MATH3733 – Part 2, continuous time

"beta– version-a" 30.11.2008

In this part of the course we will develop some tools to analyze stochastic market *directly* in continuous time, limiting procedure as in the Central Limit Theorem being performed in advance. For the discrete CRR model the evolution of the stock price could have been considered as a *random walk* along the nodes arranged as a tree with branches. Accordingly, in continuous time this evolution will be modeled by a continuous analogue of discrete random walk, called *Wiener process*, or, equivalently, *Brownian motion*. Hence, the plan is as follows. Firstly, we present some description of Wiener process (WP) and formulate the model of the market based on it, called BMS model for Black–Merton–Scholes.

Secondly, we will need to learn *integrate with respect to WP*: the notion of *stochastic integral* will be introduced. This has deep analogy with Riemenn's integral, and, hence, it is recommended to revise your Riemenn's integration skills, first of all definitions and main theorems as to why, e.g., any continuous function is integrable on a closed bounded interval, etc., however, cannot be reduced to "non-random" analysis, and will use some purely probabilistic notions of convergence for integral sums[1].

Thirdly, stochastic integration developed, we may introduce such object as *stochastic differential*, which is just a short notation for corresponding integral equations.

Fourthly, we will learn the stochastic analogue of composite function differentiation rule, called *Itô's formula*, in several versions.

Fifthly, once we have stochastic differentials, we may and ought to consider *stochastic differential equations* (SDE's), because our stock price, being some exponential expression, will satisfy some SDE, and, hence, it is important to learn how to solve such equations.

Sixthly, we will learn how to change measures in order to add or subtract drift in our SDE's, having in mind that in discrete model, the most useful tool was implied probability measure, under which all prices, asset's and options', were all discounted martingales. Now with the help of this new tool, Cameron–Martin–Girsanov's (CMG) theorem, we will lean how construct *implied measure* on BMS market. This tool will lead us straight to Black–Scholes formula, the goal of the module.

Seventhly, BS formula already established, yet an additional section will give some links to *Black–Scholes' PDE* of parabolic type, and it will be shown how solve this PDE via linear SDE's. This was the original method by Black and Scholes, and it does not use CMG theorem. Most of this section is advanced reading, except the section 24.

Eighthly, some short introduction to American option pricing will be provided, by using SDE's and *optimal stopping*. This whole section is an advanced reading.

[1] The presentation in this module is some simplification of the real matter which should be based on *Lebesgue's* integral instead of Riemenn's. However, discussion of this is beyond the level of our course.

NB: All *advanced reading* notices are given only for information, not required for the exam. *Further advanced reading* notes show links to recent achievements or open problems of science.

1 Random Walk very important

Wiener process or Brownian motion (BM) is a continuous time analogue of a random walk with very small steps. There is a strict theorem like CLT behind this.

Random walk $W_n(t)$ with steps $\delta = 1/n$ is defined by induction as

$$W_n(0) = 0;$$

$$W_n((k+1)/n) := W_n(k/n) + \xi_k/\sqrt{n};$$

$$W_n(t) := W_n([nt]/n),$$

where $[a] = \max(n = 0, \pm 1, \pm 2 : \ n \le a)$ (called *integer part of* a), and IID r.v.'s

$$\xi_k = \pm 1, \qquad \text{with probability } 1/2, \quad k = 1, 2, \ldots$$

We are going to study the limiting behaviour of the random walk $W_n(t)$. There are **three main properties** which will remain valid in the limit:

- $W_n(t) \approx \mathcal{N}(0, t)$,
- $W_n(t') - W_n(t) \approx \mathcal{N}(0, |t' - t|)$, and
- increments $W_n(j'/n) - W_n(j/n)$ (j, j' integers) are independent for disjoint intervals $(j/n, j'/n)$.

Let us see what occurs in the limit.

1. **What happens to the value $W_n(t)$ for fixed $t > 0$, as $n \to \infty$?** Denote $m = \sqrt{[nt]}$. By definition,

$$W_n(t) = \sum_{k=1}^{[nt]} \xi_k/\sqrt{n} = \frac{1}{\sqrt{n}} \sum_{k=1}^{[nt]} \xi_k$$

$$= t^{1/2} \frac{1}{\sqrt{nt}} \sum_{k=1}^{[nt]} \xi_k = t^{1/2} \frac{\sqrt{[nt]}}{\sqrt{nt}} \frac{1}{\sqrt{m}} \sum_{k=1}^{m} \xi_k.$$

Notice that

$$\frac{\sqrt{[nt]}}{\sqrt{nt}} \to 1, \quad n \to \infty.$$

By CLT (CLT-1),

$$\frac{1}{\sqrt{m}} \sum_{k=1}^{m} \xi_k \Longrightarrow \eta \sim \mathcal{N}(0,1).$$

Remind that $\Longrightarrow$ means weak convergence (revise this notion, handout 9!). Hence,

$$W_n(t) \Longrightarrow t^{1/2}\eta \sim t^{1/2}\mathcal{N}(0,1) \equiv \mathcal{N}(0,t), \quad n \to \infty.$$

2. **What happens to the increment $W_n(t) - W_n(s)$, $t > s$, as $n \to \infty$?**

   A similar reasoning implies a weak convergence

   $$W_n(t) - W_n(s) \Longrightarrow |t-s|^{1/2}\mathcal{N}(0,1) \equiv \mathcal{N}(0, |t-s|), \quad n \to \infty.$$

3. An essential property of "pre-limiting increments" $W_n(t) - W_n(s)$ is **independence** (follows from construction), that is,

   - if $i_1 < j_1 < i_2 < j_2 < \ldots < i_k < j_k$, then independent are the r.v.'s

     $$W_n(j_1/n) - W_n(i_1/n), \quad W_n(j_2/n) - W_n(i_2/n), \quad \ldots, \quad W_n(j_k/n) - W_n(i_k/n).$$

The properties 1–3 are preserved in the limit, hence, the following definition is approved:

2 Brownian Motion = Wiener process key topic

Definition 1 **Brownian motion** $(W(t),\, t \geq 0)$ *is a random process such that:*

1. $W(0) = 0$;
2. *the increments of $W(t)$ on disjoint intervals are independent; any increment $W(t) - W(s)$, $t > s$, is a Gaussian r.v. $\sim \mathcal{N}(0, t-s)$; in particular,*

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

3. *paths of $W(t)$ are continuous.*

(The latter does not follow directly from properties of W_n but can be derived from *2.*)

3 Black – Merton – Scholes stock price model key topic

We assume (compare to formula (2), lecture 10) that the stock price is represented as

$$\boxed{S_t = S_0 \exp(\mu t + \sigma W(t)).}$$

Here μ is called **drift** or **trend**, σ^2 is called **volatility** of the market. The main goal remains to value European options for this market model, and as prompted by the lecture 10, our guess is (since $\sqrt{T} \times \mathcal{N}(0,1) \sim \mathcal{N}(0,T)$),

$$C_0 = e^{-rT} \tilde{E} \left(S_0 e^{\left((r-\sigma^2/2)T + \sigma W_T\right)} - K \right)_+ .$$

However, there is a question, *how understand $\tilde{E}$ here?* In the other words, *what is implied measure on our continuous time market?* The answer to this question will be given below, in the Cameron–Martin–Girsanov theorem.

4 Wiener process properties key topic

Let us learn main properties of WP, which follow from the definition.

1. $$EW(t) = 0, \qquad EW(t)^2 = t.$$

   Exercise 1 *Show this by using the definition of WP.*

2. Paths of $W(t)$ are **continuous**. (Just believe that it may be proved[2].)

3. Paths of $W(t)$ are **not differentiable**. [Hence, $\int f(s)dW(s)$ cannot be defined as $\int f(s)W'(s)ds$.]
   Hint[3]. *Suppose they are differentiable. Consider partitions of some interval, say, $[0,T]$, into equal subintervals, and the sum $\sum_k (\Delta W(t_k))^2$. As partitions tend to zero, this sum should tend to zero since $\sum_k (\Delta W(t))^2 \approx \sum_k (W'(t_k))^2 (\Delta t_k)^2 \approx (\Delta t_k) \sum_k (W'(t_k))^2 (\Delta t_k) \approx (\Delta t_k)(\int_0^1 (W'(t)^2 dt) \to 0$. On the other hand, this sum has a constant mean value T, and by CLT tends weakly to T. [In fact, due to the Law of Large Numbers.]*

4. BM is a **Markov process**. Follows from similar property for $W_n(t)$. Proof not required.

[2]The key tool here is *Kolmogorov's theorem on continuous modification* of a process (advanced reading).
[3]This hint is advanced reading.

5. BM is a **martingale**. Follows from 4 (however, proof not required), the definition, and the property of conditional expectations for independent r.v.'s ($t' > t$):

$$E(W(t') - W(t)|W(t)) = E(W(t') - W(t)) = 0.$$

5 Stochastic Integrals (SI) key topic

- For any *continuous* stochastic process $f_t \in \mathcal{F}_t^W$ (remind that this **denotes** the assumption that f_t is a function determined by values W_s, $s \leq t$), such that $\int_0^t E f_s^2 \, ds < \infty$,

$$\int_0^t f_s \, dW_s := l.i.m._{\lambda \to 0} \sum_i f_{s_i} \left(W_{s_{i+1}} - W_{s_i} \right).$$

- Here $l.i.m.$ = limit in (square) mean, which is defined as

$$E \left| \int_0^t f_s \, dW_s - \sum_i f_{s_i} \left(W_{s_{i+1}} - W_{s_i} \right) \right|^2 \to 0, \quad \lambda \to 0.$$

- **Important:** for any continuous, bounded (and some unbounded[4]) f_t's $\in F_t^W$, the SI does exist. *I.e., the l.i.m. does exist: this is a Theorem (proof not required).*

- **Important:** f_{s_i} is taken at the left point s_i of each partition.

6 Properties of Stochastic Integrals key topic

Three Main Properties of SI *(you ought to learn by heart the properties and to understand the proofs provided)*

1. $$E \int_0^t f_s \, dW_s = 0.$$

2. $$E(\int_0^t f_s \, dW_s)^2 = \int_0^t E \, f_s^2 \, ds.$$

3. $$\int_0^t f_s \, dW_s \qquad \text{is a martingale (short notation "mart").}$$

[4]and many discontinuous (advanced reading)

The three main properties read,

- *1 – expectation of any SI equals zero.*
- *2 – variance of a SI has a representation via Riemenn's integral in the r.h.s.*
- *3 – any SI is a martingale as a function of the upper limit (time t.)*

We will <u>not</u> consider processes f without the property[5] $E\int_0^t f_s^2\,ds < \infty$**.**

Some elements of proofs.

1. - The SI $\int_0^t f_s\,dW_s$ is a l.i.m. of integral sums,

$$\sum_i f_{s_i}\Delta W_{s_i}.$$

   So it suffices to show the same property for such sums.

   - **Exercise 2** *Show it. <u>Hint:</u> use the following*

   Lemma 1 (Cauchy-Bouniakovsky-Schwarz) *For any r.v. η,*

$$|E\eta| \leq \sqrt{E\eta^2}.$$

   Exercise 3 *Show the lemma. <u>Hint:</u> use the property*

$$0 \leq var\,(\eta) = E\eta^2 - (E\eta)^2.$$

   - By properties of expectation,

$$E\sum_i f_{s_i}\Delta W_{s_i} = \sum_i E f_{s_i}\Delta W_{s_i},$$

   so it is sufficient to consider only one term.

[5] but in advanced theories they may be considered (advanced material)

- We have,

$$\begin{aligned}
Ef_{s_i}\Delta W_{s_i} &= EE[f_{s_i}\Delta W_{s_i}|F^W_{s_i}] \\
&= Ef_{s_i}E[\Delta W_{s_i}|F^W_{s_i}] \\
&\overset{\text{(by Markov property)}}{=} Ef_{s_i}E[\Delta W_{s_i}|W_{s_i}] \\
&= E\left(f_{s_i}\times 0\right) = 0.
\end{aligned}$$

The fourth equality holds true because the increments of BM are independent, in particular, so are

$$W_{s_{i+1}} - W_{s_i} \text{ and } W_{s_i} - W_0,$$

and remember that $W_0 = 0$. This means that ΔW_{s_i} does not depend on W_{s_i}. Hence, its conditional expectation equals non-expectation, that is,

$$E[\Delta W_{s_i}|W_{s_i}] = E[\Delta W_{s_i}] = 0.$$

So, the first property is a consequence of the fact that

$$EW(t) = 0, \quad \forall t.$$

2. (a) Let us show the property for **just one term**, that is,

$$E\left(f_{s_i}\Delta W_{s_i}\right)^2 = Ef^2_{s_i}(s_{i+1} - s_i).$$

Indeed,

$$\begin{aligned}
E\left(f_{s_i}\Delta W_{s_i}\right)^2 &= E[E[(f_{s_i}\Delta W_{s_i})^2\,|F^W_{s_i}]] \\
&= E\left[f^2_{s_i}E[(\Delta W_{s_i})^2\,|F^W_{s_i}]\right] \\
&\overset{\text{(by Markov property)}}{=} E\left[f^2_{s_i}E[(\Delta W_{s_i})^2\,|W_{s_i}]\right] \\
&\overset{\text{(as } \Delta W_{s_i} \text{ and } W_{s_i} \text{ are independent)}}{=} E\left[f^2_{s_i}E[(\Delta W_{s_i})^2]\right] \\
&\overset{\text{(as } E(\Delta W_{s_i})^2 = \Delta s_i)}{=} E\left[f^2_{s_i}\left(\Delta s_i\right)\right].
\end{aligned}$$

Informally, the property is based upon the fact that $EW^2(t) = t, \quad \forall t.$

(b) Consider two terms[6]. If we repeat the calculus for

$$E\left(f_{s_1}\Delta W_{s_1} + f_{s_2}\Delta W_{s_2}\right)^2,$$

we square the brackets $(\ldots)^2$ out and get similarly

$$E\left[f_{s_1}^2\left(\Delta s_1\right)\right] + E\left[f_{s_2}^2\left(\Delta s_2\right)\right],$$

and

$$2\,E\left[f_{s_1}\left(\Delta W_{s_1}\right) f_{s_2}\left(\Delta W_{s_2}\right)\right]$$

which is equal to zero. Indeed,

$$E\left[f_{s_1}\left(\Delta W_{s_1}\right) f_{s_2}\left(\Delta W_{s_2}\right)\right]$$

$$= E\left[E\left[f_{s_1}\left(\Delta W_{s_1}\right) f_{s_2}\left(\Delta W_{s_2}\right)\right]|\mathcal{F}_{s_2}^W\right]$$

$$= E\left[f_{s_1}\left(\Delta W_{s_1}\right) f_{s_2} E\left[\left(\Delta W_{s_2}\right)\right]|\mathcal{F}_{s_2}^W\right] = 0,$$

since

$$E\left[\left(\Delta W_{s_2}\right)|\mathcal{F}_{s_2}^W\right] = E\left[\left(\Delta W_{s_2}\right)|W_{s_2}\right] = 0,$$

because ΔW_{s_2} and W_{s_2} are independent.

3. We leave the third property **without proof**.

Exercise 4 *Let* $\eta \sim \mathcal{N}(0, \sigma^2)$. *Show that*

$$E\eta^4 = 3\sigma^4.$$

Hint: calculate the integral $\int x^4 (2\pi)^{-1/2} e^{-x^2/2}\,dx$, *using by parts integration.*

Exercise 5 *Using the definition and previous Exercise, show*[7] *that*

$$\int_0^t W_s\,dW_s = W_t^2/2 - t/2 \quad \left(\underline{\textbf{not}} \quad W_t^2/2\right).$$

[6] advanced reading

[7] *Without* the Itô formula the proof will not be required as advanced reading; but **with** the latter this is a must.

7 Stochastic differential key topic

Definition 2 *We say that the process X_t has a* **stochastic differential**

$$dX_t = b_t\,dt + \sigma_t\,dW_t$$

iff

$$X_t = X_0 + \int_0^t b_s\,ds + \int_0^t \sigma_s\,dW_s$$

where the first integral is a "usual" Riemann one while the second one is a **stochastic** *one. In particular, we require that σ is continuous.*

So, a stochastic *differential* is just a convention: all is understood in an **integral form**. **Nonetheless, one can understand it, roughly, as**

$$\Delta X_t \equiv X_{t+\delta} - X_t \approx b_t\delta + \sigma_t(W_{t+\delta} - W_t).$$

Example 1 *In the BMS model the stock price will be chosen as*

$$S_t = S_0 \exp(\mu t + \sigma W_t).$$

Q: *does S_t possess a stochastic differential? To answer this question we will need Itô's formula.*

8 Squared increments of BM very important

We discuss two sides of an important property of the increments of BM. The first one says that $(W_{t+\delta} - W_t)^2$ is approximately equal to δ for small $\delta > 0$; in the terms of **stochastic differentials**, it roughly means that

$$dW_t = \sqrt{dt}$$

(yet, the sign of ΔW_t may be negative as well as positive, which implies that $\sqrt{dt}$ – or rather $\pm\sqrt{dt}$ – should be treated with precautions). Indeed, such a notation was in use in this theory in 1960s. The second one states that the sum of squared increments of BM until time t equals approximately t.

1. $$(W_{t+\delta} - W_t)^2 \approx \delta, \quad \delta \to 0,$$

   in the following strict sense:

   $$E\left|(W_{t+\delta} - W_t)^2 - \delta\right|^2 = o(\delta), \quad \delta \to 0.$$

Indeed, $(W_{t+\delta} - W_t)^2 - \delta$ is a r.v. with a mean value zero and a variance

$$E\left|(W_{t+\delta} - W_t)^2 - \delta\right|^2 = 3\delta^2 - 2\delta \times \delta + \delta^2 = 2\delta^2 = o(\delta).$$

Exercise 6 *Show the statements. Hint: calculate Gaussian integrals.*

2.

$$\sum_{s_i \le t} (\Delta W_{s_i})^2 \approx t, \quad \lambda \approx 0.$$

Indeed ($\lambda := \max_i \Delta s_i$),

$$E \sum_{s_i \le t} (\Delta W_{s_i})^2 = \sum_{s_i \le t} E(\Delta W_{s_i})^2 = \sum_{s_i \le t} (\Delta s_i) = t,$$

while

$$var\left(\sum_{s_i \le t} (\Delta W_{s_i})^2\right) = \sum_{s_i \le t} var\left((\Delta W_{s_i})^2\right)$$

$$= \sum_{s_i \le t} 2(\Delta s_i)^2 \le 2 \sum_{s_i \le t} \lambda(\Delta s_i) = 2\lambda t \to 0, \quad \lambda \to 0.$$

So, the sum $\sum_{s_i \le t} (\Delta W_{s_i})^2$ is a random variable with a mean value t and a very small variance (i.e., $var \to 0$). Hence, it approximately equals t.

9 Itô's formula – 1 extremely important, to be learnt by heart

Consider a twice bounded differentiable function $g(w)$ and function $f(t, w)$ with one derivative in t and two in w (all bounded and continuous).

Theorem 1 (Itô's formula – 1)

$$d\, g(W_t) = g'(W_t) dW_t + (1/2) g''(W_t) dt$$

and

$$d\, f(t, W_t) = \frac{\partial f}{\partial t}(t, W_t) dt + \frac{\partial f}{\partial w}(t, W_t) dW_t + (1/2) \frac{\partial^2 f}{\partial w^2}(t, W_t) dt.$$

Proof: the idea. We say that

$$dg(W_t) = g'(W_t)dW_t + (1/2)g''(W_t)dt,$$

loosely speaking, if

$$\Delta g(W_t) \approx g'(W_t)\Delta W_t + (1/2)g''(W_t)\Delta t.$$

Hence, let us consider $\Delta g(W_t)$. By Taylor's expansion,

$$\Delta g(W_t) \equiv g(W_{t+\delta}) - g(W_t)$$

$$\approx g'(W_t)\Delta W_t + (1/2)g''(W_t)(\Delta W_t)^2 + o((\Delta W_t)^2).$$

Hence,

$$g(W_t) - g(W_0) \approx \sum_i g'(W_{s_i})\Delta W_{s_i} + (1/2)\sum_i g''(W_{s_i})(\Delta W_{s_i})^2 + o((\Delta W_{s_i})^2).$$

Since

$$(\Delta W_{s_i})^2 \approx \Delta s_i,$$

we get the desired assertion concerning function g. Indeed,

$$g(W_t) - g(W_0) = \sum_i g'(W_{s_i})\Delta W_{s_i} + (1/2)\sum_i g''(W_{s_i})(\Delta W_{s_i})^2$$

$$\approx \sum_i g'(W_{s_i})\Delta W_{s_i} + (1/2)\sum_i g''(W_{s_i})(\Delta s_i).$$

Here the second integral sum is an approximation to the Riemann integral $\int_0^t (1/2)g''(W_s)\,ds$ while the first one approximates the stochastic integral $\int_0^t g'(W_s)\,dW_s$. **Function $f(t,x)$ can be considered similarly.**

Both formulas can be justified for many unbounded functions such as x^2, $\exp(x)$, tx^3, $\exp(-t+x)$, *etc.*[8]

Example 2 *1. By the Itô formula,*

$$d(W_t^2/2) = W_t\,dW_t + (1/2)\,(dW_t)^2 = W_t\,dW_t + (1/2)\,dt.$$

2. So,

$$d(W_t^2/2 - t/2) = W_t\,dW_t + (1/2)\,dt - (1/2)\,dt = W_t\,dW_t.$$

[8]Proof is advanced reading, but we will use this result for such functions without special notice.

3. Hence,

$$\int_0^t W_s\, dW_s = W_t^2/2 - t/2.$$

Example 3 *A stochastic differential of a stock price* $S_t = S_0 \exp(\mu t + \sigma W_t)$ *in the BMS model is, by Itô's formula,*

$$dS_t = S_t \left(\mu dt + \sigma dW_t + (1/2)\sigma^2 dt\right).$$

This is an example of a linear **stochastic differential equation***.*

Exercise 7 *Find the expressions for* dW_t^n*,* $n = 3, 4, \ldots$

10 Itô's formula: symbolic rules _{important}

The formula may be remembered symbolically as follows. Imagine that we are writing Taylor's expansion for $dg(W_t)$ via (dW_t), $(dW_t)^2$, and so on. Then we get, formally,

$$dg(W_t) = g'(W_t)dW_t + (1/2)g''(W_t)(dW_t)^2 + (1/3!)g'''(W_t)(dW_t)^3 + \ldots$$

Now we apply **the symbolic rules:**

$$\boxed{(dW_t)^2 = dt,} \quad (dW_t)^3 = (dW_t)^4 = \ldots = 0.$$

To tackle $f(t, W_t)$ an additional rule is applied:

$$\boxed{(dW_t)(dt) = 0,} \quad \& \quad (dt)^2 = (dt)^3 = (dW_t)^2(dt) = (dW_t)(dt)^2 = \ldots = 0.$$

11 Itô's formula – 2, chain rule very important

The following result includes the previous theorem and usually is called *Itô's formula*. It applies not only to WP, but to any process with a stochastic differential. Thus, WP becomes just a special case with a stochastic differential dW_t.

Theorem 2 (Itô's formula as a chain rule) *Let process* X_t *have a stochastic differential,*

$$dX_t = b_t dt + \sigma_t dW_t.$$

Then for any function $g(x)$ *with two continuous bounded derivatives,*

$$d\, g(X_t) = g'(X_t)dX_t + \frac{1}{2}g''(W_t)(dX_t)^2$$

$$= \left[b_t g'(X_t) + \frac{1}{2}\sigma_t^2 g''(X_t)\right] dt + g'(X_t)\sigma_t dW_t$$

and for any function $f(t,x)$ with two continuous bounded derivatives in x and one in t,

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

$$= \left[\frac{\partial f}{\partial t}(t,X_t) + b_t\frac{\partial f}{\partial x}(t,X_t) + \frac{1}{2}\sigma_t^2\frac{\partial^2 f}{\partial x^2}(t,X_t)\right]dt + \frac{\partial f}{\partial x}(t,X_t)\sigma_t dW_t$$

where according to symbolic rules of Itô's calculus, *(most important!)*

$$(dX_t)^2 = (b_t dt + \sigma_t dW_t)^2 = b_t^2(dt)^2 + 2b_t\sigma_t dt dW_t + \sigma_t^2(dW_t)^2 = 0 + 0 + \sigma_t^2 dt.$$

The theorem can be also applied for g and f with *locally* bounded derivatives, if all integrals (Riemann and stochastic) are well-defined, for example, $g(x) = x^n$, $n = 1, 2, \ldots$ and $g(x) = \exp(ax)$.

Proof (the idea). We assume that b_t and σ_t are bounded and continuous. Similarly to the special case $X_t = W_t$, we say that

$$dg(X_t) = g'(X_t)dX_t + \frac{1}{2}g''(X_t)dt,$$

loosely speaking, if

$$\Delta g(X_t) \approx g'(X_t)\Delta X_t + \frac{1}{2}g''(X_t)(\Delta X_t)^2.$$

Let us consider $\Delta g(X_t)$. By Taylor's expansion,

$$\Delta g(X_t) \equiv g(X_{t+\delta}) - g(X_t) \approx g'(X_t)\Delta X_t + \frac{1}{2}g''(X_t)(\Delta X_t)^2 + o((\Delta X_t)^2).$$

Hence,

$$g(X_t) - g(X_0) \approx \sum_i g'(X_{s_i})\Delta X_{s_i} + \frac{1}{2}\sum_i g''(X_{s_i})(\Delta X_{s_i})^2 + o((\Delta X_{s_i})^2).$$

Now substitute

$$\Delta X_{s_i} \approx \sigma_{s_i}\Delta W_{s_i} + b_{s_i}\Delta s_i$$

and remember

$$(\Delta W_{s_i})^2 \approx \Delta s_i, \quad (\Delta s_i)^2 \approx 0, \quad (\Delta s_i)(\Delta W_{s_i}) \approx 0.$$

Hence,

$$g(X_t) - g(X_0)$$

$$\approx \sum_i g'(X_{s_i})(b_{s_i}\Delta s_i + \sigma_{s_i}\Delta W_{s_i}) + \frac{1}{2}\sum_i g''(X_{s_i})(b_{s_i}\Delta s_i + \sigma_{s_i}\Delta W_{s_i})^2$$

$$\approx \sum_i g'(X_{s_i})\sigma_{s_i}\Delta W_{s_i} + \sum_i g'(X_{s_i})b_{s_i}\Delta s_i + \frac{1}{2}\sum_i g''(X_{s_i})\Delta s_i.$$

Here the integral sums approximate corresponding Riemann integrals

$$\int_0^t g'(X_s)\sigma_s\,dW_s + \int_0^t g'(X_s)b_s\,ds + \frac{1}{2}\int_0^t g''(X_s)\sigma_s^2\,ds.$$

Function f is considered similarly. Remind that for many g's and f's which are unbounded and with unbounded derivatives the Itô formula can be justified as well.

Exercise 8 *Find a stochastic differential of* $X_t = \exp(aW_t + bt)$.

Solution. By the Itô formula,

$$dX_t = aX_t dW_t + bX_t dt + \frac{1}{2}a^2 X_t dt = (b + a^2/2)X_t dt + aX_t dW_t.$$

Exercise 9 *Find a solution to a stochastic differential equation (SDE)*

$$dX_t = aX_t dW_t, \quad X_0 = 1.$$

Solution. In the previous Exercise let $b = -a^2/2$; then for $X_t = \exp(aW_t - a^2t/2)$,

$$dX_t = aX_t dW_t, \quad \text{and} \quad X_0 = \exp(0) = 1.$$

12 Itô's formula for complex valued functions useful

Let $(W_t,\, t \ge 0)$ be a Wiener process, and $(f(x),\, x \in R^1)$ be some complex valued function, that is,

$$f(x) = A(x) + i\,B(x).$$

Let us assume that functions A and B have two continuous derivatives. Then, we can apply Itô's formula to $A(W_t)$ and to $B(W_t)$:

$$dA(W_t) = A'(W_t)\,dW_t + \frac{1}{2}A''(W_t)\,dt,$$

and

$$dB(W_t) = B'(W_t)\,dW_t + \frac{1}{2}B''(W_t)\,dt.$$

In the integral form, equivalently, we have,

$$A(W_t) - A(0) = \int_0^t A'(W_s)\,dW_s + \int_0^t \frac{1}{2}A''(W_s)\,ds,$$

and

$$B(W_t) - B(0) = \int_0^t B'(W_s)\,dW_s + \int_0^t \frac{1}{2}B''(W_s)\,ds.$$

Hence, multiplying the latter line by $i = \sqrt{-1}$ and taking a sum, we get the following version of Itô's formula:

$$A(W_t)+i\,B(W_t)-A(0)-i\,B(0) = \int_0^t \left(A'(W_s) + iB'(W_s)\right)\,dW_s + \int_0^t \frac{1}{2}\left(A''(W_s) + iB''(W_s)\right)\,ds.$$

Equivalently,

$$\boxed{f(W_t) - f(0) = \int_0^t f'(W_s)\,dW_s + \int_0^t \tfrac{1}{2}f''(W_s)\,ds.}$$

EXAMPLE!!! to insert

13 Itô's formula, one more application (useful)

Frequently, there is a situation where there are two processes, say, X_t and Y_t, with stochastic differentials,

$$dX_t = \sigma_t\,dW_t + b_t\,dt,$$

and

$$dY_t = \alpha_t\,dW_t + \beta_t\,dt.$$

How to find out the stochastic differential of the product X_tY_t?

Theorem [*Itô's formula for products*]

$$\boxed{d(X_t\,Y_t) = X_t\,dY_t + Y_t\,dX_t + (dX_t)(dY_t)} \tag{1}$$

$$\equiv X_t\,(\alpha_t\,dW_t + \beta_t\,dt) + Y_t\,(\sigma_t\,dW_t + b_t\,dt) + \sigma_t\alpha_t\,dt.$$

Corollary [*Itô's formula for product of two SI's*]
If

$$dX_t = \sigma_t \, dW_t, \qquad dY_t = \alpha_t \, dW_t,$$

then

$$d(X_t \, Y_t) = X_t \, dY_t + Y_t \, dX_t + (dX_t)(dY_t) \tag{2}$$

$$\equiv X_t \, \alpha_t \, dW_t + Y_t \, \sigma_t \, dW_t + \sigma_t \alpha_t \, dt.$$

In particular,

$$\boxed{E(X_t \, Y_t) - E(X_0 \, Y_0) = \int_0^t E\sigma_t \alpha_s \, ds.}$$

(Compare to the second main property of SI.)

Proof of the Theorem: the idea. To show (29), it suffices to consider finite differences,

$$\Delta(X_t Y_t) = X_{t+\Delta t} Y_{t+\Delta t} - X_t Y_t = X_{t+\Delta t} Y_{t+\Delta t} - X_{t+\Delta t} Y_t + X_{t+\Delta t} Y_t - X_t Y_t$$

$$= X_{t+\Delta t} \Delta Y_t + Y_t \Delta X_t = X_t \Delta Y_t + (X_{t+\Delta t} - X_t) \Delta Y_t + Y_t \Delta X_t$$

$$\equiv X_t \, \Delta Y_t + Y_t \, \Delta X_t + (\Delta X_t)(\Delta Y_t).$$

On an intuitive level, this suffices to treat it as (1).

14 Stochastic Differential Equations very important

A general SDE has a form

$$dX_t = \sigma(X_t) dW_t + b(X_t) dt, \quad X_0 = x. \tag{3}$$

We assume functions σ and b to be bounded and such that

$$|\sigma(x) - \sigma(x')| + |b(x) - b(x')| \le C|x - x'|, \quad \forall x, x' \tag{4}$$

(e.g., with bounded derivatives).

Definition 3 *X_t, $t \ge 0$ is a* **solution** *of the equation (8) if*

1. *X_t is continuous;*

2. for any t and any $a \in R^1$, $\{X_t \le a\} \in F_t^W$;

3. for any $t \ge 0$,

$$X_t = x + \int_0^t \sigma(X_s)\, dW_s + \int_0^t b(X_s) ds. \tag{5}$$

Theorem 3 (Itô) *Under assumptions (4) the equation (8) has a unique solution for any $t \ge 0$.*

How to solve a SDE practically? You must know two main cases in which a SDE can be solved explicitly: with constant and linear coefficients. If there is no explicit solution, one may try to solve a SDE approximately. We will not prove general theorems about existence and uniqueness; however you should know that under our assumptions equation (8) possesses a unique solution. In textbooks you may meet notions of **strong and weak solutions.** Our solutions are always strong (it means item 2 of the definition).

15 SDE with constant coefficients (non-compulsory reading)

Consider a SDE

$$dX_t = \sigma\, dW_t + b\, dt, \quad X_0 = x. \tag{6}$$

Theorem 4 *Solution of equation (6) is given by the formula*

$$X_t = x + \sigma\, W_t + b\, t, \quad t \ge 0.$$

Proof. By Itô's formula,

$$dX_t = \sigma\, dW_t + b\, dt, \quad \text{and initial data is} \quad X_0 = x.$$

Notice that if Y_t is **another** solution then

$$Y_t - X_t = x + \sigma\, W_t + b\, t - x - \sigma\, W_t - b\, t = 0,$$

that is, Y_t coincides with X_t, that is, solution is unique.

16 SDE with linear coefficients (compulsory and important)

Consider a SDE

$$dX_t = a\, X_t\, dW_t + b\, X_t\, dt, \quad X_0 = x. \tag{7}$$

Theorem 5 *Solution of equation (7) is given by the formula*

$$X_t = x\ \exp\Big((b - a^2/2)t + aW_t\Big), \quad t \geq 0.$$

Proof. Initial data is $X_0 = x$. By Itô's formula,

$$dX_t = x\, d\exp\Big((b - a^2/2)t + aW_t\Big)$$

$$= x\exp\Big((b - a^2/2)t + aW_t\Big)\Big((b - a^2/2)\, dt + a\, dW_t + \frac{1}{2}a^2\, dt\Big)$$

$$= X_t\ (b\, dt + a\, dW_t) = b\, X_t\, dt + a\, X_t\, dW_t.$$

Notice that if Y_t is another solution then by Itô's formula, $d(Y_t/X_t) = 0$ (*Exercise: show this!*[9]), hence, $Y_t/X_t = const$, hence, $Y_t/X_t = Y_0/X_0 = x/x = 1$ (if $x \neq 0$), that is, Y_t coincides with X_t, i.e. solution is unique. If $x = 0$ then $X_t = 0$, and **it can be shown** that any other solution Y_t to $Y_t = \int_0^t \sigma Y_s dW_s + \int_0^t bY_s ds$ must be identically zero (*Exercise, not compulsory*).

17 Approximate solutions of SDEs (only needed for Practical)

We consider the **Euler method** of approximation solution: fix some $h > 0$ and define

$$X_0 = x,$$

$$X_h = X_0 + \sigma(X_0)(W_h - W_0) + b(X_0)h, \quad \text{etc.},$$

$$X_{(k+1)h} = X_{kh} + \sigma(X_{kh})(W_{(k+1)h} - W_{kh}) + b(X_{kh})h, \quad k \geq 0.$$

If h is small, this is a reasonable approximation.

If a solution for *any* t is needed, the Euler method may be extended as

$$X_{kh+s} = X_{kh} + \sigma(X_{kh})(W_{kh+s} - W_{kh}) + b(X_{kh})s, \quad 0 \leq s \leq h.$$

This method is widely used for SDE simulations, and you will be asked to do it using **R**.

Note for the Practical. There might be problems with solution if $|\sigma(x)|, |b(x)|$ increase faster than $C|x|$ as $|x| \to \infty$: for some times solutions may not exist. Simulations may show the divergence after a finite number of steps, so if you meet such a case, it is perhaps not a fault of a programme.

For further advanced reading see [Mil, Ver03], etc.

[9] Not compulsory

18 Generator (to learn the definition)

For any SDE

$$dX_t = \sigma(X_t)\,dW_t + b(X_t)\,dt, \quad X_0 = x,$$

we call *a generator* of the process X_t the following operator,

$$L := \frac{\sigma^2(x)}{2}\frac{\partial^2}{\partial x^2} + b(x)\frac{\partial}{\partial x}.$$

This is frequently used while applying *Itô's formula, which may be rewritten in the form,*

$$dg(X_t) = Lf(X_t)\,dt + g'(X_t)\sigma(X_t)\,dW_t.$$

Emphasize that the term generator relates only to an SDE.

19 Cameron-Martin-Girsanov transformation (key topic)

In the CRR model, we have seen how to calculate the implied probabilities at each step and for each branch of the binomial tree. We can say that we **transform** the real probabilities p and $1-p$ for each one-step branch into implied ones, and this procedure gives us the probability for any "multi-step branch", which is the outcome for our probability space. Hence, we can find the implied probability of any event, and calculate the implied expectation on this sample space.

In the BMS continuous time model, the real probabilities are defined by Wiener process. The corresponding probability measure P can be calculated, for the event $A = (a < W_T \leq b)$, as

$$P(A) = P\left(a < W_T \leq b\right) = \int_a^b (2\pi T)^{-1/2} \exp\left(-\frac{z^2}{2T}\right)\,dz.$$

The definition of the expectation of any function g of W_T, $g(W_T)$ is

$$Eg\left(W_T\right) = \int_{-\infty}^{\infty} g(z)(2\pi T)^{-1/2} \exp\left(-\frac{z^2}{2T}\right)\,dz.$$

We may say that W_T's are outcomes (as far as event is defined via the value W_T). Hence, we may ask how to find the analogue for implied probabilities in this model. It should be a new "implied measure" $\tilde{P}$ under which the stock price

$$S_t = S_0 \exp(\mu t + \sigma W_t), \quad 0 \leq t \leq T,$$

is a discounted martingale. We are looking for an implied probability $\tilde{P}(A)$ by the formula

$$\tilde{P}(A) \equiv \tilde{E}1(A) := E\gamma_T 1(A),$$

with some density γ_T which is to be determined. For example, for $A = (a < W_T \leq b)$ the last formula would look like

$$\tilde{P}(a < W_T \leq b) \equiv \tilde{E}1(a < W_T \leq b) = E\gamma_T 1(a < W_T \leq b),$$

and the expectation of $g(W_T)$ w.r.t. $\tilde{P}$ would have a form

$$\tilde{E}g\left(W_T\right) = E\gamma_T g\left(W_T\right).$$

The theorem due to Cameron - Martin and Girsanov describes the general form of possible densities γ_T. In our case of BMS model with constant σ and b, this general form is

$$\gamma_T = \exp(-cW_T - \frac{c^2}{2}T).$$

So, the aim is to choose a constant c, so that the process S_t would be a discounted martingale. Further, the call price C_t must be a discounted martingale as well, as it was in the pre-limiting CRR model.

Given a density γ_T, the formulas for "tilde-probability" and "tilde-expectation" read

$$\tilde{P}(a < W_T \leq b) = E\gamma_T 1(a < W_T \leq b)$$

$$= \int_a^b \exp\left(-cz - c^2T/2\right)(2\pi T)^{-1/2}\exp\left(-\frac{z^2}{2T}\right)\,dz$$

and

$$\tilde{E}g\left(W_T\right) = E\gamma_T g\left(W_T\right)$$

$$= \int_{-\infty}^{\infty} \exp\left(-cz - c^2T/2\right) g(z)(2\pi T)^{-1/2}\exp\left(-\frac{z^2}{2T}\right)\,dz.$$

Definition 4 *We call a process $W_t + at$, $0 \leq t \leq T$,* **a Wiener process with a drift***; we say that the drift is at and the drift coefficient is equal to a.*

Notice that "Wiener process with a drift" is *not a "Wiener process"*. In particular, its mean value at time t is not zero.

Theorem 6 *Under the measure $\tilde{P}$ with $\gamma_T = \exp(-cW_T - \frac{c^2}{2}T)$, the process $W_t, 0 \leq t \leq T$, is a Wiener process with a drift $-ct$, while the process*

$$\tilde{W}_t := W_t + ct$$

is a Wiener process.

We will make some steps towards a rigorous proof.

1. γ_T is a density in the sense that $\gamma_T \geq 0$ and

$$\tilde{P}(\Omega) = \tilde{E}1 = 1.$$

   Indeed, we have,

$$\tilde{E}\,1 = \frac{1}{\sqrt{2\pi\,T}} \int 1\; \exp(-cz - c^2T/2)\; \exp(-z^2/(2T))\, dz = 1$$

   (see Exercises 3).

2. A new mean value: $\tilde{E}W_T = -cT$. Indeed,

$$\begin{aligned}
\tilde{E}\,W_T &= \frac{1}{\sqrt{2\pi\,T}} \int z\; \exp(-cz - c^2T/2)\; \exp(-z^2/(2T))\, dz \\
&= \frac{1}{\sqrt{2\pi\,T}} \int (z - cT + cT)\; \exp(-(z + cT)^2/(2T))\, dz \\
&= 0 - cT\, \frac{1}{\sqrt{2\pi\,T}} \int \exp(-(z + cT)^2/(2T))\, dz = -cT.
\end{aligned}$$

3. For $t < T$ we have, $\tilde{E}W_t = -ct$. (Will follow from step 6 below.)

4. A "tilde mean value" of $\tilde{W}$: $\tilde{E}\tilde{W}_t = 0$. Let us show this for $t = T$:

$$\tilde{E}\tilde{W}_T = \tilde{E}(W_T + cT) = (\tilde{E}W_T) + cT = -cT + cT = 0.$$

5. $\tilde{E}\tilde{W}_t^2 = t$. This will follow from the next item below.

6. $\tilde{W}_t$ is Gaussian under measure $\tilde{P}$. We show this for $t = T$: the new distribution function is

$$\tilde{P}(\tilde{W}_T \le x) = \tilde{E}1(\tilde{W}_T \le x) = E\gamma_T\, 1(\tilde{W}_T \le x)$$

$$= E\exp(-cW_T - c^2T/2)\, 1(W_T \le x - cT)$$

$$= \int_{-\infty}^{x-cT} \exp(-cy - c^2T/2)\frac{1}{\sqrt{2\pi T}} \exp(-y^2/(2T))\, dy$$

$$= \int_{-\infty}^{x-cT} \frac{1}{\sqrt{2\pi T}} \exp(-(y+cT)^2/(2T))\, dy = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi T}} \exp(-y^2/(2T))\, dy.$$

Hence, a new density is

$$\tilde{p}_T(x) = \frac{\partial \tilde{P}(\tilde{W}_T \le x)}{\partial x} = \frac{1}{\sqrt{2\pi T}} \exp(-x^2/(2T)),$$

which is indeed a Gaussian density, namely, of distribution $\mathcal{N}(0, T)$.

7. Denote[10] $\gamma_{t,T} = \exp(-c(W_T - W_t) - c^2(T-t)/2)$, so that

$$\gamma_T = \gamma_t \gamma_{t,T}.$$

Then one may extend explanations in 3. – 6. to all $t < T$ using the last formula and *independence* of the random variables γ_t and $\gamma_{t,T}$, – show why, – and the fact that $E\gamma_{t,T} = 1$, – also show why. (Both questions are not compulsory exercises.)

20 Black-Scholes' formula via Cameron-Martin-Girsanov (very important)

This is a way to find the value of call option which does not require anything but Gaussian integrals from Exercises 1 and 2[11]. We assume again that the stock price is given by the equation

$$S_t = S_0 \exp(\mu t + \sigma W_t).$$

[10] This item is advanced reading.
[11] And CMG Theorem, of course!

We are going to find a new probability measure $\tilde{P}$ under which $e^{-rt}S_t$, $0 \le t \le T$, is a martingale. For this aim we represent W_t in the form (c is a constant which will be chosen later)

$$W_t = \tilde{W}_t - ct, \quad t \ge 0,$$

which is equivalent to

$$\tilde{W}_t = W_t + ct, \quad t \ge 0.$$

Due to Cameron-Martin-Girsanov theorem, ($\tilde{W}_t$, $0 \le t \le T$) is a martingale under the measure $\tilde{P}$ defined by its density

$$\gamma_T = \exp(-cW_T - c^2T/2).$$

Now, S_t may be represented using $\tilde{W}_t$ by the formula

$$S_t = S_0 \exp(\mu t + \sigma(\tilde{W}_t - ct)) = S_0 \exp((\mu - \sigma c)t + \sigma\tilde{W}_t),$$

so that

$$\exp(-rt)S_t = S_0 \exp((\mu - \sigma c - r)t + \sigma\tilde{W}_t). \tag{8}$$

Lemma 2 (very important) *Let $c = (\mu - r + \sigma^2/2)/\sigma$. Then $e^{-rt}S_t$, $0 \le t \le T$, is a martingale under $\tilde{P}$.*

Proof. By the Itô formula, the stochastic differential of $\exp(-rt)S_t$ via $\tilde{W}_t$ (see (8)) is equal to

$$d\,(\exp(-rt)S_t) = S_t[\mu - \sigma c - r + \sigma^2/2]\,dt + S_t\sigma\,d\tilde{W}_t.$$

The choice of c means that $\mu - \sigma c - r + \sigma^2/2 = 0$. So, $d\,(\exp(-rt)S_t) = S_t\sigma\,d\tilde{W}_t$. Therefore, $(\exp(-rt)S_t)$ is a martingale under $\tilde{P}$.

Assume that under the implied probability measure $\tilde{P}$ the price C_t of the European call option is a discounted martingale, that is, the process $\exp(-rt)C_t$ is a martingale. The justification of such assumption is that this is true for the CRR model with small steps $\delta > 0$; hence, it should be true in the limit.

Theorem 7 (Black-Scholes, very important**)** *The function C_t is given by the formula ($0 \le t \le T$)*

[*BS*]

$$Y(t, S) = S\Phi(d_+) - Ke^{-r(T-t)}\Phi(d_-),$$

where

$$d_\pm = d_\pm(t) = \frac{\log(S/K) + (r \pm \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}.$$

Proof:

1. It suffices to show the formula for $t = 0$.
2. All we have to do is calculate the expectation (remember that $\mu - c\sigma = r - \sigma^2/2$)

$$C_0 = e^{-rT}\tilde{E}(S_T - K)_+ \tag{9}$$

$$= e^{-rT}\tilde{E}\left(S_0 e^{\mu T + \sigma(\tilde{W}_T - cT)} - K\right)_+$$

$$= e^{-rT}\tilde{E}\left(S_0 e^{\left(r-\sigma^2/2\right)T + \sigma \tilde{W}_T} - K\right)_+ .$$

3. Since $\tilde{W}_T \sim \mathcal{N}(0, T)$ under measure $\tilde{P}$, the random variable $\tilde{W}_T$ has the same normal distribution as $\sqrt{T}\xi$ where $\xi \sim \mathcal{N}(0, 1)$ (always under measure $\tilde{P}$). So we may rewrite the last expression in the form

$$C_0 = e^{-rT}\tilde{E}\left(S_0 e^{\left(r-\sigma^2/2\right)T + \sigma\sqrt{T}\xi} - K\right)_+$$

$$= e^{-rT}\int \left(S_0 e^{\left(r-\sigma^2/2\right)T + \sigma\sqrt{T}z} - K\right)_+ p_1(z)\, dz.$$

We integrate over the half-line

$$z: \; S_0 e^{\left(r-\sigma^2/2\right)T + \sigma\sqrt{T}z} \geq K$$

(for all other z the expression in the brackets is equal to zero), or equivalently,

$$z: \; e^{\left(r-\sigma^2/2\right)T + \sigma\sqrt{T}z} \geq K/S_0,$$

or equivalently,

$$z: \; \left(r - \sigma^2/2\right)T + \sigma\sqrt{T}z \geq \log(K/S_0),$$

or equivalently,

$$z: \; \sigma\sqrt{T}z \geq \log(K/S_0) - \left(r - \sigma^2/2\right)T,$$

or, at last,

$$z \geq \frac{\log(K/S_0) - (r - \sigma^2/2)\,T}{\sigma\sqrt{T}}.$$

4. So (we write S instead of S_0),

$$C_0 = e^{-rT} \int_{\frac{\log(\frac{K}{S})-\left(r-\frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}}^{\infty} \left(S e^{\left(r-\frac{\sigma^2}{2}\right)T+\sigma\sqrt{T}z} - K \right) p_1(z)\, dz$$

$$= e^{-rT} S \int_{\frac{\log(\frac{K}{S})-\left(r-\frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}}^{\infty} e^{\left(r-\frac{\sigma^2}{2}\right)T+\sigma\sqrt{T}z} p_1(z)\, dz - e^{-rT} K \int_{\frac{\log(\frac{K}{S})-\left(r-\frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}}^{\infty} p_1(z)\, dz$$

$$= e^{-rT} S \int_{-\infty}^{-\frac{\log(\frac{K}{S})-\left(r-\frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}} e^{\left(r-\frac{\sigma^2}{2}\right)T-\sigma\sqrt{T}z} p_1(z)\, dz - e^{-rT} K \int_{-\infty}^{-\frac{\log(\frac{K}{S})-\left(r-\frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}} p_1(z)\, dz$$

$$= e^{-rT} S \int_{-\infty}^{\frac{\log(\frac{S}{K})+\left(r-\frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}} e^{rT}\, e^{\left(-\frac{\sigma^2}{2}\right)T-\sigma\sqrt{T}z} p_1(z)\, dz - e^{-rT} K \int_{-\infty}^{\frac{\log(\frac{S}{K})+\left(r-\frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}} p_1(z)\, dz$$

$$= S \int_{-\infty}^{\frac{\log(\frac{S}{K})+\left(r-\frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}} e^{\left(-\frac{\sigma^2}{2}\right)T-\sigma\sqrt{T}z} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}\, dz - e^{-rT} K \Phi\left(\frac{\log(\frac{S}{K}) + \left(r-\frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}} \right)$$

$$= S \int_{-\infty}^{\frac{\log(\frac{S}{K})+\left(r-\sigma^2/2\right)T}{\sigma\sqrt{T}}} \frac{1}{\sqrt{2\pi}} e^{-(z+\sigma\sqrt{T})^2/2}\, dz - e^{-rT} K \Phi\left(\frac{\log(S_0/K) + \left(r-\frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}} \right)$$

$$= S \int_{-\infty}^{\frac{\log(\frac{S}{K})+\left(r+\frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}} \frac{1}{\sqrt{2\pi}} e^{-(z')^2/2}\, dz' - e^{-rT} K \Phi\left(\frac{\log(\frac{S}{K}) + \left(r-\frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}} \right)$$

$$= S\Phi\left(\frac{\log(\frac{S}{K}) + (r + \sigma^2/2)\,T}{\sigma\sqrt{T}}\right) - e^{-rT}K\Phi\left(\frac{\log(\frac{S}{K}) + \left(r - \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}\right).$$

21 Solving PDE: constant coefficients – 1

(advanced reading)

Consider a Partial Differential Equation (PDE) with constant coefficients,

$$u_t + \frac{1}{2}\sigma^2 u_{xx} + b u_x = 0, \quad u(T, x) = g(x) \tag{10}$$

("heat equation with a drift"). Consider a solution to SDE

$$dX_t = \sigma dW_t + b dt,\ t \ge 0, \quad \text{with initial data } X_0 = x. \tag{11}$$

Then,

$$X_t = x + \sigma W_t + bt \sim \mathcal{N}(x + bt, \sigma^2 t). \tag{12}$$

Suppose $u(t, x)$ is a solution to (10). Apply the Itô formula to $u(t, X_t)$:

$$du(t, X_t) = [u_t(t, X_t) + \frac{1}{2}\sigma^2 u_{xx}(t, X_t) + b u_x(t, X_t)]\, dt + \sigma u_x(t, X_t)\, dW_t = \sigma u_x(t, X_t)\, dW_t.$$

Hence,

$$u(T, X_T) - u(0, x) = \int_0^t \sigma u_x(s, X_s)\, dW_s,$$

and therefore[12] (notation E_x means *expectation for the process starting from x*)

$$u(0, x) = E_x\, u(T, X_T) = E_x\, g(X_T); \quad \text{likewise,}\ \ u(t, x) = E_x\, g(X_{T-t}). \tag{13}$$

[12]Very advanced reading notice: The method for general coefficients, even not non-degenerate, was originally introduced by I. I. Gikhman in 1949-1950, one of the founders of stochastic analysis.

22 Solving PDE: constant coefficients – 2

(advanced reading)

Consider a PDE with constant coefficients and a "potential" $(-ru)$,

$$u_t + \frac{1}{2}\sigma^2 u_{xx} + b u_x - ru = 0, \quad u(T,x) = g(x). \tag{14}$$

Consider a solution to SDE (11). Suppose $u(t,x)$ is a solution to (14). Apply the Itô formula to $\exp(-rt)u(t,X_t)$:

$$d\exp(-rt)u(t,X_t) = \exp(-rt)[u_t(t,X_t) + \frac{1}{2}\sigma^2 u_{xx}(t,X_t) + bu_x(t,X_t) - ru(t,X_t)]\,dt$$

$$+ \exp(-rt)\sigma u_x(t,X_t)\,dW_t = \exp(-rt)\sigma u_x(t,X_t)\,dW_t.$$

Hence,

$$\exp(-rT)u(T,X_T) - u(0,x) = \int_0^t \exp(-rs)\sigma u_x(s,X_s)\,dW_s,$$

and so

$$u(0,x) = \exp(-rT)E_x\, g(X_T); \quad \text{likewise,}\ \ u(t,x) = \exp(-r(T-t))E_x\, g(X_{T-t}). \tag{15}$$

23 Solving PDE via SDE: linear coefficients – 1

(advanced reading)

Consider a PDE with coefficients $\sigma^2 x^2/2$ and $b\,x$,

$$u_t + \frac{1}{2}\sigma^2 x^2 u_{xx} + bxu_x = 0, \quad u(T,x) = g(x). \tag{16}$$

Consider a solution to linear SDE

$$dX_t = \sigma X_t dW_t + bX_t dt, \qquad X_0 = x. \tag{17}$$

Then,

$$X_t = x\exp(bt - \sigma^2 t/2 + \sigma W_t). \tag{18}$$

Suppose $u(t,x)$ is a solution to (10). Apply the Itô formula to $u(t,X_t)$:

$$du(t,X_t) = [u_t(t,X_t) + \frac{1}{2}\sigma^2 X_t^2 u_{xx}(t,X_t) + bX_t u_x(t,X_t)]\,dt + \sigma X_t u_x(t,X_t)\,dW_t$$

$$= \sigma X_t u_x(t,X_t)\,dW_t.$$

Hence, $u(T,X_T) - u(0,x) = \int_0^t \sigma X_s u_x(s,X_s)\,dW_s$, and therefore

$$u(0,x) = E_x\,g(X_T); \quad \text{likewise,} \;\; u(t,x) = E_x\,g(X_{T-t}). \tag{19}$$

24 Solving PDE via SDE: linear coefficients – 2

(important reading)

Consider a PDE with coefficients $\sigma^2 x^2/2$ and $b\,x$, and a "potential" $(-ru)$,

$$\boxed{u_t + \tfrac{1}{2}\sigma^2 x^2 u_{xx} + bxu_x - ru = 0, \quad u(T,x) = g(x).} \tag{20}$$

Consider a solution to SDE

$$\boxed{dX_t = \sigma X_t dW_t + bX_t dt, \quad X_0 = x.} \tag{21}$$

Suppose $u(t,x)$ is a solution to (23). Apply the Itô formula to $\exp(-rt)u(t,X_t)$:

$$d\exp(-rt)u(t,X_t) = \exp(-rt)[u_t(t,X_t) + \frac{1}{2}\sigma^2 X_t^2 u_{xx}(t,X_t) + bX_t u_x(t,X_t) - ru(t,X_t]\,dt$$

$$+\sigma X_t u_x(t,X_t)\,dW_t = \exp(-rt)\sigma X_t u_x(t,X_t)\,dW_t.$$

Hence, $\exp(-rT)u(T,X_T) - u(0,x) = \int_0^t \exp(-rs)\sigma X_s u_x(s,X_s)\,dW_s$, and therefore

$$u(0,x) = \exp(-rT)\,E_x\,g(X_T); \quad \text{likewise,} \;\; u(t,x) = \exp(-r(T-t))E_x\,g(X_{T-t}). \tag{22}$$

Remark. One can now check directly corresponding equations using the distributions of X_t in each case: in (13) and (15) it is a normal distribution, and in (19) and (22) a log-normal one (i.e. the logarithm of the r.v. has a normal distribution).

Remark. Sections 21–23 are marked as advanced reading only to reduce compulsory material. There is no harm to read them, and it may help understand the section 24. Similarly the next section 25 is no more difficult and may help understand the section 24.

25 Solving PDE via SDE: linear coefficients – 3

(advanced reading)

Consider a PDE with coefficients $\sigma^2 x^2/2$ and $b\,x$, and a "potential" $(-ru)$, *and a right hand side* (-1),

$$u_t + \frac{1}{2}\sigma^2 x^2 u_{xx} + bxu_x - ru = -1, \quad u(T,x) = g(x). \tag{23}$$

Consider a solution to SDE

$$dX_t = \sigma X_t dW_t + bX_t dt, \qquad X_0 = x.$$

Suppose $u(t,x)$ is a solution to (23). Apply the Itô formula to $\exp(-rt)u(t,X_t)$:

$$d\exp(-rt)u(t,X_t) = \exp(-rt)[u_t(t,X_t) + \frac{1}{2}\sigma^2 X_t^2 u_{xx}(t,X_t) + bX_t u_x(t,X_t) - ru(t,X_t]\,dt$$

$$+\sigma X_t u_x(t,X_t)\,dW_t = \exp(-rt)[-1]\,dt + \exp(-rt)\sigma X_t u_x(t,X_t)\,dW_t.$$

Hence, in the integral form,

$$\exp(-rT)u(T,X_T) - u(0,x) = -\int_0^T \exp(-rs)\,ds + \int_0^T \exp(-rs)\sigma X_s u_x(s,X_s)\,dW_s,$$

and, therefore, after taking expectation,

$$\exp(-rT)\,E_x\,g(X_T) - u(0,x) = +\frac{1}{r}\,\exp(-rs)\Big|_0^T,$$

or, equivalently,

$$\boxed{u(0,x) = \exp(-rT)\,E_x\,g(X_T) + \tfrac{1}{r}\left[1 - \exp(-rT)\right].}$$

26 Black-Scholes' equation

(advanced reading)

We assume that the stock price process is given by the formula

$$S_t = S_0 \exp(\mu t + \sigma W_t), \quad t \geq 0.$$

Assume that the (European) call option price is a function of time and S_t (and not its values in the past):

$$C_t = Y(t, S_t)$$

where $Y(t, S)$ is a function with two continuous derivatives in S and one in t.

Theorem 8 (Black-Scholes) *The function Y satisfies an equation*

$$(BS) \qquad \frac{\partial Y}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 Y}{\partial S^2} + rS\frac{\partial Y}{\partial S} - rY = 0,$$

with a terminal condition

$$Y(T, S) = (S - K)_+.$$

This equation possesses a (unique) explicit solution $(0 \le t \le T)$

$$Y(t, S) = S\Phi(d_+) - Ke^{-r(T-t)}\Phi(d_-), \tag{24}$$

where

$$d_\pm = d_\pm(t) = \frac{\log(S/K) + (r \pm \sigma^2/2)(T - t)}{\sigma\sqrt{T - t}}.$$

Notice that the equation does not depend on μ. So, the solution does not depend on it either. This effect is important in practice: the volatility σ may be determined much more easily and precisely than the trend μ. So, even if two traders evaluate the trend differently, they may agree with the value of a call since they use the same estimated volatility.

Proof of Theorem: **(further advanced reading)**

(a) Straightforward calculations show that the function given by the formula (24) does satisfy (BS) and the terminal condition (the latter in the sense that $\lim_{t\to T} Y(t, S) = (S - K)_+$), which is recommended to check. Though, in the end of the proof we will get to the formula independently.

(b) Consider a portfolio

$$X_t = \phi_t S_t + \psi_t B_t$$

with some *policy* (ϕ_t, ψ_t), $0 \le t \le T$. Similarly to CRR model, we are looking for such a policy that

$$Y(t, S_t) = X_t, \quad \forall \ 0 \le t \le T.$$

We consider only self-financing policies, that is, by definition,

$$dX_t = \phi_t dS_t + \psi_t dB_t,$$

compare to the notion of self-financing policies in discrete time. The sense is that only changes of prices S_t and B_t make changes in the price of the portfolio.

(c) Since $Y(t, S_t) \equiv X_t$, the stochastic differentials should also be equal:

$$dY(t, S_t) \equiv dX_t.$$

Remind that

$$dS_t = \sigma S_t dW_t + \mu' S_t\, dt, \quad \mu' = \mu + \sigma^2/2, \qquad dB_t = rB_t\, dt.$$

By Itô's formula,

$$dX_t = \psi_t r B_t\, dt + \phi_t \left(\sigma S_t dW_t + \mu' S_t\, dt\right)$$

and

$$dY(t, S_t) = \left(\frac{\partial Y}{\partial t} + \mu' S \frac{\partial Y}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 Y}{\partial S^2}\right) dt + \frac{\partial Y}{\partial S}\, \sigma\, S_t\, dW_t.$$

(d) So we get two equations:

$$\phi_t \sigma S_t = \sigma S_t \frac{\partial Y}{\partial S},$$

and

$$\phi_t \mu' S_t + \psi_t r B_t = \frac{\partial Y}{\partial t} + \mu' S \frac{\partial Y}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 Y}{\partial S^2}.$$

(e) From the first one,

$\phi_t = \dfrac{\partial Y}{\partial S}$ (this value is called **Delta** of a call option, important for hedging).

(f) We substitute this to the second equation and get

$$r\psi_t B_t = \frac{\partial Y}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 Y}{\partial S^2}.$$

(g) Since

$$\psi_t B_t = X_t - \phi_t S_t = Y(t, S_t) - \phi_t S_t = Y(t, S_t) - S_t \frac{\partial Y}{\partial S},$$

we finally obtain

$$rY = \frac{\partial Y}{\partial t} + rS\frac{\partial Y}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 Y}{\partial S^2}.$$

(h) From results of lecture 16 (how to solve PDEs via SDEs), it follows that solution of equation (BS) must have a form

$$Y(0, S) = \exp(-rT)E(S_0 \exp((r - \sigma^2/2)T + \sigma\sqrt{T}\eta) - K)_+, \quad \eta \sim \mathcal{N}(0, 1) \tag{25}$$

(compare with last formula from lecture 10).

(i) Calculus leading from (25) to (24) was fulfilled in **Exercises 3 and 4**, and will be repeated in the final handout.

Exercise 10 *Show that the portfolio with* $\phi_t = \partial Y(t, S_t)/\partial S$ *and* $\psi_t = (Y(t, S_t) - S_t \partial Y(t, S_t)/\partial S)/B_t$ *is indeed self-financing. Hint: show that* $X_t = Y(t, S_t)$ *and find* $dY(t, S_t)$ *taking into account the Black-Scholes equation.*

27 American options: definition

Definition 5 *An American option is a contract similar to European option with a corresponding payoff (call* $(S_t - K)_+$*, put* $(K - S_t)_+$*, general option* $g_t(S_t) \geq 0$*), but which may be exercised any time to expiry.*

28 American options in the CRR model

(further advanced reading)

- Let $\delta > 0$ be time step, $T = n\delta$ expiry, $V_T = g_T(S_T)$ payoff of the option at expiry. Then at any node x at expiry,

$$V_{n\delta}(x) = g_{n\delta}(x). \tag{26}$$

Consider time before expiry, $(n-1)\delta$. Either the option is exercised with payoff $g_{(n-1)\delta}(x)$, or it is not; in the latter case its value at the node x is $\exp(-r\delta)\tilde{V}_{(n-1)\delta}(x)$; here $\tilde{V}_{(n-1)\delta}(x)$ denotes $(\tilde{p}_x g(x^+) + (1-\tilde{p}_x)g(x^-))$ where $\tilde{p}_x$ is a corresponding implied probability at this node. So it is reasonable to exercise the option at time $(n-1)\delta$ in the node x iff $g(x) > \exp(-r\delta)\tilde{V}_{(n-1)\delta}(x)$. Hence, the value of the option in the node x at time $(n-1)\delta$ must be equal to

$$V_{(n-1)\delta}(x) = \max\left(g(x), \exp(-r\delta)\tilde{V}_{(n-1)\delta}(x))\right). \tag{27}$$

- This can be repeated by induction backward in time from T to zero: at any time $k\delta$ between 0 and $(n-1)\delta$,

$$V_{k\delta}(x) = \max\left(g(x), \exp(-r\delta)\tilde{V}_{(k+1)\delta}(x))\right), \tag{28}$$

provided $\tilde{V}_{(k+1)\delta}(x)$ has already been calculated.

- There is no explicit formula for the price V_0, like for European options, except for call and some very special options. However, the general result may be presented in the form

$$V_0 = \max_{\tau} \tilde{E}\left(e^{-r\tau} g_\tau(S_\tau)\right), \quad V_t = \max_{\tau \geq t} \tilde{E}\left(e^{-r(\tau-t)} g_\tau(S_\tau)|\mathcal{F}_t\right), \tag{29}$$

where τ is any **stopping time**, i.e. a random time which satisfies the property: for any t, the set $(\tau > t) \in \mathcal{F}_t$, that is, one must decide whether to stop at t or proceed further only using information to time t. The max in the first formula above is taken over all possible stopping times taking values $0, \delta, 2\delta, \ldots, n\delta = T$.

- We will check it only for $n = 1$. Indeed, if $t = T = \delta$, then by definition the value V_t from (29) is a correct price of the option. We only need to check that (29) gives a correct price for $t = 0$. At $t = 0$ we must decide whether to stop or wait till expiry, or, in other words, we choose between $g(x)$ and $\exp(-r\delta)(\tilde{p}_x g(x^+) + (1-\tilde{p}_x)g(x^-))$, i.e. take the maximum of these two values. This coincides with equation (27), so (29) indeed provides a fair option price.

- Once (suppose) we have found the price, the next question is when to stop, that is, how to choose τ in order to optimize the profit, actually to achieve the maximum? Notice that the correct price may not be achieved if the owner does not take care, that is, if he does not choose the optimal stopping rule. The optimal rule which follows from the algorithm above is that one must stop when for the first time $g_t(S_t) \geq \exp(-r\delta)\tilde{V}_t(S_t)$, or

$$\tau = \min(0 \leq t \leq n\delta: \; g_t(S_t) \geq e^{-r\delta}\tilde{V}_t(S_t)).$$

- An amazing fact is that for American call the price is the same as for European option (that is, given by the Black and Scholes formula) with the same parameters. Correspondingly, the optimal stopping rule for American call is wait until expiry, that is, not exercise the right to sell the option earlier. (Notice that there are modifications of the standard American call, e.g., with "discounted payoff"; the latter does not apply to all of them, in general).

29 American options in the continuous time model

(further advanced reading)

- It may be shown that the price $V = V(t, S)$ solves the following PDE (called PDE with a free boundary condition):

$$\begin{cases} \left[V_t + \frac{1}{2}\sigma^2 S^2 V_{SS} + rSV_S - rV\right] = 0, & \text{for } (t,x): \; V(t,x) > g_t(x), \\ \left[V_t + \frac{1}{2}\sigma^2 S^2 V_{SS} + rSV_S - rV\right] \leq 0, & \text{for } (t,x): \; V(t,x) \leq g_t(x). \end{cases} \tag{30}$$

  In other words, there are two regions in $[0, T] \times R$: one in which one should continue waiting ($V(t, x) > g_t(x)$), and the other where one should immediately stop ($V(t, x) \leq g_t(x)$). At expiry, $V(T, x) = g_T(x)$.

- Correspondingly, the optimal stopping rule is wait until $V(t, S_t) \leq g_t(S_t)$ for the first time, then stop.

- How intuitively derive the PDE: in the region where one should continue waiting, the no-arbitrage approach with replicating self-financing portfolio gives the same equations as for European options,

$$\left[Y_t + \frac{1}{2}\sigma^2 S^2 Y_{SS} + rSY_S - rY\right] = 0.$$

- It may be shown (using the stochastic calculus) that under certain assumptions (29) provides the solution to equation (30).

- How to solve equation (30)? Numerically: choose a small δ and adjust the binomial model (i.e. prices and probabilities of up and down jumps); then use the algorithm described in (26) - (28).
- Again, for American call the price coincides with that of European call (with all parameters equal), and the optimal stopping rule is wait until expiry.

30 Some further advanced reading

- A. N. Shiryaev, (recommended textbook on stochastic finance) Chapter VI.
- G. Peskir, A. N. Shiryaev, the monograph 2006 *(an advanced book on stochastic finance).*
- I. I. Gikhman, A. V. Skorokhod, Stochastic differential equations, the monograph *(in particular, about weak and strong solutions: one of Skorokhod's results claims that any SDE with only continuous bounded coefficients has a (possibly non-unique) weak solution.)*
- N. V. Krylov, research paper 1969 on Ito's SDEs *(one of results states that for uniformly non-degenerate diffusion coefficient and both coefficients bounded (without any continuity assumption), the equation (3) always possess a weak solution; uniqueness of distribution is known in major cases (e.g., if σ is continuous).*
- S. V. Anulova, A. Yu. Veretennikov, et al. Stochastic calculus, book: *(in particular, about most of results on weak and strong solutions known to 1980).*
- N. Abourashchi, A. Veretennikov, forthcoming preprint on exponential mixing rates for 2D stochastic systems, the section 3 on weak solutions and Girsanov's transformations *(some new version of weak solution for special SDE's arising in stochastic mechanics).*