2. Foundations of Mathematical Finance and Stochastic Calculus

This chapter presents the necessary foundations to understand the mathematical, financial and computational aspects behind this model. In the first section we start with simple interest rate necessities and go on to financial Derivatives which are necessary to understand to correctly calibrate and use the model for pricing. The section 2.2 starts with the most important aspects in stochastic calculus which is the key step to understand and work with stochastic differential equations. One additional topic in this section is the no-arbitrage Pricing which are a prerequisite to understand modern option pricing theory. The final section in this chapter gives the reader an overview about the computational aspects which are important to build this model.

2.1. Interest Rates and Derivatives

This important section is about the basic definitions which will be used through the whole thesis. All Definitions are quite similar to the definitions of standard term structure modeling and interest rate derivatives textbooks, see Brigo and Mercurio [2006], Joshi [2008], Björk [2009], Fries [2007] and Glasserman [2004].

In the section below we consider a set of increasing maturities $T_0, T_1, ..., T_N$.

$$t = T_0, j = 1, ..., N$$

Definition 2.1.1 ***(Money market account)***
We start with the evolution of a simple money market account $B(t)$ over time where we use the definition of the instantaneous short rate $r(t)$. Under the short rate one can understand the interest rate under which the money accrues when reinvested continuously. It is important to note that the short rate is a theoretical concept and cannot be directly observed at the market. We start with $B(t) = 1$. The differential equation for the evolution

C. Hackl, *Calibration and Parameterization Methods for the Libor Market Model,*
BestMasters, DOI 10.1007/978-3-658-04688-0_2, © Springer Fachmedien Wiesbaden 2014

is

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

where one solution for this differential equation is

$$B(T) = \exp\left\{\int_t^T r_s ds\right\} \tag{2.1}$$

Definition 2.1.2 ***(Zero Coupon Bond)***
A Zero Coupon Bond or just Zero Bond has no intermediate payments and guarantees its holder one unit of amount at time T_j. The natural boundaries when interest rates are positive are $P(t,T_j) < 1$ and arbitrage would be possible if not $P(t,T_j) > 0$ as zero costs at t would produce 1 income in T_j.

$$P(t,T_j) = \frac{B(t)}{B(T_j)} \tag{2.2}$$

if we use (2.1) from the instantaneous short rate, this formula leads to the stochastic Zero Coupon Bond

$$D(t,T_j) = \exp\left\{-\int_t^T r_s ds\right\} \tag{2.3}$$

Under a suitable probability measure, the expectation of the stochastic discount factors (stochastic Zero Coupon Bonds) are the Zero Coupon Bonds (deterministic discount factors).

Definition 2.1.3 ***(Forward Zero Bond)***
A theoretical Forward Zero Bond can be interpreted as the amount which has to be invested in T_{j-1} to get one unit of currency in T_j.

$$P(t,T_{j-1},T_j) = \frac{P(t,T_j)}{P(t,T_{j-1})} \tag{2.4}$$

Definition 2.1.4 ***(Spot Interest Rate)***
The spot rate or LIBOR rate is the constant interest rate which has to be applied to an amount which has to be invested at time t to get one unit at time T_j. It is defined in simple compounding convention as

$$L(t,T_j) = \frac{1}{\delta}\left(\frac{1}{P(t,T_j)} - 1\right) \tag{2.5}$$

where δ is a year fraction for typically $\frac{3}{12}, \frac{6}{12}$ or $\frac{9}{12}$. In this work the LIBOR rate, which represents the USD interest rate, is used as a synonym for other currencies, which use the same convention (eg EURIBOR for EUR, PRIBOR for CZK). The LIBOR rate is not explicitly modeled in the LMM, instead the forward Libor rate is modeled which is defined below.

Definition 2.1.5 ***(Forward Rate)***
The theoretical definition of forward rates is an interest rate which is set today for borrowing or lending for a certain period in the future.

$$F(t, T_{j-1}, T_j) = \frac{1}{T_j - T_{j-1}} \left(\frac{P(t, T_{j-1})}{P(t, T_j)} - 1 \right) \tag{2.6}$$

Definition 2.1.6 ***(Forward Libor Rate)***

$$\delta_j L(t, T_{j-1}, T_j) = \frac{P(t, T_{j-1})}{P(t, T_j)} - 1 = \frac{P(t, T_{j-1}) - P(t, T_j)}{P(t, T_j)} \tag{2.7}$$

Typically the fixed δ for a forward Libor Rate is 3 or 6 months, this market convention is also used below for the Libor Market Model. The above definition is also the definition of a FRA (Forward Rate Agreement), which is a traded instrument in financial markets.

Definition 2.1.7 ***(Swap Rate)***
A product which can be replicated out of FRAs, which are traded contracts on forward rates, is an interest rate swap. In an interest rate swap two parties exchange differently indexed payments, which is in its standard (plain vanilla) form a fixed interest rate against a floating interest rate (e.g. Libor rate) payment, at a specified future time instant. The fixed payer pays a prespecified amount $N\delta_j K$ at each instant T_j where $j \in [\alpha, ..., \beta]$, N is the notional amount, δ_j is the time increment (a year fraction) between T_j and T_{j-1}. The floating rate payer pays, an index tied interest rate, where the index is a Libor rate, of $N\delta_j L(T_{j-1}, T_j)$ at the dates $T_{\alpha+1}, ..., T_\beta$. The reset dates, which are the dates where the next floating rate is fixed, are at time $T_\alpha, ..., T_{\beta-1}$. As we have already stated above, the swap is a portfolio of forward rate agreements and hence can be valued

$$N \sum_{j=\alpha+1}^{\beta} \delta_j P(t, T_j)(K - F(t, T_{j-1}, T_j))$$

for a fixed rate receiver and for the payer we would just have to change the signs in the brackets.

To find the fixed swap rate at the contract start we have to set the two sides of the contract to zero and solve for $K = S_{\alpha,\beta}(t)$:

$$N \sum_{j=\alpha+1}^{\beta} \delta_j P(t,T_j) \left(K - \delta_j^{-1} \left(\frac{P(t,T_{j-1}) - P(t,T_j)}{P(t,T_j)} \right) \right) = 0$$

$$N \sum_{j=\alpha+1}^{\beta} [\delta_j P(t,T_j) K - P(t,T_{j-1}) + P(t,T_j)] = 0$$

$$\Rightarrow K \sum_{j=\alpha+1}^{\beta} \delta_j P(t,T_j) = P(t,T_\alpha) - P(t,T_\beta)$$

$$S_{\alpha,\beta}(t) = K = \frac{P(t,T_\alpha) - P(t,T_\beta)}{\sum_{j=\alpha+1}^{\beta} \delta_j P(t,T_j)} \tag{2.8}$$

The swap rate is equal to the forward rate if we take just one time period into consideration. As the swap rate is needed in terms of forward rates for the simulation, some algebraic manipulation of the previous expression has to be done. We start by dividing the numerator and denominator by $P(t,T_\alpha)$:

$$\begin{aligned} S_{\alpha,\beta}(t) &= \frac{1 - \frac{P(t,T_\beta)}{P(t,T_\alpha)}}{\sum_{j=\alpha+1}^{\beta} \delta_j \frac{P(t,T_j)}{P(t,T_\alpha)}} \\ &= \frac{1 - \frac{P(t,T_\beta)}{P(t,T_\alpha)} \overbrace{\prod_{i=\alpha+1}^{\beta-1} \frac{P(t,T_i)}{P(t,T_i)}}^{1}}{\sum_{j=\alpha+1}^{\beta} \delta_j \frac{P(t,T_j)}{P(t,T_\alpha)} \underbrace{\prod_{i=\alpha+1}^{j-1} \frac{P(t,T_i)}{P(t,T_i)}}_{1}} \\ &= \frac{1 - \prod_{i=\alpha+1}^{\beta} \frac{P(t,T_i)}{P(t,T_{i-1})}}{\sum_{j=\alpha+1}^{\beta} \delta_j \prod_{i=\alpha+1}^{j} \frac{P(t,T_i)}{P(t,T_{i-1})}} \end{aligned}$$

we use 2.7 to get

$$S_{\alpha,\beta}(t) = \frac{1 - \prod_{i=\alpha+1}^{\beta} \frac{1}{1+\delta_i L(t,T_{i-1},T_i)}}{\sum_{j=\alpha+1}^{\beta} \delta_j \prod_{i=\alpha+1}^{j} \frac{1}{1+\delta_i L(t,T_{i-1},T_i)}} \tag{2.9}$$

Definition 2.1.8 ***(Caplets/Floorlets)***
Like in Fries [2007] a caplet is defined as an call option on a forward (Libor) rate. A cap with a quoted tenor $T_j - t$ and strike K is the sum of all caplets c_j on the forward Libor rates until tenor $T_j - T_{j-1}$, all with strike K and the caplet volatility for the forward libor rate $L(t, T_{j-i}, T_j)$ is σ_j. The general payoff formula for caplets is

$$X_j \quad = \quad \delta_j \max[L(t, T_{j-1}, T_j) - K]^+$$

where $[]^+$ is defined as the functional form $[x]^+ := max[x, 0]$ and the Black '76 formula for caplets at time t is given by (the superscript B identifies the caplet formula as Black style quotation)

$$c_j^B(t) \quad = \quad \delta_j P(t, T_j)(L(t, T_{j-1}, T_j)\Phi(d_1) - K\Phi(d_2)) \tag{2.10}$$

where we have

$$d_1 \quad = \quad \frac{1}{\sigma_j\sqrt{T_j - t}}\left[\ln\left(\frac{L(t, T_{j-1}, T_j)}{K}\right) + \frac{1}{2}\sigma_j^2(T - t)\right] \textit{and} \tag{2.11}$$

$$d_2 \quad = \quad d1 - \sigma_j\sqrt{T_j - t} \tag{2.12}$$

the function for a cap is

$$cap \quad = \quad \sum\nolimits_{j=1}^{N} c_j(\sigma_j^B) \tag{2.13}$$

In financial markets the quotation is $cap = \sum_{j=1}^{N} c_j(\sigma^B)$ as there is no explicit quotation for caplet volatilities, which would be spot or forward volatilities, but just for cap volatilities which are flat for each cap tenor. Therefore the quoted cap volatilities have to be "bootstrapped" to get out the caplet volatilities for each tenor. It is also market practice in option trading that the quotation for caps/floors is in implied Black volatility and not in monetary terms as one would expect.
Black's Formula or very often called Black '76 is a modification of the Black-Scholes-Merton stochastic differential equation and its origin lies in pricing derivatives contracts for commodities but it is far more frequently used in pricing interest rate derivatives. It's importance for the Libor Market Model lies in the fact that it is consistent with the Black Formula which is a very nice characteristic for calibration. The pricing with the Black Formula implies that the forward (libor) rate is lognormally distributed at expiry under the

risk neutral measure. Another assumption is that the risk neutral rate is not stochastic. This assumptions are not plausible and the derivation of caplets with the Black formula which makes this point clear is shown in the Appendix.

A floor is like a cap where the signs in the payoff function is changed. It is a put option on the forward (libor) rate and like the cap is a sum of caplets, the floor is a sum of floorlets. The general payoff formula for floorlets is $X_j = \delta_j \max[K - L(t, T_{j-1}, T_j)]^+$. *One of the conclusions of the put call parity relationship is that selling a cap and buying a floor is equivalent to a receiver swap with fixed rate equal to the strike rate* K. *Prices of caps are not influenced by correlations of the underlying forward (libor) rates, which is opposite to swaption contracts, where the correlations of forward log-returns do matter.*

Definition 2.1.9 ***(Swaptions)***
Similar to a cap on forward rates, swaptions are options on swap rates. Payer swaptions provide the right to enter into a payer swap at a specific future point in time and receiver swaps provide the right to enter into a receiver swap at a specific future point in time. The swap tenor is from T_α *till* T_β *and the payoff of a payer swaption at time* t *is (where* $t \leq T_\alpha$*):*

$$
\begin{aligned}
&NP(t,T_\alpha)\left[\sum_{j=\alpha+1}^{\beta} P(T_\alpha,T_j)\delta_j(L(T_\alpha,T_{j-1},T_j)-K)\right]^+\\
=&NP(t,T_\alpha)\left[\sum_{j=\alpha+1}^{\beta} P(T_\alpha,T_j)\delta_j\frac{1}{\delta_j}\frac{P(T_\alpha,T_{j-1})-P(T_\alpha,T_j)}{P(T_\alpha,T_j)} - K\sum_{j=\alpha+1}^{\beta}\delta_j P(T_\alpha,T_j)\right]^+\\
=&NP(t,T_\alpha)\left[\sum_{j=\alpha+1}^{\beta} P(T_\alpha,T_{j-1})-P(T_\alpha,T_j) - K\sum_{j=\alpha+1}^{\beta}\delta_j P(T_\alpha,T_j)\right]^+\\
\Rightarrow=&NP(t,T_\alpha)\left[P(T_\alpha,T_\alpha)-P(T_\alpha,T_\beta)\frac{\sum_{j=\alpha+1}^{\beta}\delta_j P(T_\alpha,T_j)}{\sum_{j=\alpha+1}^{\beta}\delta_j P(T_\alpha,T_j)} - K\sum_{j=\alpha+1}^{\beta}\delta_j P(T_\alpha,T_j)\right]^+\\
\Rightarrow=&NP(t,T_\alpha)\left[S_{\alpha,\beta}(T_\alpha)\sum_{j=\alpha+1}^{\beta}\delta_j P(T_\alpha,T_j) - K\sum_{j=\alpha+1}^{\beta}\delta_j P(T_\alpha,T_j)\right]^+\\
\Rightarrow=&NP(t,T_\alpha)\sum_{j=\alpha+1}^{\beta}\delta_j P(T_\alpha,T_j)\left[S_{\alpha,\beta}(T_\alpha)-K\right]^+
\end{aligned}
$$

With this last equation we can see that we don't have, as in the cap/floor case an instrument which can be broken down further. If the option is executed at the expiry date, the swap rate has to be paid for the tenor of the swap, there is no optionality any more. This *fact also shows, that generally a payer swaption has to be less expensive than a corresponding cap contract. This can be shown by the following equation, where, on the left side of the inequality, the summation of discount factors cannot be taken out of* $[]^+$. *:*

$$\left[\sum_{j=\alpha+1}^{\beta} \delta_j P(T_\alpha, T_j)[L(T_\alpha, T_{j-1}, T_j) - K]\right]^+ \leq \sum_{j=\alpha+1}^{\beta} \delta_j P(T_\alpha, T_j)\,[L(T_\alpha, T_{j-1}, T_j) - K]^+$$

The Black quotation is also market practice for swaption contracts and they are calculated in a very similar way compared to caps/floors:

$$swaption^B_{\alpha,\beta}(t) = \quad PVBP(S_{\alpha,\beta}(0)\Phi(d_1) - K\Phi(d_2))^+ \tag{2.14}$$

where PVBP is the present value of a basis point which is also the denominator in 2.8:

$$PVBP_\beta = \quad P(0,\alpha) \sum_{j=\alpha+1}^{\beta} \delta_j P(\alpha, T_j)$$

In general it cannot be assumed that both, forward rates and swap rates, are lognormal distributed, as the swap rate measure cannot be expressed linearly in terms of the forward rate measure, which is also true for the opposite direction. It is market practice to ignore as inconsistencies are small, see Joshi [2008].

2.2. Stochastic Calculus and No-Arbitrage Pricing

In the theory of financial derivatives pricing the no-arbitrage theory is the major building block which was also the first step in deriving the famous Black-Scholes Option pricing formula. Using no-arbitrage theory the continuous time evolution of a financial instrument is modeled as a stochastic process and as a Stochastic Differential Equation (SDE). As we need the stochastic process or SDE to be driftless, we use the martingale measure which is a very important concept in modeling financial assets. Additionally we need to understand theorems like Girsanov's Theorem to change the numeraire and therefore the drift of a stochastic process or SDE to get to an equivalent martingale measure (EMM). The SDE is stated via the application of Itô Calculus .
Detailed explanations can be found in standard stochastic calculus and no-arbitrage theory textbooks which were stated at the beginning of the previous chapter and are also used for this chapter.

We define a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where the elements of the $\sigma - algebra$ $\mathcal{F}$ are the events on the sample space Ω with the probability measure $\mathbb{P}$, where $\mathbb{P}(\Omega) = 1$. As we don't know the future interest rates or in the case of the LMM the future libor forward rates, we need a stochastic process to model the future. We use the Wiener Process (or Brownian Motion), which is driven by independent standard normal random variables, for modeling the stochastic part which represents the uncertainty in our model.

Definition 2.2.1 ***(Brownian Motion)***
We define the n-dimensional continuous process $W^n(t) : t \geq s \geq 0$ on our probability space $(\Omega, \mathcal{F}^n, \mathbb{P})$ where $W^n(0) \geq 0$, the stochastic increments $W^n(s+t) - W^n(s)$ are independent of the history of $\mathcal{F}^n_s$, the increments $W^n(s+t) - W^n(s) \sim \mathcal{N}(0,t)$ distributed under the measure $\mathbb{P}$. The Filtration $\mathcal{F}^n_s$ consists of all information up to $W^n(s)$.With this property the process $W^n(s)$ is called adapted to the filtration $\mathcal{F}^n_s$.

Definition 2.2.2 ***(Stochastic Process and Stochastic Differential Equations)***
Let X be a n-dimensional continuous process with $X(t) : t \geq 0$ then we get this integral equation

$$X(t) = X(0) + \int_0^t \mu(s)\, ds + \int_0^t \sigma(s) dW(s)$$

where $\mu(t)$ is a drift vector, $dW(t)$ is a vector of brownian motions with n-dimensions and $\sigma(t)$ is a $n * n$ matrix of volatilities (or diffusion coefficients). Where X has the differential form

$$dX(t) = \mu(t)dt + \sigma(t)dW(t)$$

If σ and μ are deterministic functions and $X(t)$ depend just on $W(t)$, the differential is called a Stochastic Differential Equation (SDE).

Definition 2.2.3 ***(Martingale and Equivalent Martingale Measure)***
In Björk [2009] a stochastic process M(t) is a ***martingale*** *under a filtration $\mathcal{F}_t$ if and only if $E[M(t)] < \infty$ $\forall t$ and the following relation must hold $\forall s \leq t$*

$$E[M(t)|\mathcal{F}_s] = M(s)$$

this condition is very important as it says that the future value M given the information available today, equals the present value M. We have already stated in the introduction of this section that this driftlessness of the stochastic process is needed but not always observed. Therefore we need the concept of change of measure to get from a drift including probability measure $\mathbb{P}$ *to an equivalent driftless (martingale) measure* $\mathbb{Q}$ *to apply martingale theory.*

We use the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ *as stated above to express a measure* $\mathbb{Q}$ *which is equivalent to the measure* $\mathbb{P}$*. It needs to fulfill the following properties:*

$$\begin{aligned} \mathbb{P}(A) > 0 \Leftrightarrow \mathbb{Q}(A) > 0 \;\; \forall A \in \Omega \\ \textit{which is equivalent to} \\ \mathbb{P}(A) = 0 \Leftrightarrow \mathbb{Q}(A) = 0 \;\; \forall A \in \Omega \end{aligned}$$

Definition 2.2.4 ***(Change of Measure)***
To get from one measure to another we use the Radon-Nikodym Theorem. We use a nonnegative random variable Z with $E^{\mathbb{P}}[Z] = 1$ *then we can define a probability measure* $\mathbb{Q}$ *which is equivalent to the measure* $\mathbb{P}$*. Under the filtration* $\mathcal{F}_t$ *where* $t \in [0, T]$ *we get in this case a process where* $Z(t)$ *is the Radon-Nikodym derivative.*

$$\begin{aligned} d\mathbb{Q} \;&= Z(T)d\mathbb{P} \text{ on fixed } \mathcal{F}_T \text{ generates a process} \\ \Rightarrow \quad Z(t) &= E^{\mathbb{P}_t}[\frac{d\mathbb{Q}}{d\mathbb{P}}] \Leftrightarrow \frac{1}{Z(t)} = \frac{d\mathbb{P}}{d\mathbb{Q}} \text{ on } \mathcal{F}_t \end{aligned}$$

Definition 2.2.5 ***(Girsanov Theorem)***
When we change from a martingale process on one measure to an equivalent measure, the new process is generally not martingale under the new measure. The Girsanov Theorem tells us how we can change the drift of a process to get the new process martingale under the new measure. We do not subtract or add the drift of the process to get it to zero, rather we assign new probabilities to each event of the distribution. This concept is very important for the Libor Market Model to get all forward rate processes martingale under the terminal measure.
We have $W^{\mathbb{P}}$ *which is a n-dimensional standard Brownion Motion,* φ *which is an arbitrary n-dimensional adapted process vector and the process* $Z(t)$ *as from the change of measure definition above. Since Z(t) is a nonnegative martingale process we write Z as*

$$\begin{aligned} dZ(t) &= \varphi(t)Z(t)dW^{\mathbb{P}}(t) \qquad \textit{with Z(0) = 1} \\ \Rightarrow Z(t) &= exp\left(\int_0^t \varphi(s)dW^{\mathbb{P}}(s) - \frac{1}{2}\int_0^t \|\varphi(s)\|^2 ds\right) \end{aligned}$$

Using the theory from Definition 2.2.4

$$W^{\mathbb{Q}}(t) = W^{\mathbb{P}}(t) - \int_0^t \varphi(s)ds \tag{2.15}$$

We have assumed above that φ *is a process such that* $E^{\mathbb{P}}[Z(T)] = 1$ *or equivalently that the likelihood ratio* Z *is a martingale. A sufficient condition to guarantee that* Z *is a true martingale is the "Novikov Condition":*

$$E^{\mathbb{P}}\left[exp\left(\frac{1}{2}\int_0^T \|\varphi(s)\|^2 ds\right)\right] < \infty$$

The outcome of the application of Girsanov's Theorem is that the diffusion term stays unchanged and the drift, we use μ *for the drift under measure* $\mathbb{P}$*, ist changed to a new drift* $\mu + \sigma\varphi$ *under* $\mathbb{Q}$*. A very short example of the application of the Girsanov Theorem on the process* X *where* W *is a Wiener process under the respective measure:*

$$\begin{aligned} dX(t) &= \mu(t)dt + \sigma(t)dW^{\mathbb{P}}(t) \\ &\quad \textit{then we use 2.15 to get} \\ &= \mu(t)dt + \sigma(t)(dW^{\mathbb{Q}}(t) + \varphi(t)dt) \\ &= (\mu(t) + \sigma(t)\varphi(t))dt + \sigma(t)dW^{\mathbb{Q}}(t) \end{aligned}$$

Definition 2.2.6 ***(General (Martingale) Pricing Formula)***
"The First Fundamental Theorem of Asset Pricing" states that a financial model is arbitrage free if there exists a (local) martingale measure $\mathbb{Q}^N$ *which is equivalent to the risk neutral measure* $\mathbb{Q}$ *for the numeraire* N*. Assume* X *is a price process of a financial asset and* N *is the price process of the numeraire under the EMM* $\mathbb{Q}^N$*.*

$$\frac{X(t)}{N(t)} = E^{\mathbb{Q}^N}\left[\frac{X(T)}{N(T)}|\mathcal{F}_t\right] \tag{2.16}$$

The important conclusion of this equation is that we get an arbitrage free martingale pricing formula for the price process $Y(t, X)$ *of any price process* X*, where we set* $M(T) = \frac{X(T)}{N(T)}$*:*

$$Y(t, X) = N(t)E^{\mathbb{Q}^N}[M(T)|\mathcal{F}_t] \tag{2.17}$$

Definition 2.2.7 ***(Change of Numeraire Technique)***
The numeraire is the unit of measure in which the worth of, in the case of this thesis,

a financial instrument is measured (eg currency, discount bonds, ...). The risk neutral valuation theory uses the "risk-free" bank account as its natural numeraire. In arbitrage free markets we can choose any numeraire which is a traded asset. For the change of numeraire technique we consider two different numeraires N^1 and N^2 on equivalent martingale measures $\mathbb{Q}^{N^1}$ and $\mathbb{Q}^{N^2}$. Since the price of one asset has to be independent of the numeraire, we get the following equality:

$$N^1(t)E^{\mathbb{Q}^{N^1}}\left[\frac{X(T)}{N^1(T)}|\mathcal{F}_t\right] = N^2(t)E^{\mathbb{Q}^{N^2}}\left[\frac{X(T)}{N^2(T)}|\mathcal{F}_t\right]$$

we set $M(T) = \dfrac{X(T)}{N^1(T)}$ *to get*

$$\begin{aligned} E^{\mathbb{Q}^{N^1}}\left[M(T)|\mathcal{F}_t\right] &= \frac{N^2(t)}{N^1(t)}E^{\mathbb{Q}^{N^2}}\left[\frac{X(T)}{N^1(T)}\frac{N^1(T)}{N^2(T)}|\mathcal{F}_t\right] \\ &= E^{\mathbb{Q}^{N^2}}\left[M(T)\frac{N^1(T)}{N^1(t)}\frac{N^2(t)}{N^2(T)}|\mathcal{F}_t\right] \end{aligned}$$

From this equation we see that the expectation of the martingale M under $\mathbb{Q}^{N^1}$ is equal to the expectation of the martingale M multiplied with the Radon Nikodym Derivative or likelihood ratio under $\mathbb{Q}^{N^2}$. Therefore the Radon Nikodym Derivative that transfer the equivalent measures $\mathbb{Q}^{N^2}$ to $\mathbb{Q}^{N^1}$ is:

$$\frac{d\mathbb{Q}^{N^1}}{d\mathbb{Q}^{N^2}} = \frac{N^1(T)}{N^1(t)}\frac{N^2(t)}{N^2(T)} \tag{2.18}$$

Definition 2.2.8 ***(Forward Measure)***
With the definition that each asset divided by a numeraire is martingale under the measure associated with that numeraire, we have $P(t,T_i)$ and $P(t,T_j)$ with their respective measures $\mathbb{Q}^{T_i}$ and $\mathbb{Q}^{T_j}$. We use the Radon Nikodym and numeraire theory from above to get:

$$\begin{aligned} E^{\mathbb{Q}^{T_i}}\left[\frac{d\mathbb{Q}^{T_j}}{d\mathbb{Q}^{T_i}}\right] &= \frac{P(t,T_j)}{P(0,T_j)}\frac{P(0,T_i)}{P(t,T_j)} \\ &= \frac{P(0,T_i)}{P(0,T_j)}\prod_{k=i+1}^{j}\frac{P(t,T_k)}{P(t,T_{k-1})} \end{aligned}$$

If we use the forward rates as function of Zero Bonds from 2.6 and reformulate the *function above we get:*

$$E^{\mathbb{Q}^{T_i}}\left[\frac{d\mathbb{Q}^{T_j}}{d\mathbb{Q}^{T_i}}\right] = \frac{P(0,T_i)}{P(0,T_j)} \prod_{k=i+1}^{j} \frac{1}{1+\delta F(t,T_{k-1},T_k)} \tag{2.19}$$

Definition 2.2.9 ***(Itô's Lemma)***
We assume a vector of processes X following a SDE which we also used in Definition 2.2.2. As a next step we assume a stochastic process $Y(t) = f(t, X(t))$ on the L^2-space, which means that the random variables have finite second moment (twice integrable). Itô's Theorem provides us a similar application like the chain rule in normal calculus to find the differential but it differs due to the term accounting for quadratic covariation $d(X_i, X_j)$. To shorten notation we write f for $f(t, X(t))$.

$$df = \frac{\partial f}{\partial t}dt + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dX_i + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i x_j}(dX_i dX_j) \tag{2.20}$$

where we use the following formal calculation table (a scratch of the proof can be found in Shreve [2004]

$$\begin{cases} (dt)^2 = 0 \\ dt\ dW = 0 \\ (dW)^2 = dt \end{cases}$$

A useful special case of 2.20 is the product formula:

$$dX^1(t)X^2(t) = X^1(t)dX^2(t) + dX^1(t)X^2(t) + [dX^1 dX^2](t)$$

2.3. Monte Carlo Simulation and Computational Aspects

The main source for the following three subsections is the book of Glasserman Glasserman [2004] which is the main reference for Monte Carlo applications in finance.

2.3.1. Monte Carlo Methods

Valuation of complex derivatives by Monte Carlo Simulations involves modeling stochastic paths which should describe the evolution of an interest rate, asset price or other factor which is necessary for the valuation of the derivative. In the case of this work, Monte Carlo Integration is necessary to model the evolution of forward rates which is formalized by a stochastic differential equation.
The underlying principles of Monte Carlo Simulation are the law of large numbers, which is necessary that the estimate converges to the correct value and the central limit theorem which allows us to draw conclusions about the error of the estimate. If for example an asset price is simulated n times via Monte Carlo Simulation and with the price of each simulated path a payoff P is calculated we get n realizations of this payoff P. For this realized payoffs we want to calculate the expectation of the payoff, which is the best estimate for the correct P. This can be accomplished by $\frac{1}{n}\sum_{i=1}^{n} P_i = E[P] = \overline{P}$ where n is a sufficiently large number and index $i \in 1, 2, ..., n$.
The crucial point for the efficiency of the Monte Carlo algorithm is to find a trade-off between speed, which is accomplished in taking a lower number for n and accuracy, which is accomplished in choosing higher values for n. To analyze the cost of speed, the calculation of variance of the Monte Carlo estimate is helpful (we use the linear transformation property which will be described in the next subsection):

$$Var[P] = Var\left[\frac{1}{n}\sum_{i=1}^{n} P_i\right] = \frac{1}{n^2}\sum_{i=1}^{n} Var[P_i] = \frac{1}{n^2}\, n\sigma^2 = \frac{\sigma^2}{n} \rightarrow SD[P] = \frac{\sigma}{\sqrt{n}}$$

where σ^2 is the variance of P which is estimated $\sigma = \sqrt{\sigma^2} \approxeq s_P = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(P_i - \overline{P})^2}$.
As the solution of the last calculation shows, the standard error reduces with the square root of the number of simulations $s_P \frac{1}{\sqrt{n}}$. We have a convergence rate of $\mathcal{O}\left(n^{-\frac{1}{2}}\right)$ which leads to the conclusion that to reach a gain in accuracy of a factor 10 we need to increase the number of simulations by 100. The square root convergence stays true for problems with increasing dimensionality, which is the beautiful side of Monte Carlo simulation.

2.3.2. Random Numbers

The quality of the result out of a Monte Carlo Simulation depends heavily on the quality of the random numbers which are a major input for generating simulations. In the general Monte Carlo application, random numbers are treated as really random to be able to apply theory from statistics and probability theory. Modern pseudorandom number generators are already very good in mimicking real randomness. In modern mathematical software different random number generators are available, which generate random numbers with deterministic algorithms.

This random number generators produce a finite sequence of uniformly distributed numbers $U_1, U_2, ...$between 0 and 1. The important property of the generated random numbers is that they are mutually independent, which means that they are uncorrelated with each other and thaht U_i should not be predictable from $U_1, ..., U_{i-1}$. Another important factor of a random number generator is the ability to generate a series of random numbers fast. The Mersenne Twister, which is used for the LMM simulation, is the standard pseudorandom number generator in R, MATLAB and also available in C++. The algorithm generates high quality pseudorandom numbers, is very efficient and passes most of the tests for statistical randomness.

For most of the applications, the uniformly distributed numbers are transformed into other distributions. Especially in Financial Engineering a lot of simulations require sampling from the standard normal distribution, which is also true for this work.

The notation for a normal distributed random variable with mean μ and variance σ^2 is $X \sim N(\mu, \sigma^2)$. To get normal distributed random variable from simulated standard normal distributed random variables, just the following transformation, which results out of simple probability theory, $X_i = \sigma Z_i + \mu$ is necessary.

The normal distribution is characterized by a μ vector and a $\mathbf{\Sigma}$ covariance matrix which is abbreviated as $N(\mu, \mathbf{\Sigma})$. The covariance matrix $\mathbf{\Sigma}$ has to be positive semidefinite and symmetric. This means that $\mathbf{\Sigma} = \mathbf{\Sigma}^T$ and $\mathbf{x}^T \mathbf{\Sigma} \mathbf{x} \geq 0$ which is eqal to the requirement that the eigenvectors of $\mathbf{\Sigma}$ are nonnegative. For the simulation, standard normal distributed random numbers $Z \sim N(\mathbf{0}, \mathbf{I})$, where the mean is a zero vector and the covariance matrix is a identity matrix, were generated. With the requirements of the covariance matrix fulfilled we can write $\mathbf{\Sigma} = \mathbf{A}\mathbf{A}^\mathbf{T}$ and use the transformation from above to get $\mu + \mathbf{AZ} \sim N(\mu, \mathbf{A}\mathbf{A}^\mathbf{T})$. This transformation allows us to simulate from a general multivariate normal distribution $N(\mu, \mathbf{A}\mathbf{A}^\mathbf{T})$, if we find such a matrix $\mathbf{A}$. This can be done by applying linear algebra, where the Cholesky decomposition is the most frequently used method, which was also used in this work, an alternative method would be the QR decomposition. Another important property of a normal variable or vector is the linear transformation property which tells us that any linear transformation returns a normal variable again $\mathbf{lX} \sim N(\mathbf{l}\mu, \mathbf{l}^\mathbf{T}\mathbf{\Sigma}\mathbf{l})$, for further details see Glasserman [2004]. Singular covariance matrices are not very likely if market data is used for covariance generation but is possible if factor models (with less factors than the dimensions of the covariance matrix) are used to generate a covariance matrix.

2.3.3. Quasi-random numbers and Antithetics

In the last chapter the pseudo-random number generator called Mersenne Twister was discussed, which is a good generator of uniformly distributed random numbers for a sufficiently high number of simulations (Saito and Matsumoto [2008]). There is another possibility to generate random numbers, which are even less random than pseudo random numbers, but have other advantages, this numbers are called quasi-random numbers. In financial engineering a fast convergence rate is very important to reduce the number of simulations and therefore computation time. Quasi-random numbers have the advantage of a high level of uniformity in multidimensional space, but this comes with the disadvantage of not statistically independent random numbers, which is the case for pseudo-random numbers (Levy [2002]). Popular quasi-random numbers are Niederreiter (Niederreiter [1978]), Faure (Faure [1982]) and Sobol sequences, in this work the Sobol (Sobol' [1967]) sequence is tested for its improvement of the rate of convergence. Visual inspection of Figure 2.1 shows that the third picture where the Sobol sequence was used, looks more uniformly distributed than the first picture where random numbers were generated with the Mersenne Twister and the second picture where random numbers were generated with the Mersenne Twister and antithetic variables. It appears that on the first picture, the points are more clustered and not symmetric and on the second picture, the points are also more clustered but symmetric, due to the usage of antithetic variables. The Sobol sequence itself is antithetic as long as the number of iterations of the Monte Carlo Simulation is a power of 2.

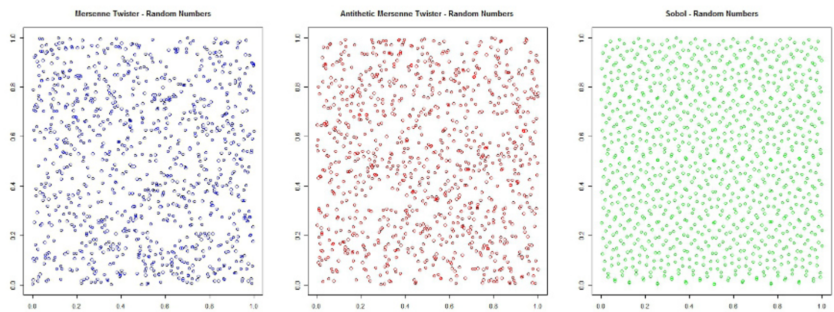

Figure 2.1.: Pseudo- and quasi-random number structure

Antithetic variables are the simplest approach to generate negatively correlated paths. This negative dependence between paths reduces the variance of the Monte Carlo Simulation and therefore increases the convergence rate. This simple approach uses the fact that if the random variable U is randomly distributed, this is also true for the variable $\widetilde{U} = 1 - U$. To generate antithetics for Gaussian distributions, we can use the inverse transform method, where $Z = F^{-1}(U)$ and $\widetilde{Z} = F^{-1}(\widetilde{U})$ are antithetic to each other and have the same distribution. This fact is used to generate antithetic random variables of the normal distribution, where the sequences $(Z_1, ..., Z_n)$ and $(\widetilde{Z}_1, ..., \widetilde{Z}_n)$ are used to generate paths of the Libor forward rate $(L_1, ..., L_n)$ and $(\widetilde{L}_1, ..., \widetilde{L}_n)$ with negative dependence. Perfectly negative correlated pairs $(L_1, \widetilde{L}_1), ..., (L_n, \widetilde{L}_n)$ are generated which can reduce total variance of the Monte Carlo simulation if (using the linear transformation property):

$$
\begin{aligned}
Var[\widetilde{C}] &< Var[C] \\
Var\left[\frac{1}{n}\sum_{i=1}^{n}\frac{C_i + \widetilde{C}_i}{2}\right] &< Var\left[\frac{1}{2n}\sum_{i=1}^{2n} C_i\right] \\
\frac{1}{4n^2} Var\left[\sum_{i=1}^{n} C_i + \widetilde{C}_i\right] &< \frac{1}{4n^2} Var\left[\sum_{i=1}^{2n} C_i\right] \\
Var\left[\sum_{i=1}^{n} C_i + \widetilde{C}_i\right] &< Var\left[\sum_{i=1}^{2n} C_i\right] \\
Var\left[C + \widetilde{C}\right] &< 2Var\left[C\right] \\
Var[C] + Var[\widetilde{C}] + 2Covar[C, \widetilde{C}] &< 2Var\left[C\right] \\
2Var[C] + 2Covar[C, \widetilde{C}] &< 2Var\left[C\right] \\
Covar[C, \widetilde{C}] &< 0
\end{aligned}
$$

where C and $\widetilde{C}$ are estimates for caps and under the assumption that the computational time to generate $\widetilde{Z}$ from Z is a marginal of the computation time of new normal random variables, see Glasserman [2004]. This leads to the fact that we can get less variance for the simulation than in doubling the number of iterations. In the final term of the equation the necessary condition is clearly stated, which is that negative correlation of the inputs produces negative correlation in the output. This condition is not alway given and therefore the use of antithetic variables can also increase the variance of the simulation.

http://www.springer.com/978-3-658-04687-3

Calibration and Parameterization Methods for the Libor Market Model
Hackl, C.
2014, IX, 64 p. 27 illus., Softcover
ISBN: 978-3-658-04687-3