Option pricing

Lars Haringa
Universiteit Utrecht

Supervisor: dr. Karma Dajani

August 31, 2014

Abstract

Using mathematical techniques at undergraduate level, an introduction to axiomatic probability theory and stochastic calculus facilitates the classic derivation of the Black-Scholes-Merton approach in valuating a European option. Brownian motion is derived as the limit of a scaled symmetric random walk and its quadratic variation is determined. This serves to evaluate the Itô integral and the Itô-Doeblin change-of-variables formula. After employing these equations to arrive at the partial differential equation for the option value, the solution is determined by the use of an equivalent risk-neutral measure.

Contents

1 Preface **2**

2 Stochastics and time **3**
2.1 Basics . . . 3
2.2 Stochastic processes . . . 6
2.3 Intermezzo: continuous discounting . . . 8

3 Stochastic calculus **9**
3.1 Brownian motion . . . 9
3.2 Properties of Brownian motion . . . 11
3.3 Comparison of computations of integrals . . . 15
3.4 Itô calculus . . . 17

4 Option pricing **24**
4.1 Derivation of Black-Scholes-Merton equation . . . 25
4.2 Solution to the option value formula . . . 30

Chapter 1

Preface

This paper was written as a bachelor's thesis under supervision of dr. Karma Dajani. It seeks to provide an introduction to the subjects of stochastic calculus as applied to the theory of mathematical finance. Steven E. Shreve's *Stochastic calculus for finance II* has been of great help and has served as a guide for this writing. An attempt has been made to link this approach to more formal theory; these subjects are however for the largest part beyond the scope of this paper. In particular, the theory of stochastic integration, such as presented in Henry McKean's *Stochastic integrals*, would be an appropriate start for further study.

Chapter 2

Stochastics and time

In this chapter, elements of probability and measure are presented to support the forthcoming theory. Most subjects are concisely stated and serve mainly as reference. For a comprehensive introduction, see for example [Ric07] on probability theory and statistics, and [Sch05] on measure theory.

2.1 Basics

Let $(\Omega, \mathcal{F}, P)$ be a probability space; that is,

- Ω is a set containing abstract *events* $\omega \in \Omega$, thought to occur randomly,
- $\mathcal{F}$ is a *σ-algebra* on Ω, a family of subsets of Ω that includes among its elements Ω and the complement of each $F \in \mathcal{F}$, and that is closed under countable union, and
- $P : \mathcal{F} \to [0, 1]$ is a *probability measure*: it satisfies $P(\Omega) = 1$ and it is countably additive on $\mathcal{F}$.

Let furthermore $\mathcal{B}$ be the *Borel σ-algebra* on $\mathbb{R}$. The Borel σ-algebra is generated by the open sets in $\mathbb{R}$ and for this reason also called the *topological σ-algebra* on $\mathbb{R}$. In general, a family of subsets is said to *generate* a σ-algebra if this σ-algebra as small as possible while containing all subsets and satisfying the σ-algebra axioms.

Let $\mathcal{C}(\mathbb{R}) := \{[a, b] \subsetneq \mathbb{R}\}$ be the set of closed real intervals, then the σ-algebra generated by $\mathcal{C}(\mathbb{R})$ is $\mathcal{B}$ [Sch05, 3.7] as well. This implies that all unions and intersections of open and/or closed sets are Borel sets. In general, a σ-algebra $\mathcal{A}$ on a set $\mathcal{S}$ generated by a family $\mathcal{T}$ of subsets of $\mathcal{S}$ is written $\sigma(\mathcal{T}) = \mathcal{A}$.

A *random variable* $X : \Omega \to \mathbb{R}$ is a function that assigns any event $\omega \in \Omega$ a real value $X(\omega) \in \mathbb{R}$. As such it may be thought of as a quantification of events. If the event ω takes place, X *results* in $X(\omega)$. The probability of X resulting in a value contained in some Borel set B is given by the *distribution measure* $\mu_X : \mathcal{B} \to [0,1]$, which itself is based on P by the definition

$$\mu_X(B) := P\left(\{\omega \in \Omega \mid X(\omega) \in B\}\right)$$

or in common shorthand with $x \in \mathrm{img}(X)$,

$$\mu_X(B) = P\left(\{x \in B\}\right).$$

Any probability measure defined on $\mathcal{C}(\mathbb{R})$ extends uniquely to $\mathcal{B}$ [Sch05, 5.7], so the distribution measure defined as $\mu_X\left([a,b]\right)$ completely specifies $\mu_X : \mathcal{B} \to \mathbb{R}$.

Usually it is mathematically convenient to work with the image of a random variable and its distribution measure, as opposed to the abstract σ-algebra and the corresponding probability measure P directly. A particular case is when the random variable X is continuous with a specified *probability distribution function* f; in this case, the probability that X results in $[a,b]$ is given by

$$\int_a^b f(x)dx,$$

which is equal to the general integral

$$\int_{\omega\in\Omega} \mathbf{1}_A dP(\omega) \text{ with } A := \{\omega \in \Omega \mid X(\omega) \in [a,b]\}.$$

Notice the appearance of an *indicator function*: with $F \in \mathcal{F}$, the indicator of F is

$$\mathbf{1}_F(\omega) := \begin{cases} 1 & \text{if } \omega \in F, \\ 0 & \text{if } \omega \notin F. \end{cases}$$

Common in statistics is the *normally distributed* random variable X with mean μ and variance σ^2, its probability distribution function f is given by

$$\text{(2.1)} \quad f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

and this fact is written $X \sim \mathcal{N}(\mu, \sigma 2)$. If in particular $X \sim \mathcal{N}(0,1)$, the random variable is said to follow *standard normal distribution*. It has a *cumulative density function* given by

$$\text{(2.2)} \quad \Phi(z) := \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx.$$

Probability distribution functions also come in handy when computing the *expectation* $\mathbf{E}$ of a function $g : \mathrm{img}(X) \to \mathbb{R}$ of X, since

$$\mathbf{E}[X] = \int_{-\infty}^{\infty} g(x)f(x)dx.$$

while the expectation in general is given by

$$\mathbf{E}[X] = \int_{\omega\in\Omega} g\left(X(\omega)\right) dP(\omega).$$

If $P(X \in B) = 1$ for $B \subseteq \mathrm{img}(X)$, then it is said that $X \in B$ *almost surely*. Notice that it is *possible* that $\exists \omega \in \Omega : X(\omega) \notin B$, but it is not *probable* that $X \notin B$; that is, $P(\{\omega \in \Omega \mid X(\omega) \notin B\}) = 0$.

When the pre-image of any Borel set is an element of $\mathcal{F}$—that is, when

$$\forall B \in \mathcal{B} : X^{-1}(B) \in \mathcal{F},$$

then X is $\mathcal{F}$*-measurable*. Notice that for any Borel set $\tilde{B}$ such that $\tilde{B} \cap \mathrm{img}(X) = \emptyset$ it holds that $X^{-1}(\tilde{B}) = \emptyset \in \mathcal{F}$, and any σ-algebra contains $\emptyset$ by definition. The definition of measurability implies that any result $x \in \mathrm{img}(X)$ can be uniquely identified with a subset of events $X^{-1}(x) \subseteq \Omega$, since $X^{-1}(x) \in \mathcal{F}$. As such, X indicates which events $\omega \in F$ may have happened. Conversely, with $F \in \mathcal{F}$, the sets $X(F)$ cover the full range of results $\mathrm{img}(X) \subseteq \mathbb{R}$. For example, take $\Omega = \mathbb{R}$ and $\mathcal{F} = \{\emptyset, \mathbb{Q}, \mathbb{R}\backslash\mathbb{Q}, \mathbb{R}\}$ (a σ-algebra), and let $\mathbf{1}_{\mathbb{Q}} : \mathbb{R} \to \{0, 1\}$ be given by

$$\mathbf{1}_{\mathbb{Q}}(\omega) = \begin{cases} 0 & \text{if } \omega \notin \mathbb{Q}, \\ 1 & \text{if } \omega \in \mathbb{Q}. \end{cases} \tag{2.3}$$

an indicator for rational numbers among the real numbers. Then $\mathbf{1}_{\mathbb{Q}}$ is $\mathcal{F}$-measurable: with $\tilde{\omega} \in \Omega$, a result $\mathbf{1}_{\mathbb{Q}}(\tilde{\omega}) \in \{0, 1\}$ indicates whether $\tilde{\omega}$ is rational or not, but does not reveal more information on $\tilde{\omega}$.

General measures $\mu : \mathcal{F} \to \mathbb{R}_{\geq 0}$ satisfy both $\mu(\emptyset)$ and countable additivity. In measure theory, measures may be employed to define a general kind of integrals, as has been implicitly suggested above. For a full treatise see for example [Sch05]. For any measure μ and any function $f : \Omega \to \mathbb{R}$, this function is μ*-integrable* if $\int |f| d\mu < \infty$. The following result is typical for the technique used in measure theory.

2.1.1 Theorem (Monotone convergence)**.** *Let $(X, \mathcal{F}, \mu)$ be a measure space and let $u_j : X \to \mathbb{R}$ with $j \in \mathbb{N}_{\geq 1}$ be μ-integrable functions such that $u_j \leq u_{j+1}$ and with (pointwise) limit $u := \lim_{j\to\infty} u_j$. Then u is μ-integrable if and only if $\lim_{j\to\infty} \int u_j d\mu < \infty$, and if so, then*

$$\int u d\mu = \lim_{j\to\infty} \int u_j d\mu.$$

Proof. See for instance [Sch05, theorem 11.1]. ∎

The above theorem serves to prove the following proposition, in turn required for the solution to the option pricing formula in chapter 4.

2.1.2 Proposition (Change of probability measure)**.** *Let $Z : \Omega \to \mathbb{R}_{\geq 0}$ be a random variable satisfying*

$$\mathbf{E}[Z] = \int_{\omega\in\Omega} Z(\omega) dP(\omega) = 1.$$

With $F \in \mathcal{F}$ it holds that

$$\tilde{P}(F) := \int_{\omega\in\Omega} \mathbf{1}_F(\omega) Z(\omega) dP(\omega)$$

is itself a probability measure.

Proof. Since Z and P are nonnegative, so is $\tilde{P}$. Furthermore

$$\tilde{P}(\Omega) = \int_{\omega\in\Omega} Z(\omega)dP(\omega) = \mathbf{E}[Z] = 1$$

by definition, so it remains to be shown that $\tilde{P}$ satisfies countable additivity.

Let $(F_n)_{n\in\mathbb{N}_{\geq 1}} \subseteq \mathcal{F}$ be disjoint and let $F := \bigcup_{n\in\mathbb{N}_{\geq 1}} F_n$. Define $B_j := \bigcup_{n=1}^{j} F_n$. Notice $\mathbf{1}_{B_j}$ is P-integrable since $\int \mathbf{1}_{B_j} dP = P(\omega \in B_j)$, and that

$$\lim_{j\to\infty} \mathbf{1}_{B_j} = \mathbf{1}_F.$$

As such,

$$\tilde{P}(F) = \int_{\omega\in\Omega} \mathbf{1}_F(\omega)Z(\omega)dP(\omega) = \lim_{j\to\infty} \int_{\omega\in\Omega} \mathbf{1}_{B_j}(\omega)Z(\omega)dP(\omega)$$

by monotone convergence 2.1.1. Since F_n are disjoint, $\mathbf{1}_{B_j}(\omega) = \sum_{n=1}^{j} \mathbf{1}_{F_n}(\omega)$. Therefore

$$\begin{aligned}\tilde{P}(F) &= \lim_{j\to\infty} \int_{\omega\in\Omega} \sum_{n=1}^{j} \mathbf{1}_{F_n}(\omega)Z(\omega)dP(\omega) \\ &= \lim_{j\to\infty} \sum_{n=1}^{j} \int_{\omega\in\Omega} \mathbf{1}_{F_n}(\omega)Z(\omega)dP(\omega) = \sum_{n=1}^{\infty} \tilde{P}(F_n),\end{aligned}$$

showing $\tilde{P}$ satisfies countable additivity, so it is a probability measure. ■

2.2 Stochastic processes

Stochastic phenomena through time—as opposed to the static framework of the previous section—may be modeled by stochastic *processes* adapted to a *filtration*. Both depend on the same time variable t, and intuitively, a filtration reflects the information available at this time. The process is random for all time 'after' t, and nonrandom (known) for all time up to t. These concepts will be used to simulate the random path of an asset's value.

Take $T \in \mathbb{R}_{\geq 0}$ and $t \in [0, T]$. A *filtration* is an inclusion of σ-algebras $\mathcal{F}(t)$ on Ω:

$$\forall t, \forall s \in [0, t] : \mathcal{F}(s) \subseteq \mathcal{F}(t).$$

Notice for each t there is a probability space $(\Omega, \mathcal{F}(t), P)$. An *adapted stochastic process* is a sequence $(X(t))_{t\in[0,T]}$ ($X(t)$ for short) of random variables $X(t) : \Omega \to \mathbb{R}$ such that $\forall t$ it holds that $X(t)$ is $\mathcal{F}(t)$-measurable. Notice the variable t does *not* correspond to the domain Ω; a more correct (but in this paper rarely used) notation is $X_t : \Omega \to \mathbb{R}$ with realization $X_t(\omega)$. As t increases, if some inclusions of the filtration are strict, the σ-algebra $\mathcal{F}(t)$ becomes *finer*—contains more, thus smaller, subsets of Ω. Because $X(t)$ is $\mathcal{F}(t)$-measurable, this implies that for some event $\tilde{\omega} \in \Omega$, learning a result $X_t(\tilde{\omega}) \in \operatorname{img}(X_t)$ reveals more information on $\tilde{\omega}$ as t increases. Often $\mathcal{F}(0) = \{\emptyset, \Omega\}$, the trivial σ-algebra, and $\mathcal{F}(T) = \sigma\left(\{X_t^{-1}(B) \mid B \in \mathcal{B} \wedge t \in [0,T]\}\right)$, such that $X(0)$ must be a constant (noninformative) function and $X(T)$ reveals the most information of all $X(t)$ about the event that its result indicates.

It may be useful to imagine Ω to consist of blueprints for functions $[0,T] \to \mathbb{R}$ that are realizations of the random variable $X(t)$. As such, the path of $X(t)$ is 'chosen' deterministically by one random event in Ω, but the outcome is only gradually revealed as t increases.

The above concepts can be applied for $t \in S$ for any ordered index set S, discrete or continuous. To reflect the assumed characteristics of financial trade, this paper is restricted to the continuous interval $S = [0,T]$.

With a filtration gradually uncovering information, the expectation of a stochastic process $X(t)$ is dependent on the time of evaluation. With $s \in [0,t]$, the *conditional expectation*

$$\mathbf{E}\left[X(t)|\mathcal{F}(s)\right]$$

represents the expectation if the information contained in $\mathcal{F}(s)$ is revealed. If this information is not relevant to the expectation, the random variable $X(t)$ is *independent* of $\mathcal{F}(s)$, and

$$\mathbf{E}\left[X(t)|\mathcal{F}(s)\right] = \mathbf{E}\left[X(t)\right].$$

In finance, an important class of stochastic processes is that of the martingales. A stochastic process $(X(t))_{t\in[0,T]}$ adapted to $(\mathcal{F}(t))_{t\in[0,T]}$ is a *martingale* if for all $s \in [0,t]$ it holds that

$$\mathbf{E}\left[X(t)|\mathcal{F}(s)\right] = X(s); \tag{2.4}$$

that is, at any time s the expectation of future values is equal to the current value.

Likewise significant to finance is the Markov property. Let $(X(t))_{t\in[0,T]}$ be a stochastic process adapted to $(\mathcal{F}(t))_{t\in[0,T]}$. Take $s, t \in [0,T]$ with $s \leq t$ and let $f : \mathbb{R} \to \mathbb{R}_{\geq 0}$ be a nonnegative and Borel-measurable function. If for all such s, t, and f, there exists another Borel-measurable function $g : \mathbb{R} \to \mathbb{R}$ such that

$$\mathbf{E}\left[f\left(X(t)\right)|\mathcal{F}(s)\right] = g\left(X(s)\right) \tag{2.5}$$

then $X(t)$ is called a *Markov process*. Essentially this means that for any $t \geq s$, the expected value of $X(t)$ at time s depends only on the current time s and the value of $X(s)$. In particular, any values $X(u)$ with $u \in [0, s[$ are not relevant for the expectation; hence, the Markov property is sometimes referred to as *memorylessness*.

In further chapters, a stochastic process $(X(t))_{t\in[0,T]}$ will often be denoted by the shorthand $X(t)$.

2.3 Intermezzo: continuous discounting

Finance, dealing with money, is frequently concerned with calculating profit from interest on value. A typical approach with mathematical attractiveness (and limited real-world applicability) is presented in this section.

Take $p \in \mathbb{N}_{\geq 0}$ and $r \in \mathbb{R}_{\geq 0}$. Call p a period and let r be the *nominal* or *flat* rate of return per period of some value V_p, meaning that

$$V_p = (1+r)^p \cdot V_0.$$

Notice the convention $1+r$, which illustrates that an interest rate is not an interest factor but rather a factorial increment.

An interest rate is said to *compound* m times per p if the value of V_p is instead given by

$$V_p = \left(1 + \frac{r}{m}\right)^{mp} \cdot V_0.$$

The quantity $\left(1 + \frac{r}{m}\right)^m - 1$ is called the *effective interest rate* and it is strictly increasing in m. It is therefore important in practical applications to indicate whether an interest rate is nominal or effective.

As the frequency of compounding approaches infinity, the interest factor approaches a familiar function.

$$\lim_{m \to \infty} \left(1 + \frac{r}{m}\right)^{mp} = e^{rp}.$$

The quantity $e^{rp} - 1$ is referred to as the *maximal effective interest rate*, and if interest is calculated this way it is said to be *continuously compounding*. As such it makes more sense to express a continuously compounding value V_t in continuous time $t \in \mathbb{R}$ as

$$V_t = e^{rt} V_0.$$

For $s \in \mathbb{R}_{\geq 0}$, the value V_s is easily expressed in terms of the value V_t.

$$V_s = e^{-r(t-s)} V_t.$$

If $s < t$, this is known as continuous discounting, and it will be employed in chapter 4 with respect to interest rates on money and stock.

An alternative approach is to take a single interest factor $(1+r)^t$; however this function doesn't share the nice properties of e^{rt} with respect to integration and differentiation.

As a sidenote, if a continuously compounding interest rate $r(t)$ is a known function of time, the interest factor from s to t is calculated as

$$e^{\int_s^t r(\tau) d\tau}.$$

Chapter 3

Stochastic calculus

The subject of this chapter is the calculus developed specifically to deal with integrals of stochastic variables. Specifically, the value aggregated by a stochastic process through time, with respect to another stochastic process, may be computed by the aid of the Itô integral presented in section 3.4. The first section concerns the stochastic process central to the model presented by Black and Scholes [BS73] to model asset value through time, the Brownian motion. First conceived by Robert Brown upon observing the apparently indeterministic, random movements of particles through water, Brownian motion is also called the Wiener process after Norbert Wiener, who granted the Brownian motion with a solid mathematical foundation—not included in this paper.

3.1 Brownian motion

In this section Brownian motion is introduced as the limiting case of the symmetric random walk. A symmetric random walk is a discrete-time binomial stochastic process that with equal probability increases or decreases one unit per timestep. Brownian motion is obtained as the length of the longest timestep approaches 0. As such, Brownian motion is a continuous-time binomial stochastic process, which itself serves as the basis of the geometric Brownian motion, which was used by Fischer Black and Myron Scholes in their derivation of a risk-neutral option hedging strategy.

Consider the instantaneous events u and d, which may be thought of as representing an "up" and a "down" move. Let Ω be the event space consisting of infinite sequences of u and d:

$$\omega \in \Omega \Leftrightarrow \omega = \omega_1\omega_2...\omega_j... \text{ such that } \omega_j \in \{\mathrm{u}, \mathrm{d}\}, \text{ with } j \in \mathbb{N}_{\geq 1},$$

and let the probabilities that the j-th element of any $\omega \in \Omega$ be either u or d be equal; that is, $\frac{1}{2}$.

Let then the j-th *step*, either up or down, be the random variable $X_j : \Omega \to \{-1, 1\}$ with

$$X_j(\omega) = \begin{cases} -1 & \text{if } \omega_j = \mathrm{d}, \\ 1 & \text{if } \omega_j = \mathrm{u}, \end{cases}$$

all independent, and define the *symmetric random walk* for $k \in \mathbb{N}_{\geq 0}$ by

$$M_k := \sum_{j=1}^{k} X_j \text{ for } k \in \mathbb{N}_{\geq 1} \text{ and } M_0 := 0$$

the 'position' after k 'steps'. Notice since the steps are random, so is the entire position; hence the name random walk. Let $(\mathcal{F}_k)_{k \in \mathbb{N}_{\geq 0}}$ be a filtration such that $(M_k)_{k \in \mathbb{N}_{\geq 0}}$ is adapted, and such that for all $k < \ell$ it holds that M_ℓ is *not* $\mathcal{F}_k$-measurable—this means that at step k the 'future' path is still random, and it implies that the *increment* $M_{k+1} - M_k$ is independent of $\mathcal{F}_k$. Two nonequal increments are also independent, as follows from the independence of all X_j. It is easily seen that M_k is a martingale; any step is either likely to be up or down, thus the expectation is to remain stationary. Lastly, the process $(M_{k+\ell} - M_k)_{\ell \in \mathbb{N}_{\geq 0}}$ is again a random walk.

Brownian motion is obtained as a limit case of the *scaled symmetric random walk*

$$W^{(n)}(t) := \frac{1}{\sqrt{n}} M_{nt}. \tag{3.1}$$

Notice the introduction of $t \in [0, T]$; if nt is not an integer then M_{nt} is defined by linear interpolation between the floor $\lfloor nt \rfloor$ and ceiling $\lceil nt \rceil$ entiers of nt,

$$M_{nt} = (\lceil nt \rceil - nt) M_{\lfloor nt \rfloor} + (nt - \lfloor nt \rfloor) M_{\lceil nt \rceil},$$

from which it follows that a path of $W^{(n)}(t)$ is continuous. The process is a martingale; take $s \in [0, t]$, then

$$\begin{aligned}
\mathbf{E}\left[W^{(n)}(t) \middle| \mathcal{F}(s)\right] =& \mathbf{E}\left[W^{(n)}(t) - W^{(n)}(s) + W^{(n)}(s) \middle| \mathcal{F}(s)\right] \\
=& \mathbf{E}\left[W^{(n)}(t) - W^{(n)}(s) \middle| \mathcal{F}(s)\right] + \mathbf{E}\left[W^{(n)}(s) \middle| \mathcal{F}(s)\right] \\
=& \mathbf{E}\left[W^{(n)}(t) - W^{(n)}(s)\right] + \mathbf{E}\left[W^{(n)}(s) \middle| \mathcal{F}(s)\right] \\
=& 0 + W^{(n)}(s),
\end{aligned}$$

since $W^{(n)}(t) - W^{(n)}(s)$ is independent of $\mathcal{F}(s)$ and its expectation is zero because the expectations of all X_j are zero.

Donsker's theorem asserts that if $(X_j)_{j \in \mathbb{N}_{\geq 1}}$ is any independent and identically distributed sequence of random variables with $\mathbf{E}[X_j] = 1$ and variance $\mathbf{V}[X_j] = 1$, then as $n \to \infty$ the scaled sum with linear interpolation 3.1 converges in distribution to the *Brownian motion* $W(t)$—short for $(W(t))_{t \in [0,T]}$. Since by definition

$$\mathbf{V}[X_j] = \mathbf{E}\left[X_j^2 - (\mathbf{E}[X_j])^2\right] = \mathbf{E}\left[X_j^2 - 0\right] = \frac{1}{2} \cdot (-1)^2 + \frac{1}{2} \cdot 1^2 = 1,$$

Brownian motion is indeed the (probability) limit case of the scaled symmetric random walk. A thorough exposition of this subject is [Pol84]. Like the scaled symmetric random walk $W(t)$ is a martingale, is continuous, has independent (non-overlapping) increments, and the expectation of any increment $W(t) - W(s)$ for $s \in [0,t]$ is zero. Notably, increments are distributed normally with variance $t - s$:

$$W(t) - W(s) \sim \mathcal{N}(0, t - s)$$

as shown in [Shr10, theorem 3.2.1] by use of the moment-generating function of the normal distribution. As with the random walk, $(W(t+u) - W(t))_{u \in [0,T-t]}$ is again a Brownian motion. The limit $n \to \infty$ suggests, intuitively, that the outcome space Ω of binary sequences is 'squeezed' into a continuum. For a more formal discussion, see [McK69, section 1.1], who defines Brownian motion as the space of all continuous paths $t \mapsto W(t)$ (equivalent to the outcome space) with imposed probabilities such that these paths form a *Gaussian family*—a generalization of random variables with a normal (Gaussian) distribution. Such statements were proved by Paul Lévy [Lév48] and simplified by Zbigniew Ciesielski [Cie61].

For each $t \in [0,T]$, W_t is a random variable $\Omega \to \mathcal{C}_t$, with $W_t(\omega) \in \mathcal{C}_t$ a realization (path) up until t, and $\mathcal{C}_t$ the set of all those paths: $\mathcal{C}_t := \{W_t(\omega) \mid \omega \in \Omega\}$. A typical filtration for a Brownian motion is the family of σ-algebras generated by the pre-images of $(W(t))_{t \in [0,T]}$. Such a filtration is given by

$$\left(\sigma\Big(\{W_t^{-1}(C) \mid C \in \mathcal{C}_t\}\Big)\right)_{t \in [0,T]}.$$

Throughout, $W(t)$ (as a random variable) denotes a Brownian motion adapted to a filtration $(\mathcal{F}(t))_{t \in [0,T]}$ which is such that any increment $W(t) - W(s)$ is independent of $\mathcal{F}(s)$. Together with the probability measure P derived from the equal probability of up and down steps and the event space Ω, $(\Omega, \mathcal{F}(t), P)_{t \in [0,T]}$ is a family of probability spaces rendering $W(t)$ formally $\mathcal{F}(t)$-measurable. In this paper the Brownian motion's filtration is of little relevance aside from its formal introduction. In particular, the different sets of information that the filtration may present are not considered (and are relatively straightforward anyway). In the financial application, the containment of consecutive σ-algebras corresponds to the perception of the unknown future gradually unfolding as time advances. Instead of emphasizing the characterization of Brownian motion as a stochastic *process*, it may be more helpful to consider it as a stochastic *function*. Each $\omega \in \Omega$ (continuous coin-toss space) may be regarded as a 'blueprint' for a full Brownian motion path (realization) $[0,T] \to \mathbb{R}$, which then motivates the notation $W(t)$ for W as a function of $t \in [0,T]$. It should be noted, however, that $W(t)$ remains unknown throughout the following theory even when regarding the final time T—in fact, this is an elementary property of financial analysis.

3.2 Properties of Brownian motion

Take $[0,t] \subsetneq \mathbb{R}$ and let

(3.2) $\Pi = \{t_0 = 0, t_1, ..., t_n = t\} \subsetneq [0,t], \forall j \in \{0, 1..., n-1\} : t_j < t_{j+1}$

be a *partition* of $[0,t]$. Define furthermore the *mesh*

$$\|\Pi\| := \max\{t_{j+1} - t_j \mid j \in \{0,1,...,n_1\}\}.$$

Notice that $\|\Pi\| \to 0$ implies, informally, that $\Pi \to [0,t]$. The limit procedure consists of taking increasingly *finer* partitions—adding more points in $[0,t]$ to the partition, increasing n—such that the mesh approaches zero. With $f : \mathbb{R} \supseteq \mathrm{dom}(f) \to \mathbb{R}$ any function defined on $[0,t]$, the *quadratic variation* of f at t

$$[f,f](t) := \lim_{\|\Pi\|\to 0} \sum_{j=0}^{n-1} \left(f(t_{j+1}) - f(t_j)\right)^2 \tag{3.3}$$

and as such, appears to be related the length of the path $f(\tau)$ as τ increases from 0 to t. However, for such f that have a continuous derivative on $]0,t[$, the quadratic variation turns out to be 0, as follows from the mean value theorem of analysis [Shr10, p.101]. This theorem does not hold for Brownian motion, which is differentiable nowhere, as shown by Wiener. Instead the following proposition holds.

3.2.1 Proposition. *The quadratic variation of Brownian motion W at time t is*

$$[W,W](t) = t$$

almost surely.

Proof. Define with the partition and mesh of 3.2 the partition quadratic variation

$$Q_\Pi := \sum_{j=0}^{n-1} \big(W(t_{j+1}) - W(t_j)\big)^2$$

and notice that $\|\Pi\| \to 0$ implies that $Q_\Pi \to [W,W](t)$. As Q_Π is a random variable, to prove that Q_Π converges to T, it will be shown that the expectation of Q_Π is T, while its variance approaches 0.

In general the following equations with regards to the variance $\mathbf{V}[X]$ of a random variable X hold.

$$\begin{aligned}
\mathbf{V}[X] =& \mathbf{E}\left[(X - \mathbf{E}[X])^2\right] \\
=& \mathbf{E}\left[X^2\right] - 2\mathbf{E}[X]\,\mathbf{E}[X] + (\mathbf{E}[X])^2 = \mathbf{E}\left[X^2\right] - (\mathbf{E}[X])^2 \\
=& \mathbf{E}\left[X^2 - (\mathbf{E}[X])^2\right]
\end{aligned}$$

For the expected value of the partition quadratic variation, observe that

$$\mathbf{E}\left[W(t_{j+1}) - W(t_j)\right] = 0,$$

so by the properties of Brownian motion, indeed

$$\begin{aligned}
\mathbf{E}\left[Q_\Pi\right] =& \sum_{j=0}^{n-1} \mathbf{E}\left[\left(W(t_{j+1}) - W(t_j)\right)^2\right] \\
=& \sum_{j=0}^{n-1} \mathbf{E}\left[\left(W(t_{j+1}) - W(t_j) - \left(\mathbf{E}\left[W(t_{j+1}) - W(t_j)\right]\right)\right)^2\right] \\
=& \sum_{j=0}^{n-1} \mathbf{V}\left[W(t_{j+1}) - W(t_j)\right] \\
=& \sum_{j=0}^{n-1} (t_{j+1} - t_j) = t_0 + t_n = t,
\end{aligned}$$

independent of $\|\Pi\|$.

With regards to the variance, note first that

$$\mathbf{E}\left[(W(t_{j+1}) - W(t_j))^4\right] = 3\left(t_{j+1} - t_j\right)^2,$$

which follows from computing the *kurtosis* of any normal random variable, such as $W(t_{j+1}) - W(t_j)$ (see for instance [Shr10, exercise 3.3]). Using again

$$\mathbf{E}\left[(W(t_{j+1}) - W(t))^2\right] = \mathbf{V}\left[W(t_{j+1}) - W(t)\right],$$

it follows that

$$\begin{aligned}
\mathbf{V}\left[Q_\Pi\right] =& \sum_{j=0}^{n-1} \mathbf{V}\left[(W(t_{j+1}) - W(t))^2\right] \\
=& \sum_{j=0}^{n-1} \mathbf{E}\left[\left((W(t_{j+1}) - W(t))^2 - (\mathbf{E}\left[W(t_{j+1}) - W(t_j)\right])^2\right)^2\right] \\
=& \sum_{j=0}^{n-1} \mathbf{E}\left[(W(t_{j+1}) - W(t_j))^4\right] \\
& - 2\sum_{j=0}^{n-1} (t_{j+1} - t_j)\, \mathbf{E}\left[(W(t_{j+1}) - W(t_j))^2\right] + \sum_{j=0}^{n-1} (t_{j+1} - t_j)^2 \\
=& 2\sum_{j=0}^{n-1} (t_{j+1} - t_j)^2 \leq 2\|\Pi\| t.
\end{aligned}$$

Clearly $\lim_{\|\Pi\|\to 0} \mathbf{V}\left[Q_\Pi\right] = 0$, so $\lim_{\|\Pi\|\to 0} Q_\Pi = [W, W](t) = t$ almost surely. ■

It is worth noting that the Brownian motion quadratic variation $[W, W](t)$ is in fact random itself; however the probability that it has a realization different from t is 0.

As mentioned in the previous section, the Brownian motion increments $W(t) - W(s)$ follow $\mathcal{N}(0, t - s)$. In particular $\mathbf{E}\left[W(t)\right] = \mathbf{E}\left[W(t) - 0\right] = \mathbf{E}\left[W(t) - W(0)\right] = 0$. Remarkably, this quality is implied by any process which adheres to four basic properties, as shown by Lévy.

3.2.2 Theorem (Lévy in one dimension)**.** *A martingale $w(t)$ with continuous paths, starting at zero, with quadratic variation $[w, w](t) = t$ is a Brownian motion.*

Proof. See for example [Shr10, theorem 4.6.4]. ■

Brownian motion is a Markov process—one motivation for its application in financial modeling.

3.2.3 Proposition. *Brownian motion is a Markov process, which means by 2.5 that for any $s, t \in [0, T]$ with $s \leq t$ and any nonnegative Borel-measurable function $f : \mathbb{R} \to \mathbb{R}_{\geq 0}$, there exists a Borel-measurable function $g : \mathbb{R} \to \mathbb{R}$ such that*

$$\mathbf{E}\left[f\left(W(t)\right)|\mathcal{F}(s)\right] = g\left(s, W(s)\right).$$

Proof. Rewrite

$$\mathbf{E}\left[f\left(W(t)\right)|\mathcal{F}(s)\right]$$

as

$$\mathbf{E}\left[f\left(W(t) - W(s) + W(s)\right)|\mathcal{F}(s)\right]$$

and notice $A := W(t) - W(s)$ is distributed $N(0, t-s)$ and independent of $\mathcal{F}(s)$ by the properties of Brownian motion, while $B := W(s)$ is $\mathcal{F}(s)$-measurable, hence, nonrandom. Therefore, integrating with the normal distribution function,

$$\mathbf{E}\left[f\left(A + B\right)|\mathcal{F}(s)\right] = \int_{-\infty}^{\infty} f(a+B)\frac{1}{\sqrt{2\pi(t-s)}}e^{-\frac{a^2}{2(t-s)}}\,da =: g(B)$$

and notice g only depends on the value of W at time s by $B = W(s)$. Since g is continuous, it is Borel-measurable. This completes the proof. ■

To summarize, the Brownian motion $\left(W(t)\right)_{t\in[0,T]}$ satisfies the following properties.

- $W(0) = 0$.
- $\mathbf{E}\left[W(t) - W(s)\right] = 0$.
- A path $W(t)$ is continuous almost surely.
- Brownian motion is a martingale.
- Brownian motion is a Markov process.
- $[W, W](t) = t$.
- $W(t) - W(s) \sim \mathcal{N}(0, t-s)$.
- The paths of a Brownian motion are not differentiable.

Since stock value is strictly positive, to model stock value behavior often a process called *geometric Brownian motion* is employed, introduced in chapter 4.

3.3 Comparison of computations of integrals

Basically, integrals measure signed area, enclosed by a function's graph and its argument's axis. In many cases it is convenient to measure this area with respect to a quantity with more meaning than the function's argument. For example, in the option pricing formula to follow in the next chapter, the area enclosed by the graph of the (positive) amount of stock held at time t and the t-axis itself is a number with little financial relevance, but if this area is measured by the stock's *value* at time t, the integral represents the total value of the investment up until its right bound of integration. In the next section a method will be developed to evaluate such integrals with respect to stochastic processes, like the Brownian motion. This section presents a brief overview of Riemann and Lebesgue integrals, and an introduction to the stochastic integral known as the Itô integral.

Let $f : \mathbb{R} \rightarrowtail \mathbb{R}$ be a function. If f is bounded and (Lebesgue) almost everywhere continuous on the nonsingleton interval $[a, b] \subseteq \mathrm{dom}(f)$, the Riemann integral exists and is computed using directly the relationship between $x \in [a, b]$ and $f(x)$ as prescribed by the function. Adjacent rectangles of fixed width $h \in \mathbb{R}_{>0}$ inscribed between the graph of f and the x-axis, with their left edge at x in this case, measure approximately $hf(x)$, and their sum becomes the exact area as $h \to 0$. More specifically, if and only if a function is Riemann integrable, both the sum of the infimum rectangle areas and the sum of the supremum rectangle areas converge to the same value, which is the integral. For instance, the Riemann integral $\int_0^1 x^p dx$, for $p \in \mathbb{N}_{\geq 1}$, may be computed using Faulhaber's formula (see for example [Knu93])

$$\sum_{k=1}^{n} k^p = \frac{1}{p+1} \sum_{\ell=0}^{p} (-1)^\ell \binom{p+1}{\ell} B_\ell n^{p+1-\ell}, \text{ which also equals } \sum_{k=0}^{n} k^p$$

with Bernoulli numbers B_ℓ and $B_1 = -\frac{1}{2}$. The Bernoulli numbers have several connections to number theory and may be given by many expressions. The Bernoulli number B_1 is either $\frac{1}{2}$ or $-\frac{1}{2}$ (the others are unambiguously defined) and the Bernoulli numbers with $B_1 = \frac{1}{2}$ are called the *first Bernoulli numbers*, while the other case is referred to as the *second Bernoulli numbers*. An explicit definition of the first Bernoulli numbers is

$$B_\ell = \sum_{w=0}^{\ell} \sum_{v=0}^{w} (-1)^v \binom{w}{v} \frac{v^\ell}{w+1}$$

and the first eight (first and second) Bernoulli numbers are $B_0 = 1$, $\pm\frac{1}{2}$, $\frac{1}{6}$, 0, $-\frac{1}{30}$, 0, $\frac{1}{42}$, and $B_7 = 0$. For $\ell > 1$, all Bernoulli numbers with odd ℓ are zero.

The Riemann integral may be computed as

$$\begin{aligned}\int_0^1 x^p dx = \lim_{n\to\infty} \sum_{j=0}^{n-1} \frac{1}{n} \left(\frac{j}{n}\right)^p &= \lim_{n\to\infty} \frac{\frac{1}{p+1} \sum_{\ell=0}^{p} (-1)^\ell \binom{p+1}{\ell} B_\ell (n-1)^{p+1-\ell}}{n^{p+1}} \\ &= \frac{1}{p+1} \lim_{n\to\infty} \sum_{\ell=0}^{p} (-1)^\ell \binom{p+1}{\ell} B_\ell \frac{(n-1)^{p+1-\ell}}{n^{p+1}}.\end{aligned}$$

In the summation, the $\ell = 0$ converges to 1 when $n \to \infty$, while all other terms converge to 0 because the highest power of n is in the denominator. The summation uses the minimum rectangle area $\frac{1}{n}\left(\frac{j}{n}\right)^p$ for the j-th rectangle; it is easily seen that using the maximum area $\frac{1}{n}\left(\frac{j+1}{n}\right)^p$ makes no difference; in this case

$$\begin{aligned}\int_0^1 x^p dx = &\lim_{n\to\infty}\sum_{j=0}^{n-1}\frac{1}{n}\left(\frac{j+1}{n}\right)^p = \lim_{n\to\infty}\sum_{j=1}^{n}\frac{1}{n}\left(\frac{j}{n}\right)^p \\ =&\frac{1}{p+1}\lim_{n\to\infty}\sum_{\ell=0}^{p}(-1)^\ell\binom{p+1}{\ell}B_\ell n^{-\ell}.\end{aligned}$$

Because the Riemann integral depends on the convergence of the infimum and supremum of f within contracting intervals, some functions that are particularly discontinuous cannot be integrated using this method. Continuing the above example, multiplying x^p by the rational indicator $\mathbf{1}_{\mathbb{Q}}$ 2.3 yields a function that takes minimum value 0 and maximum value equal to the upper bound of the contracting interval in $[0,1]$. The Lebesgue integral of f over $[a,b]$ is a limit procedure that partitions $f([a,b])$ into increasingly smaller subsets of function values and measures those by the Lebesgue measure on their pre-image. Thus can be computed the Lebesgue integral

$$\int_0^1 \mathbf{1}_{\mathbb{Q}} x^p d\lambda(x) = \lim_{n\to\infty}\sum_{j=0}^{n-1}\frac{j}{n}\lambda\left(\left\{x\in[0,1]\middle| x\in\mathbb{Q}\wedge x^p\in\left[\frac{j}{n},\frac{j+1}{n}\right[\right\}\right)$$

which is 0 because the set of rational numbers in any real interval has Lebesgue measure 0. The single value $1 \in f([0,1])$ has been discarded; this makes no difference for the integral as the pre-image of 1 is the set $\{1\}$, which also has Lebesgue measure 0. Observe that although the Lebesgue integral also partitions an interval $[0,1]$, this is in the image of f and not in its domain, as was the case with the Riemann integral. Note that Lebesgue integrals on $\mathbb{R}$ are equivalent to Riemann integrals whenever the latter are defined [Sch05, 11.8].

As a final example, the *Riemann-Stieltjes integral* is an intermediary construction, using the Riemann approach to integrate a real function f with respect to another real function g. In this example the integration is over $[0,t]$, but any real interval is satisfactory. Define a partition Π as in 3.2, then the Riemann-Stieltjes integral is given by

$$\int_0^t f(\tau)dg(\tau) := \lim_{\|\Pi\|\to 0}\sum_{j=0}^{n-1} f(t^*)\left(g(t_{j+1}-t_j)\right)$$

with $t^* \in [t_j, t_{j+1}]$. If g is differentiable and $\mathbf{D}[g]$ is continuous, then the Riemann-Stieltjes integral is equal to the Riemann integral

$$\int_0^t f(\tau)\mathbf{D}[g](\tau)d\tau.$$

A sufficient condition for the existence of the Riemann-Stieltjes integral is that f is continuous and the *variation* (see next section) of g is finite.

In the next section it will be argued that the above procedures are insufficient to evaluate integrals with respect to a Brownian motion.

3.4 Itô calculus

Take henceforth $T > 0$ such that $[0, T]$ is not a singleton and let $W(t)$ be a Brownian motion adapted to some filtration $\mathcal{F}(t)$. Recall $(\Omega, \mathcal{F}(t))$ are measure spaces for each $t \in [0, T]$.

With $\Delta(t)$ a stochastic process also adapted to $\mathcal{F}(t)$, the *Itô integral*

$$I(t) := \int_0^t \Delta(\tau) dW(\tau)$$

evaluates the aggregate value of $\Delta(t)$ over $[0, t]$ as measured by $W(t)$. Notice $\Delta(t)$ is $\mathcal{F}(t)$-measurable by the definition of an adapted process, as is $W(t)$. As such, $(I(t))_{t\in[0,T]}$ is a stochastic process adapted to $\mathcal{F}(t)$ itself.

If $W(t)$ has *bounded variation* then $I(t)$ can be evaluated using a Riemann-Stieltjes integral. Consider a partition again a partition 3.2. The aggregated (first-order) variation $\mathbf{FO}$ of $W(t)$ on the interval $[0, t]$ is defined as

$$\mathbf{FO}[W](t) := \lim_{\|\Pi\|\to 0} \sum_{j=0}^{n-1} |W(t_{j+1}) - W(t)|,$$

with again a partition with mesh approaching zero, as in 3.2. However, the following proposition shows that the variation of $W(t)$ is almost surely unbounded.

3.4.1 Proposition. *For any $t \in]0, T]$ it holds that $\mathbf{FO}[W](t) = \infty$ almost surely.*

Proof. Suppose instead that $W(t)$ has bounded variation on $[0, t]$; that is, that $\mathbf{FO}[W](t) < \infty$. For the quadratic variation it follows that

$$\begin{aligned} [W, W](t) &= \lim_{\|\Pi\|\to 0} \sum_{j=0}^{n-1} \left(W(t_{j+1}) - W(t_j)\right)^2 \\ &\leq \lim_{\|\Pi\|\to 0} \max_{j\in\{0,1,\ldots,n-1\}} \{|W(t_{j+1}) - W(t_j)|\} \\ &\quad \cdot \sum_{j=0}^{n-1} |(W(t_{j+1}) - W(t_j)|. \end{aligned}$$

As $W(t)$ is continuous, $\max_{j\in\{0,1,\ldots,n-1\}}\{|W(t_{j+1}) - W(t_j)|\}$ converges to 0, while

$$\lim_{\|\Pi\|\to 0} \sum_{j=0}^{n-1} |(W(t_{j+1}) - W(t_j)| = \mathbf{FO}[W](t)$$

is finite. It then follows that $[W, W](t) \leq 0$, which is almost surely in contradiction with proposition 3.2.1. ■

Furthermore, since $W(t)$ has no derivative,

$$\int_0^t \Delta(\tau) \mathbf{D}[W] d\tau$$

does not exist. For these reasons Kiyoshi Itô developed a different approach to evaluating $I(t)$. The construction shows some similarities to the Lebesgue integral. The full formal approach is beyond the scope of this paper; see for instance [McK69, section 2.2]. Summarizing that discussion, first note a *nonanticipating* process $(\Delta(t))_{t\in[0,T]}$ satisfies: $\forall s \in [0,t] : \Delta(t) - \Delta(s)$ is independent of $\mathcal{F}(s)$. For the definition of the integral, it is required that $\Delta(t)$ is nonanticipating, $\mathcal{F}(t)$-measurable, and that $\Delta(t)$ is square-integrable, meaning

$$\mathbf{E}\left[\int_0^t (\Delta(\tau))^2\, d\tau\right] = \int_{\omega\in\Omega}\left(\int_0^t (\Delta_\tau(\omega))^2 d\tau\right)dP(\omega) < \infty. \tag{3.4}$$

Notice this expectation is taken with respect to the probability space $(\Omega, \mathcal{F}(t), P)$.

It can then be shown that there is a sequence $(\Delta_n(t))_{n\in\mathbb{N}_{\geq 1}}$ of $\mathcal{F}(t)$-measurable nonanticipating *simple functionals*, such that these $\Delta_n(t)$ are constant on half-open subintervals

$$\left[\frac{jt}{n}, \frac{(j+1)t}{n}\right[\subsetneq [0,t] \text{ with } j \in \{0,1,...,n-1\},$$

and such that this sequence $(\Delta_n(t))_{n\in\mathbb{N}_{\geq 1}}$ is defined as

$$\Delta_n(\tau) = \Delta\left(\lfloor\frac{n\tau}{t}\rfloor\right) \text{ with } \lfloor\cdot\rfloor \text{ the floor entier,} \tag{3.5}$$

while satisfying

$$\lim_{n\to\infty}\mathbf{E}\left[\int_0^t |\Delta(\tau) - \Delta_n(\tau)|^2 d\tau\right] = 0,$$

which is L^2-convergence. Namely, from square integrability 3.4 it can be shown that there exists a sequence of processes $(\Delta'_n(t))_{n\in\mathbb{N}_{\geq 1}}$,

$$\Delta'_n(t) := \min\left\{n, \max\{-n, \Delta(t)\}\right\},$$

that are nonanticipating, $\mathcal{F}(t)$-measurable, and *bounded*, and L^2-converge to $\Delta_n(t)$:

$$\lim_{n\to\infty}\mathbf{E}\left[\int_0^t |\Delta_n(\tau) - \Delta'_n(\tau)d\tau|^2\right] = 0,$$

by dominated convergence, These $\Delta'_n(t)$ may be L^2 approached by *continuous* nonanticipating, $\mathcal{F}(t)$-measurable, bounded $\Delta''_n(t)$ by bounded convergence, which are in turn L^2-approached by $\Delta_n(t)$ in 3.5, again by bounded convergence.

Hence let $\Delta_n(t)$ be such a sequence; it can be shown that

$$\int_0^t \Delta_n(\tau)dW(\tau) := \sum_{j=0}^{n-1}\Delta_n\left(\frac{j}{n}t\right)\left(W\left(\frac{j+1}{n}t\right) - W\left(\frac{j}{n}t\right)\right)$$

has a limit in L^2, as $n\to\infty$, that is independent of the choice of Δ_n. Therefore, the square integrability condition 3.4 allows defining the unique Itô integral as

$$\begin{aligned} I(t) = \int_0^t \Delta(\tau)dW(\tau) &:= \lim_{n\to\infty}\int_0^t \Delta_n(\tau)dW(\tau) \\ &= \lim_{n\to\infty}\sum_{j=0}^{n-1}\Delta_n\left(\frac{j}{n}t\right)\left(W\left(\frac{j+1}{n}t\right) - W\left(\frac{j}{n}t\right)\right). \end{aligned} \tag{3.6}$$

The square-integrability requirement 3.4 is assumed throughout this paper and appropriate for all practical applications.

The following proposition identifies a property of Itô integrals that is fundamental to the risk-neutral framework in mathematical finance.

3.4.2 Proposition. *Itô integrals are martingales.*

Proof. This is an illustration; the proof boils down to splitting the expectation

$$\begin{aligned}\mathbf{E}\left[I(t)|\mathcal{F}(s)\right] =& \mathbf{E}\left[\int_0^t \Delta(\tau)dW(\tau)\middle|\mathcal{F}(s)\right]\\ =& \mathbf{E}\left[\lim_{n\to\infty}\sum_{j=0}^{n-1}\Delta_n\left(\frac{j}{n}t\right)\left(W\left(\frac{j+1}{n}t\right)-W\left(\frac{j}{n}t\right)\right)\middle|\mathcal{F}(s)\right]\end{aligned}$$

into terms that are $\mathcal{F}(s)$-measurable and terms that are not. The part that is $\mathcal{F}(s)$-measurable (with $\tau \leq s$) is nonrandom and has value $I(s)$. The part that is not $\mathcal{F}(s)$-measurable is independent of $\mathcal{F}(s)$ and breaks down in terms

$$\begin{aligned}&\mathbf{E}\left[\Delta_n\left(\frac{j}{n}t\right)\left(W\left(\frac{j+1}{n}t\right)-W\left(\frac{j}{n}t\right)\right)\middle|\mathcal{F}(s)\right] \text{ where } \frac{j}{n}t > s\\ &= \mathbf{E}\left[\Delta_n\left(\frac{j}{n}t\right)\right]\mathbf{E}\left[\left(W\left(\frac{j+1}{n}t\right)-W\left(\frac{j}{n}t\right)\right)\right],\end{aligned}$$

which are zero by the properties of Brownian motion. Thus $\mathbf{E}\left[I(t)|\mathcal{F}(s)\right] = I(s)$, which says that Itô processes are martingales. ■

Being based on Brownian motion, Itô integrals aggregate nonzero quadratic variation themselves. As the next proposition shows, $\Delta(t)$ amplifies the Brownian motion's quadratic variation $[W,W](t) = t$ by its square in $I(t)$.

3.4.3 Proposition. *The quadratic variation of the Itô integral $I(t)$ of 3.6 is*

$$[I,I](t) = \int_0^t (\Delta(\tau))^2 \, d\tau$$

almost surely.

Proof. This proof is restricted to nonanticipating simple processes $\Delta_n(t)$ as in 3.5, but the result holds for all Itô-integrable stochastic processes.

Because Δ_n is constant on $\left[\frac{j}{n}t, \frac{j+1}{n}t\right[$, the integral $I(t)$ is a sum [Shr10, section 4.2.1]

$$\int_0^t \Delta_n(\tau)dW(\tau) = \sum_{j=0}^{n-1}\Delta_n\left(\frac{j}{n}t\right)\left(W\left(\frac{j+1}{n}t\right)-W\left(\frac{j}{n}t\right)\right).$$

Take a partition $\Pi := \{s_0 = \frac{j}{n}t, s_1, ..., s_m = \frac{j+1}{n}t\} \subsetneq \left[\frac{j}{n}t, \frac{j+1}{n}t\right]$ as in 3.2, then the quadratic variation aggregated on $\left[\frac{j}{n}t, \frac{j+1}{n}t\right[$, which is equal that on $\left[\frac{j}{n}t, \frac{j+1}{n}t\right]$ (namely $\frac{j+1}{n}t - \frac{j}{n}t = \frac{1}{n}t$), is

$$\begin{aligned}
&\lim_{\|\Pi\|\to 0} \sum_{i=0}^{m-1} \left(\Delta_n\left(\frac{j}{n}t\right)\left(W(s_{i+1}) - W(s_i)\right)\right)^2 \\
&= \left(\Delta_n\left(\frac{j}{n}t\right)\right)^2 \lim_{\|\Pi\|\to 0} \sum_{i=0}^{m-1} \left(W(s_{i+1}) - W(s_i)\right)^2 \\
&= \left(\Delta_n\left(\frac{j}{n}t\right)\right)^2 \left([W,W]\left(\frac{j+1}{n}t\right) - [W,W]\left(\frac{j}{n}t\right)\right) \\
&= \left(\Delta_n\left(\frac{j}{n}t\right)\right)^2 \left(\frac{j+1}{n}t - \frac{j}{n}t\right) \text{ almost surely} \\
&= \int_{\frac{j}{n}}^{\frac{j+1}{n}} \left(\Delta_n(\tau)\right)^2 d\tau \text{ almost surely.}
\end{aligned}$$

As quadratic variation is additive with respect to subintervals of $[0,T]$, it follows that the quadratic variation of $I(t)$ on $[0,t]$ is the sum

$$[I,I](t) = \sum_{j=0}^{n-1} \int_{\frac{j}{n}}^{\frac{j+1}{n}} \left(\Delta_n(\tau)\right)^2 d\tau = \int_0^t \left(\Delta_n(\tau)\right)^2 d\tau \text{ almost surely}$$

which proves the proposition. ■

In order to perform calculus with respect to the non-differentiable Brownian motion, Itô (and independently of him, Wolfgang Doeblin) devised an integral equation for a certain class of functions of $W(t)$.

3.4.4 Theorem (Itô-Doeblin formula for Brownian motion). *Let $f : \mathbb{R}_{\geq 0} \times \mathbb{R} \to \mathbb{R}$ be a function such that the partial derivatives*

- $\mathbf{D}_1[f]$,
- $\mathbf{D}_2[f]$, *and*
- $\mathbf{D}_2^2[f]$

exist and are continuous. With $t \in \mathbb{R}_{\geq 0}$ it holds that

$$\begin{aligned}
f(t, W(t)) = &f(0, W(0)) + \int_0^t \mathbf{D}_1[f](\tau, W(\tau))\, d\tau \\
&+ \int_0^t \mathbf{D}_2[f](\tau, W(\tau))\, dW(\tau) + \frac{1}{2}\int_0^t \mathbf{D}_2^2[f](\tau, W(\tau))\, d\tau.
\end{aligned}$$

Proof. To prove the equality, consider again a partition

$$\Pi = \{t_0 = 0, t_1, ..., t_n = t\} \subsetneq [0,t], \forall j \in \{0, 1..., n-1\} : t_j < t_{j+1}.$$

It will be shown that the sum of the discrete increments of f, expanded using Taylor's theorem for two variables, converges to the proposed expression.

Consider the Taylor expansion

$$\begin{aligned}
f\left(t_{j+1}, W(t_{j+1})\right) =& f\left(t_j, W(t_j)\right) \\
&+ \mathbf{D}_1[f]\left(t_j, W(t_j)\right) \cdot (t_{j+1} - t_j) \\
&+ \mathbf{D}_2[f]\left(t_j, W(t_j)\right) \cdot (W(t_{j+1}) - W(t_j)) \\
&+ \frac{1}{2}\mathbf{D}_2^2[f]\left(t_j, W(t_j)\right) \cdot (W(t_{j+1}) - W(t_j))^2 \\
&+ \mathbf{D}_{1,2}[f]\left(t_j, W(t_j)\right) \cdot (W(t_{j+1}) - W(t_j)) \cdot (t_{j+1} - t_j) \\
&+ \frac{1}{2}\mathbf{D}_1^2[f]\left(t_j, W(t_j)\right) \cdot (t_{j+1} - t_j)^2 \\
&+ \text{higher order terms.}
\end{aligned}$$

By summing the equality for all $j \in \{0, 1..., n-1\}$ and canceling on both sides the terms $f\left(t_j, W(t_j)\right)$ for $j \in \{1, 2, ..., n-2\}$, the lefthandside equals $f\left(t, W(t)\right)$, while the first term on the righthandside produces $f\left(0, W(0)\right)$. The other terms on the righthandside yield summations which will be shown to equal certain integrals as $\|\Pi\| \to 0$.

For the sum of the first appearing partial derivative, it holds in the limit that

$$\lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} \mathbf{D}_1[f]\left(t_j, W(t_j)\right) \cdot (t_{j+1} - t_j) = \int_0^t \mathbf{D}_1[f]\left(\tau, W(\tau)\right) d\lambda(\tau),$$

a Lebesgue integral by the measure $t_{j+1} - t_j$. The second appearing partial is measured by $W(t_{j+1}) - W(t_j)$ and as such yields an Itô integral by

$$\begin{aligned}
&\lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} \mathbf{D}_2[f]\left(t_j, W(t_j)\right) \cdot (W(t_{j+1}) - W(t_j)) \\
&= \int_0^t \mathbf{D}_2[f]\left(\tau, W(\tau)\right) dW(\tau).
\end{aligned}$$

Notice that the third partial $\frac{1}{2}\mathbf{D}_2^2[f]\left(t_j, W(t_j)\right) \cdot (W(t_{j+1}) - W(t_j))^2$ resembles quadratic variation summands because of the $(W(t_{j+1}) - W(t_j))^2$ term. In fact,

$$\begin{aligned}
&\lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} \frac{1}{2}\mathbf{D}_2^2[f]\left(t_j, W(t_j)\right) \cdot (W(t_{j+1}) - W(t_j))^2 \\
&= \int_0^t \mathbf{D}_2^2[f[\left(\tau, W(\tau)\right) d[W, W](\tau) = \frac{1}{2}\int_0^t \mathbf{D}_2^2[f]\left(\tau, W(\tau)\right) d\tau
\end{aligned}$$

because as $\|\Pi\| \to \infty$, the infinitesimal change in quadratic variation $d[W, W](\tau)$ is the infinitesimal change in its value τ, shown in 3.2.1. For a more elaborate argument, see for example [Shr10, remark 3.4.4].

The fourth and fifth partial, as well as all partials of higher order, contain at least one difference factor $(t_{j+1} - t_j)$ or $(W(t_{j+1}) - W(t_j))$ twice. The limits of their sums is therefore bounded by the maximum of that difference factor times some finite integral without that difference factor. For instance,

$$\begin{aligned}&\sum_{j=0}^{n-1}\frac{1}{2}\mathbf{D}_1^2[f]\,(t_j,W(t_j))\cdot(t_{j+1}-t_j)^2\\&\leq\max_{j\in\{0,1,\ldots,n-1\}}\{t_{j+1}-t_j\}\sum_{j=0}^{n-1}\frac{1}{2}\mathbf{D}_1^2[f]\,(t_j,W(t_j))\cdot(t_{j+1}-t_j)\end{aligned}$$

and this expression converges as $\|\Pi\|\to 0$ to

$$0\cdot\int_0^t\mathbf{D}_1^2[f]\,(\tau,W(\tau))\,d\tau=0.$$

Since W is continuous, similarly $\max_{j\in\{0,1,\ldots,n-1\}}\{|\,(W(t_{j+1})-W(t_j))\,|\}\to 0$, so likewise all other remaining terms in the Taylor expansion are 0 in the limit.

Summing $f\,(0,W(0))$ and the obtained integrals yields the proposed equation. ■

In general an *Itô process* $X(t)$ is understood to be a stochastic process

$$X(t)=X(0)+\int_0^t\Psi(\tau)dW(\tau)+\int_0^t\Theta(\tau)d\tau \tag{3.7}$$

such that $X(0)$ is nonrandom and $\Psi(t)$ and $\Theta(t)$ are adapted to $\mathcal{F}(t)$. Notice it consists of an Itô and a Lebesgue integral. It can be shown by continuity arguments [Shr10, lemma 4.4.4] that an Itô process attains all its quadratic variation from its Itô integral; therefore from proposition 3.4.3 it follows that

$$[X,X]\,(t)=\left[\int_0^t\Psi(\tau)dW(\tau),\int_0^t\Psi(\tau)dW(\tau)\right](t)=\int_0^t\left(\Psi(\tau)\right)^2d\tau. \tag{3.8}$$

Generalizing as well to integration along the Itô process $X(t)$, with $\Gamma(t)$ adapted to $\mathcal{F}(t)$, the *integral with respect to* $X(t)$ is defined as

$$\int_0^t\Gamma(\tau)dX(\tau):=\int_0^t\Gamma(\tau)\Psi(\tau)dW(\tau)+\int_0^\tau\Gamma(\tau)\Theta(\tau)d\tau,$$

the sum of an Itô and a Lebesgue integral. This leads to a generalization of theorem 3.4.4 as follows.

3.4.5 Theorem (Itô-Doeblin formula for Itô process)**.** *Let $\left(X(t)\right)_{t\in[0,T]}$ be an Itô-process, and let $f:\mathbb{R}_{\geq 0}\times\mathbb{R}\to\mathbb{R}$ be a function such that the partial derivatives*

- $\mathbf{D}_1\,[f]$,
- $\mathbf{D}_2\,[f]$, *and*
- $\mathbf{D}_2^2\,[f]$

exist and are continuous. With $t \in \mathbb{R}_{\geq 0}$ *it holds that*

$$\begin{aligned} f\left(t, X(t)\right) =& f\left(0, W(0)\right) + \int_0^t \mathbf{D}_1\left[f\right]\left(\tau, X(\tau)\right) d\tau \\ &+ \int_0^t \mathbf{D}_2\left[f\right]\left(\tau, X(\tau)\right) \Psi(\tau) dW(\tau) \\ &+ \int_0^t \mathbf{D}_2[f]\left(\tau, X(\tau)\right) \Theta(\tau) dt \\ &+ \frac{1}{2}\int_0^t \mathbf{D}_2^2\left[f\right]\left(\tau, X(\tau)\right) \left(\Psi(\tau)\right)^2 d\tau. \end{aligned}$$

Proof. The procedure is similar to that of theorem 3.4.4. ■

Chapter 4

Option pricing

In practical finance, the *market* consists of agents trading in *assets*: something possessing economic value. The valuation of these assets through time is determined largely by classic offer-and-demand, and proves to be unpredictable in such a way that agents cannot trust that their assets retain some value. For example, if one agent's asset is money and she regularly requires some product to facilitate her economic production, she cannot trust this product to maintain the same price relative to her money through time. The rise in value of such a product is a significant financial risk for her. In order to *hedge* such risks, there exists demand for financial products that counteract in some way a negative turn of events. One such financial product is an *option*. An option grants its purchaser the right to buy an asset at some time in the future, for some fixed price called the *strike price*. Intuitively, should the value of the corresponding asset rise through time, since the strike price of the option remains constant, the value of the option increases. If the value of the asset falls, so does the value of the option, until the asset drops below the strike price: then the option is worth nothing. The agent who requires the regular purchase of some products may also purchase an option on this product: the value of the option correlates negatively with the value of her money relative to the product, and she is said to have hedged her position.

Options exist on the financial market and are typically issued by financial institutions. The two problems that these institutions face are

1. at what price to sell the option, and
2. how to invest the money from selling the option in order to deliver the asset if the purchaser claims her right.

To ensure option-selling is reasonable business, the value of the option must be such that the issuer makes no loss if she must deliver the asset. It can furthermore be shown [Shr04] that the issuer must not be able to make a sure profit in this business. The latter condition is called a *no-arbitrage requirement*; if it is violated, agents are able to make sure profits without initial capital, which would incapacitate the economic system. These conditions determine a unique but random value that the issuer must possess from investing the unknown initial option price. The randomness arises from the uncertain path that the value of the asset will follow through time.

In this chapter, a unique price and investment strategy for the issuer are derived, echoing the methods of Robert C. Merton [Mer73] and in particular the derivation of Fischer Black and Myron Scholes [BS73], part of their joint curriculum that earned Merton and Scholes the 1997 Nobel Prize in Economics (after Black had passed away in 1995) [Nob97]. The analysis assumes and considers

- one asset, with constant rate of return $\alpha \in \mathbb{R}_{>0}$, continuously divisible, meaning the asset can be divided and traded in real positive ($\mathbb{R}_{>0}$) quantities,
- one other asset, the numeraire (money); continuously divisible, measuring all value, with constant continuously compounding interest rate $r \in \mathbb{R}_{>0}$,
- the assumption $\alpha > r$ as motivation to invest in the asset,
- one option to buy one asset at fixed *maturity* time $T \in \mathbb{R}_{>0}$ at fixed strike price $K \in \mathbb{R}_{>0}$,
- $t \in [0, T]$ time,
- $S(t)$, the asset value through time, is a geometric Brownian motion; see 4.2,
- $V(t, S(t))$, the option depending on time and on the asset value, such that $V(T, S(T)) = (S(T) - K)^{+} = \max\{S(T) - K, 0\}$,
- $\Delta(t)$, the *position* (investment) of the issuer in the asset through time
- $X(t)$, the portfolio value of the issuer through time, consists of investment in the asset and in the numeraire,
- $W(t)$, a Brownian motion, and
- $S(t)$, $V(t)$, $\Delta(t)$, $X(t)$, and $W(t)$ adapted to filtration $\mathcal{F}(t)$, the market information and history at time t.

4.1 Derivation of Black-Scholes-Merton equation

This section relies heavily on the Itô calculus developed in section 3.4. It is common practice to use differential notation for the expressions derived there, as they are much shorter and are argued to provide ‘intuition’ about how the stochastic processes change. In this paper, instead it was chosen to rely on full integral notation, as being mathematically formal, and providing an intuition of its own: the integrals from 0 to t represent aggregated value of the integrand as measured by their integrator.

A unique price and investment strategy satisfying all requirements will be derived from the assumption that through time, the asset value follows a geometric Brownian motion, given in integral notation by

$$S(t) = S(0) + \sigma \int_0^t S(\tau) dW(\tau) + \alpha \int_0^t S(\tau) d\tau. \tag{4.1}$$

This assumption expresses that the asset value starts at some initial level $S(0)$, starts continuously compounding value on itself at rate r, and either rises or falls proportionally to a Brownian motion $W(t)$ adapted to $\mathcal{F}(t)$, amplified by constant *volatility* $\sigma \in \mathbb{R}_{>0}$.

Expression 4.1 is solved by the *geometric Brownian motion*

$$S(t) = S(0) e^{\sigma W(t) + \left(\alpha - \frac{1}{2}\sigma^2\right)t}. \tag{4.2}$$

To see how, consider $S : \mathbb{R}_{\geq 0} \times \mathbb{R} \to \mathbb{R}$ given by

$$S(x, y) = S(0, 0) e^{\sigma y + \left(\alpha - \frac{1}{2}\sigma^2\right)x}$$

and notice the partial derivatives $\mathbf{D}_1[S]$, $\mathbf{D}_2[S]$, and $\mathbf{D}_2^2[S]$ exist and are continuous. Therefore it follows from applying the Itô-Doeblin formula from theorem 3.4.4 that

$$\begin{aligned} S(t, W(t)) =& S(0, W(0)) + \int_0^t \mathbf{D}_1[S](\tau, W(\tau))\, d\tau \\ &+ \int_0^t \mathbf{D}_2[S](\tau, W(\tau))\, dW(\tau) + \frac{1}{2}\int_0^t \mathbf{D}_2^2[S](\tau, W(\tau))\, d\tau \\ =& S(0, W(0)) + \left(\alpha - \frac{1}{2}\sigma^2\right) \int_0^t S(\tau, W(\tau))\, d\tau \\ &+ \sigma \int_0^t S(\tau, W(\tau))\, dW(\tau) + \frac{1}{2}\sigma^2 \int_0^t S(\tau, W(\tau))\, d\tau, \end{aligned}$$

thus by substituting t for x and $W(t)$ for y in $S(x, y)$ and the cancellation of the terms multiplied by $\frac{1}{2}\sigma^2$, the integral expression 4.1 is obtained, showing geometric Brownian motion 4.2 indeed solves 4.2.

Note $S(t)$ inherits the Markov property from $W(t)$ 3.2.3. Notice furthermore that $S(x, y)$ is convex in y, so by Jensen's inequality it follows that

$$\mathbf{E}[S(t, W(t))] \geq S(t, \mathbf{E}[W(t)]) = S(0, 0) e^{(\alpha - \frac{1}{2}\sigma^2)t}$$

and the actual expectation is in fact $\mathbf{E}[S(t)] = S(0)e^{\alpha t}$. This illustrates that geometric Brownian motion is not a martingale. Its distribution is called *log-normal* because the natural logarithm of $S(t)$ follows a normal distribution. The expected value may be interpreted as the expected financial gain over time from investing in one share of the asset. Also, $S(t)$ is strictly positive. These properties bring geometric Brownian motion in accordance with several characteristics of real-world asset value behavior. For example, Fama found strong evidence that the (past) path of an asset's value cannot be used to predict the future value of the asset [Fam65]. It is however in contrast with reality that α and σ are held fixed, and, additionally, real stock behavior shows higher probability of sudden large value changes (commonly called *fat tails*) than under log-normal distribution.

The issuer's position in the asset $\Delta(t)$ is a stochastic process adapted to $\mathcal{F}(t)$. In the upcoming discussion, an expression for $\Delta(t)$ will be derived that is free of the 'risk' of the volatile Brownian motion $W(t)$ and prescribes the option's issuers amount of stock in the asset to hold at time t, depending only on the asset value $S(t)$, to perfectly hedge her obligation to the option purchaser at any time. The issuer distributes the portfolio value (her wealth) $X(t)$ at any time t over her investment $\Delta(t)$ and her position in the numeraire $X(t)-\Delta(t)$. In particular, she does not add value to her portfolio; she works with her initial capital $X(0)$. Her portfolio process is therefore expressed as

$$X(t) = X(0)+\int_0^t \sigma\Delta(\tau)S(\tau)dW(\tau)+\int_0^t \bigg((\alpha-r)\Delta(\tau)S(\tau)+rX(\tau)\bigg)d\tau,$$

an Itô process. The Itô integral is subject to the instantaneous changes of the Brownian motion amplified by its volatility. The regular integral contains the rates of return on both her positions.

Since the value of the portfolio is expressed in units of the numeraire, it is important to account for the effect of discounting, a consequence of the numeraire market rate of return r. More concretely, the value of the numeraire is time dependent. Specifically, assuming continuous discounting, the value

$v \in \mathbb{R}_{>0}$ at time $t_j \in [0,T]$

is

$e^{-rt_j} \cdot v$ at time $t=0$;

see section 2.3. Define therefore the generic discounting function

(4.3) $\varphi(x,y) := e^{-rx}y$

with continuous partial derivatives

$$\mathbf{D}_1[\varphi](x,y) = -re^{-rx}y \qquad \mathbf{D}_2[\varphi](x,y) = e^{-rx} \qquad \mathbf{D}_2^2[\varphi](x,y) = 0,$$

notice $\varphi(0,X(0)) = X(0)$, take

$\Psi(t) = \sigma\Delta(t)S(t)$ and $\Theta(t) = (\alpha-r)\Delta(t)S(t) + rX(t)$

and consider the Itô-Doeblin expansion of theorem 3.4.5 of the *discounted* portfolio

$$\begin{aligned} e^{-rt}X(t) =& \varphi(t,X(t)) \\ =& \varphi(0,X(0)) - \int_0^t re^{-r\tau}X(\tau)d\tau + \int_0^t e^{-r\tau}\sigma\Delta(\tau)S(\tau)dW(\tau) \\ &+ \int_0^t e^{-r\tau}\left((\alpha-r)\Delta(\tau)S(\tau)+rX(\tau)\right)d\tau \\ =& X(0) + \int_0^t e^{-r\tau}\sigma\Delta(\tau)S(\tau)dW(\tau) \\ &+ \int_0^t e^{-r\tau}\left((\alpha-r)\Delta(\tau)S(\tau)\right)d\tau. \end{aligned}$$

Notice that the discounted portfolio value $e^{-rt}X(t)$ does not grow as a result the numeraire market's interest rate r.

The option value $V(t, S(t))$ depends on t and the asset value $S(t)$, and is itself an Itô process. Taking

$$\Psi(t) = \sigma S(t) \qquad \text{and} \qquad \Theta(t) = \alpha S(t),$$

then from the the general Itô-Doeblin formula it follows that

$$\begin{aligned}
V(t, S(t)) =& V(0, S(0)) + \int_0^t \mathbf{D}_1[V](\tau, S(\tau))\, d\tau \\
&+ \int_0^t \mathbf{D}_2[V](\tau, S(\tau))\, \sigma S(\tau) dW(\tau) \\
&+ \int_0^t \mathbf{D}_2[V](\tau, S(\tau))\, \alpha S(\tau) d\tau \\
&+ \frac{1}{2} \int_0^t \mathbf{D}_2^2[V](\tau, S(\tau))\, \sigma^2 (S(\tau))^2\, d\tau \\
=& V(0, S(0)) + \int_0^t \mathbf{D}_2[V](\tau, S(\tau))\, \sigma S(\tau) dW(\tau) \\
&+ \int_0^t \mathbf{D}_1[V](\tau, S(\tau)) + \mathbf{D}_2[V](\tau, S(\tau))\, \alpha S(\tau) \\
&+ \frac{1}{2} \mathbf{D}_2^2[V](\tau, S(\tau))\, \sigma^2 (S(\tau))^2\, d\tau.
\end{aligned}$$

From this expression the Itô-Doeblin expansion of the discounted option value $e^{-rt}V(t, S(t)) = \varphi(t, V(t, S(t)))$ from 4.3 may be computed. Taking

$$\begin{aligned}
\Psi(t) =& \mathbf{D}_2[V](\tau, S(\tau))\, \sigma S(\tau) \text{ and} \\
\Theta(t) =& \mathbf{D}_1[V](\tau, S(\tau)) + \mathbf{D}_2[V](\tau, S(\tau))\, \alpha S(\tau) \\
&+ \frac{1}{2} \mathbf{D}_2^2[V](\tau, S(\tau))\, \sigma^2 (S(\tau))^2,
\end{aligned}$$

the general Itô-Doeblin formula from 3.4.5 yields

$$\begin{aligned}
e^{-rt}V(t, S(t)) =& \varphi(t, V(t, S(t))) \\
=& \varphi(0, V(0, S(0))) - \int_0^t r e^{-r\tau} V(\tau, S(\tau))\, d\tau \\
&+ \int_0^t e^{-r\tau} \Psi(\tau) dW(\tau) + \int_0^t e^{-r\tau} \Theta(\tau) d\tau + 0 \\
=& V(0, S(0)) + \int_0^t e^{-r\tau} \mathbf{D}_2[V](\tau, S(\tau))\, \sigma S(\tau) dW(\tau) \\
&+ \int_0^t e^{-r\tau} \Big(- rV(\tau, S(\tau)) + \mathbf{D}_1[V](\tau, S(\tau)) \\
&+ \mathbf{D}_2[V](\tau, S(\tau))\, \alpha S(\tau) + \frac{1}{2} \mathbf{D}_2^2[V](\tau, S(\tau))\, \sigma^2 (S(\tau))^2 \Big) d\tau.
\end{aligned}$$

Thus are obtained expressions for both the discounted portfolio value and the discounted asset value in terms of initial conditions and regular and Itô integrals. Since the portfolio hedges the short (outstanding) option, it is required that

(4.4) $\forall t \in [0,T] : e^{-rt}X(t) = e^{-rt}V\left(t, S(t)\right).$

In particular, the initial values must agree since $V\left(0, S(0)\right)$ provides the issuer with the initial capital to set up the hedge $X\left(0\right)$, while the values at maturity must agree so the issuer can fulfill his obligation (if applicable) and to prevent arbitrage. Evaluating 4.4 with the expanded expressions obtained above, using initial condition $X\left(0\right) = V\left(0, S(0)\right)$ to cancel those terms, an equation arises that consists of a regular and an Itô integral on both sides. In the expansion of an Itô process, only the Itô integrals almost surely have nonzero quadratic variation, as mentioned in section 3.4. Therefore the Itô integrals are equal, yielding

$$\int_0^t e^{-r\tau}\sigma\Delta(\tau)S(\tau)dW(\tau) = \int_0^t e^{-r\tau}\mathbf{D}_2[V]\left(\tau, S(\tau)\right)\sigma S(\tau)dW(\tau)$$

almost surely. As 4.4 holds for all t, this implies that

(4.5) $\Delta(t) = \mathbf{D}_2[V]\left(t, S(t)\right)$ almost surely.

As such, at time t the derivative of V with respect to the asset value provides the position $\Delta(t)$ in the asset that the issuer must take to hedge the short option. In finance, equation 4.5 is referred to as the *delta-hedging rule*.

Equation of the Itô integrals implies the regular integrals are equal, too. Again by 4.4 this implies the integrands are equal, thus

$$\begin{aligned} &e^{-rt}(\alpha - r)\Delta(t)S(t) = e^{-rt}\Big(- rV\left(t, S(t)\right) \\ &+\mathbf{D}_1[V]\left(t, S(t)\right) + \mathbf{D}_2[V]\left(t, S(t)\right)\alpha S(t) + \frac{1}{2}\mathbf{D}_2^2[V]\left(t, S(t)\right)\sigma^2\left(S(t)\right)^2\Big) \end{aligned}$$

almost surely. Dividing by e^{-rt}, substituting $\Delta(t)$ with $\mathbf{D}_2[V]\left(t, S(t)\right)$ by 4.5, and canceling terms $\mathbf{D}_2[V]\left(t, S(t)\right)\alpha S(t)$ results in

$$\begin{aligned} &-r\mathbf{D}_2[V]\left(t, S(t)\right)S(t) = -rV\left(t, S(t)\right) \\ &+\mathbf{D}_1[V]\left(t, S(t)\right) + \frac{1}{2}\mathbf{D}_2^2[V]\left(t, S(t)\right)\sigma^2\left(S(t)\right)^2 \text{ almost surely.} \end{aligned}$$

Replacing the process $S(t)$ with $y \in \mathbb{R}_{>0}$ representing arbitrary asset value yields the partial differential equation

(4.6) $rV(t,y) = \mathbf{D}_1[V]\left(t,y\right) + ry\mathbf{D}_2[V]\left(t,y\right) + \frac{1}{2}\sigma^2 y^2\mathbf{D}_2^2[V]\left(t,y\right)$ almost surely,

typically referred to as the Black-Scholes-Merton equation. A boundary condition is the requirement that the option value, hence the portfolio value, satisfies

$$V(T,y) = (y-K)^+$$

for arbitrary asset value y. The two other boundary conditions are $V(t,0) = 0$ and $y \to \infty \Rightarrow V(t,y) \to y$, saying the option value converges to the asset value as the latter approaches infinity.

Black and Scholes recognized [BS73] the partial differential equation 4.6 as the heat-transfer equation in physics, and refer to the textbook solution by Ruel V. Churchill [Chu63, p.155]. The equation is a Cauchy-Euler equation and may be solved by a change of variables into a diffusion equation. In this paper, instead an alternative approach typical to mathematical finance is chosen, using a transformation to a risk-neutral probability measure that renders the option and portfolio value a martingale.

Notice that no terms related to the Brownian motion $W(t)$ are present in 4.6. This may be interpreted as having successfully eliminated risk from the hedging strategy $\Delta(t) = \mathbf{D}_2[V]\,(t, S(t))$ in 4.5.

4.2 Solution to the option value formula

The following proposition is a simplification of Girsanov's theorem in one dimension [Gir60]. The full theorem considers volatility, rate of return on asset, and rate of return on the numeraire as stochastic processes adapted to $\mathcal{F}(t)$. This paper considers constant σ, α, and r exclusively, so the proposition has been adjusted accordingly. The proof of the full theorem is not substantially different.

Take

$$Z(t) := e^{-\frac{\alpha-r}{\sigma}W(t)-\frac{1}{2}\left(\frac{\alpha-r}{\sigma}\right)^2 t} \qquad \text{and} \qquad Z := Z(T)$$

and define

$$\tilde{W}(t) := W(t) + \frac{\alpha - r}{\sigma}t.$$

Notice the square integrability requirement $\mathbf{E}\left[\int_0^T \left(\frac{\alpha-r}{\sigma}\right)^2 (Z(\tau))^2\,d\tau\right] < \infty$ is satisfied.

4.2.1 Proposition. *The expected value $\mathbf{E}[Z] = 1$, therefore it defines an equivalent probability measure $\tilde{P}$ as in 2.1.2. Under this measure, $\tilde{W}(t)$ is a Brownian motion.*

Proof. By Lévy's theorem 3.2.2, a stochastic process with continuous paths, starting at zero, with quadratic variation t at time t, that is a martingale, is a Brownian motion. Clearly $\tilde{W}(0) = 0$ and $\tilde{W}(t)$ is continuous because $W(t)$ and $\frac{\alpha-r}{\sigma}t$ are. The quadratic variation is equal to $[W, W]\,(t) = t$ because $\frac{\alpha-r}{\sigma}t$ is differentiable and introduces no quadratic variation. To see that $\tilde{W}(t)$ is a martingale, first consider the function

$$\varphi(x, y) := e^{-\frac{\alpha-r}{\sigma}y-\frac{1}{2}\left(\frac{\alpha-r}{\sigma}\right)^2 x}$$

and the Itô-Doeblin expansion for Brownian motion 3.4.4

$$
\begin{aligned}
Z(t) =& \varphi\left(t, W(t)\right) \\
=& \varphi\left(0, W(0)\right) + \int_0^t -\frac{1}{2}\left(\frac{\alpha - r}{\sigma}\right)^2 \varphi\left(\tau, W(\tau)\right) d\tau \\
&+ \int_0^t -\frac{\alpha - r}{\sigma}\varphi\left(\tau, W(\tau)\right) dW(\tau) \\
&+ \frac{1}{2}\int_0^t \left(\frac{\alpha - r}{\sigma}\right)^2 \varphi\left(\tau, W(\tau)\right) d\tau \\
=& Z(0) - \int_0^t \frac{\alpha - r}{\sigma}\varphi\left(\tau, W(\tau)\right) dW(\tau)
\end{aligned}
$$

which shows $Z(t)$ is a martingale because Itô integrals are by proposition 3.4.2.

Using Itô's product rule, the process $\tilde{W}(t)Z(t)$ can also be expanded into the sum of its initial value and an Itô integral, and is likewise a martingale.

Since $Z(t)$ is a martingale, it follows that

$$\mathbf{E}[Z] = \mathbf{E}\left[Z(T)\right] = Z(0) = 1$$

which justifies the probability measure $\tilde{P}$ based on Z. Furthermore

$$Z(t) = \mathbf{E}\left[Z(T)|\mathcal{F}(t)\right] = \mathbf{E}\left[Z|\mathcal{F}(t)\right]$$

so $Z(t)$ is a Radon-Nikodým derivative process [Shr10, 5.2]. Therefore, using Baye's rule and the martingale property of $\tilde{W}(t)Z(t)$, for $s \in [0, t]$,

$$\tilde{\mathbf{E}}\left[\tilde{W}(t)\middle|\mathcal{F}(s)\right] = \frac{1}{Z(s)}\mathbf{E}\left[\tilde{W}(t)Z(t)\middle|\mathcal{F}(s)\right] = \frac{\tilde{W}(s)Z(s)}{Z(s)} = \tilde{W}(s),$$

proving $\tilde{W}(t)$ is also a martingale. As such, $\tilde{W}(t)$ fulfills all conditions of Levy's theorem 3.2.2, and hence $\tilde{W}(t)$ is a Brownian motion under the probability measure $\tilde{P}$ of 2.1.2. ■

Since $\tilde{W}(t)$ is a Brownian motion, the increments $\tilde{W}(t) - \tilde{W}(s)$ with $s \in [0, t]$ are normally distributed under $\tilde{P}$, with variance $t - s$. In particular,

$$\text{(4.7)} \quad w := -\frac{\tilde{W}(t) - \tilde{W}(s)}{\sqrt{t - s}}$$

follows standard normal distribution under $\tilde{P}$. This is of particular interest in the following derivation of the Black-Scholes-Merton option valuation formula.

Using methods of stochastic calculus developed earlier, it can be shown [Shr10, section 5.2.3] that the Itô expansion of the discounted portfolio process can be rewritten using $\tilde{W}(t)$ as

$$\text{(4.8)} \quad e^{-rt}X(t) = X(0) + \int_0^t e^{-r\tau}\sigma\Delta(\tau)S(\tau)d\tilde{W}(\tau)$$

and has no regular integrals. Under $\tilde{P}$, the Itô integral is a martingale, and so is the discounted portfolio process. This means that $X(t)$ satisfies the equality

$$\tilde{\mathbf{E}}\left[e^{-rT}X(T)\middle|\mathcal{F}(t)\right] = e^{-rt}X(t)$$

and equality 4.4 and the Markov property of $W(t)$ imply

$$e^{-rt}V\left(t, S(t)\right) = \tilde{\mathbf{E}}\left[e^{-rT}V\left(T, S(T)\right)\middle|\mathcal{F}(t)\right].$$

Since e^{-rt} is nonrandom, the equation may be rewritten as

$$(4.9)\quad V\left(t, S(t)\right) = \tilde{\mathbf{E}}\left[e^{-r(T-t)}V\left(T, S(T)\right)\middle|\mathcal{F}(t)\right].$$

An exact solution for $V(t)$ in terms of σ, α, r, K, and T will be derived from 4.9 using the fact that w in 4.7 is distributed standard normally under $\tilde{P}$, hence in the risk-neutral expectation.

The asset value process $S(t)$ from 4.2 can be rewritten in terms of $\tilde{W}(t)$ by

$$\begin{aligned}
S(t) =& S(0)e^{\sigma W(t)+(\alpha-\frac{1}{2}\sigma^2)t}\\
=& S(0)e^{\sigma W(t)+\left((\alpha-r)+(r-\frac{1}{2}\sigma^2)\right)t}\\
=& S(0)e^{\sigma\left(W(t)+\frac{\alpha-r}{\sigma}t\right)+\left(r-\frac{1}{2}\sigma^2\right)t}
\end{aligned}$$

yielding

$$(4.10)\quad S(t) = S(0)e^{\sigma\tilde{W}(t)+\left(r-\frac{1}{2}\sigma^2\right)t}.$$

In particular, the final asset value $S(T)$ can be expressed as

$$\begin{aligned}
S(T) =& S(0)e^{\sigma\left(\tilde{W}(T)-\tilde{W}(t)+\tilde{W}(t)\right)+\left(r-\frac{1}{2}\sigma^2\right)(T-t+t)}\\
=& S(0)e^{\sigma W(t)+\left(r-\frac{1}{2}\sigma^2\right)t}e^{\sigma\left(\tilde{W}(T)-\tilde{W}(t)\right)+\left(r-\frac{1}{2}\sigma^2\right)(T-t)}\\
=& S(t)e^{\sigma\left(\tilde{W}(T)-\tilde{W}(t)\right)+\left(r-\frac{1}{2}\sigma^2\right)(T-t)}
\end{aligned}$$

which may be rewritten using w defined in 4.7 as

$$S(T) = S(t)e^{-\sigma w\sqrt{T-t}+\left(r-\frac{1}{2}\sigma^2\right)(T-t)}.$$

Recall that the final value of the option is $V\left(T, S(T)\right) = \left(S(T)-K\right)^+$, so substituting

$$V\left(T, S(T)\right) = \left(S(t)e^{-\sigma w\sqrt{T-t}+\left(r-\frac{1}{2}\sigma^2\right)(T-t)} - K\right)^+$$

into 4.9 yields the equation

$$V\left(t, S(t)\right) = \tilde{\mathbf{E}}\left[e^{-r(T-t)}\left(S(t)e^{-\sigma w\sqrt{T-t}+\left(r-\frac{1}{2}\sigma^2\right)(T-t)} - K\right)^+\middle|\mathcal{F}(t)\right].$$

Although unwieldy at first sight, the conditioning $\mathcal{F}(t)$ implies $S(t)$ is nonrandom and may be regarded a constant. Therefore it may be replaced by $y \in \mathbb{R}_{>0}$ to represent arbitrary asset value. Furthermore, as w follow standard normal distribution under $\tilde{P}$, the expectation is actually a regular integral over the real line with respect to the standard normal distribution function. Symbolically,

$$V(t,y) = \int_{-\infty}^{\infty} e^{-r(T-t)} \left(y e^{-\sigma u \sqrt{T-t} + \left(r - \frac{1}{2}\sigma^2\right)(T-t)} - K \right)^+ \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du.$$

The option value at maturity $V(T,y)$ is positive if and only if

$$\begin{aligned}
& y e^{-\sigma u \sqrt{T-t} + \left(r - \frac{1}{2}\sigma^2\right)(T-t)} > K \\
& \Leftrightarrow -\sigma u \sqrt{T-t} > \ln\left(\frac{K}{y}\right) - \left(r - \frac{1}{2}\sigma^2\right)(T-t) \\
& \Leftrightarrow u < \underline{d} := \frac{1}{\sigma\sqrt{T-t}} \left(\ln\left(\frac{y}{K}\right) + \left(r - \frac{1}{2}\sigma^2\right)(T-t) \right)
\end{aligned}$$

else it is zero. Therefore only those values $u < \underline{d}$ need be considered, and

$$\begin{aligned}
V(t,y) &= \int_{-\infty}^{\underline{d}} e^{-r(T-t)} \left(y e^{-\sigma u \sqrt{T-t} + \left(r - \frac{1}{2}\sigma^2\right)(T-t)} - K \right) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
&= \int_{-\infty}^{\underline{d}} e^{-r(T-t)} y e^{-\sigma u \sqrt{T-t} + \left(r - \frac{1}{2}\sigma^2\right)(T-t)} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
&\quad - \int_{-\infty}^{\underline{d}} e^{-r(T-t)} K \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
&= y \int_{-\infty}^{\underline{d}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2 - r(T-t) - \sigma u \sqrt{T-t} + \left(r - \frac{1}{2}\sigma^2\right)(T-t)} du \\
&\quad - e^{-r(T-t)} K \int_{-\infty}^{\underline{d}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
&= y \int_{-\infty}^{\underline{d}} \frac{1}{\sqrt{2\pi}} e^{-\frac{\left(u + \sigma\sqrt{T-t}\right)^2}{2}} du - e^{-r(T-t)} K \Phi(\underline{d}),
\end{aligned}$$

where Φ denotes the cumulative standard normal distribution function 2.2. The remaining integral is the cumulative distribution up until $\underline{d}$ of a random variable distributed normally with mean $-\sigma\sqrt{T-t}$ and variance 1, hence with $\overline{d} := \underline{d} + \sigma\sqrt{T-t}$ it is equal to $\Phi(\overline{d})$. With these substitutions, the Black-Scholes-Merton formula is

$$V(t,y) = y\Phi(\overline{d}) - e^{-r(T-t)} K \Phi(\underline{d})$$

providing, under the assumptions of chapter 4, for each $t \in [0,T]$ and any asset value $y \in \mathbb{R}_{>0}$ the unique no-arbitrage value of a European call option with strike price K, for fixed volatility σ, asset rate of return α, and numeraire rate of return r. The fair price at which the issuer must sell the option to facilitate his hedge and prevent arbitrage on his side is $V(0, S(0))$ and his perfect hedging strategy through time is dictated by the delta-hedging rule 4.5 $\Delta(t) = \mathbf{D}_2[V](t, S(t))$.

Bibliography

[BS73] Fischer Black and Myron S. Scholes, *The pricing of options and corporate liabilities*, Journal of Political Economy **81** (1973).

[Chu63] Ruel V. Churchill, *Fourier series and boundary value problems*, Second edition, McGraw-Hill Book Co., Inc., New York-San Francisco, Calif.-Toronto-London, 1963. MR 0149173 (26 #6665)

[Cie61] Zbigniew Ciesielski, *Hölder conditions for realizations of Gaussian processes*, Transactions of the American Mathematical Society **99** (1961), 403–413. MR 0132591 (24 #A2431)

[Fam65] Eugene F. Fama, *The behavior of stockmarket prices*, Journal of Business **38** (1965).

[Gir60] Igor V. Girsanov, *On transforming a certain class of stochastic processes by absolutely continuous substitution of measures*, Theory of Probability and Its Applications **5** (1960).

[Knu93] Donald E. Knuth, *Johann Faulhaber and sums of powers*, Mathematics of Computation **61** (1993), no. 203, 277–294. MR 1197512 (94a:11030)

[Lév48] Paul Lévy, *Processus stochastiques et mouvement Brownien. Suivi d'une note de M. Loève*, Gauthier-Villars, Paris, 1948. MR 0029120 (10,551a)

[McK69] Henry P. McKean, Jr., *Stochastic integrals*, Probability and Mathematical Statistics, No. 5, Academic Press, New York-London, 1969. MR 0247684 (40 #947)

[Mer73] Robert C. Merton, *Theory of rational option pricing*, The Rand Journal of Economics **4** (1973), 141–183. MR 0496534 (58 #15058)

[Nob97] Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel, *press release*, 1997, retrieved August 25th, 2014, from `http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/1997/press.html`.

[Pol84] David Pollard, *Convergence of stochastic processes*, Springer Series in Statistics, Springer-Verlag, New York, 1984. MR 762984 (86i:60074)

[Ric07] John A. Rice, *Mathematical statistics and data analysis*, 3 ed., Brooks/Cole, 2007.

[Sch05] René L. Schilling, *Measures, integrals and martingales*, Cambridge University Press, New York, 2005. MR 2200059 (2007g:28001)

[Shr04] Steven E. Shreve, *Stochastic calculus for finance I: the binomial asset pricing model*, Springer Finance, Springer-Verlag, New York, 2004. MR 2049045 (2004m:91003)

[Shr10] ———, *Stochastic calculus for finance II: continuous-time models*, Springer Finance, Springer-Verlag, New York, 2010. MR 2057928 (2005c:91001)