Risk-Neutral Asset Pricing

David Šiška
School of Mathematics, University of Edinburgh

Academic Year 2021/22*

Contents

*Last updated 16th January 2022

1 Essentials From Stochastic Analysis

"I turn with terror and horror from this lamentable scourge of continuous functions with no derivatives."

– Charles Hermite, 1893.

For convenience we state some results from stochastic analysis. Proofs can be found for example in Stochastic Analysis for Finance lecture notes, in [1] or [6].

1.1 Probability Space

Let us always assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space is fixed. We assume that $\mathcal{F}$ is complete which means that all the subsets of sets with probability zero are included in $\mathcal{F}$. We assume there is a filtration $(\mathcal{F}_t)_{t\in[0,T]}$ (which means $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$) such that $\mathcal{F}_0$ contains all the sets of probability zero.

1.2 Stochastic Processes, Martingales

A stochastic process $X = (X(t))_{t\geq 0}$ is a collection of random variables $X(t)$ which take values in $\mathbb{R}^d$.

We will always assume that stochastic processes are *measurable*. This means that $(\omega, t) \mapsto X(\omega, t)$ taken as a function from $\Omega \times [0,\infty)$ to $\mathbb{R}^d$ is measurable with respect to σ-algebra $\mathcal{F} \otimes \mathcal{B}([0,\infty))$. This product is defined as the σ-algebra generated by sets $E \times B$ such that $E \in \mathcal{F}$ and $B \in \mathcal{B}([0,\infty))$. From Theorem A.2 we then get that

$$t \mapsto X(\omega, t) \text{ is measurable for all } \omega \in \Omega.$$

We say X is $(\mathcal{F}_t)_{t\geq 0}$ *adapted* if for all $t \geq 0$ we have that $X(t)$ is $\mathcal{F}_t$-measurable.

Definition 1.1. *Let X be a stochastic process that is adapted to $(\mathcal{F}_t)_{t\geq 0}$ and such that for every $t \geq 0$ we have $\mathbb{E}[|X(t)|] < \infty$. If for every $0 \leq s < t \leq T$ we have*

i) $\mathbb{E}[X(t)|\mathcal{F}_s] \geq X(s)$ *a.s.then the process is called* submartingale.

ii) $\mathbb{E}[X(t)|\mathcal{F}_s] \leq X(s)$ *a.s.then the process is called* supermartingale.

iii) $\mathbb{E}[X(t)|\mathcal{F}_s] = X(s)$ *a.s.then the process is called* martingale.

For submartingales we have Doob's maximal inequality:

Theorem 1.2 (Doob's submartingale inequality)**.** *Let $X \geq 0$ be an $(\mathcal{F}_t)_{t\in[0,T]}$-submartingale with right-continuous sample paths and $p > 1$ be given. Assume $\mathbb{E}\left[X(T)^p\right] < \infty$. Then*

$$\mathbb{E}\left[\sup_{0\leq t\leq T} X(t)^p\right] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}\left[X(T)^p\right].$$

Definition 1.3 (Local Martingale)**.** *A stochastic process X is called a* local martingale *if there is a sequence of stopping times $(\tau_n)_{n\in\mathbb{N}}$ such that $\tau_n \leq \tau_{n+1}$ and $\tau_n \to \infty$ as $n \to \infty$ and if the* stopped process $(X(t \wedge \tau_n))_{t\geq 0}$ *is a martingale for every n.*

1.3 Integration Classes and Itô's Formula

Definition 1.4. *By $\mathcal{H}$ we mean all $\mathbb{R}$-valued and adapted processes g such that for any $T > 0$ we have*

$$\mathbb{E}\left[\int_0^T |g(s)|^2 ds\right] < \infty.$$

By $\mathcal{S}$ we mean all $\mathbb{R}$-valued and adapted processes g such that for any $T > 0$ we have

$$\mathbb{P}\left[\int_0^T |g(s)|^2 ds < \infty\right] = 1.$$

The importance of these two classes is that stochastic integral with respect to W is defined for all integrands in class $\mathcal{S}$ and this stochastic integral is a continous *local* martingale. For the class $\mathcal{H}$ the stochastic integral with respect to W is a martingale.

Definition 1.5. *By $\mathcal{A}$ we denote $\mathbb{R}$-valued and adapted processes g such that for any $T > 0$ we have*

$$\mathbb{P}\left[\int_0^T |g(s)| ds < \infty\right] = 1.$$

By $\mathcal{H}^{d\times n}$, $\mathcal{S}^{d\times n}$ we denote processes taking values the space of $d \times n$-matrices such that each component of the matrix is in $\mathcal{H}$ or $\mathcal{S}$ respectively. By $\mathcal{A}^d$ we denote processes taking values in $\mathbb{R}^d$ such that each component is in $\mathcal{A}$

We will need the multi-dimensional version of the Itô's formula. Let W be an m-dimensional Wiener martingale with respect to $(\mathcal{F})_{t\geq 0}$. Let $\sigma \in \mathcal{S}^{d\times m}$ and let $b \in \mathcal{A}^d$. We say that the d-dimensional process X has the stochastic differential

$$dX(t) = b(t)dt + \sigma(t)dW(t) \tag{1}$$

for $t \in [0, T]$, if

$$X(t) = X(0) + \int_0^t b(s)ds + \int_0^t \sigma(s)dW(s).$$

Such a process is also called an *Itô process*.

Theorem 1.6 (Multi-dimensional Itô formula)**.** *Let X be a d-dimensional Itô process given by* (1)*. Let $u \in C^{1,2}([0,T]\times\mathbb{R}^d)$. Then the process given by $u(t, X(t))$ has the stochastic differential*

$$\begin{aligned} du(t, X(t)) = u_t(t, X(t))dt &+ \sum_{i=1}^d u_{x_i}(t, X(t))dX^i(t) \\ &+ \frac{1}{2}\sum_{i,j=1}^d u_{x_i x_j}(t, X(t))dX^i(t)dX^j(t), \end{aligned}$$

where for $i, j = 1, \ldots, n$

$$dtdt = dtdW^i(t) = 0, \quad dW^i(t)dW^j(t) = \delta_{ij}dt.$$

Here and elsewhere δ_{ij} is the Kronecker δ. This means that $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$.

We now consider a very useful special case. Let X and Y be $\mathbb{R}$-valued Itô processes. We will apply to above theorem with $f(x, y) = xy$. Then $f_x = y$, $f_y = x$, $f_{xx} = f_{yy} = 0$ and $f_{xy} = f_{yx} = 1$. Hence from the multi-dimensional Itô formula we have

$$df(X(t), Y(t)) = Y(t)dX(t) + X(t)dY(t) + \frac{1}{2}dY(t)dX(t) + \frac{1}{2}dX(t)dY(t).$$

Hence we have the following corollary

Corollary 1.7 (Itô's product rule). *Let X and Y be $\mathbb{R}$-valued Itô processes. Then*

$$d(X(t)Y(t)) = X(t)dY(t) + Y(t)dX(t) + dX(t)dY(t).$$

1.4 Theorems of Lévy and Girsanov, Martingale Representation

Theorem 1.8 (Lévy characterization). *Let $(\mathcal{F}_t)_{t\in[0,T]}$ be a filtration. Let $X = (X(t))_{t\in[0,T]}$ be a continuous d-dimensional local martingale w.r.t. $(\mathcal{F}_t)_{t\in[0,T]}$ such that $X(0) = 0$ and $dX_i(t)dX_j(t) = \delta_{ij}dt$. Then X is a Wiener process adapted to $(\mathcal{F}_t)_{t\in[0,T]}$ and such that $W_{t'} - W_t$ is independent of $\mathcal{F}_t$ for all $t' \geq t \geq 0$.*

So essentially any continuous local martingale with the right quadratic variation is a Wiener process.

Theorem 1.9 (Girsanov). *Let $(\mathcal{F}_t)_{t\in[0,T]}$ be a filtration. Let $W = (W(t))_{t\in[0,T]}$ be a d-dimensional Wiener martingale with respect to $(\mathcal{F}_t)_{t\in[0,T]}$. Let $\varphi = (\varphi(t))_{t\in[0,T]}$ be a d-dimensional process adapted to $(\mathcal{F}_t)_{t\in[0,T]}$ such that*

$$\mathbb{E}\int_0^T |\varphi(s)|^2 ds < \infty.$$

Let

$$L(t) := \exp\left(-\int_0^t \varphi(s)^T dW(s) - \frac{1}{2}\int_0^t |\varphi(s)|^2 ds\right) \tag{2}$$

and assume that $\mathbb{E}(L(T)) = 1$. Let $\mathbb{Q}$ be a new measure on $\mathcal{F}_T$ given by the Radon-Nikodym derivative $d\mathbb{Q} = L(T)d\mathbb{P}$. Then

$$W^{\mathbb{Q}}(t) := W(t) + \int_0^t \varphi(s)ds$$

is a $\mathbb{Q}$-Wiener martingale.

We don't give proof but only make some useful observations.

1. Clearly $L(0) = 1$.

2. Applying Itô's formula to $f(x) = \exp(x)$ and

$$dX(t) = -\varphi(t)^T dW(t) - \frac{1}{2}|\varphi(t)|^2 dt$$

yields

$$dL(t) = -L(t)\varphi(t)^T dW(t).$$

This means that L is a local martingale. If we could show that it is actually a true martingale (e.g. by showing that $L\varphi \in \mathcal{H}^d$) then we would get that $\mathbb{E}[L(T)] = 1$.

3. The Novikov condition is a useful way of establishing that $\mathbb{E}[L(T)] = 1$: if

$$\mathbb{E}\left[e^{\frac{1}{2}\int_0^T |\varphi(t)|^2 dt}\right] < \infty$$

then L is a martingale (and hence $\mathbb{E}L(T) = \mathbb{E}L(0) = 1$).

Theorem 1.10 (Martingale representation). *Let $W = (W(t))_{t\in[0,T]}$ be a d-dimensional Wiener martingale and let $(\mathcal{F}_t)_{t\in[0,T]}$ be generated by W. Let $M = (M(t))_{t\in[0,T]}$ be a continuous real valued local martingale with respect to $(\mathcal{F}_t)_{t\in[0,T]}$. Then there exists a unique adapted d-dimensional process $h = (h(t))_{t\in[0,T]}$, $h \in \mathcal{S}$, such that for $t \in [0,T]$ we have*

$$M(t) = M(0) + \sum_{i=1}^{d} \int_0^t h_i(s) dW_i(s).$$

If the martingale M is square integrable then $h \in \mathcal{H}$.

For more details and proof see [6, Theorem 4.2 and Theorem 4.15]. Essentially what the theorem is saying is that we can write continuous martingales as stochastic integrals with respect to some process as long as they're adapted to the filtration generated by the process.

1.5 Exercises

Exercise 1.1. Show that $\mathcal{H} \subset \mathcal{S}$.

Exercise 1.2. Let $W_i = (W_i(t))_{t\in[0,T]}$ with $i = 1, 2$ be independent Wiener processes and let $\mu \in \mathbb{R}$ and $\rho \in [-1, 1]$ be a constant. Consider

$$dX_1(t) = \mu X_1(t)dt + \rho dW_1(t)$$

and

$$dX_2(t) = -\mu X_1(t)dt + \sqrt{1-\rho^2}dW_2(t).$$

Is the process $Z(t) := X_1(t) + X_2(t)$ a Wiener process?

Exercise 1.3. Use Itô formula to show that the following processes are local martingales with respect to the filtration generated by the real valued Wiener process W:

i) $X(t) = \exp((1/2)t)\cos W(t)$

ii) $X(t) = \exp((1/2)t)\sin W(t)$

iii) $X(t) = (W(t) + t)\exp(-W(t) - (1/2)t)$.

Which ones are martingales (rather than just local martingales)?

Exercise 1.4. Let W_1 and W_2 be two Wiener processes. Let $(\mathcal{F}_t)_{t\geq 0}$ be the filtration generated by *both* W_1, W_2 (i.e. $\mathcal{F}_t = \sigma\{W_1(s), W_2(s), s \leq t\}$). For an $\mathcal{F}_t$-adapted processes $h(t) := (h_1(t), h_2(t))$ satisfying

$$\mathbb{E}\int_0^T h_i^2(s)ds < \infty, \quad i = 1,2 \tag{3}$$

one defines the stochastic integral

$$I(h) := \int_0^T h(s)dW(s) := \int_0^T h_1(s)dW_1(s) + \int_0^T h_2(s)dW_2(s)$$

with respect to $(W(t) = (W_1(t), W_2(t))$. Let $g_1(t), g_2(t)$ also satisfy (3) (with h_i replaced by g_i) and define $I(g)$ correspondingly (i.e. g_i replacing h_i). Show that

$$\mathbb{E}I(h)I(g) = \mathbb{E}\int_0^T [h_1(s)g_1(s) + h_2(s)g_2(s)]ds.$$

Hint: Itô's isometry says that

$$\mathbb{E}I^2(f) = \mathbb{E}\int_0^T f_1^2(s) + f_2^2(s)ds \tag{4}$$

for any 2-dimensional integrand f satisfying (3). Now apply (4) with the choice $f = g + h$ and $f = g - h$.

2 Arbitrage Theory in a Model Market

"The most that can be expected from any model is that it can supply a useful approximation to reality: All models are wrong; some models are useful."

– George Box, 1976.[1]

We will assume we have a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a $\mathbb{R}^n$-valued Wiener process $W = (W(t))_{t\in[0,T]} = ((W_1(t), \dots, W_n(t))^T_{t\in[0,T]}$ generating a filtration $(\mathcal{F}_t)_{t\in[0,T]}$. Recall that this means that for each $i = 1, \dots, n$ the process W_i is a real valued Wiener process and for each $j = 1, \dots, n$ the processes W_i and W_j are independent as long as $i \neq j$.

Definition 2.1. *The probability measure $\mathbb{P}$ is called the* real-world measure.

2.1 Model of a Financial Market

We will now consider the following *model* for a simple financial market. The model will consist of a *risk-free asset* denoted S_0 and *risky assets* S_1, S_2, ..., S_m. The risk free asset is modelled as a stochastic process $S_0 = (S_0(t))_{t\in[0,T]}$ given by

$$dS_0(t) = S_0(t)r(t)dt, \quad S_0(0) = 1, \tag{5}$$

where $r = (r(t))_{t\in[0,T]}$ is assumed to be adapted and almost surely integrable.

The risky assets are modelled as stochastic processes $S_i = (S_i(t))_{t\in[0,T]}$ given by

$$dS_i(t) = S_i(t)\mu_i(t)dt + S_i(t)\sum_{j=1}^{n}\sigma_{ij}(t)dW_j(t), \quad S_i(0) > 0, \tag{6}$$

$i = 1, \dots, m$, where $\mu_i = (\mu_i(t))_{t\in[0,T]}$ and $\sigma_{ij} = (\sigma_{ij}(t))_{t\in[0,T]}$ are adapted and such that

$$\int_0^t \mu_i(s)ds + \sum_{j=1}^{n}\int_0^t \sigma_{ij}(s)dW_j(s)$$

is a well defined Itô process for every $i = 1, \dots, m$.

We will assume trading takes place in continuous time, all market participants pay the same price for the assets, fractional and negative holdings of arbitrary size are permissible and our trades do not affect the market price.

Remark 2.2. All these assumptions are violated in practice: time is not continuous, different market participants pay different prices, there is a bid-ask spread, negative holdings (short-selling) is at best expensive and large players of course move the markets.

And moreover our model does not allow jumps in asset prices. So is such a model useful? That depends and what one aims to capture ....

[1]More from George Box: *"Since all models are wrong the scientist cannot obtain a 'correct' one by excessive elaboration. On the contrary following William of Occam he should seek an economical description of natural phenomena. Just as the ability to devise simple but evocative models is the signature of the great scientist so overelaboration and overparameterization is often the mark of mediocrity."*

2.2 Martingale Measures and Arbitrage

Arbitrage in general refers to risk-less profit (in some sense). There are convincing[2] economic arguments as to why markets are *mostly* free of arbitrage. The question we answer in this section is: when is our model free of arbitrage?

To mathematically define arbitrage one has to talk about trading strategies first.

Definition 2.3. *A trading strategy $h = (h_0, h_1, \ldots, h_m)$ with $h_0 = (h_0(t))_{t\in[0,T]}$ and $(h_1, \ldots, h_m) = (h_1(t), \ldots, h_m(t))_{t\in[0,T]}$ are real valued adapted stochastic process representing the "number of units" of the assets $S_0, S_1, \ldots, S_m$.*

The portfolio value at time t is an adapted stochastic process $v(t) = h_0(t)S_0(t) + h_1(t)S_1(t) + \cdots + h_m(t)S_m(t)$.

Clearly if we know $h_1, \ldots, h_m$ and we know v then we can calculate

$$h_0 = \frac{v - h_1S_1 - \cdots - h_mS_m}{S_0}.$$

The trading strategy above might require injections (or withdrawals) of cash at any time - the definition does not forbid that. The only "thing" forbidden is to "look into the future" (because we require the strategy to be adapted).

Definition 2.4. *A trading strategy is called* self-financing *if for any $0 \le t \le t' \le T$*

$$v(t') - v(t) = \int_t^{t'} h_0(s)dS_0(s) + \sum_{i=1}^{m} \int_t^{t'} h_i(s)dS_i(s). \tag{7}$$

How to understand the "self-financing" property? One way is to think about a strategy that is constant in some time interval $[t, t')$. Then the self-financing property reduces to

$$v(t') - v(t) = h_0(t)(S_0(t') - S_0(t) + \sum_{i=1}^{m} h_i(t)(S_i(t') - S_i(t)).$$

That is the change in the value of the portfolio comes precisely from the change in the values of the asset multiplied by the number of each assets we hold.

We can also think about the "fraction of portfolio value invested in a given asset" denoted by u_i. Clearly for $i = 0, 1, \ldots, m$

$$u_i(t) = \frac{h_i(t)S_i(t)}{v(t)} \quad \Leftrightarrow \quad h_i(t) = \frac{u_i(t)v(t)}{S_i}.$$

Moreover[3]

$$\sum_{i=0}^{m} u_i(t) = \sum_{i=0}^{m} \frac{1}{v(t)} h_i(t)S_i(t) = 1.$$

Of course this is not surprising.

[2]The argument goes roughly as follows: if there is an arbitrage opportunity then market participants will trade to exploit it. This trading will move prices and this will result in the arbitrage opportunity disappearing quickly.

[3]Sometimes it is more convenient not to single-out the risk free asset and start with $i = 0$ in the sum.

Writing in the differential notation the self-financing property (7) is

$$
\begin{aligned}
dv(t) &= \sum_{i=0}^{m} h_i(t)dS_i(t) = v(t)\sum_{i=0}^{m} \frac{h_i(t)S_i(t)}{v(t)}\frac{1}{S_i(t)}dS_i(t) \\
&= v(t)\sum_{i=0}^{m} \frac{u_i(t)}{S_i(t)}dS_i(t).
\end{aligned}
$$

Thus the self-financing property is equivalent to

$$
v(t') - v(t) = \sum_{i=0}^{m} \int_t^{t'} v(s)\frac{u_i(s)}{S_i(s)}dS_i(s). \tag{8}
$$

Proposition 2.5. *A trading strategy* (h_0, h) *is self-financing if and only if*

$$
d\left(\frac{v(t)}{S_0(t)}\right) = \sum_{i=1}^{m} h_i(t)d\left(\frac{S_i(t)}{S_0(t)}\right). \tag{9}
$$

Proof. Assume first that (h_0, h) is self-financing. Initially we make note of the fact that due to the Itô product rule

$$
d\left(\frac{S_i(t)}{S_0(t)}\right) = S_i(t)d\left(\frac{1}{S_0(t)}\right) + \frac{1}{S_0(t)}dS_i(t).
$$

Using Itô product rule as well as (7) and the above identity

$$
\begin{aligned}
d\left(\frac{v(t)}{S_0(t)}\right) &= \frac{1}{S_0(t)}dv(t) + v(t)d\left(\frac{1}{S_0(t)}\right) \\
&= \frac{h_0(t)}{S_0(t)}dS_0(t) + \sum_{i=1}^{m}\frac{h_i(t)}{S_0(t)}dS_i(t) \\
&\quad + \left(h_0(t)S_0(t) + \sum_{i=1}^{m} h_i(t)S_i(t)\right)d\left(\frac{1}{S_0(t)}\right) \\
&= \frac{h_0(t)}{S_0(t)}dS_0(t) + h_0(t)S_0(t)d\left(\frac{1}{S_0(t)}\right) \\
&\quad + \sum_{i=1}^{m} h_i(t)\left[\frac{1}{S_0(t)}dS_i(t) + S_i(t)d\left(\frac{1}{S_0(t)}\right)\right] \\
&= h_0(t)r(t)dt - h_0(t)r(t)dt + \sum_{i=1}^{m} h_i(t)d\left(\frac{S_i(t)}{S_0(t)}\right).
\end{aligned}
$$

This completes the proof in one direction. To go in the other direction assume that (9) holds. Our aim is to derive (7). Using the Itô product rule and rearranging and then basically following the same calculation as above but with minor variations:

$$
\begin{aligned}
dv(t) &= S_0(t)d\left(\frac{v(t)}{S_0(t)}\right) - S_0(t)v(t)d\left(\frac{1}{S_0(t)}\right) \\
&= S_0(t)\left[\sum_{i=1}^{m} h_i(t)d\left(\frac{S_i(t)}{S_0(t)}\right) - v(t)d\left(\frac{1}{S_0(t)}\right)\right] \\
&= h_0(t)dS_0(t) + \sum_{i=1}^{m} h_i(t)dS_i(t).
\end{aligned}
$$

□

An interesting question is how do the risky assets behave relative to the risk-free asset. Let $\tilde{S}_i := S_i/S_0$. We calculate, using Itô's formula, for $i = 1, \ldots, m$,

$$d\tilde{S}_i(t) = d\left(\frac{S_i(t)}{S_0(t)}\right) = S_i(t) d\left(\frac{1}{S_0(t)}\right) + \frac{1}{S_0(t)} dS_i(t) + dS_i(t) \cdot d\left(\frac{1}{S_0(t)}\right)$$
$$= -\frac{S_i(t)}{S_0(t)} r(t)dt + \frac{S_i(t)}{S_0(t)} \left(\mu_i(t)dt + \sum_{j=1}^{n} \sigma_{ij}(t) dW_j(t)\right).$$

Thus

$$d\tilde{S}_i(t) = \tilde{S}_i(t) \left(\mu_i(t) - r(t)\right) dt + \tilde{S}_i(t) \sum_{j=1}^{n} \sigma_{ij}(t) dW_j(t). \tag{10}$$

The process $\tilde{S}_i$ will in general not be a local martingale. However if the "drift"[4] term in (10) was zero then such a process would be a local martingale. Of course in some situations we can find a new measure, using Girsanov's theorem under which $\tilde{S}_i$ would be local martingales. This motivates the following definition. First, recall that two measures are called equivalent if $\mathbb{Q}(E) = 0 \iff \mathbb{P}(E) = 0$ for every $E \in \mathcal{F}$.

Definition 2.6 (Local martingale measure / risk-neutral measure)**.** *A measure $\mathbb{Q}$ that is equivalent to $\mathbb{P}$ and such that $\tilde{S}_i$, for all $i = 1, \ldots, m$, are local martingales is called* local martingale measure *or* risk-neutral measure*.*

Proposition 2.7. *Assume that there is an $\mathbb{R}^n$ valued adapted process $(\varphi(t))_{t\in[0,T]}$ such that for all $t \in [0,T]$ we have $\mu(t) - r(t) = \sigma(t)\varphi(t)$. Moreover assume that for*

$$L(t) := \exp\left(-\int_0^t \varphi(s)^T dW(s) - \frac{1}{2}\int_0^t |\varphi(s)|^2 ds\right)$$

we have $\mathbb{E}L(T) = 1$. Then there exists a local martingale measure.

Proof. The proof is a simple application of Girsanov's theorem. Recall that due to (10)

$$d\tilde{S}_i(t) = \tilde{S}_i(t) \left[\left(\mu_i(t) - r(t)\right) dt + \sum_{j=1}^{n} \sigma_{ij}(t) dW_j(t)\right].$$

Under the assumptions of our theorem

$$W^{\mathbb{Q}}(t) := W(t) + \int_0^t \varphi(s) ds$$

is a Wiener process under the measure $\mathbb{Q}$ given by $d\mathbb{Q} = L(T)d\mathbb{P}$. Moreover

$$\sigma(t)dW(t) = \sigma(t)dW^{\mathbb{Q}}(t) - \sigma(t)\varphi(t)dt = \sigma(t)dW^{\mathbb{Q}}(t) + (r(t) - \mu(t))dt.$$

Hence

$$d\tilde{S}_i(t) = \tilde{S}_i(t) \sum_{j=1}^{n} \sigma_{ij}(t) dW_j^{\mathbb{Q}}(t). \tag{11}$$

Noting that stochastic integrals are local martingales for integrands in $\mathcal{S}$ concludes the proof. $\square$

[4]If $dX(t) = b(t)dt + \nu(t)dW(t)$ then we call $b(t)$ the drift term and $\nu(t)$ the diffusion term.

Of course the local martingale measure is not necessarily unique. Consider the following "canonical" examples.

Example 2.8. Consider what we call the Black–Scholes model: W is a real-valued Wiener process, r, μ and $\sigma > 0$ are real constants,

$$dS_0(t) = rS_0(t)dt$$

and

$$dS_1(t) = \mu S_1(t)dt + \sigma S_1(t)dW(t)$$

are the risk-free and risky asset respectively. The "process" φ in the above proposition is $\varphi = \frac{r-\mu}{\sigma}$. Since $\sigma > 0$ this is well defined. We can use the Novikov's condition to check that $\mathbb{E}L(T) = 1$. Indeed

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T |\varphi(t)|^2 dt\right)\right] = \exp\left(\frac{1}{2}\left(\frac{r-\mu}{\sigma}\right)^2 T\right) < \infty.$$

Thus a local martingale measure (or risk-neutral measure) exists.

This can be generalised to m risky assets and n-dimensional Wiener process.

Example 2.9. Assume that σ is a $m \times n$ real (constant) matrix, $\mu \in \mathbb{R}^m$ is a constant column vector and r is a real constant. Let the risk-free asset be given by

$$dS_0(t) = rS_0(t)dt$$

as before. Let the m risky assets be each given by

$$dS_i(t) = S_i(t)\mu_i(t)dt + S_i(t)\sum_{j=1}^{n} \sigma_{ij} dW_j(t).$$

One can call this the Black–Scholes model for m risky assets. We now consider three separate cases:

Case 1: $n > m$ i.e. there are fewer risky assets than components of the Wiener processes. In such a situation the system of equations $\mu - r = \sigma\varphi$ has m equations but n unknowns and will possibly have infinitely many solutions. Indeed take $m = 1$, $n = 2$ and $\sigma_1 \neq 0$, $\sigma_2 \neq 0$.

$$dS_1(t) = S_1(t)\left[\mu dt + \sigma_1 dW_1(t) + \sigma_2 dW_2(t)\right].$$

We wish to solve

$$\mu - r = \sigma_1\varphi_1 + \sigma_2\varphi_2.$$

This is satisfied if

$$\varphi_1 = \frac{\mu - r - \sigma_2\varphi_2}{\sigma_1}.$$

Again we can check Novikov's condition:

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T |\varphi(t)|^2 dt\right)\right] = \exp\left(\frac{1}{2}\left(\frac{\mu - r + \sigma_2\varphi_2}{\sigma_1}\right)^2 T + \frac{1}{2}\varphi_2^2 T\right) < \infty$$

for any $\varphi_2 \in \mathbb{R}$. Thus for each $\varphi_2 \in \mathbb{R}$ we get a measure $\mathbb{Q}^{\varphi_2}$ and under all these (uncountable many) measures $\tilde{S}_1$ is a local martingale.

Case 2: $n < m$ i.e. there are more traded assets than components of the Wiener process. In this situation it may happen that there is no local martingale measure. Take $m = 2, n = 1$ with $\mu_1 = 1, \mu_2 = 2, \sigma = 1$. There is no solution to

$$\begin{aligned} \mu_1 - r &= \sigma\varphi \\ \mu_2 - r &= \sigma\varphi. \end{aligned}$$

Case 3: $m = n$ is the situation when one may get a unique local martingale measure as long as σ^{-1} exists.

Before moving further we make a useful calculation.

Corollary 2.10 (to Proposition 2.7)**.** *If the local martingale measure $\mathbb{Q}$ exists then the dynamics of S_i under $\mathbb{Q}$ are*

$$dS_i(t) = S_i(t)r(t)dt + S_i(t)\sum_{j=1}^{n}\sigma_{ij}(t)dW_j^{\mathbb{Q}}(t). \tag{12}$$

Proof. Recall that $\tilde{S}_i(t) = S_i(t)/S_0(t)$ We now use the Itô product rule and (11) to see that

$$\begin{aligned} dS_i(t) = d(S_0(t)\tilde{S}_i(t)) &= S_0(t)d\tilde{S}_i(t) + \tilde{S}_i(t)dS_0(t) \\ &= S_i(t)\sum_{j=1}^{m}\sigma_{ij}(t)dW_j^{\mathbb{Q}}(t) + S_i(t)r(t)dt. \end{aligned}$$

□

It may be worth solving (12). We now proceed to "guess" the solution. Let $X_i(t) := \ln S_i(t)$. Then we "use" Itô formula.[5] Thus

$$\begin{aligned} dX_i(t) &= \frac{1}{S_i(t)}dS_i(t) - \frac{1}{2}\frac{1}{S_i^2(t)}dS_i(t)dS_i(t) \\ &= \left(r(t) - \frac{1}{2}\sum_{j=1}^{n}\sigma_{ij}^2(t)\right)dt + \sum_{j=1}^{n}\sigma_{ij}(t)dW_j^{\mathbb{Q}}(t). \end{aligned}$$

Hence

$$X_i(T) - X_i(t) = \int_t^T\left(r(s) - \frac{1}{2}\sum_{j=1}^{n}\sigma_{ij}^2(s)\right)dt + \sum_{j=1}^{n}\int_t^T\sigma_{ij}(s)dW_j^{\mathbb{Q}}(s).$$

And so

$$S_i(T) = S_i(t)\exp\left[\int_t^T\left(r(s) - \frac{1}{2}\sum_{j=1}^{n}\sigma_{ij}^2(s)\right)ds + \sum_{j=1}^{n}\int_t^T\sigma_{ij}(s)dW_j^{\mathbb{Q}}(s)\right]. \tag{13}$$

[5] Note that $x \mapsto \ln x$ is not defined for $x = 0$ let alone differentiable. So we cannot really use Itô formula here. Hence we're only guessing.

Here (13) represents our "guess". At this point we can apply Itô formula to the function $x \mapsto \exp(x)$ which is smooth for all x to check that S_i indeed satisfies (12).

We now know enough about trading strategies to properly define arbitrage. There are various mathematical definitions of arbitrage that effectively change how "certain" is the risk-less profit. Our definition is general enough but it is not the most general. The interested reader should see Delbaen and Schachermayer [3] for the definition "no free lunch with vanishing risk".

Definition 2.11. *A self-financing trading strategy forms* arbitrage *if the value of the portfolio corresponding to this strategy satisfies*

$$\mathbb{P}\left[v(T) \geq v(0)S_0(T)\right] = 1 \tag{14}$$

and

$$\mathbb{P}\left[v(T) > v(0)S_0(T)\right] > 0. \tag{15}$$

How to interpret this? Due to (14) we are certain to obtain no less than by investing in the risk-free asset. Due to (15) we have some strictly positive probability of obtaining strictly more than by investing in the risk-free asset. This makes attempting the strategy worthwhile.

So what is the connection between local martingale measures and arbitrage? The following proposition provides a partial answer.

Proposition 2.12. *Assume that a local martingale measure $\mathbb{Q}$ exists in our model and that $\mathbb{E}^{\mathbb{Q}}[S_i(T)^2] < \infty$. If a strategy $u^i = (u^i(t))_{t\in[0,T]}$, $i = 0, 1, \ldots, m$, which expresses the fraction of total portfolio value invested in each asset, is self-financing and if there is a constant $K > 0$ such that the value of the associated portfolio satisfies almost surely*

$$v(t) \geq -K(1 + S_1(t) + \cdots + S_m(t)) \quad \forall t \in [0, T], \tag{16}$$

then this strategy does not form arbitrage.

A self-financing trading strategy that does not satisfy (16) (i.e. it has no lower bound on value process) can always be used to construct an arbitrage by following a "doubling strategy". In the context of a roulette this is the strategy that bets one pound on black. If the outcome of a spin is black we get one pound. If the outcome is red we double our bet and bet black again. And so on. In theory we are "certain" to make risk-free profit. In practice we will either run out of money or hit the casino limit on bets. So in practice the strategy doesn't work.

This is very much the same in real-world in the sense that one cannot expect to implement such doubling strategies which require unbounded borrowing from a bank.

Proof of Proposition 2.12. For simplicity we only consider the situation when $r(t) = 0$. Thus we have $S_0 = 1$, $dS_0(t) = 0$ and $\tilde{S}_i = S_i$. Let us fix a self-financing strategy such that (16) holds and moreover assume that $v(T) \geq v(0)$ $\mathbb{P}$-almost surely. Such strategy will be arbitrage if $\mathbb{P}(v(T) > v(0)S_0(T)) > 0$. Our aim is to show that this cannot be the case. Since the strategy is self-financing we know from (8) and from $dS_0(t) = 0$ that

$$dv(t) = v(t)\left(\sum_{i=0}^m u_i(t)\frac{1}{S_i(t)}dS_i(t)\right) = v(t)\left(\sum_{i=1}^m u_i(t)\frac{1}{\tilde{S}_i(t)}d\tilde{S}_i(t)\right).$$

By assumption $\tilde{S}_i$ are local martingales under the measure $\mathbb{Q}$ and hence v is also a local martingale.

This means there is a sequence of stopping times $(\tau_k)_{k\in\mathbb{N}}$ such that $\tau_k \to \infty$ as $k \to \infty$ and $(v(t \wedge \tau_k))_{t\in[0,T]}$ are $\mathbb{Q}$ martingales.

We are assuming that $\mathbb{E}^{\mathbb{Q}}[S_i(T)^2] < \infty$. Thus Doob's martingale inequality implies

$$\mathbb{E}^{\mathbb{Q}}\left[\sup_{t\in[0,T]} S_i(t)^2\right] \leq 4\mathbb{E}^{\mathbb{Q}}\left[S_i(T)^2\right] < \infty.$$

This implies that $K\left(1 + \sum_{i=1}^m \sup_{t\in[0,T]} S_i(t)\right)$ is integrable. From (16) we moreover get

$$v(T \wedge \tau_k) + K\left(1 + \sum_{i=1}^m \sup_{t\in[0,T]} S_i(t)\right) \geq 0.$$

Thus we may apply Fatou's lemma, see Lemma A.1, and, observing also that $v(T \wedge \tau_k) \to v(T)$ $\mathbb{Q}$−a.s. as $k \to \infty$, we get

$$\begin{aligned}
&\mathbb{E}^{\mathbb{Q}}\left[v(T) + K\left(1 + \sum_{i=1}^m \sup_{t\in[0,T]} S_i(t)\right)\right] \\
&\quad\leq \liminf_{k\to\infty}\left\{\mathbb{E}^{\mathbb{Q}}[v(T\wedge\tau_k)] + \mathbb{E}^{Q}\left[K\left(1 + \sum_{i=1}^m \sup_{t\in[0,T]} S_i(t)\right)\right]\right\}.
\end{aligned}$$

Hence

$$\mathbb{E}^{\mathbb{Q}}[v(T)] \leq \liminf_{k\to\infty} \mathbb{E}^{\mathbb{Q}}[v(T\wedge\tau_k)].$$

But $(v(t \wedge \tau_k))_{t\in[0,T]}$ are $\mathbb{Q}$-martingales and so $\mathbb{E}^{\mathbb{Q}}\left[v(T\wedge\tau_k)\right] = v(0)$. Thus

$$\mathbb{E}^{\mathbb{Q}}[v(T)] \leq v(0). \tag{17}$$

The measures $\mathbb{Q}$ and $\mathbb{P}$ are equivalent and thus $v(T) \geq v(0)$ $\mathbb{P}$-a.s. is also true $\mathbb{Q}$-a.s.. But this and (17) can only hold if $\mathbb{P}(v(T) > v(0)) = 0$ and this means that such self-financing strategy is not arbitrage. ☐

In fact this is only a part of a more complicated story which is summarised in the following meta theorem[6].

Meta Theorem 2.13 (1st Fundamental Theorem of Asset Pricing)**.** *The model is arbitrage free if and only if there is a local martingale measure $\mathbb{Q}$.*

It is possible to state and prove the result in full generality. See Delbaen and Schachermayer [3]. We have proved the implication in one direction.

Lemma 2.14. *Assume there is a self-financing trading strategy h such that the value of the portfolio associated with this trading strategy is*

$$dv(t) = v(t)\rho(t)dt \tag{18}$$

for some adapted process $\rho = (\rho(t))_{t\in[0,T]}$. Then either a.s. $\rho(t) = r(t)$ for all $t \in [0,T]$ or the model admits an arbitrage.

[6]In these notes meta theorem refers to a result which is stated without proper mathematical details. This is partly because it applies in situations beyond the scope of these notes: discrete time models, models with jumps etc.

Proof. Assume the condition that a.s. $\rho(t) = r(t)$ for almost all $t \in [0,T]$ is violated. Let

$$\begin{aligned}
S_+ &:= \{(\omega,t) \in \Omega \times [0,T] : \rho(t) > r(t)\},\\
S_- &:= \{(\omega,t) \in \Omega \times [0,T] : \rho(t) < r(t)\},\\
S &:= S_+ \cup S_- .
\end{aligned}$$

Note that this means that

$$S^c = \{(\omega,t) \in \Omega \times [0,T] : \rho(t) = r(t)\}.$$

We will build new portfolio with value process $\bar{v}$ corresponding to strategies $\bar{h}_0, \bar{h}$. We take $\bar{v}(0) > 0$ and construct the new trading strategies as follows:

$$\begin{aligned}
\bar{h}_0(t) &= \mathbb{1}_S(t)\frac{\left(\bar{v} - \sum_i \bar{h}_i S_i\right)(t)}{S_0(t)} + \mathbb{1}_{S^c}(t)\frac{\bar{v}(t)}{S_0(t)},\\
\bar{h}_i(t) &= \left(\mathbb{1}_{S_+}(t)h_i(t) - \mathbb{1}_{S_-}(t)h_i(t)\right)\frac{\bar{v}(t)}{v(t)},\ \ i = 1,\dots,m.
\end{aligned}$$

In English: if v, which is given by the trading strategy h, grows strictly faster than the risk free rate ($\rho > r$) then we invest according to h. If v grows strictly slower than the risk free rate ($\rho < r$) then we do exactly the opposite of h. Otherwise we just hold the risk-free asset. Everything has to be re-scaled to match how much money we have to invest. Since we want $\bar{v}$ self-financing we must have

$$\bar{v}(t) = \bar{v}(0) + \int_0^t \bar{h}_0(s)dS_0(s) + \sum_{i=1}^m \int_0^t \bar{h}_i(s)dS_i(s)\,.$$

On S^c this will be $\bar{v}(0) + \int_0^t \bar{v}(s)r(s)\,ds$. On S_+ this is (ommiting s to save space):

$$\begin{aligned}
&\Big(\bar{v} - \sum_i \bar{h}_i S_i\Big)\frac{1}{S_0}\,dS_0 + \frac{\bar{v}}{v}\sum_i h_i\,dS_i\\
&= \bar{v}\Big(1 - \frac{1}{v}\sum_i h_i S_i\Big)\frac{1}{S_0}\,dS_0 - \frac{\bar{v}}{v}h_0\,dS_0 + \frac{\bar{v}}{v}h_0\,dS_0 + \frac{\bar{v}}{v}\sum_i h_i\,dS_i\\
&= \bar{v}\Big(1 - \frac{1}{v}\sum_i h_i S_i\Big)\frac{1}{S_0}\,dS_0 - \frac{\bar{v}}{v}\Big(v - \sum_i h_i S_i\Big)\frac{1}{S_0}\,dS_0 + \frac{\bar{v}}{v}\,dv\\
&= \frac{\bar{v}}{v}\,dv = \bar{v}\rho\,ds\,,
\end{aligned}$$

where in the 1st equality we used the definition of $\bar{h}$ and added 0 and to get the 2nd equality we used the self-financing property of the strategy (h_0, h) and the fact that we can get h_0 from v and h. On S_- this is (still ommiting s):

$$\begin{aligned}
&\Big(\bar{v} - \sum_i \bar{h}_i S_i\Big)\frac{1}{S_0}\,dS_0 - \frac{\bar{v}}{v}\sum_i h_i\,dS_i\\
&= \bar{v}\Big(1 + \frac{1}{v}\sum_i h_i S_i\Big)\frac{1}{S_0}\,dS_0 + \frac{\bar{v}}{v}h_0\,dS_0 - \frac{\bar{v}}{v}h_0\,dS_0 - \frac{\bar{v}}{v}\sum_i h_i\,dS_i\\
&= \bar{v}\Big(1 + \frac{1}{v}\sum_i h_i S_i\Big)\frac{1}{S_0}\,dS_0 + \frac{\bar{v}}{v}\Big(v - \sum_i h_i S_i\Big)\frac{1}{S_0}\,dS_0 - \frac{\bar{v}}{v}\,dv\\
&= 2\bar{v}\frac{1}{S_0}\,dS_0 - \frac{\bar{v}}{v}\,dv = 2\bar{v}r\,ds - \bar{v}\rho\,ds\,,
\end{aligned}$$

where the logic is as above but note the flip of signs because we were on S_-. Thus we get that almost surely for any $t \in [0,T]$ that

$$\bar{v}(t) = \bar{v}(0) + \int_0^t \bar{v}(s)\mathbb{1}_{S^c}(s)r(s)\,ds + \int_0^t \bar{v}(s)\mathbb{1}_{S_+}(s)\rho(s)\,ds$$
$$+ \int_0^t \bar{v}(s)\mathbb{1}_{S_-}(s)(2r(s) - \rho(s))\,ds$$

since $dS_0(t) = r(t)S_0(t)\,dt$ and $dv(t) = \rho(t)v(t)\,dt$. Next if we define

$$x(t) := r(t)\mathbb{1}_{S^c}(t) + \rho(t)\mathbb{1}_{S_+}(t) + (2r(t) - \rho(t))\mathbb{1}_{S_-}(t)$$

then we see that $x(t) > r(t)$ since in particular in the case we are on S_- we have $x(t) = (2r(t) - \rho(t)) > 2r(t) - r(t) = r(t)$. Moreover

$$\bar{v}(t) = \bar{v}(0) + \int_0^t \bar{v}(s)x(s)ds$$

and so, solving this, we get

$$\bar{v}(T) = \bar{v}(0)\exp\left(\int_0^T x(s)ds\right).$$

But

$$\bar{v}(0)\exp\left(\int_0^T x(s)ds\right) > \bar{v}(0)\exp\left(\int_0^T r(s)ds\right) = \bar{v}(0)S_0(T)$$

due to the way x was build. The strategy $\bar{h}, \bar{h}_i$ is then an arbitrage according to Definition 2.11. □

2.3 Pricing Contingent Claims

We now move toward answering the question: how to price a financial derivative without introducing arbitrage in our model? We know that financial derivatives are products whose payoff "derives" from prices of more basic assets (e.g. stocks, bonds, loans…). We slightly generalise the concept of financial derivative to that of a contingent claim.

Definition 2.15. *A contingent claim is any real-valued, $\mathcal{F}_T$-measurable random variable X such that $\mathbb{E}|X| < \infty$.*[7]

The random variable X is essentially the payoff of a derivative. To be more precise:

Definition 2.16. *A contingent claim X is called simple / derived if there is a function $g : \mathbb{R}^m \to \mathbb{R}$ such that $X = g(S(T))$.*

Example 2.17. First some simple / derived contingent claims:

i) Let $m = 1$ and $g(x) = [x - K]_+$ (or $g(x) = [K - x]_+$). Then X is the payoff of a European call (or put) option.

[7] Then with martingale representation we only get a local martingale. If we wanted a true martingale we would need $\mathbb{E}X^2 < \infty$.

ii) Let $\beta \in \mathbb{R}^m$ be a given constant vector and

$$g(x) = g(x_1, \ldots, x_m) = \left[\sum_{i=1}^{m} \beta_i x_i - K\right]_+ .$$

Then $X = g(S(T)$ is the payoff of a European basket / index call option.

iii) Let $g(x) = g(x_1, x_2) = [x_2 - x_1]_+$. Then $X = g(S_1(T), S_2(T))$ is the payoff of an exchange option.

And now non-simple contingent claims:

i) Asian arithmetic option:

$$X = \left[\frac{1}{k}\sum_{j=1}^{k} S_1(T_j) - K\right]_+ ,$$

where $T_1, \ldots, T_k$ are fixed dates agreed when the option contract starts (together with T and K).

ii) Asian geometric option:

$$X = \left[\left(\prod_{j=1}^{k} S_1(T_j)\right)^{\frac{1}{k}} - K\right]_+ ,$$

where $T_1, \ldots, T_k$ are fixed dates agreed when the option contract starts (together with T and K).

iii) Barrier option: let $m = 1$ and $M(t) := \max_{s \leq t} S(T)$. Take

$$X = [S_T - K]_+ \mathbb{1}_{M_T < B},$$

where $B > 0$ is a "barrier level" agreed when the option contract starts (together with T and K).[8]

iv) Lookback option: take, for example,

$$X = [S_T - \min_{t \leq T} S_t]_+$$

or

$$X = [\max_{t \leq T} S_t - K]_+ .$$

Let us now consider what price should contingent claims "trade" for in our model / market. Assume that there is a local martingale measure $\mathbb{Q}$ for our model (and our model consists of, as before, one risk-free asset S_0 and m risky assets $S_1, \ldots, S_m$). Due to Proposition 2.12 we know that there is no arbitrage.

[8] As you can imagine there are other barrier type options.

Let the process $p = (p(t))_{t\in[0,T]}$ represent the contingent claim price. We know only that $p(T) = X$ but what should $p(t)$ be for $t < T$? Consider now the enlarged model / market:

$$S_0, S_1, \ldots, S_m, p.$$

When is this arbitrage free? If $\tilde{p}(t) := p(t)/S_0(t)$ is not a local martingale then $\mathbb{Q}$ is no longer a local martingale measure for this enlarged model / market and so this will not necessarily be arbitrage free. One way to make sure that $\tilde{p}$ is a martingale is to take

$$\tilde{p}(t) := \mathbb{E}^{\mathbb{Q}}\left[X/S_0(T)|\mathcal{F}_t\right].$$

Note that

$$p(T) = S_0(T)\tilde{p}(T) = S_0(T)\mathbb{E}^{\mathbb{Q}}\left[X/S_0(T)|\mathcal{F}_T\right] = X,$$

since $X/S_0(T)$ is an $\mathcal{F}_T$-random variable. Thus we have shown the following proposition.

Proposition 2.18. *Assume $\mathbb{Q}$ is a local martingale measure for our model. Then the enlarged model $S_0, S_1, \ldots, S_m, p$ is arbitrage free if we take*

$$p(t) := S_0(t)\mathbb{E}^{\mathbb{Q}}\left[\frac{X}{S_0(T)}\middle|\mathcal{F}_t\right]. \tag{19}$$

Moreover $P(T) = X$ as required.

We call p given by the proposition "arbitrage price". Note that this price is not necessarily unique since the local martingale measure is in general not unique![9]

2.4 Complete Markets and Replication

The problem with pure arbitrage pricing is not just that the price is not unique. Another problem is that (19) tells us nothing about how to "hedge" or "replicate" such contingent claim.

Definition 2.19. *A contingent claim X is* replicable *(or sometimes called attainable) if there is a self-financing trading strategy h and an initial portfolio capital $v(0)$ such that $v(T) = X$.*

A market / model is said to be complete *if every contingent claim is attainable.*

Proposition 2.20. *Assume that for each t the inverse $\sigma^{-1}(t)$ of $\sigma(t)$ exists and is $\mathcal{F}_t$-measurable. Let $\varphi(t) := \sigma^{-1}(t)(\mu(t) - r(t))$ and let*

$$L(t) := \exp\left(-\int_0^t \varphi(s)^T dW(s) - \frac{1}{2}\int_0^t |\varphi(s)|^2 ds\right)$$

Assume that $\mathbb{E}L(T) = 1$. Then local martingale measure $\mathbb{Q}$ exists. Moreover any contingent claim is replicable and the value of the replicating portfolio $v(t)$ is equal to $p(t)$ i.e.

$$v(t) = p(t) = S_0(t)\mathbb{E}^{Q}\left[\frac{X}{S_0(T)}\middle|\mathcal{F}_t\right]. \tag{20}$$

Finally, the local martingale measure $\mathbb{Q}$ is unique.

[9] In practice this situation is common and these notes should be expanded to include examples.

Proof. Existence of local martingale measure follows directly from Proposition 2.7. Let us now show that there is a self-financing trading strategy h and an initial portfolio capital $v(0)$ such that $v(T) = X$. By construction the process

$$M(t) := \mathbb{E}^{\mathbb{Q}}\left[\frac{X}{S_0(T)}\middle|\mathcal{F}_t\right]$$

is a $\mathbb{Q}$-martingale. Due to martingale representation theorem there is an unique process ξ such that

$$M(t) = M(0) + \int_0^t \xi(s)^T dW^{\mathbb{Q}}(s).$$

In order to have $v(T) = X$ we can equivalently try to achieve $v(T) = S_0(T)M(T)$ since

$$S_0(T)M(T) = S_0(T)\mathbb{E}^{\mathbb{Q}}\left[\frac{X}{S_0(T)}\middle|\mathcal{F}_T\right] = X$$

as both X and $S_0(T)$ are $\mathcal{F}_T$-measurable. Of course if $v(t) = S_0(t)M(t)$ for all $t \in [0,T]$ then we will get $v(T) = S_0(T)M(T)$. So let us try to have $v(0) = M(0)$ and $dv(t) = d(S_0(t)M(t))$. We note that due to the Itô product rule

$$\begin{aligned} dv(t) &= d(S_0(t)M(t)) = S_0(t)dM(t) + M(t)dS_0(t) \\ &= S_0(t)\xi(t)^T dW^{\mathbb{Q}}(t) + M(t)dS_0(t). \end{aligned} \tag{21}$$

But because the trading strategy we seek must be self-financing we also have

$$dv(t) = h_0(t)dS_0(t) + \sum_{i=1}^{m} h_i(t)dS_i(t).$$

Moreover due to (12) this becomes

$$\begin{aligned} &dv(t) \\ &= h_0(t)dS_0(t) + \sum_{i=1}^{m} h_i(t)\left(S_i(t)r(t)dt + S_i(t)\sum_{j=1}^{n}\sigma_{ij}(t)dW_j^{\mathbb{Q}}(t)\right) \\ &= \left(h_0(t) + \sum_{i=1}^{m} h_i(t)\frac{S_i(t)}{S_0(t)}\right)dS_0(t) + \sum_{i=1}^{m} h_i(t)S_i(t)\sum_{j=1}^{n}\sigma_{ij}(t)dW_j^{\mathbb{Q}}(t). \end{aligned} \tag{22}$$

To ensure that both (21) and (22) hold we need to choose h appropriately. This means taking

$$h_0(t) = M(t) - \sum_{i=1}^{m} h_i(t)\frac{S_i(t)}{S_0(t)}$$

and for each $j = 1, \dots, n$

$$S_0(t)\xi_j(t) = \sum_{i=1}^{m} h_i(t)S_i(t)\sigma_{ij}(t).$$

This is equivalent to

$$\xi_j(t) = \sum_{i=1}^{m} h_i(t)\tilde{S}_i(t)\sigma_{ij}(t) \;\; j = 1, \dots, n. \tag{23}$$

We need to solve this for h and one way to do this is to write this in the matrix form. One can (and should) easily check that (23) is equivalent to

$$
\begin{aligned}
&\xi(t) = \sigma(t)^T \mathrm{diag}(\tilde{S}(t))h(t) \\
\iff &h(t) = \left(\sigma(t)^T \mathrm{diag}(\tilde{S}(t))\right)^{-1} \xi(t) \\
\iff &h(t) = \mathrm{diag}(\tilde{S}(t))^{-1}(\sigma(t)^T)^{-1}\xi(t) \\
\iff &h(t) = \mathrm{diag}(S_0(t)/S_1(t), \ldots, S_0(t)/S_m(t))(\sigma(t)^{-1})^T\xi(t).
\end{aligned}
$$

Since we are assuming that $\sigma^{-1}(t)$ exists for every t and since $\xi(t)$ coming from martingale representation theorem is unique we have $h_0(t)$ and $h(t)$ uniquely determined. Thus the process $v = v(t)_{t\in[0,T]}$ is uniquely determined, $dv(t) = d(S_0(t)M(t))$ and in particular $v(T) = X$ and

$$
v(t) = S_0(t)M(t) = S_0(t)\mathbb{E}^{\mathbb{Q}}\left[\frac{X}{S_0(T)}\middle|\mathcal{F}_t\right].
$$

Hence we have (20).

We must still confirm that the strategy we created is self-financing. But

$$
\begin{aligned}
h(t)^T d\tilde{S}(t) &= \left[\mathrm{diag}(\tilde{S}(t))^{-1}(\sigma(t)^T)^{-1}\xi(t)\right]^T \mathrm{diag}(\tilde{S}(t))\sigma(t)dW^{\mathbb{Q}}(t) \\
&= \xi(t)^T dW^{\mathbb{Q}}(t) = dM(t) = d\left(\frac{v(t)}{S_0(t)}\right).
\end{aligned}
$$

But this, due to Proposition 2.5, means the strategy is self-financing. Hence we have shown that the contingent claim X is replicable.

It remains to show that $\mathbb{Q}$ is unique. Let $\mathbb{Q}'$ be another local martingale measure for our model. Let $A \in \mathcal{F}_T$ and let $X = \mathbb{1}_A S_0(T)$. Then X is a contingent claim. We note that since $X/S_0(T)$ is $\mathcal{F}_T$-measurable we have

$$
\frac{X}{S_0(T)} = \mathbb{E}^{\mathbb{Q}}\left[\frac{X}{S_0(T)}\middle|\mathcal{F}_T\right].
$$

Then

$$
\begin{aligned}
\mathbb{Q}'(A) &= \mathbb{E}^{\mathbb{Q}'}[\mathbb{1}_A] = \mathbb{E}^{\mathbb{Q}'}\left[\frac{X}{S_0(T)}\right] = \mathbb{E}^{\mathbb{Q}'}\left[\mathbb{E}^{\mathbb{Q}'}\left[\frac{X}{S_0(T)}\middle|\mathcal{F}_T\right]\right] = \mathbb{E}^{\mathbb{Q}'}\left[M(T)\right] \\
&= \mathbb{E}^{\mathbb{Q}'}\left[\frac{v(T)}{S_0(T)}\right].
\end{aligned}
$$

But due to (23) we have that $d(v(t)/S_0(t)) = h^T d\tilde{S}(t)$. Moreover $\tilde{S}$ must be not only $\mathbb{Q}$-local martingale but also $\mathbb{Q}'$-local martingale. Thus

$$
\mathbb{E}^{\mathbb{Q}'}\left[\frac{v(T)}{S_0(T)}\right] = v(0) = \mathbb{E}^{\mathbb{Q}}\left[\frac{v(T)}{S_0(T)}\right] = \mathbb{E}^{\mathbb{Q}}\left[\frac{X}{S_0(T)}\right] = \mathbb{Q}(A).
$$

Hence $\mathbb{Q}'(A) = \mathbb{Q}(A)$ for any $A \in \mathcal{F}_T$ which means that $\mathbb{Q} = \mathbb{Q}'$. But that means that $\mathbb{Q}$ is unique. $\square$

We conclude this section with another "meta theorem". Again this is a result that can be shown to hold in a variety of settings (e.g. appropriate discrete space / time models).

Meta Theorem 2.21 (2nd Fundamental Theorem of Asset Pricing)**.** *Assume that the market / model is arbitrage free. Then the local martingale measure is unique if and only if the market is complete.*

We have shown that if the market is complete (any contingent claim replicable) then the local martingale measure is unique. We haven't proved this in the other direction. Instead we assumed the invertibility of σ defining our model but that is not the same thing.

2.5 Replication of Simple Claims in Complete Markets

From Proposition 2.20 we know that a replicating portfolio / hedging strategy exists. If we look in the proof we see that it is given by the formula

$$h(t) = \text{diag}(S_0(t)/S_1(t), \dots, S_0(t)/S_m(t))(\sigma(t)^{-1})^T \xi(t)$$

where ξ is the process given to us by the Martingale Representation Theorem so we know it exists and it is unique. What we do not know is what this process is. This is a nice theoretical result but it will not make your colleague on the trading desk happy. Indeed if she sells a contingent claim using a price given by your model she will not know how the hedge the risk she's exposing herself to! However in the case of simple claims and with a simplified model we can do a lot better.

First assume that there is a function $\sigma : [0,T] \times \mathbb{R}^m \to \mathbb{R}^{m\times m}$ such that $\sigma(t) = \sigma(t, S(t))$. We have to assume that (12) has a unique solution. Moreover we have to specify how we model the process $r = (r(t))_{t\in[0,T]}$. For simplicity assume that there is $r : [0,T] \to \mathbb{R}$ i.e. r is a deterministic function of time. Finally let, for $t \in [0,T]$ and $S \in [0,\infty)^m$,

$$w(t,S) := \mathbb{E}^{\mathbb{Q}}\left[\frac{S_0(t)}{S_0(T)} g(S(T)) \middle| S(t) = S\right].$$

Proposition 2.22. *Let all the assumptions in Proposition 2.20 hold. Assume X is a simple claim i.e. there is $g : \mathbb{R}^m \to \mathbb{R}$ such that $X = g(S(T))$. Then*

$$h_i(t) = \frac{\partial w}{\partial S_i}(t, S(t)).$$

Moreover

$$\frac{\partial w}{\partial t} + \sum_{i,j=1}^{m} a_{ij} \frac{\partial^2 w}{\partial S_i \partial S_j} + \sum_{i=1}^{m} b_i \frac{\partial w}{\partial S_i} - rw = 0 \;\; \textit{in} \;\; [0,T) \times [0,\infty)^m \tag{24}$$

$$w(T,S) = g(S) \;\; \forall S \in [0,\infty)^m. \tag{25}$$

Here

$$a(t,S) := \frac{1}{2} \mathit{diag}(S)\sigma(t,S)\left(\mathit{diag}(S)\sigma(t,S)\right)^T, \; b(t,S) := r(t)S.$$

Thus the proposition gives us the well known Black–Scholes partial differential equation and the usual delta hedging. We now provide a justification for Proposition 2.22. We assume that w as defined above is sufficiently regular so that all the required partial derivatives exist. There is mathematically no justification for this to be the case from what we have seen so far. This is why we are not calling what we are doing a proof. We note that

$$S_0(t) = \exp\left(\int_0^t r(s)ds\right)$$

which means that $d(1/S_0(t)) = -r(t)(1/S_0(t))dt$. Let us write, to simplify notation, $1/S_0(t) =: D(t)$. So $dD(t) = d(1/S_0(t)) = -r(t)D(t)dt$. From Itô's product rule we obtain

$$d(D(t)w(t,S(t)) = D(t)dw(t,S(t)) - w(t,S(t))r(t)D(t)dt.$$

Applying the full multi-dimensional Itô's formula to the function w and the process S and substituting above we obtain

$$\begin{aligned}
&d(D(t)w(t,S(t)) \\
&= D(t)\Bigg[\Bigg(\frac{\partial w}{\partial t}(t,S(t)) + \sum_{i,j} a_{ij}(t,S(t))\frac{\partial^2 w}{\partial S_i \partial S_j}(t,S(t)) \\
&\quad + \sum_i r(t)S_i(t)\frac{\partial w}{\partial S_i}(t,S(t)) - w(t,S(t))r(t)\Bigg)dt \\
&\quad + \sum_i S_i(t)\frac{\partial w}{\partial S_i}(t,S(t))\sum_j \sigma_{ij}dW_j^{\mathbb{Q}}(t)\Bigg].
\end{aligned}$$

But we know, from the proof of Proposition 2.20 that $D(t)w(t,S(t)) = v(t)/S_0(t) = M(t)$ is a martingale and hence

$$d(D(t)w(t,S(t)) = D(t)\sum_i S_i(t)\frac{\partial w}{\partial S_i}(t,S(t))\sum_j \sigma_{ij}(t,S(t))dW_j^{\mathbb{Q}}(t).$$

Thus (by the uniqueness part of Martingale representation)

$$\xi_j(t) = D(t)\sum_i S_i(t)\frac{\partial w}{\partial S_i}(t,S(t))\sigma_{ij}(t,S(t)).$$

Recalling that

$$h(t) = \text{diag}(S_0(t)/S_1(t),\ldots,S_0(t)/S_m(t))(\sigma(t)^{-1})^T\xi(t)$$

we get

$$h_i(t) = \frac{\partial w}{\partial S_i}(t,S(t)).$$

Moreover

$$\begin{aligned}
&\frac{\partial w}{\partial t}(t,S(t)) + \sum_{i,j} a_{ij}(t,S(t))\frac{\partial^2 w}{\partial S_i \partial S_j}(t,S(t)) \\
&\quad + \sum_i r(t)S_i(t)\frac{\partial w}{\partial S_i}(t,S(t)) - w(t,S(t))r(t) = 0.
\end{aligned}$$

Since trivially $w(T,S) = g(S)$ we get (24).

2.6 Complete Market Example: Multi-dimensional Black–Scholes

Consider a model where $\sigma_{ij}(t)$ and $r(t)$ are deterministic functions of t for all t. That is we have deterministic volatility and the risk-free-rate. Recall that due to (13) we have

$$S_i(T) = S_i(t) \exp\left[\int_t^T \left(r(s) - \frac{1}{2}\sum_{j=1}^n \sigma_{ij}^2(s)\right) ds + \sum_{j=1}^n \int_t^T \sigma_{ij}(s) dW_j^{\mathbb{Q}}(s)\right].$$

Assume that the contingent claim $X = g(S(T))$ for some function $g : \mathbb{R}^m \to \mathbb{R}$ (i.e. it is a simple claim). Then due to Proposition 2.20 the contingent claim is replicable and its price is

$$p(t) = S_0(t)\mathbb{E}^Q\left[\frac{g(S(T))}{S_0(T)}\bigg|\mathcal{F}_t\right].$$

Assuming further for simplicity that $r(t) = r$ is constant we get

$$p(t) = e^{-r(T-t)}\mathbb{E}^Q\left[g(S(T))\Big|\mathcal{F}_t\right].$$

We would like to simplify this further by considering the density of $S_i(T)$ under $\mathbb{Q}$. We note that

$$\int_t^T \sigma_{ij}(s) dW_j^{\mathbb{Q}}(s) \stackrel{d}{=} Z_j \sqrt{\int_t^T \sigma_{ij}^2(s) ds},$$

where $\stackrel{d}{=}$ is used to denote equality in distribution and $Z_1, \ldots, Z_n$ are independent standard normal random variables. Let us define the function as $\xi_i(t, z_1, \ldots, z_n)$

$$:= S_i(t) \exp\left[\int_t^T \left(r(s) - \frac{1}{2}\sum_{j=1}^n \sigma_{ij}^2(s)\right) ds + \sum_{j=1}^n z_j \sqrt{\int_t^T \sigma_{ij}^2(s) ds}\right].$$

We know that for independent random variables their joint density is just the product of their marginal densities. Thus

$$p(t) = e^{-r(T-t)} \int_{\mathbb{R}^m} g\left(\xi_1(t,z), \ldots, \xi_m(z)\right) \phi(z) dz, \tag{26}$$

where

$$\begin{aligned}\phi(z) = \phi(z_1, \ldots, z_n) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z_1^2} \cdots \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z_n^2} \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}|z|^2}.\end{aligned}$$

Note that this is just the density of a multivariate normal distribution with mean zero and identity covariance, see (50).

We conclude that in this situation (multi-asset Black–Scholes model) the price of any European-type contingent claim can be computed as a multi-dimensional integral with respect to multi-variate normal density.

Example 2.23. If $m = n = 1$, $\sigma(t) = \sigma$ is a constant then we can recover the Black–Scholes formula for puts and calls.

2.7 Complete Market Example: Option on an Asset in Foreign Currency

Fix $T > 0$ and consider the following foreign-exchange (FX) model. In the model there is a domestic zero-coupon bond (ZCB) maturing at T with price $p_D(t,T)$ at t, a foreign ZCB maturing at T with price $p_F(t,T)$ at t, a foreign risky asset with price $S(t)$ in the foreign currency at time t and finally an exchange rate allowing us to buy one unit of foreign currency at time t for $f(t)$ units of domestic currency.

With real constants r_D, r_F we have the prices of the two ZCBs given by

$$\begin{aligned} p_D(t,T) &= e^{-r_D(T-t)}, \\ p_F(t,T) &= e^{-r_F(T-t)}. \end{aligned}$$

We assume that in the real-world measure $\mathbb{P}$ the exchange rate and the risky asset have the following dynamics (with real constants γ, μ, $\rho \in (-1,1)$, $\sigma_f \neq 0$ and $\sigma_S \neq 0$):

$$\begin{aligned} df &= f\gamma dt + f\sigma_f dW_1 \\ dS &= S\mu dt + S\sigma_S \left(\rho dW_1 + \sqrt{1-\rho^2} dW_2\right), \end{aligned}$$

where W_1 and W_2 are two independent Wiener processes.

We now wish to consider this in our framework. Thus we have to figure out the following:

- Taking the domestic ZCB as a risk-free asset, what are the two risky assets?
- What are the dynamics of the two risky assets under $\mathbb{P}$?
- Is the model arbitrage-free and complete with the risky assets we have chosen?

Once we have answered those we can move to pricing options. To do that we need to know:

- What is the unique local martingale (risk-neutral) measure $\mathbb{Q}$ and $d\mathbb{Q}/d\mathbb{P}$?
- What are the dynamics of f and of the two risky assets under $\mathbb{Q}$?
- What is the price (in domestic currency) of an European call / put on the foreign asset with a strike K given in domestic currency?

The two risky assets are the domestic value of the foreign risky asset: $X(t) = f(t)S(t)$ and the domestic value of the foreign ZCB: $Y(t) = f(t)p_F(t,T)$.

Their dynamics are:

$$dY = Y(r_F dt + \gamma dt + \sigma_f dW_1)$$

and

$$\begin{aligned} dX &= f dS + S df + df \cdot dS \\ &= X\left((\mu + \gamma + \sigma_f\sigma_S\rho)dt + (\rho\sigma_S + \sigma_f)dW_1 + \sigma_S\sqrt{1-\rho^2}dW_2\right). \end{aligned}$$

We have two risky assets and $W = (W_1, W_2)^T$ is a 2-dimensional Wiener process so the model is arbitrage free and complete as long as

$$\begin{pmatrix} \sigma_f & 0 \\ \rho\sigma_s + \sigma_f & \sigma_S\sqrt{1-\rho^2} \end{pmatrix}$$

is invertible. This is indeed satisfied since we are assuming $\sigma_f \neq 0$, $\sigma_S \neq 0$ and $\rho \in (-1, 1)$ which implies that $\sigma_f\sigma_S\sqrt{1-\rho^2} \neq 0$. So the required unique $\mathbb{Q}$ exists due to Proposition 2.20.

To find $d\mathbb{Q}/d\mathbb{P}$ we need first the dynamics of the "risky-assets discounted using the risk-free asset" i.e. of $\tilde{X} = X/p_D(\cdot, T)$ and $\tilde{Y} = Y/p_D(\cdot, T)$. We calculate

$$d\tilde{Y} = -r_D Y e^{r_D(T-t)}dt + e^{r_D(T-t)}dY = \tilde{Y}\left((r_F - r_D + \gamma)dt + \sigma_f dW_1\right)$$

and

$$\begin{aligned} d\tilde{X} &= -r_D X e^{r_D(T-t)}dt + e^{r_D(T-t)}dX \\ &= \tilde{X}\left((\mu + \gamma + \sigma_f\sigma_S\rho - r_D)dt + (\rho\sigma_S + \sigma_f)dW_1 + \sigma_S\sqrt{1-\rho^2}dW_2\right). \end{aligned}$$

For these to be local martingales we need $\xi = (\xi_1, \xi_2)^T$ such that

$$\sigma_f\xi_1 + \gamma + r_F - r_D = 0 \;\text{ i.e. }\; \xi_1 = \sigma_f^{-1}(r_D - r_F - \gamma)$$

and

$$(\rho\sigma_S + \sigma_f)\xi_1 + \sigma_S\sqrt{1-\rho^2}\xi_2 + \mu + \gamma + \sigma_f\sigma_S\rho - r_D = 0$$

Then

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(\int_0^T \xi^T dW(s) - \frac{1}{2}\int_0^T |\xi|^2 dt\right).$$

Since the change of measure process is a constant we can check that $\mathbb{E}[d\mathbb{Q}/d\mathbb{P}] = 1$ and so by Girsanov's theorem $W^{\mathbb{Q}}(t) = W(t) - \int_0^t \xi ds$ is a $\mathbb{Q}$-Wiener process. Then

$$\begin{aligned} df &= f(r_D - r_F)dt + \sigma_f f dW_1^{\mathbb{Q}} \\ dY &= r_D Y dt + \sigma_f Y dW_1^{\mathbb{Q}} \\ dX &= r_D X dt + \left((\rho\sigma_S + \sigma_f)dW_1^{\mathbb{Q}} + \sigma_S\sqrt{1-\rho^2}dW_2^{\mathbb{Q}}\right). \end{aligned}$$

Let $g(x) = [x - K]_+$ for a call option and $g(f, S) = [K - x]_+$ for a put option. At time 0 the option price is

$$p = \mathbb{E}^{\mathbb{Q}}\left[e^{-r_D T} g(f(T)\, S(T))\right].$$

Let $\sigma := \sqrt{(\rho\sigma_S + \sigma_f)^2 + \sigma_S^2(1-\rho^2)}$. Let

$$Z^{\mathbb{Q}} = \frac{\rho\sigma_S + \sigma_f}{\sigma}W_1^{\mathbb{Q}} + \frac{\sigma_S\sqrt{1-\rho^2}}{\sigma}W_2^{\mathbb{Q}}.$$

We see that this is a continuous martingale starting from 0 with $dZ^{\mathbb{Q}} \cdot dZ^{\mathbb{Q}} = dt$. By the Lévy characterisation of Wiener processes we know that this must be a $\mathbb{Q}$-Wiener process. Moreover

$$dX = r_D X dt + \sigma X dZ^{\mathbb{Q}}$$

and
$$p = \mathbb{E}^{\mathbb{Q}}\left[e^{-r_D T} g(X(T))\right].$$
Thus the option price can be calculated using Black–Scholes formula with risk-free rate taken as r_D, volatility takes as
$$\sigma = \sqrt{(\rho\sigma_S + \sigma_f)^2 + \sigma_S^2(1-\rho^2)}$$
and the initial asset price $f(0)S(0)$.

2.8 Exercises

Exercise 2.1. i) Solve $dS_0(t) = r(t)S_0(t)\,dt$ for $t \geq 0$ with $S_0(0) = 1$.

ii) Are the trajectories of S_0 continuous? Why?

iii) Calculate $d(1/S_0(t))$.

Exercise 2.2. Assume that $\mu \in \mathcal{A}$ and $\sigma \in \mathcal{S}$. Let W be a real-valued Wiener martingale.

i) Solve
$$dS(t) = S(t)\left[\mu(t)dt + \sigma(t)dW(t)\right], \quad S(0) = s. \tag{27}$$
Hint: Solve this first in the case that μ and σ are real constants. You can either “guess the solution”, or you can use Itô formula to help you find the “guess” by applying it to the process S and the function $x \mapsto \ln x$. Mathematically, this is no better than guessing since the function is non-differentiable at 0. Now use Itô formula correctly to check that your guess is correct.

ii) Is the function $t \mapsto S(t)$ continuous? Why?

iii) Calculate $d(1/S(t))$, assuming $s \neq 0$.

iv) With S_0 given by $dS_0(t) = r(t)S_0(t)\,dt$ calculate $d(S(t)/S_0(t))$.

v) Solve
$$dS_i(t) = S_i(t)\mu_i(t)dt + S_i(t)\sum_{j=1}^{n}\sigma_{ij}(t)dW_j(t), \quad S_i(0) > 0,$$

vi) Are the trajectories of S_i continuous? Why?

Exercise 2.3. Recall that m denotes the number of traded assets while n denotes the number of driving Wiener processes.

If $m > n$ then there are (generally) arbitrage opportunities. To see this consider $m = 2$, $n = 1$ (two assets, one driving Wiener process) such that
$$\begin{aligned} dS_1(t) &= \alpha_1 S_1(t)dt + \sigma_1 S_1(t)dW(t), \\ dS_2(t) &= \alpha_2 S_2(t)dt + \sigma_2 S_2(t)dW(t), \end{aligned}$$
with $\alpha_1, \alpha_2 \in \mathbb{R}$, $\sigma_1, \sigma_2 > 0$ constants. We assume that the riskless rate of interest is $r = 0$. Show that there is arbitrage whenever $\alpha_1 - (\sigma_1/\sigma_2)\alpha_2 \neq 0$.

Hint. Take the (constant) relative portfolios $u_1 := 1$ and $u_2 := -\sigma_1/\sigma_2$ in the respective assets. Check that in this case

$$dv(t) = v(t)\left(\alpha_1 - \frac{\sigma_1}{\sigma_2}\alpha_2\right)dt$$

which, by the lemma in above, implies that there is arbitrage unless $\alpha_1 - (\sigma_1/\sigma_2)\alpha_2 = 0$.

Exercise 2.4. Consider the Black–Scholes model where the risky asset follows the dynamics

$$dS = \mu S dt + \sigma S dW,\ S(0) = 1,$$

where $\mu \in \mathbb{R}$, $\sigma \neq 0$ are constants and W_t is a $\mathbb{P}$-Brownian motion. The riskless asset has interest rate $r = 0$, i.e. $B(t) = 1, t \geq 0$. Fix a time horizon $T > 0$.

Let us suppose that a new financial asset is introduced in the market with price

$$Z = S^3,\ t \in [0, T]$$

and investors start to trade in this asset, too, at this price.

1. Show that there is a probability $\mathbb{Q} \sim \mathbb{P}$ under which $S(t)$, $t \in [0, T]$ is a martingale. Write down $d\mathbb{Q}/d\mathbb{P}$.
2. Using Itô's formula, write down the dynamics of Z using the $\mathbb{P}$-Brownian motion W. That is, write down a stochastic differential equation that is satisfied by Z.
3. Show that there is a probability measure $\mathbb{Q}' \sim \mathbb{P}$ under which Z_t, $t \in [0, T]$ is a martingale. Write down $d\mathbb{Q}'/d\mathbb{P}$. Notice that $\mathbb{Q}'$ is different from $\mathbb{Q}$ above.
4. Construct an explicit strategy (trading in S_t, Z_t, B_t) whose value process $V(t)$ satisfies $dV(t) = kV(t)dt$ for some $k \neq r = 0$. Conclude that there are explicit arbitrage opportunities in this model.

Exercise 2.5. Let us look at the one-dimensional case, where there is one risk-free asset $B(t) = \exp(rt)$ and one risky asset

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t).$$

The discounted price $\tilde{S}(t) = e^{-rt}S(t)$ satisfies

$$d\tilde{S}(t) = (\mu - r)\tilde{S}(t)dt + \sigma\tilde{S}(t)dW(t).$$

The market price of risk (i.e. excess return per unit variance) is $\varphi := (\mu - r)/\sigma$. Defining

$$d\mathbb{Q}/d\mathbb{P} := \exp\{-\varphi W(T) - (1/2)\varphi^2 T\},$$

we get a measure $\mathbb{Q} \sim \mathbb{P}$ such that $\tilde{S}(t)$, $t \in [0, T]$ is a $\mathbb{Q}$-martingale. $\mathbb{Q}$ is the only equivalent probability measure with this property. Under $\mathbb{Q}$ the dynamics of $\tilde{S}$ is

$$d\tilde{S}(t) = \sigma\tilde{S}(t)dW^{\mathbb{Q}}(t),$$

where $W^{\mathbb{Q}}(t) := W(t) + \varphi t$, $t \in [0, T]$ is a Q-Brownian motion.

This leads to the dynamics of the risky asset's price

$$dS(t) = rS(t)dt + \sigma S(t)dW^{\mathbb{Q}}(t),$$

under $\mathbb{Q}$.

We now know that the price at 0 of an option paying G at T is

$$p = p(G) = \mathbb{E}^Q[e^{-rT}G].$$

In the particular case where $G = \Phi(S(T))$ for some function $\Phi : \mathbb{R} \to \mathbb{R}$ we can put this into an integral form:

$$p = \int_{\mathbb{R}} e^{-rT}\Phi(se^{(r-\sigma^2/2)T+\sigma y})f(y)dy.$$

Here $s = S(0)$ is the initial price of the risky asset and

$$f(y) = \frac{e^{-y^2/(2T)}}{\sqrt{2\pi}\sqrt{T}}, y \in \mathbb{R}$$

is a $N(0,T)$ density function. Note that in notation of this course T is always the *variance* and not the *standard deviation* of the normal random variable. (In some books, including Björk's, the notation $N(a,b)$ implies that b is the standard deviation. In other books (and this is the usage here) b refers to variance, i.e. the square of standard deviation.)

1. Determine the price at time 0 of an option with payoff function

   $$\Phi(s) = \ln s,\ s > 0.$$

   (This might have negative value meaning that the option holder pays.)

2. Do likewise for the European call option

   $$\Phi(s) = (s-K)^+,\ s > 0.$$

   This gives the famous Black-Scholes formula. Try to obtain the same formula as in Proposition 7.10 of [1].

3. Consider the risky asset itself as an option. I.e. $\Phi(s) = s$. In this case one should (logically) get that the price at time 0 equals $S(0)$. Check this fact.

4. Calculate the price at time 0 of the European put option

   $$\Phi(s) = (K-s)^+,\ s > 0.$$

   Hint. You can do this directly, OR, use the so-called "put-call parity":

   $$(s-K)^+ - (K-s)^+ = s - K,$$

   and note that it is easy to price options with $\Phi(s) = s$ or $\Phi(s) = K$, see the previous exercise.

5. Calculate the price at time 0 of another digital (also called binary) option, where

$$\Phi(s) = 1_{[a,b]}(s),\ s > 0,$$

where $0 < a < b$ and 1 is the indicator function.

6. Calculate the price at time 0 of a power option, where

$$\Phi(s) = s^{\beta},\ s > 0,$$

where $\beta \in \mathbb{R}$ is a given constant.

7. Consider an option with payoff $G = \Phi(S(T))$ at $T > 0$, where

$$\Phi(x) = x,\ x \leq K, \quad \Phi(x) = 0,\ x > K,$$

for some $K > 0$. Write down a formula, involving $\mathbb{Q}$, for its arbitrage-free price at 0.

Evaluate this price. You should be able to find a formula in terms of $N(\cdot)$ (the standard Gaussian cumulative distribution function) and the parameters K, T, s, r and σ.

Exercise 2.6. Show that if $dS_0 = rS_0(t)dt$ with constant $r \in \mathbb{R}$ and

$$\begin{aligned} dS_1(t) &= \alpha_1 S_1(t)dt + \nu_1 S_1(t)dW_1(t) + \nu_2 S_1(t)dW_2(t), \\ dS_2(t) &= \alpha_2 S_2(t)dt + \nu_3 S_2(t)dW_2(t), \end{aligned}$$

with W_1, W_2 independent Wiener processes and constants $\nu_1, \nu_3 > 0$, $\alpha_1, \alpha_2, \nu_2 \in \mathbb{R}$ then this market model is arbitrage-free and complete.

Exercise 2.7. We considered a two-risky-asset model with dynamics

$$\begin{aligned} dS_1(t)/S_1(t) &= \mu_1 dt + \sigma_{11} dW_1(t) + \sigma_{12} dW_2(t), \\ dS_2(t)/S_2(t) &= \mu_2 dt + \sigma_{21} dW_1(t) + \sigma_{22} dW_2(t). \end{aligned}$$

in the real world measure. The parameters μ_1, μ_2, r and

$$\sigma := \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22}, \end{bmatrix}$$

are given and constant.

i) Calculate σ^{-1}.

ii) We know from a Proposition from lectures that a local martingale measure $\mathbb{Q}$ exists. Write down the formula for the change of drift associated with this change of measure. What is $d\mathbb{Q}/d\mathbb{P}$? This required change of drift is sometimes referred to as "market price of risk".

iii) Write down the dynamics of the $\tilde{S}_i(t) = S_i(t)/S_0(t)$ for $i = 1, 2$. That is, write down the dynamics of the discounted risky assets w.r.t. the $\mathbb{Q}$-Wiener process.

iv) Write down the dynamics of the two risky assets w.r.t. the $\mathbb{Q}$-Wiener process.

Exercise 2.8. Let $(\Omega, \mathcal{F}, \mathbb{P})$ and $T > 0$ be given. Let W be an n-dimensional Wiener process generating the filtration $(\mathcal{F}_t)$ i.e. $\mathcal{F}_t = \sigma\{W_s : s \le t\}$. Consider the following model: one risk-free asset with $S_0(0) = 1$ and $dS_0 = rS_0 dt$ and m risky assets S_i, $i = 1, \dots, m$ such that

$$dS_i(t) = \mu_i S_i(t)dt + \sum_{j=1}^{n} S_i \sigma_{ij} dW_j(t).$$

So far this is the same as in lectures.

Now assume further that each risky asset has *constant dividend yield* q_i. That is: from time s to t the amount of dividends paid by S_i is $\int_s^t q_i S_i(u)du$.

1. Assume that $m = n$ and that σ^{-1} exists. Find a local martingale measure $\mathbb{Q}$ such that the discounted gain processes

$$\tilde{G}_i(t) := \frac{S_i(t)}{S_0(t)} + \int_0^t \frac{1}{S_0(u)} q_i S_i(u)du$$

   are local martingales under $\mathbb{Q}$.

2. Derive an arbitrage free formula for valuing contingent claims of the form $X = g(S(T))$ in terms of integral with respect to (multidimensional) normal density. State carefully what this density is.

3. Assume $m = n = 1$ and derive a (modification of) the Black–Scholes formula for $g(S) = [S - K]_+$ (i.e. the Black–Scholes formula for a call on a dividend paying stock with constant dividend yield).

Exercise 2.9. In the lectures we said that if $m < n$ then the market is generally incomplete, i.e. there are contingent claims X s.t. $\mathbb{E}|X| < \infty$ that are not replicable from any initial capital by any self-replicating trading strategy. Here we provide a concrete example with $m = 1$, $n = 2$: take $W_1(t)$, $W_2(t)$ independent Wiener processes and a risky asset with the dynamics

$$dS(t) = \sigma S(t) dW_1(t)$$

with some constant $\sigma \neq 0$. Take a risk-free asset $S_0(t) = 1$ for all t (i.e. $r = 0$). Let $\mathcal{F}_t, t \ge 0$ be the filtration generated by *both* W_1, W_2. Let the contingent claim be $X := W_2(T)$. We note that this is an $\mathcal{F}_T$ measurable random variable and $\mathbb{E}|X| < \infty$. Show that there is no initial capital $v(0)$ and a self-financing strategy h that will replicate this claim.

Hint. Proceed by contradiction. That is, assume that there actually is an h and $v(0)$ such that

$$X = v(T) = v(0) + \int_0^T h(t)dS(t)$$

and proceed to derive a contradiction. For that use Exercise 1.4 to show that

$$T = \mathbb{E}W_2^2(T) = \mathbb{E}W_2(T)X = \mathbb{E}W_2(T)v(T) = \dots$$

which will give you a contradiction.

3 Numeraire Pairs and Change of Numeraire

"Young man, in mathematics you don't understand things. You just get used to them."

– John von Neumann.

Calculating prices of contingent claims (e.g. options) is sometimes easier when one considers a different measure, equivalent to some local martingale measure (and hence the real world measure). The technique of "change of numeraire" is very powerful method enabling this. Before we can use it we need a little bit of theory.

3.1 Martingales Under Change of Measure

Let $(\Omega, \mathcal{F})$ be a measurable space. Recall that we say that a measure $\mathbb{Q}$ is absolutely continuous with respect to a measure $\mathbb{P}$ if $\mathbb{P}(E) = 0$ implies that $\mathbb{Q}(E) = 0$. We write $\mathbb{Q} << \mathbb{P}$.

Proposition 3.1. *Take two probability measures $\mathbb{P}$ and $\mathbb{Q}$ such that $\mathbb{Q} << \mathbb{P}$ with*

$$d\mathbb{Q} = \Lambda d\mathbb{P}.$$

Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. Then $\mathbb{Q}$ almost surely $\mathbb{E}[\Lambda|\mathcal{G}] > 0$. Moreover for any $\mathcal{F}$-random variable X we have

$$\mathbb{E}^{\mathbb{Q}}[X|\mathcal{G}] = \frac{\mathbb{E}[X\Lambda|\mathcal{G}]}{\mathbb{E}[\Lambda|\mathcal{G}]}. \tag{28}$$

Proof. Let $S := \{\omega : \mathbb{E}[\Lambda|\mathcal{G}](\omega) = 0\}$. Then $S \in \mathcal{G}$ and so by definition of conditional expectation

$$\mathbb{Q}(S) = \int_S d\mathbb{Q} = \int_S \Lambda d\mathbb{P} = \int_S \mathbb{E}[\Lambda|\mathcal{G}] d\mathbb{P} = \int_S 0 d\mathbb{P} = 0.$$

Thus $\mathbb{Q}$-a.s. we have $\mathbb{E}[\Lambda|\mathcal{G}](\omega) > 0$.

To prove the second claim assume first that $X \geq 0$. We note that by definition of conditional expectation, for all $G \in \mathcal{G}$:

$$\int_G \mathbb{E}[X\Lambda|\mathcal{G}] d\mathbb{P} = \int_G X\Lambda d\mathbb{P} = \int_G X d\mathbb{Q} = \int_G \mathbb{E}^{\mathbb{Q}}[X|\mathcal{G}] d\mathbb{Q} = \int_G \mathbb{E}^{\mathbb{Q}}[X|\mathcal{G}]\Lambda d\mathbb{P}.$$

Now we use the definition of conditional expectation to take *another* conditional expectation with respect to $\mathcal{G}$. Since $G \in \mathcal{G}$:

$$\int_G \mathbb{E}^{\mathbb{Q}}[X|\mathcal{G}]\Lambda d\mathbb{P} = \int_G \mathbb{E}\left[\mathbb{E}^{\mathbb{Q}}[X|\mathcal{G}]\Lambda|\mathcal{G}\right] d\mathbb{P}.$$

But $\mathbb{E}^{\mathbb{Q}}[X|\mathcal{G}]$ is $\mathcal{G}$-measurable and so

$$\int_G \mathbb{E}\left[\mathbb{E}^{\mathbb{Q}}[X|\mathcal{G}]\Lambda|\mathcal{G}\right] d\mathbb{P} = \int_G \mathbb{E}^{\mathbb{Q}}[X|\mathcal{G}]\mathbb{E}\left[\Lambda|\mathcal{G}\right] d\mathbb{P}.$$

Thus for all $G \in \mathcal{G}$ we get

$$\int_G \mathbb{E}[X\Lambda|\mathcal{G}]d\mathbb{P} = \int_G \mathbb{E}^{\mathbb{Q}}[X|\mathcal{G}]\mathbb{E}\left[\Lambda|\mathcal{G}\right] d\mathbb{P}.$$

Since $X \geq 0$ (and $\Lambda \geq 0$) this means that $\mathbb{P}$-a.s. and hence $\mathbb{Q}$-a.s. we have (28).

$$\mathbb{E}[X\Lambda|\mathcal{G}] = \mathbb{E}^{\mathbb{Q}}[X|\mathcal{G}]\mathbb{E}\left[\Lambda|\mathcal{G}\right].$$

For a general X write $X = X^+ - X^-$, where $X^+ = \mathbb{1}_{\{X\geq 0\}}X \geq 0$ and $X^- = -\mathbb{1}_{\{X<0\}}X \geq 0$. Then

$$\mathbb{E}^{\mathbb{Q}}[X^+ - X^-|\mathcal{G}] = \frac{\mathbb{E}[X^+\Lambda|\mathcal{G}]}{\mathbb{E}[\Lambda|\mathcal{G}]} - \frac{\mathbb{E}[X^-\Lambda|\mathcal{G}]}{\mathbb{E}[\Lambda|\mathcal{G}]} = \frac{\mathbb{E}[X^+ - X^-\Lambda|\mathcal{G}]}{\mathbb{E}[\Lambda|\mathcal{G}]}.$$

□

Now that we know what happens with conditional expectations we can look at what happens with martingales. Let $(\mathcal{F}_t)_{t\in[0,T]}$ be a filtration such that $\mathcal{F}_T = \mathcal{F}$.

Proposition 3.2. *Let there be two probability measures $\mathbb{P}$ and $\mathbb{Q}$ on $(\Omega, \mathcal{F})$ such that $\mathbb{Q} << \mathbb{P}$ (i.e. $\mathbb{Q}$ absolutely continuous w.r.t. $\mathbb{P}$) and the Radon–Nikodym derivative given by $d\mathbb{Q} = \Lambda d\mathbb{P}$ (so $\Lambda \geq 0$, $\mathbb{E}\Lambda = 1$).*

Define a process $(L(t) := \mathbb{E}[\Lambda|\mathcal{F}_t])_{t\in[0,T]}$. A process Y is a $\mathbb{Q}$-martingale if and only if the process YL is a $\mathbb{P}$-martingale.

Proof. By Proposition 3.1 we have, for any $s \leq t$,

$$\mathbb{E}^{\mathbb{Q}}[Y(t)|\mathcal{F}_s] = \frac{\mathbb{E}[Y(t)\Lambda|\mathcal{F}_s]}{\mathbb{E}[\Lambda|\mathcal{F}_s]}.$$

Using the tower property of conditional expectation

$$\mathbb{E}[Y(t)\Lambda|\mathcal{F}_s] = \mathbb{E}[\mathbb{E}[Y(t)\Lambda|\mathcal{F}_t]|\mathcal{F}_s] = \mathbb{E}[Y(t)L(t)|\mathcal{F}_s].$$

Hence

$$\mathbb{E}^{\mathbb{Q}}[Y(t)|\mathcal{F}_s] = \frac{\mathbb{E}[Y(t)L(t)|\mathcal{F}_s]}{L(s)}.$$

If Y is a $\mathbb{Q}$ martingale then we get

$$Y(s) = \frac{\mathbb{E}[Y(t)L(t)|\mathcal{F}_s]}{L(s)}$$

which means that YL is a $\mathbb{P}$ martingale. On the other hand if we start by assuming that YL is a $\mathbb{P}$ martingale then we get

$$\mathbb{E}^{\mathbb{Q}}[Y(t)|\mathcal{F}_s] = \frac{\mathbb{E}[Y(t)L(t)|\mathcal{F}_s]}{L(s)} = Y(s).$$

□

3.2 Numeraire Pairs

Recall that we write $\mathbb{Q} \sim \mathbb{P}$ if $\mathbb{Q}$ is absolutely continuous w.r.t. $\mathbb{P}$ and vice-versa. In other words the measures are equivalent.

Definition 3.3. *A* numeraire *is any traded asset N satisfying $N(t) > 0$ for all $t \in [0, T]$.*

A numeraire pair *$(N, \mathbb{Q})$ consists of a numeraire N and a measure $\mathbb{Q} \sim \mathbb{P}$ such that for any traded asset S the process S/N is a $\mathbb{Q}$-local martingale.*

For example in Section 2 we had the risk-free asset S_0 and we have shown that in some situations there is at least one measure $\mathbb{Q}$ such that for all traded assets S/S_0 is a $\mathbb{Q}$-local martingale. This makes $(S_0, \mathbb{Q})$ a numeraire pair.

Proposition 3.4. *Let $(N_1, \mathbb{Q}_1)$ be a numeraire pair. Let N_2 be another numeraire such that for $\Lambda := \frac{N_2(T)}{N_1(T)}$ we have $\mathbb{E}^{\mathbb{Q}_1}\Lambda = 1$ (we already have $\Lambda > 0$). Then $\mathbb{Q}_2$ given by $d\mathbb{Q}_2 = \Lambda d\mathbb{Q}_1$ is a probability measure, $\mathbb{Q}_2 \sim \mathbb{Q}_1$ and $(N_2, \mathbb{Q}_2)$ is a numeraire pair.*

Proof. Since N_2 is a traded asset (otherwise it cannot be a numeraire) and since $(N_1, \mathbb{Q}_1)$ is a numeraire pair we know that $L(t) := N_2(t)/N_1(t)$ is a $\mathbb{Q}_1$-local martingale. Hence $\mathbb{E}^{\mathbb{Q}_1}[\Lambda|\mathcal{F}_t] = L(t)$ and we can apply Proposition 3.2. Let X be some traded asset. Let $Y := X/N_2$. Then

$$Y(t)L(t) = \frac{X(t)}{N_1(t)}.$$

Since $(N_1, \mathbb{Q}_1)$ is a numeraire pair YL is a $\mathbb{Q}_1$-local martingale and so, due to the Proposition 3.2, Y is a $\mathbb{Q}_2$-local martingale. $\square$

Finally we get to the interesting part: how does one price a contingent claim under a different measure?

Consider a contingent claim X and assume there is some risk-free asset S_0 and a local martingale measure $\mathbb{Q}_0$ so that its arbitrage-free price $p(t)$ is given by

$$\frac{p(t)}{S_0(t)} = \mathbb{E}^{\mathbb{Q}_0}\left[\frac{X}{S_0(T)}\middle|\mathcal{F}_t\right].$$

The reader may want to have a look at Proposition 2.18 if this expression looks unfamiliar.

Consider another traded asset $N_1 > 0$ such that

$$\mathbb{E}^{\mathbb{Q}_0}\left[\frac{N_1(T)}{S_0(T)}\right] = 1.$$

Define $\mathbb{Q}_1$ by $d\mathbb{Q}_1 = \Lambda d\mathbb{Q}_0$ with $\Lambda := N_1(T)/S_0(T)$. By Proposition 3.4 we get a numeraire pair $(N_1, \mathbb{Q}_1)$. Then, due to Proposition 3.1 we get

$$\begin{aligned}
p(t) &= S_0(t)\mathbb{E}^{\mathbb{Q}_0}\left[\frac{X}{S_0(T)}\frac{N_1(T)}{N_1(T)}\middle|\mathcal{F}_t\right] \\
&= S_0(t)\mathbb{E}^{\mathbb{Q}_0}\left[\frac{X}{N_1(T)}\Lambda\middle|\mathcal{F}_t\right] \\
&= S_0(t)\mathbb{E}^{\mathbb{Q}_1}\left[\frac{X}{N_1(T)}\middle|\mathcal{F}_t\right]\mathbb{E}^{\mathbb{Q}_0}[\Lambda|\mathcal{F}_t].
\end{aligned}$$

But

$$\mathbb{E}^{\mathbb{Q}_0}[\Lambda|\mathcal{F}_t] = \mathbb{E}^{\mathbb{Q}_0}\left[\frac{N_1(T)}{S_0(T)}\middle|\mathcal{F}_t\right] = \frac{N_1(t)}{S_0(t)}.$$

Hence

$$p(t) = N_1(t)\mathbb{E}^{\mathbb{Q}_1}\left[\frac{X}{N_1(T)}\middle|\mathcal{F}_t\right].$$

Thus we have shown the following:

Proposition 3.5. *If $(S_0, \mathbb{Q}_0)$ is a numeraire pair such that the arbitrage-free price of a contingent claim is*

$$\frac{p(t)}{S_0(t)} = \mathbb{E}^{\mathbb{Q}_0}\left[\frac{X}{S_0(T)}\middle|\mathcal{F}_t\right]$$

and if $N_1 > 0$ is another numeraire such that $\mathbb{E}^{\mathbb{Q}_0}[N_1(T)/S_0(T)] = 1$ then

i) the pair $(N_1, \mathbb{Q}_1)$ is a numeraire pair where $\mathbb{Q}_1$ given by

$$\frac{dQ_1}{dQ_0} = \frac{N_1(T)}{S_0(T)},$$

ii) an arbitrage-free price of the contingent claim X in the measure $\mathbb{Q}_1$ is given by

$$\frac{p(t)}{N_1(t)} = \mathbb{E}^{\mathbb{Q}_1}\left[\frac{X}{N_1(T)}\middle|\mathcal{F}_t\right]. \tag{29}$$

3.3 Margrabe formula: Numeraire Change Example

We will use the change of numeraire technique to derive the price of an *exchange option*. This is known as the Margrabe formula. We assume that there is a risk free asset $dS_0 = rS_0dt$ and that we are given an invertible matrix σ. We assume that there are two risky assets S_i, $i = 1, 2$, with dynamics

$$dS_i = rS_idt + \sigma_{i1}S_idW_1^{\mathbb{Q}} + \sigma_{i2}S_idW_2^{\mathbb{Q}},$$

with $W^{\mathbb{Q}}$ a Wiener process in the risk neutral measure $\mathbb{Q}$. The exchange option is a European-type option in that it can only be exercised at maturity time $T > 0$. Its payoff is

$$f(S(T)) = [S_1(T) - S_2(T)]_+.$$

We then know from Proposition 2.18 that the arbitrage-free price of this option is (at time $t = 0$):

$$\frac{p(0)}{S_0(0)} = \mathbb{E}^{\mathbb{Q}}\left[\frac{[S_1(T) - S_2(T)]_+}{S_0(T)}\right].$$

Of course we can approximate this using brute-force Monte-Carlo algorithm or by numerically integrating over the density function of $S(T)$ (2D-integration). Amazingly it turns out that there is a simpler way. Looking at the payoff we notice that in fact

$$f(S(T)) = S_2(T)\left[\frac{S_1(T)}{S_2(T)} - 1\right]_+.$$

If we can come up with an appropriate numeraire which will make the leading $S_2(T)$ "disappear" and if we can somehow simply characterise the distribution of $S_1(T)/S_2(T)$ in the corresponding measure then life would be good.

In light of Proposition 3.5 we choose $N_1(t) := \frac{1}{S_2(0)} S_2(t)$. Since $\mathbb{Q}$ is a martingale measure we know that $S_2(t)/S_0(t)$ is a martingale. Hence

$$\mathbb{E}^{\mathbb{Q}}\left[\frac{N_1(T)}{S_0(T)}\right] = \mathbb{E}^{\mathbb{Q}}\left[\frac{S_2(T)/S_2(0)}{S_0(T)}\right] = \frac{S_2(0)/S_2(0)}{S_0(0)} = 1.$$

Thus, using Proposition 3.5 we have a new measure, call it $\mathbb{Q}_1$ with $(N_1, \mathbb{Q}_1)$ a numeraire pair and

$$\frac{p(0)}{N_1(0)} = \mathbb{E}^{\mathbb{Q}_1}\left[\frac{[S_1(T) - S_2(T)]_+}{N_1(T)}\right].$$

Upon noticing that $N_1(0) = 1$ this leads to

$$p(0) = S_2(0)\mathbb{E}^{\mathbb{Q}_1}\left[\left[\tfrac{S_1(T)}{S_2(T)} - 1\right]_+\right].$$

Let $Y(t) := \frac{S_1(T)}{S_2(T)}$. We see that Y is a $\mathbb{Q}_1$ martingale. Indeed S_1/N_1 must be a martingale since $(N_1, \mathbb{Q}_1)$ is a numeraire pair. But any martingale scaled by a constant is still a martingale and so S_1/S_2 must be a martingale.

Let us calculate the dynamics of Y under $\mathbb{Q}_1$ bearing in mind that we already know it is a martingale. First, using Itô's formula, we get that under $\mathbb{Q}$:

$$\begin{aligned} d\left(\frac{1}{S_2}\right) &= -S_2^{-2} dS_2 + S_2^{-3} dS_2 \cdot dS_2 \\ &= -\frac{1}{S_2} r dt - \frac{1}{S_2}\left(\sigma_{21} dW_1^{\mathbb{Q}} + \sigma_{22} dW_2^{\mathbb{Q}}\right) + \frac{1}{S_2}\left(\sigma_{21}^2 dt + \sigma_{22}^2 dt\right). \end{aligned}$$

Collecting all the terms in front of dt together we can write

$$d\left(\frac{1}{S_2}\right) = (\ldots)dt - \frac{1}{S_2}\left(\sigma_{21} dW_1^{\mathbb{Q}} + \sigma_{22} dW_2^{\mathbb{Q}}\right).$$

Now, using Itô product rule, we get

$$dY = d\left(\frac{S_1}{S_2}\right) = S_1 d\left(\frac{1}{S_2}\right) + \left(\frac{1}{S_2}\right) dS_1 + dS_1 \cdot d\left(\frac{1}{S_2}\right).$$

The great thing is that since we know that Y is $\mathbb{Q}_1$ martingale we do not really care what the terms in front of dt are as they must all be *zero* under $\mathbb{Q}_1$. So collecting all the dt terms together under $(\ldots)dt$ (which is of course not the same as the one in the equation above) we get

$$dY = (\ldots)dt - Y\left(\sigma_{21} dW_1^{\mathbb{Q}} + \sigma_{22} dW_2^{\mathbb{Q}}\right) + Y\left(\sigma_{11} dW_1^{\mathbb{Q}} + \sigma_{12} dW_2^{\mathbb{Q}}\right).$$

And under the measure $\mathbb{Q}_1$ we have

$$dY = Y\left(\sigma_{11} - \sigma_{21}\right) dW_1^{\mathbb{Q}_1} - Y\left(\sigma_{22} - \sigma_{12}\right) dW_2^{\mathbb{Q}_1}.$$

Let X be the process starting at 0 and such that

$$dX = \bar{\sigma}^{-1}\left(\left(\sigma_{11} - \sigma_{21}\right) dW_1^{\mathbb{Q}_1} - \left(\sigma_{22} - \sigma_{12}\right) dW_2^{\mathbb{Q}_1}\right),$$

where

$$\bar{\sigma} := \sqrt{(\sigma_{11} - \sigma_{21})^2 + (\sigma_{22} - \sigma_{12})^2}.$$

We note that X is a continuous martingale starting from 0 and its quadratic variation is

$$dX \cdot dX = \frac{1}{\bar{\sigma}^2} (\sigma_{11} - \sigma_{21})^2 dt + (\sigma_{22} - \sigma_{12})^2 dt = dt.$$

Thus due to the Lévy characterization theorem, see Theorem 1.8, we know that X is in fact a $\mathbb{Q}_1$ Wiener process. But

$$dY = \bar{\sigma} Y dX$$

and

$$p(0) = S_2(0)\mathbb{E}^{\mathbb{Q}_1}\left[[Y(T) - 1]_+\right].$$

The expression $\mathbb{E}^{\mathbb{Q}_1}\left[[Y(T) - 1]_+\right]$ corresponds exactly to the price of an European call option for an asset that has current price $S_1(0)/S_2(0)$, strike 1, volatility $\bar{\sigma}$ and where the risk-free rate is 0.

3.4 Exercises

Exercise 3.1. We wish to prove Proposition 3.2 in the discrete time case. Let $N \in \mathbb{N}$ and let $(\mathcal{F}_n)_{n=0}^N$ be a filtration.

Let there be two probability measures $\mathbb{P}$ and $\mathbb{Q}$ on $(\Omega, \mathcal{F})$ such that $\mathbb{Q} << \mathbb{P}$ (i.e. $\mathbb{Q}$ absolutely continuous w.r.t. $\mathbb{P}$) and the Radon–Nikodym derivative given by $d\mathbb{Q} = \Lambda d\mathbb{P}$ (so $\Lambda \geq 0$, $\mathbb{E}\Lambda = 1$).

Define a process $(L_n := \mathbb{E}[\Lambda|\mathcal{F}_n])_{n=0}^N$. Then process $Y = (Y_n)_{n=1}^N$ is a $\mathbb{Q}$-martingale if and only if the process $YL = (Y_n L_n)_{n=1}^N$ is a $\mathbb{P}$-martingale.

Exercise 3.2. Consider a model with two risky assets (plus riskless with interest rate r),

$$\begin{aligned} dS_1(t) &= S_1(t)[\mu_1 dt + \nu_1 dW_1(t)], \\ dS_2(t) &= S_2(t)[\mu_2 dt + \nu_2 dW_1(t) + \nu_3 dW_2(t)], \end{aligned}$$

with $\nu_1, \nu_3 > 0$ and W_1, W_2 independent Wiener processes under $\mathbb{P}$. The filtration is $\mathcal{F}_t = \sigma(W_1(s), W_2(s), s \leq t)$.

i) Using the (unique) martingale measure find the price of a European call on the second asset, i.e. the payoff function is $\Phi(s_1, s_2) = [s_2 - K]_+$ for some $K > 0$.

ii) Do likewise for the option with payoff

$$\Phi(x, y) = \mathbb{1}_{\{x \geq y\}},$$

(i.e. it pays one dollar if $S_1(T)$ is not larger than $S_2(T)$.

iii) Consider the so-called “indexed option” which has payoff function

$$\Phi(s_1, s_2) = \max\{as_1 - bs_2, 0\}$$

with $a, b > 0$. Determine the price of the option at time 0 which pays $\Phi(S_1(T), S_2(T))$ at time T.

iv) Do likewise for $\Phi(s_1, s_2) = \max\{as_1, bs_2\}$.

v) Choose $\nu_1 = 1, \nu_2 = 7, \nu_3 = 2$, $r = 0$, $T = 1$, $s_1 = s_2 = 1$. Calculate the price at time 0 of an option paying

$$\frac{S_1(T)}{S_2(T)} \text{ if } 1 \leq \frac{S_1(T)}{S_2(T)} \leq 2$$

and

0 otherwise,

at time T. You should be able to find a formula that involves only numbers and $N(\cdot)$, which is the standard Gaussian cumulative distribution function.

4 Interest Rate Derivatives

“There is no harm in being sometimes wrong — especially if one is promptly found out.”

– John Maynard Keynes, 1924.

We now turn to interest rate derivatives. These include bonds, interest rate swaps, bond options, caps (capped interst rate swaps) swaptions (options to enter interest rate swaps) and other exotic derivatives. The market for interest rate derivatives is larger than that for equity, FX (foreign exchange) and commodity derivatives.

Definition 4.1 (Zero-Coupon Bond (ZCB))**.** *A T-maturity zero-coupon bond (ZCB) is an asset that pays one unit of currency at time T. We will use $p(t, T)$ to denote the value of ZCB at time $t \leq T$.*

We see immediately that $p(T, T) = 1$.

Definition 4.2 (Spot rate)**.** *The* simply compounded *spot rate at time t for the time interval $[t, T]$ is the constant rate at which an investment of of $p(t, T)$ units of currency at time t accrues with simple compounding to yield a unit of currency at time T i.e. it is the rate $L(t, T)$ such that*

$$1 = (1 + (T - t)L(t, T))p(t, T).$$

We note that the simply compounded spot rate is also known as the LIBOR rate (London inter-bank offer rate). We see immediately that

$$L(t, T) = \frac{1 - p(t, T)}{(T - t)p(t, T)}.$$

Consider three times: $t < S < T$. A forward rate agreement (FRA) agreed at time t gives the holder the right (and obligation) to exchange at time T (maturity) a payment of $(T-S)L(S,T)$ for $(T-S)K$. The rate K is called the *fixed rate* as it is agreed at time t. The rate $L(S,T)$ is a *floating rate* as it will only be observed at time $S > t$. The time T-value of the FRA is thus

$$(T-S)(K - L(S,T))$$

and using the formula for $L(S,T)$ we get

$$(T-S)K - \frac{1}{p(S,T)} + 1.$$

The amount $(T-S)K + 1$ at time T is worth $p(t,T)((T-S)K+1)$ at time t. The amount $\frac{1}{p(S,T)}$ is worth $p(S,T)\frac{1}{p(S,T)} = 1$ at time S which is in turn worth $p(t,S)$ at time t. Thus the time t value of the FRA is

$$\text{FRA}(t,S,T,K) = p(t,T)(T-S)K + p(t,T) - p(t,S).$$

There is one value K which makes this FRA "fair" (i.e. t-value of zero) and this is

$$K = \frac{1}{T-S}\frac{p(t,S)-p(t,T)}{p(t,T)} = \frac{1}{T-S}\left(\frac{p(t,S)}{p(t,T)} - 1\right).$$

This motivates the following definition:

Definition 4.3 (Simply compounded forward rate)**.** *The* simply compounded *forward rate observed at time t for the interval $[S,T]$ denoted by $L(t,S,T)$ is defined by*

$$L(t,S,T) = \frac{1}{T-S}\left(\frac{p(t,S)}{p(t,T)} - 1\right).$$

Remark 4.4. Of course we have the price of FRA in terms of the simply compounded forward rate: since

$$L(t,S,T)(T-S)p(t,T) = p(t,S) - p(t,T)$$

we get

$$\text{FRA}(t,S,T,K) = p(t,T)(T-S)\left[K - L(t,S,T)\right].$$

As $S \nearrow T$ we get the instantaneous forward rate:

Definition 4.5. *If the function $T \mapsto p(t,T)$ is continuously differentiable for any $t \in [0,T]$ the we define the* instantaneous forward rate *for time T, observed at $t < T$, as*

$$f(t,T) := \lim_{S \nearrow T} L(t,S,T).$$

We can immediately calculate

$$\begin{aligned} f(t,T) &= \lim_{S\nearrow T}\frac{1}{T-S}\left(\frac{p(t,S)}{p(t,T)} - 1\right) = \frac{1}{p(t,T)}\lim_{S\nearrow T}\frac{p(t,S)-p(t,T)}{T-S} \\ &= \frac{-1}{p(t,T)}\partial_T p(t,T) = -\partial_T \ln p(t,T). \end{aligned}$$

That is

$$f(t,T) = -\partial_T \ln p(t,T). \tag{30}$$

It is important to note that the derivative is with respect to T (rather than t). In general the function $T \mapsto p(t,T)$ can be assumed to be differentiable. The map $t \mapsto p(t,T)$ cannot be assumed to be differentiable.

Remark 4.6. Due to (30) we have

$$\int_t^T f(t,S)dS = -\int_t^T \frac{\partial}{\partial S} \ln p(t,S)dS = -\ln p(t,T) + \ln p(t,t)$$

and hence

$$\exp\left(-\int_t^T f(t,S)dS\right) = p(t,T).$$

Finally we define the *short rate*.

Definition 4.7 (Short rate)**.** *The* short rate *at time t is* $r(t) := f(t,t)$.

What we have described so far is an idealised description of the interest rate world. In fact ZCBs do *not* trade for arbitrary maturities. However one can obtain the "market price of ZCBs" for all maturities by a "stripping" those from observed forward rate agreements and interest rate swap contracts by making an assumption on how to interpolate the missing forward rates (e.g. piecewise constant, piecewise linear and continuous, splines etc.). See Filipovic [4, Chapter 3] for an excellent treatment of this issue.

4.1 Some Short Rate Models

From now on we make the following assumption:

Assumption 4.8. *There is a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*, a Wiener process* W *and a filtration* $(\mathcal{F}_t)$ *generated by* W*.*

There is a risk-free asset (bank account) with evolution given by

$$dB(t) = r(t)B(t)dt$$

and there are no risky assets.

With this assumption we trivially get existence of a martingale measure $\mathbb{Q}$. This martingale measure is clearly not unique. Nevertheless we fix this risk-neutral measure from now on. We will write $\mathbb{E}$ to mean $\mathbb{E}^{\mathbb{Q}}$. Due to Proposition 2.18 we have that the arbitrage-free price $v(t)$ of any $\mathcal{F}_T$-contingent claim X is

$$v(t) = B(t)\mathbb{E}\left[\frac{X}{B(T)}\middle|\mathcal{F}_t\right]. \tag{31}$$

Since $B(t)$ is $\mathcal{F}_t$ measurable we also have

$$v(t) = \mathbb{E}\left[e^{-\int_t^T r(s)ds} X\middle|\mathcal{F}_t\right]. \tag{32}$$

An important special case is that of $X = 1$ i.e. that of the ZCB:

$$p(t,T) = B(t)\mathbb{E}\left[\frac{1}{B(T)}\middle|\mathcal{F}_t\right] = \mathbb{E}\left[e^{-\int_t^T r(s)ds}\middle|\mathcal{F}_t\right].$$

We see from (32) that to price any interest rate derivative it is sufficient to have a model for the short rate $(r(t))_{t\geq 0}$.

There are many, see e.g. Brigo and Mercurio [2]. We will look at a few. We start by considering a basic short rate model: the Vasicek model. As we will later see this model has one major flaw.

Vasicek model

Assume that the short rate satisfies

$$dr(t) = (b - ar(t))dt + \sigma dW(t). \tag{33}$$

Here $W = (W(t))_{t\geq 0}$ is a Wiener process in some risk neutral measure $\mathbb{Q}$, the parameters $b \in \mathbb{R}$, $a > 0$ $\sigma > 0$ are constant.

We can collect the properties of the short rate under Vasicek model as follows (see Exercise 4.1):

1. Mean reverting to $\frac{b}{a}$.
2. Normally distributed with known mean and variance.
3. $\mathbb{P}(r(t) < 0) > 0$. Old books and papers on interest rate models will list as a disadvantage but in fact this is a desirable property.
4. It is possible to find simple formulae for the ZCB prices and prices of options on ZCBs. We will do this later.

Imagine now that one obtains the ZCB curve $T \mapsto p^*(0,T)$ from traded market instruments. It is immediately clear that the ZCB prices implied from the Vasicek model will not be able to match all such observed curves simply because the Vasicek model has at most *four* parameters whereas the observed ZCB curve is infinite dimensional.

Cox–Ingresoll-Ross (CIR) model

The short rate is assumed to be given by the SDE

$$dr(t) = (b - ar(t))dt + \sigma\sqrt{r(t)}dW(t).$$

We will not study this model in detail but its properties are:

1. Again mean reverting to $\frac{b}{a}$.
2. Known distribution of short rate, namely non-central χ^2-distribution.
3. $\mathbb{P}(r(t) < 0|r(0) \geq 0) = 0$. For interest rates this is not useful but if one uses CIR to model other things (default intensity, volatility) then this becomes crucial.
4. It is possible to find simple (but more complicated than Vasicek) formulae for the ZCB prices and prices of options on ZCBs.

Ho–Lee Model

The short rate is assumed to be given by the SDE

$$dr(t) = \theta(t)dt + \sigma dW(t),$$

where $t \mapsto \theta(t)$ is some deterministic, integrable function that can be chosen so that $p(0,T) = p^*(0,T)$ for all $T > 0$ i.e. the prices of ZCBs given by the model and by the market match. We see immediately that the distribution of rates implied by the model is Gaussian.

Hull–White Model

This is an extension of Vasicek model allowing full calibration to ZCB prices observed in the market:

$$dr(t) = (\theta(t) - ar(t))dt + \sigma dW(t).$$

We can solve this SDE using same technique as before to obtain

$$r(t) = e^{-at}r(0) + \int_0^t e^{-a(t-s)}\theta(s)\,ds + \int_0^t e^{-a(t-s)}\sigma dW(s).$$

The properties of the short rate under Hull-White model are similar to Vasicek:

1. Mean reverting to $\frac{\theta(t)}{a}$.
2. Normally distributed with known mean and variance.
3. $\mathbb{P}(r(t) < 0) > 0$.
4. It is possible to find simple formulae for the ZCB prices and prices of options on ZCBs.
5. It is possible to choose $t \mapsto \theta(t)$ such that the model implied ZCB prices match the market. We will do this later.

The CIR and Hull–White models (or slight generalisations) are popular in the industry. It is also possible to devise "multi-factor" models where the driving Wiener process is more than one-dimensional. They have greater flexibility at the price of having less simple formulae for ZCB price and option prices. However as long as the models have "affine term structure" the ZCB prices are given in terms of solutions to ordinary differential equations as we will see shortly.

4.2 Affine Term Structure Short Rate Models

Deriving the ZCB price for each short rate model would be rather tedious. However some of the models share common features and a lot of the work of the derivation can be carried in a common framework of *affine term structure* models.

Definition 4.9. *We say a short rate model has* affine term structure *if the price of a T-ZCB at time t can be expressed as*

$$p(t,T) = \exp\left(A(t,T) - B(t,T)r(t)\right), \tag{34}$$

for some deterministic functions A and B.

The motivation for this definition is that

$$\ln p(t,T) = A(t,T) - B(t,T)r(t)$$

is an affine function of the short rate.

Theorem 4.10. *Consider a short rate given by*

$$dr(t) = \mu(t, r(t))dt + \sigma(t, r(t))dW(t). \tag{35}$$

If this SDE has a unique strong solution and if

$$\begin{aligned} \mu(t,r) &= \alpha(t)r + \beta(t), \\ \sigma(t,r) &= \sqrt{\gamma(t)r + \delta(t)} \end{aligned} \tag{36}$$

for some continuous functions $\alpha, \beta, \gamma, \delta$ then the model has affine term structure. Moreover A and B satisfy

$$\begin{aligned} \partial_t B(t,T) &= -\alpha(t)B(t,T) + \frac{1}{2}\gamma(t)B^2(t,T) - 1, \\ \partial_t A(t,T) &= \beta(t)B(t,T) - \frac{1}{2}\delta(t)B^2(t,T), \\ B(T,T) &= 0, \quad A(T,T) = 0. \end{aligned} \tag{37}$$

Checking whether our model has the form (36) is straightforward. To solve the first equation in (37) we note that this is a Ricatti ordinary differential equation (ODE) since T plays a role of a fixed parameter. Once $t \mapsto B(t,T)$ is known solving the equation for A requires only integration. Hence for affine term structure models we know how to calculate ZCB prices.

Before we prove Theorem 4.10 we will need the following result.

Proposition 4.11. *Assume that $(r(t))$ is given uniquely by* (35) *for all $t \geq 0$. Let*

$$P(t,r) := \mathbb{E}\left[e^{-\int_t^T r(u)du} \middle| r(t) = r \right].$$

i.e. $P(t,r)$ is the price of T-ZCB if at time t the short rate is r. Then

$$\begin{aligned} \partial_t P + \mu \partial_r P + \frac{1}{2}\sigma^2 \partial_r^2 P - rP = 0 \quad &on \quad [0,T) \times \mathbb{R}, \\ P(T,r) = 1 \quad &\forall r \in \mathbb{R}. \end{aligned} \tag{38}$$

We now provide a heuristic argument why this is the case. This is not a proof because we will not show that P has one derivative in t and two derivatives in r which is what we need.

Let

$$\tilde{P}(t, r(t)) := B^{-1}(t)P(t, r(t)).$$

Note that $P(t, r(t)) = p(t,T)$. Moreover

$$d(B^{-1}(t)) = d\left(e^{-\int_0^t r(u)du}\right) = -r(t)e^{-\int_0^t r(u)du}dt = -r(t)B^{-1}(t)dt.$$

Hence, using Itô's product rule and then Itô's formula, we get

$$
\begin{aligned}
&d\tilde{P}(t,r(t)) = B^{-1}(t)dP(t,r(t)) + P(t,r(t))d(B^{-1}(t)) \\
&= B^{-1}(t)\left[dP(t,r(t)) - r(t)P(t,r(t))dt\right] \\
&= B^{-1}(t)\bigg[\partial_t P(t,r(t)dt + \partial_r P(t,r(t))\,dr(t) + \frac{1}{2}\partial_r^2(t,r(t))dr(t)\cdot dr(t) \\
&\qquad\qquad - r(t)P(t,r(t))dt\bigg] \\
&= B^{-1}(t)\bigg(\partial_t + \mu(t,r(t))\partial_r + \frac{1}{2}\sigma^2(t,r(t))\partial_r^2 - r(t)\bigg)P(t,r(t))\,dt \\
&\qquad\qquad + B^{-1}(t)\sigma(t,r(t))\partial_r P(t,r(t))\,dW(t).
\end{aligned}
\tag{39}
$$

But $\mathbb{Q}$ is a local martingale measure and since $p(t,T)$ is the arbitrage price of a ZCB we know that $(\tilde{P}(t,r(t)))_{0\leq t\leq T}$ must be a (local) martingale. Hence for the above equality to hold we must have

$$\partial_t P + \mu\partial_r P + \frac{1}{2}\sigma^2\partial_r^2 P - rP = 0 \quad \text{on} \quad [0,T)\times\mathbb{R},$$

which is exactly the PDE in the proposition. Finally we note that the terminal condition is trivial since the price of ZCB at maturity is 1. $\square$

We will later need the following result:

Corollary 4.12. *Assume that* $(r(t))$ *is given uniquely by* (35) *for all* $t\geq 0$. *Let*

$$P(t,r) := \mathbb{E}\left[e^{-\int_t^T r(u)du}\middle| r(t) = r\right].$$

i.e. $P(t,r)$ *is the price of* T*-ZCB if at time* t *the short rate is* r. *Then*

$$dP(t,r(t)) = r(t)P(t,r(t))dt + \sigma(t,r(t))\partial_r\left[P(t,r(t))\right]dW(t).$$

Proof. Using that $(\tilde{P}(t,r(t)))_{0\leq t\leq T}$ must be a (local) martingale, we get from (39), that

$$
\begin{aligned}
d\tilde{P}(t,r(t)) &= B^{-1}(t)dP(t,r(t)) + P(t,r(t))d(B^{-1}(t)) \\
&= B^{-1}(t)\sigma(t,r(t))\partial_r P(t,r(t))\,dW(t)\,.
\end{aligned}
$$

Since $d(B^{-1}(t)) = -r(t)B^{-1}(t)\,dt$ and since $B^{-1}(t)\neq 0$ we get

$$dP(t,r(t)) - r(t)P(t,r(t))\,dt = \sigma(t,r(t))\partial_r P(t,r(t))\,dW(t)\,.$$

From this the conclusion of the Corollary follows. $\square$

Now we return to proving Theorem 4.10.

Proof of Theorem 4.10. Our aim is to show that the model has affine term structure, i.e. (34) holds with A and B given by (37), whenever the short rate is given by (35) and (36). The PDE (38) is a linear parabolic equation and thus has a unique solution. Hence, if we find *some* solution to (38) then we know that, in fact, it is *the* solution.

Let us guess that

$$P(t,r) = e^{A(t,T)-B(t,T)r}$$

and let us check whether this satisfies (38). We can immediately conclude that since in (38) we have $P(T,T) = 1$ we must have $A(T,T) = B(T,T) = 0$. Calculating the partial derivatives we get

$$\begin{aligned}
\partial_t \left(e^{A(t,T)-B(t,T)r}\right) &= \left(\partial_t A(t,T) - r\partial_t B(t,T)\right) P(t,r),\\
\partial_r \left(e^{A(t,T)-B(t,T)r}\right) &= -B(t,T)P(t,r),\\
\partial_r^2 \left(e^{A(t,T)-B(t,T)r}\right) &= B^2(t,T)P(t,r).
\end{aligned}$$

Hence (substituting into (38) and dividing by $P > 0$) we see that A and B must satisfy

$$\partial_t A - r\partial_t B - \mu B + \frac{1}{2}\sigma^2 B^2 - r = 0.$$

Due to (36) this becomes

$$\partial_t A - \beta B + \frac{1}{2}\delta B^2 - r\left[1 + \partial_t B + \alpha B - \frac{1}{2}\gamma B^2\right] = 0.$$

Recall that T is fixed throughout (this is about the price of a T-maturity ZCB). If we now fix t as well then the above equation is of the form $c_1 + rc_2 = 0$ for some constants c_1 and c_2 and for all r. But this can only hold true if $c_1 = 0$ and c_2 = 0, or in other words,

$$\begin{aligned}
1 + \partial_t B + \alpha B - \frac{1}{2}\gamma B^2 &= 0,\\
\partial_t A - \beta B + \frac{1}{2}\delta B^2 &= 0.
\end{aligned}$$

This (together with the terminal condition $A(T,T) = B(T,T) = 0$) is exactly (37). Since this system of ODEs has unique solution we have unique solution to (38) and thus the model has affine term structure and (37) holds. □

You can check that all the short rate models discussed so far have affine term structure (but there are short rate models that do not). For example the CIR model: $\mu(t,r) = b - ar$ and $\sigma(t,r) = \bar{\sigma}\sqrt{r}$, where $\bar{\sigma}$ is the constant in the diffusion term. Thus we have $\beta(t) = b$, $\alpha(t) = -a$, $\gamma(t) = \bar{\sigma}$ and $\delta(t) = 0$.

4.3 Calibration to Market ZCB Prices

Of the short rate models we have discussed the Hull–White and Ho-Lee models are rich enough to be calibrated to market observed ZCB prices. We will show how to do this now. The first step is always to obtain a formula for ZCB price first but this is done by using Theorem 4.10.

Ho–Lee Model, ZCB pricing

Let us recall that for some $t \mapsto \theta(t)$ the short rate dynamics in some risk neutral measure $\mathbb{Q}$ in this model are given by

$$dr(t) = \theta(t)dt + \sigma dW(t).$$

This is an affine term structure model due to Theorem 4.10 and to obtain the ZCB price we have to solve (37) which becomes (since $\alpha(t) = 0$, $\beta(t) = \theta(t)$, $\gamma(t) = 0$ and $\delta(t) = \sigma^2$):

$$\begin{aligned} \partial_t B(t,T) = -1, \quad \partial_t A(t,T) &= \theta(t)B(t,T) - \frac{1}{2}\sigma^2 B^2(t,T) \\ B(T,T) = 0, \quad A(T,T) &= 0. \end{aligned}$$

Solving the equation for B is trivial, since we just integrate:

$$B(T,T) - B(t,T) = \int_t^T -1 ds$$

i.e. $B(t,T) = T - t$. To solve for A we integrate the corresponding equation:

$$A(T,T) - A(t,T) = \int_t^T \left[\theta(s)B(s,T) - \frac{1}{2}\sigma^2 B^2(s,T)\right] ds$$

which is

$$A(t,T) = -\int_t^T \theta(s)(T-s)ds - \frac{1}{2}\sigma^2 \int_t^T (T-s)^2 ds.$$

Thus the T-maturity ZCB price (after calculating the second integral) is

$$P(t,r) = \exp\left(\int_t^T \theta(s)(s-T)ds + \frac{1}{6}\sigma^2(T-t)^3 - (T-t)r\right).$$

Ho–Lee Model, Calibration to Observed ZCB Prices

Now we move to the question of calibration. Recall that we are given market ZCB prices observed today for various maturities as a function $T \mapsto p^*(0,T)$. We also consider the observed instantaneous forward rate, see (30):

$$f^*(0,T) = -\partial_T \ln p^*(0,T).$$

Proposition 4.13. *If $T \mapsto p^*(0,T)$ is twice continuously differentiable in T then for all constants $\sigma > 0$ we have $p(0,T) = p^*(0,T)$ for every $T > 0$ in the Ho–Lee model provided that $T \mapsto \theta(T)$ is given by*

$$\theta(T) := \sigma^2 T + \partial_T f^*(0,T).$$

Proof. We take $r(0) = r = f^*(0,0)$. The instantaneous forward rate implied by the model is

$$f(0,T) = -\partial_T \ln p(0,T) = \partial_T \left[\int_0^T \theta(s)(T-s)ds - \frac{1}{6}\sigma^2 T^3 + Tf^*(0,0)\right].$$

A simple calculation leads to

$$\begin{aligned}
f(0,T) &= \partial_T \left[T \int_0^T \theta(s)ds \right] - \partial_T \int_0^T \theta(s)sds - \frac{1}{2}\sigma^2 T^2 + f^*(0,0) \\
&= T\theta(T) + \int_0^T \theta(s)ds - \theta(T)T - \frac{1}{2}\sigma^2 T^2 + f^*(0,0) \\
&= \int_0^T \theta(s)ds - \frac{1}{2}\sigma^2 T^2 + f^*(0,0).
\end{aligned}$$

Hence

$$\partial_T f(0,T) = \theta(T) - \sigma^2 T.$$

Thus, if we take

$$\theta(T) := \sigma^2 T + \partial_T f^*(0,T)$$

then the model implied instantaneous forward rate matches the market instantaneous forward rate and hence the model ZCB prices match the market ZCB prices for any maturity T. $\square$

Hull–White Model, ZCB Pricing

Recall that the short rate in the Hull–White model is given by

$$dr(t) = (\theta(t) - ar(t))dt + \sigma dW(t).$$

Thus, in the terminology of Theorem 4.10 we have $\alpha(t) = -a$, $\beta(t) = \theta(t)$, $\gamma(t) = 0$ and $\delta(t) = \sigma^2$. Thus to solve (37) for B we have to solve

$$\partial_t B(t,T) = aB(t,T) - 1, \quad B(T,T) = 0.$$

This is an ODE of the form

$$dx(t) = [ax(t) - 1]dt, \quad x(T) = 0.$$

We have seen how to solve this: using the usual product rule we calculate

$$d(e^{-at}x(t)) = e^{-at}\left[(ax(t) - 1)dt - ax(t)dt\right] = -e^{-at}dt.$$

Hence

$$e^{-aT}x(T) - e^{-at}x(t) = -\int_t^T e^{-as}ds.$$

This, with the terminal condition, leads to

$$x(t) = \frac{1}{a}\left(1 - e^{-a(T-t)}\right).$$

Hence we have solved for B and it has the form

$$B(t,T) = \frac{1}{a}\left(1 - e^{-a(T-t)}\right).$$

Now we go back to (37) to see that the equation for A now reads as:

$$\partial_t A(t,T) = \frac{\theta(t)}{a}\left(1 - e^{-a(T-t)}\right) - \frac{1}{2}\frac{\sigma^2}{a^2}\left(1 - e^{-a(T-t)}\right)^2.$$

Integrating and using the terminal condition $A(T,T) = 0$ leads to

$$A(t,T) = \int_t^T \left[\frac{1}{2}\frac{\sigma^2}{a^2}\left(1 - e^{-a(T-s)}\right)^2 - \frac{\theta(s)}{a}\left(1 - e^{-a(T-s)}\right)\right] ds. \qquad (40)$$

If we were given the function $s \mapsto \theta(s)$ then we could evaluate this integral (numerically if we have to) and thus obtain the price of a T-maturity ZCB.

Hull–White Model, Calibration to Observed ZCB Prices

We will need the following theorem from calculus: for a function $(s,T) \mapsto g(s,T)$ which is differentiable in the 2nd argument we have

$$\partial_T \left[\int_0^T g(s,T) ds \right] = g(T,T) + \int_0^T \partial_T g(s,T) ds. \qquad (41)$$

Recall that the model ZCB price is

$$p(t,T) = e^{A(t,T) - B(t,T)r(t)}.$$

Moreover the model instantaneous forward rate is:

$$f(t,T) = -\partial_T \ln p(t,T).$$

Hence the model instantaneous forward rate at time $t = 0$ with $r(0) = r$ is (recalling the form of B we obtained earlier):

$$f(0,T) = -\partial_T \left[A(0,T) - B(0,T)r\right] = re^{-aT} - \partial_T A(0,T).$$

To calculate the derivative of the integral we apply (41) to (40), with $t = 0$, to obtain

$$\begin{aligned} f(0,T) &= re^{-aT} - \int_0^T \partial_T \left[\frac{1}{2}\frac{\sigma^2}{a^2}\left(1 - e^{-a(T-s)}\right)^2 - \frac{\theta(s)}{a}\left(1 - e^{-a(T-s)}\right)\right] ds \\ &= re^{-aT} + \int_0^T \theta(s) e^{-a(T-s)} ds - \int_0^T \frac{\sigma^2}{a^2}\left(ae^{-a(T-s)}\right)\left(1 - e^{-a(T-s)}\right) ds. \end{aligned}$$

The last integral on the right hand side can be calculated (with some effort) and thus we obtain

$$f(0,T) = re^{-aT} + \int_0^T \theta(s) e^{-a(T-s)} ds - \frac{\sigma^2}{2a^2}\left(e^{-aT} - 1\right)^2. \qquad (42)$$

Now we use (41) one more time to obtain

$$\partial_T f(0,T) = -rae^{-aT} + \theta(T) - a\int_0^T \theta(s)e^{-a(T-s)} ds - \frac{\sigma^2}{2a^2}\partial_T\left[(e^{-aT} - 1)^2\right].$$

The integral term seems troublesome until we realise that the same term appears in (42). Hence we get

$$\partial_T f(0,T) = \theta(T) - af(0,T) - \frac{\sigma^2}{2a}\left(e^{-aT} - 1\right)^2 - \frac{\sigma^2}{2a^2}\partial_T\left[(e^{-aT} - 1)^2\right].$$

Thus if we choose

$$\theta(T) := \partial_T f^*(0,T) + \frac{\sigma^2}{2a^2}\partial_T\left[(e^{-aT}-1)^2\right] + af^*(0,T) + \frac{\sigma^2}{2a}\left(e^{-aT}-1\right)^2$$

then the model implied ZCB prices and the market ZCB will match. Thus we have shown the following proposition.

Proposition 4.14. *If $T \mapsto p^*(0,T)$ is twice continuously differentiable in T then for all constants $\sigma > 0$ we can find $t \mapsto \theta(t)$ such that $p(0,T) = p^*(0,T)$ for every $T > 0$ in the Hull–White model.*

Now that we have a short rate model calibrated to observed market ZCB prices it is time to use it to price some derivatives.

4.4 Options on ZCBs in Some Short Rate Models

An European call (or put) option on a T_2-maturity ZCB with exercise date T_1, such that $t < T_1 < T_2$ and with strike K is a T_1-contingent claim with payoff

$$X = [\,p(T_1,T_2) - K]_+ \text{ in case of call,}$$
$$X = [K - p(T_1,T_2)]_+ \text{ in case of put.}$$

The arbitrage free price of this contingent claim is given by (31). For a call option this means

$$v(t) = B(t)\mathbb{E}\left[\frac{1}{B(T_1)}\left[p(T_1,T_2) - K\right]_+\middle|\mathcal{F}_t\right].$$

It is possible to provide a simple formula for this. This will be the aim of the rest of this section.

Before we start we, prepare a few things. To make life simpler we introduce the following measure:

Definition 4.15 (T-forward measure)**.** *For any fixed $T > 0$ the* T-forward measure $\mathbb{Q}^T$ *is the measure in the numeraire pair $(N, \mathbb{Q}^T)$ with $N(t) := p(t,T)/p(0,T)$.*

So in the T-forward measure the price of any asset divided by $p(\cdot,T)/p(0,T)$ is a $\mathbb{Q}^T$-local martingale.

We will later use the following:

Proposition 4.16. *Let T be arbitrary and fixed and assume that*

$$p(t,T) = P(t,r(t)) = \exp(A(t,T) - B(t,T)r(t)).$$

Then

$$dp(t,T) = r(t)p(t,T)dt - B(t,T)p(t,T)\sigma(t,r(t))\,dW(t).$$

and

$$d\left(p^{-1}(t,T)\right) = p^{-1}(t,T)\left[\left(B^2(t,T)\sigma(t,r(t))^2 - r(t)\right)dt + B(t,T)\sigma(t,r(t))\,dW(t)\right].$$

Proof. We can use Itô's formula and the dynamics of r under $\mathbb{Q}$ to obtain

$$\begin{aligned} dp(t,T) &= \partial_t P(t,r(t))dt - B(t,T)p(t,T)dr(t) + \partial_r^2 P(t,r(t))dr(t)\cdot dr(t) \\ &= (...)dt - B(t,T)p(t,T)\sigma(t,r(t))\,dW(t). \end{aligned}$$

Moreover

$$d(B^{-1}(t)) = d\left(e^{-\int_0^t r(s)ds}\right) = -r(t)B^{-1}(t)dt$$

and so

$$\begin{aligned} d\left(\frac{p(t,T)}{B(t)}\right) &= p(t,T)d(B^{-1}(t)) + B^{-1}(t)dp(t,T) \\ &= -B^{-1}(t)B(t,T)p(t,T)\sigma(t,r(t))\,dW(t), \end{aligned}$$

since we know that under $\mathbb{Q}$ the process $p(\cdot,T)/B(\cdot)$ is a $\mathbb{Q}$-martingale. Hence

$$-p(t,T)r(t)B^{-1}(t)dt + B^{-1}(t)dp(t,T) = -B^{-1}(t)B(t,T)p(t,T)\sigma(t,r(t))\,dW(t),$$

which, since $B^{-1}(t) \neq 0$ implies that

$$dp(t,T) = r(t)p(t,T)dt - B(t,T)p(t,T)\sigma(t,r(t))\,dW(t).$$

Note that we could have also obtained this from Corollary 4.12. The rest of the proof is just a calculation using Itô's formula:

$$\begin{aligned} d\left(p^{-1}(t,T)\right) &= -p^{-2}(t,T)dp(t,T) + p^{-3}(t,T)dp(t,T)\cdot dp(t,T) \\ &= p^{-1}(t,T)\left[\left(B^2(t,T)\sigma(t,r(t))^2 - r(t)\right)dt + B(t,T)\sigma(t,r(t))\,dW(t)\right]. \end{aligned}$$

□

Now we return to calculating the call option price. First we note that

$$X = [\,p(T_1,T_2) - K]_+ = [\,p(T_1,T_2) - K]\,\mathbb{1}_{\{p(T_1,T_2)\geq K\}}$$

and so we can write

$$v(t) = v_1(t) - Kv_2(t),$$

where

$$\begin{aligned} v_1(t) &:= B(t)\mathbb{E}\left[\frac{p(T_1,T_2)}{B(T_1)}\mathbb{1}_{\{p(T_1,T_2)\geq K\}}\middle|\mathcal{F}_t\right], \\ v_2(t) &:= B(t)\mathbb{E}\left[\frac{1}{B(T_1)}\mathbb{1}_{\{p(T_1,T_2)\geq K\}}\middle|\mathcal{F}_t\right]. \end{aligned}$$

We start by considering $v_2(t)$ as it looks simpler. Somehow we would like to remove the $1/B(T_1)$ term. For inspiration we read about change of Numeraire again, see Section 3, and then see what we can do. Let us introduce a new numeraire $N_1(t) = p(t,T_1)/p(0,T_1)$. This means we will work in the T_1-forward measure. We have to check that $\mathbb{E}[N_1(T_1)/B(T_1)] = 1$ before we can apply Proposition 3.5. We know that N_1/B is a $\mathbb{Q}$-martingale and so

$$\mathbb{E}\left[\frac{N_1(T_1)}{B(T_1)}\right] = \mathbb{E}\left[\frac{N_1(0)}{B(0)}\right] = \mathbb{E}\left[\frac{\frac{p(0,T_1)}{p(0,T_1)}}{1}\right] = 1.$$

Thus, from Proposition 3.5 we get

$$v_2(t) = N_1(t)\mathbb{E}^{\mathbb{Q}^{T_1}}\left[\frac{p(0,T_1)}{p(T_1,T_1)}\mathbb{1}_{\{p(T_1,T_2)\geq K\}}\middle|\mathcal{F}_t\right].$$

Noticing that $p(0,T_1)$ is deterministic and known this simply becomes

$$v_2(t) = p(0,T_1)N_1(t)\mathbb{Q}^{T_1}\left[\frac{p(T_1,T_2)}{p(T_1,T_1)}\geq K\middle|\mathcal{F}_t\right].$$

In fact $p(T_1,T_1) = 1$ so we could get rid of that as well but it turns out that it makes more sense to consider $Z(t) := p(t,T_2)/p(t,T_1)$ noting that Z is a martingale in the T_1-forward measure. Thus we are interested in

$$v_2(t) = p(0,T_1)N_1(t)\mathbb{Q}^{T_1}\left[Z(T_1)\geq K\middle|\mathcal{F}_t\right].$$

We will make use of Proposition 4.16 to calculate the dynamics of Z first in $\mathbb{Q}$ but then change into the T_1 forward measure and use the fact that it's a martingale under T_1-forward measure.

$$\begin{aligned} dZ(t) &= p(t,T_2)dp^{-1}(t,T_1) + p^{-1}(t,T_1)dp(t,T_2) + dp(t,T_2)\cdot dp^{-1}(t,T_1) \\ &= (\ldots)dt + Z(t)B(t,T_1)\sigma(t,r(t))\,dW(t) - Z(t)B(t,T_2)\sigma(t,r(t))\,dW(t). \end{aligned}$$

Under the T_1-forward measure there is no drift and hence

$$dZ(t) = Z(t)\left[B(t,T_1) - B(t,T_2)\right]\sigma(t,r(t))\,dW^{\mathbb{Q}^{T_1}}(t).$$

We can use this to calculate $v_2(0)$. Let $\sigma_Z(t) := (B(t,T_1)-B(t,T_2))\sigma(t,r(t))$. Then solving for $Z(T_1)$ leads to

$$Z(T_1) = Z(0)\exp\left(-\frac{1}{2}\int_0^{T_1}\sigma_Z(t)^2dt + \int_0^{T_1}\sigma_Z(t)dW^{\mathbb{Q}^{T_1}}(t)\right).$$

From now on we *assume* that σ is a deterministic function of t, (so independent of r). We can then use our knowledge that a stochastic integral of a deterministic function has normal distribution with mean 0 and variance given by the Itô isometry we get that $Z(T_1)$ has the same distribution as

$$Z(0)\exp\left(-\frac{1}{2}\Sigma_2^2 + \Sigma_2\xi\right),$$

where

$$\Sigma_2^2 := \int_0^{T_1}\sigma^2(t)(B(t,T_1) - B(t,T_2))^2\,dt$$

and $\xi \sim N(0,1)$. Thus (since $Z(0) = \frac{p(0,T_2)}{p(0,T_1)}$)

$$\begin{aligned} \mathbb{Q}^{T_1}\left[Z(T_1)\geq K\right] &= \mathbb{Q}^{T_1}\left[Z(0)\exp\left(-\frac{1}{2}\Sigma_2^2 + \Sigma_2\xi\right)\geq K\right] \\ &= \mathbb{Q}^{T_1}\left[\xi \geq \frac{\ln\frac{K}{Z(0)} + \frac{1}{2}\Sigma_2^2}{\Sigma_2}\right] = (1 - N(d_2)), \end{aligned}$$

where $x \mapsto N(x)$ is the distribution function of $N(0,1)$ and

$$d_2 := \frac{\ln \frac{Kp(0,T_1)}{p(0,T_2)} + \frac{1}{2}\Sigma_2^2}{\Sigma_2}.$$

At time 0 we know the values of both $p(0,T)$ for all $T > 0$ and hence this is something we can calculate easily. Thus

$$v_2(0) = p(0,T_1)(1 - N(d_2)).$$

We still have to calculate (for $t = 0$)

$$v_1(t) = B(t)\mathbb{E}\left[\frac{p(T_1,T_2)}{B(T_1)}\mathbb{1}_{\{p(T_1,T_2)\geq K\}}\middle|\mathcal{F}_t\right].$$

We now consider the numeraire N_2 given by $N_2(t) := p(t,T_2)/p(0,T_2)$. For any t we have

$$\mathbb{E}\left[\frac{N_2(t)}{B(t)}\right] = \frac{1}{p(0,T_2)}\mathbb{E}\left[\frac{p(t,T_2)}{B(t)}\right] = \frac{1}{p(0,T_2)}\frac{p(0,T_2)}{B(0)} = 1,$$

since $p(\cdot,T_2)/B(\cdot)$ is a $\mathbb{Q}$-martingale. Thus we may apply Proposition 3.5 and work in the T_2-forward measure. This yields

$$\begin{aligned}
v_1(t) &= N_2(t)\mathbb{E}^{\mathbb{Q}^{T_2}}\left[\frac{p(T_1,T_2)}{N_2(T_1)}\mathbb{1}_{\{p(T_1,T_2)\geq K\}}\middle|\mathcal{F}_t\right] \\
&= p(t,T_2)\mathbb{Q}^{T_2}\left[p(T_1,T_2)\geq K\middle|\mathcal{F}_t\right] = p(t,T_2)\mathbb{Q}^{T_2}\left[\frac{p(T_1,T_2)}{p(T_1,T_1)}\geq K\middle|\mathcal{F}_t\right] \\
&= p(t,T_2)\mathbb{Q}^{T_2}\left[\frac{p(T_1,T_1)}{p(T_1,T_2)}\leq \frac{1}{K}\middle|\mathcal{F}_t\right] = p(t,T_2)\mathbb{Q}^{T_2}\left[Y(T_1)\leq \frac{1}{K}\middle|\mathcal{F}_t\right],
\end{aligned}$$

where $Y(t) := p(t,T_1)/p(t,T_2)$. This is convenient since Y a martingale in the T_2-forward measure. Recalling the calculation performed for the dynamics of Z we have that under $\mathbb{Q}$

$$dY(t) = (\ldots)dt + Y(t)\left[B(t,T_2) - B(t,T_1)\right]\sigma(t,r(t))\,dW(t).$$

But under the T_2-forward measure

$$dY(t) = Y(t)\left[B(t,T_2) - B(t,T_1)\right]\sigma(t,r(t))\,dW^{\mathbb{Q}^{T_2}}(t).$$

To evaluate $v_1(0)$ we solve the above SDE with $Y(0)$ given.

Again we *assume* that σ is a deterministic function of t, (so independent of r). As before, we conclude that

$$Y(T_1) \stackrel{d}{=} Y(0)\exp\left(-\frac{1}{2}\Sigma_1^2 + \Sigma_1\xi\right),$$

where $\stackrel{d}{=}$ means that two random variables have identical distributions, where $\xi \sim N(0,1)$ and where

$$\Sigma_1^2 := \int_0^{T_1} \sigma(t)^2(B(t,T_2) - B(t,T_1))^2 dt.$$

Hence

$$v_1(0) = p(0,T_2)N(d_1), \text{ where } d_1 := \frac{\ln \frac{p(0,T_2)}{Kp(0,T_1)} + \frac{1}{2}\Sigma_1^2}{\Sigma_1}.$$

To summarize we have proved the following proposition.

Proposition 4.17. *Under any affine short rate model that has deterministic diffusion coefficient independent of the short rate we have the price at time* 0 *of an European call option with exercise time* T_1 *on a* T_2*-ZCB is given by*

$$v(0) = p(0,T_2)N(d_1) - Kp(0,T_1)(1 - N(d_2)),$$

where

$$d_1 := \frac{\ln \frac{p(0,T_2)}{Kp(0,T_1)} + \frac{1}{2}\Sigma_1^2}{\Sigma_1} \text{ and } d_2 := \frac{\ln \frac{Kp(0,T_2)}{p(0,T_1)} + \frac{1}{2}\Sigma_2^2}{\Sigma_2},$$

with

$$\Sigma_1 := \sigma \left(\int_0^{T_1} (B(t,T_2) - B(t,T_1))^2 dt \right)^{1/2},$$

$$\Sigma_2 := \sigma \left(\int_0^{T_1} (B(t,T_1) - B(t,T_2))^2 dt \right)^{1/2}.$$

This formula of its own is of not much more than theoretical interest. However it is possible to adapt it for pricing of bonds (rather than just simple ZCBs), see e.g. Filipovic [4, Exercise 7.9], or *swaptions* (options of interest rate swaps) using a method called Jamshidian decomposition.

This is all that we will say about short rate models. More can be found in e.g. Brigo and Mercurio [2].

4.5 The Heath–Jarrow–Morton Framework

The Heath–Jarrow–Morton (HJM) framework is a general setting in which many interest rate models (not only short rate models) may be analysed. This is done by writing down an equation describing the evolution of the entire forward curve. Only some of these definitions of evolution of the forward curve will lead to a model that is free of arbitrage. We will discuss how to check whether a given forward curve evolution leads to an arbitrage free model or not.

Recall that due to Remark 4.6 we have

$$p(t,T) = \exp\left(- \int_t^T f(t,s)ds \right). \tag{43}$$

The expression is the price, at time t of a T-maturity ZCB. This is an $\mathcal{F}_t$-measurable random variable. The integrand on the right hand side is instantaneous forward rate observed at time t for the period "$[s, s+ds)$". For each $s \geq t$ this is also an $\mathcal{F}_t$-measurable random variable. Hence if we have $f(t,s)$ for every $t \geq 0$ and $s \geq t$ then we have the evolution of the entire interest rate term structure.

Hence we postulate that for every $T \geq t$

$$df(t,T) = \alpha(t,T)dt + \sigma(t,T)dW(t),$$

with $f(0,T) = f^*(0,T)$. For each fixed T we have $(\alpha(t,T))_{t\in[0,T]}$ and $(\sigma(t,T))_{t\in[0,T]}$ adapted stochastic processes such that at least

$$\mathbb{P}\left[\int_0^T |\alpha(s,T)|ds < \infty, \int_0^T |\sigma(s,T)|^2 ds < \infty\right] = 1.$$

This ensures that the integrals are well defined (and the stochastic integral is a local martingale). Finally, W is assumed to be a Wiener process in some local martingale measure $\mathbb{Q}$. But here a problem arises: the relation (43) defines ZCB prices. Since $\mathbb{Q}$ is a local martingale measure the processes $p(\cdot,T)/B$ must be local martingales for every $T \geq 0$, where B is the bank account process:

$$B(t) = B(0)\exp\left(\int_0^t r(u)du\right) = \exp\left(\int_0^t f(u,u)du\right),$$

since $r(u) = f(u,u)$, see Definition 4.7. It is far from clear which choices of α and σ create ZCB prices which obey the property that $p(\cdot,T)/B$ are local martingales under $\mathbb{Q}$. Here is the answer.

Proposition 4.18 (HJM consistency condition)**.** *If*

$$\alpha(t,T) = \sigma(t,T)\int_t^T \sigma(t,u)du \tag{44}$$

then the instantaneous forward rates

$$df(t,T) = a(t,T)dt + \sigma(t,T)dW(t) \tag{45}$$

lead to ZCB prices such that $p(\cdot,T)/B$ are local martingales under $\mathbb{Q}$ for all $T \geq 0$ (i.e. the ZCB prices are free of arbitrage) and moreover

$$dp(t,T) = p(t,T)\left[f(t,t)dt + v(t,T)dW(t)\right] \tag{46}$$

with

$$v(t,T) := -\int_t^T \sigma(t,u)du.$$

Before we get to the proof of this we will need the following theorem (all the measurability criteria are included only for completeness, we are not working at this level of rigour in our course).

Theorem 4.19 (Fubini's Theorem for Stochastic Integrals)**.** *Consider the stochastic process $\phi = (\phi(\omega,t,s))_{0\leq t,s\leq T}$ such that the map $(\omega,t) \mapsto \phi(\omega,t,s)$ is measurable with respect to the product σ-algebra of Prog$_T$ and $\mathcal{B}([0,T])$ and such that*

$$\sup_{t,s\in[0,T]} |\phi(t,s)| < \infty \textit{ a.s. } .$$

Then the process defined by $\lambda(t) = \int_0^T \phi(t,s)ds$ belongs to $\mathcal{S}$ and there is a process ψ which is measurable with respect to the product of the σ-algebras $\mathcal{F}_T$ and $\mathcal{B}[0,T]$ and such that almost surely

$$\psi(s) = \int_0^T \phi(t,s)dW(t) \text{ and } \int_0^T \psi^2(s)ds < \infty.$$

Moreover

$$\int_0^T \psi(s)ds = \int_0^T \lambda(t)dW(t)$$

i.e.

$$\int_0^T \int_0^T \phi(t,s)\, dW(t)\, ds = \int_0^T \int_0^T \phi(t,s)\, ds\, dW(t). \tag{47}$$

For proof see e.g. Filipovic [4, Theorem 6.2] We now return to proving Proposition 4.18.

Proof of Proposition 4.18. We start with a calculation, using the definition of forward rate (45). Hence

$$\begin{aligned}\ln p(t,T) &= -\int_t^T f(t,u)du \\ &= -\int_t^T f(0,u)du - \int_t^T \int_0^t \alpha(s,u)dsdu - \int_t^T \int_0^t \sigma(s,u)dW(s)du.\end{aligned}$$

We now wish to reverse the order of integration in the 2nd and 3rd integrals on the right-hand side. Let

$$\tilde{\sigma}(s,u) := \sigma(s,u)\mathbb{1}_{\{t\le u\le T\}}\mathbb{1}_{\{0\le s\le t\}}.$$

Then, using (47),

$$\begin{aligned}\int_t^T \int_0^t \sigma(s,u)dW(s)du &= \int_0^T \int_0^T \tilde{\sigma}(s,u)dW(s)du \\ = \int_0^T \int_0^T \tilde{\sigma}(s,u)du\, dW(s) &= \int_0^t \int_t^T \sigma(s,u)du\, dW(s).\end{aligned}$$

Similarly we can use the "classical" Fubini's theorem to see that

$$\int_t^T \int_0^t \alpha(s,u)dsdu = \int_0^t \int_t^T \alpha(s,u)duds.$$

Hence

$$\begin{aligned}&\ln p(t,T) \\ &= -\int_t^T f(0,u)du - \int_0^t \int_t^T \alpha(s,u)duds - \int_0^t \int_t^T \sigma(s,u)du\, dW(s).\end{aligned}$$

We now rewrite the integrals as follows

$$\begin{aligned}&\ln p(t,T) \\ &= -\int_0^T f(0,u)du - \int_0^t \int_s^T \alpha(s,u)duds - \int_0^t \int_s^T \sigma(s,u)du\, dW(s) \\ &\quad + \int_0^t f(0,u)du + \int_0^t \int_s^t \alpha(s,u)duds + \int_0^t \int_s^t \sigma(s,u)du\, dW(s).\end{aligned}$$

At this point we use the stochastic Fubini's theorem to observe that

$$\int_0^t \int_s^t \sigma(s,u)du\, dW(s) = \int_0^t \int_0^t \sigma(s,u)\mathbb{1}_{\{s\le u\le t\}}du\, dW(s)$$
$$= \int_0^t \int_0^t \sigma(s,u)\mathbb{1}_{\{s\le u\le t\}}dW(s)du = \int_0^t \int_0^u \sigma(s,u)dW(s)du$$

and we use the classical Fubini's theorem to similarly see that

$$\int_0^t \int_s^t \alpha(s,u)du\, ds = \int_0^t \int_0^u \alpha(s,u)ds\, du$$

But we also know that

$$r(u) = f(u,u) = f(0,u) + \int_0^u \alpha(s,u)ds + \int_0^u \sigma(s,u)dW(s)$$

and so

$$\begin{aligned}
&\ln p(t,T) \\
&= -\int_0^T f(0,u)du - \int_0^t \int_s^T \alpha(s,u)duds - \int_0^t \int_s^T \sigma(s,u)du\, dW(s) \\
&\quad + \int_0^t f(0,u)du + \int_0^t \int_0^u \alpha(s,u)dsdu + \int_0^t \int_0^u \sigma(s,u)dW(s)du \\
&= -\int_0^T f(0,u)du - \int_0^t \int_s^T \alpha(s,u)duds - \int_0^t \int_s^T \sigma(s,u)du\, dW(s) \\
&\quad + \int_0^t r(u)du.
\end{aligned}$$

Let us define

$$v(s,T) := -\int_s^T \sigma(s,u)du \quad \text{and} \quad b(s,T) := -\int_s^T \alpha(s,u)du + \frac{1}{2}|v(s,T)|^2.$$

Then

$$\begin{aligned}
&\ln p(t,T) = \ln p(0,T) \\
&+ \int_0^t \left(b(s,T) - \frac{1}{2}|v(s,T)|^2\right) ds + \int_0^t v(s,T)\, dW(s) + \int_0^t r(u)du.
\end{aligned}$$

In other words Then

$$d\left(\ln p(t,T)\right) = \left(r(t) + b(t,T) - \frac{1}{2}|v(t,T)|^2\right) dt + v(t,T)dW(t).$$

We apply Itô's formula to the function $x \mapsto e^x$ and the process $\ln p(\cdot,T)$ to obtain

$$dp(t,T) = p(t,T)\left[\left(r(t) + b(t,T)\right) dt + v(t,T)dW(t)\right].$$

We know that under $\mathbb{Q}$ which is a local martingale measure we need the process given by $\tilde{p}(t,T) := p(t,T)/B(t)$ to be a local martingale. Moreover

$$d(B^{-1}(t)) = d\left(e^{-\int_0^t r(s)ds}\right) = -r(t)B^{-1}(t)dt$$

and so we see that

$$\begin{aligned}d(\tilde{p}(t,T)) &= -r(t)\tilde{p}(t,T)dt + \tilde{p}(t,T)\left[(r(t)+b(t,T))\,dt + v(t,T)dW(t)\right]\\ &= \tilde{p}(t,T)\left[b(t,T)dt + v(t,T)dW(t)\right].\end{aligned}$$

Hence we need $b(t,T)=0$ for all $0\le t\le T$ which is equivalent to

$$\int_t^T \alpha(t,u)du = \frac{1}{2}|v(t,T)|^2.$$

Differentiating this with respect to T leads to

$$\alpha(t,T) = (\partial_T v(t,T))\, v(t,T) = \sigma(t,T)\int_t^T \sigma(t,u)du.$$

Thus if (44) holds then $\tilde{p}(t,T) := p(t,T)/B(t)$ is a local martingale under $\mathbb{Q}$ and

$$dp(t,T) = p(t,T)\left[r(t)dt + v(t,T)dW(t)\right]$$

which shows (46). □

Remark 4.20. From Proposition 4.18 we see that to entirely specify the interest rate term structure evolution we need only to specify the functions $(t,T)\mapsto\sigma(t,T)$ and $T\mapsto f(0,T)$. For the latter it is natural to take $f(0,T)=f^*(0,T)$ i.e. the observed forward rates. Thus we can say that in a model given via the HJM framework the calibration to the zero coupon curve / forward curve is “automatic”.

Example 4.21. Take $\sigma(t,T)=\sigma\in\mathbb{R}^+$ a constant. Then, following Proposition 4.18 leads to

$$\alpha(t,T) = \sigma\int_t^T \sigma du = \sigma^2(T-t).$$

Now

$$\begin{aligned}r(t) = f(t,t) &= f(0,t) + \int_0^t \sigma^2(t-s)\,ds + \int_0^t \sigma\,dW(s)\\ &= f(0,t) + \frac{\sigma^2 t^2}{2} + \sigma W(t).\end{aligned}$$

If we apply Itô’s formula we get

$$dr(t) = \left(\frac{d}{dt}f(0,t) + \sigma^2 t\right)dt + \sigma dW(t).$$

Taking $\theta(t) := \frac{d}{dt}f(0,t) + \sigma^2 t$ leads exactly to the Ho–Lee model calibrated to observed zero coupon curve.

4.6 Exercises

Exercise 4.1.

1. Solve (33) *Hint:* Apply Itô’s formula to the process r and the function $(t,x)\mapsto e^{at}x$.
2. Is the function $t\mapsto r(t)$ continuous? Why?

3. Calculate $\mathbb{E}r(t)$ and $\mathbb{E}\left[r^2(t)\right]$.

4. What is the distribution of $r(t)$?

Exercise 4.2. Let $W = (W_1, W_2)$ be an $\mathbb{R}^2$-valued Wiener process. Use Itô formula to derive the solution to

$$dY(t) = (a(t) - Y(t)b(t))dt + \sigma_1(t)dW_1(t) + \sigma_2(t)dW_2(t),$$

where a, b, σ_1 and σ_2 are processes adapted to the filtration generated by W.

Exercise 4.3. Let $\mathbb{Q}$ be a risk-neutral measure and let $W^{\mathbb{Q}}$ be a $\mathbb{Q}$-Wiener process. The Ho–Lee model for short rate is

$$dr(t) = \theta(t)dt + \sigma dW^{\mathbb{Q}}(t),$$

where $\sigma \in \mathbb{R}$ is a given constant. Assume that the forward rate observed at $t = 0$ is given by[10]

$$f^*(0,T) = 1 + \sin(T).$$

Fix $\sigma > 0$. What is the calibration of the Ho–Lee model such that the theoretical forward rate matches the "observed" rate? (i.e. what is $t \mapsto \theta(t)$?)

Exercise 4.4. Let $\mathbb{Q}$ be a risk-neutral measure and let $W^{\mathbb{Q}}$ be a $\mathbb{Q}$-Wiener process. In the Heath–Jarrow–Morton (HJM) framework the instantaneous rate for time T, observed at time t, denoted $f(t,T)$, is given by

$$f(t,T) = f(0,T) + \int_0^t \alpha(s,T)ds + \int_0^t \sigma(s,T)dW^{\mathbb{Q}}(s).$$

State the HJM consistency (drift) condition.

Show that if we take $\sigma(t,T) := \sigma\exp(-a(T-t))$, where $\sigma, a > 0$ given constants, and

$$f(0,T) := \frac{\sigma^2}{2a^2}\left[2e^{-aT} - 1 - e^{-2aT}\right]$$

then the HJM framework gives the Vasicek model with $b = r_0 = 0$.

Exercise 4.5. Say $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and $(\mathcal{F}_t)_{t\geq 0}$ filtration. Let $Z = (Z(t))_{t\geq 0}$ be a local martingale. Assume that there is a r.v. Y s.t. $\mathbb{E}|Y| < \infty$ and $|Z(t)| \leq Y$ for all $t \geq 0$. Show that Z is a martingale.

Hint. Use the Dominated Convergence Theorem for conditional expectations, see Theorem A.9.

[10]Of course this is just a silly exercise. You'll never see this in the market!

A Useful Results from Other Courses

"The beginner should not be discouraged if he finds he does not have the prerequisites for reading the prerequisites."

– Paul Halmos. *Measure Theory*, 1950.

The aim of this section is to collect, mostly without proofs, results that are needed or useful for this course but that cannot be covered in the lectures i.e. prerequisites. You are expected to be able to use the results given here.

A.1 Linear Algebra

The inverse of a square real matrix A exists if and only if $\det(A) \neq 0$.

The inverse of square real matricies A and B exists if and only if the inverse of AB exists and moreover $(AB)^{-1} = B^{-1}A^{-1}$.

The inverse of a square real matrix A exists if and only if the inverse of A^T exists and $(A^T)^{-1} = (A^{-1})^T$.

If x is a vector in $\mathbb{R}^d$ then diag(x) denotes the matrix in $\mathbb{R}^{d\times d}$ with the entries of x on its diagonal and zeros everywhere else. The inverse of diag(x) exists if and only if $x_i \neq 0$ for all $i = 1, \dots, d$ and moreover

$$\text{diag}(x)^{-1} = \text{diag}(1/x_1, 1/x_2, \dots, 1/x_d).$$

A.2 Real Analysis and Measure Theory

Let $(X, \mathcal{X}, \mu)$ be a measure space (i.e. X is a set, $\mathcal{X}$ a σ-algebra and μ a measure).

Lemma A.1 (Fatou's Lemma)**.** *Let $f_1, f_2, \dots$ be a sequence of non-negative and measurable functions. Then the function defined point-wise as*

$$f(x) := \liminf_{k\to\infty} f_k(x)$$

is $\mathcal{X}$-measurable and

$$\int_X f\, d\mu \leq \liminf_{k\to\infty} \int_X f_k\, d\mu.$$

Consider sets X and Y with σ-algebras $\mathcal{X}$ and $\mathcal{Y}$. By $\mathcal{X} \times \mathcal{Y}$ we denote the collection of sets $C = A \times B$ where $A \in \mathcal{X}$ and $B \in \mathcal{Y}$. By $\mathcal{X} \otimes \mathcal{Y} = \sigma(\mathcal{X} \times \mathcal{Y})$, which is the σ-algebra generated by $\mathcal{X} \times \mathcal{Y}$.

Theorem A.2. *Let $f : X \times Y \to \mathbb{R}$ be a measurable function, i.e. measurable with respect to the σ-algebras $\mathcal{X} \otimes \mathcal{Y}$ and $\mathcal{B}(\mathbb{R})$. Then for each $x \in X$ the function $y \mapsto f(x, y)$ is measurable with respecto to $\mathcal{Y}$ and $\mathcal{B}(\mathbb{R})$. Similarly for each $y \in Y$ the function $x \mapsto f(x, y)$ is measurable with respecto to $\mathcal{X}$ and $\mathcal{B}(\mathbb{R})$.*

The proof is short and so it's easiest to just include it here.

Proof. We first consider functions of the form $f = \mathbb{1}_C$ with $C \in \mathcal{X} \otimes \mathcal{Y}$. Let

$$\mathcal{H} = \{C \in \mathcal{X} \otimes \mathcal{Y} : y \mapsto \mathbb{1}_C(x, y) \text{ is } \mathcal{F} - \text{measurable for each fixed } x \in E\}.$$

It is easy to check that $\mathcal{H}$ is a σ-algebra. Moreover if $C = A \times B$ with $A \in \mathcal{X}$ and $B \in \mathcal{Y}$ then

$$y \mapsto \mathbb{1}_C(x, y) = \mathbb{1}_A(x)\mathbb{1}_B(y).$$

As x is fixed $\mathbb{1}_A(x)$ is just a constant and since $B \in \mathcal{Y}$ the function $y \mapsto \mathbb{1}_A(x)\mathbb{1}_B(y)$ must be measurable. Hence $\mathcal{X} \times \mathcal{Y} \subseteq \mathcal{H}$ and thus $\mathcal{X} \otimes \mathcal{Y} \subseteq \mathcal{H}$. But $\mathcal{H} \subseteq \mathcal{X} \otimes \mathcal{Y}$ and so $\mathcal{H} = \mathcal{X} \otimes \mathcal{Y}$. Hence if f is a simple function then the conclusion of the theorem holds.

Now consider $f \geq 0$ and let f_n be a sequence of simple functions increasing to f. Then for a fixed x the function $y \mapsto g_n(y) = f_n(x, y)$ is measurable. Moreover since $g(y) = \lim_{n\to\infty} g_n(y) = f(x, y)$ and since the limit of measurable functions is measurable we get the result for $f \geq 0$. For general $f = f^+ - f^-$ the result follows using the result for $f^+ \geq 0, f^- \geq 0$ and noting that the difference of measurable functions is measurable. $\square$

Consider measure spaces $(X, \mathcal{X}, \mu_x)$, $(Y, \mathcal{Y}, \mu_y)$. That is, X and Y are sets, $\mathcal{X}$ and $\mathcal{Y}$ are σ-algebras and μ_x and μ_y are measures on $\mathcal{X}$ and $\mathcal{Y}$ respectively. For all details on Fubini's Theorem we refer to Kolmogorov and Fomin [7].

Theorem A.3 (Fubini)**.** *Let μ be the Lebesgue extension of $\mu_x \otimes \mu_y$. Let $A \in \mathcal{X} \otimes \mathcal{Y}$. and let $f : A \to \mathbb{R}$ be a measurable function (considering $\mathcal{B}(\mathbb{R})$, the Borel σ-algebra on $\mathbb{R}$). If f is integrable i.e. if*

$$\int_A |f(x, y)| d\mu < \infty$$

then

$$\int_A f(x, y) d\mu = \int_X \left[\int_{A_x} f(x, y) d\mu_y \right] d\mu_x = \int_Y \left[\int_{A_y} f(x, y) d\mu_x \right] d\mu_y,$$

where $A_x := \{y \in Y : (x, y) \in A\}$ and $A_y := \{x \in X : (x, y) \in A\}$.

Remark A.4. The conclusion of Fubini's theorem implies that for μ_x-almost all x the integral $\int_{A_x} f(x, y) d\mu_y$ exists which in turn implies that the function $f(x, \cdot) : A_x \to \mathbb{R}$ must be measurable. This statement also holds if we exchange x for y.

A.3 Conditional Expectation

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be given.

Theorem A.5. *Let X be an integrable random variable. If $\mathcal{G} \subseteq \mathcal{F}$ is a σ-algebra then there exists a unique $\mathcal{G}$ measurable random variable Z such that*

$$\forall G \in \mathcal{G} \quad \int_G X d\mathbb{P} = \int_G Z d\mathbb{P}.$$

The proof can be found in xxxx xxxx.

Definition A.6. *Let X be an integrable random variable. If $\mathcal{G} \subseteq \mathcal{F}$ is a σ-algebra then $\mathcal{G}$-random variable from Theorem A.5 is called the* conditional expectation *of X given $\mathcal{G}$ and write $\mathbb{E}(X|\mathcal{G}) := Z$.*

Conditional expectations are rather abstract notion so two examples might help.

Example A.7. Consider $\mathcal{G} := \{\emptyset, \Omega\}$. So $\mathcal{G}$ is just the trivial σ-algebra. For a random variable X we then have, by definition, that Z is the conditional expectation (denoted $\mathbb{E}[X|\mathcal{G}]$), if and only if

$$\int_\Omega Z d\mathbb{P} = \int_\Omega X d\mathbb{P}.$$

The right hand side of the above expression is in fact just $\mathbb{E}X$ and so the equality would be satisfied if we set $Z = \mathbb{E}X$ (just a constant). Indeed then (going right to left)

$$\mathbb{E}X = \int_\Omega X d\mathbb{P} = \int_\Omega Z d\mathbb{P} = \int_\Omega \mathbb{E}X d\mathbb{P} = \mathbb{E}X \int_\Omega d\mathbb{P} = \mathbb{E}X.$$

Example A.8. Let $X \sim N(0,1)$. Let $\mathcal{G} = \{\emptyset, \{X \leq 0\}, \{X > 0\}, \Omega\}$. One can (and should) check that this is a σ-algebra. By definition the conditional expectation is a unique random variable that satisfies

$$\begin{aligned} \int_\Omega \mathbb{1}_{\{X>0\}} Z d\mathbb{P} &= \int_\Omega \mathbb{1}_{\{X>0\}} X d\mathbb{P}, \\ \int_\Omega \mathbb{1}_{\{X\leq 0\}} Z d\mathbb{P} &= \int_\Omega \mathbb{1}_{\{X\leq 0\}} X d\mathbb{P}, \\ \int_\Omega Z d\mathbb{P} &= \int_\Omega X d\mathbb{P}\,. \end{aligned} \tag{48}$$

It is a matter of integrating with respect to normal density to find out that

$$\int_\Omega \mathbb{1}_{\{X>0\}} X d\mathbb{P} = \int_0^\infty x\phi(x)dx = \frac{1}{2}\sqrt{\frac{2}{\pi}}\,, \quad \int_\Omega \mathbb{1}_{\{X\leq 0\}} X d\mathbb{P} = -\frac{1}{2}\sqrt{\frac{2}{\pi}}. \tag{49}$$

Since Z must be $\mathcal{G}$ measurable it can only take two values:

$$Z = \begin{cases} z_1 & \text{on} \quad \{X > 0\}, \\ z_2 & \text{on} \quad \{X \leq 0\}, \end{cases}$$

for some real constants z_1 and z_2 to be yet determined. But (48) and (49) taken together imply that

$$\frac{1}{2}\sqrt{\frac{2}{\pi}} = \int_\Omega \mathbb{1}_{\{X>0\}} Z d\mathbb{P} = \int_\Omega \mathbb{1}_{\{X>0\}} z_1 d\mathbb{P} = z_1 \mathbb{P}(X > 0) = \frac{1}{2} z_1.$$

Hence $z_1 = \sqrt{2/\pi}$. Similarly we calculate that $z_2 = -\sqrt{2/\pi}$. Finally we check that the third equation in (48) holds. Thus

$$\mathbb{E}[X|\mathcal{G}] = Z = \begin{cases} \sqrt{\frac{2}{\pi}} & \text{on} \quad \{X > 0\}, \\ -\sqrt{\frac{2}{\pi}} & \text{on} \quad \{X \leq 0\}. \end{cases}$$

Here are some further important properties of conditional expectations which we present without proof.

Theorem A.9 (Properties of conditional expectations)**.** *Let X and Y be random variables. Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$.*

1. *If $\mathcal{G} = \{\emptyset, \Omega\}$ then $\mathbb{E}(X|\mathcal{G}) = \mathbb{E}X$.*
2. *If $X = x$ a. s. for some constant $x \in \mathbb{R}$ then $\mathbb{E}(X|\mathcal{G}) = x$ a.s..*
3. *For any $\alpha, \beta \in \mathbb{R}$*
$$\mathbb{E}(\alpha X + \beta Y|\mathcal{G}) = \alpha\mathbb{E}(X|\mathcal{G}) + \beta\mathbb{E}(Y|\mathcal{G}).$$
This is called linearity.
4. *If $X \leq Y$ almost surely then $\mathbb{E}(X|\mathcal{G}) \leq \mathbb{E}(Y|\mathcal{G})$a.s..*
5. $|\mathbb{E}(X|\mathcal{G})| \leq \mathbb{E}(|X|\,|\mathcal{G})$.
6. *If $X_n \to X$ a. s. and $|X_n| \leq Z$ for some integrable Z then $\mathbb{E}(X_n|\mathcal{G}) \to \mathbb{E}(X|\mathcal{G})$ a. s. . This is the "dominated convergence theorem for conditional expectation".*
7. *If Y is $\mathcal{G}$ measurable then $\mathbb{E}(XY|\mathcal{G}) = YE(X|\mathcal{G})$.*
8. *Let $\mathcal{H}$ be a sub-σ-algebra of $\mathcal{G}$. Then*
$$\mathbb{E}(X|\mathcal{H}) = \mathbb{E}(\mathbb{E}(X|\mathcal{G})|\mathcal{H}).$$
This is called the tower property. *A special case is $\mathbb{E}X = \mathbb{E}(\mathbb{E}(X|\mathcal{G}))$.*
9. *If $\sigma(X)$ is independent of $\mathcal{G}$ then $\mathbb{E}(X|\mathcal{G}) = \mathbb{E}X$.*

Definition A.10. *Let X and Y be two random variables. The* conditional expectation of X given Y *is defined as $\mathbb{E}(X|Y) := \mathbb{E}(X|\sigma(Y))$, that is, it is the conditional expectation of X given the σ-algebra generated by Y.*

Definition A.11. *Let X a random variables and $A \in \mathcal{F}$ an event. The* conditional expectation of X given A *is defined as $\mathbb{E}(X|A) := \mathbb{E}(X|\sigma(A))$. This means it is the conditional expectation of X given the sigma algebra generated by A i.e. $\mathbb{E}(X|A) := \mathbb{E}(X|\{\emptyset, A, A^c, \Omega\})$.*

We can immediately see that $\mathbb{E}(X|A) = \mathbb{E}(X|\mathbb{1}_A)$.

Recall that if X and Y are jointly continuous random variables with joint density $(x, y) \mapsto f(x, y)$ then for any measurable function $\rho : \mathbb{R}^2 \to \mathbb{R}$ such that $\mathbb{E}|\rho(X, Y)| < \infty$ we have
$$\mathbb{E}\rho(X, Y) = \int_{\mathbb{R}}\int_{\mathbb{R}} \rho(x, y)f(x, y)dydx.$$
Moreover the marginal density of X is
$$g(x) = \int_{\mathbb{R}} f(x, y)dy$$
while the marginal density of Y is
$$h(y) = \int_{\mathbb{R}} f(x, y)dx.$$

Theorem A.12. *Let X and Y be jointly continuous random variables with joint density $(x, y) \mapsto f(x, y)$. Then for any measurable function $\varphi : \mathbb{R} \to \mathbb{R}$ such that $\mathbb{E}|\varphi(Y)| < \infty$ the conditional expectation of $\varphi(Y)$ given X is*

$$\mathbb{E}(\varphi(Y)|X) = \psi(X)$$

where $\psi : \mathbb{R} \to \mathbb{R}$ is given by

$$\psi(x) = \mathbb{1}_{\{g(x)>0\}} \frac{\int_{\mathbb{R}} \varphi(y) f(x, y) dy}{g(x)}.$$

Proof. Every A in $\sigma(X)$ must be of the form $A = \{\omega \in \Omega : X(\omega) \in B\}$ for some B in $\mathcal{B}(\mathbb{R})$. We need to show that for any such A

$$\int_A \psi(X) d\mathbb{P} = \int_A \varphi(Y) d\mathbb{P}.$$

But since $\mathbb{E}|\varphi(Y)| < \infty$ we can use Fubini's theorem to show that

$$\begin{aligned}
&\int_A \psi(X) d\mathbb{P} = \mathbb{E}\mathbb{1}_A \psi(X) = \mathbb{E}\mathbb{1}_{\{X\in B\}} \psi(X) = \int_B \psi(x) g(x) dx \\
&= \int_B \int_{\mathbb{R}} \varphi(y) f(x, y) dy dx = \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbb{1}_B(x) \varphi(y) f(x, y) dx dy \\
&= \mathbb{E}\mathbb{1}_{\{X\in B\}} \varphi(Y) = \int_A \varphi(Y) d\mathbb{P}.
\end{aligned}$$

□

A.4 Multivariate normal distribution

There are a number of ways how to define a multivariate normal distribution. See e.g. [5, Chapter 5] for a more definite treatment. We will define a multivariate normal distribution as follows. Let $\mu \in \mathbb{R}^d$ be given and let Σ be a given symmetric, invertible, positive definite $d \times d$ matrix (it is also possible to consider positive semi-definite matrix Σ but for simplicity we ignore that situation here).

A matrix is positive definite if, for any $x \in \mathbb{R}^d$ such that $x \neq 0$, the inequality $x^T \Sigma x > 0$ holds. From linear algebra we know that this is equivalent to:

1. The eigenvalues of the matrix Σ are all positive.
2. There is a unique (up to multiplication by -1) matrix B such that $BB^T = \Sigma$.

Let B be a $d \times k$ matrix such that $BB^T = \Sigma$.

Let $(X_i)_{i=1}^d$ be independent random variables with $N(0, 1)$ distribution. Let $X = (X_1, \ldots, X_d)^T$ and $Z := \mu + BX$. We then say $Z \sim N(\mu, \Sigma)$ and call Σ the covariance matrix of Z.

Exercise A.1. Show that $\text{Cov}(Z_i, Z_j) = \mathbb{E}((Z_i - \mathbb{E}Z_i)(Z_j - \mathbb{E}Z_j)) = \Sigma_{ij}$. This justifies the name "covariance matrix" for Σ.

It is possible to show that the density function of $N(\mu, \Sigma)$ is

$$f(x) = \frac{1}{(2\pi)^{d/2}\sqrt{\det(\Sigma)}} \exp\left(-\frac{1}{2}((x-\mu)^T\Sigma^{-1}(x-\mu))\right). \tag{50}$$

Note that if Σ is symmetric and invertible then Σ^{-1} is also symmetric.

Exercise A.2. You will show that $Z = BX$ defined above has the density f given by (50) if $\mu = 0$.

i) Show that the characteristic function of $Y \sim N(0,1)$ is $t \mapsto \exp(-t^2/2)$. In other words, show that $\mathbb{E}(e^{itY}) = \exp(-t^2/2)$. *Hint.* complete the squares.

ii) Show that the characteristic function of a random variable Y with density f given by (50) is
$$\mathbb{E}\left(e^{i(\Sigma^{-1}\xi)^T Y}\right) = \exp\left(-\frac{1}{2}\xi^T\Sigma^{-1}\xi\right).$$
By taking $y = \Sigma^{-1}\xi$ conclude that
$$\mathbb{E}\left(e^{iy^TY}\right) = \exp\left(-\frac{1}{2}y^T\Sigma^{-1}y\right).$$
Hint. use a similar trick to completing squares. You can use the fact that since Σ^{-1} is symmetric $\xi^T\Sigma^{-1}x = (\Sigma^{-1}\xi)^T x$.

iii) Recall that two distributions are identiacal if and only if their characteristic functions are identical. Compute $\mathbb{E}\left(e^{iy^TZ}\right)$ for $Z = BX$ and $X = (X_1, \ldots, X_d)^T$ with $(X_i)_{i=1}^d$ independent random variables such that $X_i \sim N(0,1)$. Hence conclude that Z has density given by (50) with $\mu = 0$.

You can now also try to show that all this works with $\mu \neq 0$.

A.5 Stochastic Analysis Details

The aim of this section is to collect technical details in stochastic analysis needed to make the main part of the notes correct but perhaps too technical to be of interest to many readers.

Definition A.13. *We say that a process X is called* progressively measurable *if the function $(\omega, t) \mapsto X(\omega, t)$ is measurable with respect to $\mathcal{F}_t \otimes \mathcal{B}([0,t])$ for all $t \in [0,T]$.*

We will use $Prog_T$ to denote the σ-algebra generated by all the progressively measurable processes on $\Omega \times [0,T]$.

If X is progressively measurable then the processes $\left(\int_0^t X(s)ds\right)_{t\in[0,T]}$ and $(X(t \wedge \tau))_{t\in[0,T]}$ are adapted (provided the paths of X are Lebesgue integrable and provided τ is a stopping time). The important thing for us is that any left (or right) continuous adapted process is progressively measurable.

A.6 More Exercises

Exercise A.3. Say $f : \mathbb{R} \to \mathbb{R}$ is smooth and $W = (W(t))_{t\in[0,T]}$ is a Wiener process. Calculate

$$\mathbb{E}\left[f'(W(T))W(T)\right].$$

References

[1] T. Björk. Arbitrage Theory in Continous Time. Oxford University Press, 2009.

[2] D. Brigo and F. Mercurio. *Interest Rate Models - Theory and Practice*, Springer, 2006.

[3] F. Delbaen and W. Schachermayer. A general version of the fundamental theorem of asset pricing. *Math. Ann.* 300(3), 463–520, 1994.

[4] D. Filipovic. *Term Structure Models*. Springer, 2009.

[5] A. Gut. *An Intermediate Course in Probability*. Springer 2009.

[6] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer 1991.

[7] A. N. Kolmogorov and S. V. Fomin. *Introductory Real Analysis*. Dover, 1975.