STOCHASTIC CALCULUS AND BLACK-SCHOLES THEORY LECTURE NOTES

ZDZISLAW BRZEZNIAK, EVGENIY ZORIN, CHRISTIAN LITTERER

References

[1] Brzeźniak, Zdzisław and Zastawniak, Tomasz, Basic stochastic processes. A course through exercises. Springer Undergraduate Mathematics Series. Springer-Verlag London, Ltd., London, 1999. x+225 pp. ISBN: 3-540-76175-6 60-01.
[2] Marek Capinski, Ekkehard Kopp and Janusz Traple, Stochastic Calculus for Finance
[3] Marek Capinski and Ekkehard Kopp, The Black-Scholes Model
[4] Lamberton, Damien and Lapeyre, Bernard, Introduction to stochastic calculus applied to finance. Second edition. Chapman & Hall/CRC Financial Mathematics Series. Boca Raton, FL, 2008. 253 pp. ISBN: 978-1-58488-626-6 91-01 (60H10 60H30 91B28)
[5] Fouque, Jean-Pierre; Papanicolaou, George; Sircar, K. Ronnie, Derivatives in financial markets with stochastic volatility. Cambridge University Press, Cambridge, 2000. xiv+201 pp. ISBN: 0-521-79163-4
[6] Øksendal, Bernt, Stochastic differential equations. An introduction with applications. Sixth edition. Universitext. Springer-Verlag, Berlin, 2003. xxiv+360 pp. ISBN: 3-540-04758-1
[7] P. Wilmott, Derivatives. The Theory and Practice of Financial Engineering, John Wiley and Sons, New York, 1998.
[8] Steele, Michael, Stochastic Calculus and Financial Applications, Springer.
[9] Musiela, Marek and Rutkowski, Marek, Martingale Methods in Financial Modelling, Springer.

1. The stochastic integral

1.1. **Brownian Motion.** We begin by recalling the notions of a stochastic process and Brownian Motion, and some of their basic properties.

Definition 1.1. *A (real valued) stochastic process is a parametrised family of random variables*

$$\{X_t\}_{t\in S}$$

Date: January 16, 2022.

This notes are based on the previous version prepared with the help of the following students. Lecture 1 has been typed in LaTex by Chen, Lecture 2 by Anastasios Biniaris, lecture 6 by Sarah Softley, lecture 8 by Charnchai.

defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ *taking values in* $\mathbb{R}$*, where* S *is either* $S = [0, T], T > 0$ *or* $S = [0, \infty)$*.*

Definition 1.2. *A* ***Brownian Motion*** *is a stochastic process* $(W_t)_{t\geq 0}$ *on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *such that*

(i) $W_0 = 0$ *a.s.*

(ii) Independent increments:for all $k \in \mathbb{N}$ *and all* $0 = t_0 < t_1 < \cdots < t_k$ *the random variables* $W_{t_1} - W_{t_0}$*,* $W_{t_2} - W_{t_1}$*,* $\ldots$*,* $W_{t_k} - W_{t_{k-1}}$ *are independent*

(iii) $W_t - W_s \sim N(0, t-s)$ *for all* $t > s \geq 0$*, i.e.* $B_t - B_s$ *has probability density function (pdf)*

$$p\left(t - s.x\right) = \frac{1}{\left(2\pi \left(t - s\right)\right)^{1/2}} \exp\left(-\frac{|x|^2}{2\left(t - s\right)}\right).$$

(iv) Trajectories are continuous a.s., i.e. there exists $\Omega_1 \subset \Omega$ *such that* $\mathbb{P}(\Omega_1) = 1$ *and for all* $\omega \in \Omega_1$ $t \to W_t(\omega)$ *is continuous for all* $t \geq 0$*.*

Note that if W is a Brownian Motion its natural filtration is given by

$$\mathcal{F}_t = \sigma\{W_s : s \leq t\}.$$

By definition of the natural filtration W_t is $\mathcal{F}_t$-measurable, i.e. W_t is a r.v. with respect to the σ-field $\mathcal{F}_t$. Furthermore, if $t > s \geq 0$ then $W_t - W_s$ is independent of the σ-field $\mathcal{F}_s$.

1.2. **Stochastic processes and filtrations.** In this subsection I have gathered a few of the basic definitions related to stochastic processes and filtrations on probability spaces. These notions are familiar from the MMF lectures and will be used throughout the course.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space. A filtration on $(\Omega, \mathcal{F}, \mathbb{P})$ is a family of σ-fields $(\mathcal{F}_t)_{t\geq 0}$, indexed with positive real numbers $t \geq 0$ (the index t is usually understood as a time), such that

(1) $\mathcal{F}_t \subset \mathcal{F}$ for every $t \geq 0$,

(2) $\mathcal{F}_s \subset \mathcal{F}_t$ for all $0 \leq s \leq t$.

Intuitively, one can think of a filtration as the information available at time t.

Definition 1.3. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $(\mathcal{F}_t)_{t\geq 0}$ be a filtration on it. A process $(\xi_t)_{t\geq 0}$ is said to be $(\mathcal{F}_t)_{t\geq 0}$* **adapted** *(or adapted w.r.t the filtration $(\mathcal{F}_t)_{t\geq 0}$) if ξ_t is $\mathcal{F}_t$-measurable for all $t \geq 0$.*

Example 1.4. *Let $(W_t)_{t\geq 0}$ be a BM on probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $(\mathcal{F}_t)_{t\geq 0} = (\sigma\{W_s : s \leq t\})_{t\geq 0}$ be the natural filtration generated by Brownian Motion. Then $(W_t)_{t\geq 0}$ is adapted with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$.*

Example 1.5. *Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and let $(W_t^1)_{t\geq 0}$ and $(W_t^2)_{t\geq 0}$ be two independent BMs. Define*

$$\mathcal{F}_t^1 = \sigma\{W_s^1 : s \leq t\}\ ,\ \mathcal{F}_t^2 = \sigma\{W_s^2 : s \leq t\}\ ,\ \mathcal{F}_t = \sigma\{W_s^1, W_s^2 : s \leq t\}$$

Then, on the one hand $(W_t^1)_{t\geq 0}$ is an adapted process w.r.t filtration $(\mathcal{F}_t^1)_{t\geq 0}$ and $(\mathcal{F}_t)_{t\geq 0}$ but not w.r.t the filtration $(\mathcal{F}_t^2)_{t\geq 0}$. On the other hand $(W_t^2)_{t\geq 0}$ is an adapted process w.r.t the filtration $(\mathcal{F}_t^2)_{t\geq 0}$ and $(\mathcal{F}_t)_{t\geq 0}$ but not w.r.t the filtration $(\mathcal{F}_t^1)_{t\geq 0}$. Moreover, $W_t^i - W_s^i$ is independent of the σ-fields $\mathcal{F}_s^i$ (known) and $\mathcal{F}_s, i = 1, 2$.

Definition 1.6. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $(\mathcal{F}_t)_{t\geq 0}$ be a filtration on this space. Further, let $(W_t)_{t\geq 0}$ be a Brownian Motion on the same probability space. We say that $(W_t)_{t\geq 0}$ is a Brownian Motion with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$ (or $(\mathcal{F}_t)_{t\geq 0}$-Brownian motion) if:*

(1) the random variable W_t is $\mathcal{F}_t$-measurable for every $t \geq 0$,

(2) the random variable $W_t - W_s$ is independent of $\mathcal{F}_s$ for all $0 \leq s < t$.

Example 1.7. *As in Example 1.4 let* $(W_t)_{t\geq 0}$ *be a BM on probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and* $(\mathcal{F}_t)_{t\geq 0} = (\sigma\{W_s : s \leq t\})_{t\geq 0}$ *be the natural filtration generated by Brownian Motion. Then (trivially)* $(W_t)_{t\geq 0}$ *is a BM with respect to the filtration* $(\mathcal{F}_t)_{t\geq 0}$.

The above definitions of probability distributions, stochastic processes and Brownian motion can be generalised to the multi-dimensional setting.The bivariate normal distribution generalises the normal distribution to two-dimensional random vectors and is defined as follows: Consider a real valued matrix Σ given by

$$\Sigma := \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}.$$

Assume that Σ has non-negative entries σ_{ij}, is also symmetric (i.e. that means $\sigma_{21} = \sigma_{12}$) and satisfies

$$\Sigma_{11}\Sigma_{22} - \Sigma_{12}\Sigma_{21} \geq 0.$$

Definition . *A two dimensional vector of random variables* $Z := (X_1, X_2)$ *is said to have (non-degenerate) bi-variate normal distribution with co-variance matrix*

$$\Sigma := \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

and mean zero if if its density is of the form

$$f(x) = \frac{1}{2\pi} \left(\Sigma_{11}\Sigma_{22} - \Sigma_{12}\Sigma_{21}\right)^{-1/2} \exp\left(-\frac{1}{2} Z^T \Sigma^{-1} Z\right),$$

where Σ^{-1} *is the inverse of the matrix* Σ *and the superscript* T *denotes the transpose of the vector.*

Examples:

(1) If X and Y are both centred, independent (univariate) normal random variables then $Z := (X, Y)$ is a bivariate normal variable with covariance matrix

$$\Sigma = \begin{pmatrix} \mathrm{Var}\,(X) & 0 \\ 0 & \mathrm{Var}\,(Y) \end{pmatrix}.$$

(2) If the independence assumption is dropped in the previous example Z may no longer be normally distributed.

(3) If A is a non-singular (i.e. invertible) linear map from $\mathbb{R}^2$ to $\mathbb{R}^2$, then AZ is once again a bivariate normal random variable.

It's an important observation (that can be proved for example using a suitable multi-variate extension of the moment-generating functions) that if (X, Y) are jointly (bivariate) normal random variables then X and Y are independent if and only if $\mathrm{Cov}(X, Y) = 0$. Also, note that any linear combination of X and Y (strictly speaking with the appropriate conventions to cover degenerate cases)cases) is once again a univariate normal distribution. The results generalise to random n -vectors. In that case the covariance matrices are assumed to be positive-definite symmetric matrices.

In this course we will in the following focus on the scalar valued case.

1.3. **Itô integral for random step processes.**

1.3.1. *Some brief motivation: The Riemann integral*

. Let f be a function $f : [a, b] \to \mathbb{R} \quad a, b \in \mathbb{R}$. Intuitively, the Riemann integral aims to capture the "area under the graph of a function". To this end one divides the interval $[a, b]$ into a large number of small intervals e.g

$$t_0^n = a < t_1^n = a + \frac{1}{N}(b - a) < \cdots < t_n^n = b$$

and calculates the approximate area for each small interval. We will call such a subdivision of an interval in the following a partition.

Definition 1.8. *Let $k \in \mathbb{N}$. We call a sequence of times*

$$t_0 = a < t_1 < t_2 < \cdots < t_{k-1} < t_k = b$$

a partition of the interval $[a, b]$.

For $x_i^n \in [t_{i-1}^n, t_i^n]$ we define the approximate area under the graph of f by letting

$$\text{“Approximate area”} = \sum_{i=1}^{n} f(x_i^n)(t_i^n - t_{i-1}^n). \tag{1.1}$$

If $f : [a, b] \to \mathbb{R}$ is continuous it can be shown that the approximate area has a limit as $N \to \infty$ and this limit is independent of the choice of points x_i^n at which we evaluate the function in each subinterval.

Example 1.9. *Let $f : [0, 1] \to \mathbb{R}$ be the function defined by $f(x) = x^2$. Approximate area for $n = 10$, $n = 30$, $n = 50$.*

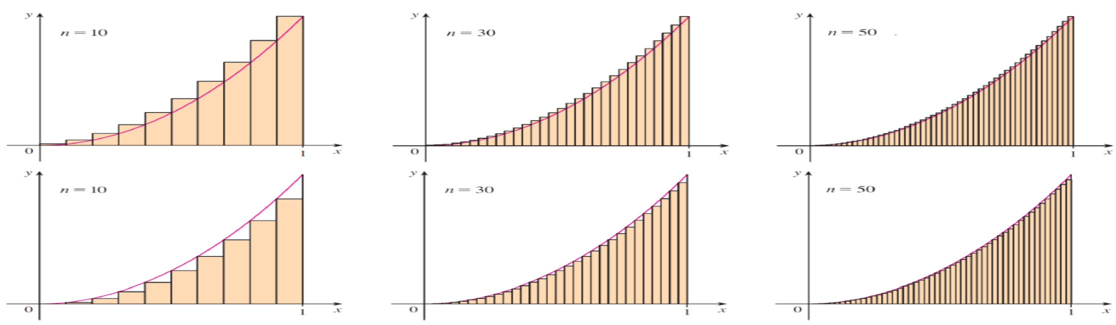

FIGURE 1. The finer the partition, the better the approximation we get

The lesson of this example is the following: If we need to integrate an integrand (such as a function, a stochastic process etc.) that changes continuously in time (in our example the function $f(x) = x^2$), then we

(a) take approximations (of the integrand) that change their values only at a finite number of times $t_0, \ldots, t_n$ (i.e. the approximations are piecewise constant in the time variable). In our example we use step functions for the approximation;
(b) define and calculate the integrals of the approximations in (a) as sums (in our example as in 1.1)
(c) show that the sums from (b) converge to a limit S as n tends to infinity. If such a limit S indeed exists and does not depend on the sequence of approximations chosen, we define the integral to be equal to S.

Whenever this strategy is viable, we obtain an *integral* of our initial object. For functions $f : [a, b] \to \mathbb{R}$, the three steps in the above construction correspond to the construction of the Riemann integral. Note that if a function is both Riemann and Lebesgue integrable the two notions of integration coincide.

In the remainder this section we execute the above strategy when the integrand is a *stochastic process*. We start by defining the objects that will be used as our approximations. Note that these processes change value only at a finite number of times (see point (a) above).

Definition 1.10. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $(W_t)_{t\geq 0}$ be a Brownian motion w.r.t a filtration $(\mathcal{F}_t)_{t\geq 0}$. A **random step process** $(\xi_t)_{t\geq 0}$ is a process satisfying the following conditions*

(i) $\mathbb{E}|\xi_t|^2 < \infty$ *for all* $t \geq 0$,

(ii) there exists a finite sequence $(t_0, \dots, t_N)$ *such that*

$0 = t_0 < t_1 < \cdots < t_N < \infty$ *and for* $i = 0, \dots, N-1$

$$\xi_t = \begin{cases} \xi_{t_i}, & \text{if } t \in [t_i, t_{i+1}) \\ 0, & \text{if } t \geq t_N. \end{cases}$$

(iii) $(\xi_t)_{t\geq 0}$ *is adapted to* $(\mathcal{F}_t)_{t\geq 0}$.

For such a process $(\xi_t)_{t\geq 0}$ *we define the integral* [1] *by letting*

$$I(\xi) := \sum_{i=0}^{N-1} \xi_{t_i}(W_{t_{i+1}} - W_{t_i}).$$

Remark *Note that the integral* $I(\xi)$ *is a random variable.*

We denote the space of random step processes by $M^2_{\text{step}}(0, \infty)$.

Example 1.11. ***(a)*** *Let* ξ_t *be given by*

$$\xi_t = \begin{cases} 1, & \text{if } t \in [0, 1) \\ W_1, & \text{if } t \in [1, 3) \\ 0, & \text{if } t \geq 3. \end{cases}$$

Is $(\xi_t)_{t\geq 0}$ *a random step process?*

[1] It is not hard to show that the integral is well defined, i.e. independent of the choice of partition

Verification of condition (iii). Is $(\xi_t)_{t\geq 0}$ adapted to $(\mathcal{F}_t)_{t\geq 0}$?

Case 1 . $t \in [0,1)$, $\xi_t = 1$. This is a constant random variable, so it is measurable with respect to any σ-field, in particular with respect to $\mathcal{F}_t$.

Case 2 . $t \in [1,3)$, $\xi_t = W_1$ is $\mathcal{F}_1$-measurable, because W_t is adapted to $\mathcal{F}_t$. But $\mathcal{F}_1 \subset \mathcal{F}_t$ (because $(\mathcal{F}_t)_{t\geq 0}$ is a filtration at $t \geq 1$). Hence, ξ_t is measurable with respect to $\mathcal{F}_t$.

Case 3 . $t \geq 3$, $\xi_t = 0$, is a constant, thus the same as Case 1. □

Verification of condition (i). $\mathbb{E}|\xi_t|^2 < \infty$?

Case 1 . $t \in [0,1)$, $\mathbb{E}(\xi_t)^2 = \mathbb{E}(1) = 1 < \infty$.

Case 3 . $t \in [3,\infty)$, $\mathbb{E}(\xi_t)^2 = 0 < \infty$.

Case 2 . $t \in [1,3)$, $\xi_t^2 = W_1^2$, so $\mathbb{E}(\xi_t)^2 = \mathbb{E}(W_1)^2 = 1$. □

Verification of condition (ii). We note that (ii) is satisfied if we take $N = 2$, $t_0 = 0, t_1 = 1, t_2 = 3$. Note that this is the natural choice to satisfy this last condition, but by no means the only one (why?). □

We deduce that $(\xi_t)_{t\geq 0}$ is a random step process and

$$\begin{aligned} I(\xi) &= \sum_{i=0}^{N-1} \xi_{t_i}(W_{t_{i+1}} - W_{t_i}) \\ &= \xi_{t_0}(W_{t_1} - W_{t_0}) + \xi_{t_1}(W_{t_2} - W_{t_1}) \\ &= 1 \cdot W_1 + W_1(W_3 - W_1). \end{aligned}$$

We have

$$\begin{aligned}
\mathbb{E}|I(\xi)|^2 &= \mathbb{E}(W_1 + W_1(W_3 - W_1))^2 \\
&= \mathbb{E}[W_1^2(W_3 - W_1 + 1)^2] \\
&= \mathbb{E}(W_1^2)\mathbb{E}(W_3 - W_1 + 1)^2 \\
&= 1 \cdot \mathbb{E}[(W_3 - W_1)^2 + 2(W_3 - W_1) + 1] \\
&= (3-1) + 2 \cdot 0 + 1 \\
&= 3.
\end{aligned}$$

Since W_1 is measurable with respect to $\mathcal{F}_1$, and $W_3 - W_1$ is independent of $\mathcal{F}_1$.

(b) *Define $(\xi_t)_{t\geq 0}$ by*

$$\xi_t = \begin{cases} 2, & \text{if } t \in [0,1) \\ W_1^2, & \text{if } t \in [1,2) \\ W_3, & \text{if } t \in [2,4) \\ 0, & \text{if } t \in [4,\infty). \end{cases}$$

Is $(\xi_t)_{t\geq 0}$ a random step process?

Three conditions need to be checked:

(i) Square integrability.

(ii) Step property.

(iii) Is this process adapted to $\mathcal{F}_t$?

Condition (ii) is satisfied with $N = 3$, $t_0 = 0, t_1 = 1, t_2 = 2, t_3 = 4$.

Is (iii) satisfied?

Case 1 . $t \in [0,1)$, $\xi_t = 2$ *is constant. Obviously, O.K.*

Case 4 . $t \geq 4$, $\xi_t = 0$ *is constant. Obviously, O.K.*

Case 2 . $t \in [1,2)$, $\xi_t = (W_1)^2$ *is $\mathcal{F}_1$-measurable. But $\mathcal{F}_1 \subset \mathcal{F}_t$, so ξ_t is also $\mathcal{F}_t$-measurable.*

Case 3 . $t \in [2,4)$, $\xi_t = W_3$ *is $\mathcal{F}_3$-measurable, but not $\mathcal{F}_2$-measurable (Exercice!).*

So, (iii) is not satisfied. ***Conclusion****: $(\xi_t)_{t\geq 0}$ is not a random step process because* *for* $t \in [2,3)$. ξ_t *is not* $\mathcal{F}_t$*-measurable.*

(c) *We want to modify example (b) so that we get a random step process.*

$$\xi_t = \begin{cases} 2, & \text{if } t \in [0,1) \\ W_1^2, & \text{if } t \in [1,2) \\ 3W_2, & \text{if } t \in [2,4) \\ 0, & \text{if } t \in [4,\infty). \end{cases}$$

It is not hard to verify all three conditions and conclude that $(\xi_t)_{t\geq 0}$ *is a random process with*

$$\begin{aligned} I(\xi) &= \sum_{i=0}^{N-1} \xi_{t_i}(W_{t_{i+1}} - W_{t_i}) \\ &= \sum_{i=0}^{2} \cdots \\ &= 2W_1 + W_1^2(W_2 - W_1) + 3W_2(W_4 - W_2) \end{aligned}$$

The calculation of $\mathbb{E}|I(\xi)|^2$ *is left as an exercise.*

The following propositions capture important properties of the space $M^2_{step}(0,\infty)$.

Proposition . $M^2_{step}(0,\infty)$ *is a vector space, i.e. if* $\xi \in M^2_{step}(0,\infty)$, $\eta \in M^2_{step}(0,\infty)$ *and* $\alpha, \beta \in \mathbb{R}$, *then* $(\alpha\xi + \beta\eta) \in M^2_{step}(0,\infty)$ *and*

$$I(\alpha\xi + \beta\eta) = \alpha I(\xi) + \beta I(\eta). \tag{1.2}$$

Here $(\alpha\xi + \beta\eta)$ *is the stochastic process defined by*

$$(\alpha\xi + \beta\eta)(t) = \alpha\xi(t) + \beta\eta(t),\ t \geq 0.$$

Note that equality (1.2) means that the integral $I(\cdot)$ is linear on $M^2_{step}(0,\infty)$. The next proposition is of fundamental importance for the construction of the Ito integral for stochastic processes.

Proposition 1.12. ***(Itô Isometry for random step processes)*** *Suppose that ξ is a random step process, i.e. $\xi \in M^2_{step}(0,\infty)$. Then,*

$$\mathbb{E}|I(\xi)|^2 = \int_0^\infty \mathbb{E}|\xi(s)|^2 \, ds.$$

Before we prove the proposition we consider a simple example.

Example 1.13. *Let ξ be as in Example 1.11 (a):*

$$\xi_t = \begin{cases} 1, & \text{if } t \in [0,1) \\ W_1, & \text{if } t \in [1,3) \\ 0, & \text{if } t \geq 3. \end{cases}$$

Calculating $\mathbb{E}|I(\xi)|^2$ using Proposition 1.12 gives

$$\mathbb{E}|I(\xi)|^2 = \int_0^\infty \mathbb{E}|\xi_t|^2 \, dt = 1 \cdot (3-0) = 3$$

which agrees with the result from Example 1.11 (a).

Proof of Proposition 1.12. If $\xi \in M^2_{step}(0,\infty)$, then we can find $N \in \mathbb{N}$ and a sequence of times

$$0 = t_0 < t_1 < \ldots < t_N$$

such that

$$\xi(t) = \begin{cases} \xi(t_i), & \text{if } t \in [t_i, t_{i+1}), \ldots \quad i = 0, \ldots, N-1 \\ 0, & \text{if } t \geq t_N. \end{cases}$$

Hence, we have on the one hand

$$
\begin{aligned}
\mathbb{E}|I(\xi)|^{2} &= \mathbb{E}\left[\sum_{i=0}^{N-1} \xi(t_i)(W(t_{i+1})-W(t_i))\right]^2 \\
&= \mathbb{E}\left[\sum_{i=0}^{n-1} \xi(t_i)^2(W(t_{i+1})-W(t_i))^2\right. \\
&+ \left. 2\sum_{0\le i<j\le N-1} \xi(t_i)\xi(t_j)[W(t_{i+1})-W(t_i)][W(t_{j+1})-W(t_j)]\right] \\
&= \sum_{i=0}^{N-1} \mathbb{E}\Big[\xi(t_i)^2(W(t_{i+1})-W(t_i))^2\Big] \\
&+ 2\sum_{0\le i<j\le N-1} \mathbb{E}\Big[\xi(t_i)\xi(t_j)[W(t_{i+1})-W(t_i)][W(t_{j+1})-W(t_j)]\Big] \qquad (1.3)
\end{aligned}
$$

Recall from MMF that if X and Y are independent random variables we have

$$\mathbb{E}\,(XY) = \mathbb{E}\,(X)\,\mathbb{E}\,(Y)\,.$$

Since $\xi(t_i)^2$ is $\mathcal{F}_{t_i}-measurable$ and $W(t_{i+1})-W(t_i)$ is independent of $\mathcal{F}_{t_i}$ we can apply this identity to the terms in the first sum in (1.3) and have for $i=0,\dots,N-1$

$$
\begin{aligned}
\mathbb{E}\Big[\xi(t_i)^2(W(t_{i+1})-W(t_i))^2\Big] &= \mathbb{E}\Big[\xi(t_i)^2\Big]\mathbb{E}\Big[(W(t_{i+1})-W(t_i))^2\Big] \\
&= \mathbb{E}\Big[\xi(t_i)^2\Big](t_{i+1}-t_i).
\end{aligned}
$$

For the final step we have used that $W(t_{i+1})-W(t_i)$ has distribution $N(0,t_{i+1}-t_i)$. For the terms in the second sum we note that the random variables $\xi(t_i), W(t_{i+1})$ and $W(t_i)$ are $\mathcal{F}_{t_j}-measurable$ and $W(t_{j+1})-W(t_j)$ is independent of $\mathcal{F}_{t_j}$. Hence,

$$
\begin{aligned}
&\mathbb{E}\Big[\xi(t_i)\xi(t_j)[W(t_{i+1})-W(t_i)][W(t_{j+1})-W(t_j)]\Big] \\
&= \mathbb{E}\Big[\xi(t_i)\xi(t_j)[W(t_{i+1})-W(t_i)]\Big]\mathbb{E}[W(t_{j+1})-W(t_j)] = 0
\end{aligned}
$$

as $\mathbb{E}\,[W(t_{j+1})-W(t_j)]=0$ by the definition of Brownian Motion. Combining the two calculations we see that

$$\mathbb{E}|I(\xi)|^{2} = \sum_{i=0}^{N-1}\mathbb{E}\Big[\xi(t_i)^2\Big](t_{i+1}-t_i).$$

On the other hand, note that

$$\mathbb{E}\,|\xi|^{\,2} = \begin{cases} \mathbb{E}|\xi(t_i)|^2, & if\ t \in [t_i, t_{i+1}),\ i \le N-1 \\ 0, & if\ t \ge t_N \end{cases}$$

and therefore

$$\begin{aligned} & \int_0^\infty \mathbb{E}|\xi_t|^{\,2} dt \\ = \quad & \mathbb{E}|\xi_{t_0}|^{\,2}(t_1 - t_0) + \mathbb{E}|\xi_{t_1}|^{\,2}(t_2 - t_1) + \ldots + \mathbb{E}|\xi_{t_{N-1}}|^{\,2}(t_N - t_{N-1}) \end{aligned}$$

completing the proof of the proposition. □

Corollary 1.14. *Suppose that $\xi, \eta \in M^2_{step}(0, \infty)$, then*

$$\mathbb{E}|I(\xi) - I(\eta)|^2 = \int_0^\infty \mathbb{E}|\xi(s) - \eta(s)|^2\, ds.$$

The proof of this corollary will be an exercise.

Example 1.15. *Consider two random step processes,*

$$\xi(t) \;=\; \begin{cases} 3, & \text{if } t \in [0,2) \\ W_2^2, & \text{if } t \in [2,5) \\ 3W_5, & \text{if } t \in [5,7) \\ 0, & \text{if } t \ge 7 \end{cases}, \quad \eta(t) \;=\; \begin{cases} 2, & \text{if } t \in [0,1) \\ W_1^3, & \text{if } t \in [1,3) \\ 2W_3^7, & \text{if } t \in [3,6) \\ 0, & \text{if } t \ge 6 \end{cases},$$

where

$$t^1_0 = 0,\ t^1_1 = 2,\ t^1_2 = 5,\ t^1_3 = 7,\ t^2_0 = 0,\ t^2_1 = 1,\ t^2_2 = 3,\ t^2_3 = 6.$$

If we let $X(t) = \xi(t) - \eta(t)$, then X is also a step process and the set of times where X can jump is exactly the union of corresponding sets of the process ξ and η $(t_0 = 0, t_1 = 1, t_2 = 2, t_3 = 3, t_4 = 5, t_5 = 6, t_6 = 7)$

Example 1.16. *Take $T > 0$ and define a process ξ by*

(•)

$$\xi(t) = \begin{cases} W(t), & if\, t < T \\ 0, & if\, t \ge T \end{cases}$$

The process $\xi(t)$ satisfies conditions (i) and (ii) in the definition of a random step process, Definition 1.10.

(ii) $\xi(t)$ *is* $\mathcal{F}_t$*-measurable for* $t \geq 0$*, i.e. the process* $(\xi_t)_{t\geq 0}$ *is* $(\mathcal{F}_t)_{t\geq 0}$*-adapted.*

(iii)

$$\begin{aligned} \mathbb{E}|\xi_t|^2 &= \begin{cases} \mathbb{E}(W_t^2), & \textit{if } t < T \\ \mathbb{E}(0), & \textit{if } t \geq T \end{cases} \\ &= \begin{cases} t, & \textit{if } t < T \\ 0, & \textit{if } t \geq T \end{cases} \\ &< \infty \end{aligned}$$

But $\xi \notin M^2_{step}(0,\infty)$ *because trajectory of* ξ_t *is not piecewise constant.*

1.4. **Itô integral, general case.** In this subsection we will carry out the construction of the Itô integral in the general case.

Corollary 1.14 suggests that we should find a sequence $\xi = (\xi^n(t))_{t\geq 0}$ of processes belonging to the class M^2_{step} such that $\xi^n(t) = 0$, if $t \geq T$ and

$$\mathbb{E}\int_0^T |\xi^n(t) - \xi(t)|^2\, dt \to 0 \text{ as } n \to \infty. \tag{1.4}$$

Assume we have found such a sequence. Does the sequence $I(\xi^n)$ converge to a random variable L as $n \to \infty$ in mean square? Does the limit depend on the choice of approximating sequence? If we can give a positive answers to both questions we can use the approximating sequence to define the Ito integral for ξ by setting

$$I(\xi) := L.$$

Using the simple inequality $(a+b)^2 \leq 2(a^2+b^2)$ we have

$$\begin{aligned} \mathbb{E}\int_0^T |\xi^n(t) - \xi^k(t)|^2\, dt &= \mathbb{E}\int_0^T |\xi^n(t) - \xi(t) + \xi(t) - \xi^k(t)|^2\, dt \\ \leq \mathbb{E}\int_0^T 2[|\xi^n(t) - \xi(t)|^2 &+ |\xi(t) - \xi^k(t)|^2]\, dt \\ = 2\mathbb{E}\int_0^T |\xi^n(t) - \xi(t)|^2\, dt &+ 2\mathbb{E}\int_0^T |\xi^k(t) - \xi(t)|^2\, dt. \end{aligned}$$

Since $\mathbb{E}\int_0^T |\xi^n(t) - \xi(t)|^2\,dt$ is small when n is large and $\mathbb{E}\int_0^T |\xi^k(t) - \xi(t)|^2]\,dt$ is small when k is large, we infer that

$$\mathbb{E}\int_0^T |\xi^n(t) - \xi^k(t)|^2\,dt \text{ is small when both } k \text{ and } n \text{ are large.} \tag{1.5}$$

Since both ξ^n and ξ^k are elements of M^2_{step} we can define $I(\xi^n)$ and $I(\xi^k)$. Moreover, since M^2_{step} is a vector space, $\xi^n - \xi^k \in M^2_{step}$ and

$$I(\xi^n - \xi^k) = I(\xi^n) - I(\xi^k).$$

So by Corollary 1.14 we have

$$\begin{aligned}
\mathbb{E}|I(\xi^n) - I(\xi^k)|^2 &= \mathbb{E}|I(\xi^n - \xi^k)|^2 = \mathbb{E}\int_0^\infty |\xi^n(t) - \xi^k(t)|^2\,dt \\
&= \mathbb{E}\int_0^T |\xi^n(t) - \xi^k(t)|^2\,dt.
\end{aligned}$$

We conclude with (1.5):

$$\mathbb{E}|\ I(\xi^n) - I(\xi^k)|^2 \text{ is small when both } k \text{ and } n \text{ are large,}$$

i.e. $I\,(\xi^n)$ is a Cauchy sequence. Using the **completeness** of the Hilbert space of square integrable random variables L^2 we can find a random variable which we denote by I(ξ) such that

$$\mathbb{E}|I(\xi^n) - I(\xi)|^2 \to 0 \text{ as } n \to \infty. \tag{1.6}$$

This random variable $I(\xi)$ does not depend on the approximating sequence and will in the following also be denoted by $\int_0^T \xi(s)\,dW(s)$. It is the Itô integral of the process $\xi(t)$ from 0 to T.

Summary To find the Itô integral we need to find a r.v. $I(\xi)$ satisfying condition (1.6).

The following two examples show how to find a sequence of random step processes that approximates a process ξ in the sense of (1.4) in two special cases. The constructions in the two cases can be generalised to a broader class of processes.

Example 1.17. *Suppose we want to define $I(\xi)$, where $T > 0$ and ξ is the process defined by*

$$\xi(t) = \begin{cases} t, & \text{if } t \in [0,T) \\ 0, & \text{if } t \geq T \end{cases}.$$

First method. This method works in general only when the process $\xi(t)$, $t \geq 0$, is continuous (i.e. it has a.s. continuous trajectories), which is obviously fine in our example.

We introduce the following notation: Let $n \in \mathbb{N}$.

$$t_0^n = 0, t_1^n = \frac{T}{n}, t_2^n = \frac{2T}{n}, \dots, t_{n-1}^n = \frac{(n-1)T}{n}, t_n^n = T.$$

and define

$$\xi^n(t) = \begin{cases} \xi(t_i^n); & \text{if } t \in [t_i^n, t_{i+1}^n), i = 0, , \cdots, , n-1 \\ 0 & \text{if } t \geq T. \end{cases}$$

For example, for $n = 5$ the new process ξ^5 is then defined as follows.
If $t \in [t_0^5, t_1^5) = [0, \frac{T}{5})$, we put $\xi^5(t) = \xi(t_0^5) = \xi(0) = 0$.
If $t \in [t_1^5, t_2^5) = [\frac{T}{5}, \frac{2T}{5})$, we put $\xi^5(t) = \xi(t_1^5) = \xi(\frac{T}{5}) = \frac{T}{5}$.
Similarly, if $t \in [t_2^5, t_3^5) = [\frac{2T}{5}, \frac{3T}{5})$, we put $\xi^5(t) = \xi(t_2^5) = \xi(\frac{2T}{5}) = \frac{2T}{5}$; if $t \in [t_3^5, t_4^5) = [\frac{3T}{5}, \frac{4T}{5})$, we put $\xi^5(t) = \xi(t_3^5) = \frac{3T}{5}$ and if $t \in [t_4^5, t_5^5) = [\frac{4T}{5}, \frac{5T}{5})$, we put $\xi^5(t) = \xi(t_4^5) = \frac{4T}{5}$. Finally, if $t \geq T$, we put $\xi^5(t) = \xi(t_5^5) = 0$.

We can show that $\xi^n(t)$ belongs to M^2_{step} and

$$\mathbb{E} \int_0^T |\xi^n(t) - \xi(t)|^2 \, dt \to 0 \text{ as } n \to \infty.$$

Second method. (Optional reading) Take for simplicity $n = 5$. Then

$$
\begin{aligned}
\xi^5(t_1^5) &= \frac{1}{t_1^5 - t_0^5}\int_{t_0^5}^{t_1^5} \xi(s)\,ds = \frac{n}{T}\int_{t_0^5}^{t_1^5} \xi(s)\,ds,\\
\xi^5(t_2^5) &= \frac{1}{t_2^5 - t_1^5}\int_{t_1^5}^{t_2^5} \xi(s)\,ds = \frac{n}{T}\int_{t_1^5}^{t_2^5} \xi(s)\,ds\\
\xi^5(t_3^5) &= \frac{1}{t_3^5 - t_2^5}\int_{t_2^5}^{t_3^5} \xi(s)\,ds = \frac{n}{T}\int_{t_2^5}^{t_3^5} \xi(s)\,ds\\
\xi^5(t_4^5) &= \frac{1}{t_5^5 - t_3^5}\int_{t_3^5}^{t_4^5} \xi(s)\,ds = \frac{n}{T}\int_{t_3^5}^{t_4^5} \xi(s)\,ds
\end{aligned}
$$

Since $\xi(t) = t$, $t \geq 0$, we thus define

$$
\begin{aligned}
\xi^5(t) &= 0, \quad \text{if} \quad t \in [0\frac{T}{5}, 1\frac{T}{5})\\
\xi^5(t) &= \xi(\frac{T}{5}) = \frac{T}{5}\int_0^{\frac{T}{5}} \xi_t\,dt = \frac{5}{T}\int_0^{\frac{T}{5}} t\,dt = \frac{5}{T}\frac{1}{2}(\frac{T}{5})^2 = \frac{1}{2}\frac{T}{5}, \quad \text{if} \quad t \in [\frac{T}{5}, 2\frac{T}{5})\\
\xi^5(t) &= \xi(\frac{2T}{5}) = \frac{5}{T}\int_{\frac{T}{5}}^{\frac{2T}{5}} \xi_t\,dt = \frac{5}{T}\int_{\frac{T}{5}}^{\frac{2T}{5}} t\,dt = \frac{5}{T}\frac{1}{2}[(\frac{2T}{5})^2 - (\frac{T}{5})^2]\\
&= \frac{5}{T}\frac{1}{2}(\frac{2T}{5} - \frac{T}{5})(\frac{2T}{5} + \frac{T}{5}) = \frac{1}{2}\frac{3T}{5} \quad \text{if} \quad t \in [2\frac{T}{5}, 3\frac{T}{5})\\
\xi^5(t) &= \xi(\frac{3T}{5}) = \frac{5}{T}\int_{\frac{2T}{5}}^{\frac{3T}{5}} \xi_t dt = \frac{5}{T}\int_{\frac{2T}{5}}^{\frac{3T}{5}} tdt = \frac{5}{T}\frac{1}{2}[(\frac{3T}{5})^2 - (\frac{2T}{5})^2]\\
&= \frac{5}{T}\frac{1}{2}(\frac{3T}{5} - \frac{2T}{5})(\frac{3T}{5} + \frac{2T}{5}) = \frac{1}{2}\frac{5T}{5}, \quad \text{if} \quad t \in [3\frac{T}{5}, 4\frac{T}{5})\\
\xi^5(t) &= \xi(\frac{4T}{5}) = \frac{5}{T}\int_{\frac{3T}{5}}^{\frac{4T}{5}} \xi_t\,dt = \frac{5}{T}\int_{\frac{3T}{5}}^{\frac{4T}{5}} t\,dt = \frac{5}{T}\frac{1}{2}[(\frac{4T}{5})^2 - (\frac{3T}{5})^2]\\
&= \frac{5}{T}\frac{1}{2}(\frac{4T}{5} - \frac{3T}{5})(\frac{4T}{5} + \frac{3T}{5}) = \frac{1}{2}\frac{7T}{5}, \quad \text{if} \quad t \in [4\frac{T}{5}, 5\frac{T}{5})
\end{aligned}
$$

From the construction it follows that our new process is adapted and in fact it belongs to M^2_{step} and moreover $\mathbb{E}\int_0^T |\xi^n(t) - \xi(t)|^2 \mathrm{dt} \to 0$ as n $\to \infty$.

Example 1.18. *In our second example we show how the above construction works in the special case of Brownian motion $W = (W_t)_{t \geq 0}$. We fix $T > 0$ and define a* *process $\xi = (\xi_t)_{t\geq 0}$ by*

$$\xi_t = \begin{cases} W_t, & \text{if } t \in [0, T), \\ 0, & \text{if } t \in [T, \infty). \end{cases}$$

We want to find $I(\xi)$.

Step 1 *Fix a natural number n. We divide interval $[0, T]$ into n smaller intervals of equal length. Define*

$$t_i^n = \frac{i}{n}T, \quad i = 0, 1, \cdots, n.$$

Then

$$0 = t_0^n < t_1^n = \frac{T}{N} < t_2^n = \frac{2T}{N} < \ldots < t_{n-1}^n = \frac{(n-1)T}{N} < t_n^n = T.$$

The length of each interval $[t_i^n, t_{i+1}^n]$ in the partition is $\frac{T}{n}$.

Step 2 *We find a sequence $\xi = (\xi^n(t))_{t\geq 0}$ of processes belonging to the class $M^2_{step}(0, \infty)$ such that $\xi^n(t) = 0$, if $t \geq T$ and*

$$\mathbb{E}\int_0^T |\xi^n(t) - \xi(t)|^2\, dt \to 0 \text{ as } n \to \infty.$$

Define

$$\xi^n(t) = \begin{cases} \xi(t_i^n), & \text{if } \ t \in [t_i^n, t_{i+1}^n),\ i = 0, 1, \cdots, n-1 \\ 0, & \text{if } \ t \geq T \end{cases}.$$

Observe that ξ^n belongs to M^2_{step}. This is true since the process ξ^n is constant on each interval $[t_i^n, t_{i+1}^n)$ and zero on the interval $[T, \infty)$.

Step 3 *Check whether, as $n \to \infty$,*

$$\mathbb{E}\int_0^T |\xi^n(t) - \xi(t)|^2\, dt \to 0?$$

We have

$$
\begin{aligned}
&\mathbb{E}\int_0^T |\xi^n(t)-\xi(t)|^2\,dt \\
&= \mathbb{E}\sum_{i=0}^{n-1}\int_{t_i^n}^{t_{i+1}^n} |\xi(t_i^n)-\xi(t)|^2\,dt \;=\; \mathbb{E}\sum_{i=0}^{n-1}\int_{t_i^n}^{t_{i+1}^n} |W(t_i^n)-W(t)|^2\,dt \\
&= \sum_{i=0}^{n-1}\int_{t_i^n}^{t_{i+1}^n} \mathbb{E}|W(t_i^n)-W(t)|^2\,dt \;=\; \sum_{i=0}^{n-1}\int_{t_i^n}^{t_{i+1}^n}(t-t_i^n)\,dt \\
&= \frac{1}{2}\sum_{i=0}^{n-1}(t_{i+1}-t_i^n)^2 \;=\; \frac{1}{2}\sum_{i=0}^{n-1}\left(\frac{T}{n}\right)^2 = \frac{T^2}{2n} \to 0 \ \textit{as } n\to\infty.
\end{aligned}
$$

Step 4 *Now we will find $I(\xi^n)$. By definition we have*

$$
\begin{aligned}
I(\xi^n) &= \sum_{i=0}^{n-1}\xi(t_i^n)(W(t_{i+1}^n)-W(t_i^n)) \\
&= \sum_{i=0}^{n-1}W(t_i^n)(W(t_{i+1}^n)-W(t_i^n))
\end{aligned}
$$

Step 5 *Find a limit of the sequence or random variables $I(\xi^n)$, $n=1,2,\cdots,\infty$.*

We have

$$
\begin{aligned}
I(\xi^n) &= \sum_{i=0}^{n-1}W(t_i^n)\big[W(t_{i+1}^n)-W(t_i^n)\big] \\
&= \frac{1}{2}\sum_{i=0}^{n-1}\big[W(t_{i+1}^n)^2-W(t_i^n)^2-(W(t_{i+1}^n)-W(t_i^n))^2\big] \\
&= \frac{1}{2}\sum_{i=0}^{n-1}\big[W(t_{i+1}^n)^2-W(t_i^n)^2\big]-\frac{1}{2}\sum_{i=0}^{n-1}\big[W(t_{i+1}^n)-W(t_i^n)\big]^2 \\
&= \frac{1}{2}(W(t_1^n)^2-W(t_0^n)^2+W(t_2^n)^2-W(t_1^n)^2+W(t_3^n)^2-W(t_2^n)^2 \\
&+ \cdots+W(t_n^n)^2-W(t_{n-1}^n)^2)-\frac{1}{2}\sum_{i=0}^{n-1}\big[W(t_{i+1}^n)-W(t_i^n)\big]^2 \\
&= \frac{1}{2}\big[W(T)^2-W(0)^2\big]-\frac{1}{2}\sum_{i=0}^{n-1}\big[W(t_{i+1}^n)-W(t_i^n)\big]^2 \\
&= \frac{1}{2}W(T)^2-\frac{1}{2}\sum_{i=0}^{n-1}\big[W(t_{i+1}^n)-W(t_i^n)\big]^2
\end{aligned}
$$

We know from the MMF module (quadratic variation of Brownian motion) that

$$\sum_{i=0}^{n-1}(W(t^n_{i+1}) - W(t^n_i))^2$$

converges in the mean square to a constant r.v. T

Conclusion *The limit of $I(\xi^n)$ is equal to $\frac{1}{2}W(T)^2 - \frac{1}{2}T$.*

So we get

$$I(\xi) = \frac{1}{2}W(T)^2 - \frac{1}{2}T.$$

We define two spaces $M^2(0,\infty)$ and $M^2_{loc}(0,\infty)$ of stochastic processes that admit (as we will see) approximations by random step processes. As a consequence these spaces will provide rich classes of examples of integrands for the Itô integral against a Brownian motion,

Definition 1.19. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(\mathcal{F}_t)_{t\geq 0}$ a filtration.*

Part I *We say that a process $(\xi_t)_{t\geq 0}$ belongs to class $M^2(0,\infty)$ if and only if*

(i) $(\xi_t)_{t\geq 0}$ *is adapted to the filtration* $(\mathcal{F}_t)_{t\geq 0}$.

(ii) the trajectories of $(\xi_t)_{t\geq 0}$ *are left continuous or right continuous a.s.*

(iii) $\mathbb{E}\int_0^\infty |\xi_t|^2 dt < \infty$.

Part II *We say that a process $(\xi_t)_{t\geq 0}$ belongs to class $M^2_{loc}(0,\infty)$ if and only if*

(i) $(\xi_t)_{t\geq 0}$ *is adapted to the filtration* $(\mathcal{F}_t)_{t\geq 0}$.

(ii) the trajectories of $(\xi_t)_{t\geq 0}$ *are left continuous or right continuous a.s.*

(iv) For every $T > 0, \mathbb{E}\int_0^T |\xi_t|^2 dt < \infty$.

Example 1.20. ***(a)*** $M^2_{step}(0,\infty) \subset M^2(0,\infty)$.

Proof. Let $\xi \in M^2_{step}(0,\infty)$.There exists a partition $t_0 = 0 < t_1 < \cdots < t_n < \infty$.

$$\xi(t) = \begin{cases} \xi(t_i), & t \in [t_i, t_{i+1}),\ i = 0, 1, \ldots n-1 \\ 0, & t \geq t_n \end{cases}$$

$\xi(t)$ is $(\mathcal{F}_t)_{t\geq 0}$ adapted, $\mathbb{E}|\xi_t|^2 < \infty$, for all $t \geq 0$.

The trajectories of ξ are continuous with the exception of finite number of times: $t_1, t_2, \dots, t_n$. Note that the trajectories are right-continuous at these points. In order to verify (iii) of Definition 1.19, it is enough to show that $\int_0^\infty \mathbb{E}|\xi_t|^2 < \infty$. But

$$\begin{aligned}\int_0^\infty \mathbb{E}|\xi_t|^2 &= \int_0^{t_1} \mathbb{E}|\xi_t|^2\,dt + \int_{t_1}^{t_2} \mathbb{E}|\xi_t|^2 dt + \cdots + \int_{t_{n-1}}^{t_n} \mathbb{E}|\xi_t|^2 dt + \int_{t_n}^\infty \mathbb{E}|\xi_t|^2 dt \\ &= \mathbb{E}|\xi_{t_0}|^2(t_1 - t_0) + \mathbb{E}|\xi_{t_1}|^2(t_2 - t_1) + \cdots + \mathbb{E}|\xi_{t_{n-1}}|^2(t_n - t_{n-1}) + 0 \\ &< \infty\end{aligned}$$

$\square$

(b) *Let ξ_t, $t \geq 0$ be the process defined by*

$$\xi_t = W_t,\ t \geq 0.$$

Is $\xi \in M^2_{loc}(0,\infty)$?

Proof. Clearly this process is adapted with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$ and its trajectories are continuous a.s.. We only need to verify that
$\int_0^T \mathbb{E}|\xi_t|^2 dt < \infty$ for $T > 0$.

We have

$$\int_0^T \mathbb{E}|\xi_t|^2 dt = \int_0^T \mathbb{E}(W_t^2)\,dt = \int_0^T t dt = \frac{T^2}{2} < \infty,$$

and $\xi \in M^2_{loc}(0,\infty)$. However,

$$\int_0^\infty \mathbb{E}(|\xi_t|^2)\,dt = \int_0^\infty t dt = \infty$$

and it follows that $\xi \notin M^2(0,\infty)$. $\square$

(c) $\xi_t = W_t^2$, *if* $t \geq 0$.

(d) $\xi_t = e^{W_t}$, *if* $t \geq 0$.

(e) $\xi_t = e^{W_t^2}$, *if* $t \geq 0$.

Checking that the process $(\xi_t)_{t\geq 0}$ *in (c) and (d) belongs to* $M^2_{loc}(0,\infty)$ *is left as an exercise. For (e) we note that adaptedness and sample path continuity are easy to check. However since*

$$\mathbb{E}\left(\exp\left(|W_t|^2\right)\right) = \infty$$

if $t \geq 1/2$ *it follows that*

$$\int_0^T \mathbb{E}(|\xi_t|^2)\, dt = \infty$$

if $T > 1/2$. *Hence, this example does not belong to* $M^2_{loc}(0,\infty)$.

The following Proposition shows that any process in $M^2(0,\infty)$ can be approximated by random step processes.

Proposition 1.21. *If* $\xi = (\xi_t)_{t\geq 0}$ *belongs to* $M^2(0,\infty)$ *we can find a sequence* ξ^n, $n = 1, 2, \cdots$, *such that*

$$\xi^n = (\xi^n_t)_{t\geq 0} \in M^2_{\text{step}}(0,\infty)$$

and

$$\mathbb{E}\int_0^\infty |\xi_t - \xi_t^n|^2\, dt \to 0, \text{ as } n \to \infty.$$

Proposition 1.22. *Let* $\xi = (\xi_t)_{t\geq 0}$ *be a stochastic process. Suppose the sequence* $\xi^n = (\xi^n_t)_{t\geq 0}$, $n = 1, 2, \cdots$, *belongs to* $M^2_{\text{step}}(0,\infty)$ *and satisfies*

$$\mathbb{E}\int_0^\infty |\xi_t - \xi_t^n|^2\, dt \to 0, \text{ as } n \to \infty. \tag{$*$}$$

Then the sequence of random variables $I(\xi^n)$, $n = 1, 2, \cdots$, *has a unique limit, i.e. there exists a unique random variable denoted by* $I(\xi)$ *such that* $\mathbb{E}(|I(\xi)|^2) < \infty$ *and*

$$\mathbb{E}|I(\xi^n) - I(\xi)|^2 \to 0, \text{ as } n \to \infty. \tag{$**$}$$

Furthermore, the random variable $I(\xi)$ *doesn't depend on the choice of the sequence* ξ^n *as long as the condition* $(*)$ *is satisfied.*

Remark By Proposition 1.21 if $\xi \in M^2(0,\infty)$ there exits a sequence $\xi^n = (\xi^n_t)_{t\geq 0} \in M^2_{\text{step}}(0,\infty)$, $n = 1,2,\cdots$, satisfying condition $(*)$. We call the limit $I(\xi)$ of the sequence $I(\xi^n)$ (taken in the mean square sense) the Itô integral of the process ξ (against the Brownian motion W) and will also denote it by $\int_0^\infty \xi_s dW_s$.

A proof of Proposition 1.22 can be traced back to our discussion of the construction of the Itô integral on page 14 where we used the inequality $(a+b)^2 \leq 2(a^2+b^2)$. We now give a variant of the proof that uses the Minkowski inequality instead.

Proof of Proposition 1.22. Using the Minkowski inequality we have

$$\begin{aligned}\Big(\mathbb{E}\int_0^T |\xi^n(t)-\xi^k(t)|^2\,dt\Big)^{1/2} &= \Big(\mathbb{E}\int_0^T |\xi^n(t)-\xi(t)+\xi(t)-\xi^k(t)|^2\,dt\Big)^{1/2}\\ \leq \Big(\mathbb{E}\int_0^T |\xi^n(t)-\xi(t)|^2\Big)^{1/2} &+ \Big(\mathbb{E}\int_0^T |\xi(t)-\xi^k(t)|^2]\,dt\Big)^{1/2}\end{aligned}$$

Since by Assumption $(*)$ $\lim_{n\to\infty}\mathbb{E}\int_0^T |\xi^n(t)-\xi(t)|^2\,dt = 0$ and $\lim_{k\to\infty}\mathbb{E}\int_0^T |\xi^k(t)-\xi(t)|^2]\,dt = 0$, we infer that

$$\lim_{k,n\to\infty}\mathbb{E}\int_0^T |\xi^n(t)-\xi^k(t)|^2\,dt = 0. \tag{1.7}$$

On the other hand, both ξ^n and ξ^k are elements of M^2_{step} and therefore we can define $I(\xi^n)$ and $I(\xi^k)$. Moreover, since M^2_{step} is a vector space, $\xi^n-\xi^k \in M^2_{step}$ and

$$I(\xi^n-\xi^k) = I(\xi^n) - I(\xi^k)$$

we have by the Itô isometry for step processes (Proposition 1.12)

$$\mathbb{E}|I(\xi^n)-I(\xi^k)|^2 = \mathbb{E}\int_0^T |\xi^n(t)-\xi^k(t)|^2\,dt$$

Thus, in view of (1.7), we conclude that

$$\lim_{k,n\to\infty}\mathbb{E}|\ I(\xi^n)-I(\xi^k)|^2 = 0,$$

i.e. that the sequence of random variable $I(\xi^n)$, $n\in\mathbb{N}$ is a Cauchy sequence in the Hilbert space of square integrable random variables $L^2(\Omega,\mathbb{P})$. By **completeness** of the space $L^2(\Omega, \mathbb{P})$ every Cauchy sequence converges. Hence, there exists a random variable denoted by I(ξ) such that $\mathbb{E}|I(\xi)|^2 < \infty$ and

$$\mathbb{E}|I(\xi^n) - I(\xi)|^2 \to 0 \text{ as } n \to \infty. \tag{1.8}$$

This random variable $I(\xi)$ does not depend on the approximating sequence. Indeed, if $\tilde{\xi}_n$ is another sequence of random step processes satisfying $(*)$, then by the proof above we can find a random variable $\tilde{I}(\xi)$ such that

$$\lim_{n\to\infty} \mathbb{E}|I(\tilde{\xi}^n) - \tilde{I}(\xi)|^2 = 0. \tag{1.9}$$

Hence, $I(\xi^n) \to I(\xi)$ and $I(\tilde{\xi}^n) \to \tilde{I}(\xi)$. On the other hand, the sequence

$$\xi^1, \tilde{\xi}^1, \xi^2, \tilde{\xi}^2, \xi^3, \tilde{\xi}^3, \cdots$$

also satisfies the condition $(*)$ and therefore arguing as before sequence

$$I(\xi^1), I(\tilde{\xi}^1), I(\xi^2), I(\tilde{\xi}^2), I(\xi^3), I(\tilde{\xi}^3), \cdots$$

is convergent. However, the last sequence contains two subsequences

$$I(\xi^1), I(\xi^2), I(\xi^3), \cdots \quad \text{and} \quad I(\tilde{\xi}^1), I(\tilde{\xi}^2), I(\tilde{\xi}^3), \cdots$$

which are both convergent, by (1.8) and respectively by (1.9), to $I(\xi)$ and $\tilde{I}(\xi)$. By the uniqueness of the limit we deduce that $I(\xi) = \tilde{I}(\xi)$, which completes the proof of the last claim. $\square$

The following proposition establishes some basic properties of the stochastic integral.

Proposition 1.23. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(W_t)_{t\geq 0}$ a Brownian Motion with respect to a filtration $(\mathcal{F}_t)_{t\geq 0}$. Assume that $\xi, \eta \in M^2(0, \infty)$ and $\alpha, \beta \in$*

$\mathbb{R}$. *Then*

$$\begin{aligned}
\mathbb{E}(I(\xi)) &= 0,\\
\mathbb{E}(|I(\xi)|^2) &= \mathbb{E}\int_0^\infty |\xi_t|^2\,dt < \infty \quad \textit{(Itô isometry)}\\
I(\alpha\xi+\beta\eta) &= \alpha I(\xi)+\beta I(\eta).
\end{aligned}$$

Proof. We prove that $\mathbb{E}(I(\xi)) = 0$, the proof of the other claims is similar.

As $\xi \in M^2(0,\infty)$, by Proposition 1.21 there exists a sequence $(\xi_t^n)_{n\in\mathbb{N}}$ of random step processes such that ξ^n converge to ξ in mean square, that is

$$\lim_{n\to\infty}\int_0^\infty \mathbb{E}(\xi_t-\xi_t^n)^2 dt = 0.$$

Moreover, by the construction of the Itô integral we have that $I(\xi_n)$ converge to $I(\xi)$ in mean square:

$$\lim_{n\to\infty}\mathbb{E}(I(\xi)-I(\xi^n))^2 = 0. \tag{1.10}$$

By Jensen's inequality (seen in the MMF module) this implies that for any random variable X,

$$(\mathbb{E}(X))^2 \le \mathbb{E}(X^2). \tag{1.11}$$

Applying inequality (1.11) to the random variable $X = I(\xi) - I(\xi^n)$, then we find

$$(\mathbb{E}[I(\xi)-I(\xi^n)])^2 \le \mathbb{E}[(I(\xi)-I(\xi^n))^2]. \tag{1.12}$$

By (1.10) we have that the right hand side of (1.12) converge to 0. Then, the left hand side of (1.12) also converges to 0, so

$$\lim_{n\to\infty}\mathbb{E}[I(\xi)-I(\xi_n)] = 0. \tag{1.13}$$

But, for any random step process ξ^n, $\mathbb{E}(I(\xi^n)) = 0$ for all $n \in \mathbb{N}$. Thus, for any $n \in \mathbb{N}$

$$\mathbb{E}[I(\xi)-I(\xi^n)] = \mathbb{E}[I(\xi)] - \mathbb{E}[I(\xi^n)] = \mathbb{E}[I(\xi)].$$

We deduce with (1.13)

$$\lim_{n\to\infty} \mathbb{E}[I(\xi)] = \mathbb{E}[I(\xi)] = 0.$$

$\square$

The following definition will allow us to define the stochastic integral over an interval $[0,T]$, i.e. define $\int_0^T W_t dW_t$.

Definition 1.24. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(W_t)_{t\geq 0}$ be a Brownian motion w.r.t. the filtration $(\mathcal{F}_t)_{t\geq 0}$. For a process $\xi \in M^2_{\mathrm{loc}}(0,\infty)$ and $t \geq 0$, we define*

$$\int_0^t \xi(s)\, dW(s) := I(1_{[0,t)}\xi),$$

where $1_{[0,t)}$ is the indicator function of the interval $[0,t)$, i.e.

$$1_{[0,t)}(s) = \begin{cases} 1, & \text{if } s \in [0,t), \\ 0, & \text{if } s \notin [0,t), \end{cases}.$$

Remark

$$[1_{[0,t)}\xi](s) = \begin{cases} \xi(s), & \text{if } s \in [0,t), \\ 0, & \text{if } s \notin [0,t), \end{cases}.$$

Example 1.25. *Recall that $W \notin M^2(0,\infty)$, but $W \in M^2_{loc}(0,\infty)$.*

Hence, by Definition 1.24 we have

$$\int_0^T W_t dW_t = I(1_{[0,T)}W) = I(\eta),$$

where

$$\eta(t) = \begin{cases} W_t, & \text{if } t \in [0,T), \\ 0, & \text{if } t \geq T, \end{cases}.$$

Similarly, we can define $\int_0^T W_t^2 dW_t$, $\int_0^T W_t^3 dW_t$, ..., $\int_0^T e^{\lambda W_t}\, dW_t$.

Recall that, $\int_0^T W_t dW_t = \frac{1}{2}(W_T^2 - T)$. A somewhat more involved, direct calculation (exercise!) allows us to find the exact value for $\int_0^T W_t^2 dW_t$. Determining more complex integrands such as $\int_0^T W_t^3 dW_t$ by the same method becomes increasingly intractible . To resolve this difficulty we will later have to prove a theorem analogous to the Fundamental Theorem of Calculus for the Riemann or Lebesgue integral. Before we go any further we we gather a number of fundamental properties of the Ito integral.

Theorem 1.26. *(Fundamental Properties of Itô integral)*
Assume that $(\Omega, \mathcal{F}, \mathbb{P})$*is a probability space,* $(\mathcal{F})_{t\geq 0}$ *is a filtration and* $(W(t))_{t\geq 0}$ *is a Brownian motion with respect to that filtration. Let* $\xi \in M^2_{loc}(0,\infty)$*. Let* $\int_0^t \xi(s)\, dW(s)$ *denote the Itô Intergral of* ξ*. Then*

(i) $\mathbb{E}|\int_0^t \xi(s)\, dW(s)|^2 = \mathbb{E}\int_0^t |\xi(s)|^2\, ds$ *(Itô isometry)*

(ii) $\mathbb{E}[\int_0^t \xi(s)\, dW(s)] = 0$

(iii) *If* $t > \tau \geq 0$ *then* $\mathbb{E}(\int_0^t \xi(s)\, dW(s) \big| \mathcal{F}_\tau) = \int_0^\tau \xi(s)\, dW(s)$ *(Martingale property)*

(iv) *If* $\eta \in M^2_{loc}(0,\infty)$*,*$\alpha, \beta \in \mathbb{R}$ *then*
$$\int_0^t (\alpha\xi(s) + \beta\eta(s))\, dW(s) = \alpha \int_0^t \xi(s)\, dW(s) + \beta \int_0^t \eta(s)\, dW(s)$$
a.s.

(v) *If* $\eta \in M^2_{loc}(0,\infty)$*, then*
$$\mathbb{E}\left[\int_0^t \xi(s)\, dW(s) \int_0^t \eta(s)\, dW(s)\right] = \int_0^t \mathbb{E}(\xi(s)\eta(s))\, ds = \mathbb{E}\int_0^t \xi(s)\eta(s)\, ds$$

Proof. To prove the Itô isometry we observe that we can approximate $1_{[0,t)}\xi$ by a sequence of processes $\xi^n \in M^2_{step}(0,\infty)$ in the sense that
$$\lim_{n\to\infty} \int_0^\infty \mathbb{E}(\xi_t - \xi_t^n)^2 dt = 0.$$
The claim now follows from the Itô isometry for random step processes and taking the limit. For properties (ii)-(iv) we similarly firstly prove all statements for $\xi, \eta \in M^2_{step}(0,\infty)$.

In the second step, we take any $\xi, \eta \in M^2_{loc}(0,\infty)$ and approximate $1_{[0,t)}\xi$, $1_{[0,t)}\eta$ by a sequence $(\xi_n)(\eta_n)$ of process from M^2_{step}. The detailed proofs are left as an exercise.

(v) This property is in fact equivalent to the property (i). We show here how to deduce (v) from (i).

We prove the equality

$$\mathbb{E}\left[\int_0^t \xi(s)\,dW(s)\int_0^t \eta(s)\,dW(s)\right] = \int_0^t \mathbb{E}(\xi(s)\eta(s))\,ds \tag{1.14}$$

the claim then follows by swapping the integral and $\mathbb{E}$ on the right hand side using Fubini's theorem.

Let $\xi, \eta \in M^2_{loc}$. Then $\xi + \eta \in M^2_{loc}$ and applying the Ito Isometry to this process we have

$$\mathbb{E}\left(\int_0^t (\xi(s)+\eta(s))\,dW(s)\right)^2 = \int_0^t \mathbb{E}\left(\xi(s)+\eta(s)\right)^2\,ds. \tag{1.15}$$

Using the binomial formula $(a+b)^2 = a^2 + 2ab + b^2$ and the fact that the integral and expectation are both linear (that is $\int(x+y) = \int x + \int y$ and $\mathbb{E}(x+y) = \mathbb{E}x + \mathbb{E}y$) the left hand side of (1.15) equals

$$\begin{aligned}\mathbb{E}\left(\int_0^t (\xi(s)+\eta(s))\,dW(s)\right)^2 &= \mathbb{E}\left(\int_0^t \xi(s)\,dW(s)\right)^2 + \mathbb{E}\left(\int_0^t \eta(s)\,dW(s)\right)^2 \\ &\quad + 2\mathbb{E}\left[\int_0^t \xi(s)\,dW(s)\int_0^t \eta(s)\,dW(s)\right]\end{aligned} \tag{1.16}$$

For exactly the same reasons the right hand side of (1.15) equals

$$\int_0^t \mathbb{E}\left(\xi(s)+\eta(s)\right)^2\,ds = \int_0^t \mathbb{E}\xi(s)^2\,ds + \int_0^t \mathbb{E}\eta(s)^2\,ds + 2\int_0^t \mathbb{E}(\xi(s)\eta(s))\,ds. \tag{1.17}$$

Hence,

$$\begin{aligned}\mathbb{E}\left(\int_0^t \xi(s)\,dW(s)\right)^2 + \mathbb{E}\left(\int_0^t \eta(s)\,dW(s)\right)^2 + 2\mathbb{E}\left[\int_0^t \xi(s)\,dW(s)\int_0^t \eta(s)\,dW(s)\right] \\ = \int_0^t \mathbb{E}\xi(s)^2\,ds + \int_0^t \mathbb{E}\eta(s)^2\,ds + 2\int_0^t \mathbb{E}(\xi(s)\eta(s))\,ds.\end{aligned} \tag{1.18}$$

Applying, again, the Ito Isometry, this time to the process ξ, we find

$$\mathbb{E}\left(\int_0^t \xi(s)\,dW(s)\right)^2 = \int_0^t \mathbb{E}\xi(s)^2\,ds.$$

Similiarly we have for the process η

$$\mathbb{E}\left(\int_0^t \eta(s)\, dW(s)\right)^2 = \int_0^t \mathbb{E}\eta(s)^2\, ds.$$

The corresponding terms in (1.18) cancel out and we are left with

$$2\mathbb{E}\left[\int_0^t \xi(s)\, dW(s) \int_0^t \eta(s)\, dW(s)\right] = 2\int_0^t \mathbb{E}(\xi(s)\eta(s))\, ds.$$

$\square$

At this point we recall from MMF the notion of a martingale.

Definition . *A stochastic process $M = (M_t)_{t\geq 0}$ is a martingale with respect to a filtration $(\mathcal{F}_t)_{t\geq 0}$ if*

(1) M is adapted w.r.t $(\mathcal{F}_t)_{t\geq 0}$

(2) $\mathbb{E}(|M_t|) < \infty$ for all $t \geq 0$

(3) $\mathbb{E}(M_t | \mathcal{F}_s) = M_s$ a.s. for all $0 \leq s < t$.

It follows from the preceding proposition that for $\xi \in M^2_{loc}(0,\infty)$ the stochastic integral $\int_0^t \xi(s)\, dW(s)$ is a martingale.

The following theorem allows us to assume that the stochastic integral has a.s. continuous trajectories.

Theorem 1.27. *Suppose $\xi \in M^2_{loc}$ and*

$$\eta_t = \int_0^t \xi_s dW_s, \quad t \geq 0.$$

Then there exists an adapted modification $\widetilde{\eta}$ of η with a.s. continuous sample paths. This modification is unique (up to equality a.s.).

From now on we always identify

$$\int_0^t \xi_s dW_s$$

with this particular modification and, hence, may assume the stochastic integral has a.s. continuous sample paths[2].

In practice, many if not most Riemann integrals are calculated using the Fundamental Theorem of Calculus (FTC) and the chain and product rule of differentiation. Suppose $f : [a, b] \to \mathbb{R}$ is continuously differentiable then

$$\begin{aligned} \int_a^b f(t) df(t) &= \int_a^b f(t) f'(t)\, dt \\ &= \frac{1}{2} \int_a^b \frac{d}{dt} f^2(t)\ dt \\ &= \frac{1}{2} f^2(b) - \frac{1}{2} f^2(a) \end{aligned}$$

Comparing this with

$$\int_0^T W_t\, dW_t = \frac{1}{2} \left(W(T)^2 - T \right)$$

we see that the Itô integral does not satisfy the FTC (there is an additional term $\frac{1}{2}T$ one the right hand side which comes from the quadratic variation of the Brownian motion). Note there is also no theory of differentiation with a chain rule in this context.

2. Itô Lemma

In the whole section,we assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space. $(\mathcal{F})_{t\geq 0}$ is a filtration of σ-field and $(W(t))_{t\geq 0}$ is a Brownian motion with respect to $(\mathcal{F})_{t\geq 0}$.

From

$$\frac{1}{2} W_t^2 = \frac{1}{2} t + \int_0^t W_s dW_s$$

we see that the image of the Itô integral $W_t = \int_0^t dW_s$ under the map $g(x) = \frac{1}{2}x^2$ is not an Itô integral but a combination of a dW_s and ds integral. Rewriting our

[2]Note that we say that η is a modification of $\widetilde{\eta}$ if $P(\eta_t = \widetilde{\eta}_t) = 1$ for all t

example
$$\frac{1}{2}\left(\int_0^t dW_s\right)^2 = \frac{1}{2}t + \int_0^t W_s dW_s.$$
It turns out that if we introduce the Itô process as sums of dW_s and ds integrals we will later see that this class of processes is stable under smooth maps, i.e. the image of an Ito process under a smooth map is another Ito process.

Definition 2.1. *An Itô process with characteristics a and b is a stochastic process $(\xi(t))_{t\geq 0}$ such that*

(1) $(\xi(t))_{t\geq 0}$ is adapted to $(\mathcal{F})_{t\geq 0}$,

(2) $(\xi(t))_{t\geq 0}$ has continuous (or piecewise continuous) trajectories a.s.

(3)
$$\xi(t) = \xi(0) + \int_0^t a(s)\, ds + \int_0^t b(s)\, dW(s),\ t \geq 0,$$
where a and b are adapted processes with (piecewise) continuous trajectories such that for any $T > 0$
$$\mathbb{E}\int_0^T |a(s)| ds < \infty, \mathbb{E}\int_0^T |b(s)|^2\, ds < \infty$$

Remark . The condition on b means that $b \in M^2_{loc}(0,\infty)$.

Definition . *If process a satisfies the conditions in Definition 2.1 we say that $a \in M^1_{loc}(0,\infty)$.*

Example 2.2. *(i) $\xi(t) = \xi(0)$ is an Itô process with characteristics $a = 0, b = 0$ if $\xi(0)$ is $\mathcal{F}_0$-measurable.*

In particular, $\xi(t) = 5, t \geq 0$ is an Ito process. At the same time, $\xi(t) = W(1), t \geq 0$ is not an Ito process because $\xi(0) = W(1)$ is not $\mathcal{F}_0$-measurable.

(ii) A process of the form $\xi(t) = \int_0^t a(s)\, ds$ is an Itô process with characteristics a and $b = 0$ provided that $a \in M^1_{loc}(0,\infty)$.

(a) $\xi(t) = \int_0^t s ds = \frac{t^2}{2}$. Notice that $a \in M^1_{loc}$ because obviously a is adapted and has

continuous trajectories and $\mathbb{E}\int_0 |a(s)|\,ds = \mathbb{E}\frac{t^2}{2} = \frac{t^2}{2} < \infty$.

(b) $\xi(t) = \int_0^t \sin(W(s))\,ds$. *Does* $a \in M^1_{loc}(0,\infty)$*?*

Since $W(s)$ *is* $\mathcal{F}_s$*-measurable and sine is a continuous function,* $(a(s))_{s\geq 0}$ *is an adapted process. Moreover by definition of Brownian motion the trajectories* $s \mapsto W(s)$ *are continuous functions a.s.. Hence, using the continuity of the sine function the trajectories of* a *are also continuous a.s. .*

In order to conclude that $a \in M^1_{loc}(0,\infty)$ *we need to check that if*

$$\begin{aligned}
\mathbb{E}\int_0^T |a(s)|\,ds &< \infty \\
\mathbb{E}\int_0^T |a(t)|dt &= \mathbb{E}\int_0^T |\sin(W(t))|dt \leq \mathbb{E}\int_0^T dt = T < \infty
\end{aligned}$$

(c) $\xi(t) = \int_0^t W(s)\,ds$. *As in (b) we can verify that* $a \in M^1_{loc}$ *and hence* ξ *is an Itô progress.*

(iii) A process of the form $\xi(t) = \int_0^t b(s)\,dW(s), t \geq 0$ *is an Itô process with characteristics* $a = 0$ *and* b *provided* $b \in M^2_{loc}(0,\infty)$.

(a) $\xi(t) = \int_0^t s dW(s)$

(b) $\xi(t) = \int_0^t \cos(W(s))\,dW(s)$

(c) $\xi(t) = \int_0^t W(s)\,dW(s)$

In each case we need to verify that the processes b *defined by*

(a) $b(s) = s, s \geq 0$

(b) $b(s) = \cos(W(s)), s \geq 0$

(c) $b(s) = W(s), s \geq 0$

possesses the following properties:

(1) b *is adapted and has piecewise continuous trajectories*

(2) $\mathbb{E}\int_0^T |b(s)|^2\,ds < \infty$ *for any* $T > 0$.

We calculated earlier that

$$\xi(t) = \int_0^t W(s)\,dW(s) = \frac{1}{2}W(t)^2 - \frac{1}{2}t$$

in other words, process $\xi(t) = \frac{1}{2}W(t)^2 - \frac{t}{2}, t \geq 0$ is an Itô process.

It is natural to ask whether if we compose Brownian motion with other smooth functions such as $g(x) = x^3$ or $g(x) = e^x$ the resulting process is still an Itô process. The Ito Lemma is a cornerstone of the stochastic calculus and gives an affirmative answer to this question.

Theorem 2.3. *(Itô Lemma in simple form)*
Assume that $f(t,x)$ is in $C^{1,2}(\mathbb{R}^2)$, i.e. the derivatives $\frac{\partial f}{\partial t}, \frac{\partial f}{\partial x}, \frac{\partial^2 f}{\partial x^2}$ exist and are continuous. Define a process η by

$$\eta(t) = f(t, W(t)),\ t \geq 0.$$

Then $\eta(t)$, $t \geq 0$ is an Itô process and

$$\eta(t) = f(0,0) + \int_0^t [\frac{\partial f}{\partial s}(s, W(s)) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(s, W(s))]ds + \int_0^t \frac{\partial f}{\partial x}(s, W(s))\,dW(s), t \geq 0.$$

Remark 2.4. *Strictly speaking we require for the Itô lemma in order to make sense of the stochastic integral that $\frac{\partial f}{\partial x}(t, W(t)) \in M^2_{loc}(0,\infty)$. The notion of the Itô integral (and correspondingly Itô processes) can be generalised to integrands more general than $M^2_{loc}(0,\infty)$ and with this generalisation the Itô lemma takes the usual form we have stated above.*

Proof. The following is a very rough sketch of the main idea of the proof. Let $t_0 < t_1 \cdots < t_n$ be a partition of $[0,t]$. Then, using Taylor expansion (for a function that has two variables)

$$\begin{aligned} f(t, W_t) - f(0, W_0) &= \sum_{j=0}^{n-1} \left(f\left(t_{j+1}, W_{t_{j+1}}\right) - f\left(t_j, W_{t_j}\right) \right) \\ &= \sum_{j=0}^{n-1} \frac{\partial f}{\partial x} \left(t_j, W_{t_j}\right) \left(W_{t_{j+1}} - W_{t_j}\right) + \sum_{j=0}^{n-1} \frac{\partial f}{\partial t} \left(t_j, W_{t_j}\right) \left(t_{j+1} - t_j\right) \\ &\quad + \sum_{j=0}^{n-1} \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \left(t_j, W_{t_j}\right) \left(W_{t_{j+1}} - W_{t_j}\right)^2 + \sum_{j=0}^{n-1} \mathrm{Rem}\,(j, n). \end{aligned}$$

As we refine the partition and take limits[3] the first term in the sum converges to $\int_0^t \frac{\partial f}{\partial x}(s, W(s))\, dW(s)$,the second to $\int_0^t \frac{\partial f}{\partial s}(s, W(s)) ds$ and the third to $\int_0^t \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(s, W(s)) ds$. The reason that this last term contributes is that the quadratic variation of the Brownian motion W_t is non-zero, in fact equals t. The term $\left(W_{t_{j+1}} - W_{t_j}\right)^2$ is of order t and produces as we refine the partition and take limits a dt contribution (if W_t was smooth this term would vanish). The remainders are sufficiently small and can be shown to vanish in the limit. □

Example 2.5. *The following examples are all applications of the Itô Lemma.*

(1) Define a process $\eta(t) = \alpha t + \beta W_t + \gamma$, $t \geq 0$, where $\alpha, \beta, \gamma \in R$ are constants. Is $\eta(t)$, $t \geq 0$ an Itô process?

Let $f(t, x) = \alpha t + \beta x + \gamma$. Then, $\eta(t) = f(t, W_t)$. In order to apply the Itô Lemma we need to verify that $f \in C^{1,2}$. We have: $\frac{\partial f}{\partial t}(t, x) = \alpha$, $\frac{\partial f}{\partial x}(t, x) = \beta$ and $\frac{\partial^2 f}{\partial x^2}(t, x) = 0$ which are all continuous functions.

From the Itô Lemma (Theorem 2.3)), we see that $\eta(t), t \geq 0$ is an Itô process. Moreover,

$$\eta(t) = \gamma + \int_0^t \alpha\, ds + \int_0^t \beta\, dW(s).$$

(2) Let $\eta(t) = W_t^2$, $t \geq 0$. Is $\eta(t)$, $t \geq 0$ an Itô process?

Define $f(t, x) = x^2$ then $\eta(t) = f(t, W_t)$, $t \geq 0$. Again, we need to verify if f is of class $C^{1,2}$.

[3] technically speaking take any sequence of partitions with mesh tending to zero

We see that $\frac{\partial f}{\partial t} = 0$, $\frac{\partial f}{\partial x} = 2x$ *and* $\frac{\partial^2 f}{\partial x^2} = 2$ *are continuous functions, so* f *is of* $C^{1,2}$ *class. From the Itô Lemma, we deduce that* $\eta(t)$, $t \geq 0$ *is an Itô process and*

$$\eta(t) = 0 + \int_0^t [0 + \frac{1}{2}2]ds + \int_0^t [2W(s)]\, dW(s),\ t \geq 0.$$

Rearranging some terms this implies that

$$\begin{aligned} W(t)^2 &= \int_0^t 1\, ds + \int_0^t 2W(s)\, dW(s) \\ &= t + 2\int_0^t W(s)\, dW(s) \\ \text{i.e. } \int_0^t W(s)\, dW(s) &= \frac{1}{2}W(t)^2 - \frac{1}{2}t^2. \end{aligned}$$

Note that we have previously obtained this expression for $\int_0^t W(s)\, dW(s)$ *directly from the definition of the Itô integral.*

(3) *Our next example takes the form* $\eta(t) = W_t^3$, $t \geq 0$. *In this case we let* $f(t, x) = x^3$ *and observe that* $\eta(t) = f(t, W_t)$, $t \geq 0$. *Again, we need to verify if* f *is of class* $C^{1,2}$.

We see that $\frac{\partial f}{\partial t} = 0$, $\frac{\partial f}{\partial x} = 3x^2$ *and* $\frac{\partial^2 f}{\partial x^2} = 6x$ *are continuous functions, so* f *is indeed in* $C^{1,2}$.

From the Itô Lemma, it follows that η *is an Itô process and*

$$\eta(t) = 0 + \int_0^t \{0 + \frac{1}{2}[6W(s)]\}\, ds + \int_0^t [3W(s)^2]\, dW(s), t \geq 0$$

$$W(t)^3 = 3\int_0^t W(s)\, ds + 3\int_0^t [W(s)^2]\, dW(s), t \geq 0,$$

which implies

$$\int_0^t W(s)^2\, dW(s) = \frac{1}{3}[W(t)^3] - \int_0^t W(s)\, ds.$$

(4) *For our final example we consider the process* $\eta(t) = e^{\alpha t + \beta W_t}$, $t \geq 0$, *where* α *and* β *are real valued constants. Again, we ask if* $\eta(t)$, $t \geq 0$ *is an Itô process.*

Letting $f(t,x) = e^{\alpha t+\beta x}$ *we see that* $\eta(t) = f(t, W_t)$, $t \geq 0$. $\frac{\partial f}{\partial t}(t,x) = \alpha e^{\alpha t+\beta W_t} = \alpha f(t,x)$, $\frac{\partial f}{\partial x}(t,x) = \beta e^{\alpha t+\beta W_t} = \beta f(t,x)$ *and* $\frac{\partial^2 f}{\partial x^2} = \beta^2 e^{\alpha t+\beta W_t} = \beta^2 f(t,x)$ *are all continuous functions.*

Hence, f *is of* $C^{1,2}$ *class. As usual, from the Itô Lemma we we see that* $\eta(t), t \geq 0$ *is an Itô process and*

$$\begin{aligned}
\eta(t) &= f(0,0) + \int_0^t [\frac{\partial f}{\partial s}(s,W(s)) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(s,W(s))]\,ds + \int_0^t \frac{\partial f}{\partial x}(s,W(s))\,dW(s) \\
&= 1 + \int_0^t [\alpha f(s,W(s)) + \frac{1}{2}\beta^2 f(s,W(s))]\,ds + \int_0^t \beta f(s,W(s))\,dW(s) \\
&= 1 + \int_0^t (\alpha + \frac{\beta^2}{2})e^{\alpha s+\beta W(s)}\,ds + \int_0^t \beta e^{\alpha s+\beta W(s)}\,dW(s).
\end{aligned}$$

Thus,

$$\eta(t) = 1 + \int_0^t (\alpha + \frac{\beta^2}{2})\eta(s)\,ds + \int_0^t \beta\eta(s)\,dW(s). \quad (*)$$

Several remarks are in order at this stage.

Remarks 2.6. (1) Itô's Lemma is commonly also referred to as Itô formula or (less commonly) as Itô rule.

(2) We will often write the Itô Lemma using the following differential shorthand

$$df\,(t,W_t) = \left(\frac{\partial f}{\partial t}(t,W(t)) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t,W(t))\right)dt + \frac{\partial f}{\partial x}(t,W(t))\,dW_t.$$

This notation is **formal** and **always** has to be interpreted using the stochastic integrals.

(3) Similarly, an Itô process $\xi(t), t \geq 0$ with characteristics $a \in M^1_{loc}(0,\infty), b \in M^2_{loc}(0,\infty)$ can be written as

$$d\xi(t) = a(t)\,dt + b(t)\,dW(t).$$

Before we generalise the Itô formula to cover general Itô processes we introduce the notion of the quadratic variation of a stochastic process. In order to make this definition we require the following theorem.

Theorem 2.7. *Let $M = (M_t)_{t\geq 0}$ be a continuous square integrable martingale. There exists a unique continuous, adapted and increasing process $\langle M\rangle$ with $\langle M\rangle_0 = 0$ such that the process defined by $M_t^2 - \langle M\rangle_t, t \geq 0$ is a martingale.*

For those interested: The proof of this theorem is a consequence of the Doob-Meyer decomposition.

Definition 2.8. *Let $M = (M_t)_{t\geq 0}$ be a continuous square integrable martingale. Denote by $\langle M\rangle$ the unique continuous, increasing, adapted process with $\langle M\rangle_0 = 0$ such that $M_t^2 - \langle M\rangle_t$ is a martingale. We call $\langle M\rangle$ the quadratic variation (process) of M.*

Remark 2.9. *The existence of the quadratic variation follows from Theorem 2.7.*

Example 2.10. *(1) Let $M = W$ a Brownian motion. Then it is easily seen from the Itô lemma (exercise!) that $\langle M\rangle_t = t, t \geq 0$.*

(2) If $M = \int_0^t b(s)\, dW_s$, $b \in M_{loc}^2$ then $\langle M\rangle_t = \int_0^t (b(s))^2\, ds$.

(3) The definition of the quadratic variation can be extended in order to cover an Itô process X with characteristics a and b. In this case $\langle X\rangle$ equals the quadratic variation of the martingale part of X,i.e.

$$\langle X\rangle_t = \int_0^t (b(s))^2\, ds.$$

Note that this is a generalisation of the definition as X is in general not a martingale.

Remark 2.11. *It has been shown in MMF that for Brownian motion W if we let $t_i^n = \frac{i}{n}$, $i = 0, \ldots, n$ we have*

$$\sum_{j=0}^{n-1} \left(W_{t_{j+1}} - W_{t_j}\right)^2 \to \langle W \rangle_t$$

in mean square as n tends to infinity. More generally, if M is a continuous square integrable martingale

$$\sum_{j=0}^{n-1} \left(M_{t_{j+1}} - M_{t_j}\right)^2 \to \langle M \rangle_t$$

where the limit is interpreted in a suitable sense, thus motivating the term quadratic variation.

The following is the general form of the Itô Lemma. It shows that the image of an Itô process under a $C^{1,2}$ map is again an Itô process.

Theorem 2.12. *(Itô Lemma in general form)*
Assume that $(\xi(t))_{t\geq 0}$ is an Itô process with characteristics $(a(t))_{t\geq 0}$ and $(b(t))_{t\geq 0}$, i.e. $\xi(t) = \xi(0) + \int_0^t a(s)\,ds + \int_0^t b(s)\,dW(s), a \in M^1_{loc}(0,\infty), b \in M^2_{loc}(0,\infty)$.
Let $f(t,x)$ be a function of $C^{1,2}$ class, i.e. $\frac{\partial f}{\partial t}, \frac{\partial f}{\partial x}, \frac{\partial^2 f}{\partial x^2}$ exist and are continuous. Define a process η by

$$\eta(t) := f(t, \xi(t)),\ t \geq 0.$$

Then the process $\eta(t)$, $t \geq 0$ is an Itô process and

$$\begin{aligned} \eta(t) &= f(0,\xi(0)) + \int_0^t [\frac{\partial f}{\partial s}(s,\xi(s)) + \frac{\partial f}{\partial x}(s,\xi(s))a(s) \\ &+ \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(s,\xi(s))b^2(s)]ds + \int_0^t \frac{\partial f}{\partial x}(s,\xi(s))b(s)\,dW(s),\ t \geq 0. \end{aligned} \tag{2.1}$$

In other words, $\eta(t), t \geq 0$ is an Itô process with characteristics

$$\begin{aligned} \tilde{a}(s) &= \frac{\partial f}{\partial t}(s,\xi(s)) + \frac{\partial f}{\partial x}(s,\xi(s))a(s) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(s,\xi(s))b^2(s) \\ \tilde{b}(s) &= \frac{\partial f}{\partial x}(s,\xi(s))b(s). \end{aligned}$$

Remarks:

(1) The simple form of the Itô formula follows from Theorem 2.12 by taking $a \equiv 0,\ b \equiv 1,\ \xi(0) \equiv 0$.

(2) In differential notation (2.1)becomes

$$df\,(t,\xi_t) = \left(\frac{\partial f}{\partial t}(t,\xi_t) + \frac{\partial f}{\partial x}(t,\xi_t)a_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t,\xi_t)b_t^2\right) dt + \frac{\partial f}{\partial x}(t,\xi_t)b_t\,dW_t,$$

where

$$d\xi_t = a_t dt + b_t dW_t. \tag{2.2}$$

Using (2.2) for $d\xi_t$ it is sometimes also convenient to regroup terms and write the Itô formula in the more compact form

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

(3) The term $\frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t,\xi_t)b_t^2 dt$ comes from the quadratic variation of the process ξ, given by $\langle \xi \rangle_t = \int_0^t (b\,(s))^2\,ds$.

In particular, in the elementary literature the quadratic variation term is sometimes also written as $d\xi_t \cdot d\xi_t$. In the context of the Ito lemma this gives

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

The term $d\xi_t \cdot d\xi_t$ is informal notation and computed by the (again informal) rules of the following multiplication table:

$$\begin{aligned} dt \cdot dt &= 0, \quad dt \cdot dW_t = 0, \\ dW_t dt &= 0, \quad dW_t \cdot dW_t = dt. \end{aligned}$$

The following integration by parts formula is a consequence of the Itô lemma and often turns out to be useful. Its proof will be covered in the exercises.

Proposition (Stochastic Integration by Parts). *Let X and Y be stochastic processes*[4] *given by $dX_t = a_1(t)\,dt + b_1(t)\,dW_t$ and $dY_t = a_2(t)\,dt + b_2(t)\,dW_t$ respectively. Then*

$$X_tY_t = X_0Y_0 + \int_0^t (X_s a_2(s) + Y_s a_1(s) + b_1(s)\,b_2(s))\,ds$$
$$+\int_0^t (X_s b_2(s) + Y_s b_1(s))\,dW_s, t \geq 0$$

or equivalently

$$d(X_tY_t) = X_t dY_t + Y_t dX_t + b_1(t)\,b_2(t)\,dt.$$

3. Stochastic Differential Equations

The following example motivates the notion of a stochastic differential equation.

Example 3.1. *Recall from a previous example that if $\eta(t) = e^{\alpha t+\beta W(t)}$, then by Ito's Lemma*

$$\eta(t) = 1+\int_0^t (\alpha + \frac{\beta^2}{2})e^{\alpha s+\beta W(s)}ds + \int_0^t \beta e^{\alpha s+\beta W(s)}\,dW_s$$

which we can rewrite in differential notation as

$$d\eta(t) = (\alpha + \frac{\beta^2}{2})\eta(t)\,dt + \beta\eta(t)\,dW_t. \tag{3.1}$$

In the example we set $\eta(t) = e^{\alpha t+\beta W(t)}$ and obtained the equation (3.1) .Conversely, we could have started with equation (3.1) and asked if there is a process η that satisfies the equation (3.1) (in our case the answer is of course yes, as $e^{\alpha t+\beta W(t)}$ does the job). This motivates the definition of a stochastic differential equation.

Definition 3.2. *Assume that*

$$F, G : [0,\infty) \times \mathbb{R} \to \mathbb{R}$$

[4]Strictly speaking we have to require that X and Y are Ito processes in order to make sense of the stochastic integrals. In latter calculations, for example when computing SDE solutions you will not be required to check this condition as we could if necessary generalise our framework of stochastic integration (and Ito processes) to a more general class of adapted, piece-wise continuous integrands

are given functions. We say that an Itô process $X(t)$, $t \geq 0$ is a solution of SDE:

$$dX_t = F(t, X_t)dt + G(t, X_t)dW_t, \quad X_0 = x_0 \tag{3.2}$$

if and only if

$$X(t) = x_0 + \int_0^t F(s, X(s))ds + \int_0^t G(s, X(s))dW_s,\ t \geq 0. \tag{3.3}$$

The following example will provide us with some intuition how we can find the solution of a stochastic differential equation in some particular cases.

Example 3.3. *We would like to find a solution to the stochastic differential equation (SDE)*

$$dX_t = aX_tdt + bX_tdW_t, \tag{3.4}$$

where $a, b \in \mathbb{R}$.

HEURISTIC ARGUMENT *We formally divide the equation by dt and get*

$$\frac{dx}{dt} = ax + bx\frac{dW}{dt}.$$

For now suppose that W_t is a deterministic function of C^1 class (keep in mind that in reality the trajectories of the Brownian motion are a.s. a.e. non-differentiable though). Then $\frac{dW}{dt} = W'(t)$ is a classical derivative and our equation becomes:

$$\frac{dx}{dt} = (a + bW'(t))x$$

This is a linear first order ODE and can be solved using separation of variables. Dividing by x and multiplying by dt we have

$$\frac{dx}{x} = (a + bW'(t))\,dt$$

and integrating

$$\ln x = at + bW(t) + c$$

$$x(t) = e^c e^{at+bW(t)} = Ce^{at+bW(t)}.$$

Notice that $x(0) = C$. *In other words, if* W_t *was differentiable the process*

$$x(t) = Ce^{at+bW(t)}$$

would solve the equation

$$\begin{cases} dx = axdt + bxdW \\ x(0) = C \end{cases}.$$

It turns out that $X(t) := Ce^{at+bW(t)}$ *derived via our heuristic calculation is unfortunately not a solution to the SDE, but, as we will see now, solves a slightly different SDE.*

Indeed, we have seen in a previous example that we can apply the Itô lemma to $F(t, W(t))$ *and obtain*

$$X(t) = X(0) + \int_0^t aX(s)\,ds + \int_0^t bX(s)\,dW(s) \tag{3.5}$$

$$= C + \int_0^t (a + \frac{b^2}{2})X(s)\,ds + \int_0^t bX(s)\,dW(s). \tag{3.6}$$

Remark . In the previous example we used the Ito Lemma in its simple form in order to derive identity (3.5). Alternatively we could have applied the Ito Lemma in its general form with $f(x) = \exp(x)$ and the Ito process $d\xi_t = adt + bdW_t$, $\xi_0 = 0$ and computed $f(\xi_t)$.

From identity (3.5) *we see that* $X(t)$ *is a solution to another SDE, namely*

$$dX_t = (a + \frac{b^2}{2})X_t dt + bX_t dW_t, \quad X_0 = C. \tag{3.7}$$

Even though the process X *defined by* $X(t) = Ce^{at+bW(t)}$, $t \geq 0$, *is not a solution to the SDE*

$$dX_t = aX_t dt + bX_t dW(t), \quad X_0 = C$$

can we still use it to find a solution to this equation? If we denote $\alpha = (a + \frac{b^2}{2})$, $\beta = b$ *equation* (3.7)*can now be written as*

$$dX_t = \alpha X_t dt + \beta X_t dW(t), \quad X_0 = C \tag{3.8}$$

Now observe that the solutions to (3.7) *and* (3.8) *are the same, i.e.*

$$X(t) = Ce^{at+bW(t)} = Ce^{(\alpha-\frac{\beta^2}{2})t+\beta W(t)}$$

solve both SDEs. There is nothing to stop us from switching the constants α with a and β with b respectively. If we do so, the fact that $X(t) = Ce^{(\alpha-\frac{\beta^2}{2})t+\beta W(t)}$ solves (3.8) *immediately allows us to draw the following conclusion:*

Conclusion: *A solution to the SDE* (3.4) *is given by the process X defined by*

$$X(t) = Ce^{(a-\frac{b^2}{2})t+bW(t)}. \tag{3.9}$$

In the previous example we have found a solution to a simple but important SDE. It is a natural question if this solution is unique. It turns out that in general an SDE may have more than one solution.

Example 3.4. *Consider the stochastic differential equation*

$$dX_t = 3X_t^{1/3}dt + 3X_t^{2/3}dW_t.$$

with the initial condition $X_0 = 0$. Clearly, the process $X_t = 0$ is a solution. But so is $X_t = W_t^3$.

A situation where a stochastic differential equation has more than one solution has several disadvantages. For instance, if we know there exists a unique solution we can deduce that any two processes that satisfy the same SDE are the same. Also, if we know that an SDE has a unique solution, we can try to find a candidate by a heuristic calculation (as done in our example) and then verify, usually using the Itô lemma, that it is indeed a solution. If we know that the solution of our SDE is unique, we are done and we do not have to deal with the more subtle question of finding *all* solutions. The following theorem shows that under suitable conditions on the functions defining the SDE one can ensure that there exists precisely one solution[5] to an SDE.

Theorem 3.5 (Existence and uniqueness of SDE solutions)**.** *Assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and $(W_t)_{t\geq 0}$ is a Brownian motion with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$* [6]*. Let $f(t,x)$, $g(t,x)$ be given functions depending on $t \in [0,T]$ (where $T > 0$ is fixed) and $x \in \mathbb{R}$. Assume that f and g satisfy the following two conditions*

(L) $\exists$ *a constant $L > 0$ such that for all $t \in [0,T]$, $x_1, x_2 \in \mathbb{R}$,*

$$\begin{aligned} |f(t,x_2) - f(t,x_1)| &\leq L|x_2 - x_1| \\ |g(t,x_2) - g(t,x_1)| &\leq L|x_2 - x_1|. \end{aligned}$$

(G) *There exists a constant C such that for all $t \in [0,T]$, $x \in \mathbb{R}$*

$$|f(t,x)| + |g(t,x)| \leq C\left(1 + |x|\right).$$

Suppose also that $X_0 : \Omega \to \mathbb{R}$ is $\mathcal{F}_0$-measurable, independent of the Brownian motion W and satisfies $\mathbb{E}|X_0|^2 < \infty$.

Then, there exists a unique solution $X(t)$, $t \in [0,T]$ of the SDE

$$\begin{cases} dX(t) = f(t,X(t))\,dt + g(t,X(t))\,dW(t) \\ X(0) = X_0 \end{cases}$$

Moreover,

$$\mathbb{E} \sup_{t\in[0,T]} |X(t)|^2 < \infty.$$

Remark 3.6. *(1) In condition (L) the L stands in fact for (globally) Lipschitz and the G in (G) is for linear growth.*

[5] Strictly speaking the theorem shows the existence of strong solutions and pathwise uniqueness.
[6] satisfying "the usual conditions"

(2) *Provided that (L) is satisfied showing that* $|f(t,0)| \leq \widetilde{C}$ *and* $|g(t,0)| \leq \widetilde{C}$ *for all* $t \in [0,T]$ *implies that (G) is satisfied as in that case*

$$\begin{aligned} |g(t,x)| &\leq |g(t,x) - g(t,0)| + |g(t,0)| \\ &\leq L|x| + \widetilde{C} \leq \max\left(L, \widetilde{C}\right)(|x|+1). \end{aligned}$$

(3) *If* $f(t,x)$ *and* $g(t,x)$ *do not depend on* t*, i.e.* $f(t,x) = f(x)$ *and* $g(t,x) = g(x)$ *then the previous part of the remark and condition (L) imply that (G) is satisfied.*

Example 3.7. *Let* $a, b \in \mathbb{R}$ *and consider the SDE*

$$dX_t = aX_t dt + bX_t dW_t, \quad X_0 = C > 0 \tag{3.10}$$

We would like to show that Theorem 3.5 applies with

$$f(t,x) = ax, \; g(t,x) = bx.$$

To see that note that both functions do not depend explicitly on t*. So, condition (G) is satisfied, if we can verify (L). We have*

$$\begin{aligned} |f(t,x_2) - f(t,x_1)| &= |ax_2 - ax_1| = |a||x_2 - x_1| \\ |g(t,x_2) - g(t,x_1)| &= |bx_2 - bx_1| = |b||x_2 - x_1|. \end{aligned}$$

Hence, (L) is satisfied with $L = \max\{|a|,|b|\}$ *and by Theorem 3.5 we conclude that the process* X *defined by*

$$X(t) = Ce^{(a - \frac{b^2}{2})t + bW(t)}, \; t \geq 0$$

is ***the unique*** *solution of SDE* (3.10) *(we found this solution in an earlier example).*

The following example contains some ordinary differential equations (ODEs) for which we can easily find an explicit solution. The example is optional reading.

Example 3.8. *(a)* $\frac{dx}{dt} = \alpha x \longrightarrow x(t) = ce^{\alpha t}$*, c is a constant.*

(b) $\frac{dx}{dt} = \sqrt{|x|} \longrightarrow x(t) \equiv 0$ *is a solution,* $x(t) = ct^2$.

$\frac{dx}{dt} = 2ct = \sqrt{|x|} = \sqrt{ct^2} = \sqrt{c}t$.

$2c = \sqrt{c},\ 4c^2 = c,\ c = \frac{1}{4}$

i.e. $x(t) = \frac{t^2}{4}$ *is also a solution. Non-uniqueness.*

(c) $\frac{dx}{dt} = x^2$, $x(t) = \frac{\pm 1}{t - \frac{1}{x_0}}$ *a solution exists only for finite time,* $t < \frac{1}{x_0}$.

We continue with a rough sketch of the proof of Theorem 3.5. The proof relies on the following theorem, which is optional reading.

Theorem 3.9 (Banach fixed point theorem). *Let* (V, d) *be a complete metric space (non-empty) and* $\varphi : V \to V$ *a contraction then* φ *has exactly one fixed point* v^*.

Remark 3.10. *(1)* $v \in V$ *is a fixed point for the function* φ *if* $\varphi(v) = v$.

(2) A map $\varphi : V \to V$ *is a contraction if for all* $x, y \in V$

$$d(\varphi(x), \varphi(y)) \leq \lambda d(x, y)$$

for some $\lambda \in [0, 1)$, *i.e.* φ *is Lipschitz with constant* $L < 1$.

(3) For all $x_0 \in V$ *the sequence* $(x_n)_{n \geq 0}$ *defined by* $x_{n+1} = \varphi(x_n)$ *tends to the fixed point* v^* *as* n *tends to infinity.*

Proof of Theorem 3.5 (Rough sketch existence part). We define a sequence (usually referred to as Picard iterations) $(\xi^n)_{n \geq 0}$ by setting $\xi^0(t) = X_0$ and

$$\xi^1(t) = X_0 + \int_0^t f(s, \xi^0(s))\, ds + \int_0^t g(s, \xi^0(s))\, dW(s), t \in [0, T]$$

$$\xi^2(t) = X_0 + \int_0^t f(s, \xi^1(s))\, ds + \int_0^t g(s, \xi^1(s))\, dW(s), t \in [0, T].$$

In general, we define the process $\xi^{n+1}(t)$ by:

$$\xi^{n+1}(t) = X_0 + \int_0^t f(s, \xi^n(s))\, ds + \int_0^t g(s, \xi^n(s))\, dW(s), t \in [0, T].$$

We can define a map Φ taking Itô processes to Itô processes by letting

$$\Phi\left(\xi\right)\left(t\right) = X_0 + \int_0^t f(s,\xi(s))\,ds + \int_0^t g(s,\xi(s))\,dW(s), t \in [0,T].$$

Using (L) and (G) one can now identify a suitable Banach space such that the map Φ is a contraction. By the Banach fixed point theorem Φ has a unique fixed point, i.e. there exists an Itô process such that

$$\Phi\left(\xi\right) = \xi$$

and moreover $\xi^n \to \xi$. We deduce that

$$\xi_t = X_0 + \int_0^t f(s,\xi(s))\,ds + \int_0^t g(s,\xi(s))\,dW(s),$$

so ξ solves our SDE. □

Example 3.11. *Consider the SDE*

$$\begin{cases} dX_t = 5dt + 7dW_t \\ X(0) = 2. \end{cases}$$

Then $X_0 = 2$ is $\mathcal{F}_0$-measurable. Because $f(t,x) = 5$ and $g(t,x) = 7$ are constant functions, they are trivially bounded. So (G) is satisfied. The condition (L) is satisfied as $|f(t,x_2) - f(t,x_1)| = |g(t,x_2) - g(t,x_1)| = 0$. Hence, the assumptions of Theorem 3.10 are satisfied and the equation has a unique solution.

The Picard iteration for this particular example is given by

$$\begin{aligned} \xi^0(t) &= x_0, \\ \xi^1(t) &= x_0 + \int_0^t f(s,\xi^0(s))\,ds + \int_0^t g(s,\xi^0(s))\,dW(s) = 2 + 5t + 7W(t) \\ \xi^2(t) &= x_0 + \int_0^t f(s,\xi^1(s))\,ds + \int_0^t g(s,\xi^1(s))\,dW(s) = 2 + 5t + 7W(t). \end{aligned}$$

Hence, for each $n \in \mathbb{N}$,

$$\xi^n(t) = 2 + 5t + 7W(t).$$

It follows that

$$\lim_{n\to\infty} \xi^n(t) = 2 + 5t + 7W(t),$$

which is the SDE solution.

Example 3.12. *We would like to solve the SDE*

$$dX_t = \sin(X_t)\,dt + \cos(X_t)\,dW_t, \quad X_0 = 0 \qquad (1).$$

We first check if the SDE satisfies the assumptions of Theorem 3.5. Let

$$f(t,x) = \sin(x),\ g(t,x) = \cos(x)$$

Assuming without loss of generality $x_1 < x_2$, we have

$$\begin{aligned} |f(t,x_2) - f(t,x_1)| &= |\sin(x_2) - \sin(x_1)| = |\int_{x_1}^{x_2} \cos(y)\,dy| \\ &\leq \int_{x_1}^{x_2} 1dy = x_2 - x_1 = |x_2 - x_1| \end{aligned}$$

Therefore, the function f satisfies condition (L) with constant $L = 1$. Similarly, we can verify that g also satisfies condition (L) with constant $L = 1$. Since, the functions f and g do not depend explicitly on t, we infer that condition (G) is satisfied.

It is not obvious that the previous example has an explicit solution. In fact, most SDEs do not have explicit solutions, i.e. a solutions given by an explicit (simple) formula. In this case we have to approximate the solution using numerical methods.

Feynman-Kac formula. The following theorem establishes an important connection between stochastic processes (in particular, SDE solutions) and partial differential equations (second order, linear, parabolic PDEs).

Theorem 3.13 (Feynman-Kac formula)**.** *Let $\phi : \mathbb{R} \to \mathbb{R}$ be a given function and let X be an Itô process defined as the solution of the SDE*

$$dX(t) = a(t,X_t)dt + b(t,X_t)dW(t). \tag{3.11}$$

Let $F(t,x)$ be a function of two variables in $C^{1,2}$ that satisfies

$$\frac{\partial F}{\partial t}(t,x) = -\frac{1}{2}b^2(t,x)\frac{\partial^2 F}{\partial x^2}(t,x) - a(t,x)\frac{\partial F}{\partial x}(t,x), \tag{3.12}$$

$$F(T,x) = \phi(x). \tag{3.13}$$

Moreover, we assume that

$$\frac{\partial F}{\partial x}(t,X(t))b(t) \in M^2_{loc}.$$

Then

$$F(t,X(t)) = \mathbb{E}(\phi(X(T))|\mathcal{F}_t), \tag{3.14}$$

where $(\mathcal{F}_t)_{t\geq 0}$ is any filtration such that $F(t,X(t))$ is $\mathcal{F}_t$-measurable for all $t \geq 0$ and the filtration $(\mathcal{F}_t)_{t\geq 0}$ can be enlarged to a filtration $(\tilde{\mathcal{F}}_t)_{t\geq 0}$ s.t. W is Brownian Motion with respect to $(\tilde{\mathcal{F}}_t)_{t\geq 0}$.

Proof. Applying the Itô lemma and using that F solves the PDE we see that

$$\begin{aligned} F(t,X_t) &= F(0,0) + \int_0^t \left[\frac{\partial F}{\partial s}(s,X_s) + a(s,X_s)\frac{\partial F}{\partial x}(s,X_s) + \frac{1}{2}b^2(s,X_s)\frac{\partial^2 F}{\partial x^2}(s,X_s)\,ds\right] \\ &\quad + \int_0^t b(s,X_s)\frac{\partial F}{\partial x}(s,X_s)\,dW_s \\ &= \int_0^t b(s,X_s)\frac{\partial F}{\partial x}(s,X_s)\,dW_s. \end{aligned}$$

Thus, noting that by definition $F(T,X(T)) = \phi(X(T))$ we have

$$\begin{aligned} \phi(X(T)) - F(0,0) &= \int_0^T b(s,X_s)\frac{\partial F}{\partial x}(s,X_s)\,dW(s) \\ F(t,X(t)) - F(0,0) &= \int_0^t b(s,X_s)\frac{\partial F}{\partial x}(s,X_s)\,dW(s) \end{aligned}$$

and subtracting gives

$$\phi(X(T)) - F(t,X(t)) = \int_t^T b(s,X_s)\frac{\partial F}{\partial x}(s,X_s)\,dW(s).$$

Taking conditional expectations with respect to $\mathcal{F}_t$ and using the martingale property of the stochastic integral for the last equality we have

$$\mathbb{E}(\phi(X(T)) - F(t,X(t))|\mathcal{F}_t) = \mathbb{E}(\int_t^T F_x(s,X(s))b(s)dW(s)|\mathcal{F}_t) = 0.$$

Since $F(t,X(t))$ is $\mathcal{F}_t$-measurable, we have

$$\mathbb{E}(F(t,X(t))|\mathcal{F}_t) = F(t,X(t))$$

completing the proof. $\square$

Remark 3.14. *Let $u \in \mathbb{R}$ be a constant and set $X_0 = u$. Denote by $X^u = (X_t^u)_{t\in[0,T]}$ the solution to the SDE* (3.11) *with intial data $X_0 = u$ and let $(\mathcal{F}_t)_{t\geq 0}$ be the natural filtration generated by X^u. Then identity* (3.14) *becomes*

$$F(0,u) = E(\varphi(X_T^u)).$$

Note that here we have used that $\mathcal{F}_0$ is the trivial sigma algebra.

The Feynman-Kac formula allows us to solve a PDE by computing the expected value of a function of an SDE solution. Later we will use the Feynman-Kac formula to solve the Black-Scholes PDE that governs the price of a (European) option in the Black-Scholes framework.

Example 3.15. *If $a(t,X_t) = 0$ and $b(t,X_t) = 1$ we have*

$$X_t = W_t + x_0$$

and the PDE in the Feynman-Kac formula becomes

$$\begin{aligned} \frac{\partial F}{\partial t}(t,x) &= -\frac{1}{2}\Delta F(t,x) \\ F(T,x) &= \varphi(x). \end{aligned}$$

Writing $u(t,x) = F(T-t,x)$ we can transform the terminal value problem into the initial value problem

$$\begin{aligned} \frac{\partial u}{\partial t}(t,x) &= \frac{1}{2}\Delta u(t,x) \\ u(0,x) &= \varphi(x), \end{aligned}$$

which is the heat equation.

3.1. **Some Revision: Absolutely continuous measures.** We begin by recalling the definition of absolute continuity for measures.

Definition 3.16. *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *and* $(\Omega, \mathcal{F}, \mathbb{Q})$ *be two probability spaces. We say that a probability measure* $\mathbb{Q}$ *is* ***absolutely continuous*** *with respect to* $\mathbb{P}$ *if and only if for every* $A \in \mathcal{F}$

$$\mathbb{P}(A) = 0 \Rightarrow \mathbb{Q}(A) = 0. \tag{3.15}$$

We write $\mathbb{Q} \ll \mathbb{P}$*. Measures* $\mathbb{P}$ *and* $\mathbb{Q}$ *are called equivalent (we write* $\mathbb{P} \sim \mathbb{Q}$*) if and only if*

$$\mathbb{Q} \ll \mathbb{P} \textit{ and } \mathbb{P} \ll \mathbb{Q}.$$

Example 3.17.

(i) *Take* $\Omega = (0,1)$*,* $\mathcal{F} = \mathcal{B}(0,1)$*, i.e. the Borel sigma algebra of* $(0,1)$*,* $\mathbb{P}$ *the Lebesgue measure and* $\mathbb{Q} = \delta_{\frac{1}{2}}$*. In that case* $\mathbb{P}((a,b)) = |b-a|$ *if* $0 \le a \le b \le 1$*, and*

$$\mathbb{Q}(A) = \begin{cases} 1, & \textit{if } \frac{1}{2} \in A, \\ 0, & \textit{if } \frac{1}{2} \notin A. \end{cases}$$

Question 1. *Does* $\mathbb{Q} \ll \mathbb{P}$ *hold?*

Consider the set

$$A = [\frac{1}{2}, \frac{1}{2}]$$

then

$$\mathbb{P}(A) = 0 \quad \textit{and } \mathbb{Q}(A) = 1.$$

Hence, $\mathbb{Q}(A) \not\ll \mathbb{P}(A)$*.*

Question 2. *Does* $\mathbb{P} \ll \mathbb{Q}$ *hold?*

Again, we give a counter example. Let A *be given by*

$$A = [\frac{2}{3}, 1)$$

Then, because $\frac{1}{2} \notin A$,

$$\mathbb{Q}(A) \quad = \quad 0.$$

On the other hand,

$$\mathbb{P}(A) \quad = \quad 1 - \frac{2}{3} = \frac{1}{3} > 0.$$

Once again we conclude that $\mathbb{P} \not\ll \mathbb{Q}$.

(ii) *For our second example we consider*

$$\mathbb{Q}(A) = \int_A \frac{1}{\sqrt{x}} dx, A \in \mathcal{B}(0,1)$$

Note that this is not a probability measure as

$$\int_\Omega \frac{1}{\sqrt{x}} dx = 2\sqrt{x}|_0^1 = 2 \neq 1.$$

We can normalise the measure (i.e. divide it by two in this case) and get

$$\mathbb{Q}(A) = \int_A \frac{1}{2\sqrt{x}} dx.$$

With this new definition $\mathbb{Q}$ *is indeed a probability measure.*

To see that $\mathbb{Q} \ll \mathbb{P}$ *holds we observe that if* $A \in \mathcal{B}(0,1)$ *with* $\mathbb{P}(A) = 0$

$$\mathbb{Q}(A) = \int_A \frac{1}{2\sqrt{x}} dx = \int_A \frac{1}{2\sqrt{x}} d\mathbb{P}(x) = 0$$

(iii) *For our final example let* $\mathbb{Q}(A) = \int_A f(x)dx, A \in B(0,1)$.

Note that $\mathbb{Q}$ *is a probability measure provided that*

(i) $f(x) \geq 0$ *for a.e.* $x \in \mathbb{R}$;

(ii) $\int_0^1 f(x)dx = 1$.

Note that the condition (ii) has to be satisfied because $\mathbb{Q}$ *is a probability measure and, hence,*

$$\mathbb{Q}(\Omega) = \int_\Omega f(x)dx = \int_0^1 f(x)dx = 1.$$

Using the same argument as in Exercise (ii), but this time replacing $\frac{1}{2\sqrt{x}}$ *by* $f(x)$ *we see that* $\mathbb{Q} \ll \mathbb{P}$.

Theorem 3.18. ***(Radon-Nikodym Theorem)***

Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ *and* $(\Omega, \mathcal{F}, \mathbb{Q})$ *are two probability spaces. Then the following two claims are equivalent*

(a) $\mathbb{Q} \ll \mathbb{P}$.

(b) There exists an $\mathcal{F}$*-measurable non-negative function* $f : \Omega \to [0, \infty)$ *such that*

$$\mathbb{Q}(A) = \int_A f(\omega) d\mathbb{P}(\omega). \tag{3.16}$$

Remark 3.19. *If (3.16) is satisfied, then*

$$\int_\Omega f(\omega) d\mathbb{P}(\omega) \quad = \quad (\mathbb{Q}(\Omega)) = 1.$$

Notation: *The function* f *from property (b) of Theorem 3.18 is denoted by* $\frac{d\mathbb{Q}}{d\mathbb{P}}$ *and called the Radon-Nikodym derivative of the measure* $\mathbb{Q}$ *with respect to measure* $\mathbb{P}$.

The Radon-Nikodym derivative$\frac{d\mathbb{Q}}{d\mathbb{P}}$ *is also often called the* ***density*** *of* $\mathbb{Q}$ *with respect to* $\mathbb{P}$.

3.2. **The Girsanov Theorem.** The Girsanov theorem states that, for Brownian motion, an absolutely continuous change of measure corresponds to changing the drift.

Theorem 3.20. ***(Girsanov Theorem)***

Assume that $(\Omega, \mathcal{F}, \mathbb{P})$ *is a probability space,* $(W_t)_{t \geqslant 0}$ *is a Brownian motion with respect to a filtration* $(\mathcal{F}_t)_{t \geqslant 0}$. *Let* $T > 0$. *Assume that* $(\theta_t)_{t \geqslant 0}$ *is an adapted process with piecewise continuous trajectories and such that*

$$\int_0^T \theta_s^2 ds < \infty \ a.s.$$

Define

$$L_t = e^{-\int_0^t \theta_s dW_s - \dfrac{1}{2} \int_0^t \theta_s^2 ds}, \ t \in [0, T] \tag{3.17}$$

and assume that $(L_t)_{t\in[0,T]}$ *is a martingale w.r.t. the filtration* $(\mathcal{F}_t)_{t\in[0,T]}$. *Then there exists a probability measure* $\mathbb{Q}_T$ *on* $(\Omega, \mathcal{F}_T)$ *such that*

$$V_t = W_t + \int_0^t \theta_s ds, \; t \in [0,T]$$

is $(\mathcal{F}_t)_{t\in[0,T]}$*-Brownian motion on* $(\Omega, \mathcal{F}_T, \mathbb{Q}_T)$ *and the measure* $\mathbb{Q}_T$ *satisfies*

$$\mathbb{Q}_T(A) := \int_A L_T(\omega) d\mathbb{P}(\omega) \quad A \in \mathcal{F}_T. \tag{3.18}$$

In particular, $\mathbb{Q}_T \ll \mathbb{P}$.

Before we consider an example for the application of the Girsanov theorem, we record the following remarks.

Remark 3.21. *(1) Note that* (3.18) *means that (restricted to* $\mathcal{F}_T$*) we have*

$$\frac{d\mathbb{Q}_T}{d\mathbb{P}} = L_T$$

and, in particular,

$$\mathbb{Q}_T(\Omega) := \int_\Omega L_T(\omega) d\mathbb{P}(\omega) = \mathbb{E}\left(L_t\right) = 1,$$

i.e. $\mathbb{Q}$ *is a probability measure.*

(2) The transformation $\mathbb{P} \to \mathbb{Q}_T$ *is called the Girsanov transformation of measures.*

Example 3.22. *We would like to apply the Girsanov theorem in the case* $\theta_t = b$, *for all* $t \geq 0$ *and we have*

$$L_t = e^{-\int_0^t \theta_s dW_s - \frac{1}{2}\int_0^t \theta_s^2 ds} = e^{-bW_t - \frac{b^2}{2}t}$$

In order to determine whether $(L_t)_{t\geqslant 0}$ *is a martingale w.r.t. the filtration* $(\mathcal{F}_t)_{t\geqslant 0}$ *we have to check the following.*

(i) $E|L_t| < \infty$.

(ii) L_t *is* $\mathcal{F}_t$ *measurable.*

(iii) $E(L_t \big| \mathcal{F}_s) = L_s$.

Checking these conditions is left as an exercise.

The following theorem often simplifies the application of the Girsanov Theorem. Notably, it replaces the hypothesis that the process $L(t)$, $t \in [0,T]$ (defined in (3.17)) is a martingale with an integrability condition, which is often easier to check.

Theorem 3.23 (Novikov Condition). *If* $\mathbb{E}\, e^{\frac{1}{2}\int_0^T \theta(t)^2\, dt} < \infty$, *then the process* $L(t)$, $t \in [0,T]$ *(defined in* (3.17)*)* is a martingale.

Remark 3.24. *The Novikov condition* $\mathbb{E}\, e^{\frac{1}{2}\int_0^T \theta(t)^2\, dt} < \infty$ *is a sufficient condition to conclude that* L_t *is a martingale. It often allows us to simplify the application of the Girsanov theorem. In order to apply Girsanov it suffices to check that* θ *is an adapted and piecewise-continuous process provided Novikov's condition holds.*

Example 3.25. *We continue the example* $\theta = b$ *from before and have*

$$\mathbb{E}\, e^{\frac{1}{2}\int_0^T \theta(t)^2\, dt} = \mathbb{E}\, e^{b^2 T} < \infty,$$

i.e. Novikov's condition holds. Adaptedness and continuity are obvious. Hence, all assumptions in Theorem 6.3 are satisfied and we conclude that

$$V(t) = W_t + \int_0^t \theta_s ds = W_t + bt, t \in [0,T]$$

is a Brownian motion on $(\Omega, \mathcal{F}_T, \mathbb{Q}_T)$ *with respect to the filtration* $(\mathcal{F}_t)_{t\in[0,T]}$.

Example 3.26. *Suppose* Y_t *is an Ito process given by*

$$dY_t = g\left(t\right) dt + dW\left(t\right), t \in \left[0,T\right],$$

where $g : \mathbb{R} \to \mathbb{R}$ *is a continuous deterministic function. Then, if* $\theta_s = g\left(s\right)$ *for* $s \in [0,T]$ *we have*

$$\int_0^T \theta(t)^2\, dt = \int_0^T g(t)^2\, dt \le C \int_0^T dt = CT,$$

where $C := \max_{t\in[0,T]} g^2(t)$ *. Thus,*

$$\mathbb{E}\, e^{\frac{1}{2}\int_0^T \theta(t)^2\,dt} = \exp\left(\frac{1}{2}CT\right) < \infty.$$

It is clear that θ *is continuous and adapted, so by applying Girsanov's theorem we conclude that* Y *is Brownian motion on* $(\Omega, \mathcal{F}_T, \mathbb{Q}_T)$ *,where* $\mathbb{Q}_T$ *is absolutely continuous with respect to* $\mathbb{P}$ *and satisfies*

$$\frac{d\mathbb{Q}_T}{d\mathbb{P}} = L_T.$$

Example 3.27. *For our next example we ask if there exists a measure* $\mathbb{Q}_T$ *on* $(\Omega,_T)$ *, absolutely continuous with respect to* $\mathbb{P}$ *such that the process*

$$W_t + t^3 + t, \quad t \in [0,T]$$

is Brownian motion on $(\Omega, \mathcal{F}_T, \mathbb{Q}_T)$ *. Applying the previous example with* $g(t) = 3t^2 + 1$ *(in that case* $\int_0^T \theta(t)^2\,dt = t^3 + t$*) we see that such a measure* $\mathbb{Q}_T$ *indeed exists.*

Example 3.28. *For our final example we ask if we can find an absolutely continuous measure* $\mathbb{Q}_T$ *(w.r.t.* $\mathbb{P}$*) such that*

$$W_t + 1, \quad t \in [0,T]$$

is Brownian motion on $(\Omega, \mathcal{F}_T, \mathbb{Q}_T)$ *. It turns out in this case the answer is no. To see this, we assume for a contradiction that such a measure* $\mathbb{Q}_T$ *exists. Consider the event* A *defined by*

$$A := \{\omega \in \Omega : W_0(\omega) = 0\}.$$

By definition of Brownian motion we have $\mathbb{P}(A) = 1$ *and, hence,* $\mathbb{P}(A^c) = 0$*. As we assumed that* $\mathbb{Q}_T \ll \mathbb{P}$ *it follows that* $\mathbb{Q}_T(A^c) = 0$ *and* $\mathbb{Q}(A) = 1$ *or in other words* $W_0 = 0$ $\mathbb{Q}_T$*-a.s.. This implies of course that* $W_0 + 1 = 1$ $\mathbb{Q}_T$*-a.s.. But, by assumption* $W_t + 1$ *is Brownian motion with respect to* $\mathbb{Q}_T$*, which implies* $W_0 + 1 = 0$

$\mathbb{Q}_T$-*a.s.. This is a contradiction and we conclude that such a measure* $\mathbb{Q}_T$ *cannot exist.*

4. Option valuation in the Black-Scholes model

4.1. **The basic setup: The Black Scholes model, trading strategies and options.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space, $W(t)$, $t \geq 0$ a Brownian motion on this space and $(\mathcal{F}_t)_{t\geq 0}$ the natural filtration generated by W. We consider a world with just one risky asset (usually a share) with price process S_t and a risk-free savings account paying interest at (interest) rate $r > 0$ with continuous compounding.

We make the following fundamental assumptions: We assume the price process S_t of the share is geometric Brownian motion. By that we mean it satisfies the SDE

$$dS_t = \mu S(t)\,dt + \sigma S(t)\,dW, \quad S_0 > 0, \quad t \geq 0 \tag{4.1}$$

for given **drift** $\mu \in \mathbb{R}$, **volatility** $\sigma > 0$ and initial price $S_0 > 0$. We want the savings account also to be expressed as a tradable asset, i.e. we should invest in it by buying a certain number of units of "something". A convenient "something" is a zero-coupon **bond** growing at rate $r > 0$ and satisfying

$$dB_t = rB\,(t)\,dt, \quad t \geq 0. \tag{4.2}$$

Above, r, μ and σ are certain fixed positive constants. The two equations (4.1) and (4.2) are in fact independent of each other and in the previous section we have already determined their exact solutions to be

$$\begin{cases} S(t) &= S(0)e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W(t)}, \quad t \geq 0, \\ B(t) &= B(0)e^{rt}, \quad t \geq 0. \end{cases}$$

In the Black-Scholes model we further assume that we have a **frictionless market** by which we mean

(1) S and B can be traded in arbitrary amounts with no transaction costs

(2) short positions are allowed, this means we can invest in, or borrow from, the riskless account at the same rate of interest r.

Admissible ways to trade in the market are formalised in the following definition.

Definition 4.1. *A* ***trading strategy*** (ψ, φ) *is a pair of adapted processes* $(\psi(t))_{t\geq 0}$ *and* $(\varphi(t))_{t\geq 0}$ *which have continuous (or piecewise continuous) trajectories a.s., such that for each* $T > 0$,

$$\int_0^T |\psi(s)|\, ds + \int_0^T |\varphi(s)|^2\, ds < \infty. \tag{4.3}$$

The corresponding wealth process $X(t)$, $t \geq 0$ *of this portfolio, is defined by*

$$X(t) := \psi(t)B(t) + \varphi(t)S(t), \quad t \geq 0. \tag{4.4}$$

Note that the wealth process X associated to a trading strategy is an Ito process.

Definition 4.2. *A trading strategy* (ψ, φ) *is* ***self-financing*** *if the corresponding wealth process satisfies*

$$dX_t = \psi(t)\, dB_t + \varphi(t)\, dS_t, \quad t \geq 0 \tag{4.5}$$

or equivalently

$$\psi_t B_t + \varphi_t S_t - (\psi_s B_s + \varphi_s S_s) = \int_s^t \varphi_u dS_u + \int_s^t \psi_u dB_u$$

for $s, t \geq 0$.

Intuitively this means the change in value of the portfolio corresponding to the trading strategy is entirely due to gains from trade.

Remark 4.3. (a) *Note that if the process* ψ *belongs to* $M^1_{\mathrm{loc}}(0,\infty)$ *and the process* φ *belongs to* $M^2_{\mathrm{loc}}(0,\infty)$ *then the pair satisfies the condition (4.3).*

(b) *To motivate the above definition of a self financing strategy we demonstrate how we can obtain it as a limit of discrete time models with time step* $\Delta t > 0$.

In the discrete case a self-financing strategy is a pair of predictable processes $(\psi(t))_{t\in\mathbb{T}}$ *and* $(\varphi(t))_{t\in\mathbb{T}}$*, where* $\mathbb{T}$ *is the set of times (e.g.* $\mathbb{T}=\mathbb{N}\Delta t$*) satisfying*

$$X(t+\Delta t)-X(t)=\psi(t)[B(t+\Delta t)-B(t)]+\varphi(t)[S(t+\Delta t)-S(t)],\quad t\in\mathbb{T}.$$

Suppose now that $T>0$ *is fixed. Choose* $N\in\mathbb{N}$ *and set* $\Delta t=\frac{T}{N}$, $t_i^N=\frac{iT}{N}$, $i=0,1,\cdots,N$*. The last formula can then be rewritten as*

$$X(t_{i+1}^N)-X(t_i^N)=\psi(t_i^N)[B(t_{i+1}^N)-B(t_i^N)]+\varphi(t_i^N)[S(t_{i+1}^N)-S(t_i^N)],\quad i=0,1,\cdots,N-1.$$

Taking the sum from $i=0$ *to* $i=N-1$*,i.e. summing up the trading gains, we get*

$$X(T)-X(0)=\sum_{i=0}^{N-1}\psi(t_i^N)[B(t_{i+1}^N)-B(t_i^N)]+\sum_{i=0}^{N-1}\varphi(t_i^N)[S(t_{i+1}^N)-S(t_i^N)].$$

Recalling the definition of the Itô integral we see that (under some appropriate assumptions)

$$\begin{aligned}\sum_{i=0}^{N-1}\psi(t_i^N)[B(t_{i+1}^N)-B(t_i^N)] &\to \int_0^T\psi(t)\,dB(t)\\ \sum_{i=0}^{N-1}\varphi(t_i^N)[S(t_{i+1}^N)-S(t_i^N)] &\to \int_0^T\varphi(t)\,dS(t)\end{aligned}$$

as $N\to\infty$*. Hence, we it follows that*

$$X(T)=X(0)+\int_0^T\psi(t)\,dB(t)+\int_0^T\varphi(t)\,dS(t),\quad T\geq 0.$$

(c) *The meaning of equation (4.5) is similar to the definition of an Itô process. It is possible to give an independent definition of the integral* $\int_0^t\varphi(s)\,dS(s)$*, but the easier way is to define it is by taking into account equation* (4.1)*. That is, we define*

$$\int_0^t\varphi(s)\,dS(s):=\int_0^t\varphi(s)\mu S(s)\,ds+\int_0^t\varphi(s)\sigma S(s)\,dW(s),\quad t\geq 0.$$

In the above definition, the first integral on the right hand side is a standard Riemann integral while the second term is an Itô integral.

Similarly, although it would be possible to give an independent definition of *the integral* $\int_0^t \psi(s)\,dB(s)$ *(as a Lebesgue-Stieltjes integral), the easiest way is to define it (by taking into account equation* (4.2)*) in the following way*

$$\int_0^t \psi(s)\,dB(s) = \int_0^t \psi(s) r B(s)\,ds, \quad t \geq 0.$$

In the above definition, the integral on the right hand side is a standard Lebesgue integral.

Definition 4.4. *Denote by* X_t *the wealth process of a trading strategy* (φ, ψ) *. An arbitrage opportunity is the existence of a self-financing trading strategy and a time* $t > 0$ *such that* $X_0 = 0$*,* $X_t \geq 0$ *a.s. and* $P[X_t > 0] > 0$ *a.s. (or equivalently* $\mathbb{E}X_t > 0$*).*

It is axiomatic that arbitrage cannot exist in the market, so no mathematical model of a market should permit arbitrage.

The following definition introduces the notion of European options.

Definition 4.5. *Consider the Black-Scholes model of a financial market. Let* T *be a positive real number.*

A European option with exercise time T *is an* $\mathcal{F}_T$*-measurable random variable of the form*

$$h(S_T), \tag{4.6}$$

where $h : \mathbb{R}_+ \to \mathbb{R}_+$ *is a Borel measurable function.*

The function h is called the payoff (function) of the European option.

Examples 4.6. *Recall the following notation* $x^+ = \max\{x, 0\}$*, i.e.*

$$x^+ = \begin{cases} x, & \text{if } x > 0, \\ 0, & \text{if } x \leq 0. \end{cases}$$

(a) *For* $K > 0$ *consider* $g(s) = (s - K)^+$*,* $s > 0$*.*

(b) *For* $K > 0$ *consider* $g(s) = (K - s)^+$, $s > 0$.

Exercise 4.7. Draw graphs for the functions in the above examples.

Remark 4.8. (a) *A European call option with exercise time* $T > 0$ *and strike* $K > 0$ *is the right (but not the obligation) to buy at time* T *a share at price* K *(called the strike). If* $S(T) > K$, *the option holder will exercise the option and buy a share at strike price* K. *He can sell it immediately afterwards for* $S(T)$ *making a net profit of*

$$S(T) - K.$$

On the other hand, if $S(T) \leq K$ *the holder will not exercise the option and his/her profit is* 0. *Hence, the payoff for the holder of a European call option with exercise time* $T > 0$ *and strike* $K > 0$ *is equal to*

$$\big(S(T) - K\big)^+ = g\big(S(T)\big)$$

where g *is as in part (a) of Example 4.6.*

(b) *A European put option with exercise time* $T > 0$ *and strike* $K > 0$ *is the right (but not the obligation) to sell at time* T *a share at price* K. *If* $S(T) < K$, *the option holder can buy the share for* $S(T)$ *and exercise the option selling the share immediately afterwards at price* K, *making a net profit of*

$$K - S(T).$$

On the other hand, if $S(T) \geq K$ *the option holder does not exercise the option and her/his profit is* 0. *Hence, the payoff for the holder of a European put option with exercise time* $T > 0$ *and strike* $K > 0$ *is equal to*

$$\big(K - S(T)\big)^+ = g\big(S(T)\big)$$

where g *is as in part (b) of Example 4.6.*

4.2. **Replication.** Consider a European option with exercise time T and payoff $h\left(S_T\right)$. Suppose there exists a self-financing trading strategy (φ, ψ) with wealth process X such that $X_T = h\left(S_T\right)$.

Definition 4.9. *We call a self-financing trading strategy (φ, ψ) with $X_T = h(S_T)$ a replicating strategy for the option with payoff $h\left(S_T\right)$.*

We will now show that under these assumptions the initial value X_0 of the wealth process associated to the replicating strategy is the unique no-arbitrage price of the option. If the price of the option is higher we can sell the option, set up the replicating strategy and invest the remainder in the risk-free bond creating arbitrage at time T (the wealth process of the replicating strategy and the pay off for the sold option cancel each other out at time T). Conversely, if the price of the option is smaller than X_0 we can buy the option, go short in the replicating strategy (sell it) and invest the positive difference in risk free bonds, once again creating arbitrage at time T. We conclude that if we assume the existence of a replicating strategy the unique no-arbitrage price of the option at time 0 is $X_{0.}$ In general: If two assets have identical cash flows in the future they must have the same value now.

Thus, to complete this argument to price a European option using the no-arbitrage principle two problems remain to be solved. First, we have to show the existence of the replicating strategy. Second, in order to apply the above arguments, we must find the processes φ, ψ giving rise to this replicating strategy.

In order to tackle the second question we make the following ansatz: We suppose the value of the option at time t is a function $C\left(t, S_t\right)$, where C is an unknown function of time t and the stock price S (note this in particular implies $C\left(T, S_T\right) = h(S_T)$). In this case we are looking for a self-financing trading strategy such that

$$X(t) = C(t, S(t)), \quad t \in [0, T]. \tag{4.7}$$

Note that though formula (4.7) suggests that $X(t)$ is independent of the bond price $B(t)$, this is not the case. To see this we note that $X(t)$ depends on the time t and t in turn is determined by $B(t)$ assuming one knows $B(0)$.

4.3. **Derivation of the Black-Scholes equation.** The following fundamental theorem shows that under appropriate assumptions the function $C(t, x)$ describing the value of a European option must necessarily satisfy the Black-Scholes equation.

Theorem 4.10. *In the Black-Scholes model consider a European option with payoff $h(S_T)$ at exercise time T. Suppose its value $C(t, x)$ at time $t < T$ is a $C^{1,2}$ function of t and x (by this we mean that $C(t, S_t)$ is the value of the option at time t,in particular $C(T, S_T) = h(S_T)$). Under these assumption the following two claims hold:*

(1) If (φ, ψ) is a self-financing strategy with wealth process $X_t = \psi_t B_t + \varphi_t S_t$ such that

$$C(t, S_t) = X_t \tag{4.8}$$

then C solves the Black-Scholes partial differential equation (PDE)

$$\frac{\partial C}{\partial t}(t, x) + rx\frac{\partial C}{\partial x}(t, x) + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 C}{\partial x^2}(t, x) - rC(t, x) = 0, \quad t \in [0, T],\ x > 0. \tag{4.9}$$

with terminal condition

$$C(T, x) = h(x) \tag{4.10}$$

and the processes φ, ψ satisfy

$$\varphi_t = \frac{\partial C}{\partial x}(t, S_t), \quad \psi_t = \frac{C(t, S_t) - S_t \frac{\partial C}{\partial x}(t, S_t)}{B_t}, \quad T > t \geq 0. \tag{4.11}$$

(2) Conversely, if φ, ψ are defined as in (4.11), X is the wealth process corresponding to the trading strategy (φ, ψ) and C satisfies the Black-Scholes PDE (4.9) with terminal condition (4.10) then identity (4.8) holds and (φ, ψ) is a self-financing strategy.

Before we embark on the proof of the theorem we gather some important observations.

Remark 4.11. *(1) The theorem does not show that such a function $C(t,x)$ with the prescribed properties exists. If it exists, however, it will solve the Black-Scholes PDE with terminal condition $h(S_T)$.*

(2) Existence of the function $C(t,x)$ will be shown later by solving the Black-Scholes equation using probabilistic methods (this will involve using both the Girsanov theorem and the Feynman-Kac formula). An alternative way to solve the Black-Scholes using analytic methods is given in the following chapter as optional reading.

Proof. On the one hand applying the Ito formula in identity (4.8) we get

$$\begin{aligned} dX_t &= \left(\frac{\partial C}{\partial t}(t,S_t) + \mu S_t \frac{\partial C}{\partial x}(t,S_t) + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 C}{\partial x^2}(t,S_t)\right) dt \\ &\quad + \sigma S_t \frac{\partial C}{\partial x}(t,S_t)\, dW_t. \end{aligned}$$

On the other hand (φ, ψ) is assumed to be self-financing and therefore

$$\begin{aligned} dX_t &= \varphi_t dS_t + \psi_t dB_t \\ &= \varphi_t(\mu S_t dt + \sigma S_t dW_t) + \psi_t r B_t dt \\ &= (\mu \varphi_t S_t + r\psi_t B_t)\, dt + \varphi_t \sigma S_t dW_t. \end{aligned}$$

Equating the drift and diffusion coefficients in these two equations we get

$$\varphi_t = \frac{\partial C}{\partial x}(t,S_t) \tag{4.12}$$

and

$$\mu \varphi_t S_t + r\psi_t B_t = \frac{\partial C}{\partial t}(t,S_t) + \mu S_t \frac{\partial C}{\partial x}(t,S_t) + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 C}{\partial x^2}(t,S_t). \tag{4.13}$$

Substituting (4.12) into (4.13) and subtracting $\mu S_t \frac{\partial C}{\partial x}(t, S_t)$ and finally dividing by rB_t we have

$$\psi_t = \frac{\frac{\partial C}{\partial t}(t, S_t) + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 C}{\partial x^2}(t, S_t)}{rB_t}. \tag{4.14}$$

Plugging both φ_t and ψ_t into (4.8) we have

$$C(t, S_t) = S_t \frac{\partial C}{\partial x}(t, S_t) + \frac{\frac{\partial C}{\partial t}(t, S_t) + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 C}{\partial x^2}(t, S_t)}{r}$$

which implies that C satisfies the Black-Scholes PDE (4.9) (strictly speaking using that S_t has a continuous positive density for $t > 0$). Similarly, since $C(T, S_T) = h(S_T)$ the terminal condition $C(T, x) = h(x)$ holds. The claim for ψ follows from (4.14) and the fact that C satisfies the Black-Scholes equation.

For the converse part we first note that the wealth process corresponding to (φ, ψ) satisfies

$$\begin{aligned} X_t &= \frac{\partial C}{\partial x}(t, S_t) S_t + \frac{C(t, S_t) - S_t \frac{\partial C}{\partial x}(t, S_t)}{B_t} B_t \\ &= C(t, S_t), \end{aligned}$$

To show that the self-financing property holds we note that

$$\begin{aligned} \int_0^t \varphi_s dS_s + \int_0^t \psi_s dB_s &= \int_0^t \frac{\partial C}{\partial x}(s, S_s) dS_s \\ &\quad + \int_0^t \frac{C(s, S_s) - S_s \frac{\partial C}{\partial x}(s, S_s)}{B_s} rB_s ds \\ &= \int_0^t \frac{\partial C}{\partial x}(s, S_s) dS_s \\ &\quad + \int_0^t \left(\frac{\partial C}{\partial s}(s, S_s) + \frac{1}{2}\sigma^2 S_s^2 \frac{\partial^2 C}{\partial x^2}(s, S_s) \right) ds \\ &= C(t, S_t) - C(0, S_0), \end{aligned}$$

where we have used that C satisfies the Black-Scholes PDE in the second step and the Ito formula applied to $C(t, S_t)$ to derive the final equality. $\square$

5. Analytic Derivation of Black-Scholes formula

Chapter 5 is left for self-reading and is not examinable.

In this chapter we show how to find the solution of the problem (4.9)-(4.10) by methods of theory of partial differential equations.

Suppose that $\sigma > 0$, $r > 0$ and $T > 0$ are fixed constants and $g(S)$, $S > 0$ is a given function. The Black-Scholes equation is a backward parabolic partial differential equation given below

$$\begin{cases} v_t(t,S) + rSv_S(t,S) + \frac{1}{2}\sigma^2 S^2 v_{SS}(t,S) = rv(t,S), \ t < T, S > 0, \\ v(T,S) = g(S), \quad S > 0, \end{cases} \tag{5.1}$$

where

$$v_t(t,S) = \frac{\partial v(t,S)}{\partial t}, \quad v_S(t,S) = \frac{\partial v(t,S)}{\partial S}$$

and

$$v_{SS}(t,S) = \frac{\partial^2 v(t,S)}{\partial S^2}.$$

Step I. We define a new function $z(t,S)$, defined for $t < T$ and $S > 0$ by the formula

$$v(t,S) = e^{rt}e^{-rT}z(t,S)$$

or equivalently

$$z(t,S) = e^{-rt}e^{rT}v(t,S)$$

Observe that the function v has partial derivatives iff the function z has the appropriate partial derivatives. Moreover,

$$\begin{aligned} v_t(t,S) &= rv(t,S) + e^{rt}e^{-rT}z_t(t,S) = rv(t,S) + e^{r(t-T)}z_t(t,S), \\ v_S(t,S) &= e^{rt}e^{-rT}z_S(t,S) = e^{r(t-T)}z_S(t,S), \\ v_{SS}(t,S) &= e^{rt}e^{-rT}z_{SS}(t,S) = e^{r(t-T)}z_{SS}(t,S), \end{aligned}$$

Inserting these into first equation 5.1 we get (note we do not change the expression on the RHS)

$$\Big\{ rv(t,S) + e^{r(t-T)}z_t(t,S) + rSe^{r(t-T)}z_S(t,S) + \frac{1}{2}\sigma^2 S^2 e^{r(t-T)}z_{SS}(t,S) = rv(t,S),$$

We see that the term $rv(t,S)$ is on both sides, hence it cancels. We get

$$\left\{ e^{r(t-T)} z_t(t,S) + rSe^{r(t-T)} z_S(t,S) + \frac{1}{2}\sigma^2 S^2 e^{r(t-T)} z_{SS}(t,S) = 0. \right.$$

Now, each term contains the same factor $e^{r(t-T)}$. Since this factor is always > 0 and hence always $\neq 0$, we can divide by it and get

$$\begin{cases} z_t(t,S) + rSz_S(t,S) + \frac{1}{2}\sigma^2 S^2 z_{SS}(t,S) = 0, \ t < T, S > 0, \\ z(T,S) = v(T,S) = g(S), \quad S > 0, \end{cases} \tag{5.2}$$

Step II. The variable S above is a positive number. The coefficients in the equation 5.2 are variable, i.e. depending on S. In order to make them constant we introduce a new variable $x = \ln S \in \mathbb{R}$ and new function $p(t,x)$ defined for $t < T$ and $x \in \mathbb{R}$ by the following formula

$$\begin{cases} p(t,x) = z(t,e^x), \quad x \in \mathbb{R} \\ z(t,S) = p(t,\ln S), \quad S > 0 \end{cases} \tag{5.3}$$

As in Step I p has partial derivatives iff z has corresponding partial derivatives and

$$\begin{aligned} z_t(t,S) &= p_t(t,\ln S) = p_t(t,x) \\ z_S(t,S) &= p_x(t,\ln S)\frac{\partial \ln S}{\partial S} = p_x(t,\ln S)\frac{1}{S} = p_x(t,x)\frac{1}{S} \\ z_{SS}(t,S) &= \frac{\partial}{\partial S}\left[p_x(t,\ln S)\frac{1}{S}\right] = p_{xx}(t,\ln S)\frac{\partial \ln S}{\partial S}\frac{1}{S} + p_x(t,\ln S)\frac{\partial \frac{1}{S}}{\partial S} \\ &= p_{xx}(t,\ln S)\frac{1}{S^2} - p_x(t,\ln S)\frac{1}{S^2} = p_{xx}(t,x)\frac{1}{S^2} - p_x(t,x)\frac{1}{S^2} \end{aligned}$$

Inserting these formuale into the first equation 5.2 we get

$$p_t(t,x) + rSp_x(t,x)\frac{1}{S} + \frac{1}{2}\sigma^2 S^2 \left[p_{xx}(t,x)\frac{1}{S^2} - p_x(t,x)\frac{1}{S^2}\right] = 0$$

Simplifying the factors and then multiplying the bracket out we get

$$p_t(t,x) + rp_x(t,x) + \frac{1}{2}\sigma^2 p_{xx}(t,x) - \frac{1}{2}\sigma^2 p_x(t,x) = 0$$

Since also $p(T,x) = z(T,e^x) = g(e^x)$ we get

$$
\begin{cases} p_t(t,x) + \left(r - \frac{1}{2}\sigma^2\right) p_x(t,x) + \frac{1}{2}\sigma^2 p_{xx}(t,x) = 0, \quad t < T; \quad x \in \mathbb{R} \\ p(T,x) = g(e^x), \quad x \in \mathbb{R} \end{cases} \tag{5.4}
$$

Step III. Now we introduce still another function denoted by $q(t,y)$, $t < T$, $y \in \mathbb{R}$ defined by formulae

$$
\begin{aligned}
p(t,x) &= q\left(t, x - (r - \frac{1}{2}\sigma^2)(t-T)\right), \\
y &= x - (r - \frac{1}{2}\sigma^2)(t-T), \\
x &= y + r - \frac{1}{2}\sigma^2(t-T), \\
q(t,y) &= p\left(t, y + (r - \frac{1}{2}\sigma^2)(t-T)\right).
\end{aligned}
$$

Assuming that the partial derivatives of q are known we can calculate the corresponding partial derivatives of p by

$$
\begin{aligned}
p_t(t,x) &= q_t(t,y) + q_y(t,y)\frac{\partial}{\partial t}\left[x - (r - \frac{1}{2}\sigma^2)(t-T)\right] \\
&= q_t(t,y) - (r - \frac{1}{2}\sigma^2)q_y(t,y) \\
p_x(t,x) &= q_y(t,y)\frac{\partial}{\partial x}\left[x - (r - \frac{1}{2}\sigma^2)(t-T)\right] = q_y(t,y) \\
p_{xx}(t,x) &= \frac{\partial p_x(t,x)}{\partial x} = \frac{\partial q_y(t,y)}{\partial x} = q_{yy}(t,y)\frac{\partial}{\partial x}\left[x - (r - \frac{1}{2}\sigma^2)(t-T)\right] = q_{yy}(t,y)
\end{aligned}
$$

Inserting the above formulae into the equation (5.4) we get

$$
q_t(t,y) - (r - \frac{1}{2}\sigma^2)q_y(t,y) + \left(r - \frac{1}{2}\sigma^2\right) q_y(t,y) + \frac{1}{2}\sigma^2 q_{yy}(t,y) = 0,
$$

Obviously the terms with $q_y(t,y)$ cancel each other and since $q(T,y) = p(T,y) = g(e^y)$ we get

$$
\begin{cases} q_t(t,y) + \frac{1}{2}\sigma^2 q_{yy}(t,y) = 0, \quad t < T; \quad y \in \mathbb{R} \\ q(T,y) = g(e^y), \quad y \in \mathbb{R} \end{cases} \tag{5.5}
$$

We proved

Theorem 5.1. *Suppose that the function $v(t,S)$, $t \leq T$, $S > 0$ solves the Black-Scholes equation (5.1), then the function $q(t,y)$, $t \leq T$, $y \in \mathbb{R}$ defined by*

$$q(t,y) = e^{r(T-t)} v(t, e^{y+(r-\frac{\sigma^2}{2})(t-T)}), \quad t \leq T, \ \ y \in \mathbb{R} \tag{$*$}$$

solves the problem (5.5).

As I have said before we have proved the above Theorem before we formulated it. However it could be more convincing if we give another, a direct proof of it. Below however, we shall prove that if the functions $v(t,S)$ and $q(t,y)$ are related by the formula $(*)$ and if q solves (5.5) then v solves (5.1). Se also Example below.

Before we do this let us first formulate

Example . Verify that the function $q(t,y)$ defined by formula $(*)$ indeed satisfies equation (5.5).

Proof. Using formula $(*)$ we will first calculate the partial derivatives of the function $q(t,y)$ in terms of the partial derivatives of the function $v(t,S)$. We have, using a shortcut notation $S = e^{y+(r-\frac{\sigma^2}{2})(t-T)}$,

$$\begin{aligned}
q_t(t,y) &= -rq(t,y) + e^{r(T-t)} \left[v_t(t,S) + v_S(t,S) \frac{\partial}{\partial t} \big(e^{y+(r-\frac{\sigma^2}{2})(t-T)} \big) \right] \\
&= -rq(t,y) + e^{r(T-t)} \left[v_t(t,S) + v_S(t,S)(r - \frac{\sigma^2}{2})\, e^{y+(r-\frac{\sigma^2}{2})(t-T)} \right] \\
q_y(t,y) &= e^{r(T-t)} v_S(t,S) \frac{\partial}{\partial y} \big(e^{y+(r-\frac{\sigma^2}{2})(t-T)} \big) \\
&= e^{r(T-t)} v_S(t,S) e^{y+(r-\frac{\sigma^2}{2})(t-T)} \\
q_{yy}(t,y) &= \frac{\partial q_y(t,y)}{\partial y} = e^{r(T-t)} v_{SS}(t,S) \left[e^{y+(r-\frac{\sigma^2}{2})(t-T)} \right]^2 + e^{r(T-t)} v_S(t,S) e^{y+(r-\frac{\sigma^2}{2})(t-T)}
\end{aligned}$$

Recalling that $S = e^{y+(r-\frac{\sigma^2}{2})(t-T)}$ and using $(*)$ we can rewrite the last formulae in a more compact way

$$
\begin{aligned}
q_t(t,y) &= -re^{r(T-t)}v(t,S) + e^{r(T-t)}\left[v_t(t,S) + (r-\frac{\sigma^2}{2})Sv_S(t,S)\right] \\
q_{yy}(t,y) &= e^{r(T-t)}S^2v_{SS}(t,S) + e^{r(T-t)}Sv_S(t,S)
\end{aligned}
$$

Inserting these formulae into the LHS of the first equation of (5.5) we get

$$
\begin{aligned}
\text{LHS of (5.5)} = -re^{r(T-t)}v(t,S) + e^{r(T-t)}\left[v_t(t,S) + (r-\frac{\sigma^2}{2})Sv_S(t,S)\right] \\
+\frac{1}{2}\sigma^2 e^{r(T-t)}\left[S^2v_{SS}(t,S) + e^{r(T-t)}Sv_S(t,S)\right].
\end{aligned}
$$

Dividing by $e^{r(T-t)}$ and using the first equation of (5.1) we get the first equation of (5.5).

We have to finish by showing that $q(T,y) = g(e^y)$. From $(*)$ we get for $t = T$ that $q(T,y) = v(T,e^y)$. Since by (5.1) we have that $v(T,S) = g(S)$ for all $S > 0$, we infer that $v(T,e^y) = g(e^y)$ for all $y \in \mathbb{R}$. Hence $q(T,y) = g(e^y)$ for all $y \in \mathbb{R}$. □

Example . Show converse to what I have just shown. In other words, prove that if the functions $v(t,S)$ and $q(t,y)$ are related by the formula $(*)$ and if q solves (5.5) then v solves (5.1).

Proof. Let $S = e^{y+(r-\frac{\sigma^2}{2})(t-T)}$,

$$
\begin{aligned}
y(t,S) &= \ln S - (r-\frac{\sigma^2}{2})(t-T), \\
y_S(t,S) &= \frac{1}{S}, \\
y_t(t,S) &= -(r-\frac{\sigma^2}{2}).
\end{aligned}
$$

So by formula $(*)$,

$$
\begin{aligned}
v(t,S) &= q\left(t, \ln S - (r-\frac{\sigma^2}{2})(t-T)\right)e^{r(t-T)}, \\
v_t &= \left(q_t - \frac{\sigma^2}{2}q_y\right)e^{r(t-T)} + rqe^{r(t-T)},
\end{aligned}
$$

$$v_S = \frac{1}{S} q_y e^{r(t-T)},$$

$$v_{SS} = -\frac{1}{S^2} q_y e^{r(t-T)} + \frac{1}{S^2} q_{yy} e^{r(t-T)},$$

Hence

$$\begin{aligned}
&v_t + rSv_S + \frac{1}{2}\sigma^2 S^2 v_{SS} \\
&= \left(q_t - \frac{\sigma^2}{2} q_y\right) e^{r(t-T)} + rqe^{r(t-T)} + rq_y e^{t(t-T)} - \frac{1}{2} q_y e^{r(t-T)} + \frac{1}{2} q_{yy} e^{r(t-T)} \\
&= e^{r(t-T)} \left(q_t + \frac{1}{2}\sigma^2 q_{yy} + rq\right) \\
&= \cdots
\end{aligned}$$

by the 1st equality in (5.5), $q_t + \frac{1}{2}\sigma^2 q_{yy} = 0$, therefore

$$\begin{aligned}
\cdots &= rqe^{r(t-T)} \\
&= rv
\end{aligned}$$

So we get the 1st equality in (5.1). And

$$v(T, S) = q(T, \ln S),$$

by the 2nd equality in (5.5),

$$q(T, \ln S) = g(e^{\ln S}) = g(S),$$

hence $v(T, S) = g(S)$, which is the 2nd equality in (5.1). This completes the proof.

$\square$

Step IV. We introduce another variable and a new function

$$\begin{aligned}
\tau &= T - t > 0, \quad t = T - \tau, \ \ \tau \geq 0, \ \ t \leq T; \\
u(\tau, y) &= q(T - \tau, y) \\
q(t, y) &= u(T - t, y)
\end{aligned}$$

We easily find that

$$\begin{aligned} q_t(t,y) &= -u_\tau(T-t,y) \\ q_y(t,y) &= u_y(T-t,y) \\ q_{yy}(t,y) &= u_{yy}(T-t,y) \end{aligned}$$

Substituting into problem (5.5) we get

$$\begin{cases} u_\tau(\tau,y) = \dfrac{1}{2}\sigma^2 u_{yy}(\tau,y), \quad \tau > 0; \;\; y \in \mathbb{R} \\ u(0,y) = g(e^y), \quad y \in \mathbb{R} \end{cases} \tag{5.6}$$

It is known that the unique solution to (5.6) is given by an explicit formula:

$$u(\tau,y) = \frac{1}{\sqrt{2\pi\,\tau}\,\sigma}\int_{-\infty}^{\infty} e^{-\frac{(y-x)^2}{2\sigma^2\,\tau}} g(e^x)\,dx, \quad \tau > 0, y \in \mathbb{R}. \tag{5.7}$$

Introducing a change of variables $z = x - y$ so that $x = y + z$ and $dx = dz$ and $z \in (-\infty, \infty)$ we get

$$u(\tau,y) = \frac{1}{\sqrt{2\pi\,\tau}\,\sigma}\int_{-\infty}^{\infty} e^{-\frac{z^2}{2\sigma^2\,\tau}} g(e^{y+z})\,dz$$

In order to get rid of the $\sigma^2\tau$ denominator in the exponent in the above formula we make one more change of variables:

$$x = \frac{z}{\sigma\sqrt{\tau}} \in (-\infty, \infty). \tag{5.8}$$

Since $dz = \sigma\sqrt{\tau}dx$ we infer that

$$u(\tau,y) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} g(e^{y+\sigma\sqrt{\tau}x})\,dx$$

Recalling the relationship between q and u we therefore have

$$q(t,y) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} g(e^{y+\sigma\sqrt{T-t}x})\,dx$$

On the other hand we can find an inverse formula to $(*)$. Putting $S = e^{y+(r-\frac{\sigma^2}{2})(t-T))}$ we get from $(*)$

$$e^{-r(T-t)}q(t,y) = v(t,S)$$

Since $y + (r - \frac{\sigma^2}{2})(t-T) = \ln S$ we get

$$y = \ln S - (r - \frac{\sigma^2}{2})(t-T)).$$

Hence,

$$v(t,S) = e^{-r(T-t)}q(t, \ln S - (r - \frac{\sigma^2}{2})(t-T)) \qquad (**)$$

Hence,

$$v(t,S) = \frac{e^{-r(T-t)}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} g(e^{\ln S-(r-\frac{\sigma^2}{2})(t-T)+\sigma\sqrt{T-t}\,x})\,dx$$

Since $e^{\ln S} = S$ we therefore get

$$v(t,S) = \frac{e^{-r(T-t)}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} g(Se^{(r-\frac{\sigma^2}{2})(T-t)+\sigma\sqrt{T-t}\,x})\,dx \tag{5.9}$$

We have therefore proven

Theorem 5.2. *The function $v(t,S)$, $t \le T$, $S > 0$ given by formula (5.9) above is a solution to the Black-Scholes equation (5.1).*

Proof. As it was mentioned just before the statement of this theorem, we have already proven it. Indeed, all our passages from one differential equation to another was equivalences so the initial value problem (5.6) is in fact equivalent to the Black-Scholes (backward parabolic) equation (5.1). Hence the unique solution to (5.6) can be transformed, passing steps from IV to I in the reversed order, to the unique solution of the Black-Scholes equation (5.1).

However there is one more way to prove this theorem, and we implement it here. First of all, we can directly calculate derivative of v and explicitly check that the equation and final state condition from (5.1) satisfied.

$$v(T,S) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} g(S) dx = g(S).$$

So we get the 2nd equation of (5.1).

Let $u = Se^{(r-\frac{\sigma^2}{2})(T-t)+\sigma\sqrt{T-t}x}$, then for $t < T$,

$$x = \frac{\ln u - \ln S - (r-\frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}, \quad dx = \frac{1}{u\sigma\sqrt{T-t}} du.$$

Hence

$$v(t,S) = \frac{e^{-r(T-t)}}{\sigma\sqrt{2\pi(T-t)}} \int_0^{\infty} e^{\frac{[\ln u - \ln S - (r-\frac{\sigma^2}{2})(T-t)]^2}{2\sigma^2(T-t)}} \frac{g(u)}{u} du.$$

Thus we can get

$$\begin{aligned} v_t = & \frac{e^{-r(T-t)r}}{\sigma\sqrt{2\pi(T-t)}} \left(r + \frac{1}{2(T-t)} \right) \int_0^{\infty} e^{\frac{[\ln u - \ln S - (r-\frac{\sigma^2}{2})(T-t)]^2}{2\sigma^2(T-t)}} \frac{g(u)}{u} du \\ & + \frac{e^{-r(T-t)r}}{\sigma\sqrt{2\pi(T-t)}} \int_0^{\infty} e^{\frac{[\ln u - \ln S - (r-\frac{\sigma^2}{2})(T-t)]^2}{2\sigma^2(T-t)}} \frac{g(u)}{u} \\ & \frac{-2[\ln u - \ln S - (r-\frac{\sigma^2}{2})(T-t)](r-\frac{\sigma^2}{2})(T-t) - [\ln u - \ln S - (r-\frac{\sigma^2}{2})(T-t)]^2}{2\sigma^2(T-t)^2} du \end{aligned}$$

$$v_s = \frac{e^{-r(T-t)}}{\sigma\sqrt{2\pi(T-t)}} \int_0^{\infty} e^{\frac{[\ln u - \ln S - (r-\frac{\sigma^2}{2})(T-t)]^2}{2\sigma^2(T-t)}} \frac{[\ln u - \ln S - (r-\frac{\sigma^2}{2})(T-t)]}{S\sigma^2(T-t)} \frac{g(u)}{u} du$$

$$\begin{aligned} v_{SS} = & \frac{e^{-r(T-t)}}{\sigma\sqrt{2\pi(T-t)}} \int_0^{\infty} e^{\frac{[\ln u - \ln S - (r-\frac{\sigma^2}{2})(T-t)]^2}{2\sigma^2(T-t)}} \left[\frac{[\ln u - \ln S - (r-\frac{\sigma^2}{2})(T-t)]^2}{S^2\sigma^4(T-t)^2} \right. \\ & \left. - \frac{1 + [\ln u - \ln S - (r-\frac{\sigma^2}{2})(T-t)]}{S^2\sigma^2(T-t)} \right] \frac{g(u)}{u} du \end{aligned}$$

Hence,

$$v_t + rSv_s + \frac{1}{2}\sigma^2 S^2 v_{SS} = rv,$$

so we get the first equation of (5.1).

This completes our proof. □

6. Derivation of the Black-Scholes formula using probabilistic methods

In this chapter we demonstrate how we can find the solution of the Black-Scholes equation (4.9)-(4.10) using probabilistic methods. Recall that in the Black-Scholes model the stock price satisfies

$$dS_t = \mu S_t dt + \sigma S_t dW.$$

Applying Theorem 3.13, i.e. the Feynman-Kac formula to the process S_t and the terminal condition g equation (3.12) becomes

$$\frac{\partial C}{\partial t}(t,x) + \mu x \frac{\partial C}{\partial x}(t,x) + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 C}{\partial x^2}(t,x) = 0, \tag{6.1}$$

$$C(T,x) = g(x). \tag{6.2}$$

Comparing the Black-Scholes PDE (4.9)-(4.10) with equation (6.1)-(6.2) we notice two significant differences. These differences stop us from immediately applying the Feynman-Kac formula to solve the B.-S. equation by means of solving an SDE. First, the coefficients of the $\frac{\partial v}{\partial x}(t,x)$ term differ: we have μx in (6.1) as opposed to rx in the Black-Scholes equation. Second, the function $rv(t,x)$ is absent in (6.1) .

To address the first difference, let $T > 0$ and consider the process

$$V_t := W_t + \frac{\mu - r}{\sigma} t, \quad t \in [0,T]$$

By a previous Example we may apply the Girsanov theorem and there exists a measure $\mathbb{Q}_T$ on the space $(\Omega, \mathcal{F}_T)$, absolutely continuous with respect to $\mathbb{P}$, such that V_t is a Brownian Motion on the probability space $(\Omega, \mathcal{F}_T, \mathbb{Q}_T)$. In the following we work on the probability space, $(\Omega, \mathcal{F}_T, \mathbb{Q}_T)$. Then, V_t is a Brownian Motion with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$ and the stock price S_t satisfies the SDE

$$dS_t = rS_t dt + \sigma S_t dV_t. \tag{6.3}$$

Hence, after the change of measure to $\mathbb{Q}_T$ the PDE associated to (6.3) in the Feynman-Kac formula is given by

$$\frac{\partial C}{\partial t}(t,x) + rx\frac{\partial C}{\partial x}(t,x) + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 C}{\partial x^2}(t,x) = 0, \tag{6.4}$$

thus removing the first obstacle. We will revisit the measure $\mathbb{Q}_T$ in the following chapter in a more general setting, where it will be known as the risk-neutral measure.

To deal with the second obstacle, we consider a modification of the function $C(t,S)$. More, precisely, we define a new function[7] $z(t,S)$, for $t < T$ and $S > 0$ by letting

$$C(t,S) = e^{r(t-T)} z(t,S), \quad t < T, S > 0.$$

Equivalently, we have

$$z(t,S) = e^{r(T-t)} C(t,S)$$

Observe, that the function C has partial derivatives if and only if the function z has the corresponding partial derivatives and

$$\begin{aligned} C_t(t,S) &= rC(t,S) + e^{r(t-T)} z_t(t,S), \\ C_S(t,S) &= e^{r(t-T)} z_S(t,S), \\ C_{SS}(t,S) &= e^{r(t-T)} z_{SS}(t,S), \end{aligned}$$

Inserting these expressions for the partial derivatives into the Black-Scholes equation (5.1) we get

$$rC(t,S) + e^{r(t-T)} z_t(t,S) + rSe^{r(t-T)} z_S(t,S) + \frac{1}{2}\sigma^2 S^2 e^{r(t-T)} z_{SS}(t,S) = rC(t,S),$$

Note, that we have not changed the expression on the right hand side. Observing that the term $rC(t,S)$ appears on both sides of the equation we may cancel it and obtain

$$e^{r(t-T)} z_t(t,S) + rSe^{r(t-T)} z_S(t,S) + \frac{1}{2}\sigma^2 S^2 e^{r(t-T)} z_{SS}(t,S) = 0.$$

[7]The motivation behind this definition will become more clear in the following chapter. $C(t,S)$ is obtained from $z(t,S)$ by discounting.

Each term in the equation contains the same factor $e^{r(t-T)}$. Since $e^{r(t-T)} > 0$ we may divide by it and get

(6.5)
$$z_t(t,S)+rSz_S(t,S)+\frac{1}{2}\sigma^2S^2z_{SS}(t,S)=0,\ t<T,S>0,z(T,S)=C(T,S)=g(S),\quad S>0.$$

Finally, we may apply Theorem 3.13 to the equation (6.5). We deduce that

$$z(t,S_t)=\mathbb{E}_{\mathbb{Q}_T}(g(S_T)|\mathcal{F}_t),$$

and consequently

$$C(t,S_t)=e^{-r(T-t)}\mathbb{E}_{\mathbb{Q}_T}(g(S_T)|\mathcal{F}_t). \tag{6.6}$$

In formula (6.6), we can take the filtration $(\mathcal{F}_t)_{t\geq 0}$ to be the natural filtration of the Brownian Motion $(V_t)_{t\geq 0}$. In particular, we may assume that $\mathcal{F}_0=\{\emptyset,\Omega\}$ and then for $t=0$ we have

$$C(0,S_0)=e^{-rT}\mathbb{E}_{\mathbb{Q}_T}(g(S(T))|\mathcal{F}_0)=e^{-rT}\mathbb{E}_{\mathbb{Q}_T}(g(S(T))).$$

For the general case $t\in[0,T)$ we consider the following example.

Example 6.1. *For $t\in[0,T]$ define two functions*

$$a(u,x)=\begin{cases}0, & \textit{if } u\in[0,t),\\ \mu x, & \textit{if } u\in[t,T].\end{cases}$$

$$b(u,x)=\begin{cases}0, & \textit{if } u\in[0,t),\\ \sigma x, & \textit{if } u\in[t,T].\end{cases}$$

Fix $S>0$ and consider the stochastic differential equation

$$d\tilde{S}_u=a(u,\tilde{S}_u)du+b(u,\tilde{S}_u)dV_u,\tilde{S}_0=S. \tag{6.7}$$

It is easy to verify that SDE (6.7) satisfies all conditions of Theorem (3.5). Denote by $\tilde{S}_t$ its solution. It is not difficult to check that the process $\tilde{S}_t$ remains constant and equal to S until time t and then starts to evolve as the stock price in the Black-Scholes equation.

The preceding example demonstrates that S is a Markov process, i.e. its future probabilities are determined by its most recent value. Putting everything together we have proved the following theorem.

Theorem 6.2. *The function $C(t,x)$ defined by the formula*

$$C(t,x) = e^{-r(T-t)}\mathbb{E}_{\mathbb{Q}_T}(g(S_T)|S_t = x) \tag{6.8}$$

is a solution to the problem (4.9)-(4.10).

We will now show how to obtain an explicit expression for the right hand side of identity (6.8). This will correspond to identity (5.9), derived in the previous chapter with analytic methods.

Theorem 6.3. *The function*

$$C(t,S) = \frac{e^{-r(T-t)}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} g(Se^{(r-\frac{\sigma^2}{2})(T-t)+\sigma\sqrt{T-t}y})\, dy \quad t \in [0,T],\ S > 0$$

is a solution to the problem (4.9)-(4.10).

Proof. Setting $V_t := W_t + \frac{\mu - r}{\sigma}t$, $t \in [0,T]$ we recall that V is Brownian motion on $(\Omega, \mathcal{F}_T, \mathbb{Q}_T)$. We are working with the Black-Scholes model, so

$$\begin{aligned} S_t &= S_0 e^{(r-\frac{\sigma^2}{2})t+\sigma V_t}, \\ \frac{S_T}{S_t} &= e^{(r-\frac{\sigma^2}{2})(T-t)+\sigma(V_T-V_t)} \end{aligned}$$

and therefore

$$S_T = S_t e^{(r-\frac{\sigma^2}{2})(T-t)+\sigma(V_T-V_t)}, \quad t \in [0,T].$$

Now, from equation (6.8) we see that for $t \in [0,T]$ we have

$$\begin{aligned} C(t,S) &= \mathbb{E}_{\mathbb{Q}_T}\left(e^{-r(T-t)}g(S_T)\big|S_t = S\right) \\ &= \mathbb{E}_{\mathbb{Q}_T}\left(e^{-r(T-t)}g(S_t e^{(r-\frac{\sigma^2}{2})(T-t)+\sigma(V_T-V_t)})\big|S_t = S\right) \end{aligned}$$

But, V is Brownian motion on $(\Omega, \mathcal{F}_T, \mathbb{Q}_T)$ and therefore $V_T - V_t$ is independent of $\mathcal{F}_t$ and

$$C\,(T,S) = \mathbb{E}_{\mathbb{Q}_T}\left(e^{-r(T-t)}g(Se^{(r-\frac{\sigma^2}{2})(T-t)+\sigma(V_T-V_t)})\right)$$

Since $V_T - V_t$ has density $p_{T-t}(x) = (2\pi(T-t))^{-1/2}e^{-\frac{x^2}{2(T-t)}}$ this expression simplifies further and we get

$$C(t,S) = \int_{-\infty}^{\infty} e^{-r(T-t)}g(Se^{(r-\frac{\sigma^2}{2})(T-t)}e^{\sigma x})(2\pi(T-t))^{-1/2}e^{-\frac{x^2}{2(T-t)}}\,dx. \tag{6.9}$$

Introducing a change of variables

$$\frac{x}{\sqrt{T-t}} = y$$

so that $x = \sqrt{T-t}y$ and $dx = \sqrt{T-t}\,dy$ we deduce that

$$C(t,S) = \frac{e^{-r(T-t)}}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}g(Se^{(r-\frac{\sigma^2}{2})(T-t)+\sigma\sqrt{T-t}y})\,dy,\ \ t\in[0,T]. \tag{6.10}$$

This last formula has been previously found by analytic methods, see (5.9).

□

We now consider the special case of a European call option with exercise time T and strike $K>0$, i.e. payoff function $g(S) = (S-K)^+$, $S>0$. Since $g(S)=0$ for $S<K$ the integrand in (5.9) vanishes unless

$$Se^{(r-\frac{\sigma^2}{2})(T-t)+\sigma\sqrt{T-t}\,x} > K \tag{6.11}$$

Dividing by S and taking logarithms we get

$$(r-\frac{\sigma^2}{2})(T-t)+\sigma\sqrt{T-t}\,x > \ln(\frac{K}{S})$$

or equivalently

$$x > \frac{\ln(\frac{K}{S}) - (r-\frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}} =: A. \tag{6.12}$$

Hence, since for $S > K$, $g(S) = S - K$, we have by linearity of the integral

$$\begin{aligned} C(t,S) &= \frac{e^{-r(T-t)}}{\sqrt{2\pi}} \int_A^\infty e^{-\frac{x^2}{2}} S e^{(r-\frac{\sigma^2}{2})(T-t)+\sigma\sqrt{T-t}\,x}\, dx - K\frac{e^{-r(T-t)}}{\sqrt{2\pi}} \int_A^\infty e^{-\frac{x^2}{2}}\, dx. \\ &= S e^{(r-\frac{\sigma^2}{2})(T-t)} \frac{e^{-r(T-t)}}{\sqrt{2\pi}} \int_A^\infty e^{-\frac{x^2}{2}} e^{\sigma\sqrt{T-t}\,x}\, dx - K\frac{e^{-r(T-t)}}{\sqrt{2\pi}} \int_A^\infty e^{-\frac{x^2}{2}}\, dx. \end{aligned} \tag{6.13}$$

For the first integral on the right hand side of (6.13) we have

$$\begin{aligned} \int_A^\infty e^{-\frac{x^2}{2}} e^{\sigma\sqrt{T-t}\,x}\, dx &= e^{\frac{\sigma^2}{2}(T-t)} \int_A^\infty e^{-\frac{1}{2}\left[x^2-2\sigma\sqrt{T-t}\,x+\sigma^2(T-t)\right]}\, dx \\ &= e^{\frac{\sigma^2}{2}(T-t)} \int_A^\infty e^{-\frac{1}{2}\left[x-\sigma\sqrt{T-t}\right]^2}\, dx \\ &= e^{\frac{\sigma^2}{2}(T-t)} \int_{A-\sigma\sqrt{T-t}}^\infty e^{-\frac{1}{2}z^2}\, dz \\ &= e^{\frac{\sigma^2}{2}(T-t)} \int_{-\infty}^{-A+\sigma\sqrt{T-t}} e^{-\frac{x^2}{2}}\, dx \end{aligned}$$

By symmetry of the integrand we have $\int_A^\infty e^{-\frac{x^2}{2}}\, dx = \int_{-\infty}^{-A} e^{-\frac{x^2}{2}}\, dx$ and we get the following formula for $C(t,S)$:

$$\begin{aligned} C(t,S) &= S e^{(r-\frac{\sigma^2}{2})(T-t)} \frac{e^{-r(T-t)}}{\sqrt{2\pi}} e^{\frac{\sigma^2}{2}(T-t)} \int_{-\infty}^{-A+\sigma\sqrt{T-t}} e^{-\frac{x^2}{2}}\, dx \\ &- K\frac{e^{-r(T-t)}}{\sqrt{2\pi}} \int_{-\infty}^{-A} e^{-\frac{x^2}{2}}\, dx \\ &= S\Phi(-A+\sigma\sqrt{T-t}) - Ke^{-r(T-t)}\Phi(-A) \end{aligned}$$

where $\Phi(t)$ is the probability distribution function of a standard normal random variable, i.e.

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{x^2}{2}}\, dx, \quad t \in \mathbb{R}.$$

We are ready to state the Black-Scholes formula for the price of a European call option.

Theorem 6.4 (Black-Scholes formula for a European call)**.** *The solution to the Black-Scholes PDE (5.3) with terminal value* $g(s) = (S-K)^+$, $S > 0$ *is given by the formula*

$$C(t,S) = S\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2),$$

where

$$\begin{aligned} d_1 &=: \frac{(r+\frac{\sigma^2}{2})(T-t)+\ln(\frac{S}{K})}{\sigma\sqrt{T-t}} \\ d_2 &:= d_1 - \sigma\sqrt{T-t} \end{aligned}$$

In particular,

$$C(0,S) = S\Phi(d_1) - Ke^{-rT}\Phi(d_2),$$

Proof. By the preceding calculations

$$C(t,S) = S\Phi(-A+\sigma\sqrt{T-t}) - Ke^{-r(T-t)}\Phi(-A),$$

where

$$-A = \frac{\ln(\frac{S}{K})+(r-\frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}} = d_2$$

and

$$-A+\sigma\sqrt{T-t} = \frac{(r+\frac{\sigma^2}{2})(T-t)+\ln(\frac{S}{K})}{\sigma\sqrt{T-t}} = d_1.$$

□

Remark 6.5. *Note that both d_1 and d_2 depend on t and on S.*

The following theorem allows us to deduce the price of a European put from the price of a European call option (and vice versa).

Theorem 6.6 (Put-Call parity)**.** *Let $T > 0$ and let $C(t)$, $t \in [0,T]$, be the value of a call option with strike price K and maturity time T. Let $P(t)$ be the value of a put option with the same strike and maturity. We denote by $S(t)$ the price of the share at time t, and r the risk-free interest rate. Then*

$$C(t) - P(t) = S(t) - K\mathrm{e}^{-r(T-t)}. \tag{6.14}$$

Proof. Suppose that at time $t = 0$ we buy a call option and sell a put option with the same maturity time T and strike price K. At time T, the payoff for this *portfolio* is $S(T) - K$, as

$$C(T) - P(T) = [S(T) - K]^+ - [K - S(T)]^+ = S(T) - K.$$

Also at time 0, we can can buy a unit of share and borrow $M := K\mathrm{e}^{-rT}$ using bonds (hence, at time T we have to return exactly $\mathrm{e}^{rT}M = \mathrm{e}^{rT}K\mathrm{e}^{-rT} = K$). The payoff for this second *portfolio* at time T is also $S(T) - K$. As we assume the *no arbitrage* principle we conclude that the prices of these two portfolios have to be the same for every $t \in [0, T]$, that is (6.14). □

Remark 6.7. *The Put-Call parity is* ***model-independent*** *in the sense that it does not rely on the particular model for the stock price we have chosen in the Black-Scholes model. Hence, it can be used to relate the prices of European put and call options in models far more general than discussed here.*

Exercise 6.8. Use the Put-Call parity theorem to prove the Black-Scholes formula for a European put option, i.e. show that

$$P\,(t, S) = Ke^{-r(T-t)}\Phi(-d_2) - S\Phi(-d_1).$$

7. Risk neutral pricing

The aim of this chapter is to show the existence of a replicating portfolio for a more general, possibly "path dependent", option h and compute the wealth process for this replicating portfolio. The initial value of the wealth process of the replicating portfolio is then, of course, the fair (no-arbitrage) price of the option. For European options we solved this problem in the previous chapters by first deriving and then solving Black-Scholes PDE. The results in this section are an alternative approach to this problem (not using the Black-Scholes PDE) that works for a broader range of options.

The following theorem allows us to find an absolutely continuous change of measure under which the discounted stock price becomes a martingale.

Theorem 7.1. *Consider the Black-Scholes model of a financial market. Recall that the stock price and bond were defined on the probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *as solutions to the equations*

$$\begin{cases} dS_t = \mu S_t dt + \sigma S_t dW_t \\ dB_t = rB_t dt \end{cases} \tag{7.1}$$

where $(W_t)_{t\geqslant 0}$ *is a Brownian motion with respect to a filtration* $(\mathcal{F}_t)_{t\geqslant 0}$.
Let $T > 0$. *Define the discounted stock price process by*

$$\tilde{S}_t := e^{-rt} S_t, \ t \geqslant 0. \tag{7.2}$$

Then, there exists a probability measure $\mathbb{Q}_T$ *such that* $\mathbb{Q}_T \ll \mathbb{P}$ *and* $(\tilde{S}_t)_{t\in[0,T]}$ *is a martingale on* $(\Omega, \mathcal{F}_T, \mathbb{Q}_T)$ *with respect to the filtration* $(\mathcal{F}_t)_{t\in[0,T]}$.
In particular,

$$\mathbb{E}_{\mathbb{Q}_T}(\tilde{S}_t) = \mathbb{E}_{\mathbb{Q}_T}(\tilde{S}_0) = S_0, \ t \in [0, T].$$

Remark 7.2. *The process* $S = (S_t)_{t\geq 0}$ *a martingale on* $(\Omega, F, \mathbb{P})$ *if and only if* $\mu = 0$. *To see that* S *cannot be a martingale if* $\mu \neq 0$ *note that if* S *were a martingale we would have*

$$\mathbb{E}S_t = \mathbb{E}S_0.$$

However,

$$\begin{aligned} \mathbb{E}S_t &= S_0 \mathbb{E} e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t} \\ &= S_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \frac{1}{2}\sigma^2 t} \\ &= S_0 e^{\mu t}. \end{aligned}$$

For the converse note that if $\mu = 0$ *we have*

$$dS_t = \sigma S_t dW_{t.}$$

The right hand side of this SDE is a stochastic integral with integrand $\sigma S \in M^2_{loc}(0,\infty)$, *which is a martingale.*

Proof of Theorem 7.1. Let $F(t,S) = e^{-rt}S$ then

$$\frac{\partial F(t,S)}{\partial t} = -re^{-rt}S, \quad \frac{\partial F}{\partial S} = e^{-rt} \quad \frac{\partial^2 F}{\partial S^2} = 0.$$

Since (S_t) is an Itô process, we can apply Itô lemma and get

$$\begin{aligned} \tilde{S}_t = F(t,S_t) &= S_0 + \int_0^t [-re^{-ru}S_u + e^{-ru}\mu S_u]du + \int_0^t \frac{\partial F}{\partial S}(u,S_u)\sigma S_u dW_u \\ &= S_0 + (\mu - r)\int_0^t \tilde{S}_u du + \sigma \int_0^t \tilde{S}_u dW_u \end{aligned} \tag{7.3}$$

and we deduce that $\tilde{S}_t$ solves the SDE

$$d\tilde{S}_t = (\mu - r)\tilde{S}_t dt + \sigma \tilde{S}_t dW(t). \tag{7.4}$$

Define the process V by

$$V_t = \frac{\mu - r}{\sigma}t + W_t, \ t \in [0,T] \tag{7.5}$$

and note that

$$\frac{\mu - r}{\sigma}t = \int_0^t \frac{\mu - r}{\sigma}du.$$

By Example 3.22 (an application of Girsanov's theorem), the process $(V_t)_{t\in[0,T)}$ is a Brownian motion on $(\Omega, \mathcal{F}_T, \mathbb{Q}_T)$ with respect to the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ where $\mathbb{Q}_T \ll \mathbb{P}$ has Radon-Nikodym derivative

$$\frac{d\mathbb{Q}_T}{d\mathbb{P}} = L_T = e^{-\frac{\mu-r}{\sigma}W_T - \frac{1}{2}\left(\frac{\mu-r}{\sigma}\right)^2 T}.$$

We have

$$dV_t = \frac{\mu - r}{\sigma}dt + dW_t$$

and, hence, we see from (7.4) that

$$\begin{aligned} d\tilde{S}_t &= \sigma\tilde{S}_t[\frac{(\mu - r)}{\sigma}dt + dW_t] \\ &= \sigma\tilde{S}_t dV_t \end{aligned}$$

and

$$\tilde{S}_t = S_0 + \int_0^t \sigma \tilde{S}_\tau dV_\tau. \tag{7.6}$$

Note that $\int_0^t \sigma \tilde{S}_\tau dV_\tau$ is an Itô integral with respect to the Brownian motion $(V_t)_{t\in[0,T]}$ on the probability space $(\Omega, \mathcal{F}_T, \mathbb{Q}_T)$ and the filtration $(\mathcal{F}_t)_{t\in[0,T]}$. Moreover, $\sigma\tilde{S}_\tau \in M^2_{loc}(0,\infty)$. Hence, $(\tilde{S}_t)_{t\in[0.T]}$ is a martingale on $(\Omega, \mathcal{F}_T, \mathbb{Q}_T)$ with respect to $(\mathcal{F}_t)_{t\in[0.T]}$.

□

The measure $\mathbb{Q}_T$ is an example of a risk-neutral measure and (as we will see) plays a crucial role in the risk-neutral approach to option valuation.

Definition 7.3. *The risk neutral measure (sometimes also called equivalent martingale measure)* $\mathbb{Q}_T$ *in the Black-Scholes model is the unique* P*-equivalent measure under which the stock price discounted by the risk free interest rate is a martingale. We have (restricting the measures to* $\mathcal{F}_T$*)*

$$\frac{d\mathbb{Q}_T}{dP} = \exp\left(-\frac{\mu - r}{\sigma} W_T - \frac{(\mu - r)^2}{2\sigma^2} T\right).$$

When working with the risk neutral measure the original measure P is usually referred to as the physical measure.

Remark 7.4. *Fixing a random element* $\omega \in \Omega$ *the Radon–Nikodym derivative* $\frac{d\mathbb{Q}_T}{dP}(\omega)$ *is small if and only if* W_T *is large. But,* W_T *is large if and only if* S_T *is large. Hence, we conclude that the risk neutral measure gives less weight than the original measure* P *to events where* S_T *is large.*

In order to avoid arbitrage we will work with the following class of admissible, self-financing strategies.

Definition 7.5. *A self-financing strategy* (φ_t, ψ_t) *is called admissible if the discounted wealth process* $\tilde{X}_t := \mathrm{e}^{-rt} X_t$ *satisfies* *(i)* $\tilde{X}_t \geq 0$ *a.s. for* $t \in [0, T)$.

(ii) $\mathbb{E}_{\mathbb{Q}_T}(\sup_{t\in[0,T]} \tilde{X}_t^2) < \infty$, *where* $\mathbb{Q}_T$ *is the risk neutral measure.*

The following equivalent condition for the self-financing property of a trading strategy was proved in the exercises.

Proposition 7.6. *Consider the Black-Scholes model of a financial market. A trading strategy* (ψ, φ) *is self-financing if and only if the wealth process* $X(t)$ *satisfies*

$$dX(t) = [rX(t) + (\mu - r)\varphi(t)S(t)]\, dt + \sigma\varphi(t)S(t)\, dW(t), \quad t \geq 0. \tag{7.7}$$

The risk neutral pricing approach allows us to consider options that are more general than the European options considered before.

Definition 7.7. *An option with expiry time* T *is a non-negative* $\mathcal{F}_T$ *- measurable random variable* h.

Example 7.8. *Consider the option* $h = g(S_T)$, *where* g *is a Borel measurable function. This example shows that - as expected - any European option with exercise time* $T > 0$ *is also an option.*

Definition 7.9. *An option* h *with expiry time* T *is called replicable if and only if there exists an admissible strategy* (φ_t, ψ_t) *such that*

$$X_T = h,$$

where $(X_t)_{t\geq 0}$ *is the wealth process of the strategy.*

The following theorem will later be required in order to show the existence of an admissible replication strategy for options in the Black-Scholes model.

Theorem 7.10 (Martingale Representation Theorem)**.** *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space,* $(W_t)_{t\geqslant 0}$ *a Brownian motion with respect to a filtration* $(\mathcal{F}_t)_{t\geqslant 0}$*. Moreover, let* $T > 0$ *and* $(M_t)_{t\geqslant 0}$ *be a martingale with respect to the filtration* $(\mathcal{F}_t)_{t\geqslant 0}$*, satisfying*

$$\mathbb{E}\int_0^T (M_t^2)dt < \infty.$$

Then there exists an $(\mathcal{F}_t)$*-adapted process* H *such that*

$$\mathbb{E}\int_0^T (H_t^2)dt < \infty$$

and

$$M_t = M_0 + \int_0^t H_s dW_s, \quad a.s., \ \forall t \in [0, T].$$

The theorem is usually proved by first showing the following proposition.

Proposition 7.11. *Suppose that* U *is an* $\mathcal{F}_T$*-measurable square integrable random variable, i.e. satisfying*

$$\mathbb{E}U^2 < \infty.$$

Then there exists a process H *satisfying*

$$\mathbb{E}\int_0^T (H_t^2)dt < \infty$$

and

$$U = \mathbb{E}(U) + \int_0^T H_t dW_t$$

The above proposition is, of course, also implied by Theorem 7.10. To see this apply the martingale representation theorem to the process $\mathbb{E}\left(U|\mathcal{F}_t\right)$, where $\mathcal{F}_t$ is as in the statement of Theorem 7.10.

In the following we work under the risk neutral measure, i.e. on the probability space $(\Omega\, , \mathcal{F}_T, \mathbb{Q}_T)$. The following lemma provides us with an equivalent characterisation of the self-financing property of a trading strategy in terms of the discounted wealth process.

Lemma 7.12. *Assume the Black-Scholes Model and recall that the discounted stock price process was defined by*

$$\tilde{S}_t := e^{-rt} S_t, \quad t \geq 0.$$

Let $(\varphi_t, \psi_t,)$ *be a trading strategy and recall that its wealth process is defined by*

$$X_t = \psi_t B_t + \phi_t S_t, \quad t \geq 0.$$

Define ***discounted wealth process*** *by*

$$\tilde{X}_t := e^{-rt} X_t, \quad t \geq 0.$$

Then the self-financing condition

$$dX_t = \psi_t dB_t + \phi_t dS_t, \quad t \geq 0, \tag{7.8}$$

is equivalent to the discounted wealth process satisfying

$$d\tilde{X} = \phi_t d\tilde{S}. \tag{7.9}$$

Proof. We work under the risk neutral measure $\mathbb{Q}_T$. Hence, the process $V_t = W_t + \frac{\mu - r}{\sigma} t$, $t \geq 0$, is a Brownian Motion on $(\Omega, \mathcal{F}_T, \mathbb{Q}_T)$. We have previously shown that on the probability space $(\Omega, \mathcal{F}, \mathbb{Q}_T)$ the processes S and B satisfy the equations

$$\begin{cases} dS &= rS(t)\,dt + \sigma S(t)\,dV_t, \quad t \geq 0, \\ dB &= rB(t)\,dt, \quad t \geq 0 \end{cases} \tag{7.10}$$

and for discounted stock price we have

$$d\tilde{S} = \sigma \tilde{S}_t dV_t. \tag{7.11}$$

We are going to apply the Itô Lemma with the function $f(t,x) = e^{-rt}x$ and the wealth process X. As we assume that (φ, ψ) is self-financing, we see from Proposition 7.6 that the wealth process satisfies

$$\begin{aligned} dX_t &= [rX_t + (\mu - r)\,\varphi_t S_t]\,dt + \sigma\varphi_t S_t dW_t \\ &= [rX_t + (\mu - r)\,\varphi_t S_t]\,dt + \sigma\varphi_t S_t \left(dV_t - \frac{\mu - r}{\sigma} dt\right) \\ &= rX_t dt + \sigma \varphi_t S_t dV_t. \qquad (7.12) \end{aligned}$$

Hence, X_t is an Itô process on $(\Omega, \mathcal{F}_T, \mathbb{Q}_T)$ with coefficients

$$a_t = rX_t$$

$$b_t = \sigma\phi_t S_t.$$

Using the Itô formula we see that

$$\tilde{X}_t = e^{-rt}X_t = f(t, X_t) = \tilde{X}_0 + \int_0^t \left[-re^{-ru}X_u + e^{-ru}rX_u + 0\right] du + \int_0^t e^{-ru}\sigma\phi_u S_u dV_u$$
$$= \tilde{X}_0 + \int_0^t e^{-ru}\sigma\phi_u S_u dV_u = \tilde{X}_0 + \int_0^t \phi_u\sigma\tilde{S}_u dV_u = \tilde{X}_0 + \int_0^t \phi_u d\tilde{S}_u,$$

where we have use (7.11) for the last equality.

For the converse we assume (7.9) and want to show that (7.8) holds. By (7.11) we see that

$$d\tilde{X}_t = \phi_t d\tilde{S}_t = e^{-rt}\phi_t\sigma S_t dV_t,$$

which shows that $\tilde{X}_t$, $t \geq 0$, is an Itô process with characteristics $a_t = 0$ and $b_t = e^{-rt}\phi_t\sigma S_t$.

Applying Itô Lemma to the function $f(t,x) = e^{rt}x$ and the Itô process $\tilde{X}_t$, $t \geq 0$ it follows that

$$X_t = e^{rt}\tilde{X}_t = f(t, \tilde{X}_t) = X_0 + \int_0^t re^{ru}\tilde{X}_u du + \int_0^t e^{ru}e^{-ru}\phi_u\sigma S_u dV_u$$
$$= X_0 + \int_0^t rX_u du + \int_0^t \phi_u\sigma S_u dV_u.$$

By (7.12) and the converse part of Proposition 7.6 our claim follows. □

Armed with these results we are ready to prove that (under some mild assumptions) options are replicable in the Black-Scholes model.

Theorem 7.13. *Under the assumptions of the Black-Scholes model suppose that h is an option with expiry time $T > 0$. Suppose moreover that*

$$\mathbb{E}_{\mathbb{Q}_T}|h|^2 < \infty,$$

and h is a non-negative random variable a.s.. Then the option h is replicable. Moreover, if

$$X_t = \varphi_t S_t + \psi_t B_t, t \in [0, T]$$

is the wealth process corresponding to the admissible replication strategy (φ_t, ψ_t) (in particular satisfying $X_T = h$), then

$$X_t = \mathbb{E}_{\mathbb{Q}_T}\left(e^{-r(T-t)}h \middle| \mathcal{F}_t\right), \quad t \in [0, T]. \tag{7.13}$$

In particular,

$$X_0 = E_{\mathbb{Q}_T}[e^{-rT}h].$$

Proof of Theorem 7.13. We begin by proving the second part of the theorem. Suppose that (φ_t, ψ_t) is an admissible trading strategy such that $X_T = h$. We will show that in this case the identity (7.13) holds true.
Recall that

$$X_t = \phi_t S_t + \psi_t B_t, \ t \geq 0$$

Multiplying this equality by e^{-rt} and denoting $\tilde{X}_t = e^{-rt}X_t$, $\tilde{S}_t = e^{-rt}S_t$ we see using $B_t = B_0 e^{rt}$ that

$$\tilde{X}_t = \phi_t \tilde{S}_t + \psi_t B_0, \ t \geq 0.$$

By Lemma 7.12 the trading strategy is self-financing if and only if

$$d\tilde{X}(t) = \phi(t)\, d\tilde{S}(t), \quad t \geq 0. \tag{7.14}$$

In the proof of Theorem 7.1 we have seen that $\tilde{S}(t)$ satisfies

$$d\tilde{S}_t = \sigma \tilde{S}_t dV_t$$

on $(\Omega\,, \mathcal{F}_T, \mathbb{Q}_T)$ and therefore

$$d\tilde{X}(t) = \phi(t)\sigma \tilde{S}_t dV_t.$$

This identity represents $\tilde{X}(t)$ as an Itô integral and we deduce that $\tilde{X}(t)$ is a martingale (both with respect to the measure $\mathbb{Q}_T$). Hence, we have

$$\tilde{X}(t) = \mathbb{E}_{\mathbb{Q}_T}\left(\tilde{X}_T \middle| \mathcal{F}_t\right), \quad t \in [0,T]. \tag{7.15}$$

By assumption $X_T = h$, which implies $\tilde{X}_T = e^{-rT}h$. Substituting the latter expression for $\tilde{X}_T$ into (7.15) we find that

$$\tilde{X}(t) = \mathbb{E}_{\mathbb{Q}_T}\left(e^{-rT}h \middle| \mathcal{F}_t\right), \quad t \in [0,T].$$

Multiplying both sides of the previous equation by e^{rt} gives (7.13).

To complete the proof we now show that h is replicable. Define the process Y by

$$Y_t = \mathbb{E}_{\mathbb{Q}_T}\left(e^{-rT}h \middle| \mathcal{F}_t\right). \tag{7.16}$$

Note that Y is a martingale and $Y_T = e^{-rT}h$. By the martingale representation theorem there exists a process g satisfying $\mathbb{E}\int_0^T g_t^2 dt$ and

$$Y_t = Y_0 + \int_0^t g_s dV_s. \tag{7.17}$$

Define φ_t by

$$\varphi_t = \frac{e^{rt}g_t}{(\sigma S_t)}$$

and ψ_t by

$$\psi_t = \frac{e^{rt}Y_t - \varphi_t S_t}{B_t},$$

the discounted wealth process $\widetilde{X}_t$ associated to this trading strategy satisfies $\widetilde{X}_t = Y_t$, in particular, we have $\widetilde{X}_0 = X_0 = \mathbb{E}_{\mathbb{Q}_T}\left(e^{-rT}h\right)$ and $X_T = h$. Using that $\varphi\widetilde{S}\sigma = g$ it follows from (7.17) that

$$d\widetilde{X}_t = \varphi_t \widetilde{S}_t \sigma dV_t = \phi(t)\, d\tilde{S}_t.$$

This implies by Lemma 7.12 that the trading strategy is self-financing. Since we assumed that h is a non-negative random variable, Y_t given by (7.16) is also non-negative $\forall t \in [0,T]$. We conclude that (φ_t, ψ_t) is an admissible strategy that replicates h. $\square$

Remark 7.14. *Note that by the usual no arbitrage arguments X_0, the initial value of the replication portfolio, is the fair, no arbitrage, price of the option h. The above result is much more general than our previous arguments, in that the option payoff can be an arbitrary, possibly "path dependent", random variable. In the special case that h is of the form $h = g(S_T)$ Theorem 7.13 implies Theorem 6.2 as a corollary (it provides an alternative proof of this result that does not use the Black-Scholes equation). On the other hand the above argument only asserts the existence of a replication portfolio and does not give an explicit formula for φ.*

Remark 7.15. *Note that under the risk neutral measure $\mathbb{Q}_T$ the stock price satisfies*

$$dS_t = rS_t dt + \sigma S_t dV_t,$$

i.e. the drift term of S under $\mathbb{Q}_T$ is $rS_t dt$ which is in line with the growth rate of the risk-free asset B that satisfies

$$dB_t = rB_t dt.$$

If the market was correctly modelled by the probability $\mathbb{Q}_T$ then the market was risk neutral. Mathematically speaking this means that $e^{-rt}S_t$ is a martingale under $\mathbb{Q}_T$.

Roughly speaking a market model admits no arbitrage if and only if the market has a risk-neutral measure.

8. Robustness of Black-Scholes Hedging[8]

8.1. **A formula for the hedging error.** If we assume the Black-Scholes model with

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \tag{8.1}$$

[8]In this section we follow the exposition by M. Davis

then the price at time t of an option with payoff $h(S_T)$ is $C_h(S_t, r, \sigma, t) = C(t, S_t)$, where $C(t, S)$ satisfies the Black-Scholes PDE with terminal condition

$$C(T, S) = h(S).$$

Suppose we sell an option at implied volatility $\hat{\sigma}$, i.e. we receive at time 0 the premium $C_h(S_0, r, \hat{\sigma}, 0)$ and we hedge under the assumption that the model (8.1) is correct with $\sigma = \hat{\sigma}$. The hedging strategy is then (as we have seen when we derived the Black-Scholes equation) "delta hedging": the number of units of the stock held at time t is the so-called option "delta" $\frac{\partial C}{\partial S}$, i.e.

$$\varphi_t = \frac{\partial C}{\partial S}(t, S_t). \tag{8.2}$$

Suppose now the model is **not** correct, but the true price process for the stock is

$$dS_t = \alpha(t, \omega) S_t dt + \beta(t, \omega) S_t dW_t, \tag{8.3}$$

where W is $(\mathcal{F}_t)_{t \geq 0}$-Brownian motion for some filtration $(\mathcal{F}_t)_{t \geq 0}$ and α_t, β_t are $\mathcal{F}_t-$adapted, say bounded, processes.[9] Using the trading strategy (8.2) (i.e. taking $\varphi_t = \frac{\partial C}{\partial S}(t, S_t)$ and $\psi_t = \frac{\left(X_t - \frac{\partial C}{\partial S}(t, S_t) S_t\right)}{B_t}$) the corresponding wealth process X_t is determined by the initial investment $X_0 = C(0, S_0)$, and the self-financing property

$$dX_t = \frac{\partial C}{\partial S}(t, S_t)\, dS_t + \left(X_t - \frac{\partial C}{\partial S}(t, S_t) S_t\right) r dt,$$

where S_t satisfies (8.3). We would like to estimate the final hedging error

$$X_T - h(S_T),$$

which is the difference of the value of the replication portfolio and the payoff of the option it was meant to replicate.

[9]Note that we do not loose generality by writing $\alpha(t, \omega) S_t$ as a more general coefficient γ can be written as

$$\gamma_t = \left(\frac{\gamma_t}{S_t}\right) S_t = \alpha_t S_t.$$

By the Ito formula $Y_t := C(t, S_t)$ satisfies

$$dY_t = \frac{\partial C}{\partial S}(t, S_t)\, dS_t + \left(\frac{\partial C}{\partial t}(t, S_t) + \frac{1}{2}\beta_t^2 S_t^2 \frac{\partial^2 C}{\partial S^2}(t, S_t)\right) dt.$$

Thus the hedging error $Z_t := X_t - Y_t$ satisfies the ordinary differential equation

$$dZ_t = \left(rX_t - rS_t\frac{\partial C}{\partial S}(t, S_t) - \frac{\partial C}{\partial t}(t, S_t) - \frac{1}{2}\beta_t^2 S_t^2 \frac{\partial^2 C}{\partial S^2}(t, S_t)\right) dt.$$

Using that C solves the Black-Scholes PDE we have

$$-\frac{\partial C}{\partial t}(t, S_t) - rS_t\frac{\partial C}{\partial S}(t, S_t) = -rC(t, S_t) + \frac{1}{2}\hat{\sigma}^2 S_t^2 \frac{\partial C}{\partial S^2}(t, S_t).$$

Substituting this expression and denoting $\Gamma_t = \Gamma(t, S_t) = \frac{\partial^2 C}{\partial S^2}(t, S_t)$ we see that

$$\frac{d}{dt}Z_t = rZ_t + \frac{1}{2}S_t^2\left(\hat{\sigma}^2 - \beta_t^2\right)\Gamma_t.$$

This is a first order ODE which can be written in the form $y' + a(t)\, y = b(t)$ and can therefore be solved using an integrating factor. Since $Z_0 = 0$, solving the ODE and noting that $Y_T = h(S_T)$ the final hedging error is given by the following key formula

$$\boxed{Z_T = X_T - h(S_T) = \int_0^T e^{r(T-t)}\frac{1}{2}S_t^2\Gamma_t\left(\hat{\sigma}^2 - \beta_t^2\right) dt.} \tag{8.4}$$

This is an important formula, as it shows that successful hedging is quite possible even under significant model error.

It is hard to imagine that the derivatives industry could exist at all without some result of this kind. We notice that successful hedging depends crucially on the relationship between the Black-Scholes implied volatility $\hat{\sigma}$ and the true "local volatility" β_t. Notice in particular, that we make a profit if we overestimate the volatility (X_T is the value of our replication portfolio at time T, $h(S_T)$ is what we need to pay the buyer of the option at exercise time and the difference $X_T - h(S_T)$ is positive in that case). The hedging error also depends on the option "convexity" Γ. If Γ is small then the hedging error is small even if the volatility has been underestimated.

8.2. **Determining the volatility.** All potential applications of the Black-Scholes formula hinge on knowledge of the volatility parameter σ in the stock price process. We briefly discuss two potential approaches to this problem.

8.2.1. *Historical volatility.* This method estimates the (future) volatility of a stock by considering its past volatility inferred from historic stock price data using statistical methods. There are several problems with this approach. Empirical studies suggest that the volatility is unstable through time: Historical precedent is a poor guide for estimating future volatility. Moreover, estimates of option prices based on historical volatility are systematically biased (options on stocks with high volatility are overpriced.

8.2.2. *Implied Volatility.* Here the idea is to infer the current market consensus on the volatility of a stock by examining the prices at which options on this stock trade at present. This means one considers the equation

$$C_t = S_t \Phi\left(d_1\left(S_t, T-t, \hat{\sigma}, K, r\right)\right) - K e^{-r(T-t)} \Phi\left(d_2\left(S_t, T-t, \hat{\sigma}, K, r\right)\right), \tag{8.5}$$

where C_t) is the current market price of a European call with exercise time T and strike K and the right hand side the Black-Scholes formula with parameters $S_t, T-t, \hat{\sigma}, K, r$. To find the implied volatility, one needs to find $\hat{\sigma}$ by solving the equation (8.5) in which all the other parameters are known. In a sense this is the inverse problem to finding the price of the call option.

Definition 8.1. *The implied volatility $\hat{\sigma}$ (of a European call option) is the value that, when put into the Black-Scholes formula, results in a model price equal to the current market price of that option.*

Strictly speaking we have defined the market-based Black-Scholes implied volatility. In practice it is quite common to quote the price of an option in terms of its implied volatility rather than its actual price (the higher the implied volatility the more expensive the option). Compare this with (8.4) and the discussion that follows.

The Black-Scholes model suggests that options with any strike/exercise time trade for the same implied volatility $\hat{\sigma}$. In actual markets this is not the case giving rise to volatility risks, and stochastic volatility models of the market that address them.

Example 8.2. *Empirical option prices suggest that if the strike price of an option is substantially higher or lower than the current price of the underlying stock the implied volatility is higher. This phenomenon is known as the volatility smile.*

9. American Options

In this section we consider American style options, which in contrast to European options can be exercised at any time $t \in (0, T]$, i.e. any time before the expiry time T.

Throughout the section we assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ is a filtration on this space and $W = (W_t)_{t\geq 0}$ is an $\mathbb{F}$-Brownian motion. We also assume that the pair (S, B) describes a Black-Scholes market model, i.e. it solves the following SDEs

$$\begin{cases} dS(t) &= \mu S(t)\,dt + \sigma S(t)\,dW(t), \quad t \geq 0, \\ dB(t) &= rB(t)\,dt, \quad t \geq 0. \end{cases} \tag{9.1}$$

Recall that $S(t)$, $t \geq 0$ denotes the stock price (risky asset), $B(t)$, $t \geq 0$ denotes the bond price (not-risky asset) and r, μ and σ are fixed positive constants.

The following definition formally introduces the notion of an American option.

Definition 9.1. *Let $T > 0$ be fixed. An American option with expiry time T is an $(\mathcal{F}_t)_{t\geq 0}$-adapted process, $h = (h(t))$, $t \in [0, T]$, such that for all $t \in [0, T]$, $h(t) \geq 0$, $\mathbb{P}$-a.s..*

We illustrate this definition with some examples.

Example 9.2. *Assume that* $g : (0, \infty) \to (0, \infty)$ *is Borel measurable, e.g. continuous, function. Then a process* h *defined by*

$$h(t) = g(S(t)), t \in [0, T]$$

is an American option with expiry time T. *In particular, if* $K > 0$ *is fixed and* $g(S) = (S - K)^+$, *then the process*

$$h(t) = (S(t) - K)^+, \ t \in [0, T]$$

is called an American call option. Similarly, if $g(S) = (K - S)^+$, *then the process*

$$h(t) = (K - S(t))^+, \ t \in [0, T]$$

is called an American put option.

In order to hedge American options we have to generalise our notion of a self-financing trading strategy.

Definition 9.3. *Let* $T > 0$ *be fixed. A trading strategy with consumption is a pair* (φ, ψ) *of* $\mathbb{F}$*-adapted process,* $\varphi = (\varphi(t))$, $t \in [0, T]$, $\psi = (\psi(t))$, $t \in [0, T]$, *such that* $\mathbb{P}$*-a.s.,*

$$\int_0^T |\psi(t)| \, dt < \infty, \tag{9.2}$$

$$\int_0^T |\varphi(t)|^2 \, dt < \infty \tag{9.3}$$

and such that there exists a continuous, non-decreasing $\mathbb{F}$*-adapted process* $C = (C(t))$, $t \in [0, T]$, *such that*

$$\begin{aligned} \varphi(t)S(t) + \psi(t)B(t) &= \varphi(0)S(0) + \psi(0)B(0) \\ &+ \int_0^t \varphi(\tau) \, dS(\tau) + \int_0^t \psi(\tau) \, dB(\tau) - C(t), \ t \in [0, T]. \end{aligned} \tag{9.4}$$

The random variable $C(t)$ *is called the cumulative consumption up to time* t *and the process* C *is called the cumulative consumption process.*

Remark 9.4. *A self-financing strategy is a trading strategy with consumption. To see this, note that if a trading strategy is self-financing the equation* (9.4) *holds with* $C(t) = 0$, $t \geq 0$.

In order to determine the price of an American option we have to be able to hedge its payoff.

Definition 9.5. *Let* $T > 0$ *be fixed and let* $h = (h(t))$, $t \in [0,T]$, *be an American option with expiry time* T. *A trading strategy with consumption* (φ, ψ) *is said to hedge the option* h *iff the corresponding wealth process* $X = (X(t))_{t\in[0,T]}$ *defined by*

$$X(t) = \varphi(t)S(t) + \psi(t)B(t), \; t \in [0,T]$$

satisfies

$$X(t) \geq h(t), \; t \in [0,T]. \tag{9.5}$$

It is natural to ask if an American option can be hedged by a trading strategy with consumption. The following theorem, which we state without proof, provides an affirmative answer to this question if the American options is of the form $h(t) = g(S(t))$. In the following $\mathbb{Q}_T$ is the risk neutral measure and $V = (V(t))_{t\in[0,T]}$ is the process defined by formula (7.5). Let me recall that V is Brownian motion on the probability space $(\Omega, \mathcal{F}, \mathbb{Q}_T)$.

Theorem 9.6. *Suppose that* $T > 0$ *and* g *is a Borel measurable function* $g : (0,\infty) \to (0,\infty)$. *Let* $h = (g(S(t)))$, $t \in [0,T]$, *be an American option with expiry time* T.
Then there exists a trading strategy with consumption $(\bar{\varphi}, \bar{\psi})$ *which hedges* h *and such the wealth process* $\bar{X}$ *corresponding to* $(\bar{\varphi}, \bar{\psi})$ *defined by*

$$\bar{X}(t) = \bar{\varphi}(t)S(t) + \bar{\psi}(t)B(t), \; t \in [0,T] \tag{9.6}$$

satisfies

$$\bar{X}(t) = u(t, S(t)),\ t \in [0, T], \tag{9.7}$$

where

$$u(t, S) := \sup_{\tau} \mathbb{E}_{Q_T}\left[e^{-r(\tau - t)} g\big(Se^{(r-\frac{\sigma^2}{2})(\tau - t)} e^{\sigma(V_\tau - V_t)}\big)\right], t > 0, S > 0, \tag{9.8}$$

where the $\sup_\tau$ *is taken over all stopping times*[10] $\tau : \Omega \to [t, T]$.

Moreover, for any trading strategy with consumption (φ, ψ) *that hedges* h*, its wealth process satisfies*

$$X(t) \geq u(t, S(t)),\ t \in [0, T]. \tag{9.9}$$

A brief discussion of stopping times. Let $(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{F}_t)$ be a filtered probability space. A random variable $\tau : \Omega \to [0, \infty]$ is a stopping time with respect to the filtration $\{\mathcal{F}_t\}$ if the event $\{\tau \leq t\} \in \mathcal{F}_t$ for every $t \geq 0$. Intuitively, this means it is possible to decide if $\{\tau \leq t\}$ has occurred based on $\mathcal{F}_t$, i.e. without looking into the future.

The following are some important examples of stopping times.

(i) If T is a nonnegative constant, then T is a stopping time.

(ii) If T, S are stopping times on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, then $T + S$, $T \wedge S$, $T \vee S$, αT, $\alpha \geq 1$ are stopping times. Here

$$\begin{aligned} [T \wedge S](\omega) &= \min\{T(\omega), S(\omega)\},\ \omega \in \Omega; \\ (T \vee S)(\omega) &= \max\{T(\omega), S(\omega)\},\ \omega \in \Omega. \end{aligned}$$

(iii) If $\{T_n\}_{n=1}^{\infty}$ is a sequence of stopping times, then the random time $\sup_{n\in\mathbb{N}} T_n$ is a stopping time.

(iv) The first hitting time of a Brownian motion is a stopping time. For example, if $\alpha \in \mathbb{R}$ then $\tau(\omega) := \{t \geq 0 | W_t(\omega) \geq \alpha\}$ is a stopping time.

---

[10] See the brief discussion of stopping times below.

Remark Since a constant $\tau = T$ is a stopping time, we deduce that the function u defined by the formula (9.8) satisfies

$$u(t,S) \geq \mathbb{E}_{Q_T}\left[e^{-r(T-t)}g\big(Se^{(r-\frac{\sigma^2}{2}(T-t))}e^{\sigma(V_T-V_t)}\big)\right], t>0, S>0, \tag{9.10}$$

We deduce that

$$u(t,S) \geq G(t,S), t>0, S>0. \tag{9.11}$$

This implies that **the value of the wealth processes corresponding to the American option with function g is larger or equal to the wealth processes corresponding to the European option with the same function g.**

We finish this section with the following result that shows that in the prices of European and American call options are actually the same in the Black-Scholes model.

Theorem 9.7. *Let $C(t)$ denote the price of the European call with expiry at $T > 0$ and strike $K > 0$ (so, the payoff is $(S_T - K)^+$). For all $t \in [0,T]$ the payoff of the American option $(S_t - K)^+$ does not exceed $C(t)$:*

$$(S_t - K)^+ \leq C(t). \tag{9.12}$$

Hence, the prices of the corresponding American and European calls are equal.

Proof of Theorem 9.7. To simplify the notation we will denote $\mathbb{E}_{\mathbb{Q}_T}$ by $\mathbb{E}^*$. By the risk neutral pricing formula we have

$$C(t) = \mathbb{E}^*\left[e^{-r(T-t)}(S_T-K)^+\big|\mathcal{F}_t\right] \tag{9.13}$$

But, because $x^+ \geq x$, $x \in \mathbb{R}$ and since $(\tilde{S}_t)_{t\in[0,T]}$ is an $\mathbb{F}$-martingale on $(\Omega, \mathcal{F}, \mathbb{Q}_T)$

$$\begin{aligned}
\mathbb{E}^*\left[e^{-r(T-t)}g(S_T)\big|\mathcal{F}_t\right] &= \mathbb{E}^*\left[e^{-r(T-t)}(S_T-K)^+\big|\mathcal{F}_t\right] \\
\geq \mathbb{E}^*\left[e^{-r(T-t)}(S_T-K)\big|\mathcal{F}_t\right] &= \mathbb{E}^*\left[e^{-r(T-t)}S_T\big|\mathcal{F}_t\right] - \mathbb{E}^*\left[e^{-r(T-t)}K\big|\mathcal{F}_t\right] \\
= e^{rt}\mathbb{E}^*\left[\tilde{S}_T\big|\mathcal{F}_t\right] - Ke^{-r(T-t)} &= e^{rt}\tilde{S}_t - Ke^{-r(T-t)} \geq S_t - K.
\end{aligned}$$

The last inequality can be seen by observing that $-Ke^{-r(T-t)} \geq -K$.

Moreover, we know that $C(t) \geq 0$ for all $t \in [0, T]$, so

$$C(t) \geq (S_t - K)^+,$$

completing the proof of (9.12).

Finally, note that (9.12) means that the payoff of an American call at any moment of time $t \in [0, T]$ is bounded above by the price of the corresponding European call (with the same expiry and strike). Hence, it is never advantageous to exercise American call before the expiry T (when its payoff precisely coincides with the payoff of the European call). Thus the best strategy hedging American call replicates the corresponding European call.

$\square$

Remark 9.8. *The same argument for American and European put options sharing a strike and expiry time breaks down. In fact, in this case early exercise of an American put can be optimal. Finding this optimal exercise time is a rather subtle question, which requires a careful analysis using the theory of stopping times. In general, it is impossible to provide a closed-form formula for the price of American put options.*

The following section is provided for self-reading. It is not examinable.

10. Barrier Options

The payoff of a barrier option is triggered on or off if the asset crosses a preset barrier for the first time. Depending on whether the barrier is to be crossed from below or from above, and whether the crossing of the barrier triggers the payoff on or off, we can talk of four type of barrier options: *up-and-in*, *up-and-out*, *down-and-in*, and *down-and-out*.

For example, the payoff of an up-and-out call option with strike price K, barrier X and expiry time T is

$$D = (S_T - K)^+ 1_{\max_{t\in[0,T]} S_t \le X},$$

and the payoff of a down-and-in put option with strike price K, barrier X and expiry time T is

$$D = (K - S_T)^+ 1_{\min_{t\in[0,T]} S_t \le X.}$$

There are various reasons why one would want to trade in barrier options. For example, the liability of the writer of a vanilla call is unlimited, but for an up-and-out call there is an upper bound on the liability. For the buyer, a barrier option can be a cheaper alternative to a vanilla option. If you have reasons to believe that the underlying price will go above a certain level, say, £9, you can buy a vanilla call with strike price £9, but if you don't expect the underlying to reach £12, an up-and-out call with a barrier at £12 will be a cheaper alternative.

To derive a formula for the price of this kind of option, we need to know the joint distribution of S_T and $\max_{t\in[0,T]} S_t$ (or $\min_{t\in[0,T]} S_t$). We begin with a simpler problem, namely finding the joint distribution of Brownian motion W_T and the maximum

$$M_T = \max_{t\in[0,T]} W_t.$$

Consider the event $\{W_T < b, M_t \ge c\}$ for some $0 < c$ and $b \le c$. Let T^c be the first time t such that $W_t = c$. Consider a new process $\bar{W}_t$ such that

$$\bar{W}_t = \begin{cases} W_t & \text{for } 0 \le t \le T^c, \\ 2c - W_t & \text{for } T^c < t. \end{cases}$$

The trajectories of $\bar{W}_t$ for $t \ge T^c$ are obtained by reflecting those of W_t about the horizontal line at c. Since $W_t - W_{T^c}$ is a Brownian motion, so is $\bar{W}_t - \bar{W}_{T^c} = -(W_t - W_{T^c})$. By independence of increments, it therefore follows that $\bar{W}_t$ for $t \ge 0$ is also a Brownian motion. This result is known as the *reflection principle*. It implies that

$$\mathbb{P}\left\{W_T < b, M_T \geq c\right\} = \mathbb{P}\{\bar{W}_T > 2c - b\} = 1 - N\left(\frac{2c-b}{\sqrt{T}}\right) = N\left(\frac{b-2c}{\sqrt{T}}\right).$$

It follows that the joint distribution function of W_T and M_T is given by

$$\begin{aligned} F(b,c) &= \mathbb{P}\left\{W_T < b, M_T < c\right\} \\ &= \mathbb{P}\left\{W_T < b\right\} - \mathbb{P}\left\{W_T < b, M_t \geq c\right\} \\ &= \mathbb{P}\left\{W_T < b\right\} - \mathbb{P}\left\{W_T > 2c - b\right\} \\ &= N\left(\frac{b}{\sqrt{T}}\right) - N\left(\frac{b-2c}{\sqrt{T}}\right) \end{aligned}$$

for any b, c such that $0 < c$ and $b \leq c$. The joint density of W_T and M_T is therefore given by

$$\begin{aligned} f(b,c) &= \frac{\partial^2 F(b,c)}{\partial b\, \partial c} \\ &= \frac{d}{dc}\left[\frac{1}{\sqrt{T}} p\left(\frac{b}{\sqrt{T}}\right) - \frac{1}{\sqrt{T}} p\left(\frac{b-2c}{\sqrt{T}}\right)\right] \\ &= \frac{2}{T} p'\left(\frac{b-2c}{\sqrt{T}}\right) \\ &= \frac{2\,(2c-b)}{\sqrt{2\pi T^3}} e^{-\frac{1}{2}\left(\frac{b-2c}{\sqrt{T}}\right)^2}, \end{aligned}$$

where

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}.$$

As an example, we shall price an up-and-out put with strike K, barrier X and expiry T. The payoff is

$$D = (K - S_T)^+ 1_{\max_{t\in[0,T]} S_t \leq X}. \tag{10.1}$$

The stock price satisfies the SDE

$$dS_t = rS_t dt + \sigma S_t\, dW_t^*,$$

so that

$$S_t = S_0 e^{\sigma W_t^* + \left(r - \frac{1}{2}\sigma^2\right)t} = S_0 e^{\sigma Y_t},$$

where

$$Y_t = W_t^* + \nu t$$

with

$$\nu = \frac{r - \frac{1}{2}\sigma^2}{\sigma}.$$

The option price can be expressed as

$$\begin{aligned}
V_0 &= e^{-rT}\mathbb{E}^{\mathbb{P}^*}(g(S_T)) \\
&= e^{-rT}\mathbb{E}^{\mathbb{P}^*}\left((K - S_T)^+ 1_{\max_{t\in[0,T]} S_t \le X}\right) \\
&= e^{-rT}\mathbb{E}^{\mathbb{P}^*}\left((K - S_T) 1_{S_T \le K, \max_{t\in[0,T]} S_t \le X}\right) \\
&= e^{-rT}K\mathbb{E}^{\mathbb{P}^*}\left(1_{S_T \le K, \max_{t\in[0,T]} S_t \le X}\right) - e^{-rT}\mathbb{E}^{\mathbb{P}^*}\left(S_T 1_{S_T \le K, \max_{t\in[0,T]} S_t \le X}\right)
\end{aligned}$$

Using the Girsanov theorem, we can introduce a probability measure $\mathbb{Q}$ such that $Y_t = W_t^* + \nu t$ is a Brownian motion under $\mathbb{Q}$. The density of $\mathbb{Q}$ with respect to $\mathbb{P}^*$ is

$$L_T = e^{-\nu W_T^* - \frac{1}{2}\nu^2 T},$$

and the density of $\mathbb{P}^*$ with respect to $\mathbb{Q}$ is

$$\frac{1}{L_T} = e^{\nu W_T^* + \frac{1}{2}\nu^2 T} = e^{\nu(Y_T - \nu T) + \frac{1}{2}\nu^2 T} = e^{\nu Y_T - \frac{1}{2}\nu^2 T}.$$

Putting

$$b = \frac{1}{\sigma}\ln\frac{K}{S_0}, \quad c = \frac{1}{\sigma}\ln\frac{X}{S_0},$$

we have

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{P}^*}\left(1_{S_T\le K,\max_{t\in[0,T]}S_t\le X}\right)=\mathbb{E}^{\mathbb{P}^*}\left(1_{Y_T\le b,\max_{t\in[0,T]}Y_t\le c}\right)\\
&=\mathbb{E}^{\mathbb{Q}}\left(e^{\nu Y_T-\frac{1}{2}\nu^2T}1_{Y_T\le b,\max_{t\in[0,T]}Y_t\le c}\right)=\int_{-\infty}^{b}e^{\nu y-\frac{1}{2}\nu^2T}\left(\int_0^c f(y,z)dz\right)dy\\
&=\int_{-\infty}^{b}e^{\nu y-\frac{1}{2}\nu^2T}\left(\frac{1}{\sqrt{T}}p\left(\frac{y}{\sqrt{T}}\right)-\frac{1}{\sqrt{T}}p\left(\frac{y-2c}{\sqrt{T}}\right)\right)dy\\
&=\frac{1}{\sqrt{2\pi T}}\int_{-\infty}^{b}e^{\nu y-\frac{1}{2}\nu^2T}e^{-\frac{y^2}{2T}}dy-\frac{1}{\sqrt{2\pi T}}\int_{-\infty}^{b}e^{\nu y-\frac{1}{2}\nu^2T}e^{-\frac{(y-2c)^2}{2T}}dy\\
&=\frac{1}{\sqrt{2\pi T}}\int_{-\infty}^{b}e^{-\frac{(y-\nu T)^2}{2T}}dy-\frac{e^{2\nu c}}{\sqrt{2\pi T}}\int_{-\infty}^{b}e^{-\frac{(y-2c-\nu T)^2}{2T}}dy\\
&=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{b-\nu T}{\sqrt{T}}}e^{-\frac{u^2}{2}}du-\frac{e^{2\nu c}}{\sqrt{2\pi}}\int_{-\infty}^{\frac{b-2c-\nu T}{\sqrt{T}}}e^{-\frac{v^2}{2}}dv\\
&=N\left(\frac{b-\nu T}{\sqrt{T}}\right)-e^{2\nu c}N\left(\frac{b-2c-\nu T}{\sqrt{T}}\right)\\
&=N\left(\frac{\frac{1}{\sigma}\ln\frac{K}{S_0}-\frac{r-\frac{1}{2}\sigma^2}{\sigma}T}{\sqrt{T}}\right)-e^{2\frac{r-\frac{1}{2}\sigma^2}{\sigma}\frac{1}{\sigma}\ln\frac{X}{S_0}}N\left(\frac{\frac{1}{\sigma}\ln\frac{K}{S_0}-2\frac{1}{\sigma}\ln\frac{X}{S_0}-\frac{r-\frac{1}{2}\sigma^2}{\sigma}T}{\sqrt{T}}\right)\\
&=N\left(d_1\right)-\left(\frac{X}{S_0}\right)^{2\frac{r}{\sigma^2}-1}N\left(d_2\right),
\end{aligned}
$$

where

$$
\begin{aligned}
d_1&=\frac{\ln\frac{K}{S_0}-\left(r-\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}},\\
d_2&=\frac{\ln\frac{S_0K}{X^2}-\left(r-\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}.
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{P}^*}\left(S_T 1_{S_T\le K,\max_{t\in[0,T]} S_t\le X}\right) = S_0\mathbb{E}^{\mathbb{P}^*}\left(e^{\sigma Y_T} 1_{Y_T\le b\max_{t\in[0,T]} Y_t\le c}\right)\\
&= S_0\mathbb{E}^{\mathbb{Q}}\left(e^{\nu Y_T-\frac{1}{2}\nu^2 T}e^{\sigma Y_T}1_{Y_T\le b,\max_{t\in[0,T]} Y_t\le c}\right) = S_0\int_{-\infty}^{b} e^{\nu y-\frac{1}{2}\nu^2 T}e^{\sigma y}\left(\int_0^c f(y,z)dz\right)dy\\
&= S_0\int_{-\infty}^{b} e^{\nu y-\frac{1}{2}\nu^2 T}e^{\sigma y}\left(\frac{1}{\sqrt{T}}p\left(\frac{y}{\sqrt{T}}\right)-\frac{1}{\sqrt{T}}p\left(\frac{y-2c}{\sqrt{T}}\right)\right)dy\\
&= S_0\frac{e^{\frac{1}{2}\sigma^2 T+\nu\sigma T}}{\sqrt{2\pi T}}\int_{-\infty}^{b} e^{-\frac{(y-\nu T-\sigma T)^2}{2T}}dy - \frac{e^{\frac{1}{2}\sigma^2 T+\nu\sigma T+2c\nu+2c\sigma}}{\sqrt{2\pi T}}\int_{-\infty}^{b} e^{-\frac{(y-2c-\nu T-\sigma T)^2}{2T}}dy\\
&= S_0\frac{e^{\frac{1}{2}\sigma^2 T+\nu\sigma T}}{\sqrt{2\pi}}\int_{-\infty}^{\frac{b-\nu T-\sigma T}{\sqrt{T}}} e^{-\frac{u^2}{2}}dy - \frac{e^{\frac{1}{2}\sigma^2 T+\nu\sigma T+2c\nu+2c\sigma}}{\sqrt{2\pi}}\int_{-\infty}^{\frac{b-2c-\nu T-\sigma T}{\sqrt{T}}} e^{-\frac{v^2}{2}}dv\\
&= S_0 e^{\frac{1}{2}\sigma^2 T+\nu\sigma T}N\left(\frac{b-\nu T-\sigma T}{\sqrt{T}}\right) - e^{\frac{1}{2}\sigma^2 T+\nu\sigma T+2c\nu+2c\sigma}N\left(\frac{b-2c-\nu T-\sigma T}{\sqrt{T}}\right)\\
&= e^{\frac{1}{2}\sigma^2 T+\frac{r-\frac{1}{2}\sigma^2}{\sigma}\sigma T}S_0 N\left(\frac{\frac{1}{\sigma}\ln\frac{K}{S_0}-\frac{r-\frac{1}{2}\sigma^2}{\sigma}T-\sigma T}{\sqrt{T}}\right)\\
&\quad - e^{\frac{1}{2}\sigma^2 T+\frac{r-\frac{1}{2}\sigma^2}{\sigma}\sigma T+2\frac{1}{\sigma}\ln\frac{X}{S_0}\frac{r-\frac{1}{2}\sigma^2}{\sigma}+2\frac{1}{\sigma}\ln\frac{X}{S_0}\sigma}S_0 N\left(\frac{\frac{1}{\sigma}\ln\frac{K}{S_0}-2\frac{1}{\sigma}\ln\frac{X}{S_0}-\frac{r-\frac{1}{2}\sigma^2}{\sigma}T-\sigma T}{\sqrt{T}}\right)\\
&= e^{rT}S_0 N(d_3) - e^{rT}\left(\frac{X}{S_0}\right)^{2\frac{r}{\sigma^2}+1}S_0 N(d_4),
\end{aligned}
$$

where

$$
\begin{aligned}
d_3 &= \frac{\ln\frac{K}{S_0}-\left(r+\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}},\\
d_4 &= \frac{\ln\frac{S_0 K}{X^2}-\left(r+\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}.
\end{aligned}
$$

It follows that the option price is

$$
\begin{aligned}
V_0 &= e^{-rT}\mathbb{E}^{\mathbb{P}^*}(D)\\
&= e^{-rT}K\mathbb{E}^{\mathbb{P}^*}\left(1_{S_T\le K,\max_{t\in[0,T]} S_t\le X}\right) - e^{-rT}\mathbb{E}^{\mathbb{P}^*}\left(S_T 1_{S_T\le K,\max_{t\in[0,T]} S_t\le X}\right)\\
&= e^{-rT}K\left[N(d_1)-\left(\frac{X}{S_0}\right)^{2\frac{r}{\sigma^2}-1}N(d_2)\right] - S_0\left[N(d_3)-\left(\frac{X}{S_0}\right)^{2\frac{r}{\sigma^2}+1}N(d_4)\right].
\end{aligned}
$$

From Theorem 7.1 we know that if $F(t,S)$ satisfies the Black-Scholes equation

$$
\frac{\partial F}{\partial t}+rS\frac{\partial F}{\partial S}+\frac{1}{2}\sigma^2 S^2\frac{\partial^2 F}{\partial S^2}=rF, \tag{10.2}
$$

then $e^{-rt}F(t,S_t)$ is a martingale under the risk neutral probability $\mathbb{P}^*$. Suppose that, in addition, $F(t,S)$ satisfies the final and boundary conditions

$$F(T,S) = (K-S)^+ \quad \text{for all } 0 < S \leq X, \tag{10.3}$$

$$F(t,X) = 0 \quad \text{for all } 0 \leq t < T. \tag{10.4}$$

We assume that $S_0 < X$ and denote by τ the first time t when $S_t = X$. Then $\min\{\tau, T\}$ is a stopping time.

By the so-called optimal stopping theorem it follows that

$$\begin{aligned}
F(0,S_0) &= \mathbb{E}^{\mathbb{P}^*}\left(e^{-r\min\{\tau,T\}}F(\min\{\tau,T\}, S_{\min\{\tau,T\}})\right) \\
&= \mathbb{E}^{\mathbb{P}^*}\left(1_{\tau<T}e^{-r\tau}F(\tau,S_\tau)\right) + \mathbb{E}^{\mathbb{P}^*}\left(1_{\tau\geq T}e^{-rT}F(T,S_T)\right) \\
&= \mathbb{E}^{\mathbb{P}^*}\left(1_{\tau\geq T}e^{-rT}(K-S_T)^+\right) \\
&= e^{-rT}\mathbb{E}^{\mathbb{P}^*}(D).
\end{aligned}$$

This means that the solution to the Black-Scholes equation (10.2) with the final and boundary conditions (10.3) and (10.4) provides the price of the up-and-out put option,

$$V_0 = e^{-rT}\mathbb{E}^{\mathbb{P}^*}(D) = F(0,S_0).$$

Finally, given the solution $F(t,S)$ to the final-boundary value problem (10.3), (10.4) for the Black-Scholes equation (10.2), we can construct a replicating strategy X_t, Y_t for the up-and-out put as follows:

$$Y_t = \frac{\partial F}{\partial S}(t,S_t), \quad X_t = \frac{F(t,S_t) - Y_t S_t}{B_t}$$

for $t \leq \min\{\tau, T\}$, and

$$Y_t = 0, \quad X_t = 0$$

for $t > \min\{\tau, T\}$. The self financing condition for $0 \leq t \leq \min\{\tau, T\}$ can be verified with Itô lemma, and for $\min\{\tau, T\} \leq t \leq T$ it follows from the fact that the strategy X_t, Y_t is stationary (does not change with t) for $\tau \leq t \leq T$ and its time τ value is nil, $V_\tau = 0$. The strategy replicates the up-and-out put payoff (10.1) since

$$V_T = X_T B_T + Y_T S_T = F(T, S_T) = (K - S_T)^+$$

if $\tau \geq T$, and

$$V_T = X_T B_T + Y_T S_T = 0$$

if $\tau < T$, so that

$$V_T = (K - S_T)^+ 1_{\max_{t\in[0,T]} S_t \leq X} = D.$$

The following section is provided for self-reading. It is not examinable.

11. Options on Dividend Paying Stock

We shall consider the simplest case of a stock paying dividends continuously at a rate $\delta \geq 0$. The bond and stock prices obey the usual SDEs

$$\begin{aligned} dB_t &= rB_t dt, \\ dS_t &= \mu S_t dt + \sigma S_t\, dW_t. \end{aligned}$$

The value of a portfolio consisting of X_t bonds and Y_t shares of stock is also given by the usual formula

$$V_t = X_t B_t + Y_t S_t.$$

However, we need to adjust the self financing condition to allow for the injection of funds (the dividends), namely,

$$dV_t = X_t dB_t + Y_t dS_t + \delta S_t dt. \tag{11.1}$$

A strategy of this kind replicates a contingent claim with payoff D and expiry T if

$$V_T = D.$$

To avoid arbitrage, the time $t \leq T$ price of such an option must be equal to V_t.

Applying the Itô formula, we have

$$\begin{aligned}
d\tilde{V}_t &= d\left[e^{-rt}V_t\right] \\
&= -re^{-rt}V_t dt + e^{-rt}dV_t \\
&= -re^{-rt}V_t dt + e^{-rt}\left(X_t dB_t + Y_t dS_t + \delta Y_t S_t dt\right) \\
&= -re^{-rt}Y_t S_t dt + e^{-rt}Y_t\left(\mu S_t dt + \sigma S_t dW_t\right) + \delta e^{-rt}Y_t S_t dt \\
&= e^{-rt}Y_t S_t\left(\left(\mu + \delta - r\right)dt + \sigma dW_t\right) \\
&= \sigma e^{-rt}Y_t S_t\, dW_t^* \\
&= \sigma Y_t \tilde{S}_t\, dW_t^*,
\end{aligned}$$

where

$$W_t^* = \frac{\mu + \delta - r}{\sigma}t + W_t.$$

Girsanov's theorem with

$$\theta_t = \frac{\mu + \delta - r}{\sigma}$$

provides a probability measure $\mathbb{P}^*$ such that W_t^* is a Brownian motion under $\mathbb{P}^*$. Then $\tilde{V}_t$ is a martingale under $\mathbb{P}^*$. It follows that the time t option price is given by

$$V_t = e^{rt}\tilde{V}_t = e^{rt}\mathbb{E}^{\mathbb{P}^*}(\tilde{V}_T\big|\mathcal{F}_t) = e^{rt}\mathbb{E}^{\mathbb{P}^*}(e^{-rT}V_T\big|\mathcal{F}_t) = e^{-r(T-t)}\mathbb{E}^{\mathbb{P}^*}(D\big|\mathcal{F}_t).$$

Next, observe that

$$\begin{aligned}
dS_t &= \mu S_t dt + \sigma S_t\, dW_t \\
&= \mu S_t dt + \sigma S_t\left(dW_t^* - \frac{\mu + \delta - r}{\sigma}dt\right) \\
&= (r - \delta)S_t dt + \sigma S_t\, dW_t^*.
\end{aligned}$$

For a sufficiently regular function $F(t,S)$ the Itô formula gives

$$\begin{aligned}
& d\left[e^{-rt}F(t,S_t)\right] \\
&= -re^{-rt}F(t,S_t)dt + e^{-rt}dF(t,S_t) \\
&= -re^{-rt}F dt + e^{-rt}\left(\frac{\partial F}{\partial t}dt + \frac{\partial F}{\partial S}dS_t + \frac{1}{2}\frac{\partial^2 F}{\partial S^2}(dS_t)^2\right) \\
&= -re^{-rt}F dt + e^{-rt}\left(\frac{\partial F}{\partial t}dt + \frac{\partial F}{\partial S}\left((r-\delta)S_t dt + \sigma S_t\, dW_t^*\right) + \frac{1}{2}\frac{\partial^2 F}{\partial S^2}\sigma^2 S_t^2 dt\right) \\
&= e^{-rt}\left(-rF + \frac{\partial F}{\partial t} + (r-\delta)S_t\frac{\partial F}{\partial S} + \frac{1}{2}\sigma^2 S_t^2\frac{\partial^2 F}{\partial S^2}dt\right)dt + e^{-rt}\sigma S_t\frac{\partial F}{\partial S}\, dW_t^*.
\end{aligned}$$

It follows that $e^{-rt}F(t,S_t)$ is a martingale under $\mathbb{P}^*$ whenever $F(t,S)$ satisfies the PDE

$$\frac{\partial F}{\partial t} + (r-\delta)S\frac{\partial F}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 F}{\partial S^2}dt = rF. \tag{11.2}$$

This is the analogue of the Black-Scholes equation for options on stock paying dividends at a rate δ.

Given a European option with payoff $D = H(S_T)$, if, in addition, $F(t,S)$ satisfies the final condition

$$F(T,S) = H(S)$$

for all $S > 0$, then it can be shown in the same way as in Chapter 4 that the time t option price is given by

$$V_t = F(t,S_t).$$

Finally, the construction of a hedging strategy (by delta hedging) is the same as before:

$$\begin{aligned}
Y_t &= \frac{\partial F}{\partial S}(t,S_t), \\
X_t &= \frac{F(t,S_t) - Y_t S_t}{B_t}
\end{aligned}$$

All that remains to be done is to verify that

$$V_t = X_t B_t + Y_t S_t = F(t,S_t)$$

satisfies (11.1). Indeed, using (11.2), we get

$$\begin{aligned}
dV_t &= dF(t, S_t) \\
&= \frac{\partial F}{\partial t}dt + \frac{\partial F}{\partial S}dS_t + \frac{1}{2}\frac{\partial^2 F}{\partial S^2}(dS_t)^2 \\
&= \frac{\partial F}{\partial t}dt + \frac{\partial F}{\partial S}\left((r-\delta)S_t dt + \sigma S_t\, dW_t^*\right) + \frac{1}{2}\frac{\partial^2 F}{\partial S^2}\sigma^2 S_t dt \\
&= rFdt + \sigma S_t \frac{\partial F}{\partial S}\, dW_t^* \\
&= rFdt + \sigma S_t \frac{\partial F}{\partial S}\left(\frac{\mu+\delta-r}{\sigma}dt + dW_t\right) \\
&= r\left(F - S_t\frac{\partial F}{\partial S}\right)dt + \frac{\partial F}{\partial S}\left(\mu S_t dt + \sigma S_t\, dW_t\right) + \delta\frac{\partial F}{\partial S}S_t dt \\
&= rX_t B_t dt + Y_t dS_t + \delta Y_t S_t dt \\
&= X_t dB_t + Y_t dS_t + \delta Y_t S_t dt.
\end{aligned}$$

Email address: christian.litterer@york.ac.uk

Department of Mathematics, The University of York, James College, room G/111, office hour Thursday 13:00, Heslington, York, YO10 5DD, U.K.