An accompaniment to a course on interest rate modeling: with discussion of Black-76, Vasicek and HJM models and a gentle introduction to the multivariate LIBOR Market Model

Vasily Nekrasov*

http://www.yetanotherquant.de

July 10, 2013

Abstract

The goal of this paper is to help the motivated students with their course on interest rate modeling and/or to help them to learn the LIBOR model by themselves. It implies the reader knows what forward rates, caps, and swap[tion]s are and has some knowledge of the quantitative finance: at least Ito Calculus, Black-Scholes-[Merton] Formula, Girsanov's Theorem (in one dimension is enough) and the risk-neutral pricing. This stuff is usually taught in the first course on continuous financial modeling and is relatively easy. The interest rate modeling is much more complicated. Still those, who carefully read the wonderful Steven Shreve's book can learn the short-rate models, change of numeraire and Heath-Jarrow-Morton framework. But not the multivariate LIBOR Model (though there is a short section on the one factor LIBOR Model and its relation to the HJM). However, the Bond/IR market is essentially multivariate and the LIBOR Model can be introduced independently. But I could not find any tutorial, which would suit me. So I decided to write my own. It concerns theory only and not the calibration and computational aspects, which are the issues for the future papers.

1 Typical problems of a quant student

If you read this paper you are probably attending an [advanced] course on interest-rate modeling. If so, you likely know some short rate models and maybe the HJM framework. Though formally not necessary for studying of the LIBOR Model, it is still very helpful to be aware of this stuff. That's why if you are not, I

*Many thanks to Dmitri Semenchenko, Dr. Frank Wittemann, Dr. Michael Busch, Dr. Michael Meyer, Dr. Thomas Mazzoni.

do recommend to read the respective chapters of Shreve's book first. Historical development was Black-76 → Short rate models → HJM → LIBOR Market Model → (LIBOR with jumps and stochastic volatility). So starting directly from the LIBOR would mean ignoring the historical roots. So we discuss the earlier models too trying to enlighten some principle ideas, which are frequently misunderstood by students.

1.1 The Black formula (Black-76)

The Black-Scholes-[Merton] formula was a breakthrough and still remains the main formula in quantitative finance. Black-76 is its modification for the *commodity(!)* market though it is far more frequently used to price the interest rate derivatives. Surprisingly, this formula often lacks attention in the lectures. For example, if the lecture notes are based on Brigo and Mercurio, it is introduced without derivation (and even motivation) and if they are based on Shreve, it might be ignored at all since Shreve derives it from LIBOR Model, which comes at the very end of the paragraph on the term structure models. We, in turn, will learn it thoroughly since the motivation behind the LIBOR is its consistence with Black-76 [for caplets].

In 1976 Fischer Black wrote his seminal "Pricing of commodity contracts". His assumptions were that the short rates are constant, the futures prices are lognormally distributed with known variance rate σ^2 and the expected change in futures prices is zero. Once again, the extension was intended for commodity market: Black, himself, did write about farmers and harvests but did not write anything about options on bonds, caps and swaptions.

But the market readily adopted his approach ignoring that e.g. bond prices and forward rates cannot be simultaneously lognormal, let alone that the assumption of the constant short rate can do for the grain market but is clearly absurd for the bond market. But in the 70th there were few models and if one has only a hammer at hand, everything around seems to be the nails. So the market felt fine using the Black-76 long before it was rigorously justified with LIBOR Model. Finally, as Black himself wrote in his another famous paper[1]: "In the end, a theory is accepted not because it is confirmed by conventional empirical tests, but because researches persuade one another that the theory is correct and relevant".

Black used the delta hedging to derive his formula (recall that by his time the risk neutral pricing was not yet discovered, so the delta hedging was an only tool to price options from no arbitrage assumption). He also appealed to the CAPM (Capital Asset Pricing Model) to motivate his assumptions empirically, stating that "if the covariance of the change in the futures price with the return on the market portfolio is zero, then the expected change in the futures price will

[1] Fischer Black, "Noise". The Journal of Finance, Volume 41, Issue 3 · July 1986

be zero" and then recalling an empirical study[2] confirming that this is [nearly] the case for the wheat, corn and soybean futures.
Now let us quickly discuss what forward and futures contracts are.

Definition 1.1. *A forward on an asset X is a contract at time t to sell X at future time T (maturity date) at the fixed price K (strike).*

Example 1.2. *Which strike makes the present value of the forward equal to zero at time t if the interest rate r is constant? (We denote such strike with $For_X(t,T)$ and call it the fair price).*

At time t you have no money. The price of the zero T-Bond is $P(t,T) = e^{-r(T-t)}$. Borrow $X(t)$ [at rate r] and buy one unit of X. At time T you have to pay back $\frac{X(t)}{P(t,T)}$ and can sell X for $X(T)$ where $X(T)$ is the spot price of X at time T. So from no arbitrage

$$For_X(t,T) \overbrace{=}^{\star} \frac{X(t)}{P(t,T)} \overbrace{=}^{\dagger} e^{r(T-t)}X(t) \tag{1}$$

The $\star$ in 1 holds even if the interest rate is not constant, all we need is a tradable T-Bond at time t. But for the $\dagger$ a constant interest rate is needed.
Forwards are the OTC contracts. Note that a forward, unlike an option, is obligatory for both counterparties.

There is a credit risk since one side can go bankrupt before T. Moreover, the contract must specify in details the quality and quantity of X by delivery. Futures on X solve both problems. Futures are traded on a [commodity] exchange and are standardized by the exchange. One can loosely consider the futures as forwards, whose fair prices are reset every day. Denote $Fut_X(t,T)$ the futures price on X at time t with delivery T. Clearly $Fut_X(T,T) = X(T)$. In order to enter the futures contract one does not need to pay $Fut_X(t,T)$, indeed it costs nothing to enter the futures contract (fair price). One still pays the safety margin but to the exchange, not to the counterparty. This is in no way the price of the contract but just a collateral to eliminate the default risk. If, for example $X(t_1,T) = 100$, $X(t_2,T) = 110$ and one went long in the futures contract at time t_1, his gain at time t_2 is 10, which is immedeately paid by the counterparty. If the counterparty cannot pay, the exchange pays from the safety margin and the contract gets closed. On maturity date T one pays the $X(t_1,T) = 100$ and receives a commodity. The sum of daily payments generated by this futures contract is

$$(X(t_2,T)-X(t_1,T))+(X(t_3,T)-X(t_2,T))+\ldots+(X(t_i,T)-X(t_{i-1},T))+\ldots$$
$$+(X(T,T)-X(t_{n-1},T) = X(T,T)-X(t_1,T) = X(T)-X(t_1,T) \tag{2}$$

[2]Dusak K., 1973 Futures trading and investor returns: An investigation of commodity market risk premiums, Journal of Political Economy 81, Nov./Dec., 1387-1406

Stop for a while and make sure you do understand the idea of a futures contract (check the Hull's book if you don't). Otherwise you will not understand 5. Note that normally there is a daily accrued interest on the account to(from) which $X(t_i, T) - X(t_{i-1}, T)$ flows on the i-th day, so 2 is *not* a profit(loss) of a futures contract unless the interest rate is equal to zero.

Now we turn back to the Black-76 formula. For simplicity denote $F(t) = Fut_X(t, T)$ (we can do so since T is arbitrary but fixed). Under the Black's assumptions(lognormality and expected zero change of the futures prices) we can write the dynamics of $F(t)$ as

$$dF(t) = \sigma F(t) dW(t) \tag{3}$$

At time t we consider a call option with a strike K and the maturity T on the futures contract with delivery at time S, $t < T < S$. For the price of this call option we write $C(t, F(t), r, T, S, K) = C(t, F)$ since r, T, S and K are arbitrary but fixed. By Ito Formula we obtain

$$\begin{aligned} dC(t, F(t)) &= C_t dt + C_F dF(t) + \frac{1}{2} C_{FF} d^2 F(t) \\ &= C_t dt + \sigma C_F F(t) dW(t) + \frac{1}{2} C_{FF} \sigma^2 F^2(t) dt \end{aligned} \tag{4}$$

We can eliminate the random factor $\sigma C_F F(t) dW(t)$ from 4 if we go short in C_F futures contract at time t. Following Black, we denote the initial value of the futures contract with $u(t)$. It follows that $u(t) = 0$ since it costs nothing to enter the contract with futures price $F(t)$. But obviously as time goes to $t + dt$, a futures contract generates a cashflow $F(t + dt) - F(t)$, i.e.

$$du(t) = dF(t) = \sigma F(t) dW(t) \tag{5}$$

So a portfolio $C(t, F) - u(t) C_F$, i.e. with a call long and C_F futures short is *riskless* (no dW factor) and evolves according to

$$d\left(C(t, F) - u(t) C_F\right) = C_t dt + \frac{1}{2} C_{FF} \sigma^2 F^2(t) dt = r\left(C(t, F) - u(t) C_F\right) \tag{6}$$

The right-hand part is due to no-arbitrage: every riskless portfolio Π must grow according to $d\Pi = r\Pi dt$. But $u(t) = 0$ and finally we get

$$C_t dt + \frac{1}{2} C_{FF} \sigma^2 F^2(t) dt = rC(t, F) \tag{7}$$

The solution of this equation is the famous Black-76 Formula

$$C(t, F(t), r, T, S, K) = F(t) e^{-r(T-t)} \mathcal{N}(d_1) - K e^{-r(T-t)} \mathcal{N}(d_2) \tag{8}$$

where

$$d_1 = \frac{1}{\sigma\sqrt{T}} \left[\ln\left(\frac{F(t)}{K}\right) + \frac{1}{2}\sigma^2 T \right]$$

$$d_2 = d_1 - \sigma\sqrt{T} = \frac{1}{\sigma\sqrt{T}}\left[\ln\left(\frac{F(t)}{K}\right) - \frac{1}{2}\sigma^2 T\right]$$

It is a useful exercise to compare the derivation of Black-76 with BS-Formula[3].

Note that there is a dependence on S (futures delivery date) only in term $F(t) := Fut_X(t, S)$ (price at time t of the futures with delivery date S). It should be clear, since what matters is the difference $F(T) - K$. An option holder may even not enter the futures contract at T but just take a cash settlement $[F(T) - K]^+$ (recall, it costs nothing to enter the futures contract).
In particular, $F(S) := Fut_X(S, S) = X(S)$ so setting $S = T$ in 8 we obtain the price of the option on commodity itself (and not on commodity futures).

Black also derived from 8 the price of the forward contract with arbitrary strike K, using the terminal condition $v(X, T, K, T) = X(T) - K$, where $v(X, t, K, T)$ is the value at time t of the forward contract on asset X with the strike K and delivery date T. Black obtained

$$v(X, t, K, T) = (Fut_X(t, T) - K)e^{r(T-t)} \tag{9}$$

Setting $K = For_X(t, T)$ yields $v(X, t, K, T) = 0$ and thus

$$For_X(t, T) = Fut_X(t, T) \tag{10}$$

so can one use the Black-76 for options on forward as well?!

In fact 10 holds only if the interest rate is deterministic!
You may read about the forward-future spread (and the definition of futures price as an expectation w.r.t. the martingale measure) in Shreve's book. But for the rest of this paper it is not necessary since even if 10 would always hold, the assumption of a *constant* interest rate *and lognormal* forward rates is itself clearly absurd.
Does it invalidate the Black Formula for caplets? Obviously not necessarily, it only shows that one cannot derive it from current assumptions. So in the LIBOR Model the assumptions are slightly modified: the forwards are still lognormal but each under its T_i-forward measure. By change of measure(and numeraire) we make the interest rate irrelevant. So if you are interested only in LIBOR model jump to 1.4. But if you would like to understand why the LIBOR model is good (besides that it makes Black-76 valid) and probably find answers to some puzzles of short-rate models then read the whole stuff.

[3] http://en.wikipedia.org/wiki/Black-Scholes (Section The BlackScholes equation).

1.2 Short-rate Models and the martingale measure

Short rate is an instantaneous time dependent interest rate r_t so that the bank account (with initial investment of 1€) grows according to

$$B(t) = \exp\left(\int_0^t r(s)ds\right) \tag{11}$$

Respectively, by the risk neutral pricing, the price of a T-Bond at time T is

$$P(t,T) = \mathbb{E}_Q\left[\frac{B_t}{B_T}|\mathcal{F}_t\right] = \mathbb{E}_Q\left[\exp\left(-\int_t^T r(s)ds\right)|\mathcal{F}_t\right] \tag{12}$$

where Q is the martingale measure. r_t can be (and really is) stochastic but at time t we know how our bank account will change in an infinitesimal timespan. On the other hand the respective change of a T-Bond price is unpredictable since it depends not only on r_t but also on future short rates up to T. We can rewrite 11 as $dB(t) = rBdt$, so $B(t)$ is differentiable and thus "less random" than a T-Bond. We can define a *discounted* T-Bond by

$$Z(t,T) = D(t)P(t,T) = P(t,T)/B(t) \tag{13}$$

The dynamics of $Z(t,T)$ must be a martingale under Q. But the dynamics of r_t need not since bonds are tradable and r_t is not.
The first short rate model is due to Vasicek (1977)[4]

$$dr(t) = (a - br_t)dt + \sigma dW_t \tag{14}$$

In modern books the model is often introduced directly under the risk neutral measure Q (so that in 14 actually stays dW_t^Q) and a student with a stock price modeling background gets puzzled: in stock market we observed a stock (a tradable asset) under the real measure than killed the drift to make it martingale. As you [should] remember, there is no arbitrage iff there is a martingale measure (so that the discounted prices of all tradable assets are martingales). And the market is complete iff the martingale measure is unique. But how to proceed with r_t, which is untradable and thus need not to be a martingale?! Moreover, it is unobservable even under the real measure[5]). Finally, does Q exist at all and if it does, is it unique?!

We will follow the original Vasicek's paper (which is far more often cited than

[4]Indeed, Vasicek started with much more general settings and considered 14 just as a special case. Merton's Model (1973) was actually published earlier but still Vasicek is somehow considered as the first, well, at least in the most of textbooks.

[5]Garatek et al (p. xiii) say that the overnight rate is a good proxy for the short rate but Filipovic (p.10) argues it is not "because the motives and needs driving overnight borrowers are very different from those of borrowers who want money for a month or more".

read)[6] and see how Vasicek have introduced the market price of risk. There is no martingale measures in his paper since the risk neutral approach was introduced several years later by Harrision, Kreps and Pliska. However, as soon as we have the market price of risk λ_t we can introduce the risk neutral measure by Girsanov's theorem and turn back to "more modern" risk neutral approach.

In a short-rate model the bond price depends on t, r_t and T so we write $P(t,T) = f(t, r_t, T)$ [7], whereas T is arbitrary but fixed. By Ito formula the dynamics of the bond price is

$$df = f_{(t)}dt + f_{(r)}dr_t + \frac{1}{2}f_{(rr)}d^2r_t \overbrace{=}^{(14)} (f_{(t)} + [a - br_t]f_{(r)} + \frac{\sigma^2}{2}f_{(rr)})dt + \sigma f_{(r)}dW_t \tag{15}$$

which we can rewrite as

$$df(t, r_t, T) = f(t, r_t, T)\,[\mu(t, r_t, T)dt + \nu(t, r_t, T)dW_t] \tag{16}$$

where

$$\mu(t, r_t, T) = \frac{f_{(t)} + [a - br_t]f_{(r)} + \frac{\sigma^2}{2}f_{(rr)}}{f(t, r_t, T)}$$

$$\nu(t, r_t, T) = \frac{\sigma f_{(r)}}{f(t, r_t, T)}$$

The formula 16 resembles us the old good lognormal dynamics of a stock price, where $\mu(t, r_t, T)$ and $\nu(t, r_t, T)$ are respectively the drift and the volatility coefficients. (A reader, who studied the multivariate stock market model might anticipate that we are going to talk about the market price of risk $[\mu(t, r_t, T) - r_t]/\nu(t, r_t, T)$)

Now consider an investor who at time t issues an amount U_1 of a bond with maturity date T_1 and simultaneously buys an amount U_2 of a bond maturing at time T_2. The total value of this portfolio is $U = U_2 - U_1$ and it evolves according to

$$dU = [U_2\mu(t, r_t, T_2) - U_1\mu(t, r_t, T_1)]dt + [U_2\nu(t, r_t, T_2) - U_1\nu(t, r_t, T_1)]dW_t \tag{17}$$

We can eliminate the stochastic risk factor in 17 choosing

$$U_1 = U\frac{\nu(t, r_t, T_2)}{\nu(t, r_t, T_1) - \nu(t, r_t, T_2)} \qquad U_2 = U + U_1 = U\frac{\nu(t, r_t, T_1)}{\nu(t, r_t, T_1) - \nu(t, r_t, T_2)}$$

Such portfolio is riskless (there is only the drift term $[U_2\mu(t, r_t, T_2) - U_1\mu(t, r_t, T_1)]dt$) and thus by no arbitrage assumption (cp. with 6)

$$U\frac{[\mu(t, r_t, T_2)\nu(t, r_t, T_1) - \mu(t, r_t, T_1)\nu(t, r_t, T_2)]}{\nu(t, r_t, T_1) - \nu(t, r_t, T_2)}dt = r_tUdt \tag{18}$$

[6] http://www.wilmott.com/messageview.cfm?catid=4&threadid=85948

[7] This holds only if r_t is a Markov process (in our case it is), otherwise $P(t,T)$ may additionally depend on r_s for all $s \in [0, t]$

or equivalently

$$\frac{\mu(t, r_t, T_2) - r_t}{\nu(t, r_t, T_2)} = \frac{\mu(t, r_t, T_1) - r_t}{\nu(t, r_t, T_1)} =: \lambda_t \tag{19}$$

where λ_t can be interpreted as the market price of risk. Note that it must be the same for all bonds and thus does note depend on bond maturity (but of course may depend on t and r_t). But what risk do we actually mean? Well, as we stated before, the bank account is "less random" than a bond, so we expect an excess return on a bond. The longer is the maturity of a bond the larger is its volatily $\nu(t, r_t, T)$ and respectively its expected excess return $\mu(t, r_t, T) - r_t$ but their *ratio* does not varies with T.
Writing 19 as

$$\mu(t, r_t, T) - r_t = \lambda_t \nu(t, r_t, T) \tag{20}$$

and substituting for $\mu(t, r_t, T)$ and $\nu(t, r_t, T)$ from 16 we get the *term structure equation*

$$f_{(t)} + [a - \lambda_t \sigma - b r_t] f_{(r)} + \frac{\sigma^2}{2} f_{(rr)} - r_t f = 0 \tag{21}$$

with the terminal condition

$$f(T, r_T, T) = 1$$

The term structure equation has a solution[8] so the model at least does not contradict the no arbitrage assumption which we engaged before. Respectively, a martingale measure exists.

On the other hand the Girsanov's theorem tells us that $dW_t^Q = \theta_t dt + dW_t$ for some process θ_t and it only remains to show that indeed $\theta_t = \lambda_t$. Rewriting 14 under Q we obtain

$$dr(t) = (a - b r_t - \sigma \theta_t) dt + \sigma dW_t^Q \tag{22}$$

Writing as before $f = f(t, r_t, T)$ for the price of a T-Bond at time t we obtain by the Ito product rule $d[D_t f] = -r_t D_t f + D_t df$ so

$$d[D_t f] = -r_t D_t f dt + D_t \left[f_{(t)} + f_{(r)}(a - b r_t - \sigma \theta_t) + \frac{\sigma}{2} f_{(rr)} \right] dt + \sigma f_{(r)} D_t dW_t^Q \tag{23}$$

Since the discounted price is a martingale the drift (dt-term) must be zero. Comparing 23 with 21 we conclude that $\theta_t = \lambda_t$ and assign $\kappa_t := a - \lambda_t \sigma$

So we showed why and how we can model directly under the risk neutral measure. However, the model does not imply any restrictions from which follows that λ_t must be unique. Within the model is just required that λ_t must be the same for all bonds but we can arbitrarily choose it. For different λ_t we will get different bond prices, all of them arbitrage free. But having only model at hand

[8] see Shreve's book how to use the Ansatz $f(t, r_t, T) = \exp(-r_t C(t, T) - A(t, T))$

we cannot choose the "right" λ_t. Thus the model as such is not complete. What you need to understand now is that the *market players, not a model*, define the martingale measure. And if the market players have the same information and the same risk preference the market will be complete. Of course it is not so in practice but we *like* assuming it because if we do, the pricing of *every* bounded contingent claim is, in principle, possible. So we assume that the market is complete and moreover, that the market price of risk is constant, i.e. $\lambda_t \equiv \lambda$. Next we realize that bonds were traded long before the quantitative models appeared :). Since the term structure equation gives us the [closed form] solution for the bond price, we can *calibrate* the model and choose the parameters (κ, b, σ) to fit the bond prices (and usually some liquid derivatives like caps and swaptions) as good as possible. Then we can use the model to price any other derivatives. Note that we can find κ but not λ directly, since the dynamics of the short rate is unobservable. This is the difference with e.g. the Black-Scholes model, in which we can [at least theoretically] find the drift μ and thus the market price of risk $\frac{\mu - r}{\sigma}$.

Note that the assumption of the constant parameters is too restrictive, so the Vasicek model generally cannot perfectly fit the current term structure. It is the difference between the equilibrium and the arbitrage free models. The former usually assume some equilibrium in demand and supply(mean reversion is a kind of such equilibrium) and generate the term structure as an output. The latter can be perfectly matched to the current term structure (but not necessarily to the current caps and swaptions prices). Vasicek, CIR are the equilibrium models and Ho Lee, Hull White, Black Karasinski, G2 are the arbitrage free models. Such fitting is achieved via a time dependent parameter, e.g. in Hull White Model κ is time dependent:

$$dr(t) = (\kappa(t) - br_t)dt + \sigma dW_t^Q \tag{24}$$

Recall that in the Vasicek Model κ is constant.

1.3 Some aspects of Heath-Jarrow-Morton

The [one factor] HJM Model is well discussed in Shreve's book and it would have made little sense to blueprint it here. Instead, we discuss some aspects which may help you to go on with Shreve.

1. Motivation. HJM models not the short rate but rather the whole forward curve. The model is complete and the drift under the risk neutral measure turns out to be fully determined by the volatility. See Shreve for more details.

2. Instantaneous forward rate. If the interest is continiously compound, the forward rate $f(t, T, T + \delta)$ is defined from

$$P(t, T) = e^{\delta f(t,T,T+\delta)} P(t, T + \delta) \quad \Leftrightarrow \quad f(t, T, T + \delta) = \frac{1}{\delta} \ln \left(\frac{P(t, T)}{P(t, T + \delta)} \right)$$

Letting $\delta \to 0$ we obtain

$$\lim_{\delta \to 0} \frac{\ln(P(t,T)) - \ln(P(t,T+\delta))}{\delta} = -\frac{\partial \ln(P(t,T))}{\partial T} =: f(t,T) \qquad (25)$$

Setting $T = t$ we get the instantaneous short rate $R(t) = f(t,t)$. Note that the instantaneous forward rate is unobservable too.

3. It turns out that the affine short rate models, i.e. those for which $P(t, r_t, T) = \exp(-r_t C(t,T) - A(t,T))$ can be embedded in HJM via the drift term condition. See Shreve for more details.

4. In Shreve's book one confronts the following expression

$$d\left(-\int_t^T f(t,s)ds\right) = f(t,t)dt - \int_t^T df(t,s)ds$$

If you need to check it but already forgot (as I did) the rule how to differentiate such integrals, you may proceed as follows

$$\frac{\partial}{dt}\left[-\int_t^T f(t,s)ds\right] = \frac{\partial}{dt}\left[\int_T^t f(t,s)ds\right] = \lim_{x\to 0}\left[\int_T^{t+x} f(t+x,s)ds - \int_T^t f(t,s)ds\right] \Big/ x$$

$$= \lim_{x\to 0} \frac{\left[\left(\int_T^{t+x} f(t+x,s)ds - \int_T^t f(t+x,s)ds\right) + \left(\int_T^t f(t+x,s)ds - \int_T^t f(t,s)ds\right)\right]}{x}$$

$$= \lim_{x\to 0} f(t+x,t) + \int_T^t \lim_{x\to 0} \frac{[f(t+x,s)ds - f(t,s)ds]}{x} = f(t,t) + \int_T^t \frac{\partial}{dt}[f(t,s)ds]$$

Here we have implicitely assumed that $f(t,T)$ is sufficiently good, so that we can exchange limit and integral.

5. It is impossible to make the instantaneous forward rate lognormal since in this case it explodes. See Shreve for more details. However, the simply compounded forward rates may be made lognormal. Shreve introduces one factor LIBOR model via HJM and I recommend to have a look on it.

1.4 Change to T-Forward measure, Radon Nykodim process

The short-rate models are usually a piece of cake (as long as you understand that a market, not a model defines a martingale measure Q, and if we calibrate a model w.r.t. the current term structure and a set of sufficiently liquid instruments - usually Caps or Swaptions - we get the short rate dynamics under

Q). The HJM is usually managable too, at least Shreve makes it so. But it is the change of numeraire / T-Forward measures what makes troubles. Often a T-Forward measure Q^T is introduced as

$$dQ^T := \frac{1}{P(0,T)B_T} dQ \tag{26}$$

which is actually

$$dQ^T := \frac{P(T,T)}{P(0,T)B_T} dQ \tag{27}$$

It is easy to prove that it is indeed a probability measure. (Can *you* do it? If not, don't worry, we will do it later). dQ^T/dQ is a Radon-Nikodym derivative from $t = 0$ and *with* T *in mind.* But we need a Radon-Nikodym (a.k.a. likelihood) process for *all* $t \in [0, T]$. It is

$$\eta_t := \frac{P(t,T)}{P(0,T)B_t} \tag{28}$$

but unfortenately it often appears just like rabbit out of the hat. Soon we will consider its derivation in detail but so far just note the term $P(t, T)$ in the numerator. So you can guess for what I state that 26 is actually 27.
As soon as we have η_t we can find at time t the price $\pi(X_T)$ of a contingent claim X_T using the T-Forward measure since

$$\begin{aligned} \mathbb{E}_{Q^T}\left[X_T|\mathcal{F}_t\right] &= \frac{1}{\eta_t}\mathbb{E}_Q\left[X_T\eta_T|\mathcal{F}_t\right] = \frac{B(t)P(0,T)}{P(t,T)}\mathbb{E}_Q\left[X_T\frac{P(T,T)}{B(T)P(0,T)}\Big|\mathcal{F}_t\right] \\ &= \frac{1}{P(t,T)}\mathbb{E}_Q\left[X_T\frac{B(t)}{B(T)}\Big|\mathcal{F}_t\right] = \frac{1}{P(t,T)}\pi(X_T) \end{aligned} \tag{29}$$

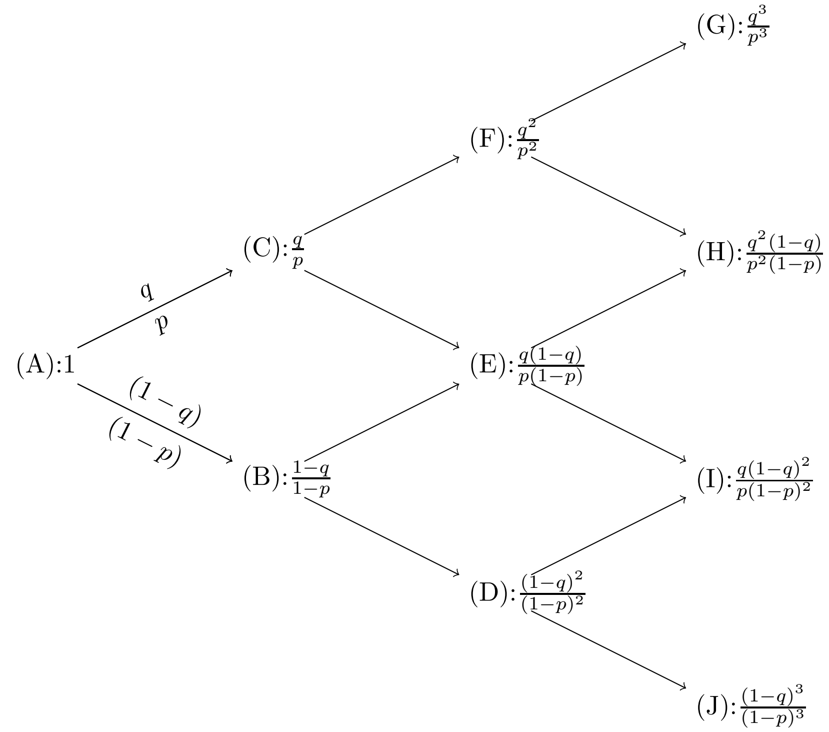
Under Q there is a quotient of X_T and B_T whereas under Q^T we have only X_T. So the formula 29 is useful when X_T and B_T are dependent, which is the case for the interest rate derivatives. However, this formula is just an elegant but useless theoretical construction until we specify the dynamics of $X(t)$ under Q^T. To do this we need to understand the *change of numeraire* for which, in turn, we need the

2 Multivariate Girsanov's Theorem and Cholesky decomposition

2.1 Preliminaries

Before we state the multivariate Girsanov's theorem, we will explain what η_t is in the discrete case (on a binary tree) and then prove two useful lemmas for η_t in continuous time.
So let us recall the change from the objective measure P to the martingale

measure Q on a binary tree. Q is not a T-Forward measure (Q is called the spot measure) but for a discussion of the Radon-Nykodim process this is irrelevant.



I noted before that the T-Forward measure is defined from $t = 0$ with T in mind. In our case T corresponds to the last four nodes: (G), (H), (I) and (J). Though Q is not a T-Forward measure, the Radon Nykodim *derivative* (in the sense of the classical Real Analysis) is still defined from $t = 0$ and with T in mind. (Or in our case it is better to say "from (A) with {(G), (H), (I), (J)} in mind"). Namely, for a measurable space $(\Omega, \mathcal{F})$ and two absolutely continuous measures Q and P on $(\Omega, \mathcal{F})$ (i.e which agree about the measure-zero sets of Ω) there is a unique[9] function f, called the Radon Nykodim derivative, such that

$$Q(A) = \int_A f(\omega)dP(\omega) \overbrace{=}^{\star} \mathbb{E}_P\left[f(\omega)1_A\right] = \mathbb{E}_P\left[f(\omega)1_A|\mathcal{F}_0\right] \qquad \forall A \in \mathcal{F} \quad (30)$$

The $\star$ in 30 is valid if P is a probability measure (in our case it is by definition). The question is in which case is Q a probability measure as well? An obvious necessary and sufficient condition is the nonnegativity of f and $\int_\Omega f(w)dP(\omega) = 1$ must hold true. Indeed, Q is a measure (i.e. $Q(\emptyset) = 0$ and Q is σ-additive) by the Radon Nykodim theorem. And since $Q(\Omega) = 1$ it is a probability measure[10].

[9] Upto a set of measure zero w.r.t. P

[10] Now I am sure you can prove that Q^T defined in 26 is a probability measure.

In our case the Radon Nykodim derivative is defined according to Table 1:

Table 1: Radon Nykodim derivative from (A) to {(G), (H), (I), (J)} in mind

ω	(G)	(H)	(I)	(J)
$f(\omega)$	$\frac{q^3}{p^3}$	$\frac{q^2(1-q)}{p^2(1-p)}$	$\frac{q(1-q)^2}{p(1-p)^2}$	$\frac{(1-q)^3}{(1-p)^3}$

Because the measures are discrete, an integral in 30 becomes just a sum and

$$Q(\Omega) = \int\limits_{\Omega} f(\omega) dP(\omega) = \frac{q^3}{p^3} p^3 + \frac{q^2(1-q)}{p^2(1-p)} 3p^2(1-p)$$

$$+\frac{q(1-q)^2}{p(1-p)^2} 3p(1-p)^2 + \frac{(1-q)^3}{(1-p)^3}(1-p)^3 = 1$$

Considering $\star$ in 30 more generally, we want to express the expectations w.r.t. Q as the expectations w.r.t. P. For any random variable $X \in \mathcal{F}$ it is actually

$$\mathbb{E}_Q[X] = \int\limits_{\Omega} X dQ(\omega) = \int\limits_{\Omega} X f(\omega) dP(\omega) = \mathbb{E}_P[f(\omega) X] \tag{31}$$

But we are interested not only in the *single* Radon Nykodim derivative. Rather we want to define the Radon Nykodim *process* η_t so that {(G), (H), (I), (J)} in mind stays but now we are looking from any node, not only from (A). It is just like replacing $P(T,T)$ and B_T in 27 with $P(t,T)$ and B_t in 28. We can define the Radon Nykodim derivative e.g. from (C) to {(G), (H), (I), (J)} in mind (Table 2). Note that it is impossible to move in (J) from (C), so the respective

Table 2: Radon Nykodim derivative from (C) to {(G), (H), (I), (J)} in mind

ω	(G)	(H)	(I)	(J)
$f(\omega)$	$\frac{q^2}{p^2}$	$\frac{q(1-q)}{p(1-p)}$	$\frac{(1-q)^2}{(1-p)^2}$	0

probability is set to zero.

However, such approach is indeed not what we need. What do we need to price derivatives? Just to take expectation w.r.t. the martingale measure and sometimes express this expectation via expectation w.r.t. the real and T-forward measures. Now recall 31! So we do not need the Radon Nykodim process as a function, we only need it in the sense of expectation, however, *conditional* expectation, given a σ-algebra $\mathcal{F}_t$.
So we define the Radon Nykodim process as

$$\eta_t = \mathbb{E}_P\left[\frac{Q}{P}\Big|\mathcal{F}_t\right] = \mathbb{E}_P\left[f(\omega)\big|\mathcal{F}_t\right] \qquad \omega \in \{(G), (H), (I), (J)\} \tag{32}$$

A conditional expectation η_t is, informally speaking, a random variable upto time t but in t is takes a concrete value and is no more random. Let's calculate the η_t for $t \in [1, .., 4]$, i.e. from start to end of our binary tree

$$t = 0 \qquad \eta_0 = \mathbb{E}_P\left[f(\omega)\middle|\mathcal{F}_0\right] = \mathbb{E}_P\left[f(\omega)\right] = 1$$

At time $t = 0$ the conditional expectation is equal to the unconditional expectation and the Randon Nykodim derivative is equal to 1 as we calculated above. Keep this point in mind, in the next section we introduce η_t more formally in continuous time and will in particular start with a nonnegative random variable Z, s.t. $\mathbb{E}_P[Z] = 1$

$$t = 2 \qquad \eta_2 = \left[p\frac{q^3}{p^3} + (1-p)\frac{q^2(1-q)}{p^2(1-p)}\right] 1_{(F)} + \left[p\frac{q^2(1-q)}{p^2(1-p)} + (1-p)\frac{q(1-q)^2}{p(1-p)^2}\right] 1_{(E)}$$

$$+ \left[p\frac{q(1-q)^2}{p(1-p)^2} + (1-p)\frac{(1-q)^3}{(1-p)^3}\right] 1_{(D)} = \frac{q^2}{p^2}1_{(F)} + \frac{q(1-q)}{p(1-p)}1_{(E)} + \frac{(1-q)^2}{(1-p)^2}1_{(D)}$$

Here $1_{(F)}$, $1_{(E)}$, $1_{(D)}$ respectively mean the event that at time $t = 2$ we land in the node (F), (E), (D). In advance we do not know which note it will be, but as time goes to $t = 2$ we will know it. If, e.g., we are in (D) by the time $t = 2$ then $\eta_2 = \frac{(1-q)^2}{(1-p)^2}$

$$t = 3 \qquad \eta_3 = \frac{q^3}{p^3}1_{(G)} + \frac{q^2(1-q)}{p^2(1-p)}1_{(H)} + \frac{q(1-q)^2}{p(1-p)^2}1_{(I)} + \frac{(1-q)^3}{(1-p)^3}1_{(J)}$$

For $t = 1$ you calculate n_1 as an exercise. A property of the nested conditional expectations $\mathbb{E}_P[\mathbb{E}_P[\eta_3|\mathcal{F}_2]|\mathcal{F}_1] = \mathbb{E}_P[\eta_3|\mathcal{F}_1]$ will make the computation easier since we already calculated $\eta_2 = \mathbb{E}_P[\eta_3|\mathcal{F}_2]$.

Now let us express the price of a contingent claim X at time $t = 1$ via the Radon Nykodim process (we write $\pi_1(X)$ for this price). For [notation] simplicity assume the interest rate is zero. Also assume w.l.o.g. that at time $t = 1$ we are in node (C) Then

$$\pi_1(X) = \mathbb{E}_Q[X_3|\mathcal{F}_1] = q^2 X_{(G)} + 2q(1-q)X_{(H)} + (1-q)^2 X_{(I)} + 0X_{(J)}$$

$$= p^2\frac{q^2}{p^2}X_{(G)} + 2p(1-p)\frac{q(1-q)}{p(1-p)}X_{(H)} + (1-p)^2\frac{(1-q)^2}{(1-p)^2}X_{(I)}$$

$$= \frac{p}{q}\left[p^2\frac{q^3}{p^3}X_{(G)} + 2p(1-p)\frac{q^2(1-q)}{p^2(1-p)}X_{(H)} + (1-p)^2\frac{q(1-q)^2}{p(1-p)^2}X_{(I)}\right]$$

$$= \frac{p}{q}\mathbb{E}_P[\eta_3 X_3|\mathcal{F}_1] = \frac{1}{\eta_1}\mathbb{E}_P[\eta_3 X_3|\mathcal{F}_1] \qquad (33)$$

Analogously assume that at time $t = 1$ we are in node (B) and show that 33 holds in this case too.

In general it holds

$$\mathbb{E}_Q[X_T|\mathcal{F}_s] = \frac{1}{\eta_s}\mathbb{E}_P[\eta_T X_T|\mathcal{F}_s] \tag{34}$$

We made a big (though a bit boring) job to understand the idea of the Radon Nykodim process by means of concrete examples. Right now we are going to rigorously discuss this idea in continuous time.
Consider a usual filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F})_t, P)$. Let Z be an a.s. positive random variable, whose value becomes known at time T, satisfying $\mathbb{E}_P[Z] = 1$. Define a new probability measure Q as (cp. 30)

$$Q(A) = \mathbb{E}_P\left[Z 1_A\right] \qquad \Leftrightarrow \qquad \frac{dQ}{dP} = Z \tag{35}$$

Define a Radon Nykodim Process[11]

$$Z(t) = \mathbb{E}_P[Z|\mathcal{F}(t)], \qquad 0 \leq t \leq T \tag{36}$$

$Z(t)$ is a martingale since by iterated conditioning

$$\mathbb{E}_P[Z(t)|\mathcal{F}(s)] = \mathbb{E}_P[\mathbb{E}_P[Z|\mathcal{F}(t)]\big|\mathcal{F}(s)] = \mathbb{E}_P[Z|\mathcal{F}(s)] = Z(s) \tag{37}$$

Now, following Shreve, we prove two lemmas in which we present further properties of $Z(t)$

Lemma 2.1. *Let t satisfying $0 \leq t \leq T$ be given and let Y be an $\mathcal{F}(t)$-measurable random variable. Then $\mathbb{E}_Q[Y] = \mathbb{E}_P[YZ(t)]$*

Proof.

$$\mathbb{E}_Q[Y] = \mathbb{E}_P[YZ] = \mathbb{E}_P[\mathbb{E}_P[YZ|\mathcal{F}_t]] \overbrace{=}^{\star} \mathbb{E}_P[Y\mathbb{E}_P[Z|\mathcal{F}_t]] = \mathbb{E}_P[YZ(t)]$$

where in $\star$ we took Y out of the conditional expectation since Y is $\mathcal{F}(t)$-measurable and thus known at t.

Lemma 2.2. *Let s and t satisfying $0 \leq s \leq t \leq T$ be given and let Y be an $\mathcal{F}(t)$-measurable random variable. Then (cp. 33)*

$$\mathbb{E}_Q[Y|\mathcal{F}(s)] = \frac{1}{Z(s)}\mathbb{E}_P[YZ(t)|\mathcal{F}(s)]$$

Proof. According to the definition of the conditional expectation, $\frac{1}{Z(s)}\mathbb{E}_P[YZ(t)|\mathcal{F}(s)]$ must be $\mathcal{F}_s$-measurable (which is obvious since $\mathcal{F}_s \subset \mathcal{F}_t$) and the partial-averaging property must hold, i.e.

$$\int\limits_A \frac{1}{Z(s)}\mathbb{E}_P[YZ(t)|\mathcal{F}_s]dQ \stackrel{!}{=} \int\limits_A Y dQ \qquad \forall A \in \mathcal{F}_s$$

[11] Here I abuse with notation a little bit writing $Z(t)$ instead of η_t in order to keep the formulae consistent with Shreve's book.

With some algebra, where in $(\star)$ we take in 1_A which is known, since $A \in \mathcal{F}_s$

$$\int\limits_A \frac{1}{Z(s)} \mathbb{E}_P[YZ(t)|\mathcal{F}_s]dQ \overset{\text{def}}{=} \mathbb{E}_Q\left[1_A \frac{1}{Z(s)} \mathbb{E}_P[YZ(t)|\mathcal{F}_s]\right] \overset{2.1}{=}$$

$$\mathbb{E}_P\left[Z(s)1_A \frac{1}{Z(s)} \mathbb{E}_P[YZ(t)|\mathcal{F}_s]\right] \overset{(\star)}{=} \mathbb{E}_P\left[\mathbb{E}_P[1_A YZ(t)|\mathcal{F}_s]\right] \overset{2.1}{=} \mathbb{E}_Q[1_A Y] \overset{\text{def}}{=} \int\limits_A Y dQ$$

2.2 Girsanov's Theorem

Exercise 2.3. *Take your Financial Math lecture notes or Shreve's book and go through the proof of the Girsanov's theorem in one dimension.*

Theorem 2.4. *(Girsanov, multiple dimension) Let T be an arbitrary but fixed positive time, $\theta(t) = [\theta_1(t), .., \theta_d(t)]$ be a d-dimensional adapted process and $W^P(t) = [W_1^P(t), .., W_d^P(t)]$ be a d-dimensional uncorrelated[12] Brownian Motion under a measure P. Define*

$$Z(t) = \exp\left(\int\limits_0^t \theta(u)dW^P(u) - \frac{1}{2}\int\limits_0^t \|\theta(u)\|^2 du\right) \tag{38}$$

$$W^Q(t) = W^P(t) + \int\limits_0^t \theta(u)du$$

and[13] assume that

$$\mathbb{E}_P\left[\int\limits_0^T \|\theta(u)\|^2 Z^2(u)du < \infty\right]$$

Set $Z = Z(T)$. Then $\mathbb{E}_P[Z] = 1$ and under the probability measure Q given by

$$Q(A) = \int\limits_A Z(\omega)dP(\omega) \qquad \forall A \in \mathcal{F} \tag{39}$$

the process $W^Q(t)$ is a d-dimensional Brownian motion. Moreover, if the $W^P(t)$ is the only source of uncertainty (i.e. it generates the filtration of our probability space) then the converse theorem holds too, i.e. all absolutely continuous changes of measure are in form of 39.

[12] In a sense, the theorem holds for correlated processes too, however, there are some nuances, see below.

[13] This is Novikov condition, which guarantees that the Radon Nykodim process $Z(t)$ does not explode.

Some remarks on notation:

$$W^Q(t) = [W_1^Q(t), .., W_d^Q(t)] \qquad W_j^Q(t) = W_j^P(t) + \int_0^t \theta_j(u)du \quad j = 1, .., d$$

$$\|\theta(u)\| = \sqrt{\sum_{j=1}^{d} \theta_j^2(u)}$$

$$\int_0^t \theta(u)dW(u) = \int_0^t \sum_{j=1}^{d} \theta_j(u)dW_j(u) = \sum_{j=1}^{d} \int_0^t \theta_j(u)dW_j(u)$$

Note that

$$dZ(t) = \theta(t)Z(t)dW^P(t) \tag{40}$$

and $\theta(t)$ is sometimes called the Girsanov kernel. If you have thoroughly done the exercise 2.3 the proof of the multidimensional case shall be nothing new but some technical details, as declared in many textbooks (Shreve p. 225, Björk p.165) and so it may be omitted. However, if the components of $W^P(t)$ are correlated, the things are not so easy: we demonstrate it by means of example with $d = 2$, $\theta = [\theta_1, \theta_2]$ and $dW_1^P(t)dW_2^P(t) = \rho dt$. In this case we have

$$\begin{aligned} dZ(t) &= Z(t)[\theta_1 dW_1^P(t) + \theta_2 dW_2^P(t) - \frac{1}{2}(\theta_1^2 + \theta_2^2)dt] \\ &+ \frac{1}{2}Z(t)[\theta_1^2 d^2W_1^P(t) + \theta_2^2 d^2W_2^P(t) + 2\theta_1\theta_2 dW_1^P(t)dW_2^P(t)] \\ &= Z(t)[\theta_1 dW_1^P(t) + \theta_2 dW_2^P(t) + \rho\theta_1\theta_2 dt] \end{aligned} \tag{41}$$

and the drift part $\rho\theta_1\theta_2 dt \neq 0$ thus $Z(t)$ is not a martingale. If we modify 38 so that $Z(t) = \exp(\int_0^t \theta(u)dW^P(s) - \frac{1}{2}\int_0^t [W^P]_t)$ (it is called the Doleans-Dade exponential, $[W]_t$ stands for the quadratic variation of $W(t)$) then $Z(t)$ is a martingale but now one should be careful with $W_j^Q(t) = W_j^P(t) + \int_0^t \theta_j(u)du\ldots$ However, we can always switch from a correlated Wiener Process to the uncorrelated one by means of

2.3 Cholesky decomposition

Let $W(t) = [W_1(t), .., W_d(t)]$ be a d-dimensional correlated Wiener process. We know that a multivariate normal random variable is completely characterized by $(\vec{\mu}, \Sigma)$, i.e. by its covariance matrix and the vector of expectations. It is also the case for the multivariate Wiener process, whereas its vector of drifts is zero and so the covariance matrix $\Sigma(t)$ is only what matters[14]. $\Sigma(t)$ may be time depen-

[14] By the Ito Isometry we can introduce the L_2-norm of $W(t)$ defined on $[0, T]$:
$\|W\|_{L_2}^2 = \sum_{i=1}^{d}\sum_{j=1}^{d}\int_0^T \rho_{ij}(s)\sigma_i(s)\sigma_j(s)ds$
Since the covariance matrix is positive definite, $\|W\|_{L_2}^2$ is nonnegative. If you like functional analysis, check that all other norm properties hold true.

dent and even random (but should be $\mathcal{F}_t$ measurable). Since $\Sigma(t)$ is positive definite, there exists a unique triangular matrix $H(t)$ such that $\Sigma(t) = H(t)H^T(t)$. (One can consider H as a square root from Σ). Let $B(t) = [B_1(t), .., B_d(t)]$ be uncorrelated standard Wiener process. Then $W^T(t) = H(t)B^T(t)$ since $H(t)B^T(t)$ has the covariance matrix $\Sigma(t)$.
We consider the simplest example with

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{bmatrix} \stackrel{!}{=} \begin{bmatrix} a & 0 \\ b & c \end{bmatrix} \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} = \begin{bmatrix} a^2 & ab \\ ab & b^2 + c^2 \end{bmatrix} \tag{42}$$

We immediately see that $a = \sigma_1$, $b = \sigma_2\rho$ and $c = \sigma_2\sqrt{1-\rho^2}$. So

$$\begin{bmatrix} W_1(t) \\ W_2(t) \end{bmatrix} = \begin{bmatrix} \sigma_1 & 0 \\ \sigma_2\rho & \sigma_2\sqrt{1-\rho^2} \end{bmatrix} \begin{bmatrix} B_1(t) \\ B_2(t) \end{bmatrix} = \begin{bmatrix} \sigma_1 B_1(t) \\ \sigma_2\rho B_1(t) + \sigma_2\sqrt{1-\rho^2} B_2(t) \end{bmatrix}$$

Exercise 2.5. *Check that this process has indeed the covariance matrix* Σ

Exercise 2.6. *Consider the cases d=3 and d=4 and calculate the Cholesky decomposition with Maple (with(linalg) then cholesky(Σ)). You may also try for d=3 manually if you like exercising in identical transformations.*
Though the elements of H will be much longer, the main idea remains the same: $W_1 = \sigma_1 B_1(t)$, $W_2(t)$ is the weighted sum of $B_1(t), B_2(t)$ and $W_3(t)$ is the weighted sum of $B_1(t), B_2(t), B_3(t)$ where the weights are the respective elements of H.

Cholesky decomposition allows us to turn a correlated Wiener process to an uncorrelated one by $B^P(t) = H^{-1}(t)W^P(t)$, commit Girsanov's transformation on $B^P(t)$, obtain $B^Q(t)$ and finally multiply it with $H(t)$.

Exercise 2.7. *Do this for the Σ from 42 and $\theta = [\theta_1, \theta_2]$.*
Check that under Q the covariance matrix remains the same. Compare it to the fact that one-dimensional Girsanov transformation does not change the volatility.

Cholesky decomposition also allows us to simulate a multivariate correlated Wiener Process. For that we just need a random number generator (e.g. Box-Muller), which generates the i.i.d standard normal random variables.

3 Change of Numeraire

Now we go on with the change of numeraire, which is practically useful only together with change of measure, for which, in turn, we need the multivariate Girsanov's theorem. Technically, a numeraire is just a strictly positive process, economically it is the asset, in which all other assets are denominated. To be a numeraire an asset should not pay dividends (they cause the price drop). A common numeraire is a bond or a bank account, so you have often considered a discounted stock price $S(t)/B(t)$ and a discounted bond price $B(t)/B(t) =$

1. But technically nothing stops us from using $S(t)$ as a numeraire. Then $S(t)/S(t) = 1$ and we must consider the dynamics of $B(t)/S(t)$. You see that the price of a numeraire-asset, "discounted" [in terms of numeraire itself] is always one.
The problem by the change of numeraire is that the risk neutral measure changes too.

Exercise 3.1. *Find the dynamics of $B(t)/S(t)$ in the Black-Scholes model under the real measure. Hint: just use Ito product rule.*
Also find the risk neutral measure, assosicated with the new numeraire. Hint: just kill the drift.

Exercise 3.2. *Consider a market with two correlated risky assets and a riskless bank account. Let the market evolve under the spot martingale measure Q as follows*

$$\begin{aligned} dS_1(t) &= S_1(t)[rdt + \sigma_1 dB_1(t)] \\ dS_2(t) &= S_2(t)[rdt + \sigma_2 \rho dB_1(t) + \sigma_2\sqrt{1-\rho^2}dB_2(t)] \\ dS_0(t) &= rS_0(t)dt \end{aligned} \tag{43}$$

Change the numeraire to S_1 and calculate the market dynamics under the new numeraire. Try to find the Girsanov kernel, which makes all assets to martingales under the new numeraire.

There are plausible economic interpretations of the change of numeraire. First of all it is the case when we deal with both domestic and foreign currency, but ok, in both cases there are just bank accounts. To be closer to the exercise 3.1, consider the following idea: as the Crisis'2008 came, many investors decided to turn their dollars and euros to gold. So sitting on a pile of gold they can estimate how many other assets (among them bonds) they can buy for their gold.
Exercise 3.1 is also a useful trick to derive the Black-Scholes-Merton formula with random interest rate (s. Shreve, p. 394, Theorem 9.4.2).

Theorem 3.3. *Let $X_0(t), X_1(t), .., X_N(t)$ be traded assets so that $\frac{X_0(t)}{X_0(t)} = 1$, $\frac{X_1(t)}{X_0(t)}, .., \frac{X_N(t)}{X_0(t)}$ are martingales (under some measure Q). Fix $T > t$ and define a new probability measure by*

$$Q^i(A) = \frac{X_0(0)}{X_i(0)} \int\limits_A \frac{X_i(T)}{X_0(T)} dQ \tag{44}$$

Then $\frac{X_0(t)}{X_i(t)} \frac{X_1(t)}{X_i(t)}, .., \frac{X_N(t)}{X_i(t)}$ are martingales under Q^i

Proof. By the Radon Nykodim theorem Q^i is a measure. We must show that it is a probability measure, i.e. that $Q^i(\Omega) = 1$. It follows

$$Q^i(\Omega) = \frac{X_0(0)}{X_i(0)} \int\limits_\Omega \frac{X_i(T)}{X_0(T)} dQ = \frac{X_0(0)}{X_i(0)} \mathbb{E}_Q \left[\frac{X_i(T)}{X_0(T)} \right] \overbrace{=}^{(\star)} \frac{X_0(0)}{X_i(0)} \frac{X_i(0)}{X_0(0)} = 1$$

In $(\star)$ we used the martingale property of $\frac{X_i(t)}{X_0(t)}$ under Q.
Now we show that $\frac{X_j(t)}{X_i(t)}$ is a martingale under Q_i. From 44 we can readily identify the Radon Nykodim derivative, it is $\frac{X_0(0)}{X_i(0)}\frac{X_i(T)}{X_0(T)}$. We want to obtain the Radon Nykodim process which is

$$\eta_t = \mathbb{E}_Q\left[\frac{X_0(0)}{X_i(0)}\frac{X_i(T)}{X_0(T)}\Big|\mathcal{F}_t\right] \overbrace{=}^{(\star)} \frac{X_0(0)}{X_i(0)}\frac{X_i(t)}{X_0(t)}$$

where in $(\star)$ we again used the martingale property. Moreover

$$\mathbb{E}_{Q^i}\left[\frac{X_j(T)}{X_i(T)}\Big|\mathcal{F}_t\right] \overset{2.2}{=} \frac{\cancel{X_i(0)}}{\cancel{X_0(0)}}\frac{X_0(t)}{X_i(t)}\mathbb{E}_Q\left[\frac{\cancel{X_0(0)}}{\cancel{X_i(0)}}\frac{\cancel{X_i(T)}}{X_0(T)}\frac{X_j(T)}{\cancel{X_i(T)}}\Big|\mathcal{F}_t\right] \tag{45}$$

Of course until we specify the dynamics of $\frac{X_j(t)}{X_i(t)}$ under Q^i the pricing formula 45 is useless for practice. Indeed, changing the numeraire we change the volatility of discounted assets (and it should be obvious since now $X_i(t)/X_i(t) = 1$ so its vola is zero but $X_0(t)/X_i(t)$ is volatile). You may look at details in Shreve's book. We, however, can proceed further with LIBOR without them.

4 The Multivariate LIBOR Market Model

Finally, we can start with our main topic. The main advantage of the LIBOR Market Model[15] (besides it provides a theoretical justification for Black-76) is that the forward rates are either directly observable on an interbank market or are calculable from [government] bond prices[16]. Let us at first recall, how the LIBOR forward rates are calculated. One fixes a tenor δ, which is a year fraction: usually one, three or six months. We consider the *simply* compounded forward rates $L(t, T, T+\delta)$ i.e.

$$P(t, T+\delta)(1 + \delta L(t, T, T+\delta)) = P(t, T) \tag{46}$$

The LIBOR model is essentially discrete, i.e. $T = \{T_0 = \delta, T_1 = 2\delta, .., T_n = (n+1)\delta\}$. For notation simplicity we write $L_j(t)$ instead of $L(t, T_{j-1}, T_j) = L(t, T_{j-1}, T_{i-j} + \delta)$. From 46 we immediately derive

$$L_j(t) = \frac{1}{\delta}\left(\frac{P(t, T_{j-1})}{P(t, T_{j-1}+\delta)} - 1\right) = \frac{1}{\delta}\left(\frac{P(t, T_{j-1})}{P(t, T_j)} - 1\right) \tag{47}$$

[15] Following an established tradition, we say "LIBOR Model" though it is actually a *forward rates* model. There is a big family of IBORs, among which the LIBOR and the EURIBOR are the most important.

[16] Unfortunately not all forward rates can be uniquely derived from bond prices, since in reality the distance between the maturities of the long term bonds are much longer than the LIBOR tenor. E.g. if there are the bonds with matirities $T_{i-1} = 10$ years, $T_i = 15$ years and the tenor is 6 months, we have to interpolate the bond prices, which is a non trivial task in practice.

Further we introduce the forward discount factor

$$P_j(t) = \frac{P(t, T_{j-1})}{P(t, T_j)} \quad \Leftrightarrow \quad L_j(t) = \frac{1}{\delta}\left(P_j(t) - 1\right) \tag{48}$$

The equation 47 tells us how to reproduce the $L_j(t)$ with a bond portfolio. So the LIBORs are tradable. On the other hand we can derive the prices of the bonds with maturities $T_1, T_2, .., T_n$ from the LIBOR rates. Thus it may (and it does) make sense to model the dynamics of LIBOR rates directly.

So in LIBOR Model we *postulate* that each LIBOR rate L_j has a dynamics under the T_j-Forward measure Q^{T_j} according to

$$dL_j(t) = \sigma_j(t) L_j(t) dW^{Q_{T_j}}(t) \tag{49}$$

Why T_j-Forward martingale measure Q^{T_j} and not simply the [spot] martingale measure Q?! Well, first of all $P(t, T_j)$ in the denominator in 47 can give us an idea about a suitable numeraire and secondly, the Black-76 is consistent under the T-Forward measure but is not under the spot measure.
If we want to price caplets/floorlets only, we do not need anything else, since the Black-76 Formula immediately follows from 49 (prove it!). Even if we need a price of Caps/Floors, 49 will still do, since they are just a sum of independent caplets/floorlets. But if we need to price e.g. swaptions, we need to consider the *correlation* between different LIBOR rates and they are obviously correlated (indeed strongly but not perfectly $\Rightarrow$ multifactor model). So we have n LIBOR rates and thus n correlated Wiener Processes. In order to decorrelate it, we use the Cholesky toolkit[17] and write 49 in vector notation as

$$\begin{bmatrix} W_1^{Q_{T_1}}(t) \\ \vdots \\ W_n^{Q_{T_n}}(t) \end{bmatrix} = \begin{bmatrix} h_{11}(t) & \dots & 0 \\ \vdots & \ddots & \vdots \\ h_{1n}(t) & \dots & h_{nn}(t) \end{bmatrix} \begin{bmatrix} B_1^{Q_{T_1}}(t) \\ \vdots \\ B_n^{Q_{T_n}}(t) \end{bmatrix} \tag{50}$$

If you carefully did exercise 2.5 you see that for each $j = [1, .., n]$ 49 follows from 50, i.e.

$$\sigma_j(t) dW^{Q_{T_j}} = \sum_{k=1}^{n} h_{jk}(t) dB_k^{Q_{T_k}}(t) \tag{51}$$

Please note that I deliberately write "n correlated Wiener Processes" and avoid writing "n-dimensional Wiener Process". As a matter of fact an n-dimensional Wiener Process should be defined under some measure, the *same* for all its components. W.l.o.g. we choose it to be the measure Q^{T_j}. From Girsanov's theorem we know that changing from one measure to another simply

[17] If we do not want to model time dependent volatilities and correlations, we can simply estimate the covariance matrix Σ from historical data and then calculate H.

means changing the drift term. So we can rewrite 50 as

$$\begin{bmatrix} W_1^{Q_{T_j}}(t) - \mu_1^j(t) \\ \vdots \\ W_j^{Q_{T_j}}(t) \\ \vdots \\ W_n^{Q_{T_j}}(t) - \mu_n^j(t) \end{bmatrix} = \begin{bmatrix} h_{11}(t) & \dots & 0 \\ \vdots & \ddots & \vdots \\ h_{1n}(t) & \dots & h_{nn}(t) \end{bmatrix} \begin{bmatrix} B_1^{Q_{T_j}}(t) - m_1^j(t) \\ \vdots \\ B_j^{Q_{T_j}}(t) - m_j^j(t) \\ \vdots \\ B_n^{Q_{T_j}}(t) - m_n^j(t) \end{bmatrix} \tag{52}$$

We do not yet know the values of $\mu_1^j(t), .., \mu_n^j(t)$ and must find them. However, we do not need to look for the explicit values of $m_1^j(t), .., m_n^j(t)$

From Theorem 3.3 we obtain

$$\eta_j^{j-1}(t) := \frac{dQ^{T_{j-1}}}{dQ^{T_j}}|\mathcal{F}_t = \frac{P(0,T_j)}{P(0,T_{j-1})}\frac{P(t,T_{j-1})}{P(t,T_j)} \tag{53}$$

Further

$$\begin{aligned} d(\eta_j^{j-1}(t)) &\overset{48}{=} \frac{P(0,T_j)}{P(0,T_{j-1})} d(P_j(t)) = \frac{P(0,T_j)}{P(0,T_{j-1})} d(1+\delta L_j(t)) \\ &= \underbrace{\frac{P(0,T_j)}{P(0,T_{j-1})}\frac{P(t,T_{j-1})}{P(t,T_j)}}_{\eta_j^{j-1}(t)} \underbrace{\frac{P(t,T_j)}{P(t,T_{j-1})}}_{\frac{1}{P_j(t)}} d(1+\delta L_j(t)) \\ = \frac{\eta_j^{j-1}(t)}{P_j(t)} \delta\sigma_j(t) L_j(t) dW^{Q_{T_j}} &= \eta_j^{j-1}(t)\frac{\delta\sigma_j(t)L_j(t)}{1+\delta L_j(t)} dW^{Q_{T_j}} \\ &= \eta_j^{j-1}(t)\frac{\delta L_j(t)}{1+\delta L_j(t)} \sum_{k=1}^n h_{jk}(t) dB^{Q_{T_j}}(t) \end{aligned}$$

$$= \eta_j^{j-1}(t)\frac{\delta L_j(t)}{1+\delta L_j(t)}[h_{j1}(t), .., h_{jn}(t)] \begin{bmatrix} dB_1^{Q_{T_j}}(t) - dm_1^j(t) \\ \vdots \\ dB_j^{Q_{T_j}}(t) - dm_j^j(t) \\ \vdots \\ dB_n^{Q_{T_j}}(t) - dm_n^j(t) \end{bmatrix} \tag{54}$$

From 40 we immediately identify $\frac{\delta\sigma_j(t)L_j(t)}{1+\delta L_j(t)} = \frac{\delta L_j(t)}{1+\delta L_j(t)}[h_{j1}(t), .., h_{jn}(t)]$ as Girsanov kernel $\theta(t)$. Respectively

$$\begin{bmatrix} dB_1^{Q_{T_j}}(t) - dm_1^j(t) \\ \vdots \\ dB_j^{Q_{T_j}}(t) - dm_j^j(t) \\ \vdots \\ dB_n^{Q_{T_j}}(t) - dm_n^j(t) \end{bmatrix} = \frac{\delta L_j(t)}{(1+\delta L_j(t))} \begin{bmatrix} h_{j1}(t) \\ \vdots \\ h_{jj}(t) \\ \vdots \\ h_{jn}(t) \end{bmatrix} + \begin{bmatrix} dB_1^{Q_{T_{j-1}}}(t) - dm_1^j(t) \\ \vdots \\ dB_j^{Q_{T_{j-1}}}(t) - dm_j^j(t) \\ \vdots \\ dB_n^{Q_{T_{j-1}}}(t) - dm_n^j(t) \end{bmatrix} \tag{55}$$

and

$$\sigma_j(t) dW^{Q_{T_j}} = [h_{j1}(t), .., h_{jj}(t), .., h_{jn}(t)] \left(\frac{\delta L_j(t)}{(1+\delta L_j(t))} \begin{bmatrix} h_{j1}(t) \\ \vdots \\ h_{jj}(t) \\ \vdots \\ h_{jn}(t) \end{bmatrix} + \begin{bmatrix} dB_1^{Q_{T_{j-1}}}(t) - dm_1^j(t) \\ \vdots \\ dB_j^{Q_{T_{j-1}}}(t) - dm_j^j(t) \\ \vdots \\ dB_n^{Q_{T_{j-1}}}(t) - dm_n^j(t) \end{bmatrix} \right)$$

$$= \frac{\delta L_j(t)}{(1+\delta L_j(t))} \sigma_j^2(t) + \sigma_j(t) dW^{Q_{T_{j-1}}}(t) \tag{56}$$

Plugging this into 49 yields together with 51

$$dL_j(t) = L_j(t) \left(\frac{\delta L_j(t)}{(1+\delta L_j(t))} \sigma_j^2(t) + \sigma_j(t) dW^{Q_{T_{j-1}}}(t) \right) \tag{57}$$

Rearranging 56 we obtain

$$\sigma_j(t) dW^{Q_{T_{j-1}}}(t) = \sigma_j(t) dW^{Q_{T_j}}(t) - \frac{\delta L_j(t)}{(1+\delta L_j(t))} \sigma_j^2(t) \tag{58}$$

or equivalently

$$dW^{Q_{T_{j-1}}}(t) = dW^{Q_{T_j}}(t) - \frac{\delta L_j(t)}{(1+\delta L_j(t))} \sqrt{\sum_{k=1}^{n} h_{jk}^2(t)} \tag{59}$$

Similarly for $dW^{Q_{T_{j-2}}}(t)$

$$dW^{Q_{T_{j-2}}}(t) = dW^{Q_{T_{j-1}}}(t) - \frac{\delta L_{j-1}(t)}{(1+\delta L_{j-1}(t))} \sqrt{\sum_{k=1}^{n} h_{(j-1)k}^2(t)} \tag{60}$$

Further

$$dL_j(t) = L_j(t)\sqrt{\sum_{k=1}^{n} h_{jk}^2(t)}\left(dW^{Q_{T_{j-1}}}(t) + \frac{\delta L_j(t)}{(1+\delta L_j(t))}\sqrt{\sum_{k=1}^{n} h_{jk}^2(t)}\right) =$$

$$L_j(t)\sqrt{\sum_{k=1}^{n} h_{jk}^2(t)}\left(dW^{Q_{T_{j-2}}}(t) + \frac{\delta L_{j-1}(t)}{(1+\delta L_{j-1}(t))}\sqrt{\sum_{k=1}^{n} h_{(j-1)k}^2(t)} + \frac{\delta L_j(t)}{(1+\delta L_j(t))}\sqrt{\sum_{k=1}^{n} h_{jk}^2(t)}\right)$$

$$= L_j(t)\left(\sigma_j(t)dW^{Q_{T_{j-2}}}(t) + \frac{\delta L_{j-1}(t)}{(1+\delta L_{j-1}(t))}\sigma_j(t)\sigma_{(j-1)}(t)\rho_{(j-1)j}(t) + \frac{\delta L_j(t)}{(1+\delta L_j(t))}\sigma_j^2(t)\right) \quad (61)$$

Proceeding so iteratively we obtain

$$dW^{Q_{T_j}}(t) = \begin{cases} dW^{Q_{T_p}}(t) - \sum\limits_{k=j+1}^{p} \frac{\delta L_k(t)}{1+\delta L_k(t)}\sigma_k(t)dt & j < p \\ dW^{Q_{T_p}}(t) & j = p \\ dW^{Q_{T_p}}(t) + \sum\limits_{k=p}^{j-1} \frac{\delta L_k(t)}{1+\delta L_k(t)}\sigma_k(t)dt & j > p \end{cases} \quad (62)$$

And

$$dL_j(t) = \begin{cases} L_j(t)\sigma_j(t)dW^{Q_{T_p}}(t) - L_j(t)\sum\limits_{k=j+1}^{p} \frac{\delta L_k(t)}{1+\delta L_k(t)}\sigma_j(t)\sigma_k(t)\rho_{jk}dt & j < p \\ L_j(t)\sigma_j(t)dW^{Q_{T_p}}(t) & j = p \\ L_j(t)\sigma_j(t)dW^{Q_{T_p}}(t) + L_j(t)\sum\limits_{k=p}^{j-1} \frac{\delta L_k(t)}{1+\delta L_k(t)}\sigma_j(t)\sigma_k(t)\rho_{jk}dt & j > p \end{cases} \quad (63)$$

Finally, it is useful to note that drifts $dm_1^j(t), .., dm_n^j(t)$ in 55 cancel out thus we can, using 62 and 63 and fixing some measure (let it be w.l.o.g Q^{T_1}), rewrite 52 as

$$\begin{bmatrix} dW_1^{Q_{T_1}}(t) \\ \vdots \\ dW_j^{Q_{T_1}}(t) \\ \vdots \\ dW_n^{Q_{T_1}}(t) \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ \sum\limits_{k=1}^{j-1} \frac{\delta L_k(t)}{1+\delta L_k(t)}\sigma_k(t)dt \\ \vdots \\ \sum\limits_{k=1}^{n-1} \frac{\delta L_k(t)}{1+\delta L_k(t)}\sigma_k(t)dt \end{bmatrix} + \begin{bmatrix} h_{11}(t) & \dots & 0 \\ \vdots & \ddots & \vdots \\ h_{1n}(t) & \dots & h_{nn}(t) \end{bmatrix} \begin{bmatrix} dB_1^{Q_{T_1}}(t) \\ \vdots \\ dB_j^{Q_{T_1}}(t) \\ \vdots \\ dB_n^{Q_{T_1}}(t)) \end{bmatrix} \quad (64)$$

The formulae 62 , 63 and 64 are the main result of this section. Of course it would be also interesting to prove that all bond prices, discounted with the T_j-bond as numeraire are indeed martingales under Q^{T_j}. But it is mainly technical question and we can spare it (or postpone till the next tutorial).
But the last task we have to do is to derive the dynamics of LIBOR rate under the spot martingale measure Q, i.e. the measure associated with the bank

account as the numeraire. For this we need to introduce the idea of a [discrete] bank account. It is as follows: assume $T_{j-1} < t < T_j$, i.e. L_{j-1} (the interest rate from T_{j-1} to T_j) is already fixed and no more random. We define the rolling over bank account at time t as follows: invest 1€ at T_0 and reinvest on each reset date until T_{j-1}. At time t the present value of such bank account is

$$B(t) = P(t, T_j) \prod_{k=0}^{j-1} (1 + \delta L_k) \tag{65}$$

We derive the dynamics of W^Q via the dynamics of $W^{Q_{T_{j+1}}}$. Inverting 28 yields

$$\eta(t) = \frac{dQ}{dQ^{Tj+1}} = \frac{B(t)P(0, T_{j+1})}{P(t, T_{j+1})} = \frac{\overbrace{P(t, T_j) \prod_{k=0}^{j-1} (1 + \delta L_k) P(0, T_{j+1})}^{:=C}}{P(t, T_{j+1})} \tag{66}$$

Analogous to 53 we obtain

$$d[\eta(t)] = Cd[P_{j+1}(t)] = Cd[1 + \delta L_{j+1}(t)] = C\delta\sigma_{j+1}(t)L_{j+1}(t)dW^{Q_{T_{j+1}}}$$
$$= C\delta\sigma_{j+1}(t)L_{j+1}(t)\frac{P(t, T_j)}{P(t, T_{j+1})}\frac{P(t, T_{j+1})}{P(t, T_j)}dW^{Q_{T_{j+1}}} = \eta(t)\frac{\delta\sigma_{j+1}(t)L_{j+1}(t)}{1 + \delta L_{j+1}(t)}dW^{Q_{T_{j+1}}} \tag{67}$$

So $\frac{\delta\sigma_{j+1}(t)L_{j+1}(t)}{1+\delta L_{j+1}(t)}$ is the Girsanov kernel and you know how proceed further.

5 What's next?

We did big (and hopefully good) job, so that now you are well prepared to your exam in interest rate modeling. However, this theory is only the top of iceberg so for your job [in order both to find and to do it] you may need to know the following things:

1). The LIBOR market model, which we have considered is actually the *forward* LIBOR model. It is not compatible with the *swap* LIBOR model (just like Black-76 for caps is not compatible with Black'76 for swaptions). However, there is a surprisingly good approximation for the swaption prices in the forward model.

2). Implied volatilities and implied correlations. As I steadily pointed out, the LIBOR market model is consistent with Black-76. Unfortunately the market prices of caps and floors are not! This phenomenon is known as volatility smile (analogously correlation smile in case of swaptions). So if you put the historical volatility into Black-76, you will hardly get the current market price of a cap or a floor. But you can proceed other way around and find an *implied* volatility, which makes Black-76 consistent with current price. This is known

as “plugging the wrong value into the wrong formula to get the right price”. Implied volatility can be time dependent and even strike dependent. It reflects the market expectation of the future volatility. There is nothing unnatural in the dependence on strike. Obviously, if the price of an asset goes too high or too low there will be a big hype and the volatility will grow.
Analogously the implied correlations can be extracted from swaptions. Swaptions and Caps(Floors) are very liquid derivatives, so one can consider them as “primary” traded assets as well. We use them to calibrate our models and then to price not so liquid derivatives.

3). Practically, it is difficult to extract the implied correlations from the market, so one introduce a parametrization. $\rho_{jk} = \exp(-\beta|T_j - T_k|)$ is a popular choice.

4). There are many models for stochastic implied volatility and the most popular is probably the SABR model. There is a whole book on the LIBOR with SABR (see References).

5). We have derived the dynamics of the LIBOR rates under different risk neutral measures. Due to pretty complicated drifts we should carefully choose the measure under which to simulate - in order to achieve good numerical accuracy.

6). The LIBOR model is essentially discrete but how about derivatives, whose maturities are not the multiples of the tenor δ? In this case one has to interpolate the LIBOR dynamics but this is in no way trivial task if we want to stay arbitrage free and keep the positivity of all rates.

7). So far we have assumed that the LIBOR rates are driven by a multidimensional Wiener Process. To model a defaultable term structure and, in general, for better reality approximation we might need to add jump processes.

8). In our model the number of Brownian Motions is equal to the number of LIBOR rates. This is indeed too many. The *Principle Component Analysis* shows that 3 random factors are enough, they are interpreted as parallel shifts, tilting(affects the slope) and flex(affect the curvature) of the forward curve.

9). After the Crisis’2008 the world of interest rates became multicurve. In particular, it means that the LIBOR or EURIBOR rates are still used as a reference for IR Swaps but for discounting one takes another curve, see Bianchetti and Carlicchi(2012). Moreover, the swaps are usually collateralized, which makes even a pricing of plain-vanilla swaps very challenging, see Brigo(2012).

6 References

Fischer Black (1976). The pricing of commodity contracts, Journal of Financial Economics, 3, 167-179.

Oldrich Vasicek (1977). An Equilibrium Characterisation of the Term Structure. Journal of Financial Economics 5 (2): 177188

John C. Hull - Options, Futures and Other Derivatives, 7th Ed. - Prentice Hall, 2008

Damiano Brigo and Fabio Mercurio - Interest Rate Models - Theory and Practice: With Smile, Inflation and Credit, 2nd Ed. - Springer, 2006

Damiano Brigo - Counterparty Risk FAQ: Credit VaR, PFE, CVA, DVA, Close-out, Netting, Collateral, Re-hypothecation, WWR, Basel, Funding, CCDS and Margin Lending, 2012 - http://arxiv.org/pdf/1111.1331v3.pdf

Marco Bianchetti and Mattia Carlicchi - Interest Rates After The Credit Crunch: Multiple-Curve Vanilla Derivatives and SABR, 2012
http://arxiv.org/ftp/arxiv/papers/1103/1103.2567.pdf

Damir Filipovic - Term Structure Models: A Graduate Course - Springer, 2009

Riccardo Rebonato, Kenneth McKay and Richard White - The SABR/LIBOR Market Model: Pricing, Calibration and Hedging for Complex Interest Rate Derivatives - Wiley, 2009

Dariusz Gatarek, Przemyslaw Bachert and Robert Maksymiuk - The LIBOR Market Model in Practice - Wiley, 2012

Thomas Björk - Arbitrage Theory in Continuous Time, 3rd Ed. - Oxford Finance, 2009

Steven E. Shreve - Stochastic Calculus for Finance II: Continuous-Time Models, 2nd Ed. - Springer, 2004

James Jardine - An Examination and Implementation of the Libor Market Model (Master Thesis), 2006 - http://www.jimme.net/Thesis.pdf

7 Solutions to Exercises

Exercise 2.5

$$cov[dW_1(t), dW_1(t)] = VAR[dW_1(t)] = \sigma_1^2 VAR[dB_1(t)] = \sigma_1^2 dt$$

$$cov[dW_2(t), dW_2(t)] = VAR[dW_2(t)] \overset{(\star)}{=} \rho^2\sigma_2^2 VAR[dB_1(t)] + (1-\rho^2)\sigma_2^2 VAR[dB_2(t)] = \sigma_2^2 dt$$

$$cov[dW_1(t), dW_2(t)] = \mathbb{E}[dW_1(t)\cdot dW_2(t)] - \underbrace{\mathbb{E}[dW_1(t)]\mathbb{E}[dW_2(t)]}_{=0}$$

$$= \mathbb{E}[\sigma_1\sigma_2\rho dB_1(t)dB_1(t) + \sigma_1\sigma_2\sqrt{1-\rho^2}\underbrace{dB_1(t)dB_2(t)}_{=0(\star)}] = \sigma_1\sigma_2\rho dt$$

where in $(\star)$ we used the independence of $B_1(t)$ and $B_2(t)$

Exercise 2.7

$$\begin{bmatrix} dB_1^Q(t) \\ dB_2^Q(t) \end{bmatrix} = \begin{bmatrix} \theta_1 dt + dB_1^P(t) \\ \theta_2 dt + dB_2^P(t) \end{bmatrix}$$

Thus

$$\begin{bmatrix} \sigma_1 & 0 \\ \sigma_2\rho & \sigma_2\sqrt{1-\rho^2} \end{bmatrix} \begin{bmatrix} dB_1^Q(t) \\ dB_2^Q(t) \end{bmatrix} = \begin{bmatrix} \sigma_1 dB_1^Q(t) \\ \sigma_2\rho dB_1^Q(t) + \sigma_2\sqrt{1-\rho^2}dB_2^Q(t) \end{bmatrix}$$

$$= \begin{bmatrix} dW_1^Q(t) \\ dW_2^Q(t) \end{bmatrix} = \begin{bmatrix} \sigma_1[\theta_1 dt + dB_1^P(t)] \\ \sigma_2\rho[\theta_1 dt + dB_1^P(t)] + \sigma_2\sqrt{1-\rho^2}[\theta_2 dt + dB_2^P(t)] \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_1\theta_1 dt + dW_1^P(t) \\ \sigma_2[\rho\theta_1 + \sqrt{1-\rho^2}\theta_2]dt + dW_2^P(t)] \end{bmatrix} =: \begin{bmatrix} \mu_1 dt + dW_1^P(t) \\ \mu_2 dt + dW_2^P(t)] \end{bmatrix}$$

Respectively, if the drifts coefficients μ_1 and μ_2 are given and you are going to kill the drifts by $W_1^P(t)$ and $W_1^P(t)$, you have to choose $\theta_1 = -\frac{\mu_1}{\sigma_1}$
and $\theta_2 = -\left(\frac{\mu_2}{\sigma_2} - \rho\theta_1\right)/\sqrt{1-\rho^2}$
Looking at this solution you are probably trying to reconcile it with section 5.4.3 of Shreve's book. In this case, do not be confused with Example 5.4.4, note that in this example there are *two* stocks but only *one* Brownian Motion.

Exercise 3.1

Recall: $dB_t = rB_t dt$ und $dS_t = S_t(\mu dt + \sigma dW_t)$

$$d\left(\frac{B_t}{S_t}\right) = \frac{dB_t}{S_t} + B_t d\left(\frac{1}{S_t}\right) = \frac{rB_t dt}{S_t} + B_t\left(-\frac{1}{S_t^2}dS_t + \frac{1}{2}\frac{2S_t}{S_t^4}d^2S_t\right)$$

$$= \frac{B_t}{S_t}\left([r-\mu+\sigma^2]dt - \sigma dW_t\right) = \frac{B_t}{S_t}\left([r-\mu+\sigma^2]dt + \sigma dW_t\right)$$

since $W(t)$ is statistically indistinguishable from $-W(t)$. We kill the drift of B_t under numeraire S_t as follows:

$$\left([r-\mu+\sigma^2]dt + \sigma dW_t\right) = \sigma\left([\frac{r-\mu}{\sigma}+\sigma]dt + dW_t\right) =: \sigma dW_t^{Q_S}$$

Now consider

$$d\left(\ln\frac{B_t}{S_t}\right) = \frac{S_t}{B_t}d\left(\frac{B_t}{S_t}\right) - \frac{1}{2}\frac{S_t^2}{B_t^2}d^2\left(\frac{B_t}{S_t}\right) = \left([r-\mu+\frac{\sigma^2}{2}]dt + \sigma dW_t\right)$$

And finally

$$\frac{B_t}{S_t} = \frac{B_0}{S_0}\exp\left([r-\mu+\frac{\sigma^2}{2}]t + \sigma W_t\right) = \frac{1}{S_0}\exp\left([r-\mu+\frac{\sigma^2}{2}]t + \sigma W_t\right)$$

Exercise 3.2

Here to keep the notation simpler I omit (t), writing e.g. S_1 instead of $S_1(t)$.

$$d\left(\frac{S_2}{S_1}\right) = S_2 d\left(\frac{1}{S_1}\right) + \frac{1}{S_1}dS_2 + dS_2 d\left(\frac{1}{S_1}\right)$$

$$= S_2\left[-\frac{1}{S_1^2}S_1\left(rdt + \sigma_1 dB_1\right) + \frac{1}{S_1}\sigma_1^2 dt\right] + \frac{S_2}{S_1}\left(rdt + \sigma_2\rho dB_1 + \sigma_2\sqrt{1-\rho^2}dB_2\right)$$

$$+\frac{S_2}{S_1}\underbrace{(-rdt - \sigma_1 dB_1 + \sigma_1^2 dt)(rdt + \sigma_2\rho dB_1 + \sigma_2\sqrt{1-\rho^2}dB_2)}_{\sigma_1\sigma_2\rho dB_1 dB_1 \text{remains only}}$$

$$= \frac{S_2}{S_1}\left[(\sigma_1^2 - \sigma_1\sigma_2\rho)dt + (\sigma_2\rho - \sigma_1)dB_1 + \sigma_2\sigma\sqrt{1-\rho^2}dB_2\right]$$

The dynamics of $\frac{S_0}{S_1}$ follows from the exercise 3.1 setting $r=\mu$, i.e.

$$d\left(\frac{S_0}{S_1}\right) = \frac{S_0}{S_1}\sigma_1\left(\sigma_1 dt + dB_1\right) = \frac{S_0}{S_1}\sigma_1 dB_1^{Q_{S_1}}$$

Respectively (recalling that we have already replaced $-dB_1$ with dB_1 in the previous equation and thus must do it in the following too) we obtain

$$\frac{S_2}{S_1}\left[(\sigma_1^2 - \sigma_1\sigma_2\rho)dt + (\sigma_2\rho - \sigma_1)dB_1 + \sigma_2\sigma\sqrt{1-\rho^2}dB_2\right]$$

$$= \frac{S_2}{S_1}\left[(\sigma_1^2 - \sigma_1\sigma_2\rho)dt + (\sigma_1 - \sigma_2\rho)(\sigma_1 dt + dB_1) - (\sigma_1 - \sigma_2\rho)\sigma_1 dt + \sigma_2\sigma\sqrt{1-\rho^2}dB_2\right]$$

$$= \frac{S_2}{S_1}\left[(\sigma_1 - \sigma_2\rho)(\sigma_1 dt + dB_1) + \sigma_2\sigma\sqrt{1-\rho^2}dB_2\right] = \frac{S_2}{S_1}\left[(\sigma_1 - \sigma_2\rho)dB_1^{Q_{S_1}} + \sigma_2\sigma\sqrt{1-\rho^2}dB_2^{Q_{S_1}}\right]$$

Note that the drift under the neu numeraire S_1 was completely absorbed with dB_1. This is because we have already started with the risk-neutral dynamics under the numeraire S_0. Had we started under the real world measure like in Exercise 3.1 it would not be the case and a part of the drift would be absorbed with dB_2.