THE DERIVATION OF THE BLACK-SCHOLES PARTIAL DIFFERENTIAL EQUATION

Learning Objectives:

- The famous Black-Schole argument is presented in this lecture - know this by heart (the argument leading to the PDE, the PDE, the formulas for European call and put options)

- Be aware that PDEs are standard mathematical objects. Therefore obtaining the price of an option in the form of a PDE is almost as good as having the price itself. Numerical methods will be applied to extract the price from the PDE.

- Work out the Greeks.

- The Black-Scholes argument is generally applicable to situations in which a derivative is hedged. We will see the argument in further action through the pricing of multi-asset options and the pricing of zero-coupon bonds.

1. The Black-Scholes PDE

The Black-Scholes partial differential equation relates the price of an European option to the parameters of the Black-Scholes Model. There are 5 parameters: S - the underlying stock price, K - the strike price, σ - the volatility of the stock price (more precisely, the volatility of the log price), r - the risk-free rate, T - the expiry. The PDE is obtained by the following argument.

First, it is assumed that we have a fixed horizon time $[0, T]$ and the stock price S_t $(t \in [0, T])$ satisfies the geometric Brownian motion:

$$dS_t/S_t = \mu dt + \sigma dW_t,$$

where W_t denotes a Brownian motion.

Consider writing an European option with price V_t at time t and which matures at T. Amongst the 5 parameters, assume that only S and t are changing over time (i.e. the risk-free rate and the volatility are assumed to be constant). Then we adopt the perspective:

$$V = V(S, t) = V(S, t; K, T, \sigma, r).$$

In order to hedge the risk in writing the option, the writer needs to maintain a portfolio of Δ shares of the underlying stock and some cash, always balancing between the two assets to keep Δ at a level that is commensurate with the risk of the option and to maintain a self-financing portfolio (i.e. no external cash in- or out-flow during the hedging process). The quantity Δ is determined by making the following portfolio

$$\Pi = V - \Delta S$$

insensitive (at the first order) to movements in the stock price.

Let's now compute infinitesimal change to Π, using Ito's Lemma:

$$\begin{aligned} d\Pi &= dV - \Delta dS \\ &= V_t dt + V_S dS + \frac{1}{2} V_{SS} dS^2 - \Delta dS \\ &= V_t dt + V_S dS + \frac{1}{2} V_{SS} (\mu S dt + \sigma S dW_t)^2 - \Delta dS \\ &= V_t dt + V_S dS + \frac{1}{2} V_{SS} S^2 \sigma^2 dt - \Delta dS \end{aligned}$$

The dependence in stock price over the time period $[t, t+dt]$ is due to the term $V_S dS$. Hence, Δ is chosen to be equal to V_S o eliminate this dependence.

Now the portfolio is riskless and must therefore grow by the risk-free rate by the principle of no-arbitrage (i.e. if equality does not hold, long one and short the other will give riskless profits):

$$d\Pi = r\Pi dt.$$

When written out, this reads

$$V_t dt + \frac{1}{2} V_{SS} S^2 \sigma^2 dt = r(V - V_S S)dt,$$

which is equivalent to

$$V_t + \frac{1}{2} S^2 \sigma^2 V_{SS} + rSV_S - rV = 0.$$

This is then the famous Black-Scholes PDE.

The boundary condition to the PDE is either one of the following:

(1) Call option: $V(S,T) = (S-K)^+$

(2) Put option $V(S,T) = (K-S)^+$

The PDE is a second order linear (in V) PDE and can be transformed into the heat equation, a standard PDE in mathematics, which can readily be solved to give the solutions:

(1) Call option price at time t:

$$V_t = S_t N(d_1) - Ke^{-r(T-t)} N(d_2);$$

(2) Put option price at time t:

$$V_t = -S_t N(-d_1) + Ke^{-r(T-t)} N(-d_2);$$

where

$$d_1 = \frac{\log(S_t/K) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}},$$

$$d_2 = \frac{\log(S_t/K) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}.$$

Example 1. *Show that* $d_1 = d_2 + \sigma\sqrt{T-t}$.

Example 2. *Verify the Put-Call Parity from the formulas above.*

Example 3. *Explain how the derivation needs to be altered if a long position in the underlying asset gives a yield of* q*, i.e. every unit gives* $qS_t dt$ *over an infinitesimal interval of length* dt*. Consequently, show that the Black-Scholes PDE modified for an underlying asset the gives a yield of* q *is*

$$V_t + \frac{1}{2}S^2\sigma^2 V_{SS} + (r-q)SV_S - rV = 0.$$

Note that this example accounts for currency options, where the yield is the interest rate of the foreign currency; as well as for futures options, where the yield is the risk-free rate - without the need to borrow money to buy shares, the situation is equivalent to having a yield equal to the carry cost.

The above argument was that put forth by Black-Scholes and Merton in their explanation of the price of European options in the framework of the principle of no-arbitrage. Scholes and Merton got their Nobel Prize for showing that the idiosyncratic risk premium of a company stock price is irrelevant to the price of an option and the framework applies to the pricing of all financial derivatives. The irrelevance of the idiosyncratic risk premium is reflected in the PDE above by the absence of the term μ which was earlier presumed to represent the fundamentals of the company in the geometric Brownian motion model for its stock price. Read carefully through the derivation to see how μ disappeared along the way!

A list of unrealistic assumptions have been made in the derivation - they are necessary for the creation of a 'nice' theory:

(1) The underlying stock price moves in accordance to the geometric Brownian motion

(2) One can long or short the stock at ease

(3) There is no position limit

(4) One is a small investor in the sense that one's trading actions do not affect the stock price

(5) There is no tax, dividend nor transaction cost issues

(6) Continuous time

(7) Perfect divisibility of assets

(8) Risk-free rate

Researchers have over the years since the Black-Scholes paper was published in 1973 to improve the model over these assumptions.

The perfect divisibility of assets and of time is a feature of a continuous model. The belief is that by discretizing a continuous model one is able to recover some degree of realism and the cost of inaccuracy far outweighs the benefits that the tools from real analysis bring to bear in a continuous framework.

To improve the risk-free rate assumption is to ask to model stochastic interest rates. This is a field on its own in quantitative finance known as Interest Rate Modelling.

There are two major aspects pertaining to the geometric Brownian motion assumption. Researchers have proposed using other models of stock price stochasticity which replace the Brownian noise factors by processes that include jumps (a general tractable class is the Levy processes). The argument is that stock prices do not always gradually diffuse over time and extraordinary jumps can be observed often. The other thread of research is via pointing out that stock price volatility is not a constant in reality. This can be seen by either computing standard deviation of returns of stocks or by looking at the implied volatility surfaces of options (they are not flat!). We will consider an extension of the Black-Scholes model by looking at a couple of simple models that include jumps or stochastic volatility later.

All these, however, do not change the fundamental fact that there is no model that can truly model the market - a facet of human activities.

The word 'model' must be interpreted differently in pure sciences and in a situation such as this.

Let us consider here the effect of discretizing the time.

Assume that the stock price satisfies the discrete version of the GBM:

$$\delta S = \mu S \delta t + \sigma S \sqrt{\delta t} Z,$$

where $Z \sim N(0,1)$.

Form the hedging portfolio:

$$\Pi = V - \Delta S.$$

Expand it up to the first order in δt:

$$\delta\Pi \approx V_t \delta t + V_S \delta S + \frac{1}{2} V_{SS} \delta S^2 - \Delta \delta S$$

$$= V_t \delta t + V_S(\mu S \delta t + \sigma S \sqrt{\delta t} Z) + \frac{1}{2} V_{SS} (\mu S \delta t + \sigma S \sqrt{\delta t} Z)^2 - \Delta \delta S$$

$$= V_t \delta t + V_S(\mu S \delta t + \sigma S \sqrt{\delta t} Z) + \frac{1}{2} V_{SS} \sigma^2 S^2 \delta t Z^2 - \Delta \delta S$$

$$= \delta t^{1/2} \sigma S (V_S - \Delta) Z + \delta t (V_t + \mu S (V_S - \Delta) + \frac{1}{2} \sigma^2 S^2 V_{SS} Z^2),$$

where we have dropped powers of δt which are greater than 1.

Remark: When one drops higher order terms, one is assuming that they are negligible. We should always bear in mind that such convenient assumptions may have significant consequences when models are used in reality, since the dropped terms may not be negligible at all. In the present framework of Brownian noise, the dropping of the higher order terms for being negligible is justified.

The choice of Δ removes two terms in the expression above, as in the continuous derivation. What differs from this setting (excluding the fact that we have already dropped higher order terms) is the term that contains Z^2 - in the continuous derivation, 1 takes the place of Z^2. In other words, the most significant effect of discretization is that the change $\delta\Pi$ in the hedging portfolio value is not deterministic.

Writing the above as

$$\delta t(V_t + \frac{1}{2}\sigma^2 S^2 V_{SS} + \frac{1}{2}\sigma^2 S^2 V_{SS}(Z^2 - 1)),$$

we see that the discrepancy from the continuous model is the term

$$\frac{1}{2}\sigma^2 S^2 V_{SS}(Z^2 - 1).$$

In other words, if one hedges faithfully in the continuous model, charging the option at the price given by the Black-Scholes formula would have completely been fair. In the discrete model, there is a discrepancy in the term $\delta\Pi$ as compared to what happens in the continuous case. The discrepancy is the term shown here. It may thus be interpreted as the discrete hedging error.

The hedging error is of the order δt. For a time horizon of $[0, T]$, one hedges a total of $T/\delta t$ times. Since the mean of $Z^2 - 1$ is 0, the expected value of the total discrete hedging error is 0 (naturally so otherwise the option price computed by the Black-Scholes formula would not be correct). The standard deviation of each discrete hedging error is on the order of δt, hence the standard deviation of the total discrete hedging error is on the order of

$$\left(\frac{T}{\delta t}\right)^{\frac{1}{2}}\delta t = \sqrt{\delta t T} = O(\sqrt{\delta t}).$$

2. WHAT'S A PDE?

A PDE is an equation that relates partial derivatives. The Black-Scholes PDE

$$V_t + \frac{1}{2}S^2\sigma^2 V_{SS} + rSV_S - rV = 0$$

involves partial derivatives of the option price V with respect to t and S. It is thus a second order PDE. It is a linear PDE because the the associated PDE operator

$$V \mapsto V_t + \frac{1}{2}S^2\sigma^2 V_{SS} + rSV_S - rV$$

is linear.

PDEs almost always come with either boundary conditions or initial conditions that describe what happens to the principal quantities at 'edge' of the domain of definition. For instance, the Black-Scholes PDE has the boundary condition $V(S,T) = (S-K)^+$ or $(K-S)^+$ (or you may call it: final condition).

Partial differential equations are how physical laws of nature express themselves. For instance, the diffusion of heat is governed by the heat equation

$$u_t = u_{xx},$$

in one dimension (i.e. on a rod), or

$$u_t = u_{xx} + u_{yy},$$

in two dimensions (i.e. on a sheet of laminar). Here u represents temperature, t represents time and x, y represents measures of spatial extent. The heat equation is also known as the diffusion equation.

The terms in the Black-Scholes PDE have both financial as well as physical interpretations. The term $V_t + \frac{1}{2}S^2\sigma^2 V_{SS}$ which represents the change in the hedging portfolio value (which comprises 'time decay' and 'hedging benefits') is close to the heat equation - it can be thought of as smoke particles slowly being diffused around in some space. The term rSV_S which represents the carry cost of the underlying stock (without the dt term) is a convection term - it can be thought of as a breeze blowing the smoke particles in a preferred direction. The term $-rV$ is a reaction term - it represents radioactive decay with half-life related to r. Put together, the Black-Scholes PDE is an instance of a reaction-convection-diffusion PDE (the equation governing the dispersion of pullutant along a flowing river with absorption by the sand is similar to it).

The Black-Scholes PDE may be transformed into the heat equation by the following transformation:

$$V(S,t) = e^{\alpha x + \beta \tau} U(x,\tau),$$

where

$$\alpha = -\frac{1}{2}\big(\frac{2r}{\sigma^2} - 1\big),\ \beta = -\frac{1}{4}\big(\frac{2r}{\sigma^2} + 1\big)^2,\ S = e^x,\ t = T - \frac{2\tau}{\sigma^2}.$$

This leads to

$$U_\tau = U_{xx}.$$

Example 4. *Verify this!*

The boundary condition for the Black-Scholes PDE is similar transformed into a boundary condition for the heat equation. Since the heat equation admits an explicit solution (which was found by mathematicians hundreds of years ago), the Black-Scholes formula can be hence derived.

We will show how the formulas are derived fully from the other approach - by probability theory and computing expectations, in Lecture 7.

Example 5. *Show that the transformation*

$$\xi = Se^{-q(T-t)},\ \Phi(\xi, t) = V(\xi e^{q(T-t)}, t)$$

transforms the modified Black-Scholes PDE for yield-giving asset to the original Black-Scholes PDE. Hence, write down the formula for the European call option when the underlying is a stock which gives dividends at yield q.

3. THE GREEKS

Greeks or sensitivities (i.e. the partial derivatives of the option price relative to the model parameters) are crucial in the Black-Scholes theory. The PDE involves V_t, V_S, V_{SS} which are related to the time-change in value, the hedging quantity and the hedging effect/benefit.

Several of the Greeks in the Black-Scholes model are named:

- Delta: V_S
- Gamma: V_{SS}
- Vega: V_σ
- Theta: V_t
- Rho: V_r

They admit formulas like the price.

Example 6. *Show that for the call option,*

(1) $V_S = N(d_1)$
(2) $V_{SS} = \frac{\phi(d_1)}{S\sigma\sqrt{T}}$
(3) $V_\sigma = S\phi(d_1)\sqrt{T}$
(4) $V_t = \frac{S\phi(d_1)\sigma}{2\sqrt{T}} + rKe^{-rT}N(d_2)$
(5) $V_r = KTe^{-rT}N(d_2)$

Here, ϕ is the density function of the standard normal.

Example 7. *Why does the notion of the implied volatility make sense?*

Example 8. *Use the Put-Call parity to give a quick derivation of the Greeks of the put option.*

4. APPLYING THE ARGUMENT TO MULTI-ASSET OPTIONS

Stochastic calculus may be extended to more than one Brownian motion process, say $W_t^1, W_t^2, \ldots, W_t^n$. The basic rules to remember, which is a consequence of Ito's Lemma in higher dimensions, are

$$dt^2 = 0,\ dtdW^i = 0,\ dW^i dW^j = \rho_{ij} dt,$$

where $\rho_{ij} = \rho_{ji}$ and $\rho_{ii} = 1$ $(i, j = 1, 2, \ldots, n)$ encode the correlation structure of these Brownian factors.

The stock prices $S_1, S_2, \ldots, S_n$ are required to satisfy the geometric Brownian motion

$$dS^i/S^i = \mu_i dt + \sigma_i dW_t^i \ (i = 1, 2, \ldots, n).$$

The derivative contract we're interest in pricing at present has a payoff of

$$V(S^1, S^2, \ldots, S^n; T) = (\sum_{i=1}^n a_i S^i - K)^+.$$

Let's form the hedging portfolio:

$$\Pi = V - \sum_{i=1}^n \Delta_i S^i.$$

Consider an infinitesimal change in the portfolio value:

$$d\Pi = dV - \sum_{i=1}^n \Delta_i dS^i$$

$$= V_t dt + \sum_{i=1}^n V_{S^i} dS^i + \sum_{i,j=1}^n \frac{1}{2} V_{S^i S^j} dS^i dS^j - \sum_{i=1}^n \Delta_i dS^i$$

$$= V_t dt + \sum_{i=1}^n V_{S^i} dS^i + \sum_{i,j=1}^n \frac{1}{2} V_{S^i S^j} \sigma_i \sigma_j \rho_{ij} S^i S^j dt - \sum_{i=1}^n \Delta_i dS^i.$$

Select each Δ^i to be:

$$\Delta^i = V_{S^i}$$

to eliminate the terms that contain changes in asset prices to render the portfolio momentarily risk-free, giving:

$$d\Pi = V_t dt + \sum_{i,j=1}^{n} \frac{1}{2} V_{S^i S^j} \sigma_i \sigma_j \rho_{ij} S^i S^j dt.$$

Applying the principle of no-arbitrage, we obtain

$$d\Pi = r\Pi dt,$$

which implies

$$V_t dt + \sum_{i,j=1}^{n} \frac{1}{2} V_{S^i S^j} \sigma_i \sigma_j \rho_{ij} S^i S^j dt = r(V - \sum_{i=1}^{n} V_{S^i} S^i) dt,$$

or

$$V_t + \sum_{i,j=1}^{n} \frac{1}{2} \sigma_i \sigma_j \rho_{ij} S^i S^j V_{S^i S^j} + r \sum_{i=1}^{n} V_{S^i} S^i - rV = 0.$$

Together with the boundary condition given by the payoff at maturity of contract, we obtain the multi-dimensional analogue of the Black-Scholes PDE.

There is no explicit solution however.

Let us clarify the meaning of the statement:

$$dW^i dW^j = \rho_{ij} dt.$$

We need the notion of the cross variation.

Definition 1. *For any two stochastic processes X, Y, the cross variation $[X, Y]$ is the stochastic process defined by*

$$[X, Y]_t := \lim_{|\Pi| \to 0} \sum_{k=0}^{n-1} (X_{t_{k+1}} - X_{t_k})(Y_{t_{k+1}} - Y_{t_k}),$$

where Π is the partition $0 = t_0 < t_1 < \cdots < t_n = t$ of $[0, t]$, provided that the limit exists. The quadratic variation of X is defined to be $[X, X]$ provided that it exists.

The informal way of writing $d[X, Y]_t$ is $dX_t dY_t$.

Proposition 1. *The cross variation is positive bilinear in X, Y (provided that cross variations involved all exist).*

Positivity means that $[X, X] \geq 0$ for all X for which the quadratic variation makes sense.

Proof. Left for exercise. $\square$

Proposition 2. *If $Z^1, Z^2, \ldots, Z^n$ are independent Brownian processes, then the following holds:*

$$[Z^i, Z^j]_t = \delta_{ij} t.$$

Here δ_{ij} is the Kronecker's delta.

Proof. First assume that $i \neq j$. Let Π be a partition of $[0, t]$: $0 = t_0 < t_1 < \cdots < t_n = t$, and let

$$C_\Pi = \sum_{k=0}^{n-1} (Z^i_{t_{k+1}} - Z^i_{t_k})(Z^j_{t_{k+1}} - Z^j_{t_k}).$$

Note that

$$E[C_\Pi] = \sum_{k=0}^{n-1} E[Z^i_{t_{k+1}} - Z^i_{t_k}] E[Z^j_{t_{k+1}} - Z^j_{t_k}] = 0.$$

And

$$\begin{aligned}
Var(C_\Pi) &= E[C_\Pi^2] - E[C_\Pi]^2 \\
&= E[\sum_{k=0}^{n-1} (Z^i_{t_{k+1}} - Z^i_{t_k})^2 (Z^j_{t_{k+1}} - Z^j_{t_k})^2 \\
&\quad + 2 \sum_{l<k}^{n-1} (Z^i_{t_{l+1}} - Z^i_{t_l})(Z^j_{t_{l+1}} - Z^j_{t_l})(Z^i_{t_{k+1}} - Z^i_{t_k})(Z^j_{t_{k+1}} - Z^j_{t_k})] \\
&= E[\sum_{k=0}^{n-1} (Z^i_{t_{k+1}} - Z^i_{t_k})^2 (Z^j_{t_{k+1}} - Z^j_{t_k})^2] \\
&= \sum_{k=0}^{n-1} (t_{k+1} - t_k)^2 \\
&< |\Pi| \sum_{k=0}^{n-1} (t_{k+1} - t_k) \\
&= |\Pi| t \to 0.
\end{aligned}$$

Hence, the random variable C_Π converges to 0.

Now assume $i = j$. Then

$$\begin{aligned} E[C_\Pi] &= \sum_{k=0}^{n-1} E[(Z^i_{t_{k+1}} - Z^i_{t_k})^2] \\ &= \sum_{k=0}^{n-1} (t_{k+1} - t_k) \\ &= t, \end{aligned}$$

and

$$\begin{aligned} Var(C_\Pi) &= E[C_\Pi^2] - E[C_\Pi]^2 \\ &= E[\sum_{k=0}^{n-1} (Z^i_{t_{k+1}} - Z^i_{t_k})^4 \\ &\quad + 2\sum_{l<k}^{n-1} (Z^i_{t_{l+1}} - Z^i_{t_l})^2 (Z^i_{t_{k+1}} - Z^i_{t_k})^2] - t^2 \\ &= E[Z^4] \sum_{k=0}^{n-1} (t_{k+1} - t_k)^2 + 2E[Z^2]^2 \sum_{l<k}^{n-1} (t_{l+1} - t_l)(t_{k+1} - t_k) - t^2 \\ &= E[Z^4] \sum_{k=0}^{n-1} (t_{k+1} - t_k)^2 + 2\sum_{k=1}^{n-1} t_k (t_{k+1} - t_k) - t^2 \\ &\to 0 + 2\int_0^t s\, ds - t^2 = 0. \end{aligned}$$

□

Now how do we obtain Brownian motions $Z^1, Z^2, \dots, Z^n$ such that

$$dZ^i_t dZ^j_t = \rho_{ij} dt\,?$$

The answer involves a standard technique. The heuristic goes like this. We search for linear combinations of the W^i's (independent Brownian motions) to do the job:

$$AW = Z,$$

i.e.

$$Z^i = \sum_{j=1}^{n} a_{ij} W^j.$$

This implies that

$$\begin{aligned}[Z^i, Z^j] &= [\sum_{k=1}^n a_{ik} W^k, \sum_{l=1}^n a_{jl} W^l] \\ &= \sum_{k,l=1}^n a_{ik} a_{jl} [W^k, W^l] \\ &= \sum_{k=1}^n a_{ik} a_{jk},\end{aligned}$$

which we want to be ρ_{ij}. In other words, we want an equality of matrices

$$(\rho_{ij})_{1 \le i,j \le n} = AA^T.$$

Note that $(\rho_{ij})_{1 \le i,j \le n}$ is not an arbitrary matrix. It is in fact positive definite, i.e.

$$\begin{aligned}\mathbf{v}^T (\rho_{ij})_{1 \le i,j \le n} \mathbf{v} &= \mathbf{v}^T AA^T \mathbf{v} \\ &= \sum_{i,j,k=1}^n v_i a_{ik} a_{jk} v_j \\ &= \sum_{i,j=1}^n v_i [Z^i, Z^j] v_j \\ &= [\sum_{i=1}^n v_i Z^i, \sum_{j=1}^n v_j Z^j] \\ &= [\mathbf{v} \bullet Z, \mathbf{v} \bullet Z] > 0\end{aligned}$$

for every non-zero vector $\mathbf{v}$.

Example 9. *Explain why the above inequality holds (recall the notion of the quadratic variation from Lecture 5).*

The following theorem comes to our rescue:

Theorem 1. *A positive definite matrix B admits a Cholesky decomposition*

$$B = CC^T.$$

The Cholesky decomposition can be computed algorithmically and is implemented in good mathematical packages such as Matlab.

Example 10. *Given independent Brownian motions* W^1, W^2*. Write down Brownian motions* Z^1, Z^2 *such that* $dZ^1 dZ^2 = \rho dt$ *in terms of* W^1, W^2*.*

5. Applying the Argument to Bonds

Let the short rate be governed by the following model:

$$dr_t = a(r_t, t)dt + b(r_t, t)dW_t$$

and consider the problem of pricing the (zero coupon) bond which matures at time T, at time t:

$$B = B^T = B(r, t; T).$$

Consider an infinitesimal change:

$$\begin{aligned} dB &= B_t dt + B_r dr + \frac{1}{2} B_{rr} dr^2 \\ &= B_t dt + B_r(a(r_t, t)dt + b(r_t, t)dW_t) + \frac{1}{2} B_{rr} b(r_t, t)^2 dt \\ &= \big(B_t + B_r a(r_t, t) + \frac{1}{2} B_{rr} b(r_t, t)^2\big)dt + b(r_t, t)B_r dW \\ &=: \mu^T B dt + \sigma^T B dW. \end{aligned}$$

In the stock and option case, we formed the hedging portfolio $\Pi = V - \Delta S$. However, in the present situation, we cannot form a similar hedging portfolio $\Pi = B - \Delta r$ because the short rate is not tradable. Instead, let's consider two bonds B^{T_1} and B^{T_2} and form the portfolio

$$\Pi = \Delta_1 B^{T_1} - \Delta_2 B^{T_2},$$

which we wish to be risk-free.

Consider an infinitesimal change:

$$\begin{aligned} d\Pi &= \Delta_1 dB^{T_1} - \Delta_2 dB^{T_2} \\ &= \Delta_1(\mu^{T_1} B^{T_1} dt + \sigma^{T_1} B^{T_1} dW) - \Delta_2(\mu^{T_2} B^{T_2} dt + \sigma^{T_2} B^{T_2} dW) \\ &= (\Delta_1 \mu^{T_1} B^{T_1} - \Delta_2 \mu^{T_2} B^{T_2})dt + (\Delta_1 \sigma^{T_1} B^{T_1} - \Delta_2 \sigma^{T_2} B^{T_2})dW. \end{aligned}$$

To be risk-free is to set

$$\Delta_1 \sigma^{T_1} B^{T_1} - \Delta_2 \sigma^{T_2} B^{T_2} = 0.$$

Invoking the Principle of No-Arbitrage, we have

$$d\Pi = r_t \Pi dt,$$

which gives

$$\begin{aligned}
\Delta_1\mu^{T_1}B^{T_1} - \Delta_2\mu^{T_2}B^{T_2} &= r_t(\Delta_1 B^{T_1} - \Delta_2 B^{T_2}) \\
\Rightarrow \sigma^{T_1}\Delta_1\mu^{T_1}B^{T_1} - \sigma^{T_1}\Delta_2\mu^{T_2}B^{T_2} &= r_t(\sigma^{T_1}\Delta_1 B^{T_1} - \sigma^{T_1}\Delta_2 B^{T_2}) \\
\Rightarrow \sigma^{T_2}\Delta_2\mu^{T_1}B^{T_2} - \sigma^{T_1}\Delta_2\mu^{T_2}B^{T_2} &= r_t(\sigma^{T_2}\Delta_2 B^{T_2} - \sigma^{T_1}\Delta_2 B^{T_2}) \\
\Rightarrow \sigma^{T_2}\mu^{T_1} - \sigma^{T_1}\mu^{T_2} &= r_t(\sigma^{T_2} - \sigma^{T_1}) \\
\Rightarrow \sigma^{T_2}(\mu^{T_1} - r_t) &= \sigma^{T_1}(\mu^{T_2} - r_t) \\
\frac{\mu^{T_1} - r_t}{\sigma^{T_1}} &= \frac{\mu^{T_2} - r_t}{\sigma^{T_2}},
\end{aligned}$$

which shows that the quantity $\frac{\mu^T - r_t}{\sigma^T}$ is independent of the (maturity of the) bond. Hence, let us define

$$\lambda(r_t, t) := \frac{\mu^T - r_t}{\sigma^T}$$

and call it the market price of risk.

Let's summarize what we have so far:

- Drift:
$$\mu^T B = B_t + B_r a(r_t, t) + \frac{1}{2} B_{rr} b(r_t, t)^2$$
- Volatility:
$$\sigma^T B = b(r_t, t) B_r$$
- Market price of risk:
$$\lambda = \frac{\mu^T - r_t}{\sigma^T}$$

The volatility and market price of risk equations imply:

$$\mu^T = \sigma^T \lambda(r_t, t) + r_t = \frac{b(r_t, t)B_r}{B} \cdot \lambda(r_t, t) + r_t.$$

Substituting that into the drift equation, we obtain the following PDE for the bond:

$$B_t + B_r a(r_t, t) + \frac{1}{2} B_{rr} b(r_t, t)^2 = b(r_t, t)\lambda(r_t, t) B_r + r_t B,$$

together with the boundary condition

$$B(T, r_t) = 1.$$

We have used the notation:

$$B = B^T = B(r_t, t; T) = B^T(r_t, t).$$

Remark: The market price of risk is a term that was invented by the academics to denote a term that is unknown which drops out of the theory in a somewhat distinguished manner. In calibrating models that involve the market price of risk, indirect methods need to used to extract this number from the market since it is not observable (academics, in being kind to themselves, say that the term is not 'directly observable' from the market). However, one should bear in mind that this is what our theory says the market should exhibit but the market certainly does not behave in a certain way just because our theory says so.

Remark: One may ask: why don't we just find the bond price from the yield curve which we may build from the usual bootstrapping method using the short rate. The answer is that the short rate model is stochastic. So instead of having one yield curve, we will end up having a whole probability space of yield curves. If we were to take the average bond price from these yield curves, it will be akin to taking the real world expectation of option payoff for the purpose of pricing an option. The Black-Scholes method requires instead that the noise that correspond to the uncertainty of the yield curve be hedged and the price of the derivative derived therefrom. In the scenario of this section, the short rate is the underlying while the bond is the derivative.