**Derivative Securities – Fall 2012– Section 6**
*Notes by Robert V. Kohn, extended and improved by Steve Allen.*
*Courant Institute of Mathematical Sciences.*

**Stochastic differential equations and the Black-Scholes PDE**. We derived the Black-Scholes formula by using no-arbitrage-based (risk-neutral) valuation in a discrete-time, binomial tree setting, then passing to a continuum limit. We started that way because binomial trees are very explicit and transparent. However the power of the discrete framework as a conceptual tool is rather limited. Therefore we now begin developing the more powerful continuous-time framework, via the Ito calculus and the Black-Scholes differential equation. This material is discussed in many places. Hull addresses it in Chapters 13 and 14, but his coverage is rather limited. Baxter & Rennie emphasizes risk-neutral expectation, avoiding almost completely the discussion of PDE's. The "student guide" by Wilmott, Howison, & Dewynne takes almost the opposite approach: it emphasizes PDE's, avoiding almost completely the discussion of risk-neutral expectation. Neftci's book provides a good introduction to Brownian motion, the Ito calculus, stochastic differential equations, and their relation to option pricing, at a level that should be accessible to students in this class. (Students taking Stochastic Calculus will learn this material and much more over the course of the semester.) A brief introduction to stochastic calculus (similar in spirit to what's here, but going somewhat deeper, with exercises and many more examples) can be found at the top of my Spring 2011 PDE for Finance course notes (on my web page).

*********************

**Why work in continuous time?**. Our discrete-time approach has the advantage of being very clear and explicit. However there is a different approach, based on Taylor expansion, the Ito calculus, and the "Black-Scholes differential equation." It has its own advantages:

- Passing to the continuous time limit is clearly legitimate for describing the stock price process. But is it legitimate for describing the value of the option, as determined by arbitrage? This is less clear, since a continuous-time hedging strategy is unattainable in practice. In what sense can we "approximately replicate" the option by trading at discrete times? The Black-Scholes differential equation will help us answer these questions.

- We assumed the stock price process (or the forward price process) was lognormal, then approximated it using a binomial tree, then applied arbitrage pricing to this tree. But what if we had approximated the stock price process by a trinomial tree? Then arbitrage pricing wouldn't have worked, since the one-period trinomial market isn't complete. It should make you nervous that we relied so heavily on the use of a particular approximation.

- The differential equation approach gives fresh insight and computational flexibility. Imagine trying to understand the implications of compound interest without using the differential equation $df/dt = rf$! (Especially: imagine how stuck you'd be if $r$ depended on $f$.)

- Differential-equation-based methods lead to efficient computational schemes (and even explicit solution formulas in some cases) not only for European options, but also for more complicated instruments such as barrier options.

**Brownian motion**. Recall our discussion of the lognormal hypothesis for stock price dynamics. It says that $\log[s(t_2)/s(t_1)]$ is a Gaussian random variable with mean $\mu(t_2 - t_1)$ and variance $\sigma^2(t_2 - t_1)$, and disjoint intervals give rise to independent random variables.

A time-dependent random variable is called a *stochastic process*. The lognormal hypothesis is related to Brownian motion $w(t)$, also known as the Wiener process, which has continuous sample paths and satisfies:

(a) $w(t_2) - w(t_1)$ is a Gaussian random variable with mean value 0 and variance $t_2 - t_1$;

(b) distinct intervals give rise to independent random variables;

(c) $w(0) = 0$.

It can be proved that these properties determine a unique stochastic process, i.e. they uniquely determine the probability distribution of any expression involving $w(t_1)$, $w(t_2)$, ..., $w(t_N)$.

An interesting feature of Brownian motion is this: for almost any realization, the function $t \mapsto w(t)$ is continuous but not differentiable. The following simple argument demonstrates that if we accept the continuity of $w(t)$ then we must expect something like non-differentiability. Given an interval $(a, b)$, divide it into subintervals by $a = t_1 < t_2 \ldots < t_N = b$. Clearly

$$\sum_{i=1}^{N-1} |w(t_{i+1}) - w(t_i)|^2 \leq \max_i |w(t_{i+1}) - w(t_i)| \cdot \sum_{i=1}^{N-1} |w(t_{i+1}) - w(t_i)|.$$

As $N \to \infty$, the left hand side has expected value $b - a$ (independent of $N$). The first term on the right tends to zero (almost surely) by continuity. So the second term on the right must tend to infinity (almost surely). Thus the sample paths of $w$ have unbounded total variation on any interval.

The process $w(t)$ can be viewed as a limit of suitably scaled random walks (we showed this in Section 4). Another important fact: writing $w(t_2) - w(t_1) = \Delta w$ and $t_2 - t_1 = \Delta t$,

$$E\left[|\Delta w|^j\right] = C_j |\Delta t|^{j/2}, \quad j = 1, 2, 3, \ldots.$$

The proof uses nothing more than calculus – and the fact that $\Delta w$ is Gaussian with mean 0 and variance $\Delta t$.

Our lognormal hypothesis can be reformulated as the statement that

$$s(t) = s(0) \exp[\mu t + \sigma w(t)].$$

**Stochastic differential equations.** Let's first review ordinary differential equations. Consider the ODE $dy/dt = f(y,t)$ with initial condition $y(0) = y_0$. It is a convenient mnemonic to write the equation in the form

$$dy = f(y,t)dt.$$

This reminds us that the solution is well approximated by its (explicit) finite difference approximation $y([j+1]\delta t) - y(j\delta t) = f(y(j\delta t), j\delta t)\delta t$, which we sometimes write more schematically as

$$\Delta y = f(y,t)\Delta t.$$

An extremely useful aspect of ODE's is the ability to use chain rule. From the ODE for $y(t)$ we can easily deduce a new ODE satisfied by any function of $y(t)$. For example, $z(t) = e^{y(t)}$ satisfies $dz/dt = e^y dy/dt = zf(\log z, t)$. In general $z = A(y(t))$ satisfies $dz/dt = A'(y)dy/dt$. The mnemonic for this is

$$dA(y) = \frac{dA}{dy}dy = \frac{dA}{dy}f(y,t)dt.$$

It reminds us of the proof, which boils down to the fact that (by Taylor expansion)

$$\Delta A = A'(y)\Delta y + \text{error of order } |\Delta y|^2.$$

In the limit as the timestep tends to 0 we can ignore the error term, because $|\Delta y|^2 \leq C|\Delta t|^2$ and the sum of $1/\Delta t$ such terms is small, of order $|\Delta t|$.

OK, now stochastic differential equations. We consider only the simplest class of stochastic differential equations, namely

$$dy = g(y,t)dw + f(y,t)dt, \quad y(0) = y_0,$$

where $w(t)$ is Brownian motion. The solution is a stochastic process, the limit of the processes obtained by the (explicit) finite difference scheme

$$y([j+1]\delta t) - y(j\delta t) = g(y(j\delta t), t)\left(w([j+1]\delta t) - w(j\delta t)\right) + f(y(j\delta t), j\delta t)\delta t,$$

which we usually write more schematically as

$$\Delta y = g(y,t)\Delta w + f(y,t)\Delta t.$$

Put differently (this is how the rigorous theory begins): we can understand the stochastic differential equation by rewriting it in integral form:

$$y(t') = y(t) + \int_t^{t'} f(y(\tau), \tau)\, d\tau + \int_t^{t'} g(y(\tau), \tau)dw(\tau)$$

where the first integral is a standard Riemann integral, and the second one is a *stochastic integral*:

$$\int_t^{t'} g(y(\tau), \tau)\, dw(\tau) = \lim_{\Delta\tau \to 0} \sum g(y(\tau_i), \tau_i)[w(\tau_{i+1}) - w(\tau_i)]$$

where $t = \tau_0 < \ldots < \tau_N = t'$.

Note: the existence of the limit defining the stochastic integral isn't obvious at all. It cannot be proved using ordinary calculus – for example, we cannot interpret it as $\int_t^{t'} g\,dw = \int_t^{t'} g\,(dw/d\tau)\,d\tau$, because $dw/dt$ does not exist! Rather, the limit defining $\int g\,dw$ must be studied using probabilistic tools. (For further discussion of this topic at a fairly elementary level see Neftci's book, or the Stochastic Calculus Review at the top of my PDE for Finance notes.)

We will often use the following key property of stochastic integrals: $\int_a^b g(\tau, y(\tau))\,dw(\tau)$ has *mean value zero*. (Fancier way of saying the same thing: the process $t \mapsto \int_a^t g(\tau, y(\tau))\,dw(\tau)$ is a *martingale*.) The explanation is easy. Recall that the stochastic integral is a limit of sums. The mean of a typical term in the sum is

$$E\left[g(t_i, y(t_i))\,(w(t_{i+1}) - w(t_i))\right] = 0$$

since $w(t_{i+1}) - w(t_i)$ is independent of all information available at time $t_i$ (so in particular it is independent of $g(t_i, y(t_i))$). Adding, the sum has mean zero. Passing to the limit, so does the stochastic integral.

It's easy to see that when $\mu$ and $\sigma$ are constant, $y(t) = \mu t + \sigma w(t)$ solves

$$dy = \sigma dw + \mu dt.$$

**Ito's lemma**. We started our discussion of SDE's by reviewing chain rule in the context of ODE's. The analogue of chain rule for SDE's is Ito's lemma. It says that if $dy = gdw + fdt$ then $z = A(y)$ satisfies the stochastic differential equation

$$dz = A'(y)dy + \tfrac{1}{2}A''(y)g^2 dt = A'(y)gdw + \left[A'(y)f + \tfrac{1}{2}A''(y)g^2\right]dt.$$

Here is a heuristic justification: carrying the Taylor expansion of $A(y)$ to second order gives

$$\begin{aligned} \Delta A &= A'(y)\Delta y + \tfrac{1}{2}A''(y)(\Delta y)^2 + \text{error of order } |\Delta y|^3 \\ &= A'(y)(g\Delta w + f\Delta t) + \tfrac{1}{2}A''(y)g^2(\Delta w)^2 + \text{errors of order } |\Delta y|^3 + |\Delta w||\Delta t| + |\Delta t|^2. \end{aligned}$$

One can show that the error terms are negligible in the limit $\Delta t \to 0$. For example, the sum of $1/\Delta t$ terms of order $|\Delta w||\Delta t|$ has expected value of order $\sqrt{\Delta t}$. Thus

$$\Delta A \approx A'(y)(g\Delta w + f\Delta t) + \tfrac{1}{2}A''(y)g^2(\Delta w)^2.$$

*Now comes the subtle part of Ito's Lemma*: the assertion that we we can replace $(\Delta w)^2$ in the preceding expression by $\Delta t$. This is sometimes mistakenly justified by saying "$(\Delta w)^2$ behaves deterministically as $\Delta t \to 0$" – which is certainly not true; in fact $(\Delta w)^2 = u^2 \Delta t$ where $u$ is a Gaussian random variable with mean value 0 and variance 1.

So why can we substitute $\Delta t$ for $(\Delta w)^2$? This can be thought of as an extension of the law of large numbers. When we solve a difference equation (to approximate a differential

equation) we must add the terms corresponding to different time intervals. So we're really interested in *sums* of the form

$$\sum_{j=1}^{N} A''(y(t_j))g^2(t_j)(\Delta w)_j^2$$

with $(\Delta w)_j = w(t_{j+1}) - w(t_j)$ and $N = T/\Delta t$. If $A''$ and $g^2$ were constant, then since the $(\Delta w)_j$ are independent, $\sum_{j=1}^{N}(\Delta w)_j^2 = \sum_{j=1}^{N} u_j^2 \Delta t$ would have mean value $N\Delta t = T$ and variance of order $N(\Delta t)^2 = T\Delta t$. Thus the sum would have standard deviation $\sqrt{T\Delta t}$, i.e. it is asymptotically deterministic. The rigorous argument is different, of course, since in truth $A''(y)g^2$ is not constant; but the essential idea is similar.

The version of Ito's Lemma stated and justified above is not the most general one – though it has all the main ideas. Similar logic applies, for example, if $A$ is a function of both $y$ and $t$. Then $z = A(y, t)$ satisfies

$$dz = \frac{\partial A}{\partial y}dy + \frac{\partial A}{\partial t}dt + \frac{1}{2}\frac{\partial^2 A}{\partial y^2}g^2 dt = \frac{\partial A}{\partial y}g\,dw + \left[\frac{\partial A}{\partial y}f + \frac{\partial A}{\partial t} + \frac{1}{2}\frac{\partial^2 A}{\partial y^2}g^2\right]dt$$

or, with a change of notation for easier reading:

$$dz = A_y\,dy + A_t\,dt + \tfrac{1}{2}A_{yy}g^2\,dt = A_y g\,dw + (A_t + A_y f + \tfrac{1}{2}A_{yy}g^2)\,dt$$

We'll also sometimes use this generalization: suppose $y_1$ and $y_2$ solve SDE's using the *same* Brownian motion, say

$$dy_1 = f_1\,dt + g_1\,dw \quad \text{and} \quad dy_2 = f_2\,dt + g_2\,dw,$$

and consider $z(t) = A(t, y_1(t), y_2(t))$. Then

$$dz = A_t\,dt + A_1\,dy_1 + A_2\,dy_2 + \tfrac{1}{2}A_{11}\,dy_1\,dy_1 + A_{12}\,dy_1\,dy_2 + \tfrac{1}{2}A_{22}\,dy_2\,dy_2$$

with the understanding that

$$A_{ij} = \frac{\partial A}{\partial y_i \partial y_j} \quad \text{and} \quad dy_i\,dy_j = g_i g_j\,dt.$$

Let's apply Ito's lemma to find the stochastic differential equation for the stock price process $s(t)$. The lognormal hypothesis says $s = e^y$ where $dy = \sigma\,dw + \mu\,dt$. Therefore $ds = e^y(\sigma\,dw + \mu\,dt) + \frac{1}{2}e^y\sigma^2\,dt$, i.e.

$$ds = (\mu + \tfrac{1}{2}\sigma^2)s\,dt + \sigma s\,dw.$$

(Warning: authors vary on what they call the "drift" of a lognormal process. Some authors (including Hull) take the convention that the lognormal process with drift $\mu$ and volatility $\sigma$ is the solution of $ds = \mu s\,dt + \sigma s\,dw$. For these authors, $\log s(t)/s(0)$ is not $\mu t + \sigma w(t)$ but rather $(\mu - \tfrac{1}{2}\sigma^2)t + \sigma w(t)$.)

5

**The Black-Scholes partial differential equation.** Consider a European option on a non-dividend-paying stock, with payoff $f(s_T)$ at maturity $T$. We have a formula for its value at time $t$, from Section 4:

$$\text{value at time t} = e^{-r(T-t)} \frac{1}{\sigma\sqrt{2\pi(T-t)}} \int_{-\infty}^{\infty} f(s_t e^x) \exp\left[\frac{-(x - [r - \frac{1}{2}\sigma^2](T-t))^2}{2\sigma^2(T-t)}\right] dx.$$

Notice that the value is a function of the present time $t$ and the present stock price $s_t$, i.e. it can be expressed in the form:

$$\text{value at time t} = V(s_t, t).$$

for a suitable function $V(s,t)$ defined for $s > 0$ and $t < T$. It's obvious from the interpretation of $V$ that

$$V(s,T) = f(s).$$

The Black-Scholes differential equation says that

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 V}{\partial s^2} + rs\frac{\partial V}{\partial s} - rV = 0.$$

It offers an alternative procedure for evaluating the value of the option, by solving the PDE "backwards in time" numerically, starting from $t = T$.

Recall that in the setting of binomial trees we had two ways of evaluating the value of an option: one by expressing it as a weighted sum over all paths; the other by working backward through the tree. Evaluating the integral formula is the continuous-time analogue of summing over all paths. Solving the Black-Scholes PDE is the continuous-time analogue of working backward through the tree. Recall also that working backward through the tree was a little more flexible – for example it didn't require that the interest rate be constant. Similarly, the Black-Scholes equation can easily be solved numerically even when the interest rate and volatility are (deterministic) functions of time.

Where does the equation come from? We'll give an explanation today based on a hedging argument. Next week we'll give a (related but somewhat different) argument involving martingales. Each argument is based on Ito's formula. Examining our arguments, you'll be able to see that the Black-Scholes PDE generalizes straightforwardly to stock price models in which the volatility and drift depend on stock price. However for simplicity we'll present the arguments in the constant-volatility, constant-drift setting

$$ds = \sigma s \, dw + (\mu + \tfrac{1}{2}\sigma^2)s \, dt$$

and we'll continue to assume that the interest rate is constant.

**Derivation of the Black-Scholes PDE by considering a hedging strategy.** Remember that when hedging in the discrete-time setting, we rebalance the portfolio so that it contains $\phi$ units of stock and the rest a risk-free bond, then we let the stock price jump to the new value. (I write $\phi$ not $\Delta$ to avoid confusion, because we have been using $\Delta$ for

6

increments.) The analogous procedure in the continuous-time setting is to rebalance at successive at time intervals of length $\delta t$, then pass to the limit $\delta t \to 0$. Suppose that after rebalancing at time $j\delta t$ the portfolio contains $\phi = \phi(s(j\delta t), j\delta t)$ units of stock. Consider the value of the option less the value of the stock during the next time interval:

$$\Pi = V - \phi s.$$

Its increment $d\Pi = \Pi([j+1]\delta t) - \Pi(j\delta t)$ is approximately

$$
\begin{aligned}
dV - \phi ds &= \frac{\partial V}{\partial s} ds + \frac{\partial V}{\partial t} dt + \tfrac{1}{2} \frac{\partial^2 V}{\partial s^2} \sigma^2 s^2 dt - \phi ds \\
&= \left( \frac{\partial V}{\partial s} - \phi \right) \sigma s \, dw + \left( \frac{\partial V}{\partial s} - \phi \right) (\mu + \tfrac{1}{2}\sigma^2) s \, dt + \left( \frac{\partial V}{\partial t} + \tfrac{1}{2} \frac{\partial^2 V}{\partial s^2} \sigma^2 s^2 \right) dt.
\end{aligned}
$$

Note that we do not differentiate $\phi$ because it is being held fixed during this time interval. We know enough to expect that the right choice of $\phi$ is $\phi(s,t) = \partial V/\partial s$. But if we didn't already know, we'd discover it now: this is the choice that eliminates the $dw$ term on the right hand side of the the last equation. Fixing $\phi$ this way, we see that

$$dV - \phi ds = \left( \frac{\partial V}{\partial t} + \tfrac{1}{2} \frac{\partial^2 V}{\partial s^2} \sigma^2 s^2 \right) dt \quad \text{is deterministic.}$$

Now, the principle of no arbitrage says that a portfolio whose return is deterministic must grow at the risk-free rate. In other words, for this choice of $\phi$ we must have

$$dV - \phi ds = r(V - \phi s)dt.$$

Combining these equations gives

$$\left( \frac{\partial V}{\partial t} + \tfrac{1}{2} \frac{\partial^2 V}{\partial s^2} \sigma^2 s^2 \right) = r(V - \phi s)$$

with $\phi = \partial V/\partial s$. This is precisely the Black-Scholes PDE.

That was fast – maybe too fast. Let's review it. Our assertion is that if $V(s,t)$ solves the Black-Scholes PDE for $t < T$ with final-time data $V(s,T) = f(s)$, then we can create an "artificial option" with initial cost $V(s(0),0)$ at time 0 and payoff $f(s(T))$ as follows: at any time $t$, hold $\phi(t)$ units of stock and a bond worth $V(s(t),t) - \phi s(t)$, where $\phi(t) = V_s(s(t),t)$. This strategy is *self-financing*, because the increment in its value from time $t$ to time $t + dt$ is exactly equal to the profit or loss associated with stock movements plus the interest on the bond. In formulas: the total value of the portfolio at time $t$ is $\phi(t)s(t) + (V(s(t),t) - \phi(t)s(t)) = V(s(t),t)$, and the portfolio is self-financing because

$$dV = \phi \, ds + r(V - \phi s) \, dt$$

as a direct consequence of Ito's lemma and our choice of $\phi$ (this is just a reorganization of the calculation done previously). The portfolio evidently has has value $V(s(0),0)$ at time 0 and $V(s(T),T) = f(s(T))$ at time $T$. Thus the proposed strategy creates an "artificial option" with payoff $f(s(T))$ and initial cost $V(s(0),0)$.

**Reduction of the Black-Scholes PDE to the linear heat equation.** The linear heat equation $u_t = u_{xx}$ is the most basic example of a parabolic PDE; its properties and solutions are discussed in every textbook on PDE's. The Black-Scholes equation is really just this standard equation written in special variables. This fact is very well-known; the following discussion follows the book by Dewynne, Howison, and Wilmott.

Recall that the Black-Scholes PDE is

$$V_t + \tfrac{1}{2}\sigma^2 s^2 V_{ss} + rsV_s - rV = 0;$$

we assume in the following that $r$ and $\sigma$ are constant. Consider the preliminary change of variables from $(s, t)$ to $(x, \tau)$ defined by

$$s = e^x, \quad \tau = \tfrac{1}{2}\sigma^2(T - t),$$

and let $v(x, \tau) = V(s, t)$. An elementary calculation shows that the Black-Scholes equation becomes

$$v_\tau - v_{xx} + (1 - k)v_x + kv = 0$$

with $k = r/(\tfrac{1}{2}\sigma^2)$. We've done the main part of the job: reduction to a constant-coefficient equation. For the rest, consider $u(x, t)$ defined by

$$v = e^{\alpha x + \beta \tau} u(x, \tau)$$

where $\alpha$ and $\beta$ are constants. The equation for $v$ becomes an equation for $u$, namely

$$(\beta u + u_\tau) - (\alpha^2 u + 2\alpha u_x + u_{xx}) + (1 - k)(\alpha u + u_x) + ku = 0.$$

To get an equation without $u$ or $u_x$ we should set

$$\beta - \alpha^2 + (1 - k)\alpha + k = 0, \quad -2\alpha + (1 - k) = 0.$$

These equations are solved by

$$\alpha = \frac{1 - k}{2}, \quad \beta = -\frac{(k + 1)^2}{4}.$$

Thus,

$$u = e^{\frac{1}{2}(k-1)x + \frac{1}{4}(k+1)^2 \tau} v(x, \tau)$$

solves the linear heat equation $u_\tau = u_{xx}$.

What good is this? Well, it can be used to give another proof of the integral formula for the value of an option (using the fundamental solution of the linear heat equation). It can also be used to understand the sense in which the value of an option at time $t < T$ is obtained by "smoothing" the payoff. Indeed, the solution of the linear heat equation is obtained by "Gaussian smoothing" of the initial data. (See Wilmott-Howison-Dewynne, or else Chapter 11 of Steele's *Stochastic Calculus and Financial Applications*, for finance-oriented discussions of the linear heat equation.)

**What about options on forwards?** Now suppose $\mathcal{F}(t)$ is a forward price process. As above, we assume for simplicity that $\mathcal{F}$ is lognormal

$$d\mathcal{F} = \mu\mathcal{F}\,dt + \sigma\mathcal{F}\,dw$$

to keep things simple. (Important note: the following discussion applies regardless of whether the underlying asset has continuous dividend yield.)

Consider a European option with payoff $f(\mathcal{F}_T)$ at time $T$. Its value at time $t$ is $V(\mathcal{F}(t), t)$ where $V$ solves the PDE

$$V_t + \tfrac{1}{2}\sigma^2\mathcal{F}^2 V_{\mathcal{F}\mathcal{F}} - rV = 0 \quad \text{for } t < T$$

with $V(\mathcal{F}, T) = f(\mathcal{F})$ at $t = T$. This is the Black-Scholes PDE for options on forwards. Its justification is parallel to argument we used for a non-dividend-paying stock. We presented that argument in two slightly different ways; let's discuss just the analogue of the second one here (leaving the analogue of the first one to the reader, as an exercise). Consider the following strategy: at time $t$ hold $\alpha(t) = e^{r(T-t)}V_{\mathcal{F}}(\mathcal{F}(t), t))$ forwards (with delivery price equal to the forward price) and a bond worth $V(\mathcal{F}(t), t)$. Note that the value of this portfolio is $V(\mathcal{F}(t), t)$ since the forward contracts have value 0. I claim that the strategy is self-financing, i.e. that the change in its value from time $t$ to time $t + dt$ is equal to the profit/loss on the forward plus interest on the bond. To see this, we must show that

$$dV(\mathcal{F}(t), t) = \alpha e^{-r(T-t)}d\mathcal{F} + rV(\mathcal{F}(t), t))\,dt.$$

The left hand side, by Ito's formula, is

$$dV = V_{\mathcal{F}}\,d\mathcal{F} + (V_t + \sigma^2\mathcal{F}^2\tfrac{1}{2}V_{\mathcal{F}\mathcal{F}})\,dt.$$

The right hand side is

$$\alpha e^{-r(T-t)}\,d\mathcal{F} + rV(\mathcal{F}(t), t))\,dt = V_{\mathcal{F}}d\mathcal{F} + rV(\mathcal{F}(t), t)$$

by our choice of $\alpha$. The two expressions are equal as a consequence of the PDE. The cost of initiating this strategy at time 0 is $V(\mathcal{F}(0), 0)$. The final-time value of the portfolio is $V(\mathcal{F}(T), T) = f(\mathcal{F}_T)$, using the final-time condition of the PDE. So this strategy produces an "artificial option" with payoff $f(\mathcal{F}_T)$ at initial cost $V(\mathcal{F}(0), 0)$.

The PDE $V_t + \tfrac{1}{2}\sigma^2\mathcal{F}^2 V_{\mathcal{F}\mathcal{F}} - rV = 0$ can also be reduced to the linear heat equation (provided $\sigma$ and $r$ are constant). The argument is very similar to the one given earlier. Starting as before, let

$$\mathcal{F} = e^x, \quad \tau = \tfrac{1}{2}\sigma^2(T - t),$$

and let $v(x, \tau) = V(\mathcal{F}, t)$. An elementary calculation shows that the PDE becomes

$$v_\tau - v_{xx} + v_x + kv = 0$$

with $k = r/(\frac{1}{2}\sigma^2)$. Now suppose $u(x,\tau)$ is defined by

$$v = e^{\alpha x + \beta \tau} u(x, \tau)$$

where $\alpha$ and $\beta$ are constants. The equation for $v$ becomes an equation for $u$, namely

$$(\beta u + u_\tau) - (\alpha^2 u + 2\alpha u_x + u_{xx}) + (\alpha u + u_x) + ku = 0.$$

To get an equation without $u$ or $u_x$ we should set

$$\beta - \alpha^2 + \alpha + k = 0, \quad -2\alpha + 1 = 0.$$

These equations are solved by

$$\alpha = \frac{1}{2}, \quad \beta = k - \frac{1}{4}.$$

Thus,

$$u = e^{\frac{1}{2}x + (k - \frac{1}{4})\tau} v(x, \tau)$$

solves the linear heat equation $u_\tau = u_{xx}$.

$$\text{*******************}$$

**More stochastic calculus**. Consider a stochastic integral of the form $\int_a^b g(s)\, dw(s)$ where $g$ is a deterministic function of $s$. It has mean value zero – we explained that in the first few pages of this section. What about its variance? The answer is simple:

$$E\left[\left(\int_a^b g(s)\, dw\right)^2\right] = \int_a^b g^2(s)\, ds.$$

Here is why. Approximating the stochastic integral by a sum, we see that the square of the stochastic integral is approximately

$$\left(\sum_i g(s_i)[w(s_{i+1}) - w(s_i)]\right)\left(\sum_j g(s_j)[w(s_{j+1}) - w(s_j)]\right)$$

$$= \sum_{i,j} g(s_i)g(s_j)[w(s_{i+1}) - w(s_i)][w(s_{j+1}) - w(s_j)] \quad .$$

For $i \neq j$ the expected value of the $i, j$th term is 0 since $[w(s_{j+1}) - w(s_j)]$ and $[w(s_{i+1}) - w(s_i)]$ are independent Gaussians, each with mean value 0. For $i = j$ the expected value of the $i, j$th term is $g^2(s_i)(s_{i+1} - s_i)$. So the expected value of the squared stochastic integral is approximately

$$\sum_i g^2(s_i)(s_{i+1} - s_i),$$

which is a Riemann sum for $\int_a^b g^2(s)\, ds$. By the way: since we are assuming that $g$ is deterministic, $\int_a^b g(s)dw(s)$ is a *Gaussian* random variable. (Proof: recall that a sum of Gaussians is Gaussian; therefore $\sum_i g(s_i)[w(s_{i+1}) - w(s_i)]$ is Gaussian. Now use the fact

that a limit of Gaussians is Gaussian.) Since we know its mean and variance, we have completely characterized this random variable.

We are in the habit of focusing on lognormal dynamics, because this is the most basic model for the price of a stock (or the forward price of a stock). Another exactly-solvable SDE is the *Ornstein-Uhlenbeck* process, which solves

$$dy = -cy\,dt + \sigma\,dw, \quad y(0) = y_0$$

with $c$ and $\sigma$ constant. (This is *not* a lognormal process, because the coefficient of $dw$ is not proportional to $y$.) Ito's lemma gives

$$d(e^{ct}y) = ce^{ct}y\,dt + e^{ct}dy = e^{ct}\sigma\,dw$$

so

$$e^{ct}y(t) - y_0 = \sigma \int_0^t e^{cs}\,dw,$$

or in other words

$$y(t) = e^{-ct}y_0 + \sigma \int_0^t e^{c(s-t)}\,dw(s).$$

We see (using the discussion at the beginning of this section) that $y(t)$ is a Gaussian random variable. So it is entirely characterized by its mean and variance. They are easy to compute: the mean is clearly

$$E[y(t)] = e^{-ct}y_0$$

since the "dw" integral has expected value 0, and the variance is

$$
\begin{aligned}
E\left[(y(t) - E[y(t)])^2\right] &= \sigma^2 E\left[\left(\int_0^t e^{c(s-t)}\,dw(s)\right)^2\right] \\
&= \sigma^2 \int_0^t e^{2c(s-t)}\,ds \\
&= \sigma^2 \frac{1 - e^{-2ct}}{2c}.
\end{aligned}
$$

The Ornstein-Uhlenbeck process is relevant to finance. One of the simplest interest-rate models is that of Vasicek, which supposes that the (short-term) interest rate $r(t)$ satisfies

$$dr = a(b - r)dt + \sigma\,dw$$

with $a$, $b$, and $\sigma$ constant. Interpretation: $r$ tends to revert to some long-term average value $b$, but noise keeps perturbing it away from this value. Clearly $y = r - b$ is an Ornstein-Uhlenbeck process, since $dy = -ay\,dt + \sigma\,dw$. Notice that $r(t)$ has a positive probability of being negative (since it is a Gaussian random variable); as a consequence the Vasicek model is not very realistic. Even so, its exact solution formulas provide helpful intuition.