Interest Rates and Credit Risk Models

Juri Hinz
University of Technology Sydney
E-Mail: juri.hinz@uts.edu.au

This module is designed for graduate students in quantitative finance. It focuses on advanced topics in interest rate theory and credit risk modelling and emphasizes their analogies. The module covers the following major topics. Products of fixed-income markets, Short rate models, Heath-Jarrow-Morton framework, LIBOR market models. Financial instruments in credit risk management, Models of default: Firm value and first passage time models, intensity based models, models of credit rating migrations. The module also provides a discussion of advantages and shortcomings of synthetic credit-linked instruments; moreover, modeling dependence structure of default events and default contagion will be treated.

Lectures: Saturdays 10:45 - 12:15

Zoom meeting

`https://us02web.zoom.us/j/83495308737?pwd=dndaQis5ekl2UkpoS0hvblNmNnhSUT09`

Meeting ID: 834 9530 8737 Passcode: 873564

Google drive:

`https://drive.google.com/drive/folders/1KMGCVazfPkvNUZdY-PFmAmRR-V0DSOi0?usp=sharing`

Books:

- *Interest Rate Modeling* L. B. G. Andersen, V. V. Piterbarg
  `https://disk.yandex.ru/i/4QMOAz1cho2zLA`
- *Stochastic Calculus for Finance II: Continuous-Time Models.* Steven E. Shreve
- *Arbitrage Theory in Continuous Time.* Thomas Björk
- *Credit Derivatives Pricing Models: Model, Pricing and Implementation.* Philipp J. Schönbucher.

1 Basic notions of fixed-income markets

1.1 Constant interest rates. Most bank accounts and loans give *compound interest rate*. If the compound annual interest rate $r \in \mathbb{R}$ is payed once a year, then the wealth W grows over the years $n = 1, 2, \ldots,$ as

$$(1+r)^n W, \qquad n = 1, 2, 3, \ldots.$$

This is an exponential increase. There is a seven-ten rule associated with this.

$$\textit{Money invested at } r = 7\% \textit{ doubles in ten years } (1.07)^{10} \approx 1.97$$
$$\textit{Money invested at } r = 10\% \textit{ doubles in seven years } (1.1)^{7} \approx 1.94$$

Compounding may be effected several times a year (for instance four times (quarterly), or twelve times (monthly)). However, it is tradition to quote the interest rate on yearly basis (for comparison reasons). This creates different notions of interest rate.

- The yearly interest rate r is compounded m times a year is called *the nominal rate*, it has the effect that the wealth W will grow over the years $n = 1, 2, \ldots,$ as

  $$((1+\frac{r}{m})^m)^n W, \qquad n = 1, 2, 3, \ldots.$$

  since $(1+\frac{r}{m})^m > (1+r)$, this is not the same as

  $$(1+r)^n W, \qquad n = 1, 2, 3, \ldots.$$

- To compare nominal rate r (compounded m times a year) to a yearly interest rate compounded once a year, the notion of *effective interest rate* is introduced. Effective interest rate $\tilde{r}$ is the yearly interest rate compounded once a year, which has the same effect

  $$1 + \tilde{r} = (1 + r/m)^m$$

  For instance $(1 + 0.08/4)^4 = 1.0824$, (nominal rate $r = 0.08$ compounded quarterly, yields effective rate of 0.0824)

Continuous compounding is obtained in the limit $m \to \infty$.

$$1 + \tilde{r} = \lim_{m\to\infty} (1 + r/m)^m = e^r$$

Say, if the nominal interest rate is 0.08 per year, continuously compounded, then the effective rate is 0.0833, since $e^{0.08} \approx 1.0833$.

Agents must take interest rate effects when investment decisions are made. More precisely, one needs concepts that consider interest rates for project valuation. Two different approaches are known

- net present value
- internal rate of return

A project is associated with a sequence of cash streams $(x_k)_{k=0}^n$ (deterministic, in this setting) effected at dates $k = 0, 1, \ldots, n$. For instance, in lumber business, one needs to invest in plantation, to wait until lumber can be harvested and to chose the right time for harvest. Possible cash streams could be

$$\begin{aligned} (x_k)_{k=0}^1 &= (-1, 2) \qquad \text{cut early} \\ (x_k)_{k=0}^2 &= (-1, 0, 3) \qquad \text{cut later} \end{aligned}$$

Under the net present value, the project associated with cash streams $(x_k)_{k=0}^n$ is assigned the present value

$$\text{PV} = \sum_{k=0}^{n} \frac{x_k}{(1+r)^k} \tag{1.1}$$

where r stands for the annual interest rate relevant for this project. Based on net present value, investment decisions are made: if net present value is positive, then the project could be realized, the higher the net present value the more valuable is the project.

Note that the net present value criterion is in line with the time value of money. Namely, in the presence of an ideal bank which provides the same interest rate r for lending and borrowing in each of the periods

$$[0, 1], [1, 2], \ldots, [n-1, n],$$

the present value of the cash flow x_k provided at the time $k = 1, \ldots, n$ (after k periods from now) must be $x_k/(1+r)^k$. Hence, present value of the cash stream $(x_k)_{k=0}^n$ is (1.1)

Surprisingly, the other method, the method of internal rate of return does not take into account the market interest rate. Instead, one calculates a hypothetical interest rate, which makes the net present value of the project equal to zero. Thus, the internal rate of return is the number r_I such that

$$0 = \sum_{k=0}^{n} \frac{x_k}{(1+r_I)^k} \tag{1.2}$$

Based on internal rate of return, investment decisions are made: if internal rate of return is higher than the actual interest rate, then the project could be realized, the higher the internal rate of return, the more valuable is the project.

The existence and uniqueness of positive solution r_I to (1.2) is ensured only under additional assumption that

first stream is negative $x_0 < 0$, all subsequent
streams are non-negative $x_k \geq 0$, $k = 1, \ldots, n$
and at least one of them is positive:
$x_k > 0$ for some $k \in \{1, \ldots, n\}$

under this condition, the continuous function

$$f : c \mapsto \sum_{k=0}^{n} x_k c^k$$

satisfies $f(0) = x_0 < 0$, the limit is $\lim_{c\uparrow+\infty} f(c) = +\infty$. By intermediate value theorem, there is $c_0 = \frac{1}{1+r_I}$ with $f(c_0) = 0$. This positive solution is unique since on $[0, \infty[$ the function f is increasing.

Example Let us show how both project valuation criteria are applied. We consider the tree project again.

Assume that $r = 0.1$, the net present value yields

$$\begin{aligned} \text{NPV} &= -1 + \frac{2}{1.1} = 0.82 \qquad \text{cut earlier,} \\ \text{NPV} &= -1 + \frac{3}{1.1^2} = 1.48 \qquad \text{cut later.} \end{aligned}$$

Hence, under net present value criterion it is better to cat later. Consider now the internal rate of return. Here we have to solve the equations

$$\begin{aligned} 0 &= -1 + 2c \qquad \Rightarrow c = \frac{1}{2} \qquad \Rightarrow r_I = 1 \qquad \text{cut earlier} \\ 0 &= -1 + 3c^2 \qquad \Rightarrow c = \frac{1}{\sqrt{3}} \qquad \Rightarrow r_I = \sqrt{3} - 1 \approx 0.72 \qquad \text{cut later} \end{aligned}$$

Hence, under internal rate of return, it is better to cut earlier.

Remark There is a debate which of two criteria is more appropriate. It seems that net present value is better. If used intelligently, it will provide consistency and rationality. One should take care on

- which kind of interest rate applies to the project
- rates for borrowing and lending may be different
- how much capital the project requires (project with the same net present value but less initial investment should be preferred)

1.2 Market for future cash. There is a huge number of opportunities in the market to transfer today's wealth to a future date in a riskless way. Some of these opportunities are realized by purchase of a financial instrument (basically an asset, which promises to pay a fixed amount of cash on a later date). Some of them are realized on bilateral basis like savings deposits (say put money into bank account and to wait). Let us point out some features of the market for future cash.

Mortgages and annuities Consider a hypothetical fixed income instrument which pays a fixed amount A at the end of each period, forever. This instrument is called perpetual annuity. For instance, the owner of such instrument may receive 1,000 USD at the first of each month, forever. (In fact, such instruments do exist in GB, and are called consols).

The present value of the perpetual annuity is

$$\sum_{k=1}^{\infty} \frac{A}{(1+r)^k} = \frac{A}{r}$$

where $r > 0$ stands for the interest rare relevant for the corresponding period. Explanation: for $q = (1+r)^{-1} \in [0, 1[$ holds

$$\sum_{k=1}^{\infty} q^k = \frac{q}{1-q} = \frac{1}{\underbrace{1/q}_{1+r} - 1} = \frac{1}{r}$$

Similarly, if the annuity terminates at the end of the nth period, we have

$$\begin{aligned} P &= \sum_{k=1}^{n} \frac{A}{(1+r)^k} = \sum_{k=1}^{\infty} \frac{A}{(1+r)^k} - \frac{1}{(1+r)^n} \sum_{k=1}^{\infty} \frac{A}{(1+r)^k} \\ &= \frac{A}{r}(1 - \frac{1}{(1+r)^n}) \end{aligned}$$

Based on this, one is able to calculate the amount A given the present price P of the annuity

$$A = \frac{r(1+r)^n P}{(1+r)^n - 1}.$$

This formula could be used to calculate the fixed amount A one needs to pay periodically to repay the loan P until n.

Savings deposits are offered by commercial banks, savings and loan institutions. Usually, there is some guarantee for the savings account.

- *demand deposit* pays a rate of interest which varies with market conditions.
- *time deposit account* pays a guaranteed interest rate. The deposit must be maintained for an agreed time period (6 months), otherwise a penalty for early withdrawal is assessed.

- *certificate of deposit (CD)* is similar to time deposit account, it is standardized in denominations of 10,000 USD. Large-denomination CDs can be sold in a market, hence CDs qualify as tradable securities.

- *Eurodollar deposit* are like CD's nominated in US-dollars but held onside US. They are issued by banks outside of US. The difference is in banking regulations and insurance.

Among tradable securities in the money market, the most important are bonds. A bond is a financial instrument, which gives its owner the right to receive

- at maturity date a pre-determined amount of cash which is called *face value* or *par value*,

- prior to maturity, there are pre-determined dates (in regular intervals, last coincides with maturity) at which the owner of the bond receives a pre-determined amount cash which is called *coupon*. Usually, the amount is a fixed fraction of the face value. This fraction is called *coupon rate*.

If the coupon payment is zero, then the bond is called zero-coupon bond, or zero-bond. Observe that each (coupon-paying) bond can be considered as a basket of zero bonds. For this reason, zero bonds turn out to be the basic building block of the entire money market theory.

There are many different types of bonds. Some of them are issued by institutions which may fail to fulfill their payment obligations. This type of assets includes corporate bonds, and bonds issued by some governments. We do not consider such bonds here since, strictly speaking, that are risky investments.

Riskless bonds are securities issued by agents that do pay back their debts. US government is among of them.

- *US Treasury bills* are issued with denominations of 10,000 (face value) with terms to maturity of 13, 26, and 52 weeks. No coupon payment is made.

- *US Treasury notes* are issued with small denominations, as 1,000. They have long maturities, 1 to 10 years. The coupon payments are every 6 months until maturity.

- *US Treasury bonds* are similar to treasury notes, but some of them are *callable*, meaning that at some scheduled coupon payment date the Treasury can force the bond holder to redeem the bond at that time for its face value.

- *US Treasury strips* are similar to treasury notes, but each coupon is represented by a zero bond and can be sold separately.

Bonds in more details Remember that the bond pays to its holder the face (par) value at maturity date. Additionally, the holder receives periodical payments, which are called coupons. In the past, actual coupons where attached to the bond certificate, to be mailed to the issuer (bank) to receive payments. Coupon payments are usually percentage of the face value. For instance, the 9 % coupon bond with face 1,000 value will have coupon payment of 90 per year. However, the period between coupon payments may be fraction of one year, for instance each 6 months. In this example, that will be 45 each half a year.
Bonds are issued to rise the capital. Bond are bought for riskless transfer of the capital to a later date. At any time, bond can be sold at the market, where demand and supply determine its price. Each agent must take all bond details when price is considered. To make a price comparison easier, the *yield* of the bond is calculated and quoted.

Consider the following table

Rate	Maturity	Bid	Asked	Yield (ask)
$4\frac{3}{4}$	Feb 97	99:31	100:00	4.64
$6\frac{3}{4}$	Feb 97	100:00	100:02	4.98
$\vdots$				

- The first column shows the coupon rate, say $6\frac{3}{4}$ percent of the face value paid per year (half of which is paid each six months).

- The second column shows maturity date

- The third column shows bid price quoted in percentage of the face value, after : are 32nd of basis point. For instance 99 : 31 mans that someone offers to buy this bond at the price
$$\frac{99\frac{31}{32}}{100} \times \text{face value}.$$

- The fourth column shows ask price quoted in percentage of the face value, after : are 32nd of basis point. For instance 100 : 02 mans that someone offers to sell this bond at the price
$$\frac{100\frac{2}{32}}{100} \times \text{face value}.$$

- the last column shows the so-called yield of the bond. This is nothing but the internal rate of return associated with the cash-stream provided by bond, where it is assumed that you buy it at the ask price. (Which ensures that the first cash flow is negative, all subsequent are positive. Internal rate of return is uniquely determined.)

The actual amount must be paid for the bond is higher than quoted, including the so-called accrued interest. This is to reimburse the owner for the lost coupon payment since last coupon payment date. Accrued interest is

$$\frac{\text{number of days since last coupon payment date}}{\text{number of days in the current coupon period}} \times \text{ coupon amount.}$$

Let us take a closer look how the yield is calculated.

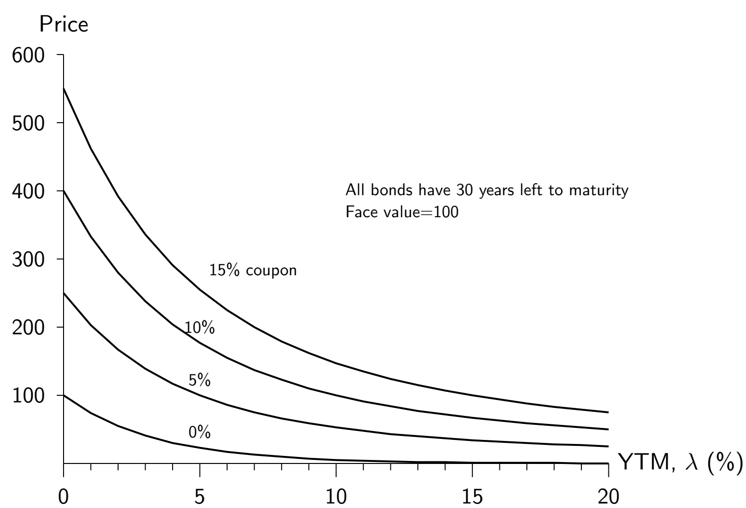
Suppose that the bond with the face value F makes $m \in \underset{\sim}{\mathbb{N}}$ coupon payment of C/m each per year and there are n periods remaining. If the current price of the bond is P then the yield to maturity is the value λ which satisfies

$$\begin{aligned} P &= \frac{F}{(1+\lambda/m)^n} + \sum_{k=1}^{n} \frac{C/m}{(1+\lambda/m)^k} \\ &= \frac{F}{(1+\lambda/m)^n} + \frac{C/m}{\lambda/m}\left(1 - \frac{1}{(1+\lambda/m)^n}\right) \end{aligned}$$

As mentioned earlier, this equation admits a unique positive solution, which must be obtained numerically. Let us discuss dependance of the yield λ on the bond price P.

Since bond is a complex product, the yield alone is not appropriate to determine whether the bond is can be bought at the offered price. However, most investors would not buy a bond at 6% yield if bank CDs for relevant maturities are offered, say at 10 %.

For typical bond examples (with different coupon rates but the same maturity), the connection between yield and price is illustrated in yield-price diagram. Note the following



- Bonds price is a decreasing and convex function of yield, approaching zero.
- If the yield is fixed, the higher coupon rate, the higher the price

- If the yield is zero, then the price equals the total payment of the bond
- If the yield is equal to the coupon rate, then the bond price equals to bonds face value (bond is at par)

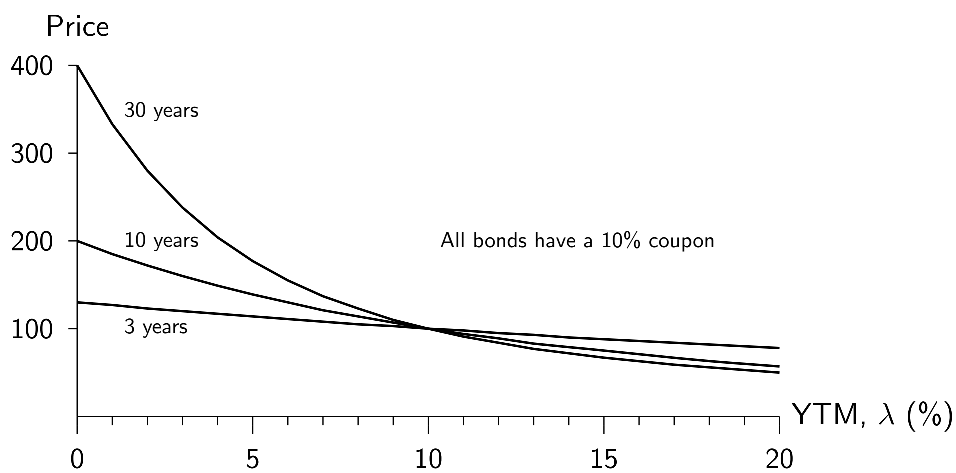
To see the last point we note that if $C = \rho F$ where $\rho > 0$ is the coupon rate paid each year in m fractions, then the above formula is

$$P = \frac{F}{(1+\lambda/m)^n} + \frac{F\rho}{\lambda}(1 - \frac{1}{(1+\lambda/m)^n})$$

For $\lambda = \rho$, we have

$$P \quad = \quad \frac{F}{(1+\rho/m)^n} + \frac{F\rho}{\rho}(1 - \frac{1}{(1+\rho/m)^n}) = F.$$

Now compare bonds with the same coupon rate but different maturities



- The common intersection is the point where the bond is traded at par (coupon rate equals the yield)
- The longer the maturity, the steeper the curve (which means that price is more sensitive to yield for long-maturity bonds)
- If yield is below coupon rate, then the longer maturity, the higher the price
- If yield is above coupon rate, then the longer maturity, the lower the price

1.3 Zero bonds and yield. Until now, the yield is characterized as a function of bond price (calculated numerically). However it is important to understand that yields of different bonds quantitatively reflect the relevant interest rate paid at the market. When interest rate changes, all bond prices change, but this change will show particular pattern, where new yields reflect the new interest rate. For this reason, practitioners tend to see yields as the true explanatory variable of bond prices, the movement of the yields has to be used as the true driver of bond prices.

However, the precise description of this mechanism is not obvious. The first step could be to visualize the recent yields for bonds. This results in the concept of the yield curve. The *yield curve* is a graphical plot of yields as function of time to maturity, depicted for particular type of bonds. (For instance, for bonds with the same coupon rate).

As mentioned earlier, zero bonds can be considered as basic building blocks of each type of coupon-paying bond. For this reason, the most convenient yield curve would be that of zero bonds. Let us see how it could be obtained. Suppose that we observe at time $t = 0$ the prices of zero bonds listed at the market. Assume that there are zero bonds with different maturities $\tau \in \{\tau_1 < \ldots, < \tau_n\} \subset \mathbb{R}_+$ with the same face value F whose prices are listed

$$B_0(\tau_1), \;\; B_0(\tau_2), \ldots, B_0(\tau_n).$$

For each bond maturing at $\tau_j \in \{\tau_1, \ldots, \tau_n\}$ calculate its yield $Y_0(\tau_j)$. Here, one needs to take care on compounding convention

- For continuous compounding, the yield $Y_0(\tau_j)$ satisfies

$$e^{Y_0(\tau_j)\tau_j} = F/B_0(\tau_j), \qquad j = 1, \ldots, n$$

- For discrete compounding, the yield $Y_0(\tau_j)$ satisfies

$$\begin{gathered}(1 + Y_0(\tau_j)/m)^{n_j} = F/B_0(\tau_j) \qquad j = 1, \ldots, n \\ n_j \text{ number of compound times within }]0, \tau_j]\end{gathered}$$

As result, one obtains a sequence

$$Y_0(\tau_1), Y_0(\tau_2), \ldots, Y_0(\tau_n)$$

indexed by time to maturity $\tau_1 < \tau_2 <, \ldots, < \tau_n$ which is plotted in a diagram. This gives the yield curve at time $t = 0$.

As the time evolves, zero bond prices change, some of zero bonds expire, new zero bonds are introduced. This will result in a different bond curve. Say, at time $t > 0$ we observe bond prices with the same face value F whose prices are available

$$B_t(\tau'_1), \;\; B_t(\tau'_2), \ldots, B_t(\tau'_{n'})$$

For each bond maturing at $\tau_j' \in \{\tau_1' < \ldots, < \tau_n'\}$, calculate its yield $Y_t(\tau_j')$. Again, take care on compounding convention

- For continuous compounding, the yield $Y_t(\tau_j)$ satisfies

$$e^{Y_t(\tau_j')(\tau_j'-t)} = F/B_t(\tau_j), \qquad j = 1, \ldots, n$$

- For discrete compounding, the yield satisfies

$$(1 + Y_t(\tau_j')/m)^{n_j} = F/B_t(\tau_j') \qquad j = 1, \ldots, n$$
$$n_j \text{ number of compound times within }]t, \tau_j']$$

As result, one obtains a sequence

$$Y_t(\tau_1'), \quad Y_t(\tau_2'), \ldots, Y_t(\tau_{n'}).$$

Indexed by time to maturity $\tau_1' - t < \tau_2' - t <, \ldots, < \tau_{n'}' - t$, we plot it in a diagram. This gives the yield curve at time t.

Such a procedure illustrates the yield curve at any time t. A typical outcome is shown in the Figure 1.

The dependence of zero bonds yields on maturity time is frequently referred to as the *term*

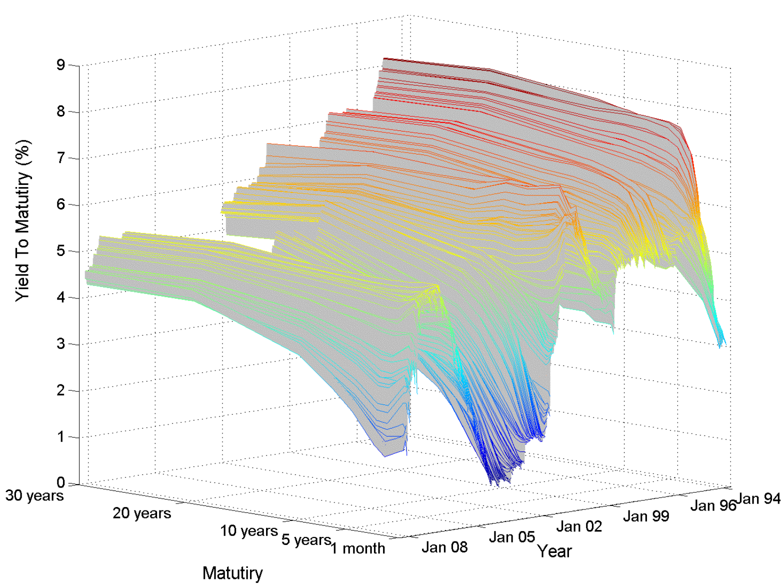


Figure 1: Yield evolution against bond's maturity

structure of interest rates. (In fact, yield curve represents *term structure*).

Let us point out typical features of the yield curve

- Generally, yields rise as the time to maturity increases (this is the so-called *normal yield curve*)

- At the beginning, the rise could be steep, followed by moderate increase

- When long-maturity bonds have lower yields than short-maturity bonds (decreasing yield curve), the curve is called *inverse yield curve.*

- When the normal yield curve curve is flattening (during its evolution), this is possible signal for a soon reversion (transition from normal to inverse curve).

The modeling of random evolution of the entire yield curve $t \mapsto (Y_t(\tau))_{\tau \geq t}$ is subject of *interest rate theory.* The basic attempt there is to establish such dynamics that the corresponding zero bond prices

$$(B_t(\tau))_{t \in [0,\tau]}, \qquad \tau \in [0, T]$$

follow an arbitrage-free price evolution.

1.4 Zero bonds determine time value of money. Consider future dates

$$0 = \tau_1 < \cdots < \tau_n.$$

Suppose that the prices

$$B_0(\tau_1),\ \ B_0(\tau_2), \ldots, B_0(\tau_n)$$

of zero bonds with face value F and maturities $\tau_1, \tau_1, \ldots, \tau_n$ are listed. Calculate the discount factors

$$d_0(\tau_k) = \frac{B_0(\tau_k)}{F}, \qquad k = 1, \ldots, n$$

The factor $d_0(\tau_k)$ is interpreted as the unique no-arbitrage price X_0 of a security whose payoff X_{τ_k} at maturity τ_k equals to one $X_{\tau_k} = 1$ (normalized zero bond).

- If such a security is listed below $X_0 < d_0(\tau_k)$, buy this security (pay X_0) and sell $1/F$ zero bonds with maturity τ_k (obtain $d_0(\tau_k)$). Pocket money $d_0(\tau_k) - X_0 > 0$. At τ_k, the positions annihilate

  $$\underbrace{X_{\tau_k}}_{=1} - \underbrace{B_{\tau_k}(\tau_k)}_{=F}/F = 0.$$

  Hence, such a strategy is an arbitrage.

- If the security is listed above $X_0 > d_0(\tau_k)$ sell this security (obtain X_0) and buy $1/F$ zero bonds with maturity τ_k (pay $d_0(\tau_k)$). Pocket money $X_0 - d_0(\tau_k) > 0$. At τ_k, the positions annihilate

  $$\underbrace{B_{\tau_k}(\tau_k)}_{=F}/F - \underbrace{X_{\tau_k}}_{=1} - = 0.$$

  Hence, such a strategy is an arbitrage.

By linearity of the pricing rule (which must hold due to no-arbitrage), any instrument promising deterministic cash flow

$$Z_{\tau_1},\ \ Z_{\tau_2},\ldots,Z_{\tau_n} \in \underline{\mathrm{R}} \qquad \text{at the dates } \tau_1,\tau_2,\ldots,\tau_n$$

must have the price today

$$\sum_{k=1}^{n} d_0(\tau_k) Z_{\tau_k} = \sum_{k=1}^{N} \frac{B_0(\tau_k)}{F} Z_{\tau_k}$$

Indeed, it is equivalent to the portfolio of zero bonds (face value F) with maturity dates $(\tau_k)_{k=1}^n$ and weights $\frac{Z_{\tau_k}}{F}$, $\quad k = 1,\ldots,n$. In particular, having zero bond prices for all relevant maturities, the prices of all coupon paying bonds follow since coupon bond is equivalent to appropriate portfolio of zero bonds. In practice, observed coupon bond prices can be used to infer on prices of zero bonds, which are not directly observable.

To determine the yield curve, we have to observe bond prices $B_0(\tau)$ for as many maturities $\tau \geq 0$ as possible. In real markets, zero bonds are listed for standardized maturity times only. Moreover, for long maturities zero bonds are not popular (may not be listed). This may result in a rare data set. Remember that zero bonds are building blocks of coupon bonds. Even if there is no zero bond with maturity time τ, it is sometimes still possible to determine the yield $Y_0(\tau)$ and the price $B_0(\tau)$ of a zero bond from coupon-paying bonds.

One begins with zero bond yield for short maturities and works forward towards long maturities.

Example Suppose yearly compounding. We have determined the yield s_1 for the zero bond maturing in one year from now with face value F^1 and bond price P^1 as

$$P^1 = \frac{F^1}{1+s_1} \quad \Rightarrow \quad s_1 = \frac{F^1}{P^1} - 1.$$

Now, we would like to determine the yield s_2 for the zero bond maturing in two years from now. There is no such bond in the market. Instead, we have a coupon bond with yearly payments of coupon C^2 with face value F^2 and price P^2 maturing in two years from now. Its price is

$$P^2 = \frac{C^2}{1+s_1} + \frac{C^2+F^2}{(1+s_2)^2}$$

a simple re-arrangement gives s_2

$$s_2 = (\frac{C^2+F^2}{P^2 - C^2/(1+s_1)})^{\frac{1}{2}} - 1.$$

By next, consider the coupon bond with yearly coupon payments maturing in three years from now to calculate s_3. Proceeding recursively, one finds yields of zero bonds with maturity times $\tau = 1,2,3$ although their prices are not listed at the exchange.

Example Another way to calculate zero bond prices is to use coupon bonds with the same maturity, same face value, and same coupon payment dates but different coupons. Say P^i is the price of a bond with face value F^i and coupon payment C^i, $i = 1, 2$ then the portfolio consisting in a long position in the first bond and short position in C^1/C^2 of the second bond yields effectively zero bond with face value $F^1 - F^2(C^1/C^2)$ the price of this portfolio is

$$P^1 - P^2 \frac{C^1}{C^2}.$$

If $F^1 - F^2(C^1/C^2) \neq 0$, then one obtains the price of zero bond with the corresponding maturity and finally calculates its yield. Here a numerical example.

Example Suppose both bonds have ten years to maturity and face value of $F = 100$. First bond pays 8% coupon rate ($C^1 = 8$) and the second pays 10% coupon rate ($C^2 = 10$). The price of the first bond is 85.89 and of the second is 98.72. Form a portfolio with first bond and -0.8 of second bond. This is a zero ten-year zero bond with face value 20. The price is

$$85.89 - 0.8 \cdot 98.72 = 6.914.$$

The yearly compounded yield s_{10} for ten-years zero bond is

$$(1 + s_{10})^{10} = \frac{20}{6.914} \qquad \Rightarrow \qquad s_{10} = (\frac{20}{6.914})^{\frac{1}{10}} - 1 \approx 0.11$$

1.5 Forward and spot rates. For simplicity, we assume that all zero bonds have face value $F = 1$. Observe today $t = 0$ the price of zero bond $B_0(\tau)$ with maturity date τ. The opportunity, for each currency unit invested at $t = 0$ to obtain at later date $1/B_0(\tau)$ currency units, can be considered as fixed-term deposit account paying constant and deterministic interest rate. However, it applies for money which is borrowed (lent) now and returned (received) at the later date τ. This leads to the definition of *spot rate* $s(\tau)$

- For continuous compounding,

$$e^{\tau s(\tau)} = B_0(\tau)^{-1}$$

- For discrete compounding,

$$(1 + s(\tau)/m)^n = B_0(\tau)^{-1}$$
$$n \text{ number of compound dates within }]0, \tau]$$

We also can schedule a deal today to obtain for each currency unit invested at $\tau_1 \geq 0$ $B_0(\tau_1)/B_0(\tau_2)$ currency units delivered at the later date $\tau_2 \geq \tau_1$. The deal for one currency unit is as follows:

- At time $t = 0$ (now) sign a short position in one zero bond maturing at τ_1, receive $B_0(\tau_1)$ invest this money to buy $B_0(\tau_1)B_0(\tau_2)^{-1}$ bonds maturing at time τ_2. The balance at the beginning is zero:

$$B_0(\tau) - B_0(\tau_1)B_0(\tau_2)^{-1} \cdot B_0(\tau_2) = 0$$

- At time $t = \tau_1$ receive the currency unit to be invested. Close short position from zero bond maturing at τ_1. Wait until τ_2.
- At time τ_2, the long fractional position in zero bond maturing at τ_2 pays $B_0(\tau_1)B_0(\tau_2)^{-1}$ currency units.

In principle, we can guarantee today a constant and deterministic interest rate for money, which is borrowed (lent) at a future time τ_1 and returned (received) at a later time τ_2. This leads to the definition of the *forward rate* $f(\tau_1, \tau_2)$

- For continuous compounding,

$$e^{(\tau_2-\tau_1)f(\tau_1,\tau_2)} = B_0(\tau_1)B_0(\tau_2)^{-1}$$

- For discrete compounding m times a year,

$$(1 + f(\tau_1, \tau_2)/m)^n = B_0(\tau_1)B_0(\tau_2)^{-1}$$
$$n \text{ number of compounding dates within }]\tau_1, \tau_2]$$

Observe that with this definition,

- forward rate where first time is now $\tau_1 = 0$ equals to the spot rate by

$$f(0, \tau) = s(\tau) \qquad \tau \geq 0.$$

  Namely, for continuous compounding we have

$$e^{(\tau_2-0)f(0,\tau_2)} = \underbrace{B_0(0)}_{=1} B_0(\tau_2)^{-1} = e^{\tau_2 s(\tau_2)}$$

  and for discrete compounding we have

$$(1 + f(0, \tau_2)/m)^{n_2} = \underbrace{B_0(0)}_{=1} B_0(\tau_2)^{-1} = (1 + s(\tau_2)/m)^{n_2}$$
$$n_2 \text{ number of compounding dates within }]0, \tau_2]$$

- forward rate can be expressed through spot rate. For continuous compounding, we have

$$f(\tau_1, \tau_2) = \frac{s(\tau_2)\tau_2 - s(\tau_1)\tau_1}{\tau_2 - \tau_1} \tag{1.3}$$

  whereas for discrete compounding (m times a year) we have

$$f(\tau_1, \tau_2) = m\left(\left[\frac{(1 + s(\tau_2)/m)^{n_2}}{(1 + s(\tau_1)/m)^{n_1}}\right]^{\frac{1}{n}} - 1\right) \tag{1.4}$$

  where n_2 are compound times within $]0, \tau_2]$, n_1 are compound times within $]0, \tau_1]$ and $n = n_2 - n_1$ are compound times within $]\tau_1, \tau_2]$.

For continuous compounding, the expression (1.3) follows from

$$\begin{aligned} e^{(\tau_2-\tau_1)f(\tau_1,\tau_2)} &= B_0(\tau_1)B_0(\tau_2)^{-1} = e^{-\tau_1 s(\tau_1)}e^{\tau_2 s(\tau_2)} \\ &= e^{\tau_2 s(\tau_2)-\tau_1 s(\tau_1)} \end{aligned}$$

For discrete compounding, the expression (1.4) follows from

$$(1+f(\tau_1,\tau_2)/m)^n = B_0(\tau_1)B_0(\tau_2)^{-1} \;=\; (1+s(\tau_1)/m)^{-n_1}(1+s(\tau_2)/m)^{n_2}.$$

Example Suppose yearly compounding and assume that $s_1 = 7\%$ and $s_2 = 8\%$ are spot rates for one and two years from now respectively. Determine the 1-year forward rate $f_{1,2}$ starting from one year from now.

$$(1+f_{1,2})^1 = (1+s_1)^{-1}(1+s_2)^2 = \frac{1.08^2}{1.07} = 1.090093 \approx 1.0901$$

This means that a 2-year 8% rate can be obtained either as a direct 2-year investment or by investing for 1 year at 7% followed by a second year at 9.01% (the second year investment must be fixed now!).

1.6 Spot and forward rates dynamics. Obviously, the market offers different conditions for riskless transfer of wealth from today ($t = 0$) to a future date $\tau \geq 0$, which can be expressed in terms of spot rate $\tau \mapsto s(\tau)$ curve (yield curve). Why does this happen? Why spot rate curve is not constant? Why it changes over time? There are thee standard explanations (theories) for this

1. Expectations theory;
2. Liquidity preference;
3. Market segmentation

Expectations Theory

Essentially it says that spot rates are determined by expectations of what rates will be in the future.

The 2-year rate is greater than 1-year rate is explained by the fact that market participants believe (or expect) that the 1-year rate will most likely go up next year. The majority beliefs that interest rate will rise translates into a market expectation.

For instance, when $s_1 = 7\%$ and $s_2 = 8\%$, the forward rate $f_{1,2} = 9.01\%$. According to the expectations hypothesis, 9.01% is the markets expected value of next years 1-year spot rate $s(1)_1$.

Liquidity preference

This theory assets that investors usually prefer short-term fixed income securities over long-term securities as they do not like to tie up capital in long-term securities which are more sensitive to interest rate changes (thus greater near-term risk).

Hence to induce investors into long-term instruments, higher rates are offered for bonds of longer maturities (spot rate curve rises as a result).

Market Segmentation

This theory argues that the market for fixed-income securities is segmented by maturity dates, and investors in different segments have different projected need for future funds or risk preference. As investors competing for short-term bonds are different from those competing for long-term bonds, the interest rates for short-term and long-term maturities are thus different.

A moderated version allows for individual investors to shift segments if the rates in an adjacent segment become more attractive. Thus adjacent rates should not be grossly out of line with each other.

For the reminder of this section, let us agree that the compounding is once a year. We discuss the spot and forward rates over the discrete time points

$$i \in \{1, 2, 3, \ldots,\} \quad \text{corresponting to years} \qquad \tau \in \{1,\ 2,\ 3, \ldots\}, \quad \text{from now}$$

Denote by

$$s_1, s_2, \ldots, s_n \quad \text{spot rates for dates } 1,\ 2,\ \ldots,\ n$$

Similarly, denote by

$$f_{i,j} \quad \text{the forward rate from time point } i \text{ to } j.$$

If one believes in market expectations, then a plausible forecast of future interest rates can be obtained as follows.

- at the beginning, observe the spot rate curve $(s_1, s_2, \ldots, s_n)$ and form forward rates $(f_{1,2}, f_{1,3}, \ldots, f_{1,n})$ as

$$f_{1,j+1} = (\frac{(1+s_{j+1})^{j+1}}{1+s_1})^{\frac{1}{j}} - 1, \qquad j = 1, \ldots, n-1$$

- according to market expectation hypothesis, $f_{1,j}$ $j = 2, \ldots, n$ should be a kind of forecast for the spot rates $(s_1^{(1)}, \ldots, s_{n-1}^{(1)})$ which will be realized at the market after one year. Hence we claim prediction as

$$s_j^{(1)} \approx f_{l,j+1}, \qquad j = 1, \ldots, n-1$$

- proceeding recursively, obtain forward rates $f_{l,l+j}$ for $j = 1, \ldots, n-l$ by

$$f_{l,j+l} = \left(\frac{(1+s_{j+l})^{j+l}}{(1+s_l)^l}\right)^{\frac{1}{j}} - 1, \qquad j = 1, \ldots, n-l$$

  and claim predictions for the spot rates at time l as

$$s_j^{(l)} \approx f_{l,j+l}, \qquad j = 1, \ldots, n-l.$$

Such a prediction strategy may be realized as the following diagram

$$\begin{array}{llllll} f_{0,1} & f_{0,2} & f_{0,3} \quad \cdots & f_{0,n-2} & f_{0,n-1} & f_{0,n} \\ f_{1,2} & f_{1,3} & f_{1,4} \quad \cdots & f_{1,n-1} & f_{1,n} & \\ f_{2,3} & f_{2,4} & f_{2,5} \quad \cdots & f_{2,n} & & \\ \cdots & \cdots & & & & \\ f_{n-2,n-1} & f_{n-2,n} & & & & \\ f_{n-1,n} & & & & & \end{array}$$

Example (A simple forecast)

To forecast next years spot rate curve from the current yearly spot rate curve

$$(s_j)_{j=1}^7 = (6.00\%, 6.45\%, 6.80\%, 7.10\%, 7.36\%, 7.56\%, 7.77\%),$$

we use expectations dynamics. Define

$$s_j^{(1)} = f_{1,j+1} = \left(\frac{(1+s_{j+1})^{j+1}}{1+s_1}\right)^{\frac{1}{j}} - 1, \qquad j = 1, \ldots, 6$$

which gives

$$(s_j^{(1)})_{j=1}^6 = (6.90\%, 7.20\%, 7.47\%, 7.70\%, 7.88\%, 8.06\%)$$

If l-step-prediction is needed, define

$$s_j^{(l)} = f_{l,j+l} = \left(\frac{(1+s_{j+l})^{j+l}}{(1+s_l)^l}\right)^{\frac{1}{j}} - 1, \qquad j = 1, \ldots, 7-l$$

and calculate

$$\begin{aligned} (s_j^{(1)})_{j=1}^6 &= (0.0690, 0.0720, 0.0747, 0.0770, 0.0787, 0.0807) \\ (s_j^{(2)})_{j=1}^5 &= (0.0750, 0.0775, 0.0797, 0.0812, 0.0830) \\ (s_j^{(3)})_{j=1}^4 &= (0.0801, 0.0821, 0.0833, 0.0850) \\ (s_j^{(4)})_{j=1}^3 &= (0.0841, 0.0849, 0.0867) \\ (s_j^{(5)})_{j=1}^2 &= (0.0857, 0.0880) \\ (s_j^{(6)})_{j=1}^1 &= (0.0904) \end{aligned}$$

Let us mention the so-called *Invariance Theorem* which claims the following:

Suppose that interest rates evolve according to expectations dynamics. Then (assuming a yearly compounding convention) a sum of money invested in the interest rate market for k years will grow by a factor of $(1+s_k)^k$ independently of the investment and reinvestment strategy (so long as all funds are fully invested).

We try to prove this theorem later

Next, we consider the short rates. The short rate r_k is defined by

$$r_k = f_{k,k+1} \qquad k = 0, 1, 2, \ldots$$

and stands for yearly-compounded interest rate offered today for money borrowed at time point k and returned at $k+1$. Given short rates $(r_k)_{k=0}^{n-1}$, the spot rates $(s_k)_{k=1}^{n}$ are specified completely, since

$$\begin{aligned}(1+s_k)^k &= B_0(k)^{-1} = \underbrace{\frac{B_0(0)}{B_0(1)}}_{1+f_{0,1}} \underbrace{\frac{B_0(1)}{B_0(2)}}_{1+f_{1,2}} \cdot \ldots \underbrace{\frac{B_0(k-1)}{B_0(k)}}_{1+f_{k-1,k}} \\ &= (1+r_0)(1+r_1) \cdot \ldots \cdot (1+r_{k-1}).\end{aligned}$$

holds for $k = 1, \ldots, n$. Observe that the interest rate offered for time from 0 to k can be earned by investing money for each of the subsequent k years, at rates that are offered today.

Similarly, the forward rate $f_{i,j}$ can be written in terms of short rates

$$\begin{aligned}(1+f_{i,j})^{j-i} = \frac{B_0(i)}{B_0(j)} = &= \underbrace{\frac{B_0(i)}{B_0(i+1)}}_{1+f_{i,i+1}} \underbrace{\frac{B_0(i+1)}{B_0(i+2)}}_{1+f_{i+1,i+2}} \cdot \ldots \underbrace{\frac{B_0(j-1)}{B_0(j)}}_{1+f_{j-1,j}} \\ &= (1+r_i)(1+r_{i+1}) \cdot \ldots \cdot (1+r_{j-1}).\end{aligned}$$

2 Fixed income instruments

2.1 Market participants and rates. For a given entity, the cost of borrowing depends on credit quality. Let us put aside the potential loss from default and consider *interbank money market* where banks provide unsecured short term credits to each other.

- **LIBOR:** (London Interbank Offered Rate, definition later). This rate is also available for deposits in different currencies.
- **Fed funds rate:** If a bank (in US) dows not have sufficient balances, it can borrow from the central bank. The rate charged overnight is called (effective) FedFunds rate. It is considered as the best proxy for a risk free USD rate. It is very close to the overnight Libor rate, with spread in single basis points.

- **Eonia and Sonia:** are similar to FedFunds. Eonia (Euro OverNight Index Average) and Sonia (Sterling OverNight Index Average) are determined by actual overnight lending/borrowing transactions from qualified banks.

While Libor represnts bank's estimates conditions at which borrowing/lending might take place, the overnight rates FedFunds/Eonia/Sonia represent transactions that have occurred. During the financial crisis $2007-2009$, there have been concerns whether Libors reflect actual costs of funding for the banks. There have been calls to scrap the Libor rates altogether. Since then, there attempts to drop Libor rate and to use overnight rates instead. This lecture will elaborate on the adjustment of the interest rate theory, required by this development. However, let us postpone this topic for the moment.

2.2 LIBOR-linked instruments. LIBOR approach takes into account that there is a finite number of maturity dates $\{0 < \tau_1 <, \dots, < \tau_{n+1}\}$, where zero bond terminate. In most cases, they are equidistant $\tau_{i+1} = \tau_i + \Delta$ for $i = 1, \dots, n$, for instance Δ corresponds to three months. LIBOR rate is defined through prices of zero bonds by

$$L_t(\tau, \tau') = \frac{B_t(\tau) - B_t(\tau')}{(\tau' - \tau)B_t(\tau')}, \qquad 0 \leq t \leq \tau \leq \tau' \leq T.$$

LIBOR rates are listed, they are used as underlyings for a large variety of sophisticated financial instruments.

If we decide at time t in a riskless way to invest the capital x within $[\tau, \tau']$, then the growth rate which can be fixed with certainty at this time t is $L_t(\tau, \tau')$. This is seen as follows: At time t

- go short for x zero bonds maturing at τ
- invest the received capital $xB_t(\tau)$ to create long position of $xB_t(\tau)/B_t(\tau')$ zero bonds maturing at τ'.
- Such a portfolio requires no capital until τ where x is used to close short position. At time τ' one receives from the long position $xB_t(\tau)/B_t(\tau')$ currency unit. The return rate of this deal is

  $$B_t(\tau)/B_t(\tau') - 1$$

  earned within $[\tau, \tau']$. To make a reasonable yearly interest rate from this, we need to annualize it. Here, one defines

  $$L_t(\tau, \tau') = (\frac{B_t(\tau)}{B_t(\tau')} - 1)/(\tau' - \tau)$$

  (implicitly assuming that there is no compounding at all) instead of continuous or discrete compounding

  $$\ln(B_t(\tau)/B_t(\tau'))/(\tau' - \tau), \qquad (B_t(\tau)/B_t(\tau'))^{\frac{1}{n}} - 1.$$

Now, consider the following financial instruments:

- Interest rate swap (IRS) exchanges cash flows at fixed dates $\tau_i \in \{\tau_{m+1}, \dots, \tau_n\}$ between two legs
  - fixed leg pays/ receives
$$\text{at } \tau_i \text{ the amount } \underbrace{K}_{\text{fixed rate}} \times \underbrace{N}_{\text{nominal amount}} \times \underbrace{(\tau_i - \tau_{i-1})}_{\text{duration}}$$
  - floating leg receives/pays
$$\text{at } \tau_i \text{ the amount } \underbrace{L_{\tau_{i-1}}(\tau_{i-1}, \tau_i)}_{\text{floating rate}} \times \underbrace{N}_{\text{nominal amount}} \times \underbrace{(\tau_i - \tau_{i-1})}_{\text{duration}}$$

  One distinguishes receiver IRS (fixed leg receives) and payer IRS (fixed leg pays). Note that payer IRS protects the owner against interest rate movements, fixing the effective interests to K.

- It can be shown that at time $t \le \tau_m$ the fair price of the payer IRS is
$$N \cdot \sum_{i=m+1}^{n} B_t(\tau_i)(L_t(\tau_{i-1}, \tau_i) - K)(\tau_i - \tau_{i-1}). \tag{2.1}$$
  The price (2.1) of the IRS is equal to zero if and only if
$$\underbrace{\sum_{i=m+1}^{n} (B_t(\tau_{i-1}) - B_t(\tau_i))}_{B_t(\tau_m) - B_t(\tau_n)} - K \sum_{i=m+1}^{n} B_t(\tau_i)(\tau_i - \tau_{i-1}) = 0$$
  That is, at time t the interest rate, which makes the IRS costless is given by
$$K_{ATM}(t) = \frac{B_t(\tau_m) - B_t(\tau_n)}{\sum_{i=m+1}^{n} B_t(\tau_i)(\tau_i - \tau_{i-1})}$$
  and is called the *at-the-money interest rate.*

- A cap is used similarly to the payer IRS, giving a protection to the obligor. At times $\tau_i = \tau_{m+1}, \dots, \tau_n$, it pays
$$\text{at } \tau_i \text{ the amount } \underbrace{(L_{\tau_{i-1}}(\tau_{i-1}, \tau_i) - K)^+}_{\text{caplet's payoff}} \times \underbrace{N}_{\text{nominal amount}} \times \underbrace{(\tau_i - \tau_{i-1})}_{\text{duration}}$$

- The counterpart to a cap is a floor. At times $\tau_i = \tau_{m+1}, \dots, \tau_n$, it pays
$$\text{at } \tau_i \text{ the amount } \underbrace{(K - L_{\tau_{i-1}}(\tau_{i-1}, \tau_i))^+}_{\text{floorlet's payoff}} \times \underbrace{N}_{\text{nominal amount}} \times \underbrace{(\tau_i - \tau_{i-1})}_{\text{duration}}$$

- A swaption gives the right to enter a payer IRS at a given future date (usually at the first settlement date τ_m of the IRS). Thus, options payoff at maturity date τ_m is

$$N(\sum_{i=m+1}^{n} B_{\tau_m}(\tau_i)(L_{\tau_m}(\tau_{i-1},\tau_i) - K)(\tau_i - \tau_{i-1}))^+$$

Note that we have utilized the IRS price from (2.1).

- A Forward rate agreement (FRA) is a forward contact written on

$$FRA_\tau(\tau,\tau') = (\tau' - \tau)\frac{L_\tau(\tau,\tau') - k}{1 + (\tau' - \tau)L_\tau(\tau,\tau')} \quad \text{at time } \tau$$

Here $[\tau,\tau']$ is the spanning interval, and k is the rate of the FRA. It turns out that FRA has a unique fair price which is model independent and is given by

$$FRA_t(\tau,\tau') = B_t(\tau) - B_t(\tau') - (\tau' - \tau)kB_t(\tau')$$

On this account the rate, for which the FRA value is zero at time t in $[0,\tau]$is simply the LIBOR rate

$$k = L_t(\tau,\tau'), \quad t \in [0,\tau]$$

- A Eurodollar future (ED) is similar in spirit to FRA. It is a futures contract written on

$$100(1 - FRA_\tau(\tau,\tau')) \quad \text{at time } \tau.$$

Unlike FRAs, EDs are standardized.

- Similarly to the above IRS, there are other swaps. Usually they exchange a fixed rate against a variable rate, different from LIBOR rate. For instance, the *Libor-in-arrears swap* has variable rates which is based on Libor spanning the next period. Furthermore the *Averaging swap* is based on variable rates referring to certain averages of Libor rates. These averages are determined by a selection of spanning intervals for the underlying Libor rates. A *floating-floating single currency swap* exchanges two different floating rates, for instance FedFunds against Libor. The floating leg can also be based on the so-called CMS rate (Constant Maturity Swap rate, which is identical to K_{ATM} defined above.) Exotic swaps have a function applied to some rate in the floating leg. A swap can be callable, meaning that the issuer can call off (redeem) the contract at any time.

- Similarly to the above European swaption, there are other swaptions. Let us consider a Bermudan swaption. It can be exercised at any time of the tenor structure of the underlying IRS. Upon exercise, the holder will receive the remaining cash flow from the IRS. It turns out that issuing a callable interest rate swap is equivalent to issuing a non-callable swap while buying an appropriate Bermudan swaption.

- An overview over a range of fixed-income instruments is presented in Chapter 5 of *Interest Rate Modeling* L. B. G. Andersen, V. V. Piterbarg.

3 Exercises: Basic notions

Exercise 1 From a $250,000 insurance settlement, a widow wants to take an income of $3000 per month. If the money can be invested at 10.75% (monthly compounding), how many months of income can she receive? The last payment will be smaller than $3000. Determine this final payment.

Solution Apply the annuity formula

$$P = \frac{A}{r}(1 - \frac{1}{(1+r)^n}) \quad \Rightarrow \quad n = -\ln(1 - \frac{P \cdot r}{A})/\ln(1+r)$$

for monthly payment

$$n = -\ln(1 - \frac{P \cdot r/12}{A})/\ln(1 + r/12)$$

where the nominal monthly compounded interest rate is $r = 0.1075$, $P = \$250,000$ and $A = \$3000$. This gives $n = 153.8946$ which means that there are $n-1 = 153$ full payments and the last payment is

$$\begin{aligned}
&(1+r)^n(P - \frac{A}{r}(1 - \frac{1}{(1+r)^{n-1}})) \\
&= (1 + 0.1075/12)^{154}(250,000 - \frac{3,000}{0.1075/12}(1 - \frac{1}{(1+0.1075/12)^{153}})) \\
&= 2685.206 \approx 2685.21.
\end{aligned}$$

Exercise 2 A debt of $25,000 is to be amortized over 7 years at 7% interest (compounded monthly). What value of monthly payments will achieve this?

Solution Apply the formula $A = \frac{r(1+r)^n P}{(1+r)^n - 1}$ for monthly payment and compounding, then we have

$$A = \frac{\frac{0.07}{12}(1 + \frac{0.07}{12})^{7\times 12} \times 25000}{(1 + \frac{0.07}{12})^{7\times 12} - 1} = 377.317.$$

Exercise 3 (The biweekly mortgage) Here is a proposal that has been advanced as a way for homeowners to save thousand of dollars on mortgage payments: pay biweekly instead of monthly. Specifically, if monthly payments are x, it is suggested that one instead pay $x/2$ every two weeks (for a total of 26 payments per year). This will pay down the mortgage faster, saving interest. The savings are surprisingly dramatic for this seemingly minor modificationoften cutting the total interest payment by over one-third. Assume a loan amount of $100,000 for 30 years at 10% interest, compounded monthly.

(a) Under a monthly payment program, what are the monthly payments and the total interest paid over the course of 30 years?

(b) Using the biweekly program, when will the loan be completely repaid, and what are the savings in total interest paid over the monthly program? (You may assume biweekly compounding for this part.)

Solution Apply the formula $A = \frac{r(1+r)^n P}{(1+r)^n - 1}$ for monthly payment to obtain

$$A = \frac{\frac{0.1}{12}(1+\frac{0.1}{12})^{30\times 12} \times 100000}{(1+\frac{0.1}{12})^{30\times 12} - 1} \approx 877.5716$$

The total interest paid is $360 \times 877.57 - 100,000 = 215925.8$. For bi-weekly program, the payment would be $A = 877.5716/2 = 438.78$. Suppose the same nominal interest rate, but bi-weekly compounded. Under this program, we would have

$$\begin{aligned} n &= -\ln(1 - \frac{P \cdot r/26}{A})/\ln(1 + r/26) \\ &= -\ln(1 - \frac{100000 \cdot 0.1/26}{438.78})/\ln(1 + 0.1/26) = 544.9331. \end{aligned}$$

which means that 545 payments of 438.78 are more than enough to pay back the mortgage. This gives the total interest of $545 \times 438.78 - 100000 = 139108.9$. Compared to monthly program, the saved interest is at least $215925.8 - 139108.9 = 76816.9$.

Exercise 4 Consider a bond with a 10% coupon and with yield to maturity= 8%. If the bonds yield to maturity remains constant, then in one year, will the bond price be higher, lower or unchanged? Why?

Solution Lower. The formula

$$P(\lambda, n) = \frac{F}{(1+\lambda/m)^n} + \sum_{k=1}^{n} \frac{C/m}{(1+\lambda/m)^k}$$

shows that the difference

$$\begin{aligned} P(\lambda, n) - P(\lambda, n-1) &= \frac{F}{(1+\lambda/m)^n}(1 - (1+\lambda/m)) + \frac{C/m}{(1+\lambda/m)^n} \\ &= \frac{1}{(1+\lambda/m)^n}(C/m - F\lambda/m) = \frac{1}{(1+\lambda/m)^n}(\rho F/m - F\lambda/m) \end{aligned}$$

is positive if coupon rate ρ is greater than yield λ. This is the case here $\rho = 10\%$, $\lambda = 8\%$.

Exercise 5 A Treasury bond paying an 8% coupon rate with semiannual coupons currently sells at par value (the last coupon has just been paid). What coupon rate would they have to pay in order to sell at par if they paid coupons annually? (Hint: The same effective annual rate of return was to be maintained.)

Solution Strictly speaking, the problem is ill-posed. Under the present assumptions there is no way to determine the price for yearly coupon bond from the price of half-yearly coupon bond. One could try to argue that the yearly coupon bond sells at par if and only if the yield, calculated on yearly compounding assumption equals to the coupon rate. Further, yearly compounded yield should be equivalent to half-yearly yield of 8%, which makes $(1 + 0.08/2)^2 - 1 = 0.816$. Thus, to possible sell the the annual coupon bond at par, coupons should be paid at the rate of 8.16%.

Exercise 6 A newly issued 5% 20-year bond pays its coupons annually. Its yield to maturity is 7%.

(i) Suppose the bond is selling at a yield to maturity of 6% at the end of the year after its first coupon is paid, and you sell the bond then. Find the rate of return for this one-year investment period.

(ii) Suppose you sell the bond after its second coupon is paid, and the bonds yield to maturity is 6% then. Determine the realized compound yield if coupons can be reinvested for one year at 3% interest rate.

Hint: The realized yield y_{realized} of an investment is defined by

$$(1 + y_{\text{realized}}/m)^n = \frac{V}{P}$$

where P is the money invested at the beginning, V money returned at the end, $m \in \underline{\mathbb{N}}$ yearly compounding, n number of compound periods during the investment.

Solution

(i) Money, invested in this strategy is

$$\begin{aligned} P &= \frac{F}{(1+0.07)^{20}} + \sum_{k=1}^{20} \frac{F\rho}{(1+0.07)^k} \\ &= F((1+0.07)^{-20} + \frac{\rho}{0.07}(1-(1+0.07)^{-20})) \end{aligned}$$

where $\rho = 0.05$ is the coupon rate. Money, returned from this strategy is

$$V = F((1+0.06)^{-19} + \frac{\rho}{0.06}(1-(1+0.06)^{-19})) + \rho F$$

We calculate the return

$$\frac{(1+0.06)^{-19} + \frac{0.05}{0.06}(1-(1+0.06)^{-19}) + 0.05}{(1+0.07)^{-20} + \frac{0.05}{0.07}(1-(1+0.07)^{-20})} \approx 1.1907$$

(ii) Money, returned after two years is

$$V = F((1+0.06)^{-18} + \frac{\rho}{0.06}(1-(1+0.06)^{-18})) + \rho F \times (1+0.03) + \rho F$$

We calculate the realized yield from

$$
\begin{aligned}
(1 + y_{\text{realized}})^2 &= \frac{(1+0.06)^{-18} + \frac{0.05}{0.06}(1-(1+0.06)^{-18}) + 0.05 \times (1+0.03) + 0.05}{(1+0.07)^{-20} + \frac{0.05}{0.07}(1-(1+0.07)^{-20})} \\
\Rightarrow \quad y_{\text{realized}} &\approx 0.1226064
\end{aligned}
$$

Exercise 7 Consider a market with zero bond prices $B_t(\tau)$ listed prior $t \leq \tau$ to their maturity dates

$$\tau \in \{1, 2, 3, \ldots, n\}$$

measured in years. Suppose that the assumption of market expectation hypothesis is satisfied exactly.

a) Determine the evolution of bond prices.

b) Suppose that at time $t \in \{1, 2, 3, \ldots, n\}$ a capital $W > 0$ is being invested in zero bonds (which are still listed at t) and at time $t+1$ all positions are closed. Calculate the return of this strategy.

c) Show the invariance theorem.

Solution

a) According to the market expectation hypothesis, forward rate gives a prediction of the spot rate. Moreover, this prediction is realized exactly

$$s_j^{(l)} = \left(\frac{1}{B_l(l+j)}\right)^{\frac{1}{j}} - 1, \qquad l = 1, \ldots, n-1, \quad j = 1, \ldots, n-l$$

Calculating the left-hand side,

$$s_j^{(l)} = \left(\frac{B_0(l)}{B_0(l+j)}\right)^{\frac{1}{j}} - 1, \qquad l = 1, \ldots, n-1, \quad j = 1, \ldots, n-l$$

we obtain

$$\frac{1}{B_l(l+j)} = \frac{B_0(l)}{B_0(l+j)}, \qquad l = 1, \ldots, n-1, \quad j = 1, \ldots, n-l$$

which yields the dynamics of bond prices under the market expectation hypothesis as

$$B_l(l+j) = \frac{B_0(l+j)}{B_0(l)} \qquad l = 1, \ldots, n-1, \quad j = 1, \ldots, n-l.$$

We can write this in recursive form

$$B_{t+1}(k) = \frac{B_0(t)}{B_0(t+1)} B_t(k), \qquad \text{for all } t < k.$$

as the following calculation shows:

$$\begin{aligned}
B_1(k) &= \frac{B_0(k)}{B_0(1)} \\
B_2(k) &= \frac{B_0(k)}{B_0(2)} = \frac{B_0(1)}{B_0(2)} \frac{B_0(k)}{B_0(1)} \\
B_3(k) &= \frac{B_0(k)}{B_0(3)} = \frac{B_0(2)}{B_0(3)} \frac{B_0(k)}{B_0(2)} \\
\vdots &= \\
B_{t+1}(k) &= \frac{B_0(k)}{B_0(t+1)} = \frac{B_0(t)}{B_0(t+1)} \underbrace{\frac{B_0(k)}{B_0(t)}}_{B_t(k)}
\end{aligned}$$

b) At the time t the bonds with maturities $t+j$ for $j = 1, \dots, n-t$ are listed. For $j = 1, \dots, n-t$ denote by θ_j the capital invested at time t in zero bond maturing at $t+j$. Then we have allocated the initial capital as

$$\sum_{j=1}^{n-t} \theta_j B_t(t+j) = W.$$

This gives the terminal capital

$$\begin{aligned}
\sum_{j=1}^{n-t} \theta_j B_{t+1}(t+j) &= \sum_{j=1}^{n-t} \theta_j \frac{B_0(t)}{B_0(t+1)} B_t(t+j) \\
&= \frac{B_0(t)}{B_0(t+1)} \underbrace{\sum_{j=1}^{n-t} \theta_j B_t(t+j)}_{W} = \frac{B_0(t)}{B_0(t+1)} W.
\end{aligned}$$

Hence the return is $B_0(t)/B_0(t+1)$ independently on the allocation of the capital W along available zero bonds.

c) Let us recursively apply the result from b). At time $t = 0$ we invest capital W in available zero bonds, at time $t = 1$ the capital has to be re-invested. This can be modeled by a strategy which closes all positions at time $t = 1$ giving the capital

$$\frac{B_0(0)}{B_0(1)} W$$

which is used to open new positions held until time $t = 2$ at which we obtain a capital

$$\frac{B_0(1)}{B_0(2)} \frac{B_0(0)}{B_0(1)} W.$$

Proceeding recursively, the capital at time $t+1$ is

$$\frac{B_0(t)}{B_0(t+1)} \frac{B_0(t-1)}{B_0(t)} \cdot \ldots \cdot \frac{B_0(1)}{B_0(2)} \frac{B_0(0)}{B_0(1)} W = \frac{1}{B_0(t+1)} W = (1 + s_{t+1})^{t+1} W$$

which is the assertion of the invariance theorem.

Exercise 8 Consider a market for zero bonds maturing at times $0 < \tau_1 < \tau_2, \ldots, < \tau_n$. Suppose that IRS provides a sequence of cash flows, given as following: At time τ_i the owner receives the amount

$$N(L_{\tau_{i-1}}(\tau_{i-1}, \tau_i) - K)(\tau_i - \tau_{i-1}), \qquad i = m+1, \ldots, n.$$

a) Is the above IRS a payer or a receiver interest rate swap?

b) Show that the unique no-arbitrage price at time $t \in [0, \tau_m]$ of the above IRS is given by

$$N \cdot \sum_{i=m+1}^{n} B_t(\tau_i)(L_t(\tau_{i-1}, \tau_i) - K)(\tau_i - \tau_{i-1}).$$

c) Consider swaption giving its owner the right to enter swap contract without payment at τ_m. Show that the unique no-arbitrage price of this swaption at time τ_m is given by

$$(N \cdot \sum_{i=m+1}^{n} B_{\tau_m}(\tau_i)(L_{\tau_m}(\tau_{i-1}, \tau_i) - K)(\tau_i - \tau_{i-1}))^+$$

d) Consider a cap which provides a sequence of cash flows, given as following: At time τ_i the owner receives the amount

$$N(L_{\tau_{i-1}}(\tau_{i-1}, \tau_i) - K)^+(\tau_i - \tau_{i-1}), \qquad i = m+1, \ldots, n$$

At time τ_m, compare the values of IRS, cap, and swaption.

Solution a) In this contract, the floating leg receives and the fixed leg pays. According to the definition, we have a payer IRS.

b) By linearity of the pricing rule, it suffices to show that at time $t \in [0, \tau_m]$ the unique no-arbitrage price of an asset which yields cash flow

$$L_{\tau_{i-1}}(\tau_{i-1}, \tau_i) - K \qquad \text{at } \tau_i$$

is given by

$$B_t(\tau_i)(L_t(\tau_{i-1}, \tau_i) - K) = \frac{B_t(\tau_{i-1}) - B_t(\tau_i)}{\tau_i - \tau_{i-1}} - B_t(\tau_i)K.$$

Using time value of money, we deduce that the price of $-K$ delivered at τ_i is given at t by $-B_t(\tau_i)K$. Hence, it remains to show that the price of $L_{\tau_{i-1}}(\tau_{i-1}, \tau_i)$ delivered at τ_i is given at t by $(B_t(\tau_i) - B_t(\tau_{i-1}))/(\tau_i - \tau_{i-1})$. Indeed, a security which pays at τ_i an amount

$$(\tau_i - \tau_{i-1})L_{\tau_{i-1}}(\tau_{i-1}, \tau_i) = \frac{1}{B_{\tau_{i-1}}(\tau_i)} - 1$$

is realized by a strategy which creates a short position in one bond maturing at τ_i (giving -1 at τ_i) and long position in bond maturing at τ_{i-1} which is re-invested at τ_{i-1} in bond maturing at τ_i (giving 1 at τ_{i-1} and $B_{\tau_{i-1}}(\tau_i)^{-1}$ at τ_i). The initial endowment of such strategy is $B_t(\tau_{i-1}) - B_t(\tau_i)$. For this reason,

$$\frac{B_t(\tau_{i-1}) - B_t(\tau_i)}{\tau_i - \tau_{i-1}}$$

is the unique no-arbitrage price at time $t \in [0, \tau_m]$ of a security which pays

$$L_{\tau_{i-1}}(\tau_{i-1}, \tau_i) \qquad \text{at } \tau_i .$$

c) A swaption gives its owner the right to enter IRS with no payment at time τ_m. Clearly, the unique no-arbitrage price of the swaption at τ_m will be equal to the swap price if it is positive. Otherwise, the swaption is worthless.

d) From the above argumentation it follows that at time τ_m the price of swaption is higher than that of the swap. Now, compare swaption against cap. Cap provides higher cashflow than swap, hence price of the cap must be greater or equal than the price of the swap, also in situations where swap price is positive. If the price of the swap is positive at τ_m, then swaption price equals to the swap price and must be equal or below the cap price. In the situation that swap price is negative at τ_m, swaption price is zero which is again below or equal to the cap price since the latter is always non-negative (indeed, cap provides non-negative cash flow). Thus, we conclude that at τ_m

$$\text{swap price} \leq \text{swaption price} \leq \text{cap price}.$$

Exercise 9 Consider a floating-rate note. This is a coupon-paying bond whose coupon payments depend on market conditions and are linked to the LIBOR rate. Let us consider the floating rate note which is defined as follows.

- at maturity time τ_n the owner receives the face value F in addition to the last coupon payment $F \times L_{\tau_{n-1}}(\tau_{n-1}, \tau_n)(\tau_n - \tau_{n-1})$.
- prior to maturity time, at times $\tau_k \in \{\tau_1, \tau_2, \ldots, \tau_{n-1}\}$ the owner of the bond receives coupon payment $F \times L_{\tau_{k-1}}(\tau_{k-1}, \tau_k)(\tau_k - \tau_{k-1})$

Calculate the price of the above floating-rate note at any time $\tau_m \in \{\tau_0, \ldots, \tau_{n-1}\}$ after the recent coupon is paid.

Solution At time $\tau_m \in \{\tau_0, \ldots, \tau_{n-1}\}$, we decompose the contract into face value payment with price

$$F B_{\tau_m}(\tau_n)$$

and coupon payments with price

$$F \times \sum_{k=m+1}^{n} B_{\tau_m}(\tau_k) L_{\tau_m}(\tau_{k-1}, \tau_k)(\tau_k - \tau_{k-1}) \quad = \quad F \times \underbrace{\sum_{k=m+1}^{n} (B_{\tau_m}(\tau_{k-1}) - B_{\tau_m}(\tau_k))}_{B_{\tau_m}(\tau_m) - B_{\tau_m}(\tau_n)}$$

That is, the price of the floating rate note at time τ_m is

$$FB_{\tau_m}(\tau_n) + F \times (B_{\tau_m}(\tau_m) - B_{\tau_m}(\tau_n)) = FB_{\tau_m}(\tau_m) = F$$

Hence, immediately after the recent coupon is paid, the floating-rate note is always traded at par.

Exercise 10 Consider the at-the-money interest rate derived from the swap price

$$K_{ATM}(t) = \frac{B_t(\tau_m) - B_t(\tau_n)}{\sum_{i=m+1}^{n} B_t(\tau_i)(\tau_i - \tau_{i-1})}$$

a) Derive the dynamics of $(K_{ATM}(\tau_k))_{k=1}^{m}$ under the market expectation hypothesis.

b) Express at-the-money interest rate $K_{ATM}(t)$ through LIBOR rates at time t.

Solution a) Given the initial bond curve $(B_0(\tau_i))_{i=1}^{n}$, the bond curve at time τ_k under the market expectation hypothesis is given by

$$(B_{\tau_k}(\tau_i) = \frac{B_0(\tau_i)}{B_0(\tau_k)})_{i=k+1}^{n}$$

Hence

$$\begin{aligned} K_{ATM}(\tau_k) &= \frac{B_{\tau_k}(\tau_m) - B_{\tau_k}(\tau_n)}{\sum_{i=m+1}^{n} B_{\tau_k}(\tau_i)(\tau_i - \tau_{i-1})} \\ &= \frac{(B_0(\tau_m) - B_0(\tau_n))/B_0(\tau_k)}{(\sum_{i=m+1}^{n} B_0(\tau_i)(\tau_i - \tau_{i-1}))/B_0(\tau_k)} \\ &= \frac{B_0(\tau_m) - B_0(\tau_n)}{\sum_{i=m+1}^{n} B_0(\tau_i)(\tau_i - \tau_{i-1})} \end{aligned}$$

which means that $(K_{ATM}(\tau_k))_{k=1}^{m}$ is constant under the market expectation hypothesis.
b) In the formula

$$K_{ATM}(t) = \frac{B_t(\tau_m) - B_t(\tau_n)}{\sum_{i=m+1}^{n} B_t(\tau_i)(\tau_i - \tau_{i-1})}$$

we divide the numerator and the denominator by $B_t(\tau_m)$ to obtain

$$K_{ATM}(t) = \frac{1 - \frac{B_t(\tau_n)}{B_t(\tau_m)}}{\sum_{i=m+1}^{n} \frac{B_t(\tau_i)}{B_t(\tau_m)}(\tau_i - \tau_{i-1})}.$$

Further, express each fraction using LIBOR rates

$$\begin{aligned} \frac{B_t(\tau_i)}{B_t(\tau_m)} &= \frac{B_t(\tau_i)}{B_t(\tau_{i-1})} \cdot \ldots \cdot \frac{B_t(\tau_{m+2})}{B_t(\tau_{m+1})} \frac{B_t(\tau_{m+1})}{B_t(\tau_m)} \\ &= (1 + L_t(\tau_{i-1}, \tau_i)(\tau_i - \tau_{i-1}))^{-1} \cdot \ldots \cdot (1 + L_t(\tau_m, \tau_{m+1})(\tau_{m+1} - \tau_m))^{-1} \\ &= \Pi_{j=m+1}^{i}(1 + L_t(\tau_{j-1}, \tau_j)(\tau_j - \tau_{j-1}))^{-1} \end{aligned}$$

Hence

$$K_{ATM}(t) = \frac{1 - \Pi_{j=m+1}^{n}(1 + L_t(\tau_{j-1}, \tau_j)(\tau_j - \tau_{j-1}))^{-1}}{\sum_{i=m+1}^{n} \Pi_{j=m+1}^{i}(1 + L_t(\tau_{j-1}, \tau_j)(\tau_j - \tau_{j-1}))^{-1}(\tau_i - \tau_{i-1})}.$$

4 Stochastic calculus

In this section, we review stochastic calculus with respect to *continuous semimartingales*, for details the interested reader is referred to

Introduction to Option Pricing Theory
Gopinath KallianpurRajeeva L. Karandikar

4.1 Continuous semimartingales. The total variation of a function $f : [0,T] \to \mathbb{R}$ is defined by

$$\mathrm{Var}(f)_t = \sup\{\sum_{k=0}^{n-1} |f_{s_{k+1}} - f_{s_k}| \, : \, 0 \leq s_0 \leq s_1 \leq \dots, \leq s_n \leq t, \;\; n \in \mathbb{N}\}.$$

The variation $t \mapsto \mathrm{Var}(f)_t$ is non-decreasing. We say that f has bounded variation, if $\mathrm{Var}(f)_T < \infty$, which is the case if and only if the function f can be expressed as difference of two non-decreasing functions. A process $(M_t)_{t\in[0,T]}$ is called a local martingale, if there exists a sequence $(\tau_n)_{n\in\mathbb{N}}$ of stopping times which increases to infinity $\lim_{n\to\infty} \tau_n = +\infty$ such that

$$\text{each of the stopped processes } (M_{t\wedge\tau_n})_{t\in[0,T]} \text{ follows a martingale.}$$

Each martingale is a local martingale.

By definition, the class of continuous semimartingales consists of two types of processes,

$$\mathcal{A}_c = \{A = (A_t)_{t\in[0,T]} \, : \, A \text{ is continuous with bounded variation } \}$$

and

$$\mathcal{M}_c = \{M = (M_t)_{t\in[0,T]} \, : \, M \text{ is a continuous local martingale } \}$$

in the sense that for each $X = (X_t)_{t\in[0,T]} \in \mathcal{S}_c$ there exist $A = (A_t)_{t\in[0,T]} \in \mathcal{A}_c$ and $M = (M_t)_{t\in[0,T]} \in \mathcal{M}_c$ such that $X = A + M$ meaning that

$$X_t = A_t + M_t \quad \text{for all } t \in [0,T].$$

It turns out that the intersection of both classes is trivial

$$\mathcal{A}_c \cap \mathcal{M}_c = \{\text{processes that are constant}\}.$$

4.2 Idea of stochastic integration. The theory of stochastic integration introduces the integrals of the type

$$\int_0^t h_s dX_s \qquad t \in [0,T]$$

for a large classes of integrands $(h_s)_{s\in[0,T]}$ and integrators $(X_s)_{s\in[0,T]}$. It turns out that the integration can not be defined pathwise. More precisely, the attempt to define

$$\int_0^t h_s(\omega)dX_s(\omega) \qquad t\in[0,T]$$

for each $\omega\in\Omega$ for functions $s\mapsto h_s(\omega)$, $s\mapsto X_s(\omega)$ fails, since the most interesting integrators (like Brownian motion) exhibit singular properties (unbounded variation) which does not allow for a straight forward definition of such an integral. Nevertheless, it is possible to define the integral in the sense of Ito, which is based on an appropriate limit of the sum of the type

$$\sum_k h_{t_k}(X_{t_{k+1}}-X_{t_k}),$$

Note here that the value h_{t_k} of the integrant is taken on the left point t_k of the interval, $[t_k,t_{k+1}]$ over which the increment $X_{t_{k+1}}-X_{t_k}$ of the integrator is calculated.

In what follows, we sketch the basic steps in the construction of the stochastic integral. Because of significant mathematical complexity, we stay on the 'symbolic' level, explaining its properties and rules, however.

4.3 Construction of Ito integral. Suppose that the integrator

$$X=(X_t)_{t\in[0,T]} \quad \text{is continuous}$$

and the integrand is a simple process

$$h_t=\sum_{k=0}^n H^k 1_{]t_k,t_{k+1}]}(t)$$

which is constant on the intervals of the partition $0=t_0\le\cdots\le t_n\le t_{n+1}=T$ with values $H^1,\ldots,H^n$, which are random variables. For such situation, the integral is defined as

$$\int_0^T h_s dX_s=\sum_{k=0}^n H^k(X_{t_{k+1}}-X_{t_k}).$$

The integral with variable upper boundary is defined as a stochastic process by

$$\int_0^t h_s dX_s := \int_0^T h_s 1_{[0,t]}(s)dX_s, \qquad t\in[0,T]$$

As indicated above, the extension to a more general class of integrands is not obvious and depends on the integrator. The general extension idea is the following: Given integrand $h=(h_s)_{s\in[0,T]}$ and integrator $X=(X_s)_{s\in[0,T]}$ chose a sequence $(h^m)_{m\in\mathbb{N}}$ of simple processes $h^m=(h^m_s)_{s\in[0,T]}$ which approximates the integrand

$$\lim_{m\to\infty} h^m = h \quad \text{where the limit is understood in appropriate sense.}$$

If the sequence of integrals $(\int_0^T h_s^m dX_s)_{m\in\mathbb{N}}$ converges for $m \to \infty$ and the limit does not depend on the choice of the approximating sequence $(h^m)_{m\in\mathbb{N}}$, then it makes sense to define the integral as

$$\int_0^T h_s dX_s := \lim_{m\to\infty} \int_0^T h_s^m dX_s$$

The realization of this idea involves several steps.

- It turns out that the class of integrator processes, for which this idea works for a reasonable large integrand class, is

  $$\mathcal{S}_c = \{X = (X_t)_{t\in[0,T]} \,:\, X \text{ is a continuous semimartingale } \}.$$

  Given $X \in \mathcal{S}_c$, the precise class of integrands, where $\int_0^T h_s dX_s$ is defined, depends on X. However, for each X the integral $\int_0^T h_s dX_s$ is well-defined if $(h_s)_{s\in[0,T]}$ is left-continuous with existing right limits. This yields a large class of common integrands.

- For $A = (A_t)_{t\in[0,T]} \in \mathcal{A}_c$ and $M = (M_t)_{t\in[0,T]} \in \mathcal{M}_c$ the integrals

  $$\int_0^T h_s dA_s, \qquad \int_0^T h_s dM_s$$

  are defined in different ways. For the sum

  $$X_t = A_t + M_t \qquad t \in [0,T]$$

  the integral is introduced componentwise

  $$\int_0^T h_s dX_s = \int_0^T h_s dA_s + \int_0^T h_s dM_s$$

  This yields a consistent definition since other decompositions $X = A' + M'$ with $A' \in \mathcal{A}_c$, $M' \in \mathcal{M}_c$ differ from the previous by constant processes $A - A' = M' - M$ and the integral vanishes for the constant integrator.

- The stochastic integral enjoys many useful properties, among them are

  $$\begin{aligned} (\int_0^t h_s dA_s)_{t\in[0,T]} &\in \mathcal{A}_c \qquad && \text{for } (A_s)_{s\in[0,T]} \in \mathcal{A}_c \\ (\int_0^t h_s dM_s)_{t\in[0,T]} &\in \mathcal{M}_c \qquad && \text{for } (M_s)_{s\in[0,T]} \in \mathcal{M}_c \\ (\int_0^t h_s dX_s)_{t\in[0,T]} &\in \mathcal{S}_c \qquad && \text{for } (X_s)_{s\in[0,T]} \in \mathcal{S}_c \end{aligned}$$

  which hold for each h where the integral is defined. (Note that the last property is a consequence of the previous both.)

- A natural chain rule holds

$$\int_0^t h_s d(\int_0^\cdot h'_u dX_u)_s = \int_0^t h_s h'_s dX_s \qquad \text{for all } t \in [0,T].$$

- For sufficiently smooth function f, the process $(f(X_t))_{t\in[0,T]}$ follows a continuous semimartingale, if $(X_t)_{t\in[0,T]}$ is continuous semimartingale. The so-called Ito formula helps to decompose this semimartingale $(f(X_t))_{t\in[0,T]}$ into a continuous processes of bounded variation and a local martingale and to identify these components.

- When dealing with semimartingales, one of the most important quantities is the *quadratic covariation.* It turns out that for continuous semimartingales

$$X = (X_t)_{t\in[0,T]} \in \mathcal{S}_c, \qquad Y = (Y_t)_{t\in[0,T]} \in \mathcal{S}_c$$

there exists the limit

$$\langle X, Y\rangle_t = \lim_{|\Pi^n|\to 0} \sum_{t_i^n \in \Pi^n, t_i^n < t} (X_{t_{i+1}^n} - X_{t_i^n})(Y_{t_{i+1}^n} - Y_{t_i^n})$$

which is understood in appropriate sense and is taken over partitions

$$\Pi^n = \{0 \le t_0^n <, \dots, t_n^n \le t_{n+1}^n = T\}, \qquad n \in \underline{\mathrm{N}}$$

of increasing tightness

$$|\Pi^n| = \max_{t_k^n \in \Pi^n} |t_{k+1}^n - t_k^n| \to 0 \quad \text{for } n \to \infty.$$

- Quadratic covariation $\langle X, Y\rangle = (\langle X, Y\rangle_t)_{t\in[0,T]}$ enjoys the following properties

$$\begin{aligned}
\langle Y, X\rangle &= \langle X, Y\rangle \quad \text{for all } X, Y \in \mathcal{S}_c \\
\langle Y, X\rangle &\in \mathcal{A}_c \quad \text{for all } X, Y \in \mathcal{S}_c \\
\langle Y + Y', X\rangle &= \langle Y, X\rangle + \langle Y', X\rangle \quad \text{for all } Y, Y', X \in \mathcal{S}_c \\
\langle A, X\rangle &= 0 \quad \text{for all } X \in \mathcal{S}_c,\ A \in \mathcal{A}_c \\
\langle \int_0^\cdot h_s dY_s, X\rangle_t &= \int_0^t h_s d\langle Y, X\rangle_s \quad \text{for all } X, Y \in \mathcal{S}_c \\
& \text{ and } (h_s)_{s\in[0,T]},\ t \in [0,T]. \\
\langle \int_0^\cdot h_s^X dX_s, \int_0^\cdot h_s^Y dY_s\rangle_t &= \int_0^t h_s^X h_s^Y d\langle X, Y\rangle_s \quad \text{for all } X, Y \in \mathcal{S}_c \\
& \text{ and } (h_s^X)_{s\in[0,T]},\ (h_s^Y)_{s\in[0,T]},\ t \in [0,T].
\end{aligned}$$

- The quadratic variation is introduced as

$$\langle X\rangle = \langle X, X\rangle$$

4.4 Ito's formula. This is a very useful result. It helps to decide whether a function of a vector of semimartingales follows (local) martingale.

For continuous semimartingales

$$X^1 = (X^1_t)_{t\in[0,T]} \in \mathcal{S}_c, \ \ \ldots, \ \ X^d = (X^d_t)_{t\in[0,T]} \in \mathcal{S}_c$$

and a function $f : \underset{\sim}{\mathrm{R}}^d \to \underset{\sim}{\mathrm{R}}$ which is sufficiently smooth, holds

$$\begin{aligned} f(X^1_t, \ldots, X^d_t) &= f(X^1_0, \ldots, X^d_0) + \sum_{i=1}^{d} \int_0^t \partial_i f(X^1_s, \ldots, X^d_s) dX^i_s \\ &\quad + \frac{1}{2} \sum_{i,j=1}^{d} \int_0^t \partial_i \partial_j f(X^1_s, \ldots, X^d_s) d\langle X^i, X^j \rangle_s, \quad t \in [0, T]. \end{aligned}$$

The idea behind the Ito's formula is based on Taylor approximation, taking into account that the second-order terms yield a contribution due to non-vanishing covariation for semimartingale paths. To see how it may work consider one dimensional case $d = 1$ and chose a partition $0 = t_0 \leq t_1 \leq \ldots \leq t_n \leq t_{n+1} = t$

$$\begin{aligned} f(X_t) - f(X_0) &= \sum_{k=0}^{n} (f(X_{t_{k+1}}) - f(X_{t_k})) \\ &= \sum_{k=0}^{n} \left(f'(X_{t_k})(X_{t_{k+1}} - X_{t_k}) + \frac{1}{2} f''(X_{t_k})(X_{t_{k+1}} - X_{t_k})^2 + R_k \right) \end{aligned}$$

By appropriate estimations, one obtains Itos result in the limit for increasingly tight partitions.

4.5 Brownian motion. Let $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t\in[0,T]})$ be a filtered probability space.
Definition *The adapted process* $(W_t)_{t\in[0,T]}$ *is called Brownian motion if*

(i) $W_0 = 0$, $(W_t)_{t\in[0,T]}$ *is continuous,*

(ii) for each $0 \leq s \leq t \leq T$, *the increment* $W_t - W_s$ *follows normal distribution* $N(0, t-s)$ *(with mean zero and variance* $t - s$*)*

(iii) for each $0 \leq s \leq t \leq T$, *the increment* $W_t - W_s$ *and the* σ*-algebra* $\mathcal{F}_s$ *are independent.*

The d-dimensional ($d \in \underset{\sim}{\mathrm{N}}$) Brownian motion is defined as $\underset{\sim}{\mathrm{R}}^d$-valued process

$$(W_t = (W^1_t, \ldots, W^d_t))_{t\in[0,T]}$$

where each component $(W^i_t)_{t\in[0,T]}$, $i = 1, \ldots, d$ follows a Brownian motion and $(W^1_t)_{t\in[0,T]}$, $\ldots$, $(W^d_t)_{t\in[0,T]}$ are independent. The existence of the process of Brownian motion is ensured by appropriate construction procedure (first discovered by N. Wiener).

The one-dimensional Brownian motion is a prime example of a martingale since for $0 \leq s \leq t \leq T$

$$\mathbb{E}(W_t - W_s \,|\, \mathcal{F}_s) = \mathbb{E}(W_t - W_s) = 0$$

and so

$$\mathbb{E}(W_t \,|\, \mathcal{F}_s) = \mathbb{E}(W_t - W_s + W_s \,|\, \mathcal{F}_s) = \mathbb{E}(W_t - W_s \,|\, \mathcal{F}_s) + W_s = W_s$$

Furthermore, other martingales are associated with one-dimensional Brownian motion $(W_t)_{t\in[0,T]}$.

- the process

  $$M_t = W_t^2 - t \qquad t \in [0,T]$$

  is a martingale since

  $$\begin{aligned}
  \mathbb{E}(M_t \,|\, \mathcal{F}_s) &= \mathbb{E}(W_t^2 - t \,|\, \mathcal{F}_s) = \mathbb{E}((W_t - W_s + W_s)^2 - ((t-s)+s) \,|\, \mathcal{F}_s) \\
  &= \mathbb{E}((W_t - W_s)^2 - (t-s) + 2(W_t - W_s)W_s + \underbrace{W_s^2 - s}_{M_s} \,|\, \mathcal{F}_s) \\
  &= \underbrace{\mathbb{E}((W_t - W_s)^2 \,|\, \mathcal{F}_s)}_{=\mathbb{E}((W_t - W_s)^2) = t-s} - (t-s) + 2W_s \underbrace{\mathbb{E}(W_t - W_s | \,\mathcal{F}_s)}_{=0} + M_s \\
  &= M_s
  \end{aligned}$$

- the process

  $$M_t = \exp(\sigma W_t - \frac{\sigma^2}{2} t) \qquad t \in [0,T]$$

  is a martingale since

  $$\begin{aligned}
  \mathbb{E}(M_t \,|\, \mathcal{F}_s) &= \mathbb{E}(e^{\sigma(W_t - W_s) - \frac{\sigma^2}{2}(t-s)} M_s \,|\, \mathcal{F}_s) \\
  &= \mathbb{E}(e^{\sigma(W_t - W_s) - \frac{\sigma^2}{2}(t-s)} \,|\, \mathcal{F}_s) M_s \\
  &= \mathbb{E}(e^{\sigma(W_t - W_s) - \frac{\sigma^2}{2}(t-s)}) M_s \\
  &= \int_{\mathbb{R}} e^{\sigma x - \frac{\sigma^2}{2}(t-s)} e^{-\frac{x^2}{2(t-s)}} \frac{dx}{\sqrt{2\pi(t-s)}} M_s \\
  &= \int_{\mathbb{R}} e^{-\frac{1}{2(t-s)}(x^2 - 2x\sigma(t-s) + \sigma^2(t-s)^2)} \frac{dx}{\sqrt{2\pi(t-s)}} M_s \\
  &= \underbrace{\int_{\mathbb{R}} e^{-\frac{1}{2(t-s)}(x - \sigma(t-s))^2} \frac{dx}{\sqrt{2\pi(t-s)}}}_{=1} M_s = M_s
  \end{aligned}$$

4.6 The special role of the Brownian motion. Let $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t\in[0,T]})$ be a filtered probability space which supports the process $(W_t)_{t\in[0,T]}$ of one-dimensional Brownian motion. Because of the following remarkable properties, Brownian motion plays a central role in the stochastic calculus

- As pointed out above, the quadratic variation of one-dimensional Brownian motion $(W_t)_{t\in[0,T]}$ is given by

$$\langle W, W\rangle_t = t \qquad t \in [0,T].$$

The *Levy's characterization* states the converse: If a continuous local martingale $M = (M_t)_{t\in[0,T]}$ satisfies

$$\langle M, M\rangle_t = t \qquad t \in [0,T].$$

then $M = (M_t)_{t\in[0,T]}$ follows a Brownian motion.

- Given a Brownian motion $(W_t)_{t\in[0,T]}$, the stochastic integral

$$\int_0^t h_s dW_s \qquad \text{where } \mathbb{P}(\int_0^t |h|_s^2 ds < \infty) = 1$$

follows a local martingale. In particular, for appropriate integrand $(h_t)_{t\in[0,T]}$ the process $(\int_0^t h_s dW_s)_{t\in[0,T]}$ is a true martingale. The question which martingales appear in this way can be clarified under the additional assumptions that the filtration is generated by $(W_t)_{t\in[0,T]}$

$$\mathcal{F}_t \qquad \left\{ \begin{array}{l} \text{is the smallest } \sigma\text{-algebra such that} \\ W_s \text{ is } \mathcal{F}_t\text{-measurable for each } s \le t \end{array} \right. \qquad \text{for all } t \in [0,T]$$

The *martingale representation theorem* states that if $(\mathcal{F}_t)_{t\in[0,T]}$ is generated by $(W_t)_{t\in[0,T]}$, then for each martingale $(M_t)_{t\in[0,T]}$ there exists $(h_s)_{s\in[0,T]}$ such that holds

$$M_t = M_0 + \int_0^t h_s dW_s, \qquad t \in [0,T].$$

4.7 Stochastic calculus in the framework of diffusion processes. Given filtered probability space $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t\in[0,T]})$ which supports d-dimensional Brownian motion

$$W_t = (W_t^1, \dots, W_t^d), \qquad t \in [0,T]$$

One can show that the quadratic covariation between components of Brownian motion $(W_t)_{t\in[0,T]}$ fulfills

$$\langle W^i, W^j\rangle_t = \left\{ \begin{array}{ll} 0 & \text{if } i \ne j \\ t & \text{for } i = j. \end{array} \right.$$

In many applications, it suffices to considers semimartingales of a special structure

- the processes of bounded variation are of the type

$$(\int_0^t h_s ds)_{t\in[0,T]} \qquad \text{where } \mathbb{P}(\int_0^t |h|_s ds < \infty) = 1$$

- local martingales are of the type

$$(\int_0^t h_s dW_s^i)_{s\in[0,T]} \qquad \text{where } \mathbb{P}(\int_0^t |h|_s^2 ds < \infty) = 1$$

Composed from these ingredients, a semimartingale is given by the diffusion-type process

$$X_t = X_0 + \int_0^t h_s ds + \sum_{i=1}^{d} \int_0^t h_s^i dW_s, \qquad t \in [0,T]$$

Many models in financial mathematics are formulated in this framework.

4.8 Stochastic differential equations. Stochastic differential equations are appropriate to construct a large variety of processes. A stochastic differential equation addresses stochastic process $X = (X_t)_{t\in[0,T]}$ which satisfies

$$X_{t'} = X_t + \int_t^{t'} \mu(X)_s ds + \int_t^{t'} \sigma(X)_s dW_s, \qquad t < t' \tag{4.1}$$

where for each $s \in [0,T]$, $\mu(X)_s$ and $\sigma(X)_s$ are appropriate functionals which depend on the path $(X_u)_{u\in[0,s]}$ of the process until s. The short-hand notation for (4.1) is

$$dX_t = \mu(X)_t dt + \sigma(X)_t dW_t \tag{4.2}$$

In most applications, one deals with explicitly given functions

$$\mu(X)_u = \mu(u, X_u), \qquad \sigma(X)_u = \sigma(u, X_u).$$

The process $(W_t)_{t\in[0,T]}$ stands for a Brownian motion. Here one has to distinguish two situations. If the filtered probability space and the Brownian motion are given, then a process $(X_t)_{t\in[0,T]}$ which satisfies (4.2) is called *strong solution*, on the contrary to weak solution which is a process $(X_t)_{t\in[0,T]}$ on appropriate filtered probability space which supports a Brownian motion $(W_t)_{t\in[0,T]}$ such that (4.2) holds. Let us illustrate these notions by examples

- General linear diffusion. Consider

  $$dX_u = (a_u + b_u X_u)du + (\gamma_u + \sigma_u X_u)dW_u, \quad X_0 = X_0^* \tag{4.3}$$

  on a given filtered probability space $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t\in[0,T]})$ equipped with Brownian motion $(W_t)_{t\in[0,T]}$. Suppose for the moment that $(a_u)_{u\in[0,T]} \equiv 0$, $(\gamma_u)_{u\in[0,T]} \equiv 0$. Then a strong solution to

  $$dZ_u = b_u Z_u du + \sigma_u Z_u dW_u, \quad Z_0 = 1$$

  is given by

  $$Z_u = e^{\int_0^u b_s ds + \int_0^u \sigma_s dW_s - \frac{1}{2}\int_0^u \sigma_s^2 ds} \qquad u \in [0,T]$$

  (to be directly verified by the Ito formula). It turns out that one can correct this process to obtain a solution to (4.3). Define

  $$Y_u = \int_0^u \frac{a_s - \sigma_s \gamma_s}{Z_s} ds + \int_0^s \frac{\gamma_s}{Z_s} dW_s + X_0^*$$

Then by construction

$$dY_u = \frac{a_u - \sigma_u \gamma_u}{Z_u} du + \frac{\gamma_u}{Z_u} dW_u$$

and the Ito formula yields

$$\begin{aligned} d(Y_u Z_u) &= Y_u dZ_u + Z_u dY_u + d\langle Y, Z \rangle_u \\ &= Y_u (b_u Z_u du + \sigma_u Z_u dW_u) + Z_u \left(\frac{a_u - \sigma_u \gamma_u}{Z_u} du + \frac{\gamma_u}{Z_u} dW_u \right) + \frac{\gamma_u}{Z_u} \sigma_u Z_u du \\ &= (a_u + b_u Y_u Z_u) du + (\gamma_u + \sigma_u Y_u Z_u) dW_u \end{aligned}$$

Example: Suppose that $\sigma_u = 0$ for all $u \in [0, T]$ then we obtain

$$Z_u = e^{\int_0^u b_s ds}, \qquad u \in [0, T]$$

In this case

$$Y_u = X_0^* + \int_0^u a_s e^{-\int_0^s b_v dv} ds + \int_0^s \gamma_s e^{-\int_0^s b_v dv} dW_s$$

giving

$$X_u = Z_u Y_u = e^{\int_0^u b_s ds} X_0^* + \int_0^u a_s e^{\int_s^u b_v dv} ds + \int_0^s \gamma_s e^{\int_s^u b_v dv} dW_s$$

For instance, for the stochastic differential equation

$$dr_t = (b - a r_t) dt + \sigma dW_t, \qquad r_0 = r_0^*$$

with deterministic $a, b, \sigma \in \underset{\sim}{\mathrm{R}}$ we have

$$a_u = b, \quad b_u = -a, \quad \gamma_u = \sigma, \quad \sigma_u = 0$$

which gives

$$\begin{aligned} r_u &= e^{-\int_0^u a ds} r_0^* + \int_0^u b e^{-\int_s^u a dv} ds + \int_0^u \sigma e^{-\int_s^u a dv} dW_s \\ &= r_0 e^{-au} + \int_0^u e^{-a(u-s)} b ds + \int_0^u e^{-a(u-s)} \sigma dW_s \\ &= r_0 e^{-au} + \frac{b}{a}(1 - e^{-au}) + \int_0^u e^{-a(u-s)} \sigma dW_s \end{aligned}$$

In the case $b = 0$ one obtains the so-called Ornstein-Uhlenbeck process

$$r_u = r_0 e^{-au} + \int_0^t e^{-a(u-s)} \sigma dW_s$$

- Square-root process ($a, b, \sigma > 0$ appropriate)

$$dR_t = (a - bR_t) dt + \sigma \sqrt{R_t} dW_t, \qquad R_0 = R_0^* > 0 \tag{4.4}$$

Rather than giving the process on a fixed filtered probability space (as for the linear diffusion above), we construct $(R_t)_{t \in [0,T]}$ on an appropriate probability space (equipped with a Brownian motion $(W_t)_{t\in[0,T]}$).

Consider $d \in \underline{\mathrm{N}}$ independent Brownian motions $(W^1_t)_{t\in[0,T]} \dots, (W^d_t)_{t\in[0,T]}$. For each the solution to

$$dX^j_t = -\frac{b}{2}X^j_t dt + \frac{1}{2}\sigma dW^j_t, \quad X_0 = X^*_0 \in \underline{\mathrm{R}}, \qquad j = 1, \dots, d$$

given by the so-called Ornstein-Uhlenbeck process

$$X^j_t = e^{-\frac{b}{2}t}X_0 + \int_0^t e^{-\frac{b}{2}(t-s)}\frac{1}{2}\sigma dW^j_s.$$

Note that

$$X^j_t \text{ is normal with mean } e^{-\frac{b}{2}t}X_0 \text{ and variance } \frac{\sigma^2}{4}\frac{1-e^{-bt}}{b}$$

Define

$$R_t = \sum_{j=1}^{d}(X^j_t)^2 \quad t \in [0, T].$$

What relation holds for dR_t? Use the Ito formula

$$\begin{aligned}
dR_t &= \sum_{j=1}^{d}(2X^j_t dX^j_t + d\langle X^j, X^j\rangle_t) \\
&= \sum_{j=1}^{d}(2X^j_t(\frac{-b}{2}X^j_t dt + \frac{\sigma}{2}dW^j_t) + \frac{\sigma^2}{4}dt) \\
&= (\underbrace{\frac{d\sigma^2}{4}}_{a} - b\underbrace{\sum_{j=1}^{d}(X^j_t)^2}_{R_t})dt + \sigma\sum_{j=1}^{d}X^j_t dW^j_t \\
&= (a - bR_t)dt + \sigma\sqrt{R_t}\underbrace{\sum_{j=1}^{d}\frac{X^j_t}{\sqrt{R_t}}dW^j_t}_{dW_t} \\
&= (a - bR_t)dt + \sigma\sqrt{R_t}dW_t
\end{aligned}$$

Surprisingly, the process

$$W_t = \sum_{j=1}^{d}\int_0^t \frac{X^j_s}{\sqrt{R_s}}dW^j_s, \quad t \in [0, T]$$

turns out to be a Brownian motion: in view of Levy's characterization, it suffices to verify that

- it is continuous: granted by construction of stochastic integral
- it is local martingale: granted by construction, since the integrators $(W^j_t)_{t\in[0,T]}$ are martingales

– quadratic variation is $(\langle W, W\rangle_t = t)_{t\in[0,T]}$.

$$d\langle W, W\rangle_t = \sum_{j=1}^{d} \frac{(X_t^j)^2}{R_t} dt = dt$$

Conclusion: There exists a (weak) solution to (4.4) and the distribution of $R_t = \sum_{j=1}^{d} X_t^j$ seems to be equal to the sum of squares of independent Gaussian variables. (Indeed one can show that that the square root process follows a χ^2-distribution, with appropriate degrees of freedom).

4.9 Illustration: Using Ito formula. Suppose that the filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in\mathbb{R}_+})$ supports a Brownian motion $(B_t)_{t\in\mathbb{R}_+}$. Given an Ito process $(X_t)_{t\in\mathbb{R}_+}$, consider its decomposition

$$X_t = X_0 + \int_0^t a_s ds + \int_0^t m_s dB_s, \qquad t \in \mathbb{R}_+.$$

Use Ito formula to determine $(a_t)_{t\in\mathbb{R}_+}$ for the following processes $(X_t)_{t\geq 0}$:

(a) $X_t = \cos(B_t)\sqrt{t} \quad t \in \mathbb{R}_+$ (4 marks)

(b) $X_t = B_t^2 \int_0^t e^{-s} dB_s \quad t \in \mathbb{R}_+$ (3 marks)

(c) $X_t = \int_0^t B_s^2 ds e^{-B_t} \quad t \in \mathbb{R}_+$ (3 marks)

(a)

$$dX_t = -\sin(B_t)\sqrt{t}dB_t + \cos(B_t)\frac{1}{2\sqrt{t}}dt - \frac{1}{2}\cos(B_t)\sqrt{t}dt$$

That is

$$a_t = \cos(B_t)\frac{1}{2\sqrt{t}} - \frac{1}{2}\cos(B_t)\sqrt{t} \qquad t \in \mathbb{R}_+.$$

(b)

$$dX_t = (2B_t \int_0^t e^{-s}dB_s)dB_t + B_t^2 e^{-t}dB_t + \frac{1}{2}((2\int_0^t e^{-s}dB_s)dt + 2\cdot 2B_t e^{-t}dt)$$

That is

$$a_t = \int_0^t e^{-s}dB_s + 2B_t e^{-t} \qquad t \in \mathbb{R}_+.$$

(c)

$$dX_t = e^{-B_t}B_t^2 dt - \int_0^t B_s^2 ds \cdot e^{-B_t}dB_t + \frac{1}{2}\int_0^t B_s^2 ds e^{-B_t}dt$$

That is

$$a_t = e^{-B_t}B_t^2 + \frac{1}{2}\int_0^t B_s^2 ds e^{-B_t} \qquad t \in \mathbb{R}_+.$$

5 Short rate models and affine term structure

5.1 No-arbitrage and martingale modelling. The mathematical theory of arbitrage has developed over last 20 years.

- In this theory, one considers price processes and trading strategies which are modeled by the so-called semimartingales. Adaptations are needed to exclude pathological cases of trading strategies (similar to doubling strategies in a casino).
- Unlike in the discrete time case, there is no direct equivalence between no-arbitrage and the martingale property. However, the assumption that price process follows, after appropriate numeraire change, a martingale with respect to an equivalent measure, is still sufficient to exclude arbitrage.
- The common practical viewpoint is that one should consider such stochastic models for asset price evolution, which exhibit a built-in no-arbitrage. That is, when proposing a stochastic model, one ensures already by construction, that there exists a discounting asset and an equivalent martingalizing measure. This is referred to as the *martingale modeling approach*.

5.2 Bond modeling with short rate. Given a filtered probability space

$$(\Omega, \mathcal{F}, \mathbb{Q}, (\mathcal{F}_t)_{t \in [0,T]}),$$

let us agree on the following assumptions

- $\mathbb{Q}$ is supposed to describe a risk neutral measure
- the probability space supports the process $(W_t)_{t \in [0,T]}$ of Brownian motion
- we denote the conditional expectation by $\mathbb{E}_t(\cdot) = \mathbb{E}(\cdot | \mathcal{F}_t)$ for $t \in [0,T]$

To establish a reasonable model of money market, one needs to introduce the dynamics of zero coupon bonds

$$(B_t(\tau))_{t \in [0,\tau]}, \qquad \text{with marurities } \tau \in [0,T]$$

such that there is no arbitrage for trading bonds with different maturities. Following the ideas of martingale modeling it suffices to introduce a numeraire $(C_t)_{t \in [0,T]}$ such that

$$\left(\frac{B_t(\tau)}{C_t}\right)_{t \in [0,T]} \qquad \text{follows martingale with respect to } \mathbb{Q}.$$

In the *short rate modeling* approach, one considers the process $(r_t)_{t \in [0,T]}$ with the interpretation that r_t describes an instantaneous continuously compounded short rate at time t. With this interpretation, one defines the so-called standard savings account

$$B_t = e^{\int_0^t r_s ds} \qquad t \in [0,T]$$

which is used as numeraire $(C_t)_{t\in[0,T]}$ to establish an arbitrage-free dynamics of all zero coupon bonds

$$\frac{B_t(\tau)}{B_t} = \mathbb{E}^{\mathbb{Q}}\left(\frac{B_\tau(\tau)}{B_\tau} \,|\, \mathcal{F}_t\right) = \mathbb{E}^{\mathbb{Q}}\left(\frac{1}{B_\tau} \,|\, \mathcal{F}_t\right)$$

which yields the definition

$$B_t(\tau) = \mathbb{E}^{\mathbb{Q}}_t\left(e^{-\int_t^\tau r_s ds}\right), \qquad 0 \le t \le \tau \le T.$$

At this point, the standard savings account $(B_t)_{t\in[0,T]}$ can be considered as an abstract numeraire. However, it is important to point out that it can be considered as wealth of a trading strategy which invests one currency unit in just maturing bond, at maturity of this bond the wealth is entirely invested in the next maturing bond. Proceeding recursively, the wealth is rolled over the next maturing bond. Supposing that the market lists zero coupon bonds with any maturity time $\tau \in [0, T]$, the wealth of such strategy can be approximatively described by $(B_t)_{t\in[0,T]}$. That is, let us assume that $(B_t)_{t\in[0,T]}$ represents the price dynamics of a tradable security, which can be synthetically created from dynamic wealth reallocation from maturing bond to the next maturing bond.

In the short rate modeling of money markets, the short rate evolution is supposed to follow a stochastic differential equation of the form

$$dr^\alpha_t = \mu^\alpha(t, r_t)dt + \sigma^\alpha(t, r_t)dW_t, \qquad r_0 = r^*_0 \in \underline{\mathrm{R}} \tag{5.1}$$

whose coefficients

$$\begin{aligned} [0, T] \times \underline{\mathrm{R}} \to \underline{\mathrm{R}}, \qquad & (t, r) \mapsto \mu^\alpha(t, r) \\ [0, T] \times \underline{\mathrm{R}} \to \underline{\mathrm{R}} \qquad & (t, r) \mapsto \sigma^\alpha(t, r) \end{aligned}$$

are parameterized by α in appropriate parameter space.

Having determined the dynamics

$$B^\alpha_t(\tau) = \mathbb{E}^{\mathbb{Q}}_t\left(e^{-\int_t^\tau r^\alpha_s ds}\right), \qquad 0 \le t \le \tau \le T,$$

the model needs to be calibrated to the market data. Usually, one tries to chose the parameters in the coefficients and possibly the initial condition r^*_0 such that

- the initially observed bond curve $(B^*_0(\tau))_{\tau\in[0,T]}$ is explained as good as possible,
- listed prices of related derivatives (for instance, bond options) are explained as good as possible.

Problems that may occur are

- Inconsistency with the initial data: It may be not possible to match the bond curve by any choice of parameters and initial short rate

$$(B_0^{\alpha}(\tau))_{\tau\in[0,T]} \neq (B_0^*(\tau))_{\tau\in[0,T]} \quad \text{for all } \alpha$$

- Inconsistency with time evolution: For model parameters α^*, which explain the initial bond curve at the beginning

$$(B_0^{\alpha^*}(\tau))_{\tau\in[0,T]} = (B_0^*(\tau))_{\tau\in[0,T]},$$

  the future market bond curve is not explained

$$(B_t^{\alpha^*}(\tau))_{\tau\in[t,T]} \neq (B_t^*(\tau))_{\tau\in[t,T]}.$$

The inconsistency with initial curve is a real drawback of the model. Usually, by a choice of finite number of (real) parameters, the consistency with initial data can not be achieved. However, if parameterizations includes functions, then there are sufficient degrees of freedom to fit the initial bond curve.

The inconsistency with time evolution is an accepted problem within short rate modeling. Similar situation occurs in commodity modeling where the dynamics of futures curves is hardly captured by a model. Usually, one needs to re-calibrate the model periodically, when the market futures curve is not explained by parameters chosen previously. This is of course not satisfactory since the idea is that the chosen model should be able to describe the dynamics for all times.

The advantage of short rate models is that most of they provide simple closed form expressions for zero bonds and for bond options.

In the above setting and under slight natural assumptions, it turns out that the conditional expectation

$$B_t(\tau) = \mathbb{E}^{\mathbb{Q}}\left(e^{-\int_t^\tau r_s ds} \,|\, \mathcal{F}_t\right)$$

can be written conditioning on the recent short rate only

$$B_t(\tau) = \mathbb{E}^{\mathbb{Q}}\left(e^{-\int_t^\tau r_s ds} \,|\, \sigma(r_t)\right).$$

The reason is that strong solution to stochastic differential equations like (5.1) are Markov processes. Due to factorization property of conditional expectation, there exists a function

$$f^\tau : [0,\tau] \times \underline{\mathrm{R}} \to \underline{\mathrm{R}}_+$$

such that

$$B_t(\tau) = \mathbb{E}_t^{\mathbb{Q}}\left(e^{-\int_t^\tau r_s ds}\right) = \mathbb{E}^{\mathbb{Q}}\left(e^{-\int_t^\tau r_s ds} \,|\, \sigma(r_t)\right) = f^\tau(t, r_t)$$

for all $t \in [0,\tau]$. This function turns out to be solution of an appropriate partial differential equation, which results from the so-called Feynmann-Kac formula.

5.3 Bond prices from Feynmann-Kac formula. Consider the strong solution $(r_t)_{t\in[0,T]}$ to the stochastic differential equation

$$dr_t = \mu(t, r_t)dt + \sigma(t, r_t)dW_t, \qquad r_0 = r_0^* \in \underline{\mathrm{R}}$$

The Feynmann-Kac formula shows how to calculate the function f^τ from

$$E_t(e^{-\int_t^\tau R(r_s)ds} g(r_\tau)) = f^\tau(t, r_t)$$

with given coefficient functions μ, σ and appropriate functions R and g on $\underline{\mathrm{R}}$. Although the exact formulation of conditions and proof are technically involving, the idea is simple.

One realizes that $f^\tau(t, r_t)$ can be written as

$$f^\tau(t, r_t) = \underbrace{E_t(e^{-\int_0^\tau R(r_s)ds} g(r_\tau))}_{M_t} e^{\int_0^t R(r_s)ds}$$

where for appropriate functions R, g

the process $(M_t)_{t\in[0,T]}$ follows a martingale

obviously

the process $(e^{\int_0^t R(r_s)ds})_{t\in[0,T]}$ is of bounded variation.

Thus, Ito formula suggests that the drift term for $df^\tau(t, r_t)$ is $R(r_t)f^\tau(t, r_t)dt$ since

$$\begin{aligned} df^\tau(t, r_t) &= d(M_t e^{\int_0^t R(r_s)ds}) = R(r_t)e^{\int_0^t R(r_s)ds} M_t dt + e^{\int_0^t R(r_s)ds} dM_t \\ &= R(r_t)f^\tau(t, r_t)dt + e^{\int_0^t R(r_s)ds} dM_t. \end{aligned}$$

On the other hand, Ito formula indicates another drift expression

$$\begin{aligned} df^\tau(t, r_t) &= f^\tau_{(1,0)}(t, r_t)dt + f^\tau_{(0,1)}(t, r_t)dr_t + \frac{1}{2} f^\tau_{(0,2)}(t, r_t)d\langle r\rangle_t \\ &= \left(f^\tau_{(1,0)}(t, r_t) + f^\tau_{(0,1)}(t, r_t)\mu(t, r_t) + \frac{1}{2} f^\tau_{(0,2)}(t, r_t)\sigma(t, r_t)^2 \right) dt \\ &\quad + f^\tau_{(0,1)}(t, r_t)\sigma(r_t)dW_t \end{aligned}$$

From the canonical decomposition of semimartingales we know that

$$R(r_t)f^\tau(t, r_t) = f^\tau_{(1,0)}(t, r_t) + f^\tau_{(0,1)}(t, r_t)\mu(t, r_t) + \frac{1}{2} f^\tau_{(0,2)}(t, r_t)\sigma(t, r_t)^2$$

which suggests to try out a partial differential equation on $[0, \tau] \times \underline{\mathrm{R}}$

$$f^\tau_{(1,0)}(t, r) + f^\tau_{(0,1)}(t, r)\mu(t, r) + \frac{1}{2} f^\tau_{(0,2)}(t, r)\sigma(t, r)^2 - R(r)f^\tau(t, r) = 0$$

subject to the terminal condition

$$f^\tau(\tau, r) = g(r) \qquad \text{for all } r \in \underline{\mathrm{R}}.$$

Clearly in our case of bond price calculation the functions are

$$R(r) = r, \quad g(r) = 1 \qquad r \in \underline{\mathrm{R}}$$

which results in the following partial differential equation

$$f^{\tau}_{(1,0)}(t,r) + f^{\tau}_{(0,1)}(t,r)\mu(t,r) + \frac{1}{2} f^{\tau}_{(0,2)}(t,r)\sigma(t,r)^2 - r f^{\tau}(t,r) = 0$$

subject to the terminal condition

$$f^{\tau}(\tau, r) = 1 \qquad \text{for all } r \in \underline{\mathrm{R}}.$$

It turns out that for certain choices of coefficient functions μ and σ, the solution of the above partial differential equation reduces to a system of ordinary differential equations. This is an important case, since these ordinary differential equations can frequently be solved explicitly. Even if there is no explicit solution, the numerical treatment of ordinary differential equations is by far simpler than that of partial differential equations.

The situation, where this reduction is possible correspond to the so-called *affine term structure* situation, a case where the solution is given by

$$f^{\tau}(t,r) = e^{\mathcal{A}^{\tau}(t) - \mathcal{B}^{\tau}(t) \cdot r}$$

with deterministic functions

$$\mathcal{A}^{\tau}, \mathcal{B}^{\tau} : [0, \tau] \mapsto \underline{\mathrm{R}}.$$

One obtains the affine term structure with coefficients

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

where α, β, γ, and δ are deterministic functions of $[0, T]$.Let us formulate the statement.

Assume that the coefficient functions μ and σ are as above, define the function

$$\mathcal{B}^{\tau} : [0, \tau] \mapsto \underline{\mathrm{R}}$$

as solutions to the following ordinary differential equation

$$\begin{aligned} \dot{\mathcal{B}}^{\tau}(t) + \alpha(t)\mathcal{B}^{\tau}(t) - \frac{1}{2}\gamma(t)\mathcal{B}^{\tau}(t)^2 &= -1 \\ \text{with boundary condition} \quad \mathcal{B}^{\tau}(\tau) &= 0 \end{aligned}$$

Having obtained the function $\mathcal{B}^{\tau}$, the function $\mathcal{A}^{\tau} : [0, \tau] \mapsto \underline{\mathrm{R}}$ is given as solution to

$$\begin{gathered} \dot{\mathcal{A}}^{\tau}(t) = \beta(t)\mathcal{B}^{\tau}(t) - \frac{1}{2}\delta(t)\mathcal{B}^{\tau}(t)^2 \\ \text{with boundary condition} \quad \mathcal{A}^{\tau}(\tau) = 0 \end{gathered}$$

All we need is to prove that

$$f^\tau(t,r) = e^{\mathcal{A}^\tau(t) - \mathcal{B}^\tau(t)\cdot r}$$

solves the partial differential equation. Calculating the derivatives

$$\begin{aligned}
\partial_{(1,0)} f^\tau(t,r) &= (\dot{\mathcal{A}}^\tau(t) - \dot{\mathcal{B}}^\tau(t) r) f^\tau(t,r) \\
\partial_{(0,1)} f^\tau(t,r) &= -\mathcal{B}^\tau(t) f^\tau(t,r) \\
\frac{1}{2}\partial_{(0,2)} f^\tau(t,r) &= \frac{1}{2}\mathcal{B}^\tau(t)^2 f^\tau(t,r)
\end{aligned}$$

we see that

$$f^\tau_{(1,0)}(t,r) + f^\tau_{(0,1)}(t,r)\mu(t,r) + \frac{1}{2} f^\tau_{(0,2)}(t,r)\sigma(t,r)^2 - r f^\tau(t,r) = 0$$

holds if

$$(\dot{\mathcal{A}}^\tau(t) - \dot{\mathcal{B}}^\tau(t) r) + (-\mathcal{B}^\tau(t))(\alpha(t) r + \beta(t)) + \frac{1}{2}\mathcal{B}^\tau(t)^2(\gamma(t) r + \delta(t)) - r = 0$$

holds for each $r \in \underset{\sim}{\mathrm{R}}$, $t \in [0,\tau]$ and $\tau \in [0,T]$, This is satisfied if

$$\begin{aligned}
-\dot{\mathcal{B}}^\tau(t) - \alpha(t)\mathcal{B}^\tau(t) + \frac{1}{2}\gamma(t)\mathcal{B}^\tau(t)^2 - 1 &= 0 \\
\dot{\mathcal{A}}^\tau(t) - \beta(t)\mathcal{B}^\tau(t) + \frac{1}{2}\delta(t)\mathcal{B}^\tau(t)^2 &= 0
\end{aligned}$$

holds for each $r \in \underset{\sim}{\mathrm{R}}$, $t \in [0,\tau]$ and $\tau \in [0,T]$. Furthermore due to

$$\mathcal{B}^\tau(\tau) = 0, \qquad \mathcal{A}^\tau(\tau) = 0$$

the required boundary condition

$$f^\tau(\tau, r) = 1, \qquad r \in \underset{\sim}{\mathrm{R}}$$

holds. Having obtained the function f^τ in this way, one obtains

$$f^\tau(t, r_t) = \mathbb{E}(e^{-\int_t^\tau r_s ds} \mid \mathcal{F}_t) \qquad t \in [0,\tau], \;\; \tau \in [0,T].$$

Thus the zero bond prices are given as required

$$B_t(\tau) = f^\tau(t, r_t) \qquad t \in [0,\tau], \;\; \tau \in [0,T].$$

Note that the ordinary differential equation

$$\begin{aligned}
\dot{\mathcal{B}}^\tau(t) + \alpha(t)\mathcal{B}^\tau(t) - \frac{1}{2}\gamma(t)\mathcal{B}^\tau(t)^2 &= -1 \\
\text{with boundary condition} \quad \mathcal{B}^\tau(\tau) &= 0
\end{aligned}$$

includes a quadratic term, this is the so-called Ricatti equation, its solution can be obtained explicitly if the coefficients are constant.

5.4 Standard short rate models. In what follows, we present a choice of standard models, defined by the corresponding stochastic differential equation for the short rate. In all models the coefficients a, b, σ are deterministic and $(\theta(t))_{t\in[0,T]}$ $(\sigma(t))_{t\in[0,T]}$ are deterministic functions.

1. Vasicek model
$$dr_t = (b - ar_t)dt + \sigma dW_t$$
2. Cox-Ingersoll-Ross model (CIR)
$$dr_t = a(b - r_t)dt + \sigma\sqrt{r_t}dW_t$$
3. Dothan model
$$dr_t = ar_t dt + \sigma r_t dW_t$$
4. Black-Derman-Toy model
$$dr_t = \theta(t)r_t dt + \sigma(t)dW_t$$
5. Hull-White I (extended Vasicek) model
$$dr_t = (\theta(t) - ar_t)dt + \sigma(t)dW_t$$
6. Ho-Lee model
$$dr_t = \theta(t)dt + \sigma dW_t$$
7. Hull-White II (extended CIR) model
$$dr_t = (\theta(t) - a(t)r_t)dt + \sigma(t)\sqrt{r_t}dW_t$$

Example let us calculate the bond prices within Vasicek model. Obviously, Vaicek model yields affine term structure. Let us identify the coefficients

$$dr_t = (\underbrace{b}_{=\beta(t)} \underbrace{-a}_{\alpha(t)} r_t)dt + \underbrace{\sigma}_{\sqrt{\delta(t)}} dW_t$$

where $\gamma(t) = 0$ for all $t \in [0, T]$. With this, we write the ordinary differential equation for $(\mathcal{B}^\tau(t))_{t\in[0,T]}$.

$$\dot{\mathcal{B}}^\tau(t) + \alpha(t)\mathcal{B}^\tau(t) - \frac{1}{2}\gamma(t)\mathcal{B}^\tau(t)^2 = -1, \qquad \mathcal{B}^\tau(\tau) = 0$$

is given as

$$\dot{\mathcal{B}}^\tau(t) - a\mathcal{B}^\tau(t) = -1, \qquad \mathcal{B}^\tau(\tau) = 0$$

with the solution

$$\mathcal{B}^\tau(t) = \frac{1 - e^{-a(\tau - t)}}{a}.$$

Now, let us turn to the calculation of $\mathcal{A}^\tau$. The ordinary differential equation

$$\dot{\mathcal{A}}^\tau(t) = \beta(t)\mathcal{B}^\tau(t) - \frac{1}{2}\delta(t)\mathcal{B}^\tau(t)^2, \qquad \mathcal{A}^\tau(\tau) = 0$$

can be solved by integration.

$$\mathcal{A}^\tau(t) = -\int_t^\tau \left(\beta(u)\mathcal{B}^\tau(u) - \frac{1}{2}\delta(u)\mathcal{B}^\tau(u)^2\right) du$$

In our case of $\beta(t) = b$ and $\delta(t) = \sigma^2$ for all $t \in [0, T]$, we obtain

$$\begin{aligned}
\mathcal{A}^\tau(t) &= -\int_t^\tau \left(b\mathcal{B}^\tau(u) - \frac{1}{2}\sigma^2\mathcal{B}^\tau(u)^2\right) du \\
&= -b\int_t^\tau \mathcal{B}^\tau(u)du + \frac{\sigma^2}{2}\int_t^\tau \mathcal{B}^\tau(u)^2 du
\end{aligned}$$

After a calculation, the result is

$$\begin{aligned}
B_t(\tau) &= e^{\mathcal{A}^\tau(t) - \mathcal{B}^\tau(t) r_t} \qquad t \in [0, \tau], \;\; \tau \in [0, T] \\
\mathcal{B}^\tau(t) &= \frac{1 - e^{-a(\tau - t)}}{a} \\
\mathcal{A}^\tau(t) &= \frac{(\mathcal{B}^\tau(t) - \tau + t)(ab - \frac{\sigma^2}{2})}{a^2} - \frac{\sigma^2\mathcal{B}^\tau(t)^2}{4a}.
\end{aligned}$$

(Please check!). The Vasicek model may be not consistent with a given initial bond curve, since four parameters (r_0, b, a, σ) define a particular initial bond curve family. A market bond curve may be not matched. This is not the case for the Ho-Lee model, since it involves a function $(\theta(t))_{t\in[0,T]}$ giving enough flexibility.

Example Consider Ho-Lee model. Here the short rate follows

$$dr_t = \underbrace{\theta(t)}_{\beta(t)} dt + \underbrace{\sigma}_{\sqrt{\delta(t)}} dW_t$$

where obviously $\alpha(t) = 0$ and $\gamma(t) = 0$ for all $t \in [0, T]$. Thus, the the differential equation for $(\mathcal{B}^\tau(t))_{t\in[0,T]}$ becomes

$$\dot{\mathcal{B}}^\tau(t) + \underbrace{\alpha(t)}_{=0}\mathcal{B}^\tau(t) - \frac{1}{2}\underbrace{\gamma(t)}_{=0}\mathcal{B}^\tau(t)^2 = -1, \qquad \mathcal{B}^\tau(\tau) = 0$$

and is given as

$$\dot{\mathcal{B}}^\tau(t) = -1, \qquad \mathcal{B}^\tau(\tau) = 0$$

with the solution $\mathcal{B}^\tau(t) = \tau - t$. The function $\mathcal{A}^\tau$ is calculated as

$$\begin{aligned}
\mathcal{A}^\tau(t) &= \int_t^\tau \left(-\underbrace{\beta(u)}_{\theta(u)}\mathcal{B}^\tau(u) + \frac{1}{2}\underbrace{\delta(u)}_{\sigma^2}\mathcal{B}^\tau(u)^2\right) du \\
&= -\int_t^\tau \theta(u)(\tau - u)du + \int_t^\tau \frac{1}{2}\sigma^2(\tau - u)^2 du
\end{aligned}$$

with this the bond price dynamics in the Ho-Lee model is given by

$$B_t(\tau) = e^{-\int_t^\tau \theta(u)(\tau-u)du + \frac{\sigma^2}{2}\frac{(\tau-t)^3}{3} - (\tau-t)r_t} \qquad t \in [0,\tau], \;\; \tau \in [0,T].$$

Example Hull White I model, here the short rate follows

$$dr_t = (\underbrace{\theta(t)}_{=\beta(t)} \underbrace{-a}_{\alpha(t)} r_t)dt + \underbrace{\sigma(t)}_{\sqrt{\delta(t)}} dW_t$$

where obviously $\gamma(t) = 0$ for all $t \in [0,T]$. With this, the differential equation for $(\mathcal{B}^\tau(t))_{t\in[0,T]}$ becomes

$$\dot{\mathcal{B}}^\tau(t) + \underbrace{\alpha(t)}_{-a} \mathcal{B}^\tau(t) - \frac{1}{2} \underbrace{\gamma(t)}_{=0} \mathcal{B}^\tau(t)^2 = -1, \qquad \mathcal{B}^\tau(\tau) = 0$$

is given as

$$\dot{\mathcal{B}}^\tau(t) - a\mathcal{B}^\tau(t) = -1, \qquad \mathcal{B}^\tau(\tau) = 0$$

with the solution

$$\mathcal{B}^\tau(t) = \frac{1 - e^{-a(\tau-t)}}{a}$$

Now, let us turn to the calculation of $\mathcal{A}^\tau$. The ordinary differential equation

$$\dot{\mathcal{A}}^\tau(t) = \beta(t)\mathcal{B}^\tau(t) - \frac{1}{2}\delta(t)\mathcal{B}^\tau(t)^2, \qquad \mathcal{A}^\tau(\tau) = 0$$

can be solved by integration.

$$\mathcal{A}^\tau(t) = -\int_t^\tau \left(\beta(u)\mathcal{B}^\tau(u) - \frac{1}{2}\delta(u)\mathcal{B}^\tau(u)^2\right) du$$

In our case of $\beta(t) = \theta(t)$ and $\delta(t) = \sigma(t)^2$ for all $t \in [0,T]$, we obtain

$$\begin{aligned}
\mathcal{A}^\tau(t) &= -\int_t^\tau \left(\theta(u)\mathcal{B}^\tau(u) - \frac{1}{2}\sigma(u)^2\mathcal{B}^\tau(u)^2\right) du \\
&= -\int_t^\tau \theta(u)\mathcal{B}^\tau(u)du + \frac{1}{2}\int_t^\tau \sigma(u)^2\mathcal{B}^\tau(u)^2 du
\end{aligned}$$

6 Forward rate models and HJM approach

6.1 The idea of forward rate modeling. Remember that the short rate approach claims that

$$B_t(\tau) = \mathbb{E}_t^{\mathbb{Q}}(e^{-\int_t^\tau r_s ds}), \qquad t \in [0,\tau], \;\; t \in [0,T]$$

where $(r_t)_{t\in[0,T]}$ is an exogenously given process describing the continuously compounded instantaneous short rate evolution. Beyond its advantages, the framework of short rate modeling suffers from the following drawbacks

- inconsistency (either with initial data or with the evolution of the bond curve)
- unrealistic description of money market dynamics in terms of few factors. Actually, only one factor, the short short rate r_t (in the generic case). There are extensions that suggest to model r_t as a difference (or linear transformation) of a multi-variate process which is defined as solution to an appropriate mutli-dimensional stochastic differential equation (for instance, linear diffusion).
- short rate models become increasingly complicated, when it comes to the introduction of several explanatory factors (beyond short rate). In particular, the fit to the initial bond curve becomes difficult.

As an alternative way, the bond curve dynamics can be expressed as

$$B_t(\tau) = e^{-\int_t^\tau f_t(s)ds}, \qquad t \in [0, \tau], \ \ \tau \in [0, T]$$

Here the quantities $f_t(s)$ are interpretable as instantaneous continuously compounded forward rates. Their dynamics must be modeled by stochastic processes

$$(f_t(\tau))_{t\in[0,\tau]}, \qquad \text{for each } \tau \in [0, T].$$

Strictly speaking, this requires a parallel modeling of an infinite number of processes describing the instantaneous forward rates at the same time.

6.2 The HJM approach. Heath, Jarrow and Morton suggested to describe this forward rates directly, by solution to a family of stochastic differential equations

$$df_t(\tau) = \alpha_t(\tau)dt + \sigma_t(\tau)dW_t, \qquad t \in [0, \tau]$$

(family is indexed by τ) where the coefficients $\alpha_t(\tau)$, $\sigma_t(\tau)$ must be specified for all $0 \le t \le \tau \le T$ accordingly.

- usually, the coefficients $\sigma_t(\tau)$ and the Brownian motion are multi-dimensional, which means that the dynamics is actually given as

$$\begin{aligned} df_t(\tau) &= \alpha_t(\tau)dt + [\sigma_t^1(\tau), \dots, \sigma_t^d(\tau)]d \begin{bmatrix} W_t^1 \\ \vdots \\ W_t^d \end{bmatrix} \\ &= \alpha_t(\tau)dt + \sum_{i=1}^d \sigma_t^i(\tau)dW_t^i \end{aligned}$$

- one decides to introduce the dynamics of forward rates directly under the risk neutral measure $\mathbb{Q}$. In this framework the (possible multi-dimensional) process $(W_t)_{t\in[0,T]}$ is assumed to follow a Brownian motion under $\mathbb{Q}$.

- in this situation, it turns out that the stochastic processes $(\alpha_t(\tau))_{t\in[0,\tau]}$ $\tau \in [0,T]$ and $(\sigma_t(\tau))_{t\in[0,\tau]}$ $\tau \in [0,T]$ can not be chosen arbitrary. In order to preserve no-arbitrage, if the forward rate volatility structure $(\sigma_t(\tau))_{t\in[0,\tau]}$ $\tau \in [0,T]$ is specified, the dynamics $(\alpha_t(\tau))_{t\in[0,\tau]}$ $\tau \in [0,T]$ follows.

- In the HJM setting, the calibration of the model to the initial data becomes trivial, since having observed the bond curve $(B_0^*(\tau))$ from the market data, one can match it by
$$B_0^*(\tau) = e^{-\int_0^\tau f_0^*(u)du}$$
which is ensured by
$$f_0^*(\tau) = -\frac{\partial}{\partial\tau}\ln(B_0^*(\tau))$$
Using this initial forward rates $(f_0^*(\tau))_{\tau\in[0,T]}$ as starting values in the dynamics
$$df_t(\tau) = \alpha_t(\tau)dt + \sigma_t(\tau)dW_t, \qquad f_0(\tau) = f_0^*(\tau),$$
one obtains a perfect fit of the initial bond prices to the initial bond curve.

6.3 The HJM drift condition. The most important result ist the so-called HMJ drift condition
$$\boxed{\alpha_t(\tau) = \sigma_t(\tau)\int_t^\tau \sigma_t(s)ds \qquad 0 \le t \le \tau \le T}$$
Note that if the volatilities of the forward rates are multi dimensional, then the product is interpreted as scaler product, thus the condition reads as
$$\alpha_t(\tau) = \sum_{i=1}^d \sigma_t^i(\tau)\int_t^\tau \sigma_t^i(s)ds \qquad 0 \le t \le \tau \le T$$
The HJM drift condition ensures that the bond prices defined through forward rates as above yield an arbitrage free price evolution of the money market. Let us show how it works.

Given the short rate dynamics
$$df_t(\tau) = \alpha_t(\tau)dt + \sigma_t(\tau)dW_t, \qquad f_0(\tau) = f_0^*(\tau)$$
we introduce the process
$$r_t = f_t(t) \qquad t \in [0,T]$$
and define a numeraire
$$B_t = e^{\int_0^t r_u du} \qquad t \in [0,T].$$
Let us show that HJM drift condition ensures no arbtirage implying that

(6.1) $$\left(\frac{B_t(\tau)}{B_t}\right)_{t\in[0,\tau]} \quad \text{follows a martingale with respect to the measure } \mathbb{Q}. \qquad \tau \in [0,T]$$

Before we enter proof, let us perform some calculations.

Consider the stochastic process

$$X_t(\tau) = \int_t^\tau f_t(s)ds, \qquad t \in [0,\tau], \;\; \tau \in [0,T]$$

and calculate its differential

$$dX_t(\tau) = -f_t(t)dt + \int_t^\tau df_t(s)ds$$

This step requires some detailed argumentation, since the stochastic and the (weak, process-valued) integral need to be interchanged. Let us avoid technicalities and proceed further.

$$\begin{aligned} dX_t(\tau) &= -r_t dt + \int_t^\tau (\alpha_t(u)dt)du + \int_t^\tau (\sigma_t(u)dW_t)du \\ &= -r_t dt + (\int_t^\tau \alpha_t(u)du)dt + (\int_t^\tau \sigma_t(u)du)dW_t \end{aligned}$$

From this, it follows that

$$d\langle X_\cdot(\tau)\rangle_t = (\int_t^\tau \sigma_t(u)du)^2 dt$$

Now we enter the proof of (9.3), using the Ito formula to derive

$$\begin{aligned} dB_t(\tau) &= de^{-X_t(\tau)} = -e^{-X_t(\tau)}dX_t(\tau) + \frac{1}{2}e^{-X_t(\tau)}d\langle X_\cdot(\tau)\rangle_t \\ &= -B_t(\tau)\left(-r_t dt + (\int_t^\tau \alpha_t(u)du)dt + (\int_t^\tau \sigma_t(u)du)dW_t\right) \\ &\quad +\frac{1}{2}B_t(\tau)(\int_t^\tau \sigma_t(u)du)^2 dt \end{aligned}$$

Now observe that if HJM drift condition holds, then

$$\int_t^\tau \alpha_t(u)du = \frac{1}{2}(\int_t^\tau \sigma_t(u)du)^2 \qquad 0 \le t \le \tau \le T$$

Indeed, as functions in τ both expressions have the same initial value of 0 at $\tau = t$. Furthermore, both are differentiable with respect to τ and have the same derivatives due to HJM drift condition

$$\alpha_t(\tau) = \sigma_t(\tau)\int_t^\tau \sigma_t(u)du.$$

For the case of multi-dimensional forward rate volatilities, the proof adapts accordingly (scalar product with itself instead of squares!)

Thus, if HJM drift condition holds then

$$\begin{aligned} dB_t(\tau) &= -B_t(\tau)(-r_t dt + (\int_t^\tau \sigma_t(u)du)dW_t) \\ &= B_t(\tau)(r_t dt - (\int_t^\tau \sigma_t(u)du)dW_t) \end{aligned}$$

Use now the numeraire $(B_t)_{t\in[0,T]}$ to conclude that

$$\begin{aligned} d(B_t(\tau)B_t^{-1}) &= B_t^{-1}dB_t(\tau) + B_t(\tau)dB_t^{-1} \\ &= B_t^{-1}B_t(\tau)(r_t dt - (\int_t^\tau \sigma_t(u)du)dW_t) + B_t(\tau)B_t^{-1}(-r_t)dt \\ &= B_t^{-1}B_t(\tau)\cdot\left(-(\int_t^\tau \sigma_t(u)du)dW_t\right) \end{aligned}$$

That is, the HJM drift condition ensures that the discounted bond prices follow positive-valued local martingales. Under appropriate conditions on $(\sigma_t(\tau))_{t\le\tau}$ the discounted bond prices are true martingales, which excludes arbitrage. In this situation,

$$\frac{B_t(\tau)}{B_t} = \mathbb{E}_t^{\mathbb{Q}}(\frac{1}{B_\tau}) = \mathbb{E}_t^{\mathbb{Q}}(e^{-\int_0^\tau r_s ds})$$

holds, which implies that the bonds are given as previously, in terms of the short rate

$$B_t(\tau) = \mathbb{E}_t^{\mathbb{Q}}(e^{-\int_t^\tau r_s ds}), \qquad 0\le t\le\tau\le T.$$

6.4 Using HJM models. Let us present a stylized procedure which can be used in the implementation of the HJM model.

1. Specify, on your own choice, the volatility structure $(\sigma_t(\tau))_{t\in[0,\tau]}$ with $\tau\in[0,T]$ of forward rates. Good results can be achieved with deterministic volatilities with three dimensions. Some authors argue that the volatility of forward rates may decrease ($\tau\mapsto\sigma_t(\tau)$ decreasing?).

2. Determine the drift structure $(\alpha_t(\tau))_{t\in[0,\tau]}$ with $\tau\in[0,T]$ of forward rates from the HJM condition

$$\alpha_t(\tau) = \sigma_t(\tau)\int_t^\tau \sigma_t(u)du \quad \text{for all } 0\le t\le\tau\le T.$$

3. Go to the market and observe today's forward rate structure

$$f_0^*(\tau) = -\frac{\partial}{\partial\tau}\ln(B_0^*(\tau))$$

4. Compute bond prices using the formula

$$B_t(\tau) = e^{-\int_t^\tau f_t(u)du} \qquad 0\le t\le\tau\le T$$

   where $f_t(\tau)$ is a solution to

$$df_t(\tau) = \alpha_t(\tau)dt + \sigma_t(\tau)dW_t, \qquad f_0(\tau) = f_0^*(\tau).$$

5. Use the obtained dynamics $(B_t(\tau))_{t\in[0,\tau]}$, $\tau\in[0,T]$ to determine the derivatives prices (options written on bonds, caps, floors, caplets, swaptions.)

Example The Ho-Lee (short term model) turns out to be the simplest case of an HJM model. Specify the forward rate volatility

$$\sigma_t(\tau) = \sigma \in]0, \infty[\quad \text{for all } t \leq \tau$$

then the drift is determined by

$$\alpha_t(\tau) = \sigma_t(\tau) \int_t^\tau \sigma_t(s)ds = \sigma \int_t^\tau \sigma ds = \sigma^2(\tau - t)$$

Determine then the initial forward rates

$$f_0^*(\tau) = -\frac{\partial}{\partial \tau} \ln(B_0^*(\tau))$$

Calculate now the forward rates

$$\begin{aligned} f_t(\tau) &= f_0^*(\tau) + \int_0^t \alpha_u(\tau)du + \int_0^t \sigma_u(\tau)dW_u \\ &= f_0^*(\tau) + \int_0^t \sigma^2(\tau - u)du + \int_0^t \sigma dW_u \\ &= f_0^*(\tau) + \sigma^2 t(\tau - \frac{t}{2}) + \sigma W_t \end{aligned}$$

In particular, short rate is

$$r_t = f_t(t) = f_0^*(t) + \frac{\sigma^2}{2}t^2 + \sigma W_t$$

The stochastic dynamics of the short rate is

$$dr_t = (\dot{f}_0^*(t) + \sigma^2 t)dt + \sigma dW_t, \qquad r_0 = f_0^*(0).$$

This is a typical short rate dynamics of the Ho-Lee model, already fitted to the initial bond curve.

6.5 Initial forward rates estimation. In practice the market lists at any time only a finite number of bond prices

$$1 = B_0^*(0), \;\; B_0^*(\tau_1), \;\; \ldots, \;\; B_0^*(\tau_n)$$

From this, the frequent problem is that of the estimation of the continuous (differentiable) bond curve

$$B_0^*(\tau), \qquad \tau \in [0, T] \setminus \{\tau_1, \ldots, \tau_n\}.$$

This problem is solved by appropriate interpolation procedure. First consider a simple solution where the fit is given in terms of a piecewise constant forward rate curve

$$f_0^*(\tau) = \sum_{k=1}^n c_k 1_{]\tau_{k-1}, \tau_k]}(\tau), \qquad \tau \in [0, T].$$

The coefficients must be determined from

$$e^{-\int_{\tau_{k-1}}^{\tau_k} f_0^*(\tau)d\tau} = \frac{B_0^*(\tau_k)}{B_0^*(\tau_{k-1})}, \qquad \text{for all } k = 1, \dots n$$

That is, the coefficients are given by

$$c_k = \frac{\ln(\frac{B_0^*(\tau_{k-1})}{B_0^*(\tau_k)})}{\tau_k - \tau_{k-1}} \qquad \text{for all } k = 1, \dots n$$

A smooth fit is usually obtained from a parameterized family. That is one decides to model $(f_0^*(\tau))_{\tau\in[0,T]}$ by a member from a pre-specified family $\mathcal{V}$ of smooth curves (polynomials, splines) which is appropriately paremeterized. Then one needs to solve the problem

$$f_0^*(\cdot) := \operatorname{argmin}_{g\in\mathcal{V}} \left(\lambda \int_0^T |\ddot{g}(\tau)|^2 d\tau + \sum_{k=1}^n \beta_k \left| \frac{B_0^*(\tau_k)}{B_0^*(\tau_{k-1})} - e^{-\int_{\tau_{k-1}}^{\tau_k} g(\tau)d\tau} \right|^2 \right)$$

with appropriate positive parameters $\lambda, \beta_1, \dots, \beta_n$ which can be adjusted to obtain a desirable results.

6.6 Disatvantages of HJM models. Beyond its appealing style and ease of calibration, the forward rate based modeling in terms of HJM framework exhibits also clear disadvantages.

- Although fully consistent with the initial forward rates, the model suffers from the inconsistency with the bond curve evolution.

- To increase flexibility in the bond curve movements, one needs to increase the number of stochastic factors (the dimension of driving Brownian motion = dimension of the forward rate volatilities). In practice of money market modeling, three factors yield satisfactory results in many cases.

- Unlike in the setting of short term models, one loses the Markov property of the dynamics, in general. To understand this, remember that the short rate models gave the bond prices as

  $$B_t(\tau) = f^\tau(t, r_t), \quad \text{for } t \in [0, \tau] \text{ with appropriate function } f^\tau$$

  That is, knowing r_t determines all bond prices at time t in the market. This happens since the short rate dynamics $(r_t)_{t\in[0,T]}$ is a Markov process (being a strong solution of a stochastic differential equation!). That is, the knowledge of r_t determines all distributions of all bond-related random variables at time t. This is very important when it comes to the valuation of American type contracts, where optimal stopping (or more general optimal control) problems need to be solved. Basically, optimal control techniques work only for Markov processes (Backward SDE technique may be one exception). On this account, it is important to know under which conditions the short rate dynamics, derived from the HJM setting is still a Markov process.

6.7 Markov property of HJM models. Remember that defining

$$r_t = f_t(t), \qquad t \in [0,T]$$

the bond prices from HJM model satisfy

$$B_t(\tau) = \mathbb{E}^{\mathbb{Q}}(e^{-\int_t^\tau r_s ds} \,|\, \mathcal{F}_t), \qquad t \in [0,T].$$

If $(r_t)_{t\in[0,T]}$ were Markovian, then

$$\begin{aligned} B_t(\tau) &= \mathbb{E}^{\mathbb{Q}}(e^{-\int_t^\tau r_s ds} \,|\, \mathcal{F}_t) \\ &= \mathbb{E}^{\mathbb{Q}}(e^{-\int_t^\tau r_s ds} \,|\, \sigma(r_t)) = f^\tau(t, r_t) \end{aligned}$$

So we need to find conditions which ensure that

$$r_t = f_0^*(t) + \int_0^t \alpha_u(t)du + \int_0^t \sigma_u(t)dW_u$$

is a Markov process. In the setting of Gaussian HJM, where

the forward rate volatilities are $\sigma_t(\tau)$
and deterministic for all $0 \le t \le \tau \le T$

the Markov property can be ensured by the following assumption on the structure of forward rate volatilities

there exist deterministic functions ξ and ψ
such that $\sigma_s(t) = \xi(s)\psi(t)$ holds for all $s \le t$

To show this result, we need to discuss only the Markovian property of the stochastic part of $(r_t)_{t\in[0,T]}$, this ist the process

$$Z_t = \int_0^t \sigma_s(t)dW_s, \qquad t \in [0,T].$$

Consider the calculation

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}}(Z_t \,|\, \mathcal{F}_u) &= \mathbb{E}^{\mathbb{Q}}(\int_0^t \sigma_s(t)dW_s \,|\, \mathcal{F}_u) \\ &= \mathbb{E}^{\mathbb{Q}}(\int_0^u \sigma_s(t)dW_s \,|\, \mathcal{F}_u) + \underbrace{\mathbb{E}^{\mathbb{Q}}(\int_u^t \sigma_s(t)dW_s \,|\, \mathcal{F}_u)}_{=0} \\ &= \int_0^u \sigma_s(t)dW_s = \psi(t)\int_0^u \xi(s)dW_s \\ &= \frac{\psi(t)}{\psi(u)}\int_0^u \underbrace{\xi(s)\psi(u)}_{\sigma_s(u)}dW_s = \frac{\psi(t)}{\psi(u)}Z_u \end{aligned}$$

Using the tower property of conditional expectation, we come close to the Markovian property:

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}(Z_t \mid \sigma(Z_u)) &= \mathbb{E}^{\mathbb{Q}}(\mathbb{E}^{\mathbb{Q}}(Z_t \mid \mathcal{F}_u) \mid \sigma(Z_u)) \\ &= \mathbb{E}^{\mathbb{Q}}(\frac{\psi(t)}{\psi(u)} Z_u \mid \sigma(Z_u)) \\ &= \frac{\psi(t)}{\psi(u)} Z_u = \mathbb{E}^{\mathbb{Q}}(Z_t \mid \mathcal{F}_u)\end{aligned}$$

Exercise Complete the proof of Markovian property by showing that

$$\mathbb{E}^{\mathbb{Q}}(h(Z_t) \mid \mathcal{F}_u) = \mathbb{E}^{\mathbb{Q}}(h(Z_t) \mid \sigma(Z_u))$$

holds for each bounded measurable function h.

That is,

$$\mathbb{E}^{\mathbb{Q}}(Z_t \mid \mathcal{F}_u) = \frac{\psi(t)}{\psi(u)} Z_u = \mathbb{E}^{\mathbb{Q}}(Z_t \mid \sigma(Z_u)) \quad \text{for all } u \leq t$$

which shows that $(Z_t)_{t\in[0,T]}$ and so $(r_t)_{t\in[0,T]}$ are indeed Markovian processes.

Finally, consider the relation of one-dimensional Markovian HJM models to the Hull-White short term models. Given forward rate volatilities as above, under the factorization condition

$$\sigma_s(t) = \xi(s)\psi(t), \qquad s \leq t$$

calculate the drift from HJM condition

$$\begin{aligned}\alpha_s(t) &= \sigma_s(t) \int_s^t \sigma_s(u)du \\ &= \xi(s)\psi(t) \int_s^t \xi(s)\psi(u)du = \xi(s)^2\psi(t) \int_s^t \psi(u)du\end{aligned}$$

Hence the integral is

$$\int_0^t \alpha_s(t)ds = \psi(t) \int_0^t \left(\xi(s)^2 \int_s^t \psi(u)du\right) ds$$

then

$$r_t = f_t(t) = \underbrace{f_0^*(t) + \int_0^t \alpha_s(t)ds}_{U(t)} + \underbrace{\int_0^t \sigma_s(t)dW_s}_{\psi(t)\int_0^t \xi(s)dW_s}$$

with this

$$\begin{aligned}dr_t &= \dot{U}_t dt + \underbrace{\dot{\psi}(t)\left(\int_0^t \xi(s)dW_s\right)}_{(r_t - U(t))\frac{\dot{\psi}(t)}{\psi(t)}} dt + \psi(t)\xi(t)dW_t \\ &= \left[\underbrace{\dot{U}(t) - U(t)\frac{\dot{\psi}(t)}{\psi(t)}}_{\theta(t)} + r_t \underbrace{\frac{\dot{\psi}(t)}{\psi(t)}}_{-a(t)}\right] dt + \underbrace{\psi(t)\xi(t)}_{\sigma(t)} dW_t\end{aligned}$$

The dynamics $dr_t = (\theta(t) - a(t)r_t)dt + \sigma(t)dW_t$ is addressed within Hull-White setting (slight extension of Hull-White I).

7 The Girsanov transform and the change of numeraire

7.1 Equivalent probability measures. Two probability measures $\mathbb{P}$ and $\mathbb{Q}$ on a measure space $(\Omega, \mathcal{F})$ are called equivalent (written as $\mathbb{P} \sim \mathbb{Q}$) if they have the same zero sets

$$\mathbb{P}(F) = 0 \iff \mathbb{Q}(F) = 0 \text{ for each } F \in \mathcal{F}$$

A non-trivial result about equivalent measures is that they are always connected by Radom-Nikody density. That is if $\mathbb{P} \sim \mathbb{Q}$ then there exists a positive-valued random variable

$$\frac{d\mathbb{P}}{d\mathbb{Q}} : \Omega \to]0, \infty[$$

which satisfies

$$\mathbb{E}^{\mathbb{Q}}(\frac{d\mathbb{P}}{d\mathbb{Q}} 1_F) = \mathbb{E}^{\mathbb{P}}(1_F) \qquad \text{for each } F \in \mathcal{F}.$$

The Radon-Nikodym derivative is uniquely defined by this equation. That is, if a random variable X satisfies

$$\mathbb{E}^{\mathbb{Q}}(X 1_F) = \mathbb{E}^{\mathbb{P}}(1_F) \qquad \text{for each } F \in \mathcal{F}.$$

then

$$X = \frac{d\mathbb{P}}{d\mathbb{Q}} \quad \text{almost surely.}$$

An interesting situation occurs when the measures are restricted. Say, given a sub-sigma-algebra $\mathcal{G} \subset \mathcal{F}$, we can restrict the domains of definition of $\mathbb{P}$ and $\mathbb{Q}$ to $\mathcal{G}$ obtaining the measures $P|_{\mathcal{G}}$ and $\mathbb{Q}|_{\mathcal{G}}$ on $(\Omega, \mathcal{G})$ in this way

$$\mathbb{P}|_{\mathcal{G}}(G) = \mathbb{P}(G) \quad \mathbb{Q}|_{\mathcal{G}}(G) = \mathbb{Q}(G) \quad \text{for all } G \in \mathcal{G}$$

Obviously, $\mathbb{P}|_{\mathcal{G}}$ and $\mathbb{Q}|_{\mathcal{G}}$ are equivalent, since they share the same zero sets. According to the existence of the Radon-Nikodym derivative, there must be a $\mathcal{G}$-measurable random variable $\frac{d\mathbb{P}|_{\mathcal{G}}}{d\mathbb{Q}|_{\mathcal{G}}}$ such that

$$\mathbb{E}^{\mathbb{Q}}(\frac{d\mathbb{P}|_{\mathcal{G}}}{d\mathbb{Q}|_{\mathcal{G}}} 1_G) = \mathbb{E}^{\mathbb{P}}(1_G) \qquad \text{for each } G \in \mathcal{G}.$$

It turns out that it is given by the conditional expectation (with respect to $\mathbb{Q}$) of the previous Radon-Nikodym derivative

$$\frac{d\mathbb{P}|_{\mathcal{G}}}{d\mathbb{Q}|_{\mathcal{G}}} = \mathbb{E}^{\mathbb{Q}}(\frac{d\mathbb{P}}{d\mathbb{Q}} \,|\, \mathcal{G})$$

To verify this assertion, it ensures to check that the property of Radon-Nikodym derivative $\frac{d\mathbb{P}|_{\mathcal{G}}}{d\mathbb{Q}|_{\mathcal{G}}}$ is fulfilled by the above conditional expectation.

$$\begin{aligned}
\mathbb{E}^{Q|_{\mathcal{G}}}(\mathbb{E}^{Q|_{\mathcal{G}}}(\frac{d\mathbb{P}}{d\mathbb{Q}}\,|\,\mathcal{G})1_G) &= \mathbb{E}^{Q|_{\mathcal{G}}}(\mathbb{E}^{Q|_{\mathcal{G}}}(\frac{d\mathbb{P}}{d\mathbb{Q}}1_G\,|\,\mathcal{G})) \\
&= \mathbb{E}^{\mathbb{Q}}(\frac{d\mathbb{P}}{d\mathbb{Q}}1_G) \\
&= \mathbb{E}^{\mathbb{P}}(1_G) = \mathbb{E}^{P|_{\mathcal{G}}}(1_G)
\end{aligned}$$

Example Let $(\Omega, \mathcal{F})$ be finite measure space, consider two measures P and Q which satisfy

$$P(\{\omega\}) > 0, \quad \mathbb{Q}(\{\omega\}) > 0, \quad \text{for all } \omega \in \Omega$$

Hence P and Q are equivalent. Consider a finite partition of Ω

$$\{G_1, \ldots, G_n\} \subset \mathcal{F}, \quad G_i \cap G_j = \emptyset \quad \text{if } i \neq j, \quad \cup_{i=1}^n G_i = \Omega$$

which yields sigma-algebra $\mathcal{G}$ consisting of all possible unions of partition sets. Now, we have

$$\begin{aligned}
\frac{d\mathbb{P}}{d\mathbb{Q}} &= \sum_{\omega\in\Omega} \frac{P(\{\omega\})}{\mathbb{Q}(\{\omega\})} 1_{\{\omega\}} \\
\frac{d\mathbb{P}|_{\mathcal{G}}}{d\mathbb{Q}|_{\mathcal{G}}} &= \sum_{i=1}^{n} \frac{P(G_i)}{\mathbb{Q}(G_i)} 1_{G_i} \\
\mathbb{E}^{\mathbb{Q}}(\frac{d\mathbb{P}}{d\mathbb{Q}}\,|\,\mathcal{G}) &= \sum_{i=1}^{n} \frac{\mathbb{E}^{\mathbb{Q}}(\frac{d\mathbb{P}}{d\mathbb{Q}}1_{G_i})}{\mathbb{Q}(G_i)} 1_{G_i} \\
&= \sum_{i=1}^{n} \frac{P(G_i)}{\mathbb{Q}(G_i)} 1_{G_i}
\end{aligned}$$

The important result about measure change is the so-called Bayes Formula, which states that to calculate conditional expectation of a random variable X with respect to measure $\mathbb{P} \sim \mathbb{Q}$, one can calculate the conditional expectation of X with respect to $\mathbb{Q}$ and then to correct the result.

$$\mathbb{E}^{\mathbb{P}}(X\,|\,\mathcal{G}) = \frac{\mathbb{E}^{\mathbb{Q}}(\frac{d\mathbb{P}}{d\mathbb{Q}}X\,|\,\mathcal{G})}{\mathbb{E}^{\mathbb{Q}}(\frac{d\mathbb{P}}{d\mathbb{Q}}\,|\,\mathcal{G})}$$

The proof is based on the above insights, how Radon-Nikodym derivative changes, when the measures are restricted. Obviously, we need to show

$$\mathbb{E}^{\mathbb{P}}(X\,|\,\mathcal{G})\mathbb{E}^{\mathbb{Q}}(\frac{d\mathbb{P}}{d\mathbb{Q}}\,|\,\mathcal{G}) = \mathbb{E}^{\mathbb{Q}}(\frac{d\mathbb{P}}{d\mathbb{Q}}X\,|\,\mathcal{G})$$

a sufficient condition therefore is that for each $G \in \mathcal{G}$ holds

$$\mathbb{E}^{\mathbb{Q}}\left(\mathbb{E}^{\mathbb{P}}(X\,|\,\mathcal{G})\mathbb{E}^{\mathbb{Q}}(\frac{d\mathbb{P}}{d\mathbb{Q}}\,|\,\mathcal{G})1_G\right) = \mathbb{E}^{\mathbb{Q}}\left(\mathbb{E}^{\mathbb{Q}}(\frac{d\mathbb{P}}{d\mathbb{Q}}X\,|\,\mathcal{G})1_G\right)$$

The left-hand side yields

$$
\begin{aligned}
\mathbb{E}^{\mathbb{Q}}\left(\mathbb{E}^{\mathbb{P}}(X\,|\,\mathcal{G})\mathbb{E}^{\mathbb{Q}}(\frac{d\mathbb{P}}{d\mathbb{Q}}\,|\,\mathcal{G})1_G\right) &= \mathbb{E}^{\mathbb{Q}}\left(\mathbb{E}^{\mathbb{P}}(X\,|\,\mathcal{G})\frac{d\mathbb{P}|_{\mathcal{G}}}{d\mathbb{Q}|_{\mathcal{G}}}1_G\right) \\
&= \mathbb{E}^{Q|_{\mathcal{G}}}\left(\mathbb{E}^{\mathbb{P}}(X\,|\,\mathcal{G})\frac{d\mathbb{P}|_{\mathcal{G}}}{d\mathbb{Q}|_{\mathcal{G}}}1_G\right) \\
&= \mathbb{E}^{P|_{\mathcal{G}}}\left(\mathbb{E}^{\mathbb{P}}(X\,|\,\mathcal{G})1_G\right) = \mathbb{E}^{\mathbb{P}}\left(X1_G\right)
\end{aligned}
$$

The right-hand side yields

$$
\begin{aligned}
\mathbb{E}^{\mathbb{Q}}\left(\mathbb{E}^{\mathbb{Q}}(\frac{d\mathbb{P}}{d\mathbb{Q}}X\,|\,\mathcal{G})1_G\right) &= \mathbb{E}^{\mathbb{Q}}\left(\mathbb{E}^{\mathbb{Q}}(\frac{d\mathbb{P}}{d\mathbb{Q}}X1_G\,|\,\mathcal{G})\right) \\
&= \mathbb{E}^{\mathbb{Q}}\left(\frac{d\mathbb{P}}{d\mathbb{Q}}X1_G\right) = \mathbb{E}^{\mathbb{P}}\left(X1_G\right)
\end{aligned}
$$

7.2 Change of measure in the setting of Girsanov theorem. Suppose that the filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in[0,T]})$ supports the process $(W_t)_{t\in[0,T]}$ of (multivariate) Brownian motion. Consider the process $(\varphi_t)_{t\in[0,T]}$, which we call *Girsanov kernel*. With the Girsanov kernel we asscociate the process

$$
L_t = e^{\overbrace{\int_0^t \varphi_s dW_s - \frac{1}{2}\int_0^t |\varphi_s|^2 ds}^{}_{X_t}}.
$$

It turns out that this process is a local martingale, which is seen from its stochastic differential

$$
\begin{aligned}
dL_t &= de^{X_t} = e^{X_t}dX_t + \frac{1}{2}e^{X_t}d\langle X\rangle_t \\
&= L_t(\varphi_t dW_t - \frac{1}{2}|\varphi_t|^2dt + \frac{1}{2}|\varphi_t|^2dt) = L_t\varphi_t dW_t
\end{aligned}
$$

An important question is to decide whether $(L_t)_{t\in[0,T]}$ is a true martingale. It turns out that this is a difficult question. The Novikov condition

$$
\mathbb{E}(e^{\frac{1}{2}\int_0^T|\varphi_s|^2ds}) < \infty
$$

is sufficient to ensure the martingale property of $(L_t)_{t\in[0,T]}$. If $(L_t)_{t\in[0,T]}$ is a true martingale, then obviously

$$
\mathbb{E}(L_T) = \mathbb{E}(L_0) = L_0 = 1
$$

In this settings, we can use L_T as Radon-Nikodym derivative, to define another equivalent measure $\mathbb{Q} \sim \mathbb{P}$ by

$$
Q(F) := \mathbb{E}^{\mathbb{P}}(L_T 1_F) \qquad \text{for all } F \in \mathcal{F}.
$$

with this, we have

$$
\frac{d\mathbb{Q}}{d\mathbb{P}} = L_T.
$$

Naturally, it turns out that with respect to $\mathbb{Q}$ the process $(W_t)_{t\in[0,T]}$ is not a Brownian motion. But it can be corrected to a Brownian motion. Namely, the Grisanov theorem states that the process

$$W_t^{\mathbb{Q}} = W_t - \int_0^t \varphi_s ds, \qquad t \in [0,T]$$

is a Brownian motion with respect to $\mathbb{Q}$.

The idea in the verification of this result is to check

- that

  (7.1) $\qquad (W_t^{\mathbb{Q}})_{t\in[0,T]}$ follows a local martingale with respect to $\mathbb{Q}$.

  This will be addressed later

- the quadratic variation satisfies

  (7.2) $$\langle W^{\mathbb{Q}}\rangle_t = \langle W\rangle_t = t$$

  This is obvious, since the processes $(W_t^{\mathbb{Q}})_{t\in[0,T]}$ and $(W_t)_{t\in[0,T]}$ differ only by the process $(\int_0^t \varphi_s ds)_{t\in[0,T]}$ which is a bounded variation process.

Combining the assertion (7.1) and (7.2), we obtain with Levy's characterization that $(W_t^{\mathbb{Q}})_{t\in[0,T]}$ is a Brownian motion with respect to $\mathbb{Q}$.

Now let us argue why (7.1) holds. Basically, to show that $(W_t^{\mathbb{Q}})_{t\in[0,T]}$ follows a martingale with respect to $\mathbb{Q}$, it suffices to show that

(7.3) $\qquad (W_t^{\mathbb{Q}} L_t)_{t\in[0,T]}$ is a local martingale with respect to $\mathbb{P}$,

since in general it holds that

(7.4) $\qquad$ if $(Y_t L_t)_{t\in[0,T]}$ is a (local) martingale with respect to $\mathbb{P}$,
then $(Y_t)_{t\in[0,T]}$ is a (local) martingale with respect to $\mathbb{Q}$

Let us show (7.4) under the assumption that $(Y_t L_t)_{t\in[0,T]}$ is a true martingale with respect to $\mathbb{P}$, the local martingale case can be carried out using stopping techniques.

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}}(Y_t \,|\, \mathcal{F}_s) &= \frac{\mathbb{E}^{\mathbb{P}}(\frac{d\mathbb{Q}}{d\mathbb{P}} Y_t \,|\, \mathcal{F}_s)}{\mathbb{E}^{\mathbb{P}}(\frac{d\mathbb{Q}}{d\mathbb{P}} \,|\, \mathcal{F}_s)} \\
&= \frac{\mathbb{E}^{\mathbb{P}}(L_T Y_t \,|\, \mathcal{F}_s)}{\mathbb{E}^{\mathbb{P}}(L_T \,|\, \mathcal{F}_s)} = \frac{\mathbb{E}^{\mathbb{P}}(L_t Y_t \,|\, \mathcal{F}_s)}{L_s} = \frac{L_s Y_s}{L_s} = Y_s
\end{aligned}$$

Finally, a direct inspection of the stochastic differential $d(W_t^{\mathbb{Q}} L_t)$ shows that it is a local martingale with respect to P.

$$\begin{aligned}
d(W_t^{\mathbb{Q}} L_t) &= W_t^{\mathbb{Q}} dL_t + L_t dW_t^{\mathbb{Q}} + d\langle W_\cdot^{\mathbb{Q}}, L_\cdot\rangle_t \\
&= W_t^{\mathbb{Q}} dL_t + L_t dW_t - L_t\varphi_t dt + d\langle W_\cdot, L_\cdot\rangle_t \\
&= W_t^{\mathbb{Q}} dL_t + L_t dW_t - L_t\varphi_t dt + L_t\varphi_t dt \\
&= W_t^{\mathbb{Q}} dL_t + L_t dW_t
\end{aligned}$$

Example Given the process $(W_t)_{t\in[0,T]}$ of a Brownian motion with respect to P, consider the process $(W_t - t)_{t\in[0,T]}$. Obviously it does not behave like a Brownian motion since too many trajectories drift down. Still looking at a single trajectory one realizes that it looks like an (unlikely) trajectory of Brownian motion. The Girsanov theorem suggest to change the probabilistic view on trajectories: correct the occurrence of down-drifting trajectories, to make them 'less likely' then, from the new perspective the process $(W_t - t)_{t\in[0,T]}$ looks like a Brownian motion again. That is, basically the Radon-Nikodym density

$$L_T = e^{W_T - \frac{1}{2}T}$$

penalizes the trajectories drifting down: L_T is high in the events when W_T is high and low in the events when W_T is low, the factor $e^{-\frac{1}{2}T}$ stands only for the normalization.

7.3 Change of numeraire revisited. Remember that in order to exclude arbitrage, asset prices are modeled such that such that the wealth of the trading strategy, after discounting by numeraire asset follows a martingale (with respect to appropriate risk-neutral measure). Having established such a model, is some situations, one needs to change the numeraire asset, choosing one of the tradable securities as the new numeraire. Here, one needs to understand how this affects the risk neutral measure.

Consider positive-valued asset price evolutions $(N^1_t)_{t\in[0,T]}$, $(N^2_t)_{t\in[0,T]}$, and $(X_t)_{t\in[0,T]}$. We interpret $(N^1_t)_{t\in[0,T]}$ as a native numeraire asset, whereas $(X_t)_{t\in[0,T]}$ and $(N^2_t)_{t\in[0,T]}$ are tradable securities with

$$(\tfrac{X_t}{N^1_t})_{t\in[0,T]}, \quad (\tfrac{N^2_t}{N^1_t})_{t\in[0,T]} \text{ are } \mathbb{Q}^1\text{-martingales}$$

Assume that $(N^2_t)_{t\in[0,T]}$ is proposed as new numeraire asset. So we will discount the price evolutions as

$$(\tfrac{X_t}{N^2_t})_{t\in[0,T]}, \quad (\tfrac{N^1_t}{N^2_t})_{t\in[0,T]}$$

The question is, with respect to which measure this processes are martingales. The change of numeraire technique answers this question as follows

$$\begin{gathered}\text{if } (\tfrac{X_t}{N^1_t})_{t\in[0,T]}, \quad (\tfrac{N^2_t}{N^1_t})_{t\in[0,T]} \text{ are } \mathbb{Q}^1\text{-martingales, then} \\ (\tfrac{X_t}{N^2_t})_{t\in[0,T]}, \quad (\tfrac{N^1_t}{N^2_t})_{t\in[0,T]} \text{ are } \mathbb{Q}^2\text{-martingales} \\ \text{where } d\mathbb{Q}^2 = \tfrac{N^2_T}{N^1_T}\tfrac{N^1_0}{N^2_0} d\mathbb{Q}^1.\end{gathered}$$

The proof is straight forward and utilizes the behavior of the conditional expectation under change of measure

$$\mathbb{E}^{\mathbb{Q}^2}_t(X) = \frac{\mathbb{E}^{\mathbb{Q}^1}_t(\frac{d\mathbb{Q}^2}{d\mathbb{Q}^1}X)}{\mathbb{E}^{\mathbb{Q}^1}_t(\frac{d\mathbb{Q}^2}{d\mathbb{Q}^1})}$$

Namely, for $s \leq t$

$$
\begin{aligned}
\mathbb{E}_s^{\mathbb{Q}^2}(\frac{X_t}{N_t^2}) &= \frac{\mathbb{E}_s^{\mathbb{Q}^1}(\frac{N_T^2}{N_T^1}\frac{\cancel{N_0^1}}{\cancel{N_0^2}}\frac{X_t}{N_t^2})}{\mathbb{E}_s^{\mathbb{Q}^1}(\frac{N_T^2}{N_T^1}\frac{\cancel{N_0^1}}{\cancel{N_0^2}})} \\
&= \frac{\mathbb{E}_s^{\mathbb{Q}^1}(\frac{\cancel{N_t^2}}{N_t^1}\frac{X_t}{\cancel{N_t^2}})}{\frac{N_s^2}{N_s^1}} = \frac{\frac{X_s}{N_s^1}}{\frac{N_s^2}{N_s^1}} = \frac{X_s}{N_s^2}
\end{aligned}
$$

In the framework of Ito processes, the change of numearaire technique can be very powerful if one utilizes the concrete structure of positive-valued martingales as exponential martingales and uses the Girsanov transform.

Note that the above change of numeraire can be equivalently formulated as

$$
\begin{gathered}
\text{If } (X_t)_{t\in[0,T]}, \quad (N_t)_{t\in[0,T]} \text{ are } \mathbb{Q}\text{-martingales} \\
\text{then } (X_t N_t^{-1})_{t\in[0,T]}, \quad (N_t^{-1})_{t\in[0,T]} \text{ are } \mathbb{Q}^N\text{-martingales} \\
\text{where } d\mathbb{Q}^N = N_T N_0^{-1} dQ.
\end{gathered} \tag{7.5}
$$

Indeed, put N_t instead of N_t^2/N_t^1, and X_t instead of X_t/N_t^1. Alternatively, put $N_t^1 = 1$ and $N_t^2 = N_t$.

In the case of Ito processes, this change-of-numeraire transformation has more structure. Namely, a positive valued martingale $(Y_t)_{t\in[0,T]}$ frequently follows

$$
dY_t = Y_t \sigma_t^Y dW_t \tag{7.6}
$$

with an appropriate process $(\sigma_t^Y)_{t\in[0,T]}$. Processes of the form (7.6) are called exponential martingales (exponential local martingales if $(Y_t)_{t\in[0,T]}$ is local martingale). One can show that if the filtration is generated by Brownian motion (as in the martingale representation theorem), then each positive-valued (local) martingale is of this form. In this setting the process $(\sigma_t^Y)_{t\in[0,T]}$ is called the *stochastic logarithm.*

Suppose that $(W_t)_{t\in[0,T]}$ is a Brownian motion with respect to $\mathbb{Q}$. If both, $(X_t)_{t\in[0,T]}$ and $(N_t)_{t\in[0,T]}$ in (7.5) are exponential martingales with stochastic logarithms $(\sigma_t^X)_{t\in[0,T]}$ and $(\sigma_t^N)_{t\in[0,T]}$ respectively

$$
dX_t = X_t \sigma_t^X dW_t, \quad dN_t = N_t \sigma_t^N dW_t,
$$

then the fraction $(X_t N_t^{-1})_{t\in[0,T]}$ and the reciprocal $(N_t^{-1})_{t\in[0,T]}$ are also an exponential martingales

$$
d(X_t N_t^{-1}) = (X_t N_t^{-1})(\sigma_t^X - \sigma_t^N) dW_t^N, \quad dN_t^{-1} = N_t^{-1}(-\sigma_t^N) dW_t^N \tag{7.7}
$$

where

$$
(W_t^N = W_t - \textstyle\int_0^t \sigma_u^N du)_{t\in[0,T]} \text{ is a } \mathbb{Q}^N\text{-Brownian motion} \tag{7.8}
$$

Remark Note that taking fraction of exponential martingales yields exponential martingale whose stochastic logarithm equals to the difference of the stochastic logarithm of the numerator less the stochastic logarithm of the denominator. However, take care that the measure changes and the Brownian motion changes, too.

Let us verify the statements (7.7) and (7.8). The assertion (7.8) follows from Girsanov's theorem. The statement (7.7) is directly verified by using the Ito's formula

$$\begin{aligned}
dN_t^{-1} &= (-1)N_t^{-2}dN_t + \frac{1}{2}2N_t^{-3}d\langle N\rangle_t \\
&= -N_t^{-2}N_t\sigma_t^N dW_t + N_t^{-3}N_t^2\sigma_t^N \circ \sigma_t^N dt \\
&= N_t^{-1}(-\sigma_t^N)\underbrace{(dW_t - \sigma_t^N dt)}_{dW_t^N} \\
d(X_tN_t^{-1}) &= X_t dN_t^{-1} + N^{-1}dX_t + d\langle X, N^{-1}\rangle_t \\
&= X_tN_t^{-1}(-\sigma_t^N)dW_t^N + N_t^{-1}X_t\sigma_t^X(dW_t^N + \sigma_t^N dt) + X_tN_t^{-1}\sigma_t^X \circ (-\sigma_t^N)dt \\
&= X_tN_t^{-1}(\sigma_t^X - \sigma_t^N)dW_t^N
\end{aligned}$$

7.4 Forward and short rate. Remember the expectation hypothesis. It claims that the forward rate equals to the expectation of the short rate. In the framework of HJM modeling, this would mean

$$f_0(\tau) = \mathbb{E}(\underbrace{f_\tau(\tau)}_{r_\tau}) \qquad 0 \le t \le \tau \le T.$$

This is, of course, not true. Neither with respect to the objective probability measure (what is this?) nor with respect to the risk neutral measure $\mathbb{Q}$ underlying the construction of the HJM models. Surprisingly, the expectation hypothesis is true in the HJM framework, if one takes the so-called forward measure $\mathbb{Q}^\tau$ associated with the numeraire $(B_t(\tau))_{t\in[0,\tau]}$.

Namely, the numeraire $(B_t(\tau))_{t\in[0,\tau]}$ corresponds to the measure $\mathbb{Q}^\tau$ determined by

$$d\mathbb{Q}^\tau = \frac{B_\tau(\tau)}{B_\tau}\frac{B_0}{B_0(\tau)}dQ = \frac{1}{B_\tau}\frac{1}{B_0(\tau)}dQ$$

With this, we obtain

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}^\tau}(r_\tau) &= \mathbb{E}^{\mathbb{Q}}(\frac{1}{B_\tau}\frac{1}{B_0(\tau)}r_\tau) = \frac{1}{B_0(\tau)}\mathbb{E}^{\mathbb{Q}}(\frac{r_\tau}{B_\tau}) \\
&= \frac{1}{B_0(\tau)}\mathbb{E}^{\mathbb{Q}}(r_\tau e^{-\int_0^\tau r_s ds}) = -\frac{1}{B_0(\tau)}\mathbb{E}^{\mathbb{Q}}(\frac{\partial}{\partial x}|_{x=\tau} e^{-\int_0^x r_s ds}) \\
&= -\frac{1}{B_0(\tau)}\frac{\partial}{\partial x}|_{x=\tau}\underbrace{\mathbb{E}^{\mathbb{Q}}(e^{-\int_0^x r_s ds})}_{B_0(x)} = -\frac{1}{B_0(\tau)}\frac{\partial}{\partial x}|_{x=\tau} e^{-\int_0^x f_0(s)ds} \\
&= f_0(\tau).
\end{aligned}$$

7.5 Call pricing under stochastic interest rates. Consider the valuation of a Call written on the terminal value of the asset $(S_t)_{t\in[0,T]}$ under the assumption that the interest rates are stochastic. The option price at time $t = 0$ is given by

$$
\begin{aligned}
\mathbb{E}^{\mathbb{Q}}(\frac{1}{B_T}(S_T - K)^+) &= \mathbb{E}^{\mathbb{Q}}(\frac{1}{B_T}(S_T - K)1_{\{S_T \geq K\}}) \\
&= S_0 \mathbb{E}^{\mathbb{Q}}(\frac{S_T}{B_T}\frac{B_0}{S_0}1_{\{S_T \geq K\}}) - K\mathbb{E}^{\mathbb{Q}}(\frac{B_T(T)}{B_T}\frac{B_0}{B_0(T)}1_{\{S_T \geq K\}})B_0(T) \\
&= S_0 Q^S(S_T \geq K) - KB_0(T)\mathbb{Q}^T(S_T \geq K) \qquad (7.9)
\end{aligned}
$$

Note that

- the native risk measure $\mathbb{Q}$, which corresponds to the discounting asset $(B_t)_{t\in[0,T]}$ (continuously compounded savings account) is called the *spot martingale measure*
- the measure $\mathbb{Q}^T$, which corresponds to the discounting asset $(B_t(T))_{t\in[0,T]}$ (zero bond with maturity T) is called the *T-forward martingale measure*
- the measure Q^S, which corresponds to the discounting asset $(S_t)_{t\in[0,T]}$ (stock price) does not have any specific name, we called it *S-martingale measure*

The formula (7.9) becomes useful when both, the expectation under the spot and under the forward measure can be calculated.

For instance, suppose that under the spot martingale measure holds

$$
\begin{gathered}
d(\tfrac{S_t}{B_t}) = (\tfrac{S_t}{B_t})\sigma_t dW_t \qquad d(\tfrac{B_t(T)}{B_t}) = (\tfrac{B_t(T)}{B_t})\sigma_t(T) dW_t \\
\text{such that } (\Sigma_t = \sigma_t - \sigma_t(T))_{t\in[0,T]} \text{ is deterministic}
\end{gathered} \qquad (7.10)
$$

where $(W_t)_{t\in[0,T]}$ follows a Brownian motion with respect to the spot martingale measure $\mathbb{Q}$. Then one can calculate the option price (7.9) explicitly.

First consider $\mathbb{Q}^T(S_T \geq K)$. According to the change of numeraire, the process $(\frac{S_t}{B_t(T)})_{t\in[0,T]}$ follows a martingale with respect to $\mathbb{Q}^T$ with terminal value

$$
S_T = \frac{S_T}{B_T(T)} = \frac{S_0}{B_0(T)} e^{\int_0^T \Sigma_u dW_u^T - \frac{1}{2}\int_0^T |\Sigma_u|^2 du}
$$

where $(W_u^T)_{t\in[0,T]}$ is a process of Brownian motion under $\mathbb{Q}^T$. Hence, we obtain

$$
\begin{aligned}
\mathbb{Q}^T(S_T \geq K) &= \mathbb{Q}^T(\frac{S_0}{B_0(T)} e^{\int_0^T \Sigma_u dW_u^T - \frac{1}{2}\int_0^T |\Sigma_u|^2 du} \geq K) \\
&= \mathbb{Q}^T(\int_0^T \Sigma_u dW_u^T \geq \ln(\frac{KB_0(T)}{S_0}) + \frac{1}{2}\int_0^T |\Sigma_u|^2 du) \\
&= \mathbb{Q}^T(\underbrace{\frac{-\int_0^T \Sigma_u dW_u^T}{\sqrt{\int_0^T |\Sigma_u|^2 du}}}_{\text{standard normal}} \leq \underbrace{\frac{\ln(\frac{S_0}{KB_0(T)}) - \frac{1}{2}\int_0^T |\Sigma_u|^2 du}{\sqrt{\int_0^T |\Sigma_u|^2 du}}}_{d_2}) \\
&= \Phi(d_2)
\end{aligned}
$$

Now, we turn to

$$Q^S(S_T \geq K) = Q^S(\frac{B_T(T)}{S_T} \leq \frac{1}{K})$$

According to the change of numeraire transformation, the process $(\frac{B_t(T)}{S_t})_{t\in[0,T]}$ follows a martingale with respect to Q^S with terminal value

$$\frac{1}{S_T} = \frac{B_T(T)}{S_T} = \frac{B_0(T)}{S_0} e^{-\int_0^T \Sigma_u dW_u^S - \frac{1}{2}\int_0^T |\Sigma_u|^2 du}$$

where $(W_u^S)_{t\in[0,T]}$ is a process of Brownian motion under the measure Q^S which is given by

$$dQ^S = \frac{S_T}{B_T}\frac{B_0}{S_0} dQ.$$

Hence, we obtain

$$\begin{aligned}
Q^S(\frac{1}{S_T} \leq \frac{1}{K}) &= Q^S(\frac{B_0(T)}{S_0} e^{-\int_0^T \Sigma_u dW_u^S - \frac{1}{2}\int_0^T |\Sigma_u|^2 du} \leq \frac{1}{K}) \\
&= Q^S(-\int_0^T \Sigma_u dW_u^S \leq \ln(\frac{S_0}{KB_0(T)}) + \frac{1}{2}\int_0^T |\Sigma_u|^2 du) \\
&= Q^S(\underbrace{\frac{-\int_0^T \Sigma_u dW_u^S}{\sqrt{\int_0^T |\Sigma_u|^2 du}}}_{\text{standard normal}} \leq \underbrace{\frac{\ln(\frac{S_0}{KB_0(T)}) + \frac{1}{2}\int_0^T |\Sigma_u|^2 du}{\sqrt{\int_0^T |\Sigma_u|^2 du}}}_{d_1}) \\
&= \Phi(d_1)
\end{aligned}$$

With this, the option price under stochastic interest rates and under the assumption (7.10) is

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}}(\frac{1}{B_T}(S_T - K)^+) &= S_0 Q^S(S_T \geq K) - KB_0(T)\mathbb{Q}^T(S_T \geq K) &\quad (7.11)\\
&= S_0\Phi(d_1) - KB_0(T)\Phi(d_2) &\quad (7.12)
\end{aligned}$$

7.6 Bond option in the Hull-White model. Consider the short rate dynamics from the Hull-White model

$$dr_t = (\theta(t) - ar_t)dt + \sigma dW_t, \qquad \begin{array}{c} \sigma, a > 0 \text{ deterministic}, \\ (\theta(t))_{t\in[0,T]} \text{ deterministic function} \end{array}$$

Due to affine term structure, the bond prices are given as

$$B_t(\tau) = e^{\mathcal{A}^\tau(t) - \mathcal{B}^\tau(t) r_t}, \qquad t \in [0, \tau], \ \ \tau \in [0, T]$$

where

$$\mathcal{B}^\tau(t) = \frac{1 - e^{-a(\tau - t)}}{a} \qquad t \in [0, \tau], \ \ \tau \in [0, T].$$

Note that this term determines the volatility of the bond, since according to the Ito formula

$$\begin{aligned}
dB_t(\tau) &= (\dot{\mathcal{A}}^\tau(t) - \dot{\mathcal{B}}^\tau(t) r_t) B_t(\tau) dt - \mathcal{B}^\tau(t) B_t(\tau) dr_t + \frac{1}{2} B_t(\tau) \mathcal{B}^\tau(t)^2 d\langle r_\cdot \rangle_t \\
&= B_t(\tau) r_t dt - \mathcal{B}^\tau(t) B_t(\tau) \sigma dW_t
\end{aligned}$$

Thus, zero bond with maturity τ discounted by the standard saving account follows martingale dynamics

$$d(\frac{B_t(\tau)}{B_t}) = (\frac{B_t(\tau)}{B_t})(-\mathcal{B}^\tau(t)\sigma)dW_t$$

Consider European Call written on zero bond with the following parameters

- option's expiry date T
- bond's maturity $\tau > T$
- strike price $K > 0$

In this setting, we apply the formula for option pricing under stochastic interest rate. Here the underlying is zero bond maturing at τ. Obviously, the difference in volatilities between this underlying and zero bond maturing at options expiry date is deterministic:

$$\begin{aligned}\Sigma_t &= \underbrace{(-\sigma)\frac{1-e^{-a(\tau-t)}}{a}}_{\text{underlying's vola}} - \underbrace{(-\sigma)\frac{1-e^{-a(T-t)}}{a}}_{\text{bond's vola}} \\ &= \frac{\sigma}{a}(e^{-a(\tau-t)} - e^{-a(T-t)})\end{aligned}$$

Thus the assumption (7.10) is satisfied. The integral is given as

$$\Sigma^2 := \int_0^T |\Sigma_u|^2 du = \frac{\sigma^2}{2a^3}(e^{-a\tau} - e^{-aT})^2(e^{aT} - 1)$$

and the option price results from (7.12) as

$$\mathbb{E}^{\mathbb{Q}}(\frac{1}{B_T}(B_T(\tau) - K)^+) = B_0(\tau)\Phi(d_1) - KB_0(T)\Phi(d_2)$$

where

$$\begin{aligned}d_1 &= \frac{\ln(\frac{B_0(\tau)}{KB_0(T)}) + \frac{1}{2}\Sigma^2}{\sqrt{\Sigma^2}} \\ d_2 &= \frac{\ln(\frac{B_0(\tau)}{KB_0(T)}) - \frac{1}{2}\Sigma^2}{\sqrt{\Sigma^2}}\end{aligned}$$

Note that the function $(\theta(t))_{t\in[0,T]}$ does not enter the bond option formula!

7.7 Bond option in the Gaussian HJM-setting. In the framework of HJM-modeling, the assumption that

the forward rate volatilities $\sigma_t(\tau)$
are deterministic for all $0 \le t \le \tau \le T$

we obtain Gaussian forward rates

$$f_t(\tau) = f_0^*(\tau) + \int_0^t \alpha_s(\tau)ds + \int_0^t \sigma_s(\tau)dW_s$$

As calculated above, the bond prices follow

$$dB_t(\tau) = B_t(\tau)\left(f_t(t)dt - (\int_t^\tau \sigma_t(u)du)dW_t\right)$$

That is, the volatility of the bond $(B_t(\tau))_{t\in[0,\tau]}$ is given by

$$\text{deterministic quantity } -\textstyle\int_t^\tau \sigma_t(u)du, \qquad t\in[0,\tau]. \tag{7.13}$$

Once again, under the assumption (7.13), for deterministic forward rate volatilities, forward rates are Gaussian (here forward rates and short rate may become negative!) and bond prices are log-normally distributed. With this, the valuation of bond options can be treated as in the setting of Hull-White model.

Namely, the difference in volatilities of the bond maturing at $\tau > T$ and the bond maturing at T is given by

$$\begin{aligned}\Sigma_t &= \left(-\int_t^\tau \sigma_t(u)du\right) - \left(-\int_t^T \sigma_t(u)du\right)\\ &= -\int_T^\tau \sigma_t(u)du \qquad t\in[0,T]\end{aligned}$$

Calculating the integral

$$\Sigma^2 = \int_0^T \|\Sigma_t\|^2 dt = \int_0^T \|\int_T^\tau \sigma_t(u)du\|^2 dt$$

and

$$\begin{aligned}d_1 &= \frac{\ln(\frac{B_0(\tau)}{KB_0(T)}) + \frac{1}{2}\Sigma^2}{\sqrt{\Sigma^2}}\\ d_2 &= \frac{\ln(\frac{B_0(\tau)}{KB_0(T)}) - \frac{1}{2}\Sigma^2}{\sqrt{\Sigma^2}},\end{aligned}$$

the European call with expiry date T and strike price K written on zero bond with maturity $\tau > T$ can be priced as above, by

$$\mathbb{E}^{\mathbb{Q}}(\frac{1}{B_T}(B_T(\tau) - K)^+) = B_0(\tau)\Phi(d_1) - KB_0(T)\Phi(d_2).$$

7.8 Modeling of asset prices. Remember that the evolution of any asset price must be modeled such that arbitrage from any possible trading of this asset is impossible. Therefore, one needs to model the wealth evolution of self-financed trading strategies. This can be more complicated that it sounds, since many asset pay dividends (stocks) or require storage costs (commodities), or do not admit short positions (commodities). Furthermore, when entering positions in forwards or futures, no cash payment is required (actually, one needs to provide a margin, but in most of the cases it does not have to be cash). The change of numeraire technique can be very helpful to gain an idea how to model the evolution of an asset. Basically, we have to ensure that the wealth from self-financed trading, discounted by standard savings account follows a martingale with respect to the spot martingale measure. Of course, we are free in the choice of the numeraire (but then the measure has to be changed). First, let us consider a standard situation.

Stock. This is a generic case, the outcome will be as expected. Consider for a stock price $(S_t)_{t\in[0,T]}$ the self-financed trading strategy $\pi = (\pi_t)_{t\in[0,T]}$ where π_t stands for the number of asset units (also negative and fractional) in the portfolio at time $t \in [0, T]$. To ensure that the trading is self-financed, we suppose that $(X_t^\pi - \pi_t S_t)$ is the amount invested in standard savings account at time t. Let us agree that the wealth $(X_t^\pi)_{t\in[0,T]}$ from self-financed trading follows the dynamics

$$\begin{aligned} dX_t^\pi &= \pi_t dS_t + r_t \underbrace{(X_t^\pi - \pi_t S_t)}_{\text{savings account}} dt \\ &= r_t X_t^\pi dt + \pi_t(dS_t - r_t S_t dt) \\ &= r_t X_t^\pi dt + \pi_t B_t d(S_t B_t^{-1}) \end{aligned}$$

Namely,

$$B_t d(S_t B_t^{-1}) = B_t(B_t^{-1} dS_t - r_t S_t B_t^{-1} dt) = dS_t - r_t S_t dt.$$

Now consider the wealth, discounted by the standard savings account

$$\begin{aligned} d(X_t^\pi B_t^{-1}) &= B_t^{-1} dX_t^\pi + X_t^\pi dB_t^{-1} \\ &= B_t^{-1}(r_t X_t^\pi dt + \pi_t B_t d(S_t B_t^{-1})) + X_t^\pi(-r_t) B_t^{-1} dt \\ &= \pi_t d(S_t B_t^{-1}) \end{aligned}$$

That is, if

$$\begin{array}{c} \text{stock price, discounted as } (S_t B_t^{-1})_{t\in[0,T]} \\ \text{follows a martingale with respect} \\ \text{to the spot martingale measure } \mathbb{Q} \sim \mathbb{P} \end{array} \tag{7.14}$$

then, for appropriate strategies π the discounted wealth $(X_t^\pi B_t^{-1})_{t\in[0,T]}$ is martingale with respect to $\mathbb{Q}$. This excludes arbitrage for such strategies and shows that it is correct to model stock prices is (7.14).

Dividend-Paying stock If a stock with the price process $(S_t)_{t\in[0,T]}$ pays dividends, the situation will be different. Consider strategies $\pi = (\pi_t)_{t\in[0,T]}$ as above, suppose that there is no initial endowment and that the capital is dynamically invested in the standard savings account. Suppose that the stock continuously pays dividends at a rate, which we described by a fraction $(a_t)_{t\in[0,T]}$ of the stock price. Then the wealth $(X^\pi_t)_{t\in[0,T]}$ follows the dynamics

$$\begin{aligned} dX^\pi_t &= \pi_t dS_t + r_t \underbrace{(X^\pi_t - \pi_t S_t)}_{\text{savings account}} dt + \pi_t a_t S_t dt \\ &= r_t X^\pi_t dt + \pi_t (dS_t - (r_t - a_t) S_t dt) \\ &= r_t X^\pi_t dt + \pi_t \tilde{B}_t d(S_t \tilde{B}_t^{-1}) \end{aligned}$$

Where $(\tilde{B}_t = e^{\int_0^t (r_s - a_s) ds})_{t\in[0,T]}$, since

$$\tilde{B}_t d(S_t \tilde{B}_t^{-1}) = \tilde{B}_t (\tilde{B}_t^{-1} dS_t - (r_t - a_t) S_t \tilde{B}_t^{-1} dt) = dS_t - (r_t - a_t) S_t dt$$

Hence

$$\begin{aligned} d(X^\pi_t B_t^{-1}) &= B_t^{-1} dX^\pi_t + X^\pi_t dB_t^{-1} \\ &= B_t^{-1}(r_t X^\pi_t dt + \pi_t \tilde{B}_t d(S_t \tilde{B}_t^{-1})) + X^\pi_t (-r_t) B_t^{-1} dt \\ &= B_t^{-1} \pi_t \tilde{B}_t d(S_t \tilde{B}_t^{-1}) \end{aligned}$$

That is, if

$$\begin{gathered} \text{stock price, discounted as } (S_t \tilde{B}_t^{-1})_{t\in[0,T]} \\ \text{follows a martingale with respect} \\ \text{to the spot martingale measure } \mathbb{Q} \sim \mathbb{P} \end{gathered} \tag{7.15}$$

then, for appropriate strategies π the discounted wealth $(X^\pi_t B_t^{-1})_{t\in[0,T]}$ is martingale with respect to $\mathbb{Q}$. This excludes arbitrage for such strategies and shows that it is correct to model stock prices is in (7.15).

Futures: Consider a future maturing at time T and denote its price process by $(E_t(T))_{t\in[0,T]}$. According to the definition of a futures contract,

- No capital is needed to take position π_t at time t. The position can be fractional and negative.
- Futures trading yields continuous cash settlement. When futures price increases long party receives a payment and short party has to pay. When futures price decreases, vice versa.
- At maturity, futures price $E_T(T)$ matches the spot price S_T.

Let us agree that the wealth $(X^\pi_t)_{t\in[0,T]}$ of a futures trading strategy $\pi = (\pi_t)_{t\in[0,T]}$ satisfies

$$dX^\pi_t = r_t X^\pi_t dt + \pi_t dE_t(T).$$

This means that the cash capital is always invested in the standard savings account, and that the change of the futures price contributes to the wealth of the strategy proportionally to the position in the futures contract. After discounting by standard savings account, we obtain

$$
\begin{aligned}
d(X_t^\pi B_t^{-1}) &= B_t^{-1} dX_t^\pi + X_t^\pi dB_t^{-1} \\
&= B_t^{-1} r_t X_t^\pi dt + B_t^{-1} \pi_t dE_t(T) + X_t^\pi(-r_t) B_t^{-1} dt \\
&= B_t^{-1} \pi_t dE_t(T)
\end{aligned}
$$

That is, if

undiscounted futures price $(E_t(T))_{t\in[0,T]}$
follows a martingale with respect
to the spot martingale measure $\mathbb{Q} \sim \mathbb{P}$

then, for appropriate strategies π the discounted wealth $(X_t^\pi B_t^{-1})_{t\in[0,T]}$ is a martingale with respect to $\mathbb{Q}$. This excludes arbitrage for such strategies and shows that it is correct model futures prices as

$$
E_t(T) = \mathbb{E}^{\mathbb{Q}}(S_T \mid \mathcal{F}_t), \qquad t \in [0,T].
$$

Forward price. Consider a forward contract maturing at time T and denote its price evolution by $(F_t(T))_{t\in[0,T]}$. According to the definition of the forward contract,

- No capital needed to take position. Position can be fractional and negative.
- Having entered a position at time t, the seller (buyer) is obliged to sell (to buy) the underlying good at time T at the previously agreed forward price $F_t(T)$.
- At maturity, the forward price $F_T(T)$ equals to the spot price S_T.

Let us agree that the terminal wealth X_T^π of the forward trading strategy $\pi = (\pi_t)_{t\in[0,T]}$ behaves as follows

$$
\begin{aligned}
(\pi_s = 1_{]t,T]}(s))_{s\in[0,T]} &\implies X_T^\pi = S_T - F_t(T) \\
(\pi_s = 1_{]t,t+\Delta]}(s))_{s\in[0,T]} &\implies X_T^\pi = (S_T - F_t(T)) - (S_T - F_{t+\Delta}(T)) \\
&\qquad\quad = F_{t+\Delta}(T) - F_t(T)
\end{aligned}
$$

That is, for appropriate strategies, we have

$$
X_T^\pi = \int_0^T \pi_s dF_s(T)
$$

To determine the wealth of the strategy at an intermediate time $t \in [0,T]$, we consider the payoff $\int_0^t \pi_s dF_s(T)$ gained from trading $(\pi_s)_{s\in[0,t]}$ until time t and transfer it from the future date T to the recent t using the time value of money. Thus

$$
X_t^\pi = B_t(T) \int_0^t \pi_s dF_s(T)
$$

where $(B_t(T))_{t\in[0,T]}$ stands for the zero bond maturing at time T. With this, we obtain

$$d(X_t^\pi B_t(T)^{-1}) = \pi_t dF_t(T)$$

That is, if

undiscounted forward price $(F_t(T))_{t\in[0,T]}$
follows a martingale with respect
to the forward martingale measure $\mathbb{Q}^T \sim \mathbb{P}$

then, for appropriate strategies π the process $(X_t^\pi B_t(T)^{-1})_{t\in[0,T]}$ follows martingale with respect to $\mathbb{Q}^T$ and so $(X_t^\pi B_t^{-1})_{t\in[0,T]}$ follows martingale with respect to the spot martingale measure $\mathbb{Q}$, which excludes arbitrage for trading forward contract.

Forward-Future spread The difference between forward and futures price is explained by interest rate effects.

$$\begin{aligned} F_0(T) - E_0(T) &= \mathbb{E}^{\mathbb{Q}^T}(S_T) - \mathbb{E}^{\mathbb{Q}}(S_T) = \mathbb{E}^{\mathbb{Q}}(\frac{d\mathbb{Q}^T}{d\mathbb{Q}}S_T) - \mathbb{E}^{\mathbb{Q}}(\frac{d\mathbb{Q}^T}{d\mathbb{Q}})\mathbb{E}^{\mathbb{Q}}(S_T) \\ &= \mathrm{Cov}^Q(\frac{B_T(T)}{B_T}\frac{B_0}{B_0(T)}, S_T) = \frac{\mathrm{Cov}^Q(B_T^{-1}, S_T)}{B_0(T)} \end{aligned}$$

It turns out that if the dynamics of short rate is independent on spot price of the underlying, then forward and futures price must coincide. However, the interpretation of what it means to be independent with respect to the spot martingale measure is not obvious. At least, we know now that for deterministic interest rates, the forward and the futures price must agree.

7.9 Using change of numeraire in the commodity price modeling. Commodities differ from typical financial assets in the following aspects

- storage of physical commodities is difficult and costly
- long positions require cost-of-carry
- short positions (in the physical commodity) are impossible

For this, a significant trading volume in commodity business is done on the futures/forward basis. In practice, most of the positions are closed before it comes to delivery. (Obviously agents have their own supply agreements but need to hedge themselves against price movements).

The so-called spot-based commodity price models attempt to describe the evolution of commodity spot prices $(S_t)_{t\in[0,T]}$ by a diffusion process. Given futures maturity dates $\tau \in [0, T]$, one models the commodity future price processes $(E_t(\tau))_{t\in[0,\tau]}$, $\tau \in [0, T]$ in line with the following insights:

- At maturity, futures price matches the spot price

$$E_\tau(\tau) = S_\tau, \qquad \tau \in [0,T].$$

- There is no arbitrage for futures trading. That is, given the spot price dynamics $(S_t)_{t\in[0,T]}$ under the spot martingale measure $\mathbb{Q}$, the futures prices are given by

$$E_t(\tau) = \mathbb{E}^{\mathbb{Q}}(S_\tau \,|\, \mathcal{F}_t), \qquad 0 \le t \le \tau \le T.$$

Some other criteria must be considered, too. For instance

- Consistency with the initial futures curve

$$E_0(\tau) = E_0^*(\tau), \qquad \tau \in [0,T]$$

- Consistency with evolution of the futures curve. (At least, a sufficient flexibility of the futures curve)

It turns out that the expertise from the interest rate theory may be very helpful here.

Suppose that the commodity spot process is modeled by a diffusion process

$$dS_t = S_t(-\mu_t dt + \sigma_t dW_t), \qquad t \in [0,T],$$

where the process $(W_t)_{t\in[0,T]}$ is assumed to follow a Brownian motion with respect to the spot martingale measure $\mathbb{Q}$. Then, we have

$$S_\tau = S_0 e^{-\int_0^\tau \mu_s ds} \underbrace{e^{\int_0^\tau \sigma_s dW_s - \frac{1}{2}\int_0^\tau \sigma_s^2 ds}}_{D_\tau}.$$

Assume now that the process

(7.16) $$D_t = e^{\int_0^t \sigma_s dW_s - \frac{1}{2}\int_0^t \sigma_s^2 ds}, \qquad t \in [0,T]$$
follows a true martingale

This can be achieved by appropriate assumption on spot price volatility $(\sigma_t)_{t\in[0,T]}$, for instance by imposing Novikov condition

(7.17) $$\mathbb{E}(e^{\frac{1}{2}\int_0^T \sigma_s^2 ds}) < \infty.$$

If (7.16) holds, then the futures price can be calculated as follows

$$\begin{aligned}
E_t(\tau) &= \mathbb{E}^{\mathbb{Q}}(S_\tau \,|\, \mathcal{F}_t) \\
&= \mathbb{E}^{\mathbb{Q}}(S_t e^{-\int_t^\tau \mu_s ds} \underbrace{e^{\int_t^\tau \sigma_s dW_s - \frac{1}{2}\int_t^\tau \sigma_s^2 ds}}_{D_\tau/D_t} \,|\, \mathcal{F}_t) \\
&= S_t \frac{\mathbb{E}^{\mathbb{Q}}(D_\tau e^{-\int_t^\tau \mu_s ds} \,|\, \mathcal{F}_t)}{D_t} = S_t \frac{\mathbb{E}^{\mathbb{Q}}(D_T e^{-\int_t^\tau \mu_s ds} \,|\, \mathcal{F}_t)}{\mathbb{E}^{\mathbb{Q}}(D_T \,|\, \mathcal{F}_t)} \\
&= S_t \mathbb{E}^{\tilde{Q}}(e^{-\int_t^\tau \mu_s ds} \,|\, \mathcal{F}_t)
\end{aligned}$$

where the measure $\tilde{Q}$ is given by

$$d\tilde{Q} = D_T dQ = \left(e^{\int_0^T \sigma_s dW_s - \frac{1}{2}\int_0^T \sigma_s^2 ds}\right) dQ.$$

Note that the process

$$\tilde{W}_t = W_t - \int_0^t \sigma_s ds \qquad t \in [0,T]$$

is a Brownian motion under the measure $\tilde{Q}$.

Now, recognize that if $(\mu_s)_{s\in[0,T]}$ is modeled as a short rate of an interest rate model with respect to $\tilde{Q}$ as spot martingale measure, then

$$\tilde{B}_t(\tau) = \mathbb{E}^{\tilde{Q}}(e^{-\int_t^\tau \mu_s ds} \,|\, \mathcal{F}_t) \qquad 0 \le t \le \tau \le T$$

would be interpretable of as zero bond prices of a 'virtual' money market. Hence futures price can be written as spot price multiplied by such a virtual bond price

$$E_t(\tau) = S_t \tilde{B}_t(\tau) \qquad 0 \le t \le \tau \le T.$$

Using this expression, the consistency with initial futures curve

$$E_0(\tau) = E_0^*(\tau) \qquad \tau \in [0,T]$$

is fulfilled if and only if the bond market

$$(\tilde{B}_t(\tau))_{t\in[0,\tau]} \;\; \tau \in [0,T] \tag{7.18}$$

is modeled such that a perfect fit to the initial bond curve is possible

$$\tilde{B}_0(\tau) = E_0^*(\tau)/S_0^* \qquad \tau \in [0,T]$$

Furthermore, the flexibility in changes of the futures curve reduces to the flexibility of the bond curve of the virtual bond market (7.18).

Following this approach, following construction of a commodity market is possible:

1. Chose the spot martingale measure $\mathbb{Q}$, a Brownian motion $(W_t)_{t\in[0,T]}$ under Q and a spot price volatility process $(\sigma_t)_{t\in[0,T]}$ such that

   $$D_t = e^{\int_0^t \sigma_s dW_s - \frac{1}{2}\int_0^t \sigma_s^2 ds} \qquad t \in [0,T]$$

   is a true martingale.

2. Define the equivalent measure $\tilde{Q}$ by $d\tilde{Q} = D_T dQ$.

3. Construct a short rate process $(\mu_t)_{t\in[0,T]}$ imagining that $\tilde{Q}$ is the spot martingale measure and define

$$\tilde{B}_t(\tau) = \mathbb{E}^{\tilde{Q}}(e^{-\int_t^\tau \mu_s ds} \mid \mathcal{F}_t) \qquad 0 \le t \le \tau \le T$$

Such a virtual bond market could be obtained using a standard short rate or a forward rate model, where the process

$$\tilde{W}_t = W_t - \int_0^t \sigma_s ds \qquad t \in [0, T].$$

is used instead of the generic driving Brownian motion underlying the standard model.

4. Define the futures price evolution as

$$E_t(\tau) = S_t \tilde{B}_t(\tau) \qquad 0 \le t \le \tau \le T.$$

Having constructed virtual bond prices by one of the standard models, we obtain

$$d\tilde{B}_t(\tau) = \tilde{B}_t(\tau)\left(\mu_t dt - \tilde{\sigma}_t(\tau) d\tilde{W}_t\right)$$

where the bond volatilities

$$-\tilde{\sigma}_t(\tau) \qquad 0 \le t \le \tau \le T$$

are determined by particular choice of interest rate model with the short rate $(\mu_t)_{t\in[0,T]}$ and spot martingale measure $\tilde{Q}$. Let us determine the volatility structure of commodity futures constructed in this way

$$\begin{aligned}
dE_t(\tau) &= d(S_t \tilde{B}_t(\tau)) \\
&= S_t d\tilde{B}_t(\tau) + \tilde{B}_t(\tau) dS_t + d\langle \tilde{B}_.(\tau), S_. \rangle_t \\
&= S_t(\tilde{B}_t \mu_t dt - \tilde{B}_t(\tau)\tilde{\sigma}_t(\tau) d\tilde{W}_t) + \tilde{B}_t(\tau)(S_t(-\mu_t)dt + S_t \sigma_t dW_t) - S_t \sigma_t \tilde{B}_t(\tau) \tilde{\sigma}_t(\tau) dt \\
&= S_t(-\tilde{B}_t(\tau))\tilde{\sigma}_t(\tau)(dW_t - \sigma_t dt) + \tilde{B}_t(\tau) S_t \sigma_t dW_t - S_t \sigma_t \tilde{B}_t(\tau)\tilde{\sigma}_t(\tau) dt \\
&= \underbrace{S_t \tilde{B}_t(\tau)}_{E_t(\tau)} \underbrace{(\sigma_t - \tilde{\sigma}_t(\tau))}_{\Sigma_t(\tau)} dW_t
\end{aligned}$$

Now observe that if

$$\begin{array}{c}\text{the spot price volatility } (\sigma_t)_{t\in[0,T]} \\ \text{and bond volatilities } (-\tilde{\sigma}_t(\tau))_{t\in[0,\tau]} \\ \text{are deterministic for all } \tau \in [0, T]\end{array} \tag{7.19}$$

then the futures follow

$$dE_t(\tau) = E_t(\tau) \underbrace{\Sigma_t(\tau)}_{\text{deterministic}} dW_t$$

log-normally distributed martingales

$$E_t(\tau) = E_0(\tau)e^{\int_0^t \Sigma_u(\tau)dW_u - \frac{1}{2}\int_0^t \|\Sigma_u(\tau)\|^2 du} \qquad 0 \leq t \leq \tau \leq T.$$

In this situation, the valuation of European Calls on commodity futures is a straight-forward procedure. For simplicity, suppose that the interest rate $r \in]0, \infty[$ is constant and deterministic. Consider a European Call with strike price K with maturity τ written on commodity futures price with expiry $\tau' > \tau$. At time t the fair price of this call is obtained as follows

$$\begin{aligned}
&e^{-r(\tau-t)}\mathbb{E}^{\mathbb{Q}}((E_\tau(\tau') - K)^+ \,|\, \mathcal{F}_t) = \\
&= \; e^{-r(\tau-t)}\mathbb{E}^{\mathbb{Q}}((E_t(\tau')e^{\int_t^\tau \Sigma_u(\tau')dW_u - \frac{1}{2}\int_t^\tau \|\Sigma_u(\tau')\|^2 du} - K)^+ \,|\, \mathcal{F}_t) \\
&= \; e^{-r(\tau-t)}\left(E_t(\tau')\Phi(d_1) - K\Phi(d_2)\right)
\end{aligned}$$

where

$$\begin{aligned}
d_1 &= \frac{\ln(\frac{E_t(\tau')}{K}) + \frac{1}{2}\Sigma^2}{\sqrt{\Sigma^2}} \\
d_2 &= \frac{\ln(\frac{E_t(\tau')}{K}) - \frac{1}{2}\Sigma^2}{\sqrt{\Sigma^2}} \\
\Sigma^2 &= \int_t^\tau \|\Sigma_u(\tau')\|^2 du
\end{aligned}$$

Furthermore, a calendar spread can be priced explicitly. A calendar spread pays the difference of futures prices with different maturities. More precisely let us consider a European option with payoff

$$(E_\tau(\tau') - KE_\tau(\tau''))^+ \text{ at maturity } \tau < \tau' < \tau'', \; K \geq 0$$

Define a measure Q^E by

$$dQ^E = \frac{E_{\tau''}(\tau'')}{E_0(\tau'')}dQ,$$

the fair price of this spread is calculated at time $t \leq \tau$ as

$$\begin{aligned}
&e^{-r(\tau-t)}\mathbb{E}^{\mathbb{Q}}((E_\tau(\tau') - KE_\tau(\tau''))^+ \,|\, \mathcal{F}_t) = \\
&= \; e^{-r(\tau-t)}\mathbb{E}^{\mathbb{Q}}(E_\tau(\tau'')(\frac{E_\tau(\tau')}{E_\tau(\tau'')} - K)^+ \,|\, \mathcal{F}_t) \\
&= \; e^{-r(\tau-t)}\mathbb{E}^{\mathbb{Q}}(\frac{E_\tau(\tau'')}{E_0(\tau'')}(\frac{E_\tau(\tau')}{E_\tau(\tau'')} - K)^+ \,|\, \mathcal{F}_t)E_0(\tau'') \\
&= \; e^{-r(\tau-t)}\mathbb{E}^{Q^E}((\frac{E_\tau(\tau')}{E_\tau(\tau'')} - K)^+ \,|\, \mathcal{F}_t)\underbrace{\mathbb{E}(\frac{E_\tau(\tau'')}{E_0(\tau'')} \,|\, \mathcal{F}_t)E_0(\tau'')}_{E_t(\tau'')} \\
&= \; e^{-r(\tau-t)}E_t(\tau'') \times \\
&\quad \times\mathbb{E}^{Q^E}((\frac{E_t(\tau')}{E_t(\tau'')}e^{\int_t^\tau(\tilde{\sigma}_u(\tau'')-\tilde{\sigma}_u(\tau'))dW_u^E - \frac{1}{2}\int_t^\tau \|\tilde{\sigma}_u(\tau'')-\tilde{\sigma}_u(\tau')\|^2 du} - K)^+ \,|\, \mathcal{F}_t) \\
&= \; e^{-r(\tau-t)}E_t(\tau'')\left(\frac{E_t(\tau')}{E_t(\tau'')}\Phi(d_1) - K\Phi(d_2)\right)
\end{aligned}$$

where

$$
\begin{aligned}
d_1 &= \frac{\ln(\frac{E_t(\tau')}{E_t(\tau'')K}) + \frac{1}{2}\Sigma^2}{\sqrt{\Sigma^2}} \\
d_2 &= \frac{\ln(\frac{E_t(\tau')}{E_t(\tau'')K}) - \frac{1}{2}\Sigma^2}{\sqrt{\Sigma^2}} \\
\Sigma^2 &= \int_t^\tau \|\tilde{\sigma}_u(\tau'') - \tilde{\sigma}_u(\tau')\|^2 du
\end{aligned}
$$

8 LIBOR-based market models (LMM)

8.1 The philosophy of LMM. The idea of this approach is to model the dynamics of LIBOR rates directly, avoiding the construction of zero bonds prices.

The LIBOR market modeling assums that there exists a finite number of bond's maturity dates

$$0 < \tau_1 < \tau_2 < \ldots < \tau_{n+1} = T$$

The goal of LIBOR market modeling is to introduce the evolution of LIBOR rates

$$(L_t(\tau_i, \tau_{i+1}))_{t \in [0,\tau_i]}, \qquad i = 1, \ldots, n$$

which behave such as they were derived by

$$L_t(\tau_i, \tau_{i+1}) = \frac{B_t(\tau_i) - B_t(\tau_{i+1})}{(\tau_{i+1} - \tau_i) B_t(\tau_{i+1})} = \frac{1}{\tau_{i+1} - \tau_i} \left(\frac{B_t(\tau_i)}{B_t(\tau_{i+1})} - 1 \right)$$

from an arbitrage-free model for zero bonds

$$(B_t(\tau_i))_{t \in [0,\tau_i]}, \qquad i = 1, \ldots, n.$$

In particular, such that

- The resulting dynamics for tradable LIBOR-related assets is free of arbitrage
- The relevant financial instruments can be efficiently priced
- The model can be calibrated to the market data, which means that appropriate model parameters can be chosen in such way that the observed Caplet prices (or, more importantly, the observed Swaption prices) are explained.

8.2 Black's caplet formula. Consider a Caplet, which provides

$$\text{a payoff of } (L_{\tau_i}(\tau_i, \tau_{i+1}) - K)^+ \text{ at the time } \tau_{i+1}.$$

Basically, a caplet can be considered as a European Call with strike price K, written on the terminal payoff

$$L_{\tau_i}(\tau_i, \tau_{i+1}) = \frac{1}{\tau_{i+1} - \tau_i}\left(\frac{1}{B_{\tau_i}(\tau_{i+1})} - 1\right) \quad \text{at } \tau_{i+1}$$

of the single LIBOR $(L_t(\tau_i, \tau_{i+1}))_{t\in[0,\tau_i]}$ at time τ_{i+1}. A caplet is an important instrument. Within interest rate models discussed so far, a valuation of caplet would result in the calculation of

$$B_t\mathbb{E}^{\mathbb{Q}}(\frac{(L_{\tau_i}(\tau_i, \tau_{i+1}) - K)^+}{B_{\tau_{i+1}}} \mid \mathcal{F}_t)$$

which could give a complicated formula. However, practitioners have used Black's 76 formula to calculate the fair price of a caplet. In certain sense, this practice is based on to the assumption that $(L_t(\tau_i, \tau_{i+1}))_{t\in[0,T]})_{t\in[0,T]}$ follows a geometric Brownian motion. We shall try to find conditions, which justify such a valuation.

Let us switch to the forward measure $\mathbb{Q}^{\tau_{i+1}}$ in the caplet valuation

$$\begin{aligned}
B_t\mathbb{E}^{\mathbb{Q}}_t(\frac{(L_{\tau_i}(\tau_i, \tau_{i+1}) - K)^+}{B_{\tau_{i+1}}}) &= B_t(\tau_{i+1})\mathbb{E}^{\mathbb{Q}^{\tau_{i+1}}}_t(\frac{(L_{\tau_i}(\tau_i, \tau_{i+1}) - K)^+}{\underbrace{B_{\tau_{i+1}}(\tau_{i+1})}_{=1}}) \\
&= B_t(\tau_{i+1})\mathbb{E}^{\mathbb{Q}^{\tau_{i+1}}}_t((L_{\tau_i}(\tau_i, \tau_{i+1}) - K)^+)
\end{aligned}$$

With this in mind, look at the evolution of the LIBOR

$$L_t(\tau_i, \tau_{i+1}) = \frac{B_t(\tau_i) - B_t(\tau_{i+1})}{(\tau_{i+1} - \tau_i)B_t(\tau_{i+1})} = \frac{1}{\tau_{i+1} - \tau_i}\left(\frac{B_t(\tau_i)}{B_t(\tau_{i+1})} - 1\right)$$

and realize that within a generic bond market model

$$\text{the process } (L_t(\tau_i, \tau_{i+1}))_{t\in[0,\tau_i]} \text{ follows a positive-valued martingale under the forward measure } \mathbb{Q}^{\tau_{i+1}}. \tag{8.1}$$

Under appropriate conditions (martingale representation theorem must hold), the insight (8.1) is expressed as

$$dL_t(\tau_i, \tau_{i+1}) = L_t(\tau_i, \tau_{i+1})\gamma_t(\tau_i)dW_t(\tau_{i+1}), \tag{8.2}$$

where

$$(W_t(\tau_{i+1}))_{t\in[0,\tau_{i+1}]} \text{ is a Brownian motion under } \mathbb{Q}^{\tau_{i+1}}. \tag{8.3}$$

and $(\gamma_t(\tau_i))_{t\in[0,\tau_i]}$ is the stochastic logarithm of $(L_t(\tau_i, \tau_{i+1}))_{t\in[0,\tau_i]}$, as in (8.2). In principle, it may happen that

$$(\gamma_t(\tau_i))_{t\in[0,\tau_i]} \text{ is a deterministic process,} \tag{8.4}$$

In such case, the Black's Caplet formula holds exactly.

$$
\begin{aligned}
B_0 \mathbb{E}^{\mathbb{Q}}\left(\frac{(L_{\tau_i}(\tau_i, \tau_{i+1}) - K)^+}{B_{\tau_{i+1}}}\right) &= B_0(\tau_{i+1})\left(L_0(\tau_i, \tau_{i+1})\Phi(d_1) - K\Phi(d_2)\right) \\
d_1 &= \frac{\ln(\frac{L_0(\tau_i, \tau_{i+1})}{K}) + \frac{1}{2}\Sigma^2}{\sqrt{\Sigma^2}} \\
d_2 &= \frac{\ln(\frac{L_0(\tau_i, \tau_{i+1})}{K}) - \frac{1}{2}\Sigma^2}{\sqrt{\Sigma^2}} \\
\Sigma^2 &= \int_0^{\tau_i} \|\gamma_u(\tau_i)\|^2 du
\end{aligned}
$$

Question: Can we construct interest rate models, where Black's Caplet formula holds exactly?

Answer: Yes, surprisingly. They are addressed in the BGM-framework (named after Brace, Gaterek, Musiela). Their idea is to specify exogenously

- d-dimensional deterministic volatilities $(\gamma_t(\tau_i))_{t\in[0,\tau_i]})$, $i = 1, \ldots, n$
- d-dimensional Brownian motion $(W_t(\tau_{n+1}))_{t\in[0,\tau_{n+1}]}$
- construct all LIBOR rate dynamics $(L_t(\tau_i, \tau_{i+1}))_{t\in[0,\tau_i]}$, $i = 1, \ldots, n$ such they behave as they were derived from an arbitrage-free bond market. Thereby, the bond dynamics is not specified and is not determined by model ingredients. But it exists in the sense that by specification of additional ingredients it can be determined.

8.3 The idea of BGM. Consider a bond market model

$$(B_t(\tau_i))_{t\in[0,\tau_i]}, \qquad i = 1, \ldots, n+1 \tag{8.5}$$

which is generic in the sense that the bond prices follow diffusion processes

$$dB_t(\tau_i) = B_t(\tau_i)(r_t dt - \sigma_t(\tau_i))dW_t, \qquad i = 1, \ldots, n+1 \tag{8.6}$$

driven by

(8.7)
the process $(W_t)_{t\in[0,T]}$
of Brownian motion with respect
to the spot martingale measure $\mathbb{Q}$

In what follows, we show that the resulting dynamics of LIBORs

$$dL_t(\tau_i, \tau_{i+1}) = L_t(\tau_i, \tau_{i+1})\gamma_t(\tau_i)dW_t(\tau_{i+1}), \qquad i = 1, \ldots, n \tag{8.8}$$

can be algebraically written in terms of

- the Brownian motion $(W_t(\tau_{n+1}))_{t\in[0,\tau_{n+1}]}$ with respect to $\mathbb{Q}^{\tau_{n+1}}$
- volatilities $(\gamma_t(\tau_i))_{t\in[0,\tau_i]}$, $i = 1, \dots, n$.

This is not trivial, since in (8.8) the Brownain motion $(W_t(\tau_{i+1}))_{t\in[0,\tau_{i+1}]}$ must be obtainable from both ingredients $(W_t(\tau_{n+1}))_{t\in[0,\tau_{n+1}]}$ and $(\gamma_t(\tau_i))_{t\in[0,\tau_i]}$, $i = 1, \dots, n$. Let us do some calculations.

Under the assumptions (8.5) — (8.7) of a generic bond market, let us consider the forward measures

$$d\mathbb{Q}^{\tau_{i+1}} = \left(\frac{B_{\tau_{i+1}}(\tau_{i+1})}{B_{\tau_{i+1}}}\frac{B_0}{B_0(\tau_{i+1})}\right) dQ, \qquad i = 1, \dots, n$$

According to the Girsanov theorem,

$$\begin{aligned} W_t(\tau_{i+1}) &= W_t - \int_0^t (-\sigma_u(\tau_{i+1}))du \\ &= W_t + \int_0^t \sigma_u(\tau_{i+1})du, \qquad t \in [0, \tau_{i+1}] \end{aligned}$$

follows a process of Brownian motion with respect to $\mathbb{Q}^{\tau_{i+1}}$. With this, we can reach each Brownian motion $(W_t(\tau_j))_{t\in[0,\tau_j]}$ by

$$\begin{aligned} W_t(\tau_j) &= \underbrace{W_t}_{W_t(\tau_{n+1})-\int_0^t \sigma_u(\tau_{n+1})du} + \int_0^t \sigma_u(\tau_j)du \\ &= W_t(\tau_{n+1}) + \int_0^t (\sigma_u(\tau_j) - \sigma_u(\tau_{n+1}))du \\ &= W_t(\tau_{n+1}) + \sum_{i=j}^{n} \int_0^t (\sigma_u(\tau_i) - \sigma_u(\tau_{i+1}))du \end{aligned} \tag{8.9}$$

Using differences in volatilities of subsequently maturing bonds is needed, since the the volatility of the LIBOR contains terms of this type, as we see now.

Using the representation

$$(\tau_{i+1} - \tau_i)L_t(\tau_i, \tau_{i+1}) = B_t(\tau_i)/B_t(\tau_{i+1}) - 1 \qquad \text{for all } t \in [0, \tau_i]$$

we obtain

$$\begin{aligned} (\tau_{i+1} - \tau_i)dL_t(\tau_i, \tau_{i+1}) &= d(\frac{B_t(\tau_i)}{B_t(\tau_{i+1})} - 1) = d(\frac{B_t(\tau_i)}{B_t(\tau_{i+1})}) \\ &= (\frac{B_t(\tau_i)}{B_t(\tau_{i+1})})(\sigma_t(\tau_{i+1}) - \sigma_t(\tau_i))dW_t(\tau_{i+1}) \\ &= ((\tau_{i+1} - \tau_i)L_t(\tau_i, \tau_{i+1}) + 1)(\sigma_t(\tau_{i+1}) - \sigma_t(\tau_i))dW_t(\tau_{i+1}). \end{aligned}$$

Thus, we express the LIBOR volatility in terms of the bond volatilities for $i = 1, \dots, n$

$$dL_t(\tau_i, \tau_{i+1}) = L_t(\tau_i, \tau_{i+1}) \underbrace{\frac{1 + (\tau_{i+1} - \tau_i)L_t(\tau_i, \tau_{i+1})}{(\tau_{i+1} - \tau_i)L_t(\tau_i, \tau_{i+1})}(\sigma_t(\tau_{i+1}) - \sigma_t(\tau_i))}_{\gamma_t(\tau_i)} dW_t(\tau_{i+1}).$$

Fix this result as

$$\gamma_t(\tau_i) = \frac{1 + (\tau_{i+1} - \tau_i)L_t(\tau_i, \tau_{i+1})}{(\tau_{i+1} - \tau_i)L_t(\tau_i, \tau_{i+1})}(\sigma_t(\tau_{i+1}) - \sigma_t(\tau_i)), \quad t \in [0, \tau_i]$$

and rewrite as

$$\sigma_t(\tau_{i+1}) - \sigma_t(\tau_i) = \frac{\gamma_t(\tau_i)(\tau_{i+1} - \tau_i)L_t(\tau_i, \tau_{i+1})}{1 + (\tau_{i+1} - \tau_i)L_t(\tau_i, \tau_{i+1})}$$

to conclude that (8.9) can be expressed as

$$W_t(\tau_j) = W_t(\tau_{n+1}) - \sum_{i=j}^{n} \int_0^t \frac{\gamma_u(\tau_i)(\tau_{i+1} - \tau_i)L_u(\tau_i, \tau_{i+1})}{1 + (\tau_{i+1} - \tau_i)L_u(\tau_i, \tau_{i+1})} du \quad t \in [0, \tau_i]. \tag{8.10}$$

Finally, deduce the dynamics

$$\begin{aligned} dL_t(\tau_j, \tau_{j+1}) &= \gamma_t(\tau_j)L_t(\tau_j, \tau_{j+1})dW_t(\tau_{j+1}) \\ &= \gamma_t(\tau_j)L_t(\tau_j, \tau_{j+1}) \left(dW_t(\tau_{n+1}) - \sum_{i=j+1}^{n} \frac{\gamma_t(\tau_i)(\tau_{i+1} - \tau_i)L_t(\tau_i, \tau_{i+1})}{1 + (\tau_{i+1} - \tau_i)L_t(\tau_i, \tau_{i+1})} dt \right) \end{aligned}$$

Note that due to this recursion, one can work backward:

i) Specify $(W_t(\tau_{n+1}))_{t\in[0,\tau_{n+1}]}$ and $(\gamma_t(\tau_n))_{t\in[0,\tau_n]}$ to calculate $(L_t(\tau_n, \tau_{n+1}))_{t\in[0,\tau_n]}$.

ii) Given $(\gamma_t(\tau_i))_{t\in[0,\tau_i]}$ and $(L_t(\tau_i, \tau_{i+1}))_{t\in[0,\tau_i]}$ for all $i = j+1, \dots, n$, calculate $(L_t(\tau_j, \tau_{j+1}))_{t\in[0,\tau_j]}$.

Note that to to carry out this procedure, one needs only

- Brownian motion $(W_t(\tau_{n+1}))_{t\in[0,\tau_{n+1}]}$
- LIBOR volatilities $(\gamma_t(\tau_j))_{t\in[0,\tau_j]}$, $j = 1, \dots, n$.

In this context, an important question is

(8.11) Given $(L_t(\tau_i, \tau_{i+1}))_{t\in[0,\tau_i]}$ for $i = 1, \dots, n$ constructed from $(W_t(\tau_{n+1}))_{t\in[0,\tau_{n+1}]}$ and $(\gamma_t(\tau_j))_{t\in[0,\tau_j]}$, $j = 1, \dots, n$, is there a bond market model (8.5)-(8.7) with such LIBORs?

The answer is not obvious. Note that the LIBOR dynamics determines the bond prices at maturity dates $\tau_1, \dots, \tau_n$ only.

$$\begin{aligned} B_{\tau_i}(\tau_j) &= \frac{\cancel{B_{\tau_i}(\tau_{i+1})}}{\underbrace{B_{\tau_i}(\tau_i)}_{=1}} \cdots \frac{\cancel{B_{\tau_i}(\tau_{j-1})}}{\cancel{B_{\tau_i}(\tau_{j-2})}} \frac{B_{\tau_i}(\tau_j)}{\cancel{B_{\tau_i}(\tau_{j-1})}} \\ &= (1 + (\tau_{i+1} - \tau_i)L_{\tau_i}(\tau_i, \tau_{i+1}))^{-1} \dots (1 + (\tau_j - \tau_{j-1})L_{\tau_i}(\tau_{j-1}, \tau_j))^{-1} \end{aligned}$$

For the intermediate dates $t \in [0, T] \setminus \{\tau_1, \dots, \tau_{n+1}\}$, the bond prices are not determined.

However, following additional specifications can determine all bond prices

- Specify explicitly the bond volatilities in front of maturity

$$(-\sigma_t(\tau_{i+1}))_{t\in[\tau_i,\tau_{i+1}]}, \qquad i=1,\dots,n$$

- Specify short rate dynamics $(r_s)_{s\in[t_i,\tau_{i+1}]}$ which fits the requirements

$$B_{\tau_{i+1}}(\tau_{i+1}) = 1 = B_{\tau_i}(\tau_{i+1})e^{\int_{\tau_i}^{\tau_{i+1}} r_s ds - \int_{\tau_i}^{\tau_{i+1}} \sigma_s(\tau_{i+1})dW_s - \frac{1}{2}\int_{\tau_i}^{\tau_{i+1}} \|\sigma_s(\tau_{i+1})\|^2 ds}$$

  this is equivalent to

$$e^{\int_{\tau_i}^{\tau_{i+1}} r_s ds} = B_{\tau_i}^{-1}(\tau_{i+1})e^{\int_{\tau_i}^{\tau_{i+1}} \sigma_s(\tau_{i+1})dW_s - \frac{1}{2}\int_{\tau_i}^{\tau_{i+1}} |\sigma_s(\tau_{i+1})|^2 ds}$$

  How to determine the process $(r_s)_{s\in[\tau_i,\tau_{i+1}]}$ from this?

Problem: Given an $\mathcal{F}_\tau$-measurable random variable X_τ. Is there an adapted process $(x_s)_{s\in[0,\tau]}$ such that $X_\tau = \int_0^\tau x_s ds$?

8.4 Calibration of BGM models. Note that BGM approach yields bond prices at bond maturities only. On this account, we can address consistency with initial data and consistency with time evolution on bond maturities.
Given initial bond curve $(B_0^*(\tau_i))_{i=1}^{n+1}$, the BGM model automatically fits to this data, starting the evolution of the LIBORs at values given by initial bond prices

$$L_0(\tau_i,\tau_{i+1}) = \frac{B_0^*(\tau_i) - B_0^*(\tau_{i+1})}{(\tau_{i+1}-\tau_i)B_0^*(\tau_{i+1})} \qquad i=1,\dots,n$$

At time τ_j the LIBORs have evolved. They fit to the bond curve $(B_{\tau_j}^*(\tau_i))_{i=j}^{n+1}$ if LIBOR rates follow an evolution which yields

$$L_{\tau_j}(\tau_i,\tau_{i+1}) = \frac{B_{\tau_j}^*(\tau_i) - B_{\tau_j}^*(\tau_{i+1})}{(\tau_{i+1}-\tau_i)B_0^*(\tau_{i+1})} \qquad i=j,\dots,n.$$

If we suppose that the interest rate is always positive, then there will be a positive number on the righ-hand side of this equation. Beyond this, we can not assume anything else. That is, BGM model is consistent with time evolution if the joint distribution

$$(L_{\tau_i}(\tau_j,\tau_{j+1}))_{j=i}^n$$

is equivalent to the Lebesgue measure on $]0,\infty[^{n-j}$, considering the evolution

$$dL_t(\tau_j,\tau_{j+1}) = L_t(\tau_j,\tau_{j+1})\gamma_t(\tau_j)dW_t(\tau_{j+1}), \quad j=i,\dots,nj$$

we see that this is the case, for instance, if

$$(\gamma_t(\tau_j))_{t\in[0,\tau_i]}, \;\; j=i,\dots,n \quad \text{are deterministic and linearly independent}$$

Hence, by appropriate choice of LIBOR volatilities, the time consistency of BGM model can be achieved.

However, in the practice, the models are required not only to explain prices of underlying assets but also to fit the prices of derivatives. In particular, interest rate models are supposed to explain the listed swaption prices. Note that even within BGM framework swaptions do not have closed-form expressions. However, the market practice is to apply Black-Scholes formula to swap rates (at-the-money interest rates) and to compare the result to the listed swaption price in order to infer the implied volatility. Thereafter, the parameters of the BGM model are adapted, such that the volatility of the swap rate comes close to the implied Black-Scholes volatility. Let us consider a simplified procedure illustrating this philosophy.

Remember the IRS contract (interest rate swap) with price at $t \in [0, \tau_m]$ prior to the first settlement date τ_m of the swap.

$$\mathrm{IRS}_t(K) = \sum_{i=m+1}^{n} B_t(\tau_i)(L_t(\tau_{i-1}, \tau_i) - K)(\tau_i - \tau_{i-1})$$

Remember that the swap rate S_t (at the money interest rate) is defined at time t as the strike K at which the contract value is zero

$$S_t = \frac{B_t(\tau_m) - B_t(\tau_n)}{\sum_{i=m+1}^{n} B_t(\tau_i)(\tau_i - \tau_{i-1})}, \qquad t \in [0, \tau_m]$$

In this context, the process

$$P_t = \sum_{i=m+1}^{n} B_t(\tau_i)(\tau_i - \tau_{i-1}) \qquad t \in [0, \tau_m]$$

is called the *annuity numearaire*. The reason for this name is that P_t represents the time t value of the contract, which pays $(\tau_i - \tau_{i-1})$ currency units at τ_i and is similar to an annuity contract. A simple algebraic manipulation shows that

$$\mathrm{IRS}_t(K) = P_t(S_t - K) \qquad t \in [0, \tau_m].$$

Namely

$$\begin{aligned}
P_t(S_t - K) &= S_t P_t - K P_t \\
&= \underbrace{B_t(\tau_m) - B_t(\tau_n)}_{=\sum_{i=m+1}^{n} B_t(\tau_i)L_t(\tau_{i-1},\tau_i)(\tau_i-\tau_{i-1})} - K \sum_{i=m+1}^{n} B_t(\tau_i)(\tau_i - \tau_{i-1}) \\
&= \sum_{i=m+1}^{n} B_t(\tau_i)(L_t(\tau_{i-1}, \tau_i) - K)(\tau_i - \tau_{i-1}).
\end{aligned}$$

Hence the payoff of the Swaption

$$(\mathrm{IRS}_{\tau_m}(K))^+ \qquad \text{at maturity } \tau_m \text{ of the Swaption}$$

can be equivalently written as

$$(\mathrm{IRS}_{\tau_m}(K))^+ = P_{\tau_m}(S_{\tau_m} - K)^+ \qquad \text{at maturity } \tau_m \text{ of the Swaption}$$

and looks similarly to the payoff of an ordinary European Call written on swap rate $(S_t)_{t\in[0,\tau_m]}$ with strike price K and maturity τ_m. That is, the fair price of the Swaption at time $t = 0$ is given by

$$\mathbb{E}^{\mathbb{Q}}(\frac{P_{\tau_m}}{B_{\tau_m}}(S_{\tau_m} - K)^+)$$

There is no closed-form expression for this. Of course, given any bond market model model, such contract can be priced by Monte-Carlo simulation. However, to be able calibrating the model to Swaption prices, one needs a closed-form approximation.

To derive such approximation, we switch from the spot martingale measure $\mathbb{Q}$ to the so-called *annuity measure* Q^P via change-of-numeraire

$$dQ^P = \frac{P_{\tau_m}}{B_{\tau_m}}\frac{B_0}{P_0}dQ \tag{8.12}$$

to obtain

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}(\frac{P_{\tau_m}}{B_{\tau_m}}(S_{\tau_m} - K)^+) &= \mathbb{E}^{\mathbb{Q}}(\frac{P_{\tau_m}}{B_{\tau_m}}\frac{B_0}{P_0}(S_{\tau_m} - K)^+)\frac{P_0}{B_0} \\ &= \mathbb{E}^{Q^P}((S_{\tau_m} - K)^+)\frac{P_0}{B_0}\end{aligned}$$

Now let us approximate the dynamics of

$$\begin{aligned}S_t &= \frac{B_t(\tau_m) - B_t(\tau_n)}{P_t} \\ &= \frac{\sum_{i=m+1}^n B_t(\tau_i)L_t(\tau_{i-1},\tau_i)(\tau_i - \tau_{i-1})}{P_t}\end{aligned}$$

by a geometric Brownian motion with respect to Q^P. Note that $(S_t)_{t\in[0,\tau_m]}$ follows a positive valued process (since in the BGM setting, the bond curve is always decresing!). Furthermore, by the change-of-numeraire, $(S_t)_{t\in[0,\tau_m]}$ follows a martingale. With this, there exists a stochastic process $(\sigma^S(t))_{t\in[0,\tau_m]}$

$$dS_t = S_t\sigma_t^S dW_t^P$$

where $(W_t^P)_{t\in[0,T]}$ is a Brownian motion with respect to the annuity measure. Of course, the process $(\sigma^S(t))_{t\in[0,\tau_m]}$ is not deterministic, however the distribution of

$$S_{\tau_m} = S_0 e^{\int_0^{\tau_m}\sigma_t^S dW_t^P - \frac{1}{2}\int_0^{\tau_m}|\sigma_t^S|^2 dt}$$

could be approximatively described by

$$\overline{S}_{\tau_m} = S_0 e^{\overline{\sigma}^S W_{\tau_m} - \frac{1}{2}|\overline{\sigma}^S|^2\tau_m}$$

with some Brownian motion $(W_t)_t$, if the constant $\overline{\sigma}^S \in]0,\infty[$ is chosen appropriately, approximating

$$|\overline{\sigma}^S|^2 \approx \frac{1}{\tau_m}\int_0^{\tau_m}|\sigma_t^S|^2 dt = \frac{1}{\tau_m}\langle\int_0^{\cdot}\sigma_t^S dW_t^P\rangle_{\tau_m} = \frac{1}{\tau_m}\langle\int_0^{\cdot}\frac{1}{S_t}dS_t\rangle_{\tau_m}.$$

Note that such an approximation is perfect if the stochastic logarithm $(\sigma_t^S)_{t\in[0,\tau_m]}$ of the swap rate is deterministic.

To approximate the quadratic variation of the local martingale $(\int_0^t S_t^{-1}dS_t)_{t\in[0,\tau_m]}$, we first consider an approximation of the stochastic differential dS_t

$$\begin{aligned} dS_t &= d\left(\frac{\sum_{i=m+1}^n B_t(\tau_i)L_t(\tau_{i-1},\tau_i)(\tau_i-\tau_{i-1})}{P_t}\right) \\ &\approx \sum_{i=m+1}^n (\tau_i-\tau_{i-1})\frac{B_t(\tau_i)}{P_t}dL_t(\tau_{i-1},\tau_i). \end{aligned}$$

This estimation takes into account that the fraction $B_t(\tau_i)P_t^{-1}$ changes slowly, compared to the changes of LIBORs. That is

$$dS_t \approx S_t\sum_{i=m+1}^n (\tau_i-\tau_{i-1})\frac{B_t(\tau_i)}{P_t}\frac{L_t(\tau_{i-1},\tau_i)}{S_t}\gamma_t(\tau_{i-1})dW_t(\tau_i)$$

With this, we obtain

$$\begin{aligned} &\langle\int_0^{\cdot}\frac{1}{S_t}dS_t\rangle_{\tau_m} \approx \langle\int_0^{\cdot}\sum_{i=m+1}^n (\tau_i-\tau_{i-1})\frac{B_t(\tau_i)}{P_t}\frac{L_t(\tau_{i-1},\tau_i)}{S_t}\gamma_t(\tau_{i-1})dW_t(\tau_i)\rangle_{\tau_m} \\ &\approx \int_0^{\tau_m}\sum_{i,j=m+1}^n (\tau_i-\tau_{i-1})(\tau_j-\tau_{j-1})\frac{B_t(\tau_i)}{P_t}\frac{B_t(\tau_j)}{P_t}\frac{L_t(\tau_{i-1},\tau_i)}{S_t}\frac{L_t(\tau_{j-1},\tau_j)}{S_t}\gamma_t(\tau_{i-1})\gamma_t(\tau_{j-1})dt \end{aligned}$$

The reason for quadratic covariation

$$\langle W_\cdot(\tau_i), W_\cdot(\tau_j)\rangle_t = t$$

is that the Brownian motions $(W_t(\tau_j)_{t\in[0,\tau_j]}$, $j=1,\dots n$ differ from the original Brownian motion $(W_t)_{t\in[0,T]}$ by a process of bounded variation, due to Girsanov theorem.

The integral

$$\int_0^{\tau_m}\sum_{i,j=m+1}^n (\tau_i-\tau_{i-1})(\tau_j-\tau_{j-1})\frac{B_t(\tau_i)}{P_t}\frac{B_t(\tau_j)}{P_t}\frac{L_t(\tau_{i-1},\tau_i)}{S_t}\frac{L_t(\tau_{j-1},\tau_j)}{S_t}\gamma_t(\tau_{i-1})\gamma_t(\tau_{j-1})dt$$

contains random variables, whose realizations are not known at time $t=0$. That is, one suggests the following deterministic approximation:

$$\sum_{i,j=m+1}^n (\tau_i-\tau_{i-1})(\tau_j-\tau_{j-1})\frac{B_0(\tau_i)}{P_0}\frac{B_0(\tau_j)}{P_0}\frac{L_0(\tau_{i-1},\tau_i)}{S_0}\frac{L_0(\tau_{j-1},\tau_j)}{S_0}\int_0^{\tau_m}\gamma_t(\tau_{i-1})\gamma_t(\tau_{j-1})dt$$

and defines the value $|\overline{\sigma}^S|^2$ normalizing by duration τ_m

$$(8.13) \qquad \left\{ \begin{array}{rcl} |\overline{\sigma}^S|^2 &=& \frac{1}{\tau_m}\sum_{i,j=m+1}^n(\tau_i-\tau_{i-1})(\tau_j-\tau_{j-1})\frac{B_0(\tau_i)}{P_0}\frac{B_0(\tau_j)}{P_0}\times \\ && \times\frac{L_0(\tau_{i-1},\tau_i)}{S_0}\frac{L_0(\tau_{j-1},\tau_j)}{S_0}\int_0^{\tau_m}\gamma_t(\tau_{i-1})\gamma_t(\tau_{j-1})dt \end{array}\right.$$

Note that this quantity depends on the initial bond curve and on the choice of LIBOR volatilities

$$\gamma = (\gamma_t(\tau_i)_{t \in [0,\tau_i]}, i = 1, \ldots, n)$$

which determine the BGM model uniquely.

Having fixed the initial bond curve, we consider the dependence on LIBOR volatilities That is, the approximation depends on them

$$\overline{\sigma}^S = \overline{\sigma}^S(\gamma)$$

With this approximation, the following stylized model calibration procedure can be carried out

1. Model LIBOR volatilities, for instance by a family of deterministic functions, appropriately parameterized by $\vartheta \in \theta$

$$\gamma(\vartheta) = (\gamma_t^{\vartheta}(\tau_i)_{t \in [0,\tau_i]}, i = 1, \ldots, n), \qquad \vartheta \in \theta$$

2. Observe listed Swaption price C_0. Calculate the implied Black's volatility σ^{Bl}, which corresponds to this price via Black's caplet formula:

$$(8.14) \qquad \left\{ \begin{array}{lcl} C_0 & = & \mathbb{E}^{Q^P}((S_{\tau_m} - K)^+)\frac{P_0}{B_0} = P_0(S_0 N(d_1) - K N(d_2)) \\ d_1 & = & \frac{\ln(S_0/K) + \frac{1}{2}(\sigma^{Bl})^2 \tau_m}{\sqrt{(\sigma^{Bl})^2 \tau_m}} \\ d_2 & = & \frac{\ln(S_0/K) - \frac{1}{2}(\sigma^{Bl})^2 \tau_m}{\sqrt{(\sigma^{Bl})^2 \tau_m}} \end{array} \right.$$

3. Find a parameter ϑ^* which best explains this implied volatility

$$\vartheta^* = \text{argmin}_{\vartheta \in \theta} |\overline{\sigma}^S(\gamma(\vartheta)) - \sigma^{Bl}|^2$$

In a generic situation, there are many Swaptions, traded at the market, say with prices $S_0^1, \ldots, S_J^1$, $J \in \underline{\mathbb{N}}$ with different times to first settlement and different contract durations. In this case, the implied Black volatilities are parameterized by two time parameters, giving the so-called *volatility surface*. One tries to approximate this surface in the sense that

$$\vartheta^* = \text{argmin}_{\vartheta \in \theta} \sum_{j=1}^{J} |\overline{\sigma}^S(\gamma(\vartheta), j) - \sigma^{Bl,j}|^2.$$

Note that both, the implied volatility $\sigma^{Bl,j}$ and the approximation $\overline{\sigma}^S(\gamma(\vartheta), j)$ depend on the parameters of the Swaption j, their calculation must be adopted from (8.14) and (8.13) accordingly.

9 Credit risk

9.1 Credit Derivatives. These are securities

- whose payoff is affected by *credit event*
- used to transfer, to hedge or to manage credit risk

A typical structure of a credit derivative is

- Party A (protection buyer)
- Party B (protection seller)
- Contract between A and B encompasses
  - reference credit
  - credit event definition
  - cash flows (protection fee and claim at credit event)

Example Consider the so-called digital swap

- reference credit: Eurobond issued by Brazil
- credit event: missed coupon or principal payment
- claim at credit event: B pays 10^6 USD to A

Other typical credit-linked instruments include

Loans: These are bilateral contracts and are still most important instruments to raise the investment capital. Within a typical loan,

- at origination, τ_0 lender pays K to borrower
- at maturity τ_n borrower returns K to the lender
- in between, at times τ_i, $i = 1, \ldots, n$, borrower pays agreed amounts c_i, $i = 1, \ldots n$ to the lender

Bonds: These are securities version of loans, they are tradable in small dominations.

- At issuance the investors buy (defaultable) bonds from the issuer
- At coupon payment dates τ_i, $i = 1, \ldots, n$ each bond provides a coupon payment c_i.

- At maturity, the issuer pays the last coupon and the principle

Advantage of bond with respect to loans are

- marketability (can be sold/bought in between)
- access to a large class of lenders
- value add-on to the bond due to flexibility
- transparency because of the market price
- hedge against default risk (short positions)

However, if the amount of required capital is small, loans are preferred to bonds.

There are different types of (risky, defaultable) bonds.

- zero bonds (no coupon payments)
- fixed-coupon bonds (coupon payments are fixed), these bonds are most common
- par floaters: principal payment is fixed, but coupons are related to the market interest rate. Usually they refer to LIBOR rates.
- convertable bonds: the holder of the bond has the right to transfer the bond into a pre-specified number of shares. These type of bonds are popular for risky start-up companies.

Another type of instruments are *asset swaps* which convert cash flow from a defaultable bond into LIBOR-based stream plus a spread.

Example A sells to B asset swap, then

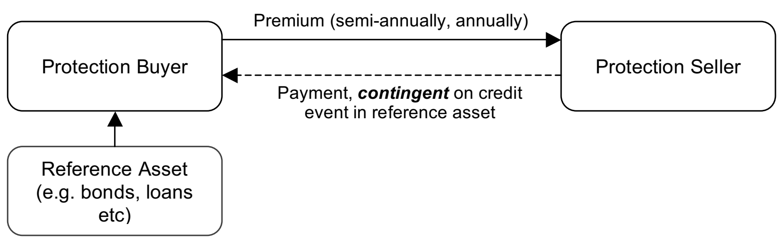
- B provides to A cash-flows from coupons of a pre-specified defaultable bond
- A provides to B cash-flows from coupons of an appropriate floating rate note plus an additional amount, called asset swap spread

In certain sense, asset swap spread reimburses the owner for default risk. At issuance of the asset swap, the swap spread is adjusted accordingly. Asset swap does not protect the buyer from potential loss in the credit event. If the bond defaults, then the payments (on both sides) stop.

Now, let us turn to one of the most important assets in the credit risk management, the so-called *credit default swap* (CDS). A single-name credit default swap encompasses the following structure

- B agrees to pay a claim to A in the case that a credit event of the reference entity occurs
- A pays regularly a fee for this protection

Thereby, the credit even can be defined broadly. For instance, a credit spread widening or rating downgrade.



The details of CDS are

- at default, different alternatives for protection payment are possible
  - physical delivery of another asset
  - payment of a notional amount less the post default market value of the reference asses
  - a pre-agreed fixed payoff (as in the case of digital swap)
- the protection fee payment may have following alternatives
  - is regularly, at agreed times
  - fixed, once upfront
  - if default occurs between payment times, the party A may have to pay a fraction of the next payment

Example Consider a credit default swap

- reference asset: Daimler
- term of the swap: 5 years

- notional 20×10^6 USD
- fee $= 166$ bp

In this case, prior to default, A pays to B

$$\frac{\text{fee} \ \times \ \text{notional}}{m} \qquad \begin{array}{c} m \text{ times a year, in} \\ \text{regular intervals} \end{array}$$

In our case, this would be

$$116 \times 10^{-4} \times 20 \times 10^6 \times \frac{1}{2} = 116,000 \quad \text{USD, semi-annually}$$

At occurrence of default event, A notifies B and pays the last protection fraction

$$116,000 \times \frac{\text{number of days to next payment}}{\text{number of days between payments}}$$

Then B pays to A a certain amount. Usually, this amount is intended to reimburse for a loss which is caused by credit event on 20×10^6 USD, invested in Daimler. A typical situation here is that one considers a loss for 20×10^6 USD invested in a bond. For instance, if the bond with face value of 1000 USD is traded at 430 USD after default, then the loss on this bond is

$$\frac{1000 - 430}{1000} = 0.57$$

In this case, B pays to A $0.57 \times 20 \times 10^6 = 11.4$ USD.

In practice, many problems are to be solved, before in the case of a credit event the protection buyer receives its protection

- Determining the fact of default. (Is restructuring of a company a default? What happens if the company is taken over or merged?)
- Determining the payment in the case of default. Due to liquidity issues, it is difficult to determine the after-default price. Furthermore, if default payment is based on delivery of a physical asset, then problems with delivery may occur (again, because of liquidity).

For the issuer of CDS, it is very important to hedge against potential loss in the case of default. A typical strategy which reduces exposure from CDS issuance is to take short positions in the appropriate bonds. However, the way of creating short positions must be worked out. There are two possibilities to create short positions

- One way is the so-called repo-transactions, which are typical for government bonds. Here, the bonds are bought from a party, which intends to hold them. At the same time, both parties enter a bilateral re-purchase agreement, where the bonds are returned at a pre-determined future date

- The other way is to sign short positions at the forward market. Here one needs a counter party which agrees to hold long forward postilion, this may be a problem for large volumes.

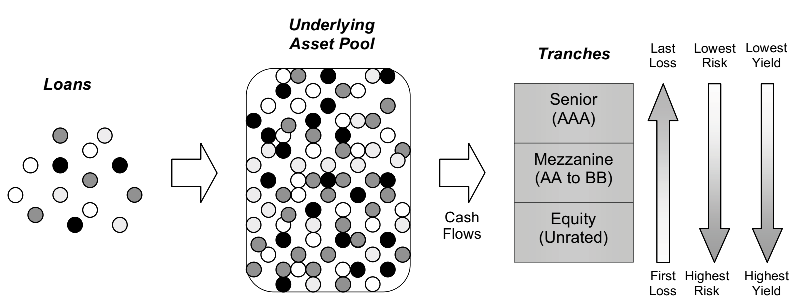
Other credit-linked derivatives are

- options on defaultable bonds such as Calls and Puts of European, American or Bermudian style

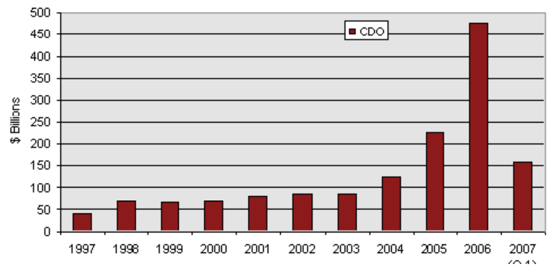
- credit spread options. Here, the yield difference between a risk-free and risky bond is important. (Note that the yield of risky bond is calculated in the same way as for non-risky bonds, but for promised payments). Such an option may give the right to buy a risky bond at certain date at a pre-specified yield, given by the yield of a similar riskless reference bond plus a fixed spread.

Collateralized debt obligations (CDO's) Collaterized Debt Obligations, or CDOs for short, are investment-grade securities backed by a pool of bonds, loans and other assets. These assets are sold to a specially designed company which issues notes whose payment depends on the losses in the pool. Thereby, the notes refer to different tranches in the portfolio loss.

The advantage of CDOs is to create higher-quality investment instruments (senior tranche) from from a pool of low-quality, not rated, or speculative-grade investments. This could include corporate bonds, government bonds, asset-backed securities, bank loans, mortgage-backed securities, and even tranches from other CDOs. The credit quality increase should be given by creation of protection tranches. This could help to attract capital from agents who would not invest in low-quality assets.

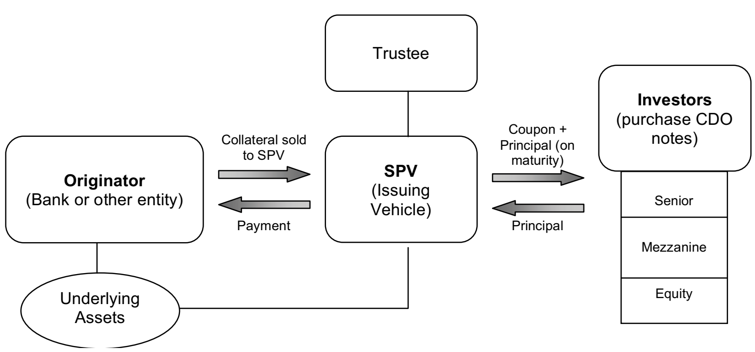


These instruments have represented the fastest growing segments in the credit derivatives *before the financial crisis* and and are under discussion now.



There are different types of CDOs with different structures, the basic concept remains the same repackage the risk profile of the underlying pool of assets into securities having different degrees of seniority in their claims to the underlying assets. A CDO is typically constructed as follows:

1. The CDO Originator (usually an Investment Bank) acquires an inventory of assets.
2. Cash flows from the underlying asset pool are repackaged and sold to investors along with the associated risks. These sold rights are called tranches, each with its own risk and reward exposure depending on the contractual terms of the CDOs.



The senior tranche is the first to receive any income, and the last to suffer any losses. The equity tranche is the first to take any losses should borrowers default on loans in the underlying asset pool and the last to receive any income. Each tranche of the CDO may be given a rating by credit rating agencies such as Standard & Poor's or Moody's which reflect the likelihood that interest and principal payments will be made on time. Senior tranches of

CDOs often receive very high ratings (AAA), followed by Mezzanine tranches (AA to BB), and finally equity tranches (unrated/junk). The varying risk profiles and coupon payments naturally give the tranches different values to investors, and consequently they are sold at different prices and traded at varying credit spreads. Tranching enables CDO originators to create additional classes of securities whose ratings are higher than the average rating of the underlying asset pool.

9.2 Firm value based modeling. The simplest (and the first) among credit risk models is due to Merton. Here one assumes that a firm issues at the initial time $t = 0$ two types of securities

- debt: this is a zero coupon bond with face value D and maturity date T
- equity: this are shares of the company

The price processes are given by

- firm value $(V_t)_{t\in[0,T]}$
- risky bond $(\overline{B}_t(T))_{t\in[0,T]}$
- shares $(S_t)_{t\in[0,T]}$

In Merton's model, one supposes that the default of the company can occur only at the final date T. The default event is defined as a situation, where the company fails to repay its debt. In this case

- the company is liquidated
- the bond holder receive their payment first
- the remaining part is paid to share holders

With this definition, the terminal values of bond and shares are related to the terminal value of the firm as follows

$$
\begin{aligned}
\overline{B}_T(T) &= D - (D - V_T)^+ \\
&= V_T - (V_T - D)^+ \qquad (9.1)\\
S_T &= (V_T - D)^+ \qquad (9.2)
\end{aligned}
$$

In the framework of Merton's model, the focus is on determining the no-arbitrage price dynamics of all three assets

$$(V_t)_{t\in[0,T]}, \qquad (\overline{B}_t(T))_{t\in[0,T]}, \qquad (S_t)_{t\in[0,T]}.$$

Assuming a constant and deterministic short rate $r > 0$, the firm value is supposed to folllow

$$dV_t = V_t(r dt + \sigma^V dW_t), \qquad t \in [0, T]$$

where $\sigma^V \in]0, \infty[$ represents the volatility of the firm value and $(W_t)_{t \in [0,T]}$ stands for the process of Brownian motion with respect to the spot martingale measure $\mathbb{Q}$.

With the assumption (9.2), we obtain the shares price

$$\frac{S_t}{e^{rt}} = \mathbb{E}^{\mathbb{Q}}\left(\frac{(V_T - D)^+}{e^{rT}} \,|\, \mathcal{F}_t\right), \qquad t \in [0, T]$$

which means that the shares price must be given by the ordinary Black-Scholes formula with strike price D, applied to the firm value

$$\begin{aligned}
S_t &= \mathrm{BS}(V_t, D, \sigma^V, r, T - t) \qquad t \in [0, T] \\
&= V_t \Phi(d_1) - e^{-r(T-t)} D \Phi(d_2) \\
d_1 &= \frac{\ln(V_t/D) + r(T-t) + \frac{1}{2}(\sigma^V)^2 (T-t)}{\sigma^V \sqrt{T-t}} \\
d_2 &= \frac{\ln(V_t/D) + r(T-t) - \frac{1}{2}(\sigma^V)^2 (T-t)}{\sigma^V \sqrt{T-t}}
\end{aligned}$$

Furthermore, using (9.1), we obtain

$$\overline{B}_t(T) = V_t - \mathrm{BS}(V_t, D, \sigma^V, r, T-t) \qquad t \in [0, T] \tag{9.3}$$

At this point let us derive some observation about the dependence of the price $\overline{B}_t(T)$ of the risky bond on model quantities. If all other quantities remain constant, then the bond price is

- increasing in firm value V_t

  $$\overline{B}_t(T) = e^{-r(T-t)} D - e^{-r(T-t)} \cdot \mathbb{E}^{\mathbb{Q}}((D - V_T)^+ \,|\, \mathcal{F}_t)$$

  since the price of a put is a decreasing function in the price of the underlying.

- non-decreasing in the face value of the bond D (=firm's debt). This is clear since the terminal payoff of the bond

  $$\overline{B}_T(T) = V_T - (V_T - D)^+$$

  is non-decreasing in D.

- is decreasing in the interest rate r. (Why?)

- is increasing in time to maturity, because in (9.3) the Black-Scholes price is decreasing in time to maturity.

- is decreasing in the volatility of the firm's value σ^V, since in (9.3) the Black-Scholes price is increasing in volatility. (Black-Scholes Vega is always positive).

An interesting consequence of the last observation is known as *asset substitution*. In reality, it is possible to increase the riskyness of the firm by changing its capital structure. Namely, by selling firm's assets and investing them in high-volatility securities. According to the theory, this does not change the firm value but changes its volatility. What bond holder lose by the increase in firm's value volatility is given to the shareholders, since

$$\begin{aligned}
\overline{B}_t(T) &= V_t - \mathrm{BS}(V_t, D, \sigma^V, r, T-t) \\
S_t &= \mathrm{BS}(V_t, D, \sigma^V, r, T-t)
\end{aligned}$$

This feature is an important reason for covenants in bonds, since bond holders need to exercise some control over the investment decisions of the firm, or at least being able to transfer their bonds into firm's shares.

9.3 A practical implementation of Merton's model. In practice, it is almost impossible to observe the value of the firm, because of large degree of freedom in the asset valuation and due to different accounting procedures. One way out is to take the market capitalization of the firm (this is total value of all its shares). Furthermore, it is convenient to estimate the volatility of the firm from the exchange trading of its shares.

Thus, according to the Merton's model,

$$S_t = \mathrm{BS}(V_t, D, \sigma^V, r, T-t) \qquad t \in [0, T]$$

the Ito's formula show that

$$dS_t = S_t r dt + \partial_V \mathrm{BS}(V_t, D, \sigma^V, r, T-t)\sigma^V V_t dW_t.$$

On the other hand, it seems possible to estimate the volatility σ^S of the stock price from historical share prices directly, using the approximation

$$dS_t \approx S_t(r_t dt + \sigma^S dW_t)$$

Combining this estimates, we would obtain the equation

$$S_t \sigma^S \approx \partial_V \mathrm{BS}(V_t, D, \sigma^V, r, T-t)\sigma^V V_t$$

Therefore, one obtains two equations, which hold approximatively

$$\begin{aligned}
S_t &= \mathrm{BS}(V_t, D, \sigma^V, r, T-t) \\
S_t \sigma^S &= \partial_V \mathrm{BS}(V_t, D, \sigma^V, r, T-t)\sigma^V V_t
\end{aligned}$$

On the left-hand side, we have observable quantities, stock price S_t and historical stock volatility σ^S. On the right-hand side, we have only two non-observable quantities, the firm's value V_t and the firm's volatility σ^V. Although these equations are non-linear, numerical root searchers usually quickly determine the solution (V_t, σ_t^V). Thus, it is in principle possible to reconstruct firm's value and firm's volatility from stock price observations, given interest rate r, firm's debt D and time $T-t$.

Finally, let us mention an interesting application of Merton's model in the implementation of Moody's valuation procedure KMV. According to the Merton's model, the value of the firm at final time T is given by

$$V_T = V_t e^{(T-t)r - \frac{1}{2}(T-t)(\sigma^V)^2 + \sigma^V(W_T - W_t)}$$

Thus, the probability of finishing below D at T, given the firm's value V_t is

$$\begin{aligned}
\mathbb{Q}(V_T \le D \,|\, \mathcal{F}_t) &= \mathbb{Q}((T-t)r - \frac{1}{2}(T-t)(\sigma^V)^2 + \sigma^V(W_T - W_t) \le \ln(D/V_t) \,|\, \mathcal{F}_t) \\
&= \mathbb{Q}(\frac{\sigma^V(W_T - W_t)}{\sigma^V\sqrt{T-t}} \le \frac{\ln(D/V_t) - (T-t)r + \frac{1}{2}(T-t)(\sigma^V)^2}{\sigma^V\sqrt{T-t}} \,|\, \mathcal{F}_t) \\
&= \Phi(\frac{\ln(D) - \ln(V_t) - (T-t)r + \frac{1}{2}(T-t)(\sigma^V)^2}{\sigma^V\sqrt{T-t}}) \\
&= \Phi(\frac{\ln(D) - \mathbb{E}^{\mathbb{Q}}(\ln(V_T) \,|\, \mathcal{F}_t)}{\sigma^V\sqrt{T-t}})
\end{aligned}$$

In other words, the risk-neutral probability, that the firm survives is given by

$$\begin{aligned}
1 - \mathbb{Q}(V_T \le D \,|\, \mathcal{F}_t) &= 1 - \Phi(\frac{\ln(D) - \mathbb{E}^{\mathbb{Q}}(\ln(V_T) \,|\, \mathcal{F}_t)}{\sigma^V\sqrt{T-t}}) \\
&= \Phi(\frac{\mathbb{E}^{\mathbb{Q}}(\ln(V_T) \,|\, \mathcal{F}_t) - \ln(D)}{\sigma^V\sqrt{T-t}}).
\end{aligned}$$

The higher is this

$$\text{distance to default} \quad \frac{\mathbb{E}^{\mathbb{Q}}(\ln(V_T) \,|\, \mathcal{F}_t) - \ln(D)}{\sigma^V\sqrt{T-t}} \tag{9.4}$$

the more credible should be the firm. One basic idea of Moody's KMV is to use the distance to default as a key accounting ratio, based on which the default probabilities are estimated from historical data. That is, Moody's KMV does not follow Merton's model, but uses it's outcome as a statistical scoring method.

9.4 An extension of Merton's model. Consider now the case where the interest rate is stochastic and default can occur prior to the terminal time T. Suppose that

- default free zero bond with maturity T follows

$$dB_t(T) = B_t(T)\left(r_t dt - \sigma^r \cdot (T-t) dW_t^1\right)$$

- firm value follows

$$dV_t = V_t(r_t dt + \sigma^V(\rho dW_t^1 + \sqrt{1-\rho^2}dW_t^2))$$

where $(W_t^1, W_t^2)_{t\in[0,T]}$ is a process of Brownian motion with respect to the spot martingale measure $\mathbb{Q}$ and $\sigma^r, \sigma^V \in]0,\infty[$ are volatility parameters. In this framework, the dependence of firm's value on interest rates is represented by a correlation parameter $\rho \in]0,1[$.

Suppose that the default of the firm may occur at a random time, if the firm's value falls below the time value of its debt. Define stopping time

$$\begin{aligned}\overline{\tau} &= \inf\{t \in [0,T] \,:\, V_t \le DB_t(T)\} \\ &= \inf\{t \in [0,T] \,:\, V_t/B_t(T) \le D\}\end{aligned}$$

and agree to model the default time of the firm by $\overline{\tau}$. Assume that the risky bond is normalized. In the case of no-default, it pays 1 currency unit at maturity T, whereas in the default case the bond holder receives $1-c$ currency units at T. Thereby, $c \in [0,1]$ can be considered as the *protection rate*.

Thus, we consider the no-arbitrage price dynamics $(\overline{B}_t(T))_{t\in[0,T]}$ of defaultable bond , which is given by

$$\begin{aligned}\frac{\overline{B}_t(T)}{B_t(T)}1_{\{\overline{\tau}>t\}} &= \mathbb{E}^{\mathbb{Q}^T}(\frac{\overline{B}_T(T)}{B_T(T)}\,|\,\mathcal{F}_t)1_{\{\overline{\tau}>t\}} \\ &= \mathbb{E}^{\mathbb{Q}^T}(1_{\{\overline{\tau}>T\}} + (1-c)1_{\{\overline{\tau}\le T\}}\,|\,\mathcal{F}_t)1_{\{\overline{\tau}>t\}} \\ &= (1 - c\mathbb{Q}_t^T(\overline{\tau} \le T))1_{\{\overline{\tau}>t\}}\end{aligned}$$

Let us calculate the pre-default price dynamics of the defaultable bond

$$\overline{B}_t(T)1_{\{\overline{\tau}>t\}} = B_t(T)(1 - c\mathbb{Q}_t^T(\overline{\tau} \le T))1_{\{\overline{\tau}>t\}} \qquad \text{for all } t \in [0,T].$$

Since

$$\begin{aligned}\{\overline{\tau} > T\} \cap \{\overline{\tau} > t\} &= \{V_u > B_u(T)D \,:\, u \in [t,T]\} \cap \{\overline{\tau} > t\} \\ &= \{\frac{V_u}{B_u(T)} > D \,:\, u \in [t,T]\} \cap \{\overline{\tau} > t\}\end{aligned}$$

we need to focus on the dynamics of the process

$$\tilde{V}_u = \frac{V_u}{B_u(T)} \quad u \in [t,T]$$

with respect to the forward measure $\mathbb{Q}^T$. According to the change of numearaire technique, it holds that

$$\begin{aligned}d\tilde{V}_u &= \tilde{V}_u\left(\left[\sigma^V\rho, \sigma^V\sqrt{1-\rho^2}\right] - [-\sigma^r(T-u), 0]\right) d\begin{bmatrix} W_t^{1,T} \\ W_u^{2,T}\end{bmatrix} \\ &= \tilde{V}_u\left((\sigma^V\rho + \sigma^r(T-u))dW_u^{1,T} + \sigma^V\sqrt{1-\rho^2}dW_u^{2,T}\right)\end{aligned}$$

That is,

$$\tilde{V}_u = \tilde{V}_t e^{\int_t^u \Sigma_s dW_s^T - \frac{1}{2}\int_t^u |\Sigma_s|^2 ds} \qquad u \in [t,T]$$

where $(W_t^T)_{t\in[0,T]}$ is a two-dimensional Brownian motion under the forward measure $\mathbb{Q}^T$ and the deterministic process $(\Sigma_t)_{t\in[0,T]}$ represents the difference of volatilities

$$\Sigma_u = \begin{bmatrix} \sigma^V \rho \\ \sigma^V \sqrt{1-\rho^2} \end{bmatrix} - \begin{bmatrix} -\sigma^r(T-u) \\ 0 \end{bmatrix} \qquad u \in [t,T]$$

Now, let us conclude that

$$\begin{aligned}
\{\tilde{V}_u > D \,:\, u \in [t,T]\} &= \{\ln(\tilde{V}_u/\tilde{V}_t) > \ln(D/\tilde{V}_t) \,:\, u \in [t,T]\} \\
&= \{M_u - \frac{1}{2}\langle M \rangle_u > \ln(D/\tilde{V}_t) \,:\, u \in [t,T]\}
\end{aligned}$$

where the martingale

$$M_u = \int_t^u \Sigma_s dW_s^T, \qquad u \in [t,T]$$

possesses the quadratic covariation

$$\begin{aligned}
\langle M \rangle_u &= \int_t^u \|\Sigma_s\|^2 ds \\
&= \int_t^u \| \left[\sigma^V \rho + \sigma^r(T-s), \sigma^V \sqrt{1-\rho^2}\right] \|^2 ds \\
&= \int_t^u \left((\sigma^V)^2 + 2\rho\sigma^V\sigma^r(T-s) + (\sigma^r)^2(T-s)^2\right) ds
\end{aligned}$$

It is well-known that each continuous martingale can be represented as a time-changed Browinan motion. Thus, there exists a process $(\tilde{W}_t)_{t\geq 0}$ of one-dimensional Brownian motion with respect to $\mathbb{Q}^T$ such that

$$\tilde{W}_{\langle M \rangle_u} = M_u, \qquad u \in [t,T].$$

In this special case, the time change is deterministic, since $(\langle M \rangle_u)_{u\in[t,T]}$ is deterministic.

Now, write

$$\begin{aligned}
\mathbb{Q}^T(\overline{\tau} > T \,|\, \mathcal{F}_t) 1_{\{\overline{\tau} > t\}} &= \mathbb{Q}^T\left(M_u - \frac{1}{2}\langle M \rangle_u \geq \ln(\frac{D}{\tilde{V}_t}) \text{ for } u \in [t,T]\right) 1_{\{\overline{\tau} > t\}} \\
&= \mathbb{Q}^T\left(\tilde{W}_u - \frac{1}{2}u \geq \ln(\frac{D}{\tilde{V}_t}) \text{ for } u \in [0, \langle M \rangle_T]\right) 1_{\{\overline{\tau} > t\}}
\end{aligned}$$

where

$$\begin{aligned}
\langle M \rangle_T &= \int_t^T \left[(\sigma^V)^2 + 2\rho\sigma^V\sigma^r(T-s) + (\sigma^r)^2(T-s)^2\right] ds \\
&= (\sigma^V)^2(T-t) + \rho\sigma^V\sigma^r(T-t)^2 + \frac{1}{3}(\sigma^r)^2(T-t)^3
\end{aligned}$$

Finally, let us cite the following result: Given a Brownian motion $(W_t^*)_{t\in[0,x]}$, the probability that the process $(W_t^* - \frac{1}{2}t)_{t\in[0,x]}$ does not exceed the boundary $k < 0$ is well-known and is given by

$$P(\inf_{t\in[0,x]}(W_t^* - \frac{1}{2}t) \geq k) = \Phi(\frac{-k - \frac{1}{2}x}{\sqrt{x}}) - e^k\Phi(\frac{k - \frac{1}{2}x}{\sqrt{x}})$$

Thus, we can put all together. The price evolution of the defaultable bond is given by

$$\overline{B}_t(T)1_{\{\overline{\tau}>t\}} = B_t(T)(1 - c(1 - p_t))1_{\{\overline{\tau}>t\}}$$

where

$$\begin{aligned} p_t &= \Phi(\frac{-k_t - \frac{1}{2}x_t}{\sqrt{x_t}}) - e^{k_t}\Phi(\frac{k_t - \frac{1}{2}x_t}{\sqrt{x_t}}) \\ k_t &= \ln(\frac{D}{\tilde{V}_t}) = \ln(\frac{B_t(T)D}{V_t}) \\ x_t &= (\sigma^V)^2(T-t) + \rho\sigma^V\sigma^r(T-t)^2 + \frac{1}{3}(\sigma^r)^2(T-t)^3 \end{aligned}$$

9.5 Intensity based approach to credit risk. So far, we have seen that the firm value based models attempt to describe the default event in terms of fundamental model quantities, such as firm value and firm's debt. On contrary to this, the intensity based model abstract from concrete situation. They do not attempt to model the occurrence of credit event using fundamentals. Instead the default time is modeled as a stopping time on an appropriate filtered probability space.

Let us consider an ad-hoc approach. based on concrete example. Here a citation from Dow Jones Newswires on 26 September 2007

GM's 8.375% notes due 2033 fell 28.75 points to 86.5 as the UAW strike entered its second day. Still, the cost of protecting GM bonds has barely budged on the week. GM credit default swaps were quoted at 565 basis points Tuesday, 20 basis points wider than where they began the week, according to Joe Patire, managing director at YieldQuest Securities in Atlanta. That means it now costs $565,000 per year to protect $10 million of GM bonds for five years.

Let's check.

- yield change in basis points $+28.77$, price falls to 86.5
- protection in basis points 565 means protection fee for ten millions: $565 \times 10^{-4} \times 10 \times 10^6 = 565 \times 10^3$
- protection increase for ten millions because of the strike: $20 \times 10^{-4} \times 10 \times 10^6 = 20 \times 10^3$

Note that because of strike

- Market realizes a higher default probability of GM
- Market price of GM bonds drops, the (promised) yield increases
- Market traded CDS on GM increase in fair spread.

Consider now an abstract situation. Given a filtered probability space $(\Omega, \mathcal{F}, \mathbb{Q}, (\mathcal{F}_t)_{t\in[0,T]})$ where Q represents the spot martingale measure. Suppose that the dynamics of non-risky zero bonds is described in terms of

$$(B_t(\tau))_{t\in[0,\tau]}, \quad \tau \in [0,T], \qquad (B_t)_{t\in[0,T]} \tag{9.5}$$

and suppose that the default time (of a certain entity) is described in terms of a stopping time

$$\left\{ \begin{array}{c} \overline{\tau} : \Omega \to [0,T] \cup \{\infty\} \\ \text{which is independent on all} \\ \text{random variables in (9.5).} \end{array} \right. \tag{9.6}$$

For the moment we suppose that it follows exponential distribution

$$Q(\overline{\tau} \le t) = 1 - e^{-\lambda t}, \quad t \in [0,T], \qquad \mathbb{Q}(\overline{\tau} > T) = e^{-\lambda T}$$

We suppose that the parameter $\lambda > 0$ is fixed but not known yet and attempt to adjust it in such way, that it explains the market situation.

Remark The parameter $\lambda > 0$ can be interpreted as default intensity in the following sense.

$$\begin{aligned} Q(\overline{\tau} \in]t, t+\Delta] \,|\, \tau > t) &= \frac{\mathbb{Q}(\{\overline{\tau} \in]t, t+\Delta]\} \cap \{\tau > t\})}{\mathbb{Q}(\overline{\tau} > t)} \\ &= \frac{\mathbb{Q}(\overline{\tau} > t) - \mathbb{Q}(\overline{\tau} > t+\Delta))}{\mathbb{Q}(\overline{\tau} > t)} \\ &= \frac{e^{-\lambda t} - e^{-\lambda(t+\Delta)}}{e^{-\lambda t}} \\ &= 1 - e^{-\lambda\Delta} \approx \lambda\Delta \end{aligned}$$

That is, the probability that the company defaults over next short period is proportional to the duration of the period, the proportionality factor is λ.

Consider a particular CDS contract with payment dates $\tau_1 < \tau_2 < \ldots, < \tau_n$. That is,

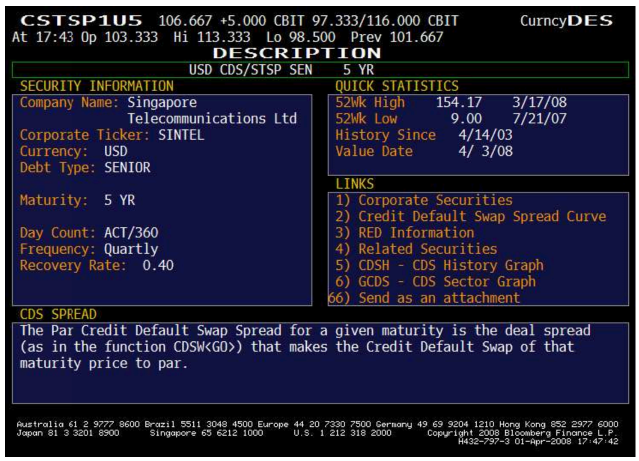
- the premium leg requires payment

  $$x \times (\tau_i - \tau_{i-1}) \quad \text{at } \tau_i$$

  if no default occurred within $[0, \tau_i]$. In the case of default, premium leg does not require payments.

- the protection leg pays δ currency units at the default time $\overline{\tau}$, if default occurred within protection time $[0, \tau_n]$.
- the value x^* is called *fair CDS spread* if the market price of CDS contract with premium x^* is quoted at zero.

A typical problem in the intensity based approach is to calibrate a certain model to the market data. In our simplified situation, we need to determine the parameter $\lambda^* \in]0, \infty[$ which explains the quoted CDS spread x^*. As an illustration, we consider the following contract



whose fair spread, quoted on annual basis and paid quarterly, is 101.67 basis points.

First calculate the value of the premium leg, prior to default. The payment at τ_i of $1_{\{\overline{\tau} > \tau_i\}}$ can be considered as payoff of defaultable zero bond with price evolution is $(\overline{B}_t(\tau_i))_{t \in [0, \tau_i]}$. Prior to default we have

$$
\begin{aligned}
\overline{B}_0(\tau_i) &= \mathbb{E}^{\mathbb{Q}}\left(\frac{1_{\{\overline{\tau} > \tau_i\}}}{B_{\tau_i}}\right) \\
&= B_0(\tau_i) \mathbb{E}^{\mathbb{Q}}(1_{\{\overline{\tau} > \tau_i\}})
\end{aligned}
$$

because of independence. Note that the risk-neutral *survival probability* $Q(\overline{\tau} > \tau_i)$ explains the spread between the risky and non-risky bond. In our context of exponentially distributed default time, we have

$$
\overline{B}_0(\tau_i) = B_0(\tau_i) e^{-\lambda \tau_i}, \qquad i = 1, \ldots, n
$$

That is, the initial value of the premium leg is given as

$$
V_0^{\text{prem}}(x) = x \times \sum_{i=1}^{n} B_0(\tau_i) e^{-\lambda \tau_i} (\tau_i - \tau_{i-1})
$$

Now, calculate the value of the default leg. Note that CDS pays protection δ at $\overline{\tau}$ if default occurs prior to maturity τ_n. If we equivalently transfer this payoff to the maturity τ_n (by investing in standard savings account), then we obtain

$$\delta 1_{\{\overline{\tau}\leq\tau_n\}}\frac{B_{\tau_n}}{B_{\overline{\tau}}} \quad \text{at } \tau_n.$$

That is, the value of the default leg is given as

$$\begin{aligned}
V_0^{\text{def}} &= \mathbb{E}^{\mathbb{Q}}\left(\frac{\delta 1_{\{\overline{\tau}\leq\tau_n\}}B_{\tau_n}/B_{\overline{\tau}}}{B_{\tau_n}}\right) \\
&= \delta\mathbb{E}^{\mathbb{Q}}\left(1_{\{\overline{\tau}\leq\tau_n\}}\frac{1}{B_{\overline{\tau}}}\right) \\
&= \delta\mathbb{E}^{\mathbb{Q}}\left(\int_0^{\tau_n}\lambda e^{-\lambda s}B_s^{-1}ds\right) \\
&= \delta\int_0^{\tau_n}\lambda e^{-\lambda s}\mathbb{E}^{\mathbb{Q}}(B_s^{-1})ds \\
&= \delta\int_0^{\tau_n}\lambda e^{-\lambda s}B_0(s)ds
\end{aligned}$$

where we have utilized the independence assumption (9.6). In particular, the fair CDS spread at time $t=0$ is given as solution x^* to $V_0^{\text{prem}}(x)=V_0^{\text{def}}$ giving

$$x\times\sum_{i=1}^n B_0(\tau_i)e^{-\lambda\tau_i}(\tau_i-\tau_{i-1})=\delta\int_0^{\tau_n}\lambda e^{-\lambda s}B_0(s)ds$$

hence

$$x^*=\frac{\delta\int_0^{\tau_n}\lambda e^{-\lambda s}B_0(s)ds}{\sum_{i=1}^n B_0(\tau_i)e^{-\lambda\tau_i}(\tau_i-\tau_{i-1})}$$

Note that the following approximation may hold

$$x^*=\lambda\delta\frac{\int_0^{\tau_n}e^{-\lambda s}B_0(s)ds}{\underbrace{\sum_{i=1}^n B_0(\tau_i)e^{-\lambda\tau_i}(\tau_i-\tau_{i-1})}_{\int_0^{\tau_n}e^{-\lambda s}B_0(s)ds}}\approx\lambda\delta. \tag{9.7}$$

For the case of CDS on SingTel depicted above, we assume that the interest rate $r=0.03$ and that the protection is paid quarterly to calculate λ^* as follows:

Premium leg value:

$$\begin{aligned}
V^{\text{prem}}(x^*) &= \underbrace{\text{Notional}}_{=1}\times\underbrace{0.0101667}_{\text{fee}}\sum_{k=1}^n e^{-r\tau_k}e^{-\lambda\tau_k}\underbrace{(\tau_k-\tau_{k-1})}_{=1/4} \\
&= \frac{0.0101667}{4}\sum_{k=1}^{20}e^{(-0.03-\lambda)k/4}
\end{aligned}$$

Protection leg value:

$$\begin{aligned}
V^{\text{def}} &= \underbrace{\text{Notional}}_{=1}\underbrace{\delta}_{0.6}\int_0^{\tau_n}\lambda e^{-\lambda s}e^{-rs}ds \\
&= 0.6\lambda\int_0^5 e^{-(0.03+\lambda)s}ds \\
&= \frac{0.6\lambda}{0.03+\lambda}(1-e^{-5\cdot(0.03+\lambda)})
\end{aligned}$$

equating both, we obtain

$$\frac{0.0101667}{4}\sum_{k=1}^{20}e^{(-0.03-\lambda)k/4} = \frac{0.6\lambda}{0.03+\lambda}(1-e^{-5\cdot(0.03+\lambda)})$$

whose solution is $\lambda^* \approx 0.01684547$.

Remark Note that the approximation (9.7) in this case is

$$\delta\times\lambda^* = 0.6\times 0.01684547 = 0.01010728 \approx 0.0101667 = x^*$$

In the previous case of GM's strike, we may conclude that the change $\Delta x^* = 20$ (basis points) in the fair CDS spread is approximately equal to the protection rate times the change in default intensity giving $\delta\cdot(\Delta\lambda)$. On the other hand, $(\Delta\lambda)$ could be an approximation to the yield change, thus $(\Delta\lambda)\approx 28.75$. The recovery rate, which explains this would be

$$\delta = \frac{20}{28.77}\approx 0.7.$$

Having calibrated our model to the listed CDS spread, other credit linked instruments (all types of coupon bonds!) can be priced.

With this over-simplistic modeling of default by exponentially distributed independent random time, following problems will occur:

- If another fair spread is listed, for instance where all parameters are the same but the maturity is 10 years, instead of 5 years. Then, most likely, the determined parameter $\lambda^*\approx 0.01684547$ will not correspond to the fair spread of this CDS. In other words, the same calibration procedure, applied to the new CDS will result in other value for λ^*. This shows that instead of a single parameter, a curve $(\lambda_s)_{s\in[0,T]}$ of somehow time-changing default intensities needs to be considered.

- Assuming time dependent deterministic intensity $(\lambda_s)_{s\in[0,T]}$, we may define the distribution of the default time as

$$Q(\overline{\tau}>t) = e^{-\int_0^t\lambda_s ds}, \qquad \text{for all } t\in[0,T]$$

  This may help to fit the model to all fair CDS spreads at the beginning. As the time evolves, the spreads will change. The previously chosen default intensity $(\lambda_s)_{s\in[0,T]}$ will not explain the market situation after a while.

- This suggests to model the default intensity as a stochastic process. A realization of this insight yields an important class of credit risk models, the so-called *intensity based* approach. It turns out that the valuation of credit-linked derivatives within the intensity-based approach is very similar to the techniques known from the from interest rate theory.

9.6 Ingredients of intensity based approach. Now, we consider credit-linked instruments which are modeled within time horizon $[0, T]$ using the following ingredients:

- The default time is modeled by random variable $\overline{\tau}$ with values in $[0, T] \cup \{+\infty\}$, in the most cases as a stopping time within a given filtration.

- If the default occurs before or at maturity T, then the owner of the contract receives a recovery payment. The recovery is modeled by explicitly specified stochastic process $(Z_t)_{t \in [0,T]}$. The recovery is paid at default time $\overline{\tau}$ and is given by $Z_{\overline{\tau}}$.

- If no default occurs within $[0, T]$, then the owner of the contract receives a promised claim X_T at maturity T. This could be a deterministic payoff (for instance face value of a bond) or another security which may depend on prices of other assets at maturity time T

The modeling of recovery process is a difficult task. In reality, bond holders have a certain priority over share holders. That is, if the company defaults, then the firm value is used to pay the debt to the bond holders first, the remaining part is then given to the shareholders. However, depending on the definition of the defaultable security, different recovery rules may apply. (It is also possible that the instrument does not pay any recovery). Historical recovery rates show that the usual recovery can be assumed around 40% - 60% of the promised payoff of the defaultable asset. The detailed modeling of recovery procedure may be very involving. For this reason, the theory uses three approximative procedures. These rules do not describe one of the common practices in the real-world markets but are used as approximations.

- Recovery of the face value. At default, the investor receives a particular fraction of the promised payoff. For instance, a defaultable zero bond with recovery of the face value is given by

$$(Z_t = \underbrace{1 - q}_{\delta})_{t \in [0,T]}, \qquad X_T = 1$$

- Recovery of the treasury. It is applicable for securities where there exists a non-risky asset, similar to the defaultable asset. At default, the owner receives a fraction of this non-risky asset. For instance, a defaultable zero bond with recovery of the treasury is given by

$$(Z_t = (1 - q) B_t(T))_{t \in [0,T]}, \qquad X_T = 1$$

where $(B_t(T))_{t\in[0,T]}$ stands for the price process of non-defaultable zero bond.

- Recovery of the market value. It is applicable for a broad class of assets, but can not be modeled by an exogenously specified recovery process $(Z_t)_{t\in[0,T]}$. At default, the owner receives a fraction of the pre-default market price of the defaulted security.

9.7 A Deterministic situation. Now let us consider a toy example which we use to demonstrate the concept. Consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{Q}, (\mathcal{F}_t)_{t\in[0,T]})$ where $\mathbb{Q}$ represents the spot martingale measure. Assume that

- the short rate $(r_t)_{t\in[0,T]}$ is deterministic
- the recovery process $(Z_t)_{t\in[0,T]}$ is deterministic
- the promised claim X_T is deterministic
- the probability space supports an exponentially distributed random varaiable

$$E : \Omega \to [0, \infty[, \qquad \mathbb{Q}(E \leq t) = 1 - e^{-t}, \;\; t \in [0, \infty[$$

Given a function

$$[0, T] \to]0, \infty[, \quad t \mapsto \lambda_t,$$

define the default time $\overline{\tau} : \Omega \to [0, T] \cup \{\infty\}$ by

$$\overline{\tau} := \inf\{t \in [0, T] \, : \, \int_0^t \lambda_s ds \geq E\}$$

With this definition, it holds that

$$\{\overline{\tau} \leq t\} = \{E \leq \int_0^t \lambda_s ds\} \quad \text{for all } t \in [0, T[\tag{9.8}$$

Indeed, to see $\subseteq$, suppose that $\omega \in \{\overline{\tau} \leq t\}$, hence

$$\inf\{u \in [0, T] \, : \, \int_0^u \lambda_s ds \geq E(\omega)\} \leq t$$

thus for each $\varepsilon > 0$ we have $\int_0^{t+\varepsilon} \lambda_s ds \geq E(\omega)$ which means that $\int_0^t \lambda_s ds \geq E(\omega)$. To see $\supseteq$, assume that $\int_0^t \lambda_s ds \geq E(\omega)$ hence

$$t \in \{u \in [0, T] \, : \, \int_0^u \lambda_s ds \geq E(\omega)\}$$

that is

$$t \geq \inf\{u \in [0, T] \, : \, \int_0^u \lambda_s ds \geq E(\omega)\} = \overline{\tau}(\omega)$$

Note that if $(\lambda_s)_{s\in[0,T]}$ is not constant, then $\overline{\tau}$ is not exponentially distributed on $[0, T]$, because

$$Q(\overline{\tau} \leq t) = Q(E \leq \int_0^t \lambda_s ds) = 1 - e^{-\int_0^t \lambda_s ds} \quad \text{for all } t \in [0, T[$$

The function $(\lambda_s)_{s\in[0,T]}$ is interpreted as time-varying default intensity due to

$$
\begin{aligned}
Q(\overline{\tau} \in]t, t+\Delta] | \overline{\tau} > t) &= \mathbb{Q}(\int_0^t \lambda_u du < E \leq \int_0^{t+\Delta} \lambda_u du \,|\, E > \int_0^t \lambda_u du) \\
&= \frac{\mathbb{Q}(\int_0^t \lambda_u du < E \leq \int_0^{t+\Delta} \lambda_u du)}{\mathbb{Q}(E > \int_0^t \lambda_u du)} \\
&= \frac{\mathbb{Q}(E \leq \int_0^{t+\Delta} \lambda_u du) - Q(E \leq \int_0^t \lambda_u du)}{\mathbb{Q}(E > \int_0^t \lambda_u du)} \\
&= \frac{(1 - e^{-\int_0^{t+\Delta} \lambda_u du}) - (1 - e^{-\int_0^t \lambda_u du})}{e^{-\int_0^t \lambda_u du}} \\
&= 1 - e^{-\int_t^{t+\Delta} \lambda_u du} \approx \int_t^{t+\Delta} \lambda_u du \approx \lambda_t \Delta
\end{aligned}
$$

Introduce the default-indicator process

$$H_t = 1_{\{\overline{\tau} \leq t\}} \qquad \text{for all } t \in [0, T],$$

obviously

$$H_t(\omega) = 1_{[\overline{\tau}(\omega), T]}$$

it starts at zero and jumps to one at default. Consider the filtration generated by the default indicator process $(H_t)_{t\in[0,T]}$

$$\mathcal{H}_t = \sigma(H_s \,:\, s \leq t), \quad t \in [0, T].$$

Essentially, the sigma-algebra $\mathcal{H}_t$ represents the information, whether the default occurred within $[0, t]$ and is generated by the events

$$\{\overline{\tau} \leq s\}, \qquad s \in [0, t]$$

If default occurred within $[0, t]$, then $\mathcal{H}_t$ contains also the information at which time it happened. More precisely, conditioned on $\mathcal{H}_t$ we have the following distribution of $\overline{\tau}$.

- in the case $\overline{\tau} \leq t$, the distribution is concentrated at the point $\overline{\tau}$
- in the case $\overline{\tau} > t$, the distribution is concentrated on $]t, T] \cup \{\infty\}$ and coincides with

$$
\begin{aligned}
Q(\overline{\tau} \leq s \,|\, \overline{\tau} > t) &= (1 - e^{-\int_t^s \lambda_u du}) 1_{]t,T]}(s) \quad \text{for all } s \in [0, T] \\
Q(\overline{\tau} = \infty \,|\, \overline{\tau} > t) &= e^{-\int_t^T \lambda_u du}
\end{aligned}
$$

Formally, we can write the conditional distribution using the notion δ_z of point measure, concentrated at point z as

$$
\begin{aligned}
Q(\overline{\tau} \in ds \,|\, \mathcal{H}_t) &= 1_{\{\overline{\tau} \leq t\}} \delta_{\overline{\tau}}(ds) + \\
&\quad 1_{\{\overline{\tau} > t\}} \left(\lambda_s e^{-\int_t^s \lambda_u du} 1_{]t,T]}(s) ds + e^{-\int_t^T \lambda_u du} \delta_\infty(ds) \right)
\end{aligned}
$$

In what follows, we are only interested in the pre-default situation and consider the distribution on the remaining time interval

$$
\begin{aligned}
(9.9)\qquad 1_{\{\overline{\tau}>t\}}\mathbb{Q}(\overline{\tau}\in ds\,|\,\mathcal{H}_t) &= 1_{\{\overline{\tau}>t\}}\lambda_s e^{-\int_t^s \lambda_u du}ds \qquad s\in]t,T] \\
(9.10)\qquad 1_{\{\overline{\tau}>t\}}\mathbb{Q}(\overline{\tau}=\infty\,|\,\mathcal{H}_t) &= 1_{\{\overline{\tau}>t\}}e^{-\int_t^T \lambda_s ds}
\end{aligned}
$$

For a credit-linked instrument with deterministic recovery process $(Z_t)_{t\in[0,T]}$ and deterministic terminal claim X_T, we obtain the fair instrument price V_t at time $t\in[0,T]$ as follows

In the case the entity defaulted, the recovery is received and invested in the savings account

$$
V_t 1_{\{\overline{\tau}\le t\}} = \frac{B_t}{B_{\overline{\tau}}} Z_{\overline{\tau}} 1_{\{\overline{\tau}\le t\}}
$$

this is not an interesting case, since we deal with pre-default valuation. In the case that no default occurred, then $V_t 1_{\{\overline{\tau}>t\}}$ is the fair price of the

recovery cash stream $Z_{\overline{\tau}} 1_{\{\overline{\tau}\le T\}}$ at $\overline{\tau}$

and of the

promised claim cash stream $X_T 1_{\{\overline{\tau}>T\}}$ at T

Let us calculate the pre-default value of both cash streams.

Recovery cash stream: Upon default, the recovery is received and invested in savings account until maturity, hence

$$
\frac{B_T}{B_{\overline{\tau}}} Z_{\overline{\tau}} 1_{\{\overline{\tau}>t\}} 1_{\{\overline{\tau}\le T\}} \quad \text{at maturity } T
$$

That is, the fair pre-default price of this cash stream is given by

$$
\begin{aligned}
1_{\{\overline{\tau}>t\}} B_t \mathbb{E}^{\mathbb{Q}}\left(\frac{\frac{B_T}{B_{\overline{\tau}}} Z_{\overline{\tau}} 1_{\{\overline{\tau}>t\}} 1_{\{\overline{\tau}\le T\}}}{B_T}\,|\,\mathcal{H}_t\right) &= \\
&= 1_{\{\overline{\tau}>t\}} \mathbb{E}^{\mathbb{Q}}\left(\frac{B_t}{B_{\overline{\tau}}} Z_{\overline{\tau}} 1_{\{t<\overline{\tau}\le T\}}\,|\,\mathcal{H}_t\right) \\
&= 1_{\{\overline{\tau}>t\}} \int_t^T \frac{B_t}{B_s} Z_s \mathbb{Q}(\overline{\tau}\in ds\,|\,\mathcal{H}_t) \\
&= 1_{\{\overline{\tau}>t\}} \int_t^T e^{-\int_t^s r_u du} Z_s \lambda_s e^{-\int_t^s \lambda_u du} ds \\
(9.11)\qquad &= 1_{\{\overline{\tau}>t\}} \int_t^T e^{-\int_t^s (r_u+\lambda_u) du} Z_s \lambda_s ds
\end{aligned}
$$

Note that the effect of possible default comes as an additional discount factor added to the short rate.

Promised claim cash stream. It occurs at maturity T in the case of no-default, hence its pre-default value at time $t \in [0,T]$ is given by

$$\begin{aligned}
1_{\{\overline{\tau}>t\}} B_t \mathbb{E}^{\mathbb{Q}}(\frac{X_T}{B_T} 1_{\{\overline{\tau}>T\}} \,|\, \mathcal{H}_t) &= 1_{\{\overline{\tau}>t\}} X_T e^{-\int_t^T r_s ds} \mathbb{E}^{\mathbb{Q}}(1_{\{\overline{\tau}>T\}} \,|\, \mathcal{H}_t) \\
&= 1_{\{\overline{\tau}>t\}} X_T e^{-\int_t^T r_s ds} Q(\overline{\tau} = \infty \,|\, \mathcal{H}_t) \\
&= 1_{\{\overline{\tau}>t\}} X_T e^{-\int_t^T r_s ds} e^{-\int_t^T \lambda_s ds} \\
&= 1_{\{\overline{\tau}>t\}} X_T e^{-\int_t^T (r_s+\lambda_s) ds}
\end{aligned}$$

Combine now (9.11) and (9.12) to write the fair pre-default value of the instrument as

$$V_t 1_{\{\overline{\tau}>t\}} = \left(\int_t^T e^{-\int_t^s (r_u+\lambda_u) du} Z_s \lambda_s ds + X_T e^{-\int_t^T (r_s+\lambda_s) ds} \right) 1_{\{\overline{\tau}>t\}} \tag{9.12}$$

Remark Observe that up to the factor $1_{\{\overline{\tau}>t\}}$, the right-hand side of the above equation yields fair price of an instrument which continuously pays at intensity $(Z_s\lambda_s)_{s\in[0,T]}$ in addition to the payoff X_T at maturity T in a market, where the short rate is given by $(r_s+\lambda_s)_{s\in[0,T]}$. This observations shows that pricing methodologies form fixed-income markets can be beneficially utilized to value credit-related instruments. In particular, we elaborate on the use of affine term structure short rate models in what follows.

9.8 Extension towards random intensities. The deterministic setting from above serves a starting point to extend the methodology towards

$$\begin{cases} \text{random short rate } (r_t)_{t\in[0,T]} \\ \text{random intensity } (\lambda_t)_{t\in[0,T]} \\ \text{random recovery } (Z_t)_{t\in[0,T]} \\ \text{random promised claim } X_T \end{cases} \tag{9.13}$$

The idea is to use several information structures.

- the information flow $(\mathcal{F}_t)_{t\in[0,T]}$ comes from the observation of the quantities (9.13),
- the information flow $(\mathcal{H}_t)_{t\in[0,T]}$ comes from the observation of the default indicator $(H_t = 1_{\{\overline{\tau}\le t\}})_{t\in[0,T]}$
- the information $(\mathcal{G}_t)_{t\in[0,T]}$ available to the agents is defined by the composition of both, $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{H}_t$ for all $t \in [0,T]$.

Given the information $\mathcal{F}_T$, all quantities from (9.13) behave such as they were deterministic. By appropriate construction of $\overline{\tau}$, the valuation under condition $\mathcal{H}_t \vee \mathcal{F}_T$ becomes exactly the same as in the deterministic setting. Having obtained $\mathbb{E}^{\mathbb{Q}}(\cdot \,|\, \mathcal{H}_t \vee \mathcal{F}_T)$, the conditioning on $\mathcal{G}_t$ follows the usual trick of double conditioning

$$\mathbb{E}^{\mathbb{Q}}(\cdot \,|\, \underbrace{\mathcal{H}_t \vee \mathcal{F}_t}_{\mathcal{G}_t}) = \mathbb{E}^{\mathbb{Q}}(\mathbb{E}^{\mathbb{Q}}(\cdot \,|\, \mathcal{H}_t \vee \mathcal{F}_T) \,|\, \underbrace{\mathcal{H}_t \vee \mathcal{F}_t}_{\mathcal{G}_t})$$

This information decomposition makes the intensity based model accessible by the tools from interest rate theory. However, the basic trick is the construction of random default time $\overline{\tau}$ driven by stochastic default intensity, which is referred to as the Cox's construction.

9.9 Cox's construction. Is natural way to extend the above methodology is to stochastic default intensities. The idea of Cox's construction is to introduce a separate random source which drives the default intensity (and possible other factors, like short rate evolution). Given the path of default intensity, the random default time is constructed as previously, from an exponentially distributed random variable which is independent of random source driving the default intensity. We now mimic the deterministic-intensity construction with appropriate changes.

- Consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{Q}, (\mathcal{F}_t)_{t\in[0,T]})$ where the measure $\mathbb{Q}$ is supposed to represent the spot martingale measure and $(\mathcal{F}_t)_{t\in[0,T]}$ stands for the available information (coming for the observation of the short rate and of factors responsible for the default intensity evolution)
- Assume that the short rate $(r_t)_{t\in[0,T]}$ is $(\mathcal{F}_t)_{t\in[0,T]}$-adapted.
- Suppose that the recovery process $(Z_t)_{t\in[0,T]}$ is $(\mathcal{F}_t)_{t\in[0,T]}$-adapted.
- Suppose that the intensity process

  $$(\lambda_t)_{t\in[0,T]} \text{ is positive-valued and } (\mathcal{F}_t)_{t\in[0,T]}\text{–adapted.}$$

- Assume that the probability space supports an exponentially distributed random variable

  $$E : \Omega \to]0, \infty[, \qquad \mathbb{Q}(E \leq t) = 1 - e^{-t},$$

  such that the following crucial assumption holds:

  $$E \text{ and } \mathcal{F}_T \text{ are independent} \tag{9.14}$$

As previously, define the default time $\overline{\tau} : \Omega \to [0, T] \cup \{\infty\}$ by

$$\overline{\tau} := \inf\{t \in [0, \infty[: \int_0^t \lambda_s ds \geq E\}$$

Introduce the default process $(H_t = 1_{\{\overline{\tau}\leq t\}})_{t\in[0,T]}$ which generates the filtration $(\mathcal{H}_t)_{t\in[0,T]}$

$$\mathcal{H}_t = \sigma(H_s \,:\, s \leq t), \quad t \in [0, T].$$

and suppose that the only available information $(\mathcal{G}_T)_{t\in[0,T]}$ comes from the original information $(\mathcal{F}_T)_{t\in[0,T]}$ and from the potential default observation. Thus, define the filtration $(\mathcal{G}_t)_{t\in[0,T]}$ as follows

$$\mathcal{G}_t = \mathcal{H}_t \vee \mathcal{F}_t \quad t \in [0, T].$$

With this definition, $\overline{\tau}$ becomes a stopping time within filtration $(\mathcal{G}_t)_{t\in[0,T]}$.

Although E and $\mathcal{F}_T$ are independent, the random time $\overline{\tau}$ and $\mathcal{F}_T$ are not independent. This is because

$$\overline{\tau} := \inf\{t \in [0,\infty[: \int_0^t \lambda_s ds \geq E\}$$

depends on $(\lambda_s)_{s\in[0,T]}$. However all events of the type

$$\{\overline{\tau} \leq u\} = \{E \leq \int_0^u \lambda_s ds\}, \qquad u \in [0,t]$$

depend on the realization of $(\lambda_s)_{s\in[0,t]}$ of the intensity until t only. This issue and the independence of E on $\mathcal{F}_T$ is the reason why any $\mathcal{H}_t$-measurable random variable conditioned on $\mathcal{F}_t$ gives the same result as conditioned on $\mathcal{F}_T$

$$\mathbb{E}(H \,|\, \mathcal{F}_t) = \mathbb{E}(H \,|\, \mathcal{F}_T), \qquad \text{for all } \mathcal{H}_t\text{-measurable } H,\ t \in [0,T]$$

This result can be written in the following way

$$\mathbb{E}(H \,|\, \mathcal{F}_t \vee \mathcal{F}_T) = \mathbb{E}(H \,|\, \mathcal{F}_t), \qquad \text{for each } \mathcal{H}_t\text{-measurable } H.$$

and is nothing but the assertion

$$\mathcal{H}_t \text{ and } \mathcal{F}_T \text{ are } \mathcal{F}_t\text{-conditionally independent}$$

The notion of conditional independence is symmetric, that is

$$\mathcal{F}_T \text{ and } \mathcal{H}_t \text{ are } \mathcal{F}_t\text{-conditionally independent}$$

which equivalent to

$$\mathbb{E}(F \,|\, \mathcal{F}_t \vee \mathcal{H}_t) = \mathbb{E}(F \,|\, \mathcal{F}_t), \qquad \text{for each } \mathcal{F}_T\text{-measurable } F,\ t \in [0,T] \tag{9.15}$$

We may interpret this result as follows: Knowledge whether the default occurred within $[0,t]$ and if so, at which time gives only an indirect conclusion on the behavior of the default intensity $(\lambda_s)_{s\in[0,t]}$ until t, but this information is already contained in $\mathcal{F}_t$, hence $\mathcal{H}_t$ can be dropped.

Now, let us turn to the calculation of the distribution of $\overline{\tau}$, conditioned on $\mathcal{F}_T \vee \mathcal{H}_t$. Given the information $\mathcal{F}_T$ the entire trajectory $(\lambda_t)_{t\in[0,T]}$ becomes visible. That is, by the same arguments as in the deterministic setting (9.9), (9.10), we derive the conditional distribution

$$1_{\{\overline{\tau}>t\}}\mathbb{Q}(\overline{\tau} \in ds \,|\, \mathcal{F}_T \vee \mathcal{H}_t) = 1_{\{\overline{\tau}>t\}}\lambda_s e^{-\int_t^s \lambda_u du} ds \qquad s \in]t,T] \tag{9.16}$$

$$1_{\{\overline{\tau}>t\}}\mathbb{Q}(\overline{\tau} = \infty \,|\, \mathcal{F}_T \vee \mathcal{H}_t) = 1_{\{\overline{\tau}>t\}} e^{-\int_t^T \lambda_s ds}. \tag{9.17}$$

Of course, in the post-default situation $\{\overline{\tau} \leq t\}$ the distribution shrinks to a point $\overline{\tau}(\omega) \in [0,t]$, but this is not the situation we are interested in.

Finally, let us calculate the pre-default price $(1_{\{\overline{\tau}>t\}}V_t)_{t\in[0,T]}$ of the credit-linked instrument with $(\mathcal{F}_t)_{t\in[0,T]}$–adapted recovery process $(Z_t)_{t\in[0,T]}$ and $\mathcal{F}_T$-measurable terminal claim X_T. We obtain the fair instrument price V_t at time $t\in[0,T]$ as follows

In the case the entity defaulted, the recovery is received and invested in the savings account

$$V_t 1_{\{\overline{\tau}\le t\}} = \frac{B_t}{B_{\overline{\tau}}} Z_{\overline{\tau}} 1_{\{\overline{\tau}\le t\}}$$

this is not an interesting case, since we deal with pre-default valuation. In the case that no default occurred, then $V_t 1_{\{\overline{\tau}>t\}}$ is the fair price of the

recovery cash stream $Z_{\overline{\tau}} 1_{\{\overline{\tau}\le T\}}$ at $\overline{\tau}$

and of the

promised claim cash stream $X_T 1_{\{\overline{\tau}>T\}}$ at T

Let us calculate the pre-default value of both cash streams.

Recovery cash stream: Upon default, the recovery is received and invested in savings account until maturity, hence

$$\frac{B_T}{B_{\overline{\tau}}} Z_{\overline{\tau}} 1_{\{\overline{\tau}>t\}} 1_{\{\overline{\tau}\le T\}} \quad \text{at maturity } T$$

That is, the fair pre-default price of this cash stream is given by

$$\begin{aligned} 1_{\{\overline{\tau}>t\}} B_t \mathbb{E}^{\mathbb{Q}} \left(\frac{\frac{B_T}{B_{\overline{\tau}}} Z_{\overline{\tau}} 1_{\{\overline{\tau}>t\}} 1_{\{\overline{\tau}\le T\}}}{B_T} \,|\, \mathcal{F}_T \vee \mathcal{H}_t \right) &= \\ = 1_{\{\overline{\tau}>t\}} \int_t^T e^{-\int_t^s (r_u+\lambda_u)du} Z_s \lambda_s ds & \qquad (9.18) \end{aligned}$$

Now, we need to condition this random variable on $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{H}_t$, this gives

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}} \left(1_{\{\overline{\tau}>t\}} \int_t^T e^{-\int_t^s (r_u+\lambda_u)du} Z_s \lambda_s ds \,|\, \mathcal{F}_t \vee \mathcal{H}_t \right) &= \\ = 1_{\{\overline{\tau}>t\}} \mathbb{E}^{\mathbb{Q}} \left(\underbrace{\int_t^T e^{-\int_t^s (r_u+\lambda_u)du} Z_s \lambda_s ds}_{\mathcal{F}_T-\text{measurable}} \,|\, \mathcal{F}_t \vee \cancel{\mathcal{H}_t} \right) & \\ = 1_{\{\overline{\tau}>t\}} \mathbb{E}^{\mathbb{Q}} \left(\int_t^T e^{-\int_t^s (r_u+\lambda_u)du} Z_s \lambda_s ds \,|\, \mathcal{F}_t \right) & \end{aligned}$$

Note that the condition $\mathcal{H}_t$ cancels out due to conditional independence (9.15).

Promised claim cash stream. It occurs at maturity T in the case of no-default, hence its pre-default value at time $t\in[0,T]$ is given by

$$1_{\{\overline{\tau}>t\}} B_t \mathbb{E}^{\mathbb{Q}} \left(\frac{X_T}{B_T} e^{-\int_t^T \lambda_s ds} 1_{\{\overline{\tau}>T\}} \,|\, \mathcal{F}_T \vee \mathcal{H}_t \right) \;=\; 1_{\{\overline{\tau}>t\}} X_T e^{-\int_t^T (r_s+\lambda_s)ds}$$

Now, we need to condition this random variable on $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{H}_t$, this gives

$$\begin{aligned} & \mathbb{E}^{\mathbb{Q}}(1_{\{\overline{\tau}>t\}} X_T e^{-\int_t^T (r_s+\lambda_s)ds} \mid \mathcal{F}_t \vee \mathcal{H}_t) \\ = & \; 1_{\{\overline{\tau}>t\}} \mathbb{E}^{\mathbb{Q}}(\underbrace{X_T e^{-\int_t^T (r_s+\lambda_s)ds}}_{\mathcal{F}_T-\text{measurable}} \mid \mathcal{F}_t \vee \cancel{\mathcal{H}_t}) \end{aligned}$$

Note that the condition $\mathcal{H}_t$ cancels out due to conditional independence (9.15). Combine now (9.18) and (9.19) to write the fair pre-default value of the instrument as

$$\begin{aligned} 1_{\{\overline{\tau}>t\}} V_t &= 1_{\{\overline{\tau}>t\}} \mathbb{E}^{\mathbb{Q}}(\int_t^T e^{-\int_t^s (r_u+\lambda_u)du} Z_s \lambda_s ds \mid \mathcal{F}_t) \\ & \quad + 1_{\{\overline{\tau}>t\}} \mathbb{E}^{\mathbb{Q}}(X_T e^{-\int_t^T (r_s+\lambda_s)ds} \mid \mathcal{F}_t) \end{aligned}$$

9.10 Zero bonds within Cox's setting. Consider the simplest case of credit-linked instrument: the zero bond with maturity T and zero recovery. According to the above results, its pre-default price is given by

$$\begin{aligned} 1_{\{\overline{\tau}>t\}} \overline{B}_t(T) &= 1_{\{\overline{\tau}>t\}} \mathbb{E}^{\mathbb{Q}}(\underbrace{X_T}_{=1} e^{-\int_t^T (r_s+\lambda_s)ds} \mid \mathcal{F}_t) \\ &= 1_{\{\overline{\tau}>t\}} \mathbb{E}^{\mathbb{Q}}(e^{-\int_t^T (r_s+\lambda_s)ds} \mid \mathcal{F}_t) \end{aligned}$$

Thus, under zero recovery the price of defaultable zero bond is similar to that of non-defaultable zero bond where instead of the short rate we put short rate plus a default intensity. Since default intensity is a positive-valued process, the possibility of default is reflected by an additional discount. More importantly is that we can now use the machinery of interest rate models (in the short rate formulation) to describe the entire term structure of defaultable bonds. All we need is to define another short rate process describing the default-adjusted short rate $(\lambda_t + r_t)_{t\in[0,T]}$ and to model so the price evolution prior to default of defaultable zero bonds. However, this seems to be applicable under zero recovery assumption only. Remember that modeling of the recovery was addressed in terms of three stylized procedures. In particular, an interesting rule which comes close to some market practice was the recovery of the market value, where the investors are reimbursed for the loss, caused by the price drop due to default. This recovery rule could not be considered since it it can not be given in terms of an exogenously given recovery process. Fortunately, it turns out that for zero bonds, the zero recovery with an adjusted default intensity is equivalent to the recovery of the market value.

Suppose that under the recovery of the market value, the investors receive a fraction $(1-q)$ (with $q \in [0,1]$) of the zero-bond price, which was listed immediately before default. With this assumption, the pre-default price process $(\overline{B}_t(\tau))_{t\in[0,\tau]}$ of defaultable zero bond would satisfy

$$\begin{aligned} 1_{\{\overline{\tau}>t\}} \overline{B}_t(T) &= 1_{\{\overline{\tau}>t\}} \mathbb{E}^{\mathbb{Q}}(\int_t^T e^{-\int_t^s (r_u+\lambda_u)du} \overline{B}_s(T)(1-q)\lambda_s ds \mid \mathcal{F}_t) \\ & \quad + 1_{\{\overline{\tau}>t\}} \mathbb{E}^{\mathbb{Q}}(e^{-\int_t^T (r_s+\lambda_s)ds} \mid \mathcal{F}_t) \end{aligned}$$

It seems difficult to determine a solution $(\overline{B}_t(T))_{t\in[0,T]}$ to this equation. However, the following intuition may help: If within the next short time period, one loses only a fraction, say 50% of the bond, could it be the same as if one loses 100% of the bond, but with half of the probability? Could it be that zero bond with zero recovery at default intensity $(q\lambda_t)_{t\in[0,T]}$ gives the same price as zero bond with recovery $(1-q)$ and default intensity $(\lambda_t)_{t\in[0,T]}$?

Let us verify whether it is true. For simplicity, let us assume that we are back in the setting where short rate and default intensity are deterministic. Then zero bond with zero recovery and intensity $(q\lambda_t)_{t\in[0,T]}$ has a pre-default price

$$\overline{B}_t^{(q)}(T) = e^{-\int_t^T (r_s+q\lambda_s)ds} \tag{9.19}$$

and all we need is to verify that $(\overline{B}_t^{(q)}(T))_{t\in[0,T]}$ satisfies

$$\overline{B}_t^{(q)}(T) \quad = \quad \int_t^T e^{-\int_t^s (r_u+\lambda_u)du}\overline{B}_s^{(q)}(T)(1-q)\lambda_s ds + e^{-\int_t^T (r_s+\lambda_s)ds} \tag{9.20}$$

By straight-forward calculation, we see that the right-hand side of (9.20)

$$\begin{aligned}
&\int_t^T e^{-\int_t^s (r_u+\lambda_u)du}\overline{B}_s^{(q)}(T)(1-q)\lambda_s ds + e^{-\int_t^T (r_s+\lambda_s)ds} = \\
&= \int_t^T e^{-\int_t^s (r_u+\lambda_u)du}\underbrace{e^{-\int_s^T (r_u+q\lambda_u)du}}_{\overline{B}_s^{(q)}(T)}(1-q)\lambda_s ds + e^{-\int_t^T (r_s+\lambda_s)ds} \\
&= \int_t^T e^{-\int_t^T (r_u+\lambda_u)du}e^{\int_s^T (1-q)\lambda_u du}(1-q)\lambda_s ds + e^{-\int_t^T (r_s+\lambda_s)ds} \\
&= e^{-\int_t^T (r_u+\lambda_u)du}\underbrace{\int_t^T e^{\int_s^T (1-q)\lambda_u du}(1-q)\lambda_s ds}_{-e^{\int_s^T (1-q)\lambda_u du}|_{s=t}^{s=T}} + e^{-\int_t^T (r_s+\lambda_s)ds} \\
&= e^{-\int_t^T (r_u+\lambda_u)du}\left(e^{\int_t^T (1-q)\lambda_u du} - 1\right) + e^{-\int_t^T (r_s+\lambda_s)ds} \\
&= e^{-\int_t^T (r_u+q\lambda_u)du} = \overline{B}_t^{(q)}(T)
\end{aligned}$$

equals to the left-hand side of (9.20) due to (9.19). In the framework of Cox's construction, this result generalizes to the situation where the short rate and the default intensity are stochastic processes. Again, the main technique is the double conditioning

$$\mathbb{E}^{\mathbb{Q}}(\cdot \mid \mathcal{H}_t \vee \mathcal{F}_t) = \mathbb{E}^{\mathbb{Q}}(\mathbb{E}^{\mathbb{Q}}(\cdot \mid \mathcal{H}_t \vee \mathcal{F}_T) \mid \mathcal{H}_t \vee \mathcal{F}_t)$$

combined with conditional independence arguments.

In any case, a useful outcome is that zero recovery under intensity $(q\lambda_t)_{t\in[0,T]}$ gives the same prices of zero bonds as if they were priced under the fractional recovery of the market value, paying a recovery of $(1-q)$ of the pre-default market price.

9.11 Intensity modelling with short rate technique. In the two-factor Gaussian model, the short rate and the default intensity are modeled by processes which follow the dynamics as in the Vasicek (actually, Hull-White) approach. Note that such processes are Gaussian, that is, both, the short rate $(r_t)_{t\in[0,T]}$ and the default intensity $(\lambda_t)_{t\in[0,T]}$ may become negative. Although the negativity of the short rate is not considered as a major problem, the negativity of the default intensity does not have any interpretation. Nevertheless, for appropriate parameters and for short times ahead such model may give useful insights.

The two-factor Gaussian model is specified as

$$\begin{aligned} dr_t &= (k(t) - ar_t)dt + \sigma(t)dW_t \\ d\lambda_t &= (\overline{k}(t) - \overline{a}\lambda_t)dt + \overline{\sigma}(t)d\overline{W}_t \end{aligned}$$

with Brownian motions $(W_t)_{t\in[0,T]}, (\overline{W}_t)_{t\in[0,T]}$ satisfying

$$\langle W_\cdot, \overline{W}_\cdot \rangle_t = \rho t \qquad \text{for all } t \in [0,T] \text{ with } \rho \in]-1,1[.$$

In this model, both processes $(r_t)_{t\in[0,T]}$ and $(\lambda_t)_{t\in[0,T]}$ are mean-reverting with time-changing levels $(k_t)_{t\in[0,T]}$ and $(\overline{k}_t)_{t\in[0,T]}$ respectively. The numbers a and $\overline{a}$ stand for the corresponding mean reversion speeds and $(\sigma(t))_{t\in[0,T]}$ and $(\overline{\sigma}(t))_{t\in[0,T]}$ describe the stochastic impacts. The parameter $\rho \in]-1,1[$ describes the dependence in changes of short rate and default intensity. Some authors argue that the correlation between interest rates and default intensity could be negative.

In the two-factor Gaussian model the default-free zero bond prices are given by

$$B_t(\tau) = e^{\mathcal{A}^\tau(t) - \mathcal{B}^\tau(t)r_t} \qquad 0 \le t \le \tau \le T$$

where

$$\begin{aligned} r_t &= r_0 e^{-at} + \int_0^t e^{-a(t-s)}k(s)ds + \int_0^t e^{-a(t-s)}\sigma(s)dW_s \\ \mathcal{B}^\tau(t) &= \frac{1 - e^{-a(\tau-t)}}{a} \\ \mathcal{A}^\tau(t) &= -\int_t^\tau \mathcal{B}^\tau(s)k(s)ds + \frac{1}{2}\int_t^\tau \mathcal{B}^\tau(s)^2\sigma(s)^2 ds \end{aligned}$$

The pre-default prices of defaultable zero bonds under zero recovery assumption are given by

$$\overline{B}_t(\tau) = e^{\tilde{\mathcal{A}}^\tau(t) - \tilde{\mathcal{B}}^\tau(t)\lambda_t} B_t(\tau) \qquad 0 \le t \le \tau \le T$$

where

$$
\begin{aligned}
\lambda_t &= \lambda_0 e^{-at} + \int_0^t e^{-\overline{a}(t-s)}\overline{k}(s)ds + \int_0^t e^{-\overline{a}(t-s)}\overline{\sigma}(s)d\overline{W}_s \\
\tilde{\mathcal{B}}^\tau(t) &= \frac{1-e^{-\overline{a}(\tau-t)}}{\overline{a}} \\
\tilde{\mathcal{A}}^\tau(t) &= -\int_t^\tau \tilde{\mathcal{B}}^\tau(s)\tilde{k}(s)ds + \frac{1}{2}\int_t^\tau \tilde{\mathcal{B}}^\tau(s)^2\overline{\sigma}(s)^2 ds \\
\tilde{k}(s) &= \overline{k}(s) - \rho\overline{\sigma}(s)\sigma(s)\mathcal{B}^\tau(s), \quad s \in [0,T]
\end{aligned}
$$

Let us prove five last equalities. According to the previous section, the pre-default price of defaultable bond is given by

$$
\begin{aligned}
\overline{B}_t(\tau) &= \mathbb{E}^{\mathbb{Q}}(e^{-\int_t^\tau (r_u+\lambda_u)du} \,|\, \mathcal{F}_t) \\
&= \mathbb{E}^{\mathbb{Q}}(\underbrace{e^{-\int_t^\tau r_u du}}_{B_t/B_\tau} e^{-\int_t^\tau \lambda_u du} \,|\, \mathcal{F}_t) \\
&= \frac{\mathbb{E}^{\mathbb{Q}}(\frac{B_\tau(\tau)}{B_\tau}\frac{B_0}{B_0(\tau)} e^{-\int_t^\tau \lambda_u du} \,|\, \mathcal{F}_t)}{\frac{B_t(\tau)}{B_t}\frac{B_0}{B_0(\tau)}} B_t(\tau) \\
&= \mathbb{E}^{\mathbb{Q}^\tau}(e^{-\int_t^\tau \lambda_u du} \,|\, \mathcal{F}_t) B_t(\tau)
\end{aligned}
$$

Now, represent the Brownian motion $(\overline{W}_t)_{t\in[0,T]}$ as

$$
\overline{W}_t = \rho W_t + \sqrt{1-\rho^2} W_t^\perp
$$

where $(W_t, W_t^\perp)_{t\in[0,T]}$ is a two dimensional Brownian motion. Remember that the discounted bond follows martingale dynamics

$$
d\left(\frac{B_t(\tau)}{B_t}\right) = \left(\frac{B_t(\tau)}{B_t}\right)(-\sigma(t)\mathcal{B}^\tau(t))dW_t
$$

and use the Girsanov theorem to conclude that

$$
\begin{bmatrix} W_t^\tau \\ W_t^{\perp,\tau} \end{bmatrix} = \begin{bmatrix} W_t \\ W_t^\perp \end{bmatrix} - \begin{bmatrix} \int_0^t (-\sigma(u)\mathcal{B}^\tau(u))du \\ 0 \end{bmatrix} \qquad t \in [0,T]
$$

follows a Brownian motion under $\mathbb{Q}^\tau$. With this, we write the dynamics of default intensity $(\lambda_t)_{t\in[0,T]}$ as

$$
\begin{aligned}
d\lambda_t &= (\overline{k}(t) - \overline{a}\lambda_t)dt + \overline{\sigma}(t)(\rho dW_t + \sqrt{1-\rho^2}dW_t^\perp) \\
&= (\overline{k}(t) - \overline{a}\lambda_t)dt + \overline{\sigma}(t)(\rho(dW_t^\tau - \sigma(t)\mathcal{B}^\tau(t)dt) + \sqrt{1-\rho^2}dW_t^{\perp,\tau}) \\
&= (\underbrace{\overline{k}(t) - \rho\overline{\sigma}(t)\sigma(t)\mathcal{B}^\tau(t)}_{\tilde{k}(t)} - \overline{a}\lambda_t)dt + \overline{\sigma}(t)(\rho dW_t^\tau + \sqrt{1-\rho^2}dW_t^{\perp,\tau})
\end{aligned}
$$

Note that this equation is driven by

$$
\rho W_t^\tau + \sqrt{1-\rho^2} W_t^{\perp,\tau} \qquad t \in [0,T]
$$

which follows a Brownian motion under the forward measure $\mathbb{Q}^{\tau}$. According to affine terms structure methodology, we have

$$\mathbb{E}^{\mathbb{Q}^{\tau}}(e^{-\int_t^{\tau} \lambda_u du} \,|\, \mathcal{F}_t) = e^{\tilde{\mathcal{A}}^{\tau}(t) - \tilde{\mathcal{B}}^{\tau}(t)\lambda_t}$$

9.12 Intensity modeling with HJM technique. The similarity between interest rate models and intensity based techniques in credit risk allows using also HJM approach. Recall that the default-free bonds are given by

$$B_t(\tau) = \mathbb{E}^{\mathbb{Q}}(e^{-\int_t^{\tau} r_s ds} \,|\, \mathcal{F}_t), \qquad 0 \le t \le \tau \le T$$

whereas within intensity based approach, the defaultable bond prices prior to default satisfy

$$\overline{B}_t(\tau) = \mathbb{E}^{\mathbb{Q}}(e^{-\int_t^{\tau} (r_s + \lambda_s) ds} \,|\, \mathcal{F}_t), \qquad 0 \le t \le \tau \le T$$

As we know from HJM approach $(r_t)_{t \in [0,T]}$ can be equivalently introduced as

$$r_t = f_t(t) \qquad t \in [0, T]$$

where the entire forward rate evolution $(f_t(\tau))_{t \in [0,\tau]}$, $\tau \in [0, T]$ is modeled by appropriate model of HJM type. The same applies to pre-default price evolution of defaultable bonds. Using appropriate HJM model, the default-adjusted short rate can be introduced as

$$r_t + \lambda_t = \overline{f}_t(t)$$

where the forward rate evolution $(\overline{f}_t(\tau))_{t \in [0,\tau]}$, $\tau \in [0, T]$ is modeled by another model of HJM type. Basically, this idea results in a parallel realization of two interest rate models of HJM type. One structure is supposed to describe risk-free zero bonds, the other intended to describe defaultable zero bonds. A realization of this concept could be given by the following stylized procedure.

Model default-free zero bonds

1) Specify initial forward rates $(f_0^*(\tau))_{\tau \in [0,T]}$ of the default-free zero bonds

$$f_0^*(\tau) = -\frac{\partial}{\partial \tau} \ln(B_0^*(\tau)), \qquad \tau \in [0, T]$$

2) Specify d-dimensional forward rate volatility

$$(\sigma_t(\tau))_{t \in [0,\tau]}, \quad \tau \in [0, T]$$

and calculate drift

$$\alpha_t(\tau) = \sigma_t(\tau) \int_t^{\tau} \sigma_t(s) ds \qquad 0 \le t \le \tau \le T$$

3) Establish the dynamics of forward rates

$$f_t(\tau) = f_0^*(\tau) + \int_0^\tau \alpha_s(\tau)ds + \int_0^\tau \sigma_s(\tau)dW_s$$

where $(W_t)_{t\in[0,T]}$ follows a d-dimensional Brownian motion under the spot martingale measure $\mathbb{Q}$

4) Determine risk-free bond prices

$$B_t(\tau) = e^{-\int_t^\tau f_t(s)ds}, \qquad 0 \le t \le \tau \le T.$$

Model defaultable bonds

$\overline{1}$) Specify initial forward rates $(\overline{f}_0^*(\tau))_{\tau\in[0,T]}$ of the defaultable zero bonds

$$\overline{f}_0^*(\tau) = -\frac{\partial}{\partial\tau}\ln(\overline{B}_0^*(\tau)), \qquad \tau \in [0,T]$$

$\overline{2}$) Specify d-dimensional forward rate volatility

$$(\overline{\sigma}_t(\tau))_{t\in[0,\tau]}, \quad \tau \in [0,T]$$

and calculate drift

$$\overline{\alpha}_t(\tau) = \overline{\sigma}_t(\tau)\int_t^\tau \overline{\sigma}_t(s)ds \qquad 0 \le t \le \tau \le T$$

$\overline{3}$) Establish the dynamics of forward rates

$$\overline{f}_t(\tau) = \overline{f}_0^*(\tau) + \int_0^\tau \overline{\alpha}_s(\tau)ds + \int_0^\tau \overline{\sigma}_s(\tau)dW_s$$

$\overline{4}$) Determine pre-default price evolution of risky bonds

$$\overline{B}_t(\tau) = e^{-\int_t^\tau \overline{f}_t(s)ds}, \qquad 0 \le t \le \tau \le T.$$

Now, the default time needs to be modeled

5) Interpret

$$\lambda_t = \overline{f}_t(t) - f_t(t) \qquad t \in [0,T]$$

as default intensity and apply Cox's construction to obtain default time $\overline{\tau}$.

In this framework, all bond-related credit risk derivatives can be priced. Let us sketch how a Monte-Carlo procedure could be implemented.

We present a recursive scheme which propagates in time, from t to the next time $t+\Delta$ with an appropriately chosen time step $\Delta > 0$, starting at $t = 0$. As a result, in each single run, we obtain a realization of both forward rates and of the default time

$$(f_t(\tau))_{0\le t\le\tau\le T}, \quad (\overline{f}_t(\tau))_{0\le t\le\tau\le T}, \quad \overline{\tau}$$

a) Determine the model ingredients

$$
\begin{array}{ccc}
(f_0^*(\tau))_{t\in[0,T]}, & (\sigma_t(\tau))_{0\le t\le\tau\le T}, & (\alpha_t(\tau))_{0\le t\le\tau\le T} \\
(\overline{f}_0^*(\tau))_{t\in[0,T]}, & (\overline{\sigma}_t(\tau))_{0\le t\le\tau\le T}, & (\overline{\alpha}_t(\tau))_{0\le t\le\tau\le T}
\end{array}
$$

b) Initialize default intensity $\lambda_0 = \overline{f}_0^*(0) - f_0^*(0)$ and the default count-down $\gamma(0) := 0$. Draw realization E of an exponentially distributed (with intensity one) random variable. Set $t := 0$

c) Set $\gamma(t+\Delta) := \gamma(t) + \Delta\cdot\lambda_t$, if $E \in]\gamma(t), \gamma(t+\Delta)]$ then set realization of default time $\overline{\tau} := t+\Delta$.

d) Draw realization w of a standard normally distributed random variable of dimension d and define

$$
\begin{aligned}
f_{t+\Delta}(\tau) &= f_t(\tau) + \alpha_t(\tau)\Delta + \sigma_t(\tau)\sqrt{\Delta}w \\
\overline{f}_{t+\Delta}(\tau) &= \overline{f}_t(\tau) + \overline{\alpha}_t(\tau)\Delta + \overline{\sigma}_t(\tau)\sqrt{\Delta}w
\end{aligned}
$$

for $\tau \in [t+\Delta, T]$ (within an appropriate discrete set of relevant bond maturities).

e) Set $\lambda_{t+\Delta} := \overline{f}_{t+\Delta}(t+\Delta) - f_{t+\Delta}(t+\Delta)$

f) Set $t \leftarrow t+\Delta$ if $t+\Delta > T$ then finish, otherwise go to step c).

After each run of this algorithm one obtains realizations

$$
(B_t(\tau)(\omega))_{t\in[0,\tau]}, \qquad (\overline{B}_t(\tau)(\omega))_{t\in[0,\tau]}, \;\; \tau\in[0,T] \qquad \overline{\tau}(\omega)
$$

Based on this data the final payoff $\Pi_T(\omega)$ of a credit-linked instrument must be calculated. (If the default occurs prior to maturity and results in recovery payment at the default time, then it must be equivalently transformed to terminal time T.) Averaging the terminal payoff by

$$
\widehat{\pi}_N := \frac{1}{N}\sum_{n=1}^{N}\frac{\Pi_T(\omega_n)}{B_T(\omega_n)}
$$

over a large number N of Monte-Carlo simulations, gives an approximation of the initial fair price of the credit-linked instrument. The empirical variance

$$
\widehat{s}_N^2 := \frac{1}{N-1}\sum_{n=1}^{N}\left(\frac{\Pi_T(\omega_n)}{B_T(\omega_n)} - \widehat{\pi}_N\right)^2
$$

can be used to view the confidence interval. As a rule of thumb, for large N the interval

$$
\left[\widehat{\pi}_N - 2\frac{\sqrt{\widehat{s}_N^2}}{\sqrt{N}}, \quad \widehat{\pi}_N + 2\frac{\sqrt{\widehat{s}_N^2}}{\sqrt{N}}\right]
$$

is considered to contain $\mathbb{E}^{\mathbb{Q}}(\Pi_T/B_T)$ with probability of 0.95.

9.13 Intensity modeling with multi-factor CIR technique. In this approach, one considers n independent processes $(X_t^i)_{t\in[0,T]}$, $i=1,\dots,n$ and describes the short rate dynamics $(r_t)_{t\in[0,T]}$ and the default intensity $(\lambda_t)_{t\in[0,T]}$ by linear transformations of the process $(X_t^1,\dots,X_t^n)_{t\in[0,T]}$

$$\begin{aligned} r_t &= \sum_{i=1}^n w^i X_t^i \qquad t\in[0,T] \\ \lambda_t &= \sum_{i=1}^n \overline{w}^i X_t^i \qquad t\in[0,T]. \end{aligned}$$

with appropriately chosen weights $(w^i)_{i=1}^n \in [0,\infty[^n$, $(\overline{w}^i)_{i=1}^n \in [0,\infty[^n$. Thereby, the factors follows a CIR processes

$$dX_t^i = (\alpha^i - \beta^i X_t^i)dt + \sigma^i\sqrt{X^i}dW_t^i, \quad i=1,\dots,n$$

each factor $(X_t^i)_{t\in[0,T]}$ is driven by the component $(W_t^i)_{t\in[0,T]}$ of the n–dimensional Brownian motion $(W_t)_{t\in[0,T]}$ with respect to the spot martingale measure $\mathbb{Q}$.

Since the weights $(w^i)_{i=1}^n$, $(\overline{w}^i)_{i=1}^n$ are non-negative, both, the short rate and the default intensity follow non-negative processes. However, note that with this choice, the correlation between short rates and default intensity is always positive

$$\mathrm{Cov}(r_t,\lambda_t) = \sum_{i,j=1}^n w^i w^j \mathrm{Cov}(X_t^i, X_t^j) = \sum_{i=1}^n (w^i)^2 \mathrm{Var}(X_t^i) > 0$$

Since each CIR-process

$$dX_t^i = (\alpha^i - \beta^i X_t^i)dt + \sigma^i\sqrt{X^i}dW_t^i,$$

yields affine term structure and the solution to the corresponding Riccatti equation is known, we have closed-form expression for

$$\mathbb{E}^{\mathbb{Q}}(e^{-\int_t^\tau cX_s^i ds} \,|\, \mathcal{F}_t) = H_1^i(\tau-t,c)e^{-H_2^i(\tau-t,c)cX_t^i}, \qquad 0\le t\le\tau\le T$$

with functions

$$\begin{aligned} H_1^i(u,c) &= \left[\frac{2\gamma^i e^{\frac{1}{2}(\gamma^i+\beta^i)u}}{(\gamma^i+\beta^i)(e^{\gamma^i u}-1)+2\gamma^i}\right]^{\frac{2\alpha^i}{(\sigma^i)^2}} \\ H_2^i(u,c) &= \frac{2(e^{\gamma^i}-1)}{(\gamma^i+\beta^i)(e^{\gamma^i u}-1)+2\gamma^i} \\ &\quad \text{where } \gamma^i = \sqrt{(\beta^i)^2 + 2c(\sigma^i)^2} \end{aligned}$$

Because of independence of the factors, the default-free bond prices follow

$$\begin{aligned} B_t(\tau) &= \mathbb{E}^{\mathbb{Q}}(e^{-\int_t^\tau r_s ds} \,|\, \mathcal{F}_t) = \mathbb{E}^{\mathbb{Q}}(\Pi_{i=1}^n e^{-\int_t^\tau w^i X_s^i ds} \,|\, \mathcal{F}_t) \\ &= \Pi_{i=1}^n H_1^i(\tau-t,w^i)e^{-H_2^i(\tau-t,w^i)w^i X_t^i}, \quad 0\le t\le\tau\le T \end{aligned}$$

and the pre-default prices of defaultable zero bonds are given as

$$\begin{aligned}\overline{B}_t(\tau) &= \mathbb{E}^{\mathbb{Q}}(e^{-\int_t^\tau r_s ds} \,|\, \mathcal{F}_t) = \mathbb{E}^{\mathbb{Q}}(\Pi_{i=1}^n e^{-\int_t^\tau \overline{w}^i X_s^i ds} \,|\, \mathcal{F}_t) \\ &= \Pi_{i=1}^n H_1^i(\tau - t, \overline{w}^i) e^{-H_2^i(\tau - t, \overline{w}^i)\overline{w}^i X_t^i}, \quad 0 \le t \le \tau \le T.\end{aligned}$$

9.14 Basket models. In credit risk management, some of the assets are issued linking to a synthetically generated pool of defaultable instruments rather than to a single-name entity. Such constructions are referred to as credit derivatives of 'basket' type.

Basket default swap (BDS) is a generalization of CDS, except that the underlying is a basket of assets rather than a single entity. Popular versions of BDS include

- first-to-default
- n-th to default
- all-to-default

For example, consider a first-to-default BDS

- The protection buyer pays a premium to the protection seller, regularly
- The default is defined as the first credit event in the basket, typically the default of one of the assets
- At default, the premium payments terminate and the protections seller has to pay out the protection buyer accordingly to the stipulations in the contract

Similarly, n-th-to default works as first-to-default but the event is triggered when the n-th default event in the basket occurs. Note that the premium needs to be paid until the n-th default in the basket.

BDS could be useful to the investors trying to protect more than one asset. BDSs could be main building blocks of more complex structured products such as CDOs.

9.15 The Kijima's and Moromachi's approach. Consider a basket of n credits within time horizon $[0, T]$ and denote by $\overline{\tau}_j$ the default time of the j-th reference entity, $j = 1, \ldots, n$. Thereby, we target a model where

$$\overline{\tau}_j : \Omega \to [0, T] \cup \{\infty\}, \quad j = 1, \ldots, n$$

are described on appropriate probability space $(\Omega, \mathcal{F}, \mathbb{Q}, (\mathcal{F}_t)_{t\in[0,T]})$ using Cox's construction. Denote by I the random variable

$$I(\omega) = \begin{cases} 0 & \text{if} \quad \min_{k=1}^n \overline{\tau}_k(\omega) > T \\ j & \text{if} \quad \overline{\tau}_j = \min_{k=1}^n \overline{\tau}_k(\omega) \le T \end{cases}$$

Here we implicitly assume (and ensure by later construction) that more that one default in the basket at the same time occurs with probability zero.

Now consider a contract with payoff

$$X = \sum_{j=0}^{n} S_j 1_{\{I=j\}} \qquad \text{at terminal date } T.$$

That is, if no default in the basket occurs, then the owner of the contract receives $S_0 \in \underline{\mathrm{R}}$. Otherwise, the contract payoff depends on which entity defaulted first. Say if the first defaulted entity is j, then the contract holder receives $S_j \in \underline{\mathrm{R}}$.

In practice, also other contracts are known. For instance a protection against two defaults. It pays

- S_0 if no default occurs
- S_j if only one company, the company j defaults
- $S_j + S_k$ if first two defaulted companies are j and k

In what follows, we focus only on the valuation of contingent claim

$$X = \sum_{j=0}^{n} S_j 1_{\{I=j\}} \qquad \text{at terminal date } T$$

in the context of Cox's construction of default times. Crucial element here is that the default times $\overline{\tau}_1, \dots, \overline{\tau}_n$ are not independent, since each $\overline{\tau}_j$ them is driven by its own default intensity $(\lambda_t^j)_{t\in[0,T]}$ and the default intensities $(\lambda_t^1)_{t\in[0,T]}, \dots (\lambda_t^n)_{t\in[0,T]}$ are dependent, in general. Such constructions exhibits two important features

- It is possible to describe a clustering of defaults. This is important to reflect real world situations, when market conditions (financial & economic crisis?) have an impact on survival of companies.
- Having observed realizations of all default intensities $(\lambda_t^j)_{t\in[0,T]}$, $j = 1, \dots, n$, the occurrence of default times becomes independent by construction. This yields appropriate tools which are needed in the valuation of derivatives.

Let us construct the default times $\overline{\tau}_1, \dots, \overline{\tau}_n$ using Cox's idea.

On the filtered probability space $(\Omega, \mathcal{F}, \mathbb{Q}, (\mathcal{F}_t)_{t\in[0,T]})$ define adapted positive-valued intensity processes

$$(\lambda_t^j)_{t\in[0,T]}, \qquad j = 1, \dots, n$$

and realize independent identically exponentially (with parameter one) distributed random variables $E_1, \dots, E_n$ such that

$$E_1, \dots, E_n, \quad \mathcal{F}_T \qquad \text{are independent}$$

which gives random times

$$\overline{\tau}_j = \inf\{t \in [0,T] \, : \, \int_0^t \lambda_s^j ds \geq E_j\}$$

Define the default indicators

$$H_t^j = 1_{\{\overline{\tau}_j \leq t\}}, \quad t \in [0,T]$$

and define the filtration, available to the agents as

$$\mathcal{G}_t = \mathcal{F}_t \vee \bigvee_{j=1}^n \sigma(H_s^j \, : \, s \in [0,t]).$$

With this construction, the times $\overline{\tau}_1, \dots, \overline{\tau}_n$ are $\mathcal{F}_T$-conditionally independent

$$\begin{aligned} & Q(\overline{\tau}_1 \leq t_1, \overline{\tau}_2 \leq t_2, \dots, \overline{\tau}_n \leq t_n \,|\, \mathcal{F}_T) = \\ & = \quad Q(\overline{\tau}_1 \leq t_1 \,|\, \mathcal{F}_T) \cdot \mathbb{Q}(\overline{\tau}_2 \leq t_2 \,|\, \mathcal{F}_T) \cdot \ldots \cdot \mathbb{Q}(\overline{\tau}_n \leq t_n \,|\, \mathcal{F}_T) \end{aligned}$$

and the $\mathcal{F}_T$-conditioned distribution of

$$\overline{\tau} = \min_{k=1}^n \overline{\tau}_k$$

follows from

$$\begin{aligned} Q(\overline{\tau} > t \,|\, \mathcal{F}_T) &= \mathbb{Q}(\overline{\tau}_1 > t_1, \overline{\tau}_2 > t_2, \dots, \overline{\tau}_n > t_n \,|\, \mathcal{F}_T) \\ &= Q(\overline{\tau}_1 > t_1 \,|\, \mathcal{F}_T) \cdot \ldots \cdot \mathbb{Q}(\overline{\tau}_n > t_n \,|\, \mathcal{F}_T) \\ &= e^{-\int_0^t \lambda_s^1 ds} \cdot \ldots \cdot e^{-\int_0^t \lambda_s^n ds} \\ &= e^{-\int_0^t \sum_{j=1}^n \lambda_s^j ds} \end{aligned}$$

This result

$$Q(\overline{\tau} > t \,|\, \mathcal{F}_T) = e^{-\int_0^t \lambda_s ds}, \quad \text{for } t \in [0,T] \text{ where } (\lambda_s = \sum_{j=1}^n \lambda_s^j)_{s \in [0,T]}$$

shows that the first-to-default time is of the same structure as the others. It is driven by an intensity $(\lambda_s)_{s\in[0,T]}$ which is the sum of all intensities of other default times.

Now, let us calculate

$$\begin{aligned} \mathbb{Q}(I = j \,|\, \mathcal{F}_T) &= \mathbb{Q}(\{\exists \, t \in [0,T] \, : \, \tau_j \leq t, \, \tau_i > t, j \neq i\} | \, \mathcal{F}_T) \\ &= \int_0^T \mathbb{Q}(\bigcap_{i \neq j} \{\tau_i > t\} \,|\, \mathcal{F}_T, \tau_j = t) Q(\tau_j \in dt, \mathcal{F}_T) \\ &= \int_0^T e^{-\int_0^t \sum_{k \neq j} \lambda_s^k ds} \lambda_t^j e^{-\int_0^t \lambda_s^j ds} dt \\ &= \int_0^T e^{-\int_0^t \sum_{k=1}^n \lambda_s^k ds} \lambda_t^j dt = \int_0^T e^{-\int_0^t \lambda_s ds} \lambda_t^j dt \end{aligned}$$

Similarly, if no default occurred, then

$$\mathbb{Q}(I=0\,|\,\mathcal{F}_T) \quad = \quad \mathbb{Q}(\overline{\tau}_1 > T, \ldots, \overline{\tau}_n > T\,|\,\mathcal{F}_T) = e^{-\int_0^T \sum_{k=1}^n \lambda_s^k ds} = e^{-\int_0^T \lambda_s ds}$$

With these results, we can claim

$$\mathbb{E}^{\mathbb{Q}}(X_T) = S_0 \mathbb{E}^{\mathbb{Q}}(e^{-\int_0^T \lambda_s ds}) + \sum_{j=1}^n S_j \int_0^T \mathbb{E}^{\mathbb{Q}}(\lambda_t^j e^{-\int_0^t \lambda_s ds}) dt \tag{9.21}$$

The proof follows by double conditioning argument

$$\mathbb{E}^{\mathbb{Q}}(\cdot) = \mathbb{E}^{\mathbb{Q}}(\mathbb{E}^{\mathbb{Q}}(\cdot\,|\,\mathcal{F}_T)).$$

Now, let us consider a more realistic situation. where the instrument pays

$$\begin{cases} S_j \in \underline{\mathrm{R}} \text{ if } \{I=j\} \text{ at first default time } \overline{\tau} \\ S_0 \in \underline{\mathrm{R}} \text{ if } \{I=0\} \text{ at maturity time } T \end{cases} \tag{9.22}$$

In this situation, one needs to consider interest rate effects, described by a model for risk-free zero bonds.

Kishima and Moromachi suggest to model the short rate $(r_t)_{t\in[0,T]}$ as an $(\mathcal{F}_t)_{t\in[0,T]}$–adapted process. With this, the $\mathcal{F}_T$–conditioned expectation of the discounted payoff in the case of default is calculated easily

$$\begin{aligned}
&\mathbb{E}^{\mathbb{Q}}(1_{\{I=j\}}\frac{B_T}{B_{\overline{\tau}}}/B_T\,|\,\mathcal{F}_T) \\
&= \int_0^T e^{-\int_0^t r_s ds}\mathbb{Q}(\bigcap_{i\neq j}\{\tau_i > t\}\,|\,\mathcal{F}_T, \tau_j = t) Q(\tau_j \in dt, \mathcal{F}_T) \\
&= \int_0^T e^{-\int_0^t (r_s+\lambda_s)ds}\lambda_t^j dt
\end{aligned}$$

if no default occurred, the $\mathcal{F}_T$–conditioned expectation is

$$\mathbb{E}^{\mathbb{Q}}(1_{\{I=0\}}/B_T\,|\,\mathcal{F}_T) = e^{-\int_0^T r_s ds}e^{-\int_0^T \lambda_s ds} = e^{-\int_0^T (r_s+\lambda_s)ds}$$

With this, under the above assumptions, the initial fair price of the instrument (9.22) is given by

$$S_0\mathbb{E}^{\mathbb{Q}}(e^{-\int_0^T (r_s+\lambda_s)ds}) + \sum_{j=1}^n S_j \int_0^T \mathbb{E}^{\mathbb{Q}}(\lambda_t^j e^{-\int_0^t (r_s+\lambda_s)ds})dt$$

10 Rating migrations

10.1 Rating migration models. Credit rating agencies, for instance *Standard and Poors* (S&P) or *Moody's Investor Service* (Moody's) assign ratings to the issuers of debt obligations. These ratings describe the credit worthiness of the obligors. S&P and Moody's list seven classes. The S&P classes are

AAA, AA, A, BBB, BB, B, CCC

The Moody's classes are

$$\text{Aaa, Aa, A, Baa, Ba, B, Caa}$$

Within each class, a sub-division may be given (for instance, Moody's split A into A^+, A, A^-). We shall focus on 8 classes, adding D as the class of defaulted companies.

Trust to rating agencies and their accuracy is under discussion now. Ratings are very essential to companies

- Downgrades in rating cause spreads (CDS) to increase
- Some of institutions are allowed to invest in high-quality assets only. A downgrade could mean that large amounts of assets must be sold, this may have consequences.

Most of the models assume that the rating class evolution of a company behaves like a Markov process. At each time (discrete or continuous, depending on the model), the rating can change which means that the transition to a new class occurs. Looking at historical data, produced by agencies, we may estimate the following transition probabilities: For

Initial rating	AAA	AA	A	BBB	BB	B	CCC/C	D	NR
AAA	88.53	7.70	0.46	0.09	0.09	0.00	0.00	0.00	3.15
AA	0.60	87.50	7.33	0.54	0.06	0.10	0.02	0.01	3.84
A	0.04	2.07	87.21	5.36	0.39	0.16	0.03	0.06	4.67
BBB	0.01	0.17	3.96	84.13	4.03	0.72	0.16	0.23	6.61
BB	0.02	0.05	0.21	5.32	75.62	7.15	0.78	1.00	9.84
B	0.00	0.05	0.16	0.28	5.92	73.00	3.96	4.57	12.05
CCC/C	0.00	0.00	0.24	0.36	1.02	11.74	47.38	25.59	13.67

Table 1.2.1 Global Average 1 Year Transition Rates, 1981-2007 (%) (Source: S&P)

example, of the obligors rated A at the beginning of a year:

- 2.07% were upgraded to AA at the end of the year,
- 87.21% were still rated A at the end of the year,
- 5.36% were downgraded to BBB at the end of the year,
- 0.06% were rated D, they had defaulted within the year.

On this account, credit migration analysis is essential in risk management. Credit migrations are inputs to prominent models in credit risk. Their dynamics must be considered when term structure of defautable zero bonds is modeled. A typical question here is, based on

- historical credit migration matrix
- observed initial (promised) yields of risky bonds (within each rating class)

to determine a model for risk neutral evolution of credit migrations, which explains the observed risky yields. In the next step, one prices credit-linked derivatives, based on this risk-neutral dynamics. One of the standard approaches to this area of questions is the so called JLT framework (named after Jarrow, Lando, Turnbull). Since this methodology is relatively involving in continuous-time setting, we present its ideas in the discrete-time situation.

10.2 The idea of JLT approach. There are $K \in \mathbb{N}$ rating classes. The set of rating classes is defined as

$$S = \{1, \dots, K\}$$

with the interpretation that 1 is the highest class whereas K stands for the state of default. The rating of the obligor at time $t \in \{0, 1, \dots, T\}$ is represented by an S-valued random variable X_t with the interpretation that under the objective probability measure, the process $(X_t)_{t\in[0,T]}$ follows a Markov chain defined by transition probabilities

$$P(X_{t+1} = s' \mid X_t = s) = p_{s,s'}, \qquad s, s' \in S$$

which are given (for instance, estimated from historical data). This defines the transition matrix as

$$\Xi = \begin{bmatrix} p_{1,1} & p_{1,2} & \dots & p_{1,K} \\ \vdots & \vdots & \vdots & \vdots \\ p_{K-1,1} & p_{K-1,2} & \dots & p_{K-1,K} \\ 0 & 0 & \dots & 1 \end{bmatrix}.$$

Furthermore, one assumes that there exists a measure $\mathbb{Q} \sim \mathbb{P}$ which represents the spot martingale measure. It is also assumed that with respect to $\mathbb{Q}$, the non-defaultable bond price evolution

$$(B_t(\tau))_{t=0}^{\tau}, \ \ \tau \in \{0, \dots, T\} \quad \text{is independent on } (X_t)_{t=0}^{T}$$

Hence the default time

$$\overline{\tau} = \inf\{t \in \{0, \dots, T\} \,:\, X_t = K\}$$

is also independent on riskless zero bonds $\{B_t(\tau) \,:\, 0 \le t \le \tau \le T\}$. A further assumption is that with respect to $\mathbb{Q}$, the credit rating migrations $(X_t)_{t=0}^{T}$ follows a Markov process, which can be time-inhomogeneous (that is, risk-neutral transition probabilities may change from one time point to another). We assume that an arbitrage-free dynamics of zero bonds is given, and the spot martingale measure $\mathbb{Q}$ is determined.

Let $\overline{\tau}$ be the random time (10.2) at which the default occurs. Under the assumption of 'recovery of treasury' we suppose that the fraction δ of the riskless bond with the maturity τ is used to reimburse the owner of the defaulted zero bond. That is, we consider risky zero coupon bonds maturing at τ with the following payoff

$$\text{either } 1_{\{\overline{\tau}>\tau\}} \text{ at } \tau \qquad \text{or} \qquad \delta 1_{\{\overline{\tau}\leq\tau\}} \text{ at } \tau$$

where $1-\delta \in [0,1]$ stands for the recovery rate. Under these assumptions, the pre-default of the risky bond is given by

$$\begin{aligned} \frac{\overline{B}_t(\tau)}{B_t} &= \mathbb{E}^{\mathbb{Q}}\Big(\frac{\delta 1_{\{\overline{\tau}\leq\tau\}} + 1_{\{\overline{\tau}>\tau\}}}{B_\tau} \,|\, \mathcal{F}_t\Big) \\ \overline{B}_t(\tau) &= \underbrace{B_t\mathbb{E}^{\mathbb{Q}}\Big(\frac{1}{B_\tau} \,|\, \mathcal{F}_t\Big)}_{B_t(\tau)} \left(\delta \underbrace{Q_t(\overline{\tau}\leq\tau)}_{1-Q_t(\overline{\tau}>\tau)} + Q_t(\overline{\tau}>\tau)\right) \\ &= B_t(\tau)\underbrace{(\delta + (1-\delta)Q_t(\overline{\tau}>\tau))}_{\text{spread}} \end{aligned} \tag{10.1}$$

Note that the price of the risky bond depends on the starting point X_0 of the migration process (this is the class the company is rated at the beginning) and on its risk-neutral transition probabilities. These transition probabilities need to be adjusted in such way, that the risky bond prices observed at the beginning for each rating class are explained by the model.

To satisfy this requirement, time changing transition probabilities must be assumed

$$Q(X_{t+1} = s' \,|\, X_t = s) = q^{t,t+1}_{s,s'}, \qquad s, s' \in \mathcal{S}$$

which gives the transition matrices as follows

$$\Upsilon^{t,t+1} = \begin{bmatrix} q^{t,t+1}_{1,1} & q^{t,t+1}_{1,2} & \dots & q^{t,t+1}_{1,K} \\ \vdots & \vdots & \vdots & \vdots \\ q^{t,t+1}_{K-1,1} & q^{t,t+1}_{K-1,2} & \dots & q^{t,t+1}_{K-1,K} \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

Note that because of the equivalence of risk-neutral and objective measure the matrices must be constructed in such way that a positive entry in $\Upsilon^{t,t+1}$ occurs only if the corresponding entry of Ξ is positive. The JLT approach proposes to introduce the risk-neutral transition matrices as

$$\Upsilon^{t-1,t} = \mathbf{1} + \Pi(t)(\Xi - \mathbf{1}), \qquad t = 1, \dots, T$$

with diagonal matrix

$$\Pi(t) = \begin{bmatrix} \pi_1(t) & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & \pi_K(t) \end{bmatrix}$$

whose entries $(\pi_i(t))_{i=1}^K$ are non-negative. To have an interpretation of this procedure, observe that the row $j \in \{1\dots, K\}$ of the matrix $\Upsilon^{t-1,t}$ will have the form

$$[\pi_j(t)p_{j,1}, \dots, \pi_j(t)p_{j,j-1}, \pi_j(t)(p_{j,j}-1)+1, \pi_j(t)p_{j,j+1}, \dots, \pi_j(t)p_{j,K}]$$

That is, risk neutral transitions are obtained by correction of objective transition probabilities in the sense that the probability of leaving the recent state changes, but if the jump to another state occurs, then the transitions to each new state is the same as under objective measure.

Problem Which conditions must be satisfied by $(\pi_i(t))_{i=1}^K \in [0, \infty[^K$ to ensure that the matrix $\Upsilon^{t-1,t}$ is a stochastic matrix?

Now, let us turn to the question how to chose $(\Pi(t))_{t=0}^{T-1}$ such that the observed risky bond prices are explained. According to (10.1) we have

$$\overline{B}_t(\tau) = B_t(\tau)\left(\delta + (1-\delta)Q_t(\overline{\tau} > \tau)\right)$$

which gives

$$\frac{B_t(\tau) - \overline{B}_t(\tau)}{B_t(\tau)} = 1 - \delta - (1-\delta)Q_t(\overline{\tau} > \tau) = (1-\delta)Q_t(\overline{\tau} \le \tau)$$

for all $0 \le t \le \tau \le T$. With this, we see that the observation of bond prices at time t yields an exact information

$$\frac{B_t(\tau) - \overline{B}_t(\tau)}{(1-\delta)B_t(\tau)} = Q_t(\overline{\tau} \le \tau) \quad \tau = t, t+1, \qquad , T$$

on the risk-neutral distribution of the default time $\overline{\tau}$, conditioned on the information available at time t. That is the initial bond curve $(\overline{B}_0^i(\tau))_{\tau=0}^T$ of risky bonds of the rating class i gives the distribution of default time, conditioned on the starting value i of the rating process

$$Q(\overline{\tau} \le \tau \,|\, X_0 = i) = \frac{B_0(\tau) - \overline{B}_0^i(\tau)}{(1-\delta)B_0(\tau)}, \qquad \tau = 0, 1, \dots, T$$

Since the default state K is absorbing (a defaulted company stays defaulted) the event $\{X_\tau = K\}$ is the same as that $(X_t)_{t=0}^T$ reaches the state K within $t = 0, \dots, \tau$. This is nothing but the event that the default occurs within $t = 0, \dots, \tau$ which is $\{\overline{\tau} \le \tau\}$. That is,

$$Q(\overline{\tau} \le \tau \,|\, X_0 = i) = Q(X_\tau = K \,|\, X_0 = i) = (\Upsilon^{0,1}\Upsilon^{1,2} \cdot, \dots, \cdot \Upsilon^{\tau-1,\tau})_{i,K}$$

which means that

$$\frac{B_0(\tau)-\overline{B}_0^i(\tau)}{(1-\delta)B_0(\tau)} = (\Upsilon^{0,1}\Upsilon^{1,2}\cdot,\ldots,\cdot\Upsilon^{\tau-1,\tau})_{i,K}$$

for all $i=1,\ldots,K$ and $\tau=0,\ldots,T$. On this account, the problem of calculation the matrices $\Pi^1,\ldots,\Pi^{T-1}$ reduces to the solution of

$$\frac{B_0(\tau)-\overline{B}_0^i(\tau)}{(1-\delta)B_0(\tau)} = (\underbrace{\mathbf{1}+\Pi(1)(\Xi-\mathbf{1})}_{\Upsilon^{0,1}}\underbrace{\mathbf{1}+\Pi(2)(\Xi-\mathbf{1})}_{\Upsilon^{1,2}}\cdot,\ldots,\cdot\underbrace{\mathbf{1}+\Pi(\tau)(\Xi-\mathbf{1})}_{\Upsilon^{\tau-1,\tau}})_{i,K}$$

for all $i=1,\ldots,K$ and $\tau=0,\ldots,T$. This problem can be solved recursively.

At the beginning, one considers bonds with maturity $\tau=1$ which gives

$$\frac{B_0(\tau)-\overline{B}_0^i(\tau)}{(1-\delta)B_0(\tau)} = (\mathbf{1}+\Pi(1)(\Xi-\mathbf{1}))_{i,K}, \qquad i=1,\ldots,K$$

Note that on the right-hand side, there are entries of the last column of the matrix $\mathbf{1}+\Pi(1)(\Xi-\mathbf{1})$. Writing these equalities as equation system, we obtain

$$\begin{bmatrix} \frac{B_0(1)-\overline{B}_0^1(1)}{(1-\delta)B_0(1)} \\ \vdots \\ \frac{B_0(1)-\overline{B}_0^{K-1}(1)}{(1-\delta)B_0(1)} \\ 1 \end{bmatrix} = \begin{bmatrix} \pi_1(1)p_{1,K} \\ \vdots \\ \pi_{K-1}(1)p_{K-1,K} \\ 1 \end{bmatrix}$$

The last row holds since

$$\frac{B_0(1)-\overline{B}_0^K(1)}{(1-\delta)B_0(1)} = 1, \qquad \pi_K(1)(1-1)+1=1$$

Now, proceed recursively. Having obtained $\Pi(1),\ldots,\Pi(\tau-1)$ the matrices $\Upsilon^{0,1},\ldots\Upsilon^{\tau-1,\tau-1}$ are determined and $\Pi(\tau)$ must be obtained from

$$\frac{B_0(\tau)-\overline{B}_0^i(\tau)}{(1-\delta)B_0(\tau)} = \Upsilon^{0,1}\Upsilon^{1,2}\cdot,\ldots,\cdot\Upsilon^{\tau-2,\tau-1}(\mathbf{1}+\Pi(\tau))(\Xi-\mathbf{1})$$

for $i=1,\ldots,K$. Writing these equations as a linear equations system, we obtain

$$\begin{bmatrix} \frac{B_0(\tau)-\overline{B}_0^1(\tau)}{(1-\delta)B_0(\tau)} \\ \vdots \\ \frac{B_0(\tau)-\overline{B}_0^{K-1}(\tau)}{(1-\delta)B_0(\tau)} \\ 1 \end{bmatrix} = \Upsilon^{0,1}\Upsilon^{1,2}\cdot,\ldots,\cdot\Upsilon^{\tau-2,\tau-1}\begin{bmatrix} \pi_1(\tau)p_{1,K} \\ \vdots \\ \pi_{K-1}(\tau)p_{K-1,K} \\ 1 \end{bmatrix}$$

which can be solved to obtain $(\pi_1(\tau),\ldots,\pi_{K-1}(\tau))$ and so $\Pi(\tau)$. Problems that may occur are

- solution does not exist

- solution yields a non-stochastic matrix $\Upsilon^{\tau-1,\tau}$
- stochastic matrix $\Upsilon^{\tau-1,\tau}$ has positive entries where Ξ is zero
- stochastic matrix $\Upsilon^{\tau-1,\tau}$ has zero entries where Ξ is positive

In this case, the model is not consistent with initial bond curves and can not be calibrated.