$$

and that $Q(A) = E[\mathbf{1}_A L]$ is equivalent to

$$\int_\Omega \mathbf{1}_A dQ = \int_\Omega \mathbf{1}_A L dP.$$

We say that $L$ is the density of $Q$ with respect to $P$ and we write

$$\frac{dQ}{dP} = L.$$

$L$ is also the Radon-Nikodym derivative of $Q$ with respect to $P$.

The expected value of a random variable $X$, defined on the probability space $(\Omega, \mathcal{F}, P)$, with respect to $Q$, is given by the formula

$$E_Q[X] = E[XL].$$

The probabilty measure $Q$ is absolutely continuous with respect to $P$. This means that

$$P(A) = 0 \Longrightarrow Q(A) = 0.$$

When are the measures $P$ e $Q$ equivalent? In order to answer this question, we recall the definition of equivalent probability measures.

**Definition 6.2** *Two probability measures* $P$ *and* $Q$ *which apply to the same sigma-algebra* $\mathcal{F}$ *are said to be equivalent if for any event* $A \in \mathcal{F} : P(A) = 0$ *if and only if* $Q(A) > 0$*, where* $P(A)$ *and* $Q(A)$ *are the probabilities of* $A$ *under* $P$ *and* $Q$ *respectively.*

If the random variable $L$ is strictly positive ($L > 0$), then the probabilty measures $P$ and $Q$ are equivalent (or mutually absolutely continuous), which is equivalent to say that

$$P(A) = 0 \Longleftrightarrow Q(A) = 0.$$

## 6.3 Girsanov Theorem

We now discuss the simplest version of the Girsanov Theorem, which applies to a a random variable $X$ with normal distribution $N(m, \sigma^2)$. The basic question that leads to the Girsanov Theorem is: exists a probability measure $Q$, such that $X$ has a normal distribution with mean zero, i.e. $N(0, \sigma^2)$, with respect to $Q$? In order to answer this question, consider the random variable

$$L = \exp\left(-\frac{m}{\sigma^2}X + \frac{m^2}{2\sigma^2}\right).$$

It is easy to show that $E[L] = 1$. Indeed, using the probability density function of the normal distribution $N(m, \sigma^2)$, we have that

$$\begin{aligned} E[L] &= \int_{-\infty}^{+\infty} \exp\left(-\frac{m}{\sigma^2}x + \frac{m^2}{2\sigma^2}\right) \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) dx \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left(-\frac{x^2}{2\sigma^2}\right) dx = 1. \end{aligned}$$

Assume that the new measure $Q$ has density $L$ with respect to $P$. Then, in the probability space $(\Omega, \mathcal{F}, Q)$, the random variable $X$ has a characteristic function given by

$$\begin{aligned} E_Q\left[e^{itX}\right] &= E\left[e^{itX}L\right] \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left(itx - \frac{m}{\sigma^2}x + \frac{m^2}{2\sigma^2}\right) \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) dx \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left(itx - \frac{x^2}{2\sigma^2}\right) dx = e^{-\frac{\sigma^2 t^2}{2}}. \end{aligned}$$

Therefore, $X$ has distribution $N(0, \sigma^2)$.

The next version of the Girsanov Theorem is for the Brownian motion. Let $\{B_t, t \in [0, T]\}$ be a Brownian motion in the probability space $(\Omega, \mathcal{F}_T, P)$. Fix a real number $\lambda$ and consider the martingale

$$L_t = \exp\left(-\lambda B_t - \frac{\lambda^2}{2}t\right). \tag{6.2}$$

**Exercise 6.3** *Prove that the stochastic process $\{L_t, t \in [0, T]\}$, given by (6.2), is a positive martingale with expected value 1 and satisfies the stochastic differential equation*

$$\begin{aligned} dL_t &= -\lambda L_t dB_t, \\ L_0 &= 1. \end{aligned}$$

The random variable

$$L_T = \exp\left(-\lambda B_T - \frac{\lambda^2}{2}T\right)$$

is a density in the probability space $(\Omega, \mathcal{F}_T, P)$, and we can define the new probability measure

$$Q(A) = E[\mathbf{1}_A L_T],$$

for each $A \in \mathcal{F}_T$. Since $\{L_t, t \in [0,T]\}$ is a martingale then the random variable $L_t = \exp\left(-\lambda B_t - \frac{\lambda^2}{2}t\right)$ is also a density in the probabilty space $(\Omega, \mathcal{F}_t, P)$. In this probability space, the measure $Q$ has precisely the density $L_t$ with respect to $P$. Indeed, if $A \in \mathcal{F}_t$, then

$$\begin{aligned} Q(A) &= E[\mathbf{1}_A L_T] = E[E[\mathbf{1}_A L_T|\mathcal{F}_t]] \\ &= E[\mathbf{1}_A E[L_T|\mathcal{F}_t]] = E[\mathbf{1}_A L_t], \end{aligned}$$

where we have applied the properties of conditional expectation and the martingale property of $\{L_t, t \in [0,T]\}$.

**Theorem 6.4** *(Girsanov Theorem I): In the probability space $(\Omega, \mathcal{F}_T, Q)$, where $Q$ is defined by $Q(A) = E[\mathbf{1}_A L_T]$, the stochastic process*

$$\widetilde{B}_t = B_t + \lambda t$$

*is a standard Brownian motion.*

In order to prove the theorem, we need a technical lemma.

**Lemma 6.5** *Assume that $X$ is a random variable and that $\mathcal{G}$ is a $\sigma$-algebra such that:*

$$E\left[e^{iuX}|\mathcal{G}\right] = e^{-\frac{u^2\sigma^2}{2}}.$$

*Then, $X$ is independent of the $\sigma$-algebra $\mathcal{G}$ and has normal distribution $N(0,\sigma^2)$.*

For a proof of this lemma, we refer to [9] (pages 63-64).

**Proof.** (of the Girsanov Theorem) We only need to show that in the space $(\Omega, \mathcal{F}_T, Q)$, the increment $\widetilde{B}_t - \widetilde{B}_s$, with $s < t \leq T$, is independent of $\mathcal{F}_s$ and has normal distribution $N(0, t-s)$. By Lemma 6.5, this is a consequence of

$$E_Q\left[\mathbf{1}_A e^{iu\left(\widetilde{B}_t - \widetilde{B}_s\right)}\right] = Q(A)\, e^{-\frac{u^2}{2}(t-s)}, \tag{6.3}$$

for all $s < t$, $A \in \mathcal{F}_s$ and $u \in \mathbb{R}$. Indeed, if (6.3) is satisfied then, by the conditional expectation definition and by Lemma 6.5, we have that $\left(\widetilde{B}_t - \widetilde{B}_s\right)$ is independent of $\mathcal{F}_s$ and has normal distribution $N(0, t-s)$.

We now show that (6.3) is satisifed.

$$\begin{aligned} E_Q\left[\mathbf{1}_A e^{iu\left(\widetilde{B}_t - \widetilde{B}_s\right)}\right] &= E\left[\mathbf{1}_A e^{iu\left(\widetilde{B}_t - \widetilde{B}_s\right)} L_t\right] \\ &= E\left[\mathbf{1}_A e^{iu(B_t - B_s) + iu\lambda(t-s) - \lambda(B_t - B_s) - \frac{\lambda^2}{2}(t-s)} L_s\right] \\ &= E\left[\mathbf{1}_A L_s\right] E\left[e^{(iu-\lambda)(B_t - B_s)}\right] e^{iu\lambda(t-s) - \frac{\lambda^2}{2}(t-s)} \\ &= Q(A) e^{\frac{(iu-\lambda)^2}{2}(t-s) + iu\lambda(t-s) - \frac{\lambda^2}{2}(t-s)} \\ &= Q(A) e^{-\frac{u^2}{2}(t-s)}, \end{aligned}$$

where we have used the definitions of $E_Q$ and $L_t$, the independence of $(B_t - B_s)$ from $L_s$ and $A$ and the definition of $Q$. ■

Finally, we present a more general version of the Girsanov Theorem.

**Theorem 6.6** *(Teorema de Girsanov II): Let $\{\theta_t, t \in [0,T]\}$ be an adapted stochastic process that satisfies the Novikov condition:*

$$E\left[\exp\left(\frac{1}{2}\int_0^T \theta_t^2 dt\right)\right] < \infty. \tag{6.4}$$

*Then, the stochastic process*

$$\widetilde{B}_t = B_t + \int_0^t \theta_s ds$$

*is a Brownian motion with respect to the measure $Q$, defined by $Q(A) = E\left[\mathbf{1}_A L_T\right]$, where*

$$L_t = \exp\left(-\int_0^t \theta_s dB_s - \frac{1}{2}\int_0^t \theta_s^2 ds\right). \tag{6.5}$$

Note that the process $L_t$ in (6.5) satisfies the linear stochastic differential equation

$$L_t = 1 - \int_0^t \theta_s L_s dB_s.$$

In order to ensure that the process $L_t$ is a density, we need to have $E[L_t] = 1$ and the Novikov condition (6.4) is sufficient to ensure that $E[L_t] = 1$.

The second version of the Girsanov Theorem generalizes the first version. Indeed, with $\theta_t \equiv \lambda$ we obtain the first version.

A detailed discussion of the Girsanov Theorem is presented in [10].

# Bibliography


[1] Tomas Björk, Arbitrage Theory in Continuous Time, 2nd Edition. Oxford University Press, 2004.

[2] Zdzislaw Brzezniak and Tomasz Zastawniak, Basic Stochastic Processes. Springer, 1999.

[3] Marek Capinski and Ekkehard Kopp, Measure, Integral and Probability. Springer, 1998.

[4] Robert Jarrow and Philip Protter, A short history of stochastic integration and mathematical finance: The early years, 1880–1970. IMS Lecture Notes Monograph, vol. 45 (2004), pages 1–17.

[5] Ioannis Karatzas and Steven Shreve, Brownian Motion and Stochastic Calculus, 2nd edition. Springer, 1991.

[6] Peter E. Kloeden and Platen, Numerical Solution of Stochastic Differential Equations. Springer, 1992.

[7] Thomas Mikosch, Elementary Stochastic Calculus with Finance in view. World Scientific, 1998.

[8] Daniel Müller, Probabilidade e processos estocásticos: uma abordagem rigorosa com vista aos modelos em finanças. Coimbra, Almedina, 2011.

[9] David Nualart, Stochastic Calculus (Lecture notes, Kansas University), http://www.math.ku.edu/~nualart/StochasticCalculus.pdf, 2008.

[10] Bernt Oksendal, Stochastic Differential Equations, 6th Edition. Springer, 2003.

[11] Daniel Revuz and Marc Yor, Continuous martingales and Brownian motion. Third Edition. Springer, 1999.

[12] Steven Shreve, Stochastic Calculus for Finance II: Continuous-Time Models. Springer, 2004.