CHAPTER 1

From Cox and Rubinstein Model to Brownian Motion

In this chapter, we start from the simple model by Cox and Rubinstein in discrete time to introduce Gaussian processes and specifically Brownion motion. Brownion motion must be understood as a time-continuous version of the simple random walk obtained by flipping a coin at every time step and by moving according to the result of the coin. As easily guessed by the reader that knows a little bit of probability theory, Central Limit Theorem is the keystone to turn from the time-discrete to the time-continuous setting.

1. Cox and Rubinstein Model

The model by Cox and Rubinstein is the simplest one for financial a risky financial market.

1.1. Notion of Asset. In financial accounting, assets are economic resources. Anything tangible or intangible that is capable of being owned or controlled to produce value and that is held to have positive economic value is considered an asset. Simply stated, assets represent ownership of value that can be converted into cash (although cash itself is also considered an asset). See *Sullivan, Arthur; Steven M. Sheffrin (2003). Economics: Principles in action. Upper Saddle River, New Jersey 07458: Pearson Prentice Hall. pp. 272. ISBN 0-13-063085-3.*

Here are some examples of assets we will consider: stocks (or stock shares) of a company, commodities (such as petroleum or oil, cooper, coffee...), currencies. These are called *risky askets* since their prices may fluctuate with time according to uncertain scenarios. The current price of such an asset is called its *spot price*: the spot price or spot rate of a commodity, a stock or a currency is the price that is quoted for immediate (spot) settlement (payment and delivery). Because of the underlying uncertainty, a common mathematical modelling consists in assuming the spot to be driven by some randomness.

On financial markets, there also exist some *riskless assets*: a riskless asset is an asset whose future return is known with certainty. For example, savings accounts pay a pre-determined interest. Similarly, lending money at a pre-determined interest is non-risky as well: the price paid for the use of borrowed money is originally fixed by a contract.

1.2. Simple Financial Market. A simple model for a financial market consists in considering two assets only: a riskless asset and a risky one.

The price of a riskless asset is given by a deterministic rule. To simplify, we will assume that the rule is constant in time, that is driven by a time-constant interest rate. Given a deposit of 1 € at time 0, the value of the deposit at time n is $(1+r)^n$ €. The parameter $r > 0$ is called "interest rate on a time unit". Similarly, given a lend of 1 € at time 0, the money the borrower owes at time n is $(1+r)^n$ €, for the same r: to simplify, we will say that the interest a financial agent gets from the savings is the one another agent pays when borrowing.

Given an initial amount of money, say S_0^0 €, the value of the riskless asset driven by the interest rate r, is $S_0^0(1+r)^n$ €, whatever the sign S_0^0 is: if $S_0^0 > 0$, the holder of the asset is considered to deposit his money on a savings account; if $S_0^0 < 0$, the holder of the asset is considered to borrow some money from the bank.

Therefore, the dynamics of the riskless asset are given by

$$S_n^0 = (1+r)^n S_0^0, \quad n \in \mathbb{N}. \tag{1.1}$$

Examinate now the modelling of the risky asset. As already said, uncertainty is modelled by randomness. The simplest way to translate uncertainty then consists in allowing two possible random scenarios at any time step: either the spot price increases or it decreases. Denoting by $(S_n)_{n\geq 0}$ the spot price, we write:

$$S_{n+1} = S_n \times \xi_{n+1},$$

where $(\xi_n)_{n\in\mathbb{N}}$ is a sequence of random variables on some given probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with values in a two-element-set $\{u, d\}$, $u > 0$ standing for *up* and $d > 0$ for *down*, $u \geq d$. Most of the time, we will assume the variables $(\xi_n)_{n\in\mathbb{N}}$ to be identically distributed and independent. In particular, in this case, there exists some $p \in [0, 1]$ such that

$$\mathbb{P}\{\xi_n = u\} = 1 - \mathbb{P}\{\xi_n = 0\} = p, \quad n \in \mathbb{N}.$$

We will see in the next chapter that u, d and r must satisfy some suitable conditions to make the model meaningful.

Given the initial spot price S_0, (1.2) may be recasted as

$$S_n = S_0 \prod_{i=1}^{n} \xi_i, \quad n \in \mathbb{N}. \tag{1.2}$$

Such a model (1.1–1.2) with i.i.d. variables $(\xi_n)_{n\in\mathbb{N}}$ is called *Cox and Rubinstein model* for a financial market.

2. Gaussian Vectors

2.1. From Discrete to Continuous in Cox and Rubinstein. Eq. (1.2) may be written

$$S_n = S_0 \exp\left(\sum_{i=1}^{n} \ln(\xi_i)\right). \tag{2.3}$$

replace n above by N and assume that S_N represents the spot price at time 1 in real time, that is $1/N$ is a small unit of time allowing to cut the whole $[0,1]$ into small pieces. Then, $\ln(S_N)$ is expected to be macroscopic, that is

$$\mathbb{E}\left(\sum_{i=1}^{N} \ln(\xi_i)\right) \quad \text{and} \quad \mathbb{V}\left(\sum_{i=1}^{N} \ln(\xi_i)\right)$$

are of order 1. Put differently, we expect $\mathbb{E}(\ln(\xi_1))$ and $\mathbb{V}(\ln(\xi_1))$ to be of order $1/N$. We write

$$S_N = S_0 \exp\left(\mu + \sigma \sum_{i=1}^{N} \zeta_i\right),$$

where $(\zeta_i)_{1\le i\le N}$ are i.i.d. variables with zero-mean and reduced variance, i.e. $\mathbb{V}(\zeta_1) = 1$

Now, the Central Limit Theorem (CLT for short) says that, for N large, the distribution of S_N is close to the distribution of $S_0 \exp(\mu+\sigma Z)$ where $Z \sim \mathcal{N}(0,1)$. From this modelling, it seems that a reasonable modelling for the spot price at time 1 in a continuous-time market is a *log-normal* distribution, that is a variable whose logarithm has a normal or Gaussian distribution.

2.2. Gaussian Vectors. The previous paragraph is not satisfactory: it says what happens at time 1 only. Basically, we expect the same argument to say that the spot price has *marginal Gaussian distributions* when the time-step tends to 0 in the Cox and Rubinstein model. Still this is not enough. The point is determine the *joint distribution* of the price at different time-indices.

A suitable version of the CLT would show that the logarithm of the price, computed at different time-indices, has a Gaussian structure again.

DEFINITION 2.1. *A random vector* $(X_1, \dots, X_n)$ *is said to be Gaussian if any linear combination of its coordinates has a Gaussian distribution.*

Obviously, the word *linear* can be replaced by *affine*. As a basic example, we can check that, given n independent Gaussian random variables $X_1, \dots, X_n$, the vector $(X_1, \dots, X_n)$ is Gaussian. (Be careful that the word *independent* is here crucial.) As a straightforward consequence of the definition, we claim

PROPOSITION 2.2. *Given a Gaussian vector* $X = (X_1, \dots, X_n)$, *a matrix* $A \in \mathcal{M}_{m,n}(\mathbb{R})$ *and a vector* $b = (b_1, \dots, b_n) \in \mathbb{R}^n$, *the random vector* $Y = (Y_1, \dots, Y_m)$ *given by* $Y^\top = AX^\top + b^\top$ *is Gaussian as well.*

DEFINITION 2.3. *Given a Gaussian vector* $X = (X_1, \dots, X_n)$, *we call* $\mathbb{E}(X) = (\mathbb{E}(X_i))_{1 \le i \le n}$ *the mean of the vector* X *and* $\mathbb{K}(X) = (\mathrm{Cov}(X_i, X_j))_{1 \le i,j \le n}$ *the covariance matrix of the vector* X.

For example, it is well checked that the covariance matrix $\mathbb{K}(U)$ of $U = (U_1, \dots, U_n)$, where $U_1, \dots, U_n$ are n i.i.d. random variables of $\mathcal{N}(0,1)$ law, is the identity matrix of size n. It is also an easy exercise to check that, for a Gaussian vector X of size n and for any $a \in \mathbb{R}^n$, $\langle a, X \rangle_{\mathbb{R}^n}$ is a Gaussian random variable with $\langle a, \mathbb{E}(X) \rangle_{\mathbb{R}^n}$ as mean and $\langle a, \mathbb{K}(X)a \rangle_{\mathbb{R}^n}$ as variance. As a corollary, we deduce that the matrix $\mathbb{K}(X)$ is symmetric and non-negative. Conversely, given a vector $m \in \mathbb{R}^n$ and a symmetric non-negative matrix K of size n, it is well checked that $m + K^{1/2}U^\top$ is a Gaussian vector with m as mean and K as covariance matrix. We deduce

PROPOSITION 2.4. *Given a Gaussian vector* $X = (X_1, \dots, X_n)$ *of size* n, *the characteristic function of* X *has the form*

$$\varphi_X(t) = \mathbb{E}\big[\exp(i\langle t, X \rangle_{\mathbb{R}^d})\big] = \exp\Big(i\langle t, \mathbb{E}(X) \rangle_{\mathbb{R}^n} - \frac{1}{2}\langle t, \mathbb{K}(X)t \rangle_{\mathbb{R}^n}\Big),$$

$t \in \mathbb{R}^n$, *where* $i^2 = -1$.

Keep in mind that the distribution of a random vector is characterized by its characteristic function. In particular, two Gaussian vectors having the same mean and the same covariance matrix have the same distribution, so that the distribution of a Gaussian vector is characterized by its mean and its covariance matrix.

2.3. Independence and Gaussian Vectors. The specific form of the characteristic function of a Gaussian vector permits to characterize independence of the coordinates very easily.

PROPOSITION 2.5. *Given a Gaussian vector* $X = (X_1, \dots, X_n)$ *and* $1 \leq \ell_1 < \ell_2 < \cdots < \ell_p = n$, *the vectors* $(X_1, \dots, X_{\ell_1})$, $(X_{\ell_1+1}, \dots, X_{\ell_2})$, $\dots$, $(X_{\ell_{p-1}+1}, \dots, X_{\ell_p})$ *are independent if and only if the covariance matrix* $\mathbb{K}(X)$ *satisfies* $\mathbb{K}_{i,j}(X) = 0$ *whenever* $\ell_p \leq i < \ell_{p+1}$ *and* $\ell_q \leq j < \ell_{q+1}$ *for* $p \neq q$.

We will say that the blocks are independent if and only if the covariance matrix is block-diagonal.

3. Gaussian Processes

3.1. Processes.

DEFINITION 3.1. *We call a process a collection* $(X_t)_{t\in I}$ *of random variables indexed by an interval* I, *the parameter* t *being understood as a continuous time parameter.*

In finance, processes are of great importance since they permit to model the time evolution of prices on a market. Put differently, $(X_t)_{t\in I}$ may be thought as the price (or sometimes as the log-price) of a given asset over the time interval I.

DEFINITION 3.2. *The distribution of a process* $(X_t)_{t\in I}$ *is given by the family of distributions of all the possible finite-dimensional vectors formed by the process* $(X_t)_{t\in I}$, *that is all the vectors* $(X_{t_1}, \dots, X_{t_n})_{t_1,\dots,t_n\in I}$ *or* $(X_{t_1}, \dots, X_{t_n})_{t_1,\dots,t_n\in I, t_1<t_2<\cdots<t_n}$.

In particular, two processes $(X_t)_{t\in I}$ *and* $(Y_t)_{t\in I}$ *are said to have the same distribution if, for any* $t_1, \dots, t_n \in I$ *(possibly ordered), the vectors* $(X_{t_1}, \dots, X_{t_n})$ *and* $(Y_{t_1}, \dots, Y_{t_n})$ *have the same distribution.*

Similarly, two process $(X_s)_{s\in I}$ *and* $(Y_t)_{t\in J}$, J *being another interval of* $\mathbb{R}$, *are said to be independent if, for any* $s_1, \dots, s_m \in I$ *and any* $t_1, \dots, t_n \in J$ *(possibly ordered), the vectors* $(X_{s_1}, \dots, X_{s_m})$ *and* $(Y_{t_1}, \dots, Y_{t_n})$ *are independent.*

3.2. Definition of a Gaussian Process.

DEFINITION 3.3. *A process* $(X_t)_{t\in I}$ *is said to be Gaussian if, for any* $t_1, \dots, t_n \in I$ *(possibly ordered), the vector* $(X_{t_1}, \dots, X_{t_n})$ *is Gaussian.*

Given a Gaussian process $(X_t)_{t\in I}$, we call the function $I \ni t \mapsto \mathbb{E}(X_t)$ its mean function and the function $I^2 \ni (s,t) \mapsto \mathrm{Cov}(X_s, X_t)$ is covariance function. As a straightforward result, two Gaussian processes have the same law if and only if they have the same mean function and the same covariance function.

4. Brownian Motion

4.1. Definition. Brownian motion is the time-continuous version of the symmetric random walk. Its definition goes back to the work of the botanist Robert Brown in the 19th century, who tried to model the motion of grains of pollen in the air. Albert Einstein and then Norbert Wiener contributed to its modern definition.

Before we specify the definition, we understand that it must be of independent increments and it must be of stationary distribution, as the symmetric random walk is. The law of a given increment must be Gaussian since understood as a limit in the CLT. For these raeasons, we claim

DEFINITION 4.1. *A process $(B_t)_{t\geq 0}$ is a Brownian motion if*

(1) $B_0 = 0$,
(2) *for* $0 \leq s < t$, $B_t - B_s \sim \mathcal{N}(0, t-s)$,
(3) *for* $0 \leq t_1 < \cdots < t_n$, *the increments* B_{t_1}, $B_{t_2} - B_{t_1}$, $\dots$, $B_{t_n} - B_{t_{n-1}}$ *are independent,*
(4) *for any* $\omega \in \Omega$, *the path* $t \geq 0 \mapsto B_t(\omega)$ *is continuous.*

The condition (4) is not so obvious. The idea is that the paths of the Brownian motion are continuous as Brownian motion is assumed to describe the continuous motion of a particle along the line. Actually, the paths of the Brownian motion are rather irregular even if continuous: it might be shown that, with probability 1, the paths are nowhere differentiable! We will prove that, with probability 1, there exists no interval of non-zero length on which B is monotonous.

Given $x \neq 0$, we will say that $(x + B_t)_{t\geq 0}$ is a Brownian motion starting from x.

4.2. Brownian Motion as a Gaussian Process. A crucial point to charaterize the distribution of a Brownian motion is

PROPOSITION 4.2. *A process $(B_t)_{t\geq 0}$ with continuous paths is a Brownian motion if and only if it is a Gaussian process starting from* 0, *with zero mean and* $[0, +\infty)^2 \ni (s,t) \mapsto s \wedge t = \min(s,t)$ *as covariance function.*

4.3. Brownian Motion in Higher Dimension.

DEFINITION 4.3. *A process $(B_t^1, \dots, B_t^d)_{t\geq 0}$ is said to be a d-dimen- -sional Brownian motion if $(B_t^i)_{t\geq 0}$ is a Brownian motion for any $i \in \{1, \dots, d\}$ and the processes $(B_t^1)_{t\geq 0}$, $\dots$, $(B_t^d)_{t\geq 0}$ are independent.*

5. Exercices

CHAPTER 2

Stochastic Integral

In this chapter, we construct a new form of integral, called "stochastic integral". From a financial point of view, the stochastic integral describes the variations of a dynamical portfolio as the underlying risky asset evolves.

1. Discrete Stochastic Integral

In this section, we are given a discrete model for a financial market, of the same form as one we considered in Chapter 1. In particular, we will keep the same notations as in (1.1) and (1.2).

1.1. Filtration. The starting point for the construction of the integral consists in the following simple remark: given an agent on the market, the available information at time $n+1$ "larger" than the available information at time n. That is, at time $n+1$, all the prices before $n+1$ are known, whereas only the prices before n are known at time n.

The point is thus to model the available information. At time 0, only deterministic prices are known by the agent: the available information is modeled by the σ-field $\mathcal{F}_0 = \{\emptyset, \Omega\}$. (Prices that are measurable w.r.t. $\mathcal{F}_0$ are deterministic.) At time 1, all the prices writing in terms of ξ_1 are known explicitely by the agent: the available is modeled by the σ-field $\mathcal{F}_1 = \sigma(\xi_1)$. And so on... That is, we set

$$\forall n \geq 1, \quad \mathcal{F}_n = \sigma(\xi_1, \dots, \xi_n).$$

The sequence of σ-fields $(\mathcal{F}_n)_{n\geq 0}$ is called a filtration, that is a non-decreasing sequence of sub-σ-fields of $\mathcal{A}$.

1.2. Portfolio or Strategy. Assume that the agent has some amount of money to invest at time 0. At time 0, he may decide to buy ϕ_1^0 shares of the riskless asset and ϕ_1 shares of the risky asset. The price he pays to buy the whole is

$$W_0 = \phi_1^0 S_0^0 + \phi_1 S_0.$$

Let us remark the followings: (i) W_0 here stands for the initial wealth of the agent, (ii) the quantities ϕ_1^0 and ϕ_1 may be negative, that is the agent borrows some shares from the savings bank or from the company that has $(S_n)_{n\geq 0}$ as stock.

Then, the wealth at time 1 of the agent may be expressed as

$$W_1 = \phi_1^0 S_1^0 + \phi_1 S_1.$$

Clearly, the agent may decide to buy or to sell some shares at time 1 according to the new spot price of the risky asset, that is the agent may decide of a new "portfolio" at time 1. The new portfolio may be also expressed as a pair (ϕ_2^0, ϕ_2): ϕ_2^0 stands for the new number of shares in the riskless asset and ϕ_2 for the new number of shares in the risky asset. With this new portfolio, the wealth at time 2 of the agent may be expressed as

$$W_2 = \phi_2^0 S_2^0 + \phi_2 S_2.$$

And so on... We then understand a portfolio as a sequence of random variables $(\phi_n^0, \phi_n)_{n\geq 1}$ such that, at any $n \geq 1$, the pair (ϕ_n^0, ϕ_n) is $\mathcal{F}_{n-1}$-measurable. (The collection $(\phi_n^0, \phi_n)_{n\geq 1}$ is said to be a predictable process w.r.t. the filtration $(\mathcal{F}_n)_{n\geq 1}$.) We will also refer to a strategy for a portfolio.

In the case when there is no income or outcome of money, the strategy is said to be self-financing, that is the wealth at time $n \geq 1$ may be expressed in two ways:

$$W_n = \phi_n^0 S_n^0 + \phi_n S_n,$$

which is the basic writing we introduced above, and

$$W_n = \phi_{n+1}^0 S_n^0 + \phi_{n+1} S_n,$$

which says the new shares $(\phi_{n+1}^0, \phi_{n+1})$ are obtained by investing the entire wealth at time n. In other words, a strategy is said to be self-financing if, for any $n \geq 1$,

$$\phi_{n+1}^0 S_n^0 + \phi_{n+1} S_n = \phi_n^0 S_n^0 + \phi_n S_n.$$

If the strategy is not self-financing, the gap

$$C_n = \big(\phi_n^0 S_n^0 + \phi_n S_n\big) - \big(\phi_{n+1}^0 S_n^0 + \phi_{n+1} S_n\big),$$

stands for the consumption at time n, that is the difference between the wealth at time n before consumption and the wealth at time n after consumption.

1.3. Discrete Stochastic Integral. Given a self-financing strategy, we notice that the variation of the wealth of the agent on a time interval of length 1 is

$$\begin{aligned} W_{n+1} - W_n &= \big(\phi^0_{n+1} S^0_{n+1} + \phi_{n+1} S_{n+1}\big) - \big(\phi^0_n S^0_n + \phi_n S_n\big) \\ &= \big(\phi^0_{n+1} S^0_{n+1} + \phi_{n+1} S_{n+1}\big) - \big(\phi^0_{n+1} S^0_n + \phi_{n+1} S_n\big) \\ &= \phi^0_{n+1}\big(S^0_{n+1} - S^0_n\big) + \phi_{n+1}\big(S_{n+1} - S_n\big). \end{aligned}$$

Summing along the small increments, we deduce that the global variation of the wealth has the form

$$W_n - W_0 = \sum_{k=0}^{n-1} \big[\phi^0_{k+1}\big(S^0_{k+1} - S^0_k\big) + \phi_{k+1}\big(S_{k+1} - S_k\big)\big].$$

This sum appears as a discrete integral of the process $(\phi_{n+1})_{n\geq 0}$ against the variations of the prices $(S^0_n)_{n\geq 0}$ and $(S_n)_{n\geq 0}$.

2. Wiener Integral

In the end of the chapter, the objective is to construct a stochastic integral for time-continuous markets, that is for time-continuous prices driven by a Brownian motion. We call Wiener integral the case when the integrand, that is the strategy, is deterministic.

2.1. Integration of a Step Function. In this paragraph, we are given a one-dimensional Brownian motion $(B_t)_{t\geq 0}$ on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$. We are also given a deterministic left-continuous step function $f : u \in \mathbb{R}_+ \mapsto f(u)$, that vanishes for u large enough, that there exist an integer $n \geq 1$, $n+1$ reals $0 = t_0 < \cdots < t_n$ and n reals $\alpha_1, \dots, \alpha_n$ such that

$$f(u) = \sum_{k=0}^{n-1} \alpha_{k+1} \mathbf{1}_{\{(t_k, t_{k+1}]\}}(u).$$

(Pay attention that the value of f at zero is useless.)

The integral of f along the variation of B is understood as a generalization of the value of the wealth along some strategy and some variation of the stock price. We set

$$\int_0^{+\infty} f(u) dB_u = \sum_{k=0}^{n-1} \alpha_{k+1}\big(B_{t_{k+1}} - B_{t_k}\big).$$

Since the reals $\alpha_1, \dots, \alpha_n$ are deterministic, we deduce

PROPOSITION 2.1. *The law of the random variable* $\int_0^{+\infty} f(u)dB_u$ *is Gaussian, with zero as mean and*

$$\int_0^{+\infty} f^2(u)du$$

as variance.

PROOF. The proof is a straightforward consequence of the independence and of the Gaussian distribution of the increments of $(B_t)_{t\geq 0}$. $\square$

Actually, we can even prove

PROPOSITION 2.2. *Given* N *left-continuous step-functions (with a compact support)* $f_1, \ldots, f_N$*, the random vector*

$$\left(\int_0^{+\infty} f_1(u)dB_u, \ldots, \int_0^{+\infty} f_N(u)dB_u\right)$$

is a Gaussian vector with

$$K_{i,j} = \int_0^{+\infty} f_i(u)f_j(u)du, \quad 1 \leq i, j\|eqN,$$

as covariance matrix.

2.2. Integration of an L^2 Function. We have just defined an isometry between the Wiener integral of a step-function f seen as an element of $L^2(\Omega, \mathcal{A}, \mathbb{P})$ and the underlying step-function f itself seen as an element of $L^2(\mathbb{R}_+, du)$. Since the step-functions are dense in $L^2(\mathbb{R}_+, du)$ and $L^2(\Omega, \mathcal{A}, \mathbb{P})$ is complete, we can extend the definition of the Wiener integral to any deterministic function $f \in L^2(\mathbb{R}_+, du)$: the integral is then seen as the $L^2(\Omega, \mathcal{A}, \mathbb{P})$-limit of any sequence of Wiener integrals computed along any approximating sequence of step-functions of f in $L^2(\mathbb{R}_+, du)$.

By stability of Gaussian distributions along convergence in law, it is plain to see that Propositions 3.1 and 2.2 remain true in this more general framework.

2.3. The Integral as Stochastic Process. Clearly, we can let the bound of integration vary in the definition by setting, for a given $t \geq 0$,

$$\int_0^t f(u)dB_u = \int_0^{+\infty} f(u)\mathbf{1}_{(0,t]}(u)dB_u.$$

We emphasize that the function $f\mathbf{1}_{(0,t]}$ is in $L^2(\mathbb{R}_+, du)$ if the function f is $L^2_{\text{loc}}(\mathbb{R}_+, du)$.

The point is to decide what happens when letting the value of t vary. From a theoretical point of view, things are not so simple since the integral is defined as an $L^2(\Omega, \mathcal{A}, \mathbb{P})$-limit, that is it is uniquely defined up to an event of zero probability under $\mathbb{P}$. Obviously, this event of zero probability depends on t a priori, so that the union of all these events as t runs over nonnegative indices may not be an event.

We thus admit the following proposition:

PROPOSITION 2.3. *Given a function $f \in L^2_{\mathrm{loc}}(\mathbb{R}_+, du)$, there exists a continuous process denoted by*

$$\left(\int_0^t f(u)dB_u\right)_{t\geq 0},$$

such that, for any $t \geq 0$, the random variable

$$\int_0^t f(u)dB_u$$

is the stochastic integral of f along the variation of B on the interval $[0, t]$.

Then, we can prove

PROPOSITION 2.4. *Given a function $f \in L^2_{\mathrm{loc}}(\mathbb{R}_+, du)$, the process*

$$\left(\int_0^t f(u)dB_u\right)_{t\geq 0},$$

is a Gaussian process, with zero as mean, and with

$$(s, t) \in \mathbb{R}_+^2 \mapsto \int_0^{s\wedge t} f^2(u)du,$$

as covariance function. (Here, $s \wedge t = \min(s, t)$.)

3. Itô Integral

In practice, it would be too restrictive to assume the strategy to be deterministic on a time-continuous market. Therefore, there is a need for an extension of the integral to the case when the integrand is random itself. Keeping in mind the first section, we understand that the randomness in the integrand (or, equivalently, in the strategy) cannot be anticipative, that is cannot depend on the future of the stock price).

3.1. Adapted Processes. We here generalize the notion of filtration to a time-continuous market. We then call a filtration a non-decreasing family $(\mathcal{F}_t)_{t\geq 0}$ of sub-σ-fields of $\mathcal{A}$. In this framework, we say a Brownian motion $(B_t)_{t\geq 0}$ is an $(\mathcal{F}_t)_{t\geq 0}$-Brownian motion is, for any $0 \leq s \leq t$, the increment $B_t - B_s$ is independent of $\mathcal{F}_s$. As an example, a Brownian motion $(B_t)_{t\geq 0}$ is always a Brownian motion w.r.t. its own filtration $(\sigma(B_s, 0 \leq s \leq t))_{t\geq 0}$.

A process $(X_t)_{t\geq 0}$ is then said to be adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ if, for any $t \geq 0$, the r.v. X_t is $\mathcal{F}_t$-measurable. From a financial point of view, this means that there is no anticipation on the future.

3.2. Simple Processes. As a specific class of adapted processes, we can consider simple processes that are the random equivalent of left-continuous step-functions with compact supports. A process $(X_t)_{t\geq 0}$ is said to be simple if there exist an integer $n \geq 1$, $n+1$ reals $0 = t_0 < \cdots < t_n$ and n random variables $\xi_1, \dots, \xi_n$ such that

- For any $1 \leq k \leq n$, ξ_k is $\mathcal{F}_{t_{k-1}}$-measurable and square-integrable, i.e. $\mathbb{E}[\xi_{k-1}^2] < \infty$.
- For any $u \geq 0$,

$$X_u = \sum_{k=0}^{n-1} \xi_k \mathbf{1}_{(t_{k-1}, t_k]}(u).$$

We then define the integral of a simple process $(X_t)_{t\geq 0}$ as

$$\int_0^{+\infty} X_u dB_u = \sum_{k=0}^{n-1} \xi_k \big(B_{t_k} - B_{t_{k-1}} \big).$$

Then, we can prove

Proposition 3.1. *Given a simple process $(X_u)_{u\geq 0}$ as above, the stochastic integral of X along the variation of B is a square-integrable random variable and*

$$\mathbb{E}\bigg[\bigg(\int_0^{+\infty} X_u dB_u \bigg)^2 \bigg] = \mathbb{E} \int_0^{+\infty} X_u^2 du = \int_0^{+\infty} \mathbb{E}\big[X_u^2 \big] du.$$

In particular, given another simple process $(Y_u)_{u\geq 0}$,

$$\mathbb{E}\bigg[\int_0^{+\infty} X_u dB_u \int_0^{+\infty} Y_u dB_u \bigg] = \mathbb{E} \int_0^{+\infty} X_u Y_u du = \int_0^{+\infty} \mathbb{E}\big[X_u Y_u \big] du.$$

Pay attention that nothing says the distribution of the stochastic integral is Gaussian. Proposition 3.1 is referred to as "Itô's isometry".

3.3. General Case. Again, the general case is handled by a density argument. The point is to decide which processes can be approximated by simple processes in the L^2 sense. Basically, something might be said about the joint time-randomness measurability of such a process. We thus admit that any adapted process $(X_t)_{t\geq 0}$ with left-continuous path satisfying

$$\mathbb{E}\int_0^{+\infty} X_u^2 du < +\infty, \tag{3.4}$$

there exists a sequence $((X_t^n)_{t\geq 0})_{n\geq 0}$ of simple processes such that

$$\lim_{n\to+\infty} \mathbb{E}\int_0^{+\infty} |X_u - X_u^n|^2 du = 0.$$

(Pay a ttention that the integrals above are well-defined by a joint-measurability argument.) Processes that are left-continuous and adapted and that satisfy (3.4) are said to be in the space $\mathcal{M}_2(\mathbb{R}_+)$.

Then, by a completeness argument again, we can define the integral of X as the $L^2(\Omega)$-limit of the integrals of $(X^n)_{n\geq 0}$, the limit being independent of the choice of the approximating sequence. In this framework, Proposition 3.1 remains true.

3.4. Stochastic Integral as a Process. Again, we have in mind to extend the definition of the integral to intervals. A left-continuous and adapted process $(X_t)_{t\geq 0}$ is said to be in $\mathcal{M}_2([0,T])$, for some $T > 0$, if

$$\mathbb{E}\int_0^T X_u^2 du < +\infty,$$

Then, we can set

$$\int_0^T X_u dB_u = \int_0^{+\infty} \mathbf{1}_{(0,T]}(u) X_u dB_u.$$

Itô's isometry says that

$$\mathbb{E}\Big[\Big(\int_0^T X_u dB_u\Big)^2\Big] = \mathbb{E}\int_0^T X_u^2 du.$$

To let the bound T vary in a convenient way, we admit the following:

Proposition 3.2. *Given an adapted and left-continuous process* $(X_t)_{t\geq 0} \in \bigcap_{T\geq 0} \mathcal{M}_2([0,T])$*, there exists a continuous process denoted by*

$$\Big(\int_0^t X_u dB_u\Big)_{t\geq 0},$$

such that, for any $t \geq 0$, the random variable

$$\int_0^t X_u dB_u$$

is the stochastic integral of X along the variation of B on the interval $[0, t]$.

CHAPTER 3

Stochastic Integrals as Martingales

1. Martingales in Discrete Time

1.1. Conditional Expectation. We here recall the basic properties of conditional expectation.

DEFINITION 1.1. *Given a probability space* $(\Omega, \mathcal{A}, \mathbb{P})$*, a sub* σ*-algebra* $\mathcal{F} \subset \mathcal{A}$ *and an integrable random variable* X*, there exists a random variable* Y*, integrable and* $\mathcal{F}$*-measurable, such that for any bounded* $\mathcal{F}$*-measurable random variable* Z*,*

$$\mathbb{E}\big[ZX\big] = \mathbb{E}\big[ZY\big]. \tag{1.5}$$

Moreover, Y *is unique to up to an event of zero* $\mathbb{P}$*-probability. It is called* conditional expectation of X given $\mathcal{F}$ *and is denoted by* $\mathbb{E}(X|\mathcal{F})$*.*

We here specify some points in the definition above: choosing $Z = Y/|Y|\mathbf{1}_{\{Y\neq 0\}}$, we deduce that $\mathbb{E}(|Y|) \leq \mathbb{E}(|X|)$.

Eq. (1.5) also holds for any square-integrable r.v. Z whenever X is square-integrable itself. In such a case, Y is also square-integrable and, choosing $Z = Y$, we deduce that $\mathbb{E}(Y^2) \leq \mathbb{E}(X^2)$.

The conditional expectation of X given $\mathcal{F}$ must be understood as the best approximation one can give for X when having at hand all the information described by $\mathcal{F}$. In other words, when all the information described by $\mathcal{F}$ is know, the best possible approximation for X is $\mathbb{E}(X|\mathcal{F})$ precisely. In particular, when X is $\mathcal{F}$-measurable itself, the best approximation is X itself. Similarly, when X is independent of $\mathcal{F}$, the best approximation is $\mathbb{E}(X)$ since $\mathcal{F}$ gives no information on X.

The conditional expectation given $\mathcal{F}$ is linear.

1.2. Properties of Conditional Expectation. Here is a list of basic properties to compute conditional expectation

PROPOSITION 1.2. *Let* $\mathcal{F}$ *be a sub-*σ*-algebra of* $\mathcal{A}$ *and* X *be an integrable r.v. Then,*

(1) $\mathbb{E}(X|\mathcal{F}) \geq 0$ *if* $X \geq 0$*,*
(2) $\mathbb{E}(\mathbb{E}(X|\mathcal{F})) = \mathbb{E}(X)$*,*
(3) $\mathbb{E}(X|\mathcal{F}) = X$ *if* X *is* $\mathcal{F}$*-measurable,*

(4) $\mathbb{E}(X|\mathcal{F}) = \mathbb{E}(X)$ *if* X *is independent of* $\mathcal{F}$,
(5) $\mathbb{E}(ZX|\mathcal{F}) = Z\mathbb{E}(X|\mathcal{F})$ *if* Z *is bounded and* $\mathcal{F}$*-measurable,*
(6) $\mathbb{E}(\mathbb{E}(X|\mathcal{G})|\mathcal{F}) = \mathbb{E}(X|\mathcal{F})$ *if* $\mathcal{G}$ *is another sub-*σ*-algebra of* $\mathcal{A}$ *such that* $\mathcal{F} \subset \mathcal{G}$,

1.3. Martingales.

Definition 1.3. *Given a discrete filtration* $(\mathcal{F}_n)_{n\geq 0}$*, an adapted process* $(M_n)_{n\geq 0}$ *is said to be a martingale if it is integrable, that is* M_n *is integrable for any* $n \geq 0$*, and satisfies*

$$\mathbb{E}(M_{n+1}|\mathcal{F}_n) = M_n, \quad n \geq 0.$$

In particular, a martingale is of constant expectation. Actually, this property may be extended as follows:

Proposition 1.4. *We call a stopping time* τ *a r.v. with values in* $\mathbb{N} \cup \{+\infty\}$ *such that* $\{\tau \leq n\} \in \mathcal{F}_n$ *for any* $n \geq 0$.

Given a bounded stopping time τ *(i.e.* $\tau \leq K$*) and a martingale* $(M_n)_{n\geq 0}$, $\mathbb{E}(M_\tau) = \mathbb{E}(M_0)$.

A stopping time must be thought as a random time associated to the realization of some phenomenon: when occurring at time n, the phenomenon must be depend on the past before n only. For example, $\tau = \inf\{n \geq 0 : X_n \in A\}$ is a stopping time when $(X_n)_{n\geq 0}$ is an adapted process and A is a Borel subset of $\mathbb{R}$.

In what follows, we will think of positive martingales as specific wealth processes associated with some financial strategy. These are wealth processes of constant mean. Proposition 1.4 says that the mean cannot be increased by stopping the strategy at a random stopping time, that is by stopping the strategy according to the observations of the past.

1.4. Discrete Stochastic Integral as a Martingale. The notion of martingale for a wealth can be easily specified on the Cox and Rubinstein model. With the notation introduced in Chapters 1 and 2, consider indeed a self-financing integrable strategy $(\phi_n^0, \phi_n)_{n\geq 1}$. The associated wealth expands as

$$W_n = \phi_{n+1}^0 S_n^0 + \phi_{n+1} S_n = \phi_n^0 S_n^0 + \phi_n S_n.$$

Computed now the discounted local increment of the wealth:

$$(1+r)^{-1} W_{n+1} - W_n = \phi_{n+1}^0 \big((1+r)^{-1} S_{n+1}^0 - S_n^0\big) + \phi_{n+1}\big((1+r)^{-1} S_{n+1} - S_n\big).$$

Now, we know that $S_{n+1}^0 = (1+r)S_n^0$ and $S_{n+1} = \xi_{n+1} S_n$, so that

$$(1+r)^{-1} W_{n+1} - W_n = \phi_{n+1}\big((1+r)^{-1}\xi_{n+1} - 1\big) S_n.$$

In particular,

$$\mathbb{E}\big((1+r)^{-1}W_{n+1} - W_n|\mathcal{F}_n\big) = \phi_{n+1}\big[(1+r)^{-1}\mathbb{E}(\xi_{n+1}) - 1\big]S_n,$$

where $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_n = \sigma(\xi_1, \dots, \xi_n)$.

Assume in particular that $\mathbb{E}(\xi_n) = 1 + r$ for any $n \geq 1$. Then,

$$\mathbb{E}\big((1+r)^{-1}W_{n+1} - W_n|\mathcal{F}_n\big) = 0,$$

that is the discounted wealth process $((1+r)^{-n}W_n)_{n\geq 0}$ is a martingale.

We deduce the following rule: if the mean benefit rate $\mathbb{E}(\xi_1)$ of the risky asset is equal to the interest rate of the riskless asset, then the discounted wealth associated with any self-financing strategy is martingale.

1.5. Arbitrage and Neutral-risk Measure.

Definition 1.5. *An integrable self-financing strategy $(\phi_n^0, \phi^n)_{n\geq 1}$ is said to be an arbitrage opportunity if there exists a bounded stopping time τ such that the associated wealth $(W_n)_{n\geq 0}$, with $W_0 = 0$ as initial capital, satisfies at time τ: $W_\tau \geq 0$ and $\mathbb{P}\{W_\tau > 0\} > 0$.*

In other words, an arbitrage is an opportunity to make some money (in finite time) without any risk. It should be somewhat avoided in practice.

Proposition 1.6. *In the Cox and Rubinstein model, assume that $\mathbb{E}(\xi_1) = (1+r)$. Then, the market is arbitrage free.*

Proof. Any integrable self-financing strategy, with $W_0 = 0$, satisfies $\mathbb{E}((1+r)^{-\tau}W_\tau) = 0$. □

Actually, the condition $\mathbb{E}(\xi_1) = (1+r)$ is really restrictive in practice. Actually, we even expect $\mathbb{E}(\xi_1)$ to be (much) greater than $1+r$ since risk is usually awarded by a premium. To get rid of this condition, we emphasize that the notion of arbitrage opportunity is preserved by change of equivalent measure: if $\mathbb{Q}$ is another probability measure with the same zero events as $\mathbb{P}$ (that is $\mathbb{P} \sim \mathbb{Q}$), then any arbitrage under $\mathbb{P}$ is also an arbitrage under $\mathbb{Q}$ and conversely. In other words,

Proposition 1.7. *The Cox and Rubinstein model is arbitrage free if and only if there exists a probability measure, called neutral-risk probability measure, equivalent to $\mathbb{P}$, under which the r.v.'s $(\xi_n)_{n\geq 1}$ are i.i.d with $1+r$ as mean exactly.*

We observe that a necessary condition is that $d \leq 1+r \leq u$ as the expectation is a convex combination of d and u. Actually, the boundary values $1+r = d$ and $1+r = u$ should be avoided as well since $d = 1+r$

or $u = 1+r$ means that the $(\xi_n)_{n\geq 1}$ take always the same values, which is not equivalent to the original model.

PROPOSITION 1.8. *Assume that $d < 1 + r < u$, then the Cox and Rubinstein model is arbitrage free.*

PROOF. To simplify, we assume that time is bounded by some N in the Cox and Rubinstein model. Then, we can choose Ω as $\{d, u\}^N$ and $\mathbb{P}$ as $(p\delta_u + (1-p)\delta_d)^{\otimes N}$. Choosing q as the solution of the equation $qu + (1-q)d = 1 + r$, we then define $\mathbb{P}^\star = (q\delta_u + (1-q)\delta_d)^{\otimes N}$. □

2. Stochastic Integrals as a Continuous-time Martingale

2.1. Continuous-time Martingale.

DEFINITION 2.1. *Given a filtration $(\mathcal{F}_t)_{t\geq 0}$ in continuous time, an integrable and adapted process $(M_t)_{t\geq 0}$ is said to be a martingale if, for any $0 \leq s \leq t$, $\mathbb{E}[M_t|\mathcal{F}_s] = M_s$.*

A typical example is the Brownian motion itself w.r.t to any Brownian filtration. Given an $(\mathcal{F}_t)_{t\geq 0}$-Brownian motion, it is well checked that $\mathbb{E}(B_t|\mathcal{F}_s) = B_s$. Another example is $(B_t^2 - t)_{t\geq 0}$.

2.2. Martingale Property of the Stochastic Integral. We here consider a simple process $(H_t)_{t\geq 0}$ of the form:

$$H_t = \sum_{i=1}^{n} X_i \mathbf{1}_{(t_{i-1},t_i]}(t), \tag{2.6}$$

where $0 = t_0 < t_1 < \cdots < t_n$ and X_i is a square integrable $\mathcal{F}_{t_{i-1}}$-measurable r.v. for any $1 \leq i \leq n$.

Define the associated integral process:

$$M_t = \sum_{i=1}^{n} X_i \big(B_{t_i\wedge t} - B_{t_{i-1}\wedge t}\big). \tag{2.7}$$

We claim that $(M_t)_{t\geq 0}$ is a martingale w.r.t. to any filtration $(\mathcal{F}_t)_{t\geq 0}$ for which $(B_t)_{t\geq 0}$ is an $(\mathcal{F}_t)_{t\geq 0}$-Brownian motion. Integrability and adaptedness are well-checked. It thus remains to check the projective martingale property. Given $0 \leq s \leq t$, we emphasize that $\mathbb{E}(B_{t_j\wedge t}|\mathcal{F}_s) = B_{t_j\wedge s}$. The result easily follows.

Using an approximation procedure, we could prove (but we omit the proof here):

Proposition 2.2. *Given an* $(\mathcal{F}_t)_{t\geq 0}$*-Brownian motion* $(B_t)_{t\ geq0}$ *and a process* $(H_t)_{t\geq 0}$ *in* $M^2_{\mathrm{loc}}(\mathbb{R}_+)$*, the stochastic integral*

$$\left(\int_0^t H_s dB_s\right)_{t\geq 0}$$

is a martingale.

2.3. Bracket of the Stochastic Integral. Go back to (2.6). An compute the square $(M_t^2)_{t\geq 0}$. We obtain

$$\begin{aligned} M_t^2 = \sum_{i=1}^n & X_i^2\big(B_{t_i\wedge t} - B_{t_{i-1}\wedge t}\big)^2 \\ & + 2 \sum_{1\leq i<j\leq n} X_i X_j \big(B_{t_i\wedge t} - B_{t_{i-1}\wedge t}\big)\big(B_{t_j\wedge t} - B_{t_{j-1}\wedge t}\big). \end{aligned}$$

Observe that all the r.v.'s above are integrable. Notice also that

$$\begin{aligned} &\mathbb{E}\big[X_i^2\big(B_{t_i\wedge t} - B_{t_{i-1}\wedge t}\big)^2|\mathcal{F}_s\big] \\ &= X_i^2\big(B_{t_i\wedge s} - B_{t_{i-1}\wedge s}\big)^2 + \mathbb{E}\big[X_i^2\big(t_i\wedge t - t_{i-1}\vee s\big)|\mathcal{F}_s\big]. \end{aligned}$$

Similarly,

$$\begin{aligned} &\mathbb{E}\big[X_i X_j\big(B_{t_i\wedge t} - B_{t_{i-1}\wedge t}\big)\big(B_{t_j\wedge t} - B_{t_{j-1}\wedge t}\big)|\mathcal{F}_s\big] \\ &= X_i X_j\big(B_{t_i\wedge s} - B_{t_{i-1}\wedge s}\big)\big(B_{t_j\wedge s} - B_{t_{j-1}\wedge s}\big). \end{aligned}$$

Therefore,

$$\begin{aligned} \mathbb{E}\big[M_t^2|\mathcal{F}_s\big] &= M_s^2 + \sum_{i=1}^n \mathbb{E}\big[X_i^2\big(t_i\wedge t - t_{i-1}\vee s\big)|\mathcal{F}_s\big] \\ &= M_s^2 + \left[\int_s^t H_r^2 dr|\mathcal{F}_s\right]. \end{aligned}$$

That is,

$$\mathbb{E}\left[M_t^2 + \int_0^t H_r^2 dr|\mathcal{F}_s\right] = M_s^2 + \int_0^s H_r^2 dr,$$

by noting that the process $(\int_0^t H_r^2 dr)_{t\geq 0}$ is adapted.

By an approximation procedure again, we deduce that

Proposition 2.3. *Given an* $(\mathcal{F}_t)_{t\geq 0}$*-Brownian motion* $(B_t)_{t\ geq0}$ *and a process* $(H_t)_{t\geq 0}$ *in* $M^2_{\mathrm{loc}}(\mathbb{R}_+)$*, the integral*

$$\left(\int_0^t H_s^2 ds\right)_{t\geq 0}$$

is a continuous and adapted process and the process

$$\left(\int_0^t H_s dB_s - \int_0^t H_s^2 ds\right)_{t\geq 0}$$

is a martingale.

In what follows, the process $(\int_0^t H_s^2 ds)_{t\geq 0}$ will be called the bracket of the martingale $(M_t = \int_0^t H_s dB_s)_{t\geq 0}$ and will be denoted by $([M]_t)_{t\geq 0}$.

CHAPTER 4

Itô's Formula for Brownian Motion – Black and Scholes Model

1. Itô's formula

1.1. Chain Rule for Differentiable Function. Recall that when f is a C^1 function and $(x_t)_{t\geq 0}$ is a C^1 path, the infinitesimal variation of $(f(x_t))_{t\geq 0}$ expand as

$$d\big(f(x_t)\big) = f'(x_t)\dot{x}_t dt = f'(x_t)dx_t.$$

Below, the point is to investigate a similar relationship, but for the Brownian motion. The difficult point is that the Brownian paths are nowhere differentiable so that the above formula fails.

1.2. Sketch of the Proof of the Chain Rule for BM. Given a smooth function f (smooth here means that it is enough differentiable for our specific purposes with bounded derivatives). We then compute by Taylor's formula

$$\begin{aligned}f(B_{t+h}) - f(B_t) = f'(B_t)(B_{t+h} - B_t)&\\ + \frac{1}{2}f''(B_t)(B_{t+h} - B_t)^2 &+ \|f^{(3)}\|_\infty O((B_{t+h} - B_t)^3).\end{aligned}$$

We observe that the typical size for $O((B_{t+h} - B_t)^3)$ is $h^{3/2}$. We also observe that $(B_{t+h} - B_t)^2$ has h as mean. So that we can write

$$\begin{aligned}f(B_{t+h}) - f(B_t) &= f'(B_t)(B_{t+h} - B_t)\\ &\quad + \frac{1}{2}f''(B_t)h + \frac{1}{2}f''(B_t)\big[(B_{t+h} - B_t)^2 - h\big]\\ &\quad + \|f^{(3)}\|_\infty O(h^{3/2}).\end{aligned}$$

Consider now a partition $0 = t_0 < t_1 < \cdots < t_N = t$ of step $t_{i+1} - t_i = t/N$. Choosing $h = t/N$ above and making the sum over i's in

$\{0, \dots, N-1\}$, we obtain

$$
\begin{aligned}
f(B_t) - f(B_0) = &\sum_{i=0}^{N-1} f'(B_{t_i})(B_{t_{i+1}} - B_{t_i}) \\
&+ \frac{1}{2}\sum_{i=0}^{N-1} f''(B_{t_i})(t_{i+1} - t_i) \\
&+ \frac{1}{2}\sum_{i=0}^{N-1} f''(B_{t_i})\big[(B_{t_{i+1}} - B_{t_i})^2 - (t_{i+1} - t_i)\big] \\
&+ \|f^{(3)}\|_\infty O(N^{-1/2}). \\
= &T_1^N + T_2^N + T_3^N + T_4^N.
\end{aligned}
$$

We then check that T_1^N converges towards $\int_0^t f'(B_s)dB_s$ in L^2 as N tends to $+\infty$, T_2^N converges towards $(1/2)\int_0^t f''(B_s)ds$ in probability as N tends to $+\infty$, T_3^N converges towards 0 in L^2 (take the square and use the independence of the r.v.'s) and T_4^N converges towards 0 in L^2 as well.

We thus obtain,

$$
f(B_t) = f(B_0) + \int_0^t f'(B_s)dB_s + \frac{1}{2}\int_0^t f''(B_s)ds,
$$

almost-surely. By continuity, the above relationship is true almost-surely for any t.

1.3. Precise Statement.

PROPOSITION 1.1. *Given a $\mathcal{C}^2$ function f from $\mathbb{R}$ to $\mathbb{R}$ such that $(f'(B_t))_{t\geq 0}$ is in $M^2_{\text{loc}}(\mathbb{R}_+)$, we have, almost-surely,*

$$
\forall t \geq 0, \quad f(B_t) = f(B_0) + \int_0^t f'(B_s)dB_s + \frac{1}{2}\int_0^t f''(B_s)ds.
$$

Actually, we can assume f to be time-dependent. In such a case, it is necessary to assume f to be continuously differentiable in time only. We claim

PROPOSITION 1.2. *Given a $\mathcal{C}^{1,2}$ function f from $\mathbb{R}_+ \times \mathbb{R}$ to $\mathbb{R}$ such that $(f'(t, B_t))_{t\geq 0}$ is in $M^2_{\text{loc}}(\mathbb{R}_+)$, we have, almost-surely,*

$$
\begin{aligned}
\forall t \geq 0, \quad f(t, B_t) = f(0, B_0) &+ \int_0^t \partial_x f(s, B_s)dB_s \\
&+ \int_0^t \big[\partial_t f(s, B_s) + \frac{1}{2}\partial^2_{x,x} f(s, B_s)\big]ds.
\end{aligned}
$$

We will also write

$$d\big(f(B_t)\big) = \partial_x f(t,B_t)dB_t\big[\partial_t f(s,B_s) + \frac{1}{2}\partial^2_{x,x} f(s,B_s)\big]dt.$$

Above $d(f(t,B_t))$ stands for the infinitesimal variation of $f(t,B_t)$ at time t.

1.4. Typical Example. Choose $f(t,x) = \exp(rt + \sigma x)$. Then,

$$d\big(\exp(rt+\sigma B_t)\big) = \big(r + \frac{\sigma^2}{2}\big)\exp(rt+\sigma B_t)dt + \sigma\exp(rt+\sigma B_t)dB_t.$$

The M^2 condition is well-checked by using the Laplace transform of a Gaussian random variable. (Check it at home.) This is the typical model for the dynamics of a stick price. We summarize things as:

$$dS_t = \mu S_t dt + \sigma S_t dB_t, \quad t \geq 0$$

with $\mu = r + \sigma^2/2$ and

$$S_t = \exp(rt + \sigma B_t), \quad t \geq 0.$$

Taking the expectation, we understand that μ drives the mean trend of the process and that σ drives the fluctuation of the process around its mean trend. In practice, σ is referred to as the "volatility" of the stock price. The Black Scholes model is then given by the pair $(S^0_t, S_t)_{t\geq 0}$, where S^0 stands for the riskless dynamics, that is $S^0_t = S^0_0\exp(r^0 t)$, $t \geq 0$.

We observe that for $\mu = r + \sigma^2/2 = 0$, the stock price is a martingale. This situation is rare in practice but we can switch back to such a case by modifying the probability. Precisely, we observe that

$$Q_T = \exp\big(-\sigma^{-1}\mu B_T - \frac{\sigma^{-2}\mu^2}{2}T\big)$$

has 1 as expectation for a given maturity time $T > 0$. Now,

$$\begin{aligned}
&\mathbb{E}\big[Q_T S_t|\mathcal{F}_s\big] \\
&= \mathbb{E}\big[\exp\big(-\sigma^{-1}\mu B_t - \frac{\sigma^{-2}\mu^2}{2}t\big)\exp(rt+\sigma B_t)|\mathcal{F}_s\big] \\
&= \exp\big((r-\sigma^{-2}\mu^2/2)t + (\sigma-\sigma^{-1}\mu)^2(t-s)/2\big)\exp\big((\sigma-\sigma^{-1}\mu)B_s\big) \\
&= \exp\big((r-\sigma^{-2}\mu^2/2)s + [(\sigma-\sigma^{-1}\mu)^2/2 + r - \sigma^{-2}\mu^2/2](t-s)/2\big) \\
&\quad \times \exp\big((\sigma-\sigma^{-1}\mu)B_s\big). \\
&= Q_s S_s,
\end{aligned}$$

so that

$$\mathbb{E}_{\mathbb{Q}_T}\big[S_t|\mathcal{F}_s\big] = S_s,$$

where $\mathbb{Q}_T$ is the probability admitting Q_T as density w.r.t. $\mathbb{Q}$. In other words, $\mathbb{Q}_T$ must be understood as kind of neutral-risk measure. (We here say a kind of since the discount is not taken into account.)

2. Itô Processes

2.1. Definition. We call an Itô process a process $(X_t)_{t\geq 0}$ of the form

$$X_t = X_0 + \int_0^t \psi_s ds + \int_0^t \varphi_s dB_s, \tag{2.8}$$

where X_0 is an $\mathcal{F}_0$-measurable r.v., $(\psi_t)_{t\geq 0}$ is a left-continuous adapted process and $(\varphi_t)_{t\geq 0}$ is a left-continuous adapted process such that

$$\forall T > 0, \quad \mathbb{E}\int_0^T \varphi_s^2 ds < +\infty.$$

The infinitesimal variation expand as

$$dX_t = \psi_t dt + \varphi_t dB_t.$$

2.2. Self-financing Strategy. Following the discussion on the Cox and Rubinstein market, we call a self-financing strategy a pair $(\varphi_t^0, \varphi_t)_{t\geq 0}$ of bounded, left-continuous and adapted processes such that the wealth

$$W_t = \varphi_t^0 S_t^0 + \varphi_t S_t, \quad t \geq 0,$$

admits the infinitesimal expansion

$$dW_t = \varphi_t^0 dS_t^0 + \varphi_t dS_t, \quad t \geq 0.$$

In particular $(W_t)_{t\geq 0}$ is an Itô process driven by φ^0 and φ. Assume for the moment that

$$\begin{aligned} d\big(\exp(-r^0 t)W_t\big) &= -r_t^0 \exp(-r^0 t)W_t dt + \exp(-r^0 t)dW_t \\ &= -r_t^0 \exp(-r^0 t)\varphi_t S_t dt + \exp(-r^0 t)\varphi_t dS_t \end{aligned}$$

Assume now that $\exp(-r^0 t)S_t = \exp(\sigma^2 t - \sigma B_t)$, $t \geq 0$, so that it is a martingale. Then,

$$d\big(\exp(-r^0 t)W_t\big) = \varphi_t \exp(-r^0 t)S_t dB_t,$$

so that the discounted wealth is a martingale as well.

2.3. Example of Pricing. Assume that you are seeking for the price of a European option, that is the premium to cover $(S_T - K)^+$ at time T. Under a neutral-risk measure $\mathbb{P}^*$ (if exists), the stock price is a martingale so that whatever the strategy is, the initial wealth required to generate $(S_T - K)^+$ is equal to $\exp(-r^0 T)\mathbb{E}_{\mathbb{Q}}((S_T - K)^+)$.

2.4. Processes driven Noises of Dimension Greater than 2. In practice, the noise is often multi-dimensional. That is, the underlying BM $(B_t = (B_t^1, \dots, B_t^d))_{t \geq 0}$ is of dimension d, with $d \geq 2$. In such a case, we consider as filtration the family

$$\mathcal{F}_t = \sigma(B_s^1, \dots, B_s^d, s \leq t).$$

Each coordinate of B is an $(\mathcal{F}_t)_{t \geq 0}$-BM. In particular, we can consider stochastic integrals w.r.t. any of the coordinate of B. In such a framework, we call an Itô process a process of the form

$$X_t = X_0 + \int_0^t \psi_s ds + \sum_{i=1}^d \int_0^t \varphi_s^i dB_s^i, \quad t \geq 0, \tag{2.9}$$

the processes $(\psi_t)_{t \geq 0}$ and $(\varphi_t^1, \dots, \varphi_t^d)_{t \geq 0}$ being left-continuous.

2.5. Bracket of an Itô Process. By definition, we will call bracket of an Itô process as in (2.8) the process

$$\langle X \rangle_t = \int_0^t \varphi_s^2 ds.$$

When ψ is equal to zero in (2.8), this definition is motivated by the martingale property we proved in the Chapter on martingales: the process $(X_t^2 - \langle X \rangle_t)_{t \geq 0}$ is a martingale (w.r.t. to the underlying filtration). More generally, given two Itô processes X^1 and X^2 with the same kind of expansion as in (2.8), but driven by (ψ^1, φ^1) and (ψ^2, φ^2) respectively, we put

$$\langle X^1, X^2 \rangle_t = \int_0^t \varphi_s^1 \varphi_s^2 ds.$$

Again, this definition is motivated by the martingale property: $(X_t^1 X_t^2 - \langle X^1, X^2 \rangle_t)_{t \geq 0}$ is a martingale when $\psi^1 = \psi^2 = 0$

In the case when the noise has a dimension greater or equal to 2, the definition of the bracket is

$$\langle X^1, X^2 \rangle_t = \sum_{i=1}^d \int_0^t \varphi_s^{1,i} \varphi_s^{2,i} ds,$$

the definition being motivated again by a martingale property.

The basic application is the integration by parts formula

PROPOSITION 2.1. *With the notations above,*

$$d\big(X_t^1 X_t^2\big) = X_t^1 dX_t^2 + X_t^2 dX_t^1 + d\langle X^1, X^2 \rangle_t, \quad t \geq 0,$$

provided

$$\forall t \geq 0, \quad \sum_{i=1}^{d} \mathbb{E} \int_0^t \big(|X_s^1 \varphi_s^{2,i}|^2 + |X_s^2 \varphi_s^{1,i}|^2\big) ds < +\infty.$$

Again, the proof follows from a Taylor expansion at order two: assume that the processes $\psi^1, \psi^2, \varphi^{1,1}, \varphi^{2,1}, \dots, \varphi^{1,d}, \varphi^{2,d}$ are bounded and continuous and compute

$$\begin{aligned} X_{t+h}^1 X_{t+h}^2 = X_t^1 X_t^2 + X_t^1 \big(X_{t+h}^2 - X_t^2\big) + X_t^2 \big(X_{t+h}^1 - X_t^1\big) \\ + \big(X_{t+h}^2 - X_t^2\big)\big(X_{t+h}^1 - X_t^1\big). \end{aligned}$$

Now, we can write:

$$\begin{aligned} X_{t+h}^1 X_{t+h}^2 &= X_t^1 X_t^2 \\ &+ X_t^1 \psi_t^2 h + X_t^1 \sum_{i=1}^{d} \varphi_t^{2,i} (B_{t+h}^i - B_t^i) \\ &+ X_t^2 \psi_t^1 h + X_t^1 \sum_{i=1}^{d} \varphi_t^{1,i} (B_{t+h}^i - B_t^i) \\ &+ X_t^1 \int_t^{t+h} (\psi_s^2 - \psi_t^2) ds + \sum_{i=1}^{d} X_t^1 \int_t^{t+h} \big(\varphi_s^{2,i} - \varphi_t^{2,i}\big) dB_s^i \quad (= R_{t,h}^1) \\ &+ X_t^2 \int_t^{t+h} (\psi_s^1 - \psi_t^1) ds + \sum_{i=1}^{d} X_t^2 \int_t^{t+h} \big(\varphi_s^{1,i} - \varphi_t^{1,i}\big) dB_s^i \quad (= R_{t,h}^2) \\ &+ \big(X_{t+h}^2 - X_t^2\big)\big(X_{t+h}^1 - X_t^1\big). \end{aligned}$$

We notice that R^1 and R^2 vanish in L^1 when summed over a mesh of $[0,T]$ as the stepsize tends to 0. (Since the coefficients are bounded, X is clearly in L^2.)

Finally,

$$\begin{aligned} &\big(X_{t+h}^2 - X_t^2\big)\big(X_{t+h}^1 - X_t^1\big) \\ &= \sum_{i,j=1}^{d} \varphi_t^{1,i} \varphi_t^{2,j} (B_{t+h}^i - B_t^i)(B_{t+h}^j - B_t^j) + R_{t,h}^3, \end{aligned}$$

where R^3 vanishes in probability when summed over a mesh of $[0,T]$ as the stepsize tends to 0. We notice that

$$\sum_{i \neq j} \varphi_t^{1,i} \varphi_t^{2,j} (B_{t+h}^i - B_t^i)(B_{t+h}^j - B_t^j)$$

tend to 0 in L^2. The end of the proof the consists in showing that

$$\sum_{i=1}^{d} \varphi_t^{1,i} \varphi_t^{2,i} (B_{t+h}^i - B_t^i)^2$$

converges in probability towards the Riemann integral of

$$\left(\sum_{i=1}^{d} \varphi_t^{1,i} \varphi_t^{2,i} \right)_{t \geq 0}.$$

3. General Itô Formula

3.1. Itô Formula for Itô Processes. On the same model as the one used to get Itô formula for BM, we claim

THEOREM 3.1. *Let f be a $\mathcal{C}^{1,2}$ function from $\mathbb{R}_+ \times \mathbb{R}$ into $\mathbb{R}$ and X be an Itô process as in* (2.8)*, then, a.s.,*

$$\begin{aligned} \forall t \geq 0, \quad f(t, X_t) = f(0, X_0) &+ \int_0^t \partial_t f(s, X_s) ds \\ &+ \int_0^t \partial_x f(s, X_s) dX_s + \frac{1}{2} \int_0^t \partial_{x,x}^2 f(s, X_s) ds, \end{aligned}$$

provided

$$\forall t \geq 0, \quad \sum_{i=1}^{d} \mathbb{E} \int_0^t |\partial_x (s, X_s)|^2 ds < +\infty.$$

3.2. Multi-Dimensional Itô Formula. On the same model as the one used to get Itô formula for BM, we claim

THEOREM 3.2. *Let f be a $\mathcal{C}^{1,2}$ function from $\mathbb{R}_+ \times \mathbb{R}^d$ into $\mathbb{R}^d$ and X be an Itô process as in* (2.9)*, then, a.s.,*

$$\begin{aligned} \forall t \geq 0, \quad f(t, X_t) = f(0, X_0) &+ \int_0^t \partial_t f(s, X_s) ds \\ &+ \sum_{i=1}^{d} \int_0^t \partial_{x_i} f(s, X_s) dX_s + \frac{1}{2} \int_0^t \partial_{x,x}^2 f(s, X_s) ds, \end{aligned}$$

provided

$$\forall t \geq 0, \quad \sum_{i=1}^{d} \mathbb{E} \int_0^t |\partial_{x_i} (s, X_s)|^2 ds < +\infty.$$