1. Introduction

We have solved the Black and Scholes equation in Lecture 3 by transforming it into the heat equation, and using the classical solution for the initial value problem of the latter. We have a posteriori verified the correctness of that solution in Lecture 3, and encountered two more ways of finding solutions to the heat equation in the exercises of Lecture 3: the similarity method and Fourier transform. There is a third way of solving the heat equation, by using Brownian motion \( (W_t)_{t \geq 0} \) and exploiting Ito’s lemma for functions \( u(W_t, t) \) of Brownian motion:

\[
(1) \quad du(W_t, t) = \left( \frac{\partial u}{\partial t} + \frac{1}{2} \frac{\partial^2 u}{\partial w^2} \right) dt + \frac{\partial u}{\partial w} dW_t.
\]

Here the crucial observation is that the term which is proportional to \( dt \), and which is commonly called the drift term, vanishes precisely when \( u \) is a solution of the, in this case backward, heat equation. We develop this idea in section 2 below, where we will see that it leads to an expression of the solution \( u(x, t) \) of the final value problem, for the backward heat equation with final value \( f(w) \) at \( T \) as the expectation of \( f(x + W_T - W_t) \). What is more, this approach is not limited to the heat equation, but can also be directly applied to the Black and Scholes equation itself, as well as to other PDEs of so-called parabolic type, roughly characterised by having one time derivative, and up to two space derivatives with positive efficient of the second derivative. This will be detailed first for the B&S PDE for derivatives on dividend-paying assets, and finally for general parabolic PDEs.
2. Solving the heat equation using Brownian motion

Let \( u = u(w, t) \) \((w \in \mathbb{R}, t \leq T)\) be a solution of the following final value problem for the so-called backward heat equation:
\[
\begin{cases}
\frac{\partial u}{\partial t} + \frac{1}{2} \frac{\partial^2 u}{\partial w^2} = 0 \text{ for } t < T; \\
u(w, T) = f(w).
\end{cases}
\]
(2)

As observed above, Itô’s lemma then immediately implies that
\[
du(W_t, t) = \frac{\partial u}{\partial w} (W_t, t) dW_t.
\]
(3)

Integrating this differential from \( t \) to \( T \), we then see that
\[
u(W_T, T) - u(W_t, t) = \int_t^T \frac{\partial u}{\partial w} (W_s, s) dW_s,
\]
or, using the final value condition at \( T \),
\[
u(W_t, t) = f(W_T) - \int_t^T \frac{\partial u}{\partial w} (W_s, s) dW_s.
\]
(4)

The idea now is to take expectations, conditional upon fixing \( W_t = w \), leading to
\[
\begin{align*}
u(w, t) &= \mathbb{E}(f(W_T)|W_t = w) + \mathbb{E} \left( \int_t^T \frac{\partial u}{\partial w} (W_s, s) dW_s|W_t = w \right).
\end{align*}
\]
(5)

We examine the second term on the right. Since taking expectations is a linear operator, and an integral is in some sense a limit of finite sums, we can interchange integration and expectation\(^1\):
\[
\mathbb{E} \left( \int_t^T \frac{\partial u}{\partial w} (W_s, s) dW_s|W_t = w \right) = \int_t^T \mathbb{E} \left( \frac{\partial u}{\partial w} (W_s, s) dW_s|W_t = w \right)
\]
Next, using that \( dW_s \) is independent of \( W_s \) and therefore independent of any function of \( W_s \), we see that
\[
\mathbb{E} \left( \frac{\partial u}{\partial w} (W_s, s) dW_s|W_t = w \right) = \mathbb{E} \left( \frac{\partial u}{\partial w} (W_s, s)|W_t = w \right) \cdot \mathbb{E}(dW_s|W_t = w).
\]
Finally, again by independence of Brownian increments with respect to \( W_t \) up to present,
\[
\mathbb{E}(dW_s|W_t = w) = \mathbb{E}(dW_s) = 0.
\]
The conclusion is that the entire, complicated looking second term of (5) is in fact 0, and that
\[
u(w, t) = \mathbb{E}(f(W_T)|W_t = w),
\]
(6)

\(^1\) Mathematically, this needs some discussion: what we use here is some form of Fubini’s theorem for stochastic integrals: note that the integral with respect to \( W_t \) is a stochastic one.
fulfilling our promise of writing the solution as an expectation. To work out what this expectation is, write $W_T = W_t + (W_T - W_t)$. Then

$$E(f(W_T)|W_t = w) = E(f(W_t + (W_T - W_t)|W_t = w)$$

$$= E(f(w + W_T - W_t)|W_t = w)$$

$$= E(f(w + W_T - W_t)),$$

again by independence of $W_T - W_t$ and $W_t$, which makes the conditioning on $W_t = w$ irrelevant. We therefore conclude:

**Theorem 2.1.** If $u(w, t)$ is a solution to (2), then

$$u(w, t) = E(f(w + W_T - W_t)).$$

Working out the expectation is, we find the familiar formulas from Lecture 3: writing

$$W_T - W_t = \sqrt{T-t}Z = \sqrt{\tau}Z,$$

where $Z \sim N(0, 1)$ and $\tau = T-t$, we have that

$$u(w, t) = E(f(w + \sqrt{\tau}Z))$$

$$= \int_{-\infty}^{\infty} f(w + \sqrt{\tau}z)e^{-z^2/2} \frac{dz}{\sqrt{2\pi}}$$

$$= y := w + z \int_{-\infty}^{\infty} f(y)e^{-(w-y)^2/2} \frac{dy}{\sqrt{2\pi}};$$

compare theorem 2.1 from Lecture 2 (with $\sigma = 1$).

Note (ro prevent future confusion with signs) that we could also have written above that incidentally can also be written as

$$u(w, \tau) = \int_{-\infty}^{\infty} f(w - \sqrt{\tau}z)e^{-z^2/2} \frac{dz}{\sqrt{2\pi}},$$

after a change of variables replacing $z$ by $-z$ and noting that $e^{-z^2/2}$ is an even function.

3. **Solving the Black and Scholes equation using GBM**

We next show that it in fact isn’t necessary to transform the Black and Scholes equation to the heat equation at all, and that we can directly solve it by using a similar argument as the one above for the heat equation. Consider the B&S PDE for a derivative on a dividend-paying asset (to do something slightly more general than in Lecture 3):

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2S^2\frac{\partial^2 V}{\partial S^2} + (r - q)S\frac{\partial V}{\partial S} = rV,$$

with final condition

$$V(S, T) = F(S),$$
$F(S)$ being the derivative’s pay-off function. Introduce a new stochastic process $(\hat{S}_t)_{t \geq 0}$ which by definition satisfies the following SDE:

$\begin{align}
(11) \quad d\hat{S}_t &= (r - q)\hat{S}_t dt + \sigma \hat{S}_t dW_t;
\end{align}$

The ”hat” serves to distinguish this process from the real price-process $(S_t)_{t \geq 0}$ we started off with, and from which it is different. The choice of the drift-term $(r - q)(r - q)\hat{S}_t$ is informed by the coefficient of $\partial V / \partial S$ in (9), and that of the Brownian term $\sigma \hat{S}_t$ by the coefficient of $\partial^2 / \partial S^2$ in (9). In fact, one easily checks that if we apply Ito’s lemma to a function $V(\hat{S}_t, t)$, then the coefficient of $dt$ is exactly the differential expression on the left hand side of (9):

$\begin{align}
(12) \quad dV(\hat{S}_t, t) &= \left( \frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 \hat{S}_t \frac{\partial^2 V}{\partial S^2} + (r - q)\hat{S}_t \frac{\partial V}{\partial S} \right) dt + \sigma \hat{S}_t \frac{\partial V}{\partial S} dW_t,
\end{align}$

the derivatives of $V$, as always, evaluated in $(\hat{S}_t, t)$.

What about the right hand side, $rV$, of (9)? This can be included, after moving it to the left, by looking at the differential of $e^{-rt}V(\hat{S}_t, t)$ rather than of $V(\hat{S}_t, t)$: using the product rule for differentials, we easily find:

$\begin{align}
d\left( e^{-rt}V(\hat{S}_t, t) \right) &= e^{-rt}dV(\hat{S}_t, t) - re^{-rt}V(\hat{S}_t, t)dt \\
&= \left( \frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 \hat{S}_t \frac{\partial^2 V}{\partial S^2} + (r - q)\hat{S}_t \frac{\partial V}{\partial S} - rV \right) e^{-rt}dt \\
&\quad + \sigma \hat{S}_t e^{-rt} \frac{\partial V}{\partial S} dW_t.
\end{align}$

Again the conclusion is that If $V(S, t)$ satisfies the B&S equation (9), then $e^{-rt}V(\hat{S}_t, t)$ has a differential which is a pure multiple of $dW_t$:

$\begin{align}
d\left( e^{-rt}V(\hat{S}_t, t) \right) &= \sigma \hat{S}_t e^{-rt} \frac{\partial V}{\partial S} dW_t.
\end{align}$

At this point we could, for those familiar with the concept of a martingale, short-circuit the argument by saying that it follows that $e^{-rt}V(\hat{S}_t, t)$ is a martingale$^2$, and that as a consequence (in fact, the very definition of a martingale),

$\begin{align}
(13) \quad e^{-rt}V(\hat{S}_t, t) &= e^{-rT} \mathbb{E} \left( V(\hat{S}_T, T) | S_t \right) = e^{-rT} \mathbb{E} \left( F(\hat{S}_T) | S_t \right),
\end{align}$

or, conditioning upon $S_t$ being equal to $S$, $e^{-rt}V(S, t) = e^{-rT} \mathbb{E} \left( F(\hat{S}_T) | S_t = S \right)$
or

$\begin{align}
(14) \quad V(S, t) &= e^{-r(T-t)} \mathbb{E} \left( F(\hat{S}_T) | S_t = S \right),
\end{align}$

$^2$we’re riding roughshod over some delicate mathematical details here: technically speaking, for general functions $V(S, t)$ it in general is only a so-called local martingale. If however $V(S, t)$ grows at most linearly in $S$, as for the solutions of the B&S PDE we are interested in, things are OK and it is a genuine martingale.
which is the analog of (7). If we do not want to martingales, we can proceed as for the heat equation above: integrate from $t$ to $T$ to find

$$e^{-rT} V(\hat{S}_T, T) - e^{-rt} V(\hat{S}_t, t) = \int_t^T \sigma \hat{S}_u e^{-ru} \frac{\partial V}{\partial S} dW_u.$$ 

As before, take expectations conditional upon $\hat{S}_t = S$, and observe that since

$$\hat{S}_u = S_0 e^{(r-q-\frac{1}{2}\sigma^2)u} + \sigma W_u,$$

(by solving the SDE) is independent of $dW_u$,

$$\mathbb{E} \left( \sigma \hat{S}_u e^{-ru} \frac{\partial V}{\partial S} dW_u | S_t = S \right) = \mathbb{E} \left( \sigma \hat{S}_u e^{-ru} \frac{\partial V}{\partial S} | \hat{S}_t = S \right) \mathbb{E} (dW_u | \hat{S}_t = S) = \mathbb{E} (dW_u) = 0.$$

It follows that the expectation of the stochastic integral on the right hand side of (15) is 0. Hence we proved:

**Theorem 3.1.** The solution of the B&S equation is given by the discounted expectation

$$V(S, t) = e^{-r(T-t)} \mathbb{E} \left( F(\hat{S}_T) | S_t = S \right),$$

where $\hat{S}_t$ solves the SDE $d\hat{S}_t = (r-q)\hat{S}_t dt + \sigma \hat{S}_t dW_t$.

To get from this to an explicit formula, we note that if we solve the SDE from $t$ to $T$ starting of with $S$ at $t$, then

$$\hat{S}_T = S e^{(r-q-\frac{1}{2}\sigma^2)(T-t)+\sigma(W_T-W_t)}.$$

Writing $\tau = T-t$ and $W_T - W_t = \sqrt{\tau}Z$ with $Z \sim N(0, 1)$, we find

$$V(S, t) = e^{-r\tau} \mathbb{E} \left( F \left( S e^{(r-q-\frac{1}{2}\sigma^2)\tau + \sigma \sqrt{\tau} Z} \right) \right) = e^{-r\tau} \int_{-\infty}^{\infty} F \left( S e^{(r-q-\frac{1}{2}\sigma^2)\tau + \sigma \sqrt{\tau} z} \right) e^{-z^2/2} \frac{dz}{\sqrt{2\pi}}.$$

You should now also be able to do exercises 5.6.4 of chapter 5, which generalizes the method to arbitrary parabolic PDEs, and which is restated below for convenience.

**Exercise** (Feynman-Kac Formula) Consider a parabolic PDE for an unknown function $v(x, t)$ of the type

$$\partial_t v + a(x) \partial^2_x v + b(x) \partial_x v = c(x) v, \quad t < T,$$

with final condition $v(y, T) = F(y)$. The coefficients $\sigma(x)$ and $a(x)$ here are functions of $x$ (we might also allow them to depend on time $t$). The
idea is to associate to this PDE a stochastic process $X_t$ constructed from these coefficients: $X_t$ will be a solution of the SDE

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

where the function $b(x)$ is the coefficient of $\partial_x v$ in the PDE, and where the function $\sigma(x)$ is related to the coefficient $a(x)$ of the second order derivative $\partial_x^2 v$ through

$$a(x) = \frac{1}{2}\sigma(x)^2,$$

or $\sigma(c) = \sqrt{2a(x)}$; note that we clearly require $a(x)$ to be non-negative: this is part of the parabolicity of the PDE (which, technically, is called *backward parabolic*).

The Feynman-Kac formula then states that

$$v(x,t) = \mathbb{E}\left(e^{-\int_t^T c(X_u)du} F(X_T)|X_t = x}\right).$$

That is, to find $v(x,t)$, we have to solve the SDE $dX_u = a(X_u)du + \sigma(X_u)dW_u$ from time $t$ till $T$, with initial condition $X_t = x$ at $t$, and compute the expectation of $\exp(-\int_t^T c(X_u)du) F(X_T)$. Although at first sight perhaps complicated and daunting looking, this is a remarkably efficient algorithm to solve PDEs like the B&S one. This exercise sketches a proof of the Feynman-Kac theorem.

(a) Put

$$Y_t := e^{-\int_0^t c(X_u)du} v(X_t,t).$$

Show that

$$dY_t = e^{-\int_0^t c(X_u)du} \left(\partial_t v + a(X_t)\partial_x^2 v + b(X_t)\partial_x v - c(X_t)v\right) dt + e^{-\int_0^t c(X_u)du} \partial_x v dW_t,$$

all derivatives of $v$ are evaluated in $(X_t,t)$.

(b) Conclude that if $v$ solves the PDE, then $Y_t$ has zero drift and therefore is a *martingale*.

(c) From the definition of a martingale (conditional expectation of $Y_T$ at time $t = Y_t$), find the Feynman-Kac formula.