UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

VÝVOJ LIBOR MARKET MODELU

DIPLOMOVÁ PRÁCA

2020 **Máté HÉGLI**

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

VÝVOJ LIBOR MARKET MODELU

DIPLOMOVÁ PRÁCA

Študijný program:	Ekonomicko-finančná matematika a modelovanie
Študijný odbor:	1114 Aplikovaná matematika
Školiace pracovisko:	Katedra aplikovanej matematiky a štatistiky
Vedúci práce:	Mgr. Sándor Kelemen, PhD.

Bratislava 2020 **Máté HÉGLI**

77202680

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Bc. Máté Hégli
Študijný program: ekonomicko-finančná matematika a modelovanie (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: matematika
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: anglický
Sekundárny jazyk: slovenský

Názov: Evolution of LIBOR Market Model
Vývoj LIBOR Market Modelu

Anotácia:

Vedúci: Mgr. Sándor Kelemen, PhD.
Katedra: FMFI.KAMŠ - Katedra aplikovanej matematiky a štatistiky
Vedúci katedry: prof. RNDr. Marek Fila, DrSc.

Dátum zadania: 07.01.2019

Dátum schválenia: 08.01.2019 prof. RNDr. Daniel Ševčovič, DrSc.
garant študijného programu

....
študent vedúci práce

Poďakovanie Touto cestou sa chcem poďakovať svojmu vedúcemu diplomovej práce Mgr. Sándorovi Kelemenovi, PhD. za odborné poznámky a konštruktívne konzultácie aj v dobe pandemickej krízy. Ďakujem tiež firme Zurich Insurance Company za podporu a poskytnuté odborné články.

Abstrakt v štátnom jazyku

HÉGLI, Máté: Vývoj LIBOR Market Modelu [Diplomová práca], Univerzita Komenského v Bratislave, Fakulta matematiky, fyziky a informatiky, Katedra aplikovanej matematiky a štatistiky, Vedúci: Mgr. Sándor Kelemen, PhD., Bratislava, 2020, 96s.

V našej práci skúmame LIBOR Market Modely, ktoré oproti starším stochastickým modelom úrokovej miery popisujú dynamiku forwardových úrokových mier. Kvôli relatívne presnej kalibrovateľnosti k trhovým cenám capov a swapcií sú tieto modely veľmi rozšírené, obľúbené a tiež často používané aj na oceňovanie exotických derivátov úrokovej miery.

V súčasnej dobe môžu úrokové miery nadobudnúť záporné hodnoty s kladnou pravdepodobnosťou, preto namiesto štandardného lognormálneho LMM modelu, t.j. BGM modelu, skúmame jeho rozšírenia pridaním parametra posunutia a stochastickým modelovaním volatility úrokových mier. Uvedieme metodiku modelu a jeho kalibračný algoritmus navrhnutý Piterbargom & Andersenom, a predstavíme model LMMPlus, ktorý je v praxi používaný na oceňovanie finančných záväzkov v životnom poistení. Na základe predpokladov LMMPlus modelu odvodíme oceňovaciu formulu swapcií vyjadrenú pomocou momentových vytvárajúcich funkcií, a predstavíme kalibračnú metódu modelu k trhovým cenám swapcií.

Kľúčové slová: stochastický model úrokovej miery, deriváty úrokovej miery, LIBOR Market Model, LMMPlus, záporná úroková miera, stochastický model volatility, numeraire zmena, momentová vytvárajúca funkcia, cena swapcie

Abstract

HÉGLI, Máté: Evolution of LIBOR Market Model [Master Thesis], Comenius University in Bratislava, Faculty of Mathematics, Physics and Informatics, Department of Applied Mathematics and Statistics, Tutor: Mgr. Sándor Kelemen, PhD., Bratislava, 2020, 96p.

In this thesis, we investigate the LIBOR Market Models, which opposed to older stochastic interest rate models describe the dynamics of forward interest rates, that are observable on markets. These models are widely known and popular, because they can be relatively accurately calibrated to market cap and swaption prices and used for pricing exotic interest rate derivatives.

Today, interest rates can take on negative values with positive probability, therefore instead of the standard lognormal LMM model, i.e. the BGM model, we examine its extension by a displacement element and a stochastic volatility process. We present the methodology of a model and its calibration proposed by Piterbarg & Andersen and introduce the LMMPlus model, which is used in practice for the valuation of life insurance liabilities. We also derive a pricing formula for swaptions – expressed in terms of moment generating functions – under the assumptions of the LMMPlus model and present the model's calibration algorithm to market swaption prices.

Keywords: stochastic interest rate models, interest rate derivatives, LIBOR Market Model, LMMPlus, negative interest rate, stochastic volatility model, change of numeraire, moment generating function, swaption pricing

Contents

Introduction

Interest rate is one of those rare financial terms that the majority of population is familiar with. That is mainly because lending money and acquiring interest from the borrower date back several thousand years. By rewarding the act of lending with an amount of money at a certain rate, we avoid the issues of the fact that today's money is not equivalent to tomorrow's. This is the reason costumers pay interest to borrow money from a bank and gain interest on money deposited in a bank account.

Interest rate derivatives are financial instruments whose payoffs depend on the movement of interest rates. A few decades before, the number of these derivatives, such as caps or swaptions started to rapidly increase and finding proper procedures for their pricing became substantial. First efforts were made in the '70s and '80s when the premier time-homogeneous short-rate models where presented, e.g. the Vasicek (1977), Dothan (1978) or the Cox, Ingersoll and Ross model (1983). These one-factor models gave the possibility of pricing bonds and some vanilla options analytically, however, they also had some major drawbacks. For example, due to the small number of parameters in the diffusion process, they could not reproduce zero-coupon bond curves properly. They were easy to calibrate, however, since the shape of a term structure they offered was predetermined by some basic categories, their calibration to the initial zero-coupon bond curve was inaccurate. They also modelled the short-rate, which in fact is not observable on the market. Still, there is a wide range of short-rate models (see [1]) used in practice: either one-factor – assuming constant or time-dependent parameters – or two-factor models, which turned out to be more precise, especially when correlation between two rates plays relevant role. The first non-short-rate model was presented one year before the Vasicek model. Fischer Black published a paper called *The Pricing of Commodity Contracts* [2] introducing the model known as *Black-76 model*, which was "mimicking the Black and Scholes model" [1]. The first important alternative to short-rate models was the framework developed by Heath, Jarrow and Morton (1992), who chose to model the instantaneous forward interest rate. The demand constantly increased not only for standard interest rate derivatives but for a new type, the exotic interest rate derivatives, too. Pricing and hedging these derivatives required a novel approach. Also, the problem of the above-mentioned techniques – modelling interest rates that are not observable in the market – was still unsolved. The change came by Alan Brace, Dariusz Gatarek and Marek Musiela introducing the *BGM model* (1997) [3], which describes the dynamics of an interest rate directly observable in markets, i.e. the LIBOR forward rate and can deal with exotic interest rate derivatives. The BGM model has several extensions having some additional assumptions. Together, they are also known as LIBOR Market Models (LMM models).

There are several books, papers and theses discussing not only the BGM model (see [4] or [5]) and its extensions (see [1]), but their calibrations (see [6]), too. However, many of these works were written before the 2008 financial crisis, therefore most of market model extensions were not designed to handle negative interest rates. Since then, negative rates have become ordinary phenomenon in the financial market, which meant new, practicable models had to be invented allowing interest rates to reach negative values. Hence, we introduce the LMMPlus model, which is an extension of the BGM model and allows interest rates acquire negative values, furthermore, assumes a stochastic volatility model to capture the (realistic) stochastic behavior of volatility.

Our aim in this thesis is to create a comprehensive but comprehensible, easily readable guide, which enables the reader to understand LMM models in context. Therefore, besides presenting the required mathematical and financial apparatus we place an important emphasis on explaining arising ambiguities. For that, we often refer to Brigo & Mercurio (see [1]) or Piterbarg & Andersen (see [7] and [8]).

In Chapters 1 and 2, we introduce the required mathematical and economical background and define the most fundamental financial terms used throughout the thesis. We also present the zero-coupon bond together with a set of interest rates considered in this thesis. In Chapter 3 we explain the Change of Numeraire technique, which allows us to change probability spaces and measures without damaging our original market assumptions and we use it to define the forward measure. This technique will also prove its value later in Chapter 5, where we derive the forward LIBOR rate dynamics in several measures, introduce the general methodology of LMM models and a concrete model, used in practice, the LMMPlus model. In Chapter 4 we present a special derivative contract called the interest rate swap and the option (swaption) that allows to enter into a swap contract. Since the LMMPlus model is calibrated to market swaption prices, Chapter 6 is dedicated to the derivation of the swaption pricing formula expressed in terms of moment generating functions. Finally, in Chapter 7 we introduce a general grid-based calibration method proposed by Piterbarg & Andersen and a more concrete, a more easily applicable calibration algorithm of the LMMPlus model.

1 Background

In this Chapter, we provide some economical and mathematical assumptions that we will follow throughout the thesis. We introduce some fundamental definitions and build up the necessary background to define the risk-neutral probability measure.

Consider an economy where non-dividend paying securities are traded continuously in a finite horizon $[0, T]$. Assume a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ – a threesome of a sample space with outcome elements ω, a σ-algebra and a probability measure – on which prices of these securities are defined and a corresponding filtration $\mathbb{F} = \{\mathcal{F}_t : 0 \leq t \leq T\}$, where $\mathcal{F}_s \subseteq \mathcal{F}_t$ for $s \leq t$. We can think of $\mathcal{F}_t$ as the information at time t and assume, that the stochastic process of any security price is adapted to $\mathcal{F}_t$.

We also assume the *absence of arbitrage*, which simply put, disables the opportunity to invest zero today and acquire profit from that investment tomorrow. In an arbitrage-free economy the prices of two investments or financial instruments with the same payoff at time T must be equal at any time $t \leq T$. Otherwise an arbitrage opportunity arises, so that one can buy the instrument of lower value and sell the one of higher value, simultaneously. Another important assumption is, that the considered economy is *complete*, meaning that every contingent claim is replicable, or as Brigo & Mercurio stated: "a financial market is complete if and only if every contingent claim is attainable" [1].

Definition 1 Self-Financing Strategy [1] *Consider a stochastic process of security prices* $S = \{S_t : 0 \leq t \leq T\}$*, where* $S_t = (S_t^1, S_t^2, \dots, S_t^m)^\top$ *and a trading strategy process* $\phi = \{\phi_t : 0 \leq t \leq T\}$*, where* $\phi_t = (\phi_t^1, \phi_t^2, \dots, \phi_t^m)^\top$ *is a vector with predictable and locally bounded components. Define the value process to the strategy* ϕ *as*

$$V_t(\phi) = {\phi_t}^\top S_t$$

and the gains process to the strategy ϕ *as*

$$G_t(\phi) = \int_0^t {\phi_u}^\top dS_u.$$

A trading strategy ϕ *is self-financing if* $V(\phi) \geq 0$ *and its value process satisfies the equation*

$$V_t(\phi) = V_0(\phi) + G_t(\phi).$$

Definition 2 Attainable Continent Claim [1] *A contingent claim is attainable if there exists a self-financing trading strategy – defined for* $0 \leq t \leq T$ *– such that its value at time* T *equals to the value of the contingent claim.*

For a more detailed explanation of the arbitrage-free property and the completeness of economy see [1]. Brigo & Mercurio, authors of this book also discuss the connection between these economical properties and the mathematical property of existence of a unique *risk-neutral measure*, which is the result of works of Harrison & Kreps [9] and Harrison & Pliska ([10], [11]). They, in fact, proved that the economy is arbitrage free and complete if and only if there exists a unique risk-neutral measure. In addition, it not only ensures that the economy meets the requirements but also allows us to define the price of a derivative as the conditional expected value of its discounted future payoff.

However, before defining the risk-neutral measure we must demystify some other expressions used in the last paragraph. At first, we introduce the *bank account,* by which we refer to a riskless investment, where profit is accrued continuously at the risk-free rate.

Definition 3 Bank Account [1] *Let* $\beta(t)$ *be the value of a bank account for* $t \geq 0$. *We assume that* $\beta(t)$ *develops by the following differential equation:*

$$d\beta(t) = r_t \beta(t) dt,$$

where r_t *is a function of time, and satisfies* $\beta(0) = 1$.
By solving the differential equation, it is clear, that

$$\beta(t) = e^{\int_0^t r_s ds}.$$

The instantaneous rate r_t, at which the value of the bank account grows is usually referred to as the *instantaneous spot rate*, or briefly as *short-rate*. Unlike some other pricing methods, such as the Black-Scholes formula used for pricing vanilla options, the short-rate r_t is not deterministic. Therefore, in this thesis by the evolving short-rate we mean the evolution of r through a *stochastic process*. Naturally, the bank account and the discount factor (see below) are also stochastic processes.

Definition 4 Discount Factor [1] *The discount factor* $D(t,T)$ *is the amount at time* t *that is "equivalent" to one unit of currency payable at time* T. *It is given by*

$$D(t,T) = \frac{\beta(t)}{\beta(T)} = e^{-\int_t^T r_s ds} .$$

The equation above can easily be deducted from the fact, that if we wish to have one unit of currency at time T, we should invest the amount of $1/\beta(T)$. Hence, the amount at time t is going to be the discount factor $D(t,T)$. Now, that we have defined all the tools needed, we can pursue in defining the risk-neutral probability measure.

Definition 5 Risk-Neutral Measure [1] *The risk-neutral measure (or equivalent martingale measure)* $\mathbb{Q}$ *is a probability measure such that*

i) $\mathbb{Q}$ *and* $\mathbb{P}$ *are equivalent measures;*

ii) the Radon-Nikodym derivative $d\mathbb{Q}/d\mathbb{P}$ *belongs to* $L^2(\Omega, \mathcal{F}, \mathbb{P})$*;*

iii) the discounted price of a contingent claim $D(0,t)V(t)$ *is a martingale under* $\mathbb{Q}$*.*

For the definition of equivalent measures and the Radon-Nikodym derivative we refer the reader to Appendix A. Note, from the third point of *Definition 5*, we can define the price of a contingent claim as the discounted conditional expectation of its payoff under the risk-neutral measure $\mathbb{Q}$.

Definition 6 The Price of an Attainable Contingent Claim *Assume there exists a risk-neutral measure. From its third property (Definition 5) it holds that* $V(t)/\beta(t)$ *is a martingale. Hence, the following holds for all* $T > t$*:*

$$V(t) = \beta(t)\mathbb{E}^{\mathbb{Q}}\left[\frac{V(T)}{\beta(T)}\middle| \mathcal{F}_t\right], \tag{1}$$

where $V(t)$ *is the price of an arbitrary contingent claim,* $\mathbb{E}^{\mathbb{Q}}$ *is the conditional expectation under measure* $\mathbb{Q}$ *and* $\mathcal{F}_t$ *is the corresponding filtration.*

For future reference, we simplify the notation of filtration $\mathcal{F}_t$ by writing the conditional expectation as $\mathbb{E}_t^{\mathbb{Q}}$.

Since a martingale is a zero-drift stochastic process, the "martingale property" of normalized asset prices defined in *Definition 6* will often be used in next chapters, when deriving the interest rate models.

Before moving on to Chapter 2, we introduce a possible way – following mainly [5] – of finding the risk-neutral measure and deriving the equation from *Definition 6*. Assume dynamics of an arbitrary financial security price $S(t)$ as

$$dS(t) = \mu(t)S(t)dt + \sigma(t)S(t)dW(t), \tag{2}$$

where $\mu(t)$ is the drift of the process, $\sigma(t)$ is the volatility of the price and $W(t)$ is *Wiener process*.

Let (ϕ, ψ) be a self-financing strategy with its *value process*, which we define by

$$X(t) = \phi_t\beta(t) + \psi_t S(t),$$

where $\beta(t)$ is the value of a bank account. Hence, the following must hold:

$$\begin{aligned} dX(t) = \phi_t d\beta(t) + \psi_t dS(t) &= \phi_t r_t \beta(t)dt + \psi_t[\mu(t)S(t)dt + \sigma(t)S(t)dW(t)] \\ &= r_t[X(t) - \psi_t S(t)]dt + \psi_t\mu(t)S(t)dt + \psi_t\sigma(t)S(t)dW(t) \\ &= r_t X(t)dt + \psi_t\sigma(t)S(t)\left[\frac{\mu(t) - r_t}{\sigma(t)}dt + dW(t)\right]. \end{aligned}$$

By choosing a new Wiener process $W^{\mathbb{Q}}(t)$, such that

$$dW^{\mathbb{Q}}(t) = \frac{\mu(t) - r_t}{\sigma(t)}dt + dW(t), \tag{3}$$

we can rewrite the latest form of the value process dynamics as

$$dX(t) = r_t X(t)dt + \psi_t\sigma(t)S(t)dW^{\mathbb{Q}}(t).$$

All we need to do, is to show that measure $\mathbb{Q}$, under which $W^{\mathbb{Q}}(t)$ is a Wiener process satisfies the properties of the risk-neutral measure from *Definition 5*. For that, we can simply apply the Girsanov's Theorem I (see Appendix B) on

$$W^{\mathbb{Q}}(t) = \int_0^t \frac{\mu(s) - r_s}{\sigma(s)}ds + W(t),$$

the integral of Equation (3). Assuming $\frac{\mu(t)-r_t}{\sigma(t)}$ is a bounded process – such that it satisfies the Novikov's condition – the stochastic process $W^{\mathbb{Q}}(t)$ is a Wiener process on the probability space $(\Omega, \mathcal{F}, \mathbb{Q})$, furthermore measures $\mathbb{Q}$ and $\mathbb{P}$ are equivalent and their Radon-Nikodym derivative $d\mathbb{Q}/d\mathbb{P}$ is from $L^2(\Omega, \mathcal{F}, \mathbb{P})$.

Finally, we need to prove, that the discounted value process is martingale under $\mathbb{Q}$. By using the multiplication rule of the Itô's Lemma (see Appendix B) we get

$$d\left(\frac{X(t)}{\beta(t)}\right)= d\left(X(t)\frac{1}{\beta(t)}\right)=-\frac{X(t)}{\beta^2(t)}d\beta(t)+\frac{1}{\beta(t)}dX(t)$$

$$=-\frac{r_tX(t)}{\beta(t)}dt+\frac{r_tX(t)}{\beta(t)}dt+\psi_t\sigma(t)\frac{S(t)}{\beta(t)}dW^{\mathbb{Q}}(t)=\frac{S(t)}{\beta(t)}\psi_t\sigma(t)dW^{\mathbb{Q}}(t).$$

Assuming $\frac{\psi_t\sigma(t)S(t)}{\beta(t)}$ is a locally bounded process, the discounted process value is a martingale under $\mathbb{Q}$, since the drift of the process equals to zero. Hence,

$$V(t)=\beta(t)\mathbb{E}_t^{\mathbb{Q}}\left[\frac{V(T)}{\beta(T)}\right],$$

where V denotes the price of an arbitrary contingent claim associated with the self-financing strategy above.

The previous approach reveals another property, which later – when deriving LMM models – will be substantial. Note, that we moved from the real-life world (defined under measure $\mathbb{P}$) to the risk-neutral world (defined by $\mathbb{Q}$). Which leads us to the technique called the "*change of numeraire*". Using that – similarly to the risk-neutral measure – many other probability measures can be defined, while preserving the fundamental properties of contingent claims. We will discuss this later in Chapter 3.

2 Zero-Coupon Bond and Interest Rates

In the Introduction, we pointed out that most of the interest rate models consider the dynamics of short-rates. However, we also mentioned the HJM framework modelling the instantaneous forward interest rate and LMM models describing the evolution of LIBOR rates. In this Chapter, we introduce the zero-coupon bond and its price at time t, using which we define some of the spot interest rates. We also explain the Forward Rate Agreement, through which we can derive the definition of forward interest rates.

2.1 Zero-Coupon Bond

Definition 7 Zero-Coupon Bond [1] *A T-maturity zero-coupon bond is a contract that guarantees the payment of one unit of currency at time T. We denote the value of the contract at time $t < T$ as $P(t,T)$. Obviously, $P(T,T) = 1$ for any T.*

This definition seems to be suspiciously similar to the definition of discount factor from Chapter 1. However, the discount factor is a random quantity depending on the short-rate r_t, which – in the interest rate theory – is not deterministic, while the zero-coupon bond price is the value of a contract that needs to be known at time t. Still, there is a close relationship between these two quantities, which is shown in the following Lemma.

Lemma 1 Zero-Coupon Bond Price *Assuming there exists a risk-neutral measure $\mathbb{Q}$ defined by Definition 5, the price of a T-maturity zero-coupon bond at time t is given by*

$$P(t,T) = \mathbb{E}_t^{\mathbb{Q}}\left[D(t,T)\right].$$

Proof. It holds, that $P(T,T) = 1$. Now recall the definition of the risk-neutral measure, where $V(t)$ denotes the price of an asset, in this case the zero-coupon bond price. Hence, by a simple substitution $V(t) = P(t,T)$ in Equation (1) we have

$$P(t,T) = \beta(t)\mathbb{E}_t^{\mathbb{Q}}\left[\frac{1}{\beta(T)}\right] = \mathbb{E}_t^{\mathbb{Q}}[D(t,T)]\,.$$

∎

Definition 8 Zero-Bond Curve [1] *The zero-bond curve at time t is the graph of the function*

$$T \longmapsto P(t,T), \qquad T > t,$$

which considering the positivity of interest rates should (see below) be decreasing in T.

In *Definition 8*, we refer to the book [1] of Brigo & Mercurio, where the zero-bond curve (or the term-structure of discount factors) is considered to be T-decreasing, due to the positivity of interest rates (see the next Section to understand the relationship between the zero-coupon bond and interest rates). However, our experience in practice claims otherwise – in the past few years negative interest rates became common, which resulted non-monotonic zero-bond curves. As we can see in Figure 1, the term-structure of discount rates for the last quartal of 2019 used by the Zurich Insurance Company Ltd. is not decreasing. We use these bond prices also in Section 2.2 to express interest rates in terms of the zero-coupon bond price.

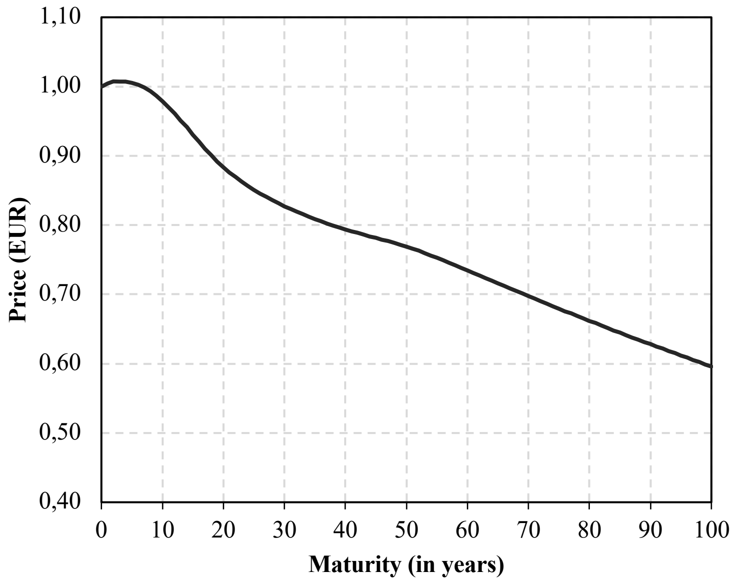

Figure 1 *Zero-Bond Curve*

2.2 Spot Interest Rates

As we mentioned earlier, any interest rate can be defined in terms of the zero-coupon bond price. Which also means, that the zero-coupon bond price can be recovered in many ways, depending on the considered type of *compounding*. Here, we present two types of spot interest rates: the continuously compounded – where an investment accrues continuously – and the simply compounded – at which an investment accrues proportionally to time. We also consider a third type, the annually-compounded interest rate, but only to be able to define the zero-coupon curve. For more, see [1].

Interest rates are characterized by the *time to maturity* of zero-coupon bonds. For that, we use the notation from the following definition.

Definition 9 Time Measure *By* $\tau(t,T)$ *we denote the time elapsed between dates* t *and* T *in years. For a time structure (or tenor structure), i.e. a set of predefined dates* $\{T_i\}_{0\le i\le N}$, *we simplify the notation to* $\tau_i = \tau(T_i, T_{i+1})$. *An interval* $\tau(t,T)$ *is known as both a tenor, when* T *is the expiry date of a contract or the time to maturity, when* T *is the maturity of a bond.*

Definition 10 Continuously-Compounded Spot Interest Rate [1] *Continuously-compounded spot interest rate at time* t *with maturity* T *is defined by*

$$R(t,T) = -\frac{\ln P(t,T)}{\tau(t,T)}.$$

From the equation above, the bond price in terms of the continuously-compounded rate $R(t,T)$ is the following:

$$P(t,T) = e^{-R(t,T)\tau(t,T)}.$$

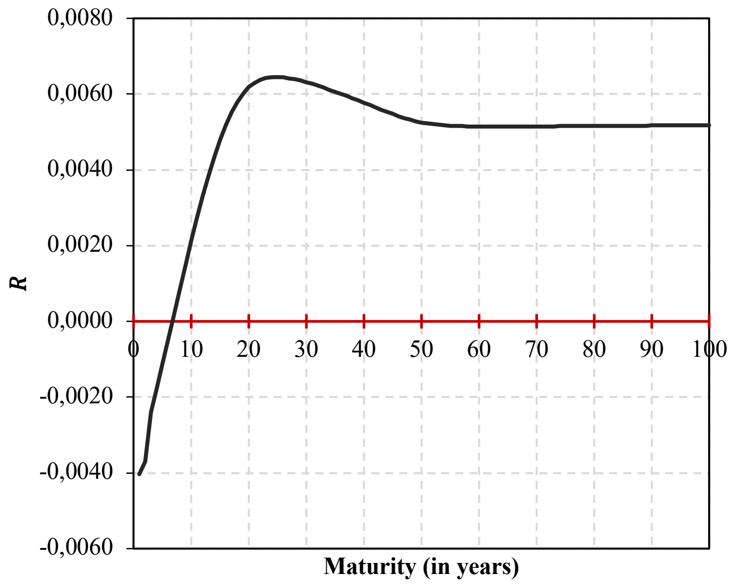

Figure 2 *Continuously-Compounded Spot Interest Rate*

Definition 11 Simply-Compounded Spot Interest Rate [1] *The simply-compounded spot interest rate at time* t *with maturity* T *is defined by*

$$L(t,T) = \frac{1 - P(t,T)}{\tau(t,T)P(t,T)}. \tag{4}$$

The most widely used and the most important interbank benchmark interest rate is the LIBOR (London Interbank Offered Rate) rate, which is a simply-compounded interest rate – hence, the notation $L(t,T)$. It is fixed in London each day by calculating the *trimmed mean* of interest rates suggested by panel banks in London. In fact, there are 35 different LIBOR rates, since they are worked out by the five main currencies of the financial markets and 7 different maturities from overnight to one year.

Again, the bond price can be expressed in terms of LIBOR as

$$P(t,T) = \frac{1}{1 + L(t,T)\tau(t,T)}.$$

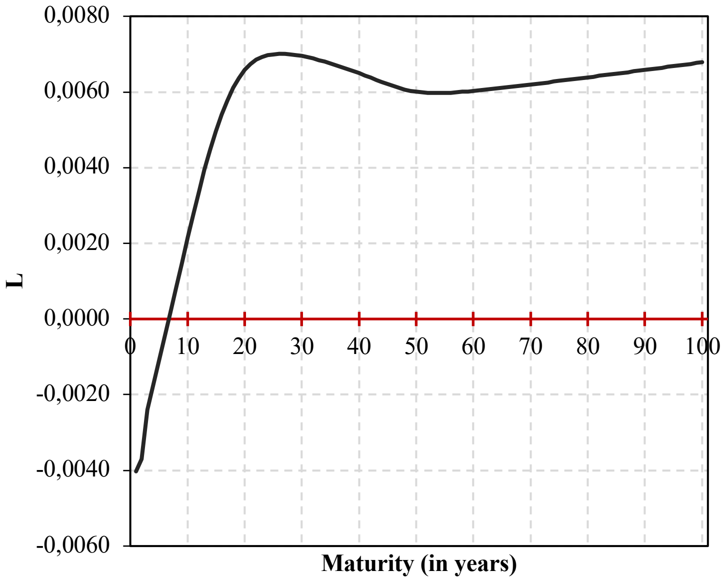

Figure 3 *Spot LIBOR Rates*

2.3 Forward Interest Rates

The forward rate can be easily defined through a Forward-Rate Agreement, which is characterized by three time instants: the current time t, the expiry time T and the maturity S, with $t < T < S$. The holder of the contract gets an interest-rate payment for the period between dates T and S. At the maturity S, a fixed payment is exchanged against a floating payment, while the former is based on a fixed rate K and the latter on the spot rate $L(T,S)$. Assuming the contract nominal value N, the value of the contract at maturity S is

$$N\tau(T,S)\big(K - L(T,S)\big). \tag{5}$$

By substituting $L(T,S)$ defined by (4), the value can be rewritten as

$$N\left(\tau(T,S)K - \frac{1}{P(T,S)} + 1\right). \tag{6}$$

Considering just the term $1/P(T,S)$ as an amount of currency at maturity time S, we can multiply it by the zero-coupon price $P(T,S)$ to obtain its value at T. Clearly, the result is one unit of currency at time T. By moving backwards in time, one can admit, at time t it is equivalent to an amount of $P(t,T)$. By multiplying the remaining two terms by $P(t,S)$ we obtain their t-time value equivalent, so we can put down the value of the contract at time t as:

$$\boldsymbol{FRA}(t,T,S,\tau(T,S),N,K) = N[P(t,S)\tau(T,S)K - P(t,T) + P(t,S)]\,.$$

Assuming there is no cash exchange defined in the contract, it is fair only if the value at time t is zero. Hence, by equating the FRA value to zero, we can find the only K, that renders the contract to be fair. The obtained value leads us to the simply-compounded forward rate.

Definition 12 Simply-Compounded Forward Interest Rate [1] *The simply-compounded forward interest rate $F(t,T,S)$ prevailing at time t for the expiry T and maturity S, with $t < T < S$ is defined by*

$$F(t,T,S) = \frac{1}{\tau(T,S)}\left(\frac{P(t,T)}{P(t,S)} - 1\right). \tag{7}$$

Now recall the value of the aforementioned FRA. By a simple rearrangement of terms in the parenthesis we can easily rewrite it by means of our newly defined interest rate:

$$\boldsymbol{FRA}(t,T,S,\tau(T,S),N,K) = NP(t,S)\tau(T,S)[K - F(t,T,S)]\,.$$

As we can see, the value of the FRA at time t is very similar to its payoff at maturity time S. By replacing the LIBOR rate $L(T,S)$ in Equation (5) with the forward rate $F(t,T,S)$ and taking the present value of the previously compiled payoff we can easily value a FRA. Notice also, that from Equations (4) and (7) the following relation holds:

$$F(T,T,T+\tau) = L(T,T+\tau). \tag{8}$$

Later, in Chapter 3 we will show, that the forward rate $F(t,T,S)$ is in fact an estimate of the spot rate $L(T,S)$ as its expectation at time t in a special probability measure.

Another type of rates, analogous of the short-rate in the future, the instantaneous forward rate can be obtained when the maturity S of a forward rate approaches the expiry time T.

Definition 13 Instantaneous Forward Interest Rate [1] *The instantaneous forward interest rate $f(t,T)$ prevailing at time* t *for maturity $T > t$ is defined as*

$$f(t,T) = \lim_{S \to T^+} F(t,T,S) = -\frac{\partial \ln P(t,T)}{\partial T}.$$

3 Change of Numeraire

As we already indicated when explaining risk-neutral measure in Chapter 1, the change of numeraire is a technique used to easily move from one probability measure to another. Although the price of a contingent claim H is obtainable as the expectation of its discounted payoff, a risk-neutral measure $\mathbb{Q}$ "is not necessarily the most natural and convenient measure for pricing the claim H " [1]. For example, LIBOR forward rates do not have a lognormal distribution in the risk-neutral measure $\mathbb{Q}$, but as we show in Chapter 5, by using the *Change of Numeraire* technique we can define a probability measure for each forward LIBOR rate, in which the property of lognormality holds.

A *numeraire* is an asset that is used as a reference to normalize other asset prices. By definition, any non-dividend paying asset can be used as a numeraire. Recall, for example, the definition of the risk-neutral measure and its "martingale-property". We proved, that the price of an arbitrary contingent claim – or rather the value of a self-financing strategy replicating it – normalized by $\beta(t)$ is a martingale under the measure $\mathbb{Q}$. In this case, we considered the bank account as the numeraire inducing measure $\mathbb{Q}$.

The *Change of Numeraire Theorem* was proposed and proved by Geman, El Karoui and Rochet [12]. It states that we can arbitrarily change numeraires and that there exists a probability measure for any of them such that the price of any other asset normalized by the numeraire is a *martingale* under the corresponding measure. An important assumption of the theorem was also proved by them, namely every self-financing strategy remains self-financing after a numeraire change, which implies that every contingent claim remains contingent claim after a numeraire change.

From our perspective, the most important is the following proposition by Geman et al. introduced in [1], that generalizes *Definition 5* to any numeraire.

Proposition 1 [1] *Consider a numeraire N and a probability measure Q^N, equivalent to the initial $\mathbb{P}$, such that the price of any traded asset X (without intermediate payments) relative to N is a martingale under Q^N, i.e.,*

$$\frac{X_t}{N_t} = \mathbb{E}^N\left[\frac{X_T}{N_T}\middle|\mathcal{F}_t\right] \quad 0 \leq t \leq \mathrm{T}.$$

Let U be an arbitrary numeraire. Then there exists a probability measure Q^U, equivalent to the initial $\mathbb{P}$, such that the price of any attainable claim Y normalized by U is a martingale under Q^U, i.e.,

$$\frac{Y_t}{U_t} = \mathbb{E}^U\left[\frac{Y_T}{U_T}\middle|\mathcal{F}_t\right] \quad 0 \le t \le \mathrm{T}.$$

Moreover, the Radon-Nikodym derivative defining the measure Q^U is given by

$$\frac{dQ^U}{dQ^N} = \frac{U_T N_0}{U_0 N_T}. \tag{9}$$

Proof. We provide a short sketch of the proof, especially for Equation (9). The first part of the proposition, i.e. the existence of the probability measure Q^U induced by numeraire U, together with martingale property of a normalized attainable claim can be proved by using the Bayes rule for conditional expectations. For that, we refer the reader to [12].

The second part, i.e. Equation (9) can be obtained by using properties of the two numeraires N and U, and a heuristic proof is introduced in [1]. Note that fraction $\frac{Z_0}{N_0}$ – where Z is an arbitrary traded asset – can be expressed in two ways, which therefore must be equal:

$$\mathbb{E}^N\left[\frac{Z_T}{N_T}\right] = \mathbb{E}^U\left[\frac{U_0}{N_0}\frac{Z_T}{U_T}\right].$$

From the definition of the Radon-Nikodym derivative (see Appendix A) we know, that

$$\mathbb{E}^N\left[\frac{Z_T}{N_T}\right] = \mathbb{E}^U\left[\frac{Z_T}{N_T}\frac{dQ^N}{dQ^U}\right].$$

Hence – according Brigo & Mercurio – by comparing the two right-hand sides

$$\frac{dQ^N}{dQ^U} = \frac{N_T U_0}{N_0 U_T},$$

which leads us directly to Equation (9).

■

Now, we know that a numeraire can be any non-dividend paying security and we also showed that there exists a unique probability measure for every such a security. Therefore, we can assume a zero-coupon bond being another numeraire with its corresponding probability measure, the T-forward risk-adjusted measure, or simply the T-forward measure.

3.1 *T*-forward Measure

Definition 14 *T*-Forward Measure *If we denote the price of an arbitrary contingent claim by V, the T-forward measure $\mathbb{Q}^T$ induced by numeraire $P(t,T)$ is the measure under which $V(t)/P(t,T)$ is a martingale. Hence the following holds for all $T > t$:*

$$V(t) = P(t,T)\mathbb{E}_t^T\left[\frac{V(T)}{P(T,T)}\right] = P(t,T)\mathbb{E}_t^T[V(T)],$$

where $\mathbb{E}_t^T$ is the conditional expectation in the T-forward measure $\mathbb{Q}^T$ and corresponding to the filtration $\mathcal{F}_t$.

In Chapter 2, we already mentioned that the forward rate is the estimate of the LIBOR spot rate. Now, that we are conscious of the existence of $\mathbb{Q}^T$, we can show that it is the conditional expectation of LIBOR rate under the measure $\mathbb{Q}^T$. For that, first we need to prove that $F(t,T,T+\tau)$ is a martingale under $\mathbb{Q}^{T+\tau}$.

Lemma 2 Martingale Property of $\boldsymbol{F(t,T,T+\tau)}$ *The forward rate $F(t,T,T+\tau)$ is a martingale under the $(T+\tau)$-forward measure $\mathbb{Q}^{T+\tau}$. In formulas:*

$$F(t,T,T+\tau) = \mathbb{E}_t^{T+\tau}[F(s,T,T+\tau)], \qquad t \le s \le T,$$

where $\tau > 0$.

Proof. Recall the definition of the forward rate,

$$F(t,T,T+\tau) = \frac{1}{\tau}\left(\frac{P(t,\ T)}{P(t,T+\tau)} - 1\right).$$

To show that $F(t,T,T+\tau)$ is a martingale under the $(T+\tau)$-forward measure, we need to prove that:

$$F(t,T,T+\tau) = \mathbb{E}_t^{T+\tau}[F(s,T,T+\tau)],$$

for any $t \le s \le T$. We can proceed as follows:

$$\mathbb{E}_t^{T+\tau}[F(s,T,T+\tau)] = \mathbb{E}_t^{T+\tau}\left[\frac{1}{\tau}\left(\frac{P(s,T)}{P(s,T+\tau)} - 1\right)\right] = \frac{1}{\tau}\left(\mathbb{E}_t^{T+\tau}\left[\frac{P(s,T)}{P(s,T+\tau)}\right] - 1\right).$$

We know, that $P(s,T)$ is a security and an arbitrary traded security normalized by a numeraire is a martingale under the corresponding measure. Therefore, $P(s,T)/P(s,T+\tau)$ is a martingale under the $(T+\tau)$-forward measure. Hence,

$$\mathbb{E}_t^{T+\tau}[F(s,T,T+\tau)] = \frac{1}{\tau}\left(\frac{P(t,\ T)}{P(t,T+\tau)} - 1\right) = F(t,T,T+\tau).$$

■

Corollary 1 *The forward rate $F(t,T,S)$ is the expectation of the spot rate $L(T,S)$ in the S-forward measure $\mathbb{Q}^S$. In formulas:*

$$F(t,T,S) = \mathbb{E}_t^S[L(T,S)],$$

where $t \le T \le S$.

Proof. From *Lemma 2* we already know, that forward rate $F(t,T,S)$ is a martingale under the S-forward measure $\mathbb{Q}^S$. Also, if we recall Equation (8) from Chapter 2, the following must hold:

$$F(T,T,S) = L(T,S).$$

Hence, the proof of the corollary is straightforward:

$$F(t,T,S) = \mathbb{E}_t^S[F(T,T,S)] = \mathbb{E}_t^S[L(T,S)]\,.$$

■

4 Swaps and Swaptions

In this Chapter, we introduce the interest rate swap and one of the most important interest derivatives, the swaption. We start by explaining and pricing the plain vanilla interest rate swap, however, in the end, our focus is on pricing the swaption, i.e. the option contract of a swap. Swaptions, together with other interest rate derivatives like caps are the most liquid and the most frequently traded derivatives on the interest rate derivative market. Therefore, in most cases LMM models are calibrated to their market prices. Since the LMMPlus model calibration methodology considers only swaptions as model target instruments, in this thesis we left out of consideration caps and their pricing formulas. However, we refer the reader to [1] or [13] where they are explained in detail.

4.1 The Plain Vanilla Interest Rate Swap

A plain vanilla interest rate swap, or – as we refer to it – a swap can be seen as a generalization of the FRA. It is an agreement to exchange payments between two legs – fixed and floating – at predefined indexed time dates. To remain simple and ease the notation in this thesis, we consider that payments of these legs occur at the same time. We derive the payoff and the price of a swap by following mainly [1] and [14].

Consider a set of predefined time dates $0 \leq T_0 < T_1 < \cdots < T_N$ starting at a given date T_0. Since in the previous Chapter we have derived some of the forward rates' properties and we have defined a discrete-time structure, it is reasonable to use the following new notation of spot LIBOR rates:

$$L(T_n, T_{n+1}) = L_n(T_n),$$

where $L_n(t)$ is the forward (LIBOR) rate defined as

$$L_n(t) = F(t, T_n, T_{n+1}) = \frac{1}{\tau_n}\left(\frac{P(t, T_n)}{P(t, T_{n+1})} - 1\right).$$

At time instant T_{i+1} the fixed leg pays out the amount

$$G\tau_i K,$$

and the floating leg pays out

$$G\tau_i L_i(T_i),$$

where G is the nominal value, K is the fixed rate, $L(T_i, T_{i+1})$ is the LIBOR interest rate and tenor τ_i is the time elapsed between T_{i+1} and T_i. Hence, the cash flow of the fixed leg payer or receiver at time T_{i+1} is given by

$$G\omega\tau_i(L_i(T_i) - K),$$

where $\omega = 1$ for the fixed rate payer and $\omega = -1$ for the fixed rate receiver. From previous chapters we know, that the price of a tradable asset can be obtainable in terms of a certain conditional expectation. We also mentioned that any forward rate $L_i(T_i)$ is a martingale under the T_{i+1}-forward measure. Therefore, since K is not time dependent, the price of a swap at time $t < T_0$ can be calculated as

$$\begin{aligned} V_{swap}(t) &= \omega G \sum_{i=0}^{N-1} \tau_i P(t, T_{i+1})\, \mathbb{E}_t^{i+1}[L_i(T_i) - K] \\ &= \omega G \sum_{i=0}^{N-1} \tau_i P(t, T_{i+1})\, (L_i(t) - K). \end{aligned}$$

Note, that the price of a swap could be derived also as the sum of multiple FRAs and the final equation would be the same.

4.2 The Swap Measure

In the following, we introduce the swap measure, for which first we must adjust the previous equation as

$$V_{swap}(t) = \omega G \sum_{i=0}^{N-1} \tau_i P(t, T_{i+1}) \left(\frac{\sum_{i=0}^{N-1} \tau_i P(t, T_{i+1})\, L_i(t)}{\sum_{i=0}^{N-1} \tau_i P(t, T_{i+1})} - K \right)$$

and define stochastic processes

$$A(t) \equiv A_{0,N}(t) = \sum_{i=0}^{N-1} \tau_i P(t, T_{i+1}),$$

$$S(t) \equiv S_{0,N}(t) = \frac{\sum_{i=0}^{N-1} \tau_i P(t, T_{i+1})\, L_i(t)}{\sum_{i=0}^{N-1} \tau_i P(t, T_{i+1})},$$

so that the swap price's equation can be rewritten as

$$V_{swap}(t) = G\omega A(t)(S(t) - K). \tag{10}$$

For future reference, $A(t)$ is known as the annuity of the swap, while $S(t)$ is the forward swap rate [14], which similarly to the forward rate in case of an FRA, can be found by following the principle of rendering the contract to be a zero priced agreement at time t.

There also exists a generalized form for both, which are used to define the price of swaps with different starting and ending dates.

Definition 15 [14] *Consider a set of predefined time dates* $0 \leq T_0 < T_1 < \cdots < T_N$. *For any integers* k, m *satisfying* $0 \leq k \leq N$, $m \geq 0$ *and* $k + m \leq N$ *the annuity rate* $A_{k,m}$ *is given by equation*

$$A_{k,m}(t) = \sum_{i=k}^{m-1} \tau_i P(t, T_{i+1}), \tag{11}$$

and the swap rate $S_{k,m}$ by equation

$$S_{k,m}(t) = \frac{P(t, T_k) - P(t, T_m)}{A_{k,m}(t)}. \tag{12}$$

Note, that the general form of the swap rate defined by Equation (12) can be obtained after rewriting the definition of $S(t)$, which is done in the following Remark.

Remark 1 *Since the forward LIBOR rate* $L_i(t)$ *is defined as*

$$L_i(t) = \frac{1}{\tau_i}\left(\frac{P(t, T_i)}{P(t, T_{i+1})} - 1\right),$$

we can substitute it in

$$S(t) = \frac{\sum_{i=0}^{N-1} \tau_i P(t, T_{i+1})\, L_i(t)}{\sum_{i=0}^{N-1} \tau_i P(t, T_{i+1})}$$

for every $i = 0, 1, \ldots, N-1$. *Hence,*

$$S(t) = \frac{\sum_{i=0}^{N-1}\big(P(t, T_i) - P(t, T_{i+1})\big)}{\sum_{i=0}^{N-1} \tau_i P(t, T_{i+1})},$$

which results

$$S(t) = \frac{P(t, T_0) - P(t, T_N)}{\sum_{i=0}^{N-1} \tau_i P(t, T_{i+1})}.$$

The denominator of the previous fraction is the annuity $A(t)$, *so the final equation of the swap rate* $S(t)$ *is given as*

$$S(t) = \frac{P(t, T_0) - P(t, T_N)}{A(t)}.$$

Hence, if we change starting dates T_0 *and* T_N *to a more general case, the swap rate is given as*

$$S_{k,m}(t) = \frac{P(t, T_k) - P(t, T_m)}{A_{k,m}(t)}.$$

The annuity of a swap is a linear combination of zero-coupon bonds, which makes it a tradable non-dividend paying asset. Therefore, we can define a new probability measure, the swap measure based on the annuity as a numeraire.

Definition 16 The Swap Measure *The swap measure $\mathbb{Q}^{k,m}$ is a probability measure induced by the numeraire $A_{k,m}$, under which the price of an arbitrary asset V normalized by the annuity $A_{k,m}$ is a martingale. In formulas*

$$V(t) = A_{k,m}(t)\mathbb{E}_t^{k,m}\left[\frac{V(T)}{A_{k,m}(T)}\right],$$

where $\mathbb{E}_t^{k,m}$ is the conditional expectation under the swap measure $\mathbb{Q}^{k,m}$ and corresponding to the filtration $\mathcal{F}_t$.

In closing this section, by proving the following Lemma we introduce a swap-rate property, that will be helpful in the following Section of swaption pricing. Note, that the swap rate is the difference of two zero-coupon bonds – a traded asset – normalized by the annuity $A_{k,m}$. Therefore, according to *Definition 16* it must be a martingale under the swap measure $\mathbb{Q}^{k,m}$.

Lemma 3 *The swap rate $S_{k,m}$ is a martingale under the swap measure $\mathbb{Q}^{k,m}$.*

Proof. For any $T \leq T_k$ – based on the fact, that the difference of two zero-coupon bonds is a traded asset – the following must hold:

$$\mathbb{E}_t^{k,m}\left[S_{k,m}(T)\right] = \mathbb{E}_t^{k,m}\left[\frac{P(T,T_k) - P(T,T_{k+m})}{A_{k,m}(T)}\right] =$$

$$= \frac{P(t,T_k) - P(t,T_{k+m})}{A_{k,m}(t)} = S_{k,m}(t).$$

■

4.3 Swaptions

A swaption gives to its holder the option to enter into a swap agreement. An option is a financial instrument, which gives the holder the right, but not the obligation to buy or sell the underlying asset for the certain price (*strike price*) at a future date. However, in case of a swap-option contract the *strike* of a swaption is the fixed rate K of the underlying swap. Similarly to vanilla options, we can differentiate two types of swaptions based on the two legs of a swap agreement: a *payer swaption* and a *receiver swaption*. In general, the expiry date of a swaption does not have to be identical with the start date of the underlying swap, but in this thesis, we assume they occur at the same time and refer to it as the *expiry date*. Hence, according to their style of exercise there are two main types:

- European swaption – in which the option can be exercised only at the expiry date.
- American swaption – in which the option can be exercised at any time during the option period.

Just like any option, swaptions can be in-the-money (ITM), at-the-money (ATM) or out-of-the-money (OTM) depending on the strike's position compared to the underlying swap rate. Considering the payer swaption, it is ITM if $S(t) > K$, while the corresponding receiver swaption is OTM. Similarly, if $S(t) < K$, the payer swaption is OTM and the receiver is ITM. A swaption is ATM if $S(t) = K$.

In this section, we derive the price of a European swaption, since, assuming the lognormality of the swap rate dynamics, it can be done analytically using the Black formula. For an American swaption the pricing must be done through numerical methods.

By the Black formula, we refer to the key result of the extended Black-Scholes-Merton model [15], [16], which is used to price European-style derivatives analytically when interest rates are stochastic. The Black-Scholes-Merton model can be derived by setting up a riskless portfolio – that consists of position both in the stock and the derivative – and assuming that its return equals to the risk-free return over a short period of time. It is because both the underlying asset price and the derivative price depend on the former's price movement. Indeed, they are perfectly correlated over a short period of time. In 1976, Fischer Black improved their primary model [2], and presented the framework, which is now known as the Black76 model and can be used to price options in terms of futures prices assuming stochastic interest rates. A detailed derivation of the whole model is out of the limits of this thesis, therefore, we refer the reader to the book of John C. Hull [13]. We only present the first steps of the valuation and directly apply the Black formula.

Recall Equation (10) defining the swap price at time t as

$$V_{swap}(t) = G\omega A(t)(S(t) - K).$$

Since we assume the T_0 is both the expiry time of the swaption and the start of the underlying swap, the payoff of the swaption at time T_0 must be

$$[\omega GA(T_0)(S(T_0)-K)]^+ = max\big(G\omega A(T_0)(S(T_0)-K)\big).$$

Hence, according to the definition of the swap measure, the swaption price at time t is given by the formula

$$V_{swaption}(0) = A(t)\mathbb{E}_0^{0,N}\left[\frac{[G\omega A(T_0)(S(T_0)-K)]^+}{A(T_0)}\right]$$
$$= GA(t)\mathbb{E}_0^{0,N}[[\omega(S(T_0)-K)]^+].$$

Assuming $S(t)$ has a lognormal distribution under the $\mathbb{Q}^{0,N}$ with standard deviation $\sigma\sqrt{T_0}$ the Black formula for the swaption price is defined in the box below.

The price of a swaption with strike K and expiry date T_0 at time 0 is given by the formula

$$V_{swaption}(0) = G\omega A(0)[S(0)\mathcal{N}(\omega d_1) - K\mathcal{N}(\omega d_2)],$$

where G is the nominal value of the underlying swap, $\omega = 1$ for the fixed rate payer and $\omega = -1$ for the fixed rate receiver,

$$d_1 = \frac{ln\left(\frac{S(0)}{K}\right) + \frac{1}{2}\sigma^2 T_0}{\sigma\sqrt{T_0}}, \qquad d_2 = d_1 - \sigma\sqrt{T_0}$$

and $\mathcal{N}(x) = \int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2}dy$ is the standard normal cumulative distribution function.

Note, this formula (besides S, K or T_0) depends on the parameter σ, which is typically derived from market data. Since such a σ is implied by the market price of an option based on a pricing model, it is called the *implied volatility*. Implied volatilities can be calculated for different strikes K and maturities T. Their mapping from K and T is known as the *volatility surface*. In the real word swaption valuation volatilities might be adjusted to incorporate expert judgment, to reflect the risk-averse prudence or alternatively some risky investment views.

As we said, pricing of swaptions using the Black formula requires lognormal swap rate dynamics, which is assumed in case of *swap market models* (see Chapter 5). In LMM models, we consider lognormal dynamics of forward LIBOR rates, which allow pricing caps using the Black formula, but prohibit us to price swaptions analytically. However, as we will see later in Chapter 6, swap rate dynamics are not far from being lognormal, therefore we can derive an approximated pricing formula for swaptions.

5 LIBOR Market Models

In many cases, some of the short-rate models, or rather their extended forms are still used in practice. There are some good reasons: they can be easily applied and also pricing with these models can offer satisfying results for many interest rate derivatives. However, they all have one major drawback: interest rates modeled by them are not directly observable in financial markets.

The most popular and most widely used models today consider the LIBOR forward rates as modeled quantities, which in contrast to the instantaneous spot and forward rates, are observable in the market. These are the Market Models, which can be calibrated and used for pricing any financial instruments whose payoff can be expressed in terms of forward rates, moreover, they are consistent with the Black formula.

In fact, there are two basic market models, which are unfortunately not compatible: the standard lognormal LIBOR Market Model (LMM model) considering lognormal dynamics of forward LIBOR rates and the lognormal Swap Rate Market Model considering lognormal dynamics of swap rates. As it can be seen in Chapter 6, if a lognormal evolution of each forward LIBOR rate is assumed under its measure, swap rates cannot be lognormal in their own measure as assumed in the lognormal swap rate model. Hence, swaptions and caps can be priced analytically with the Black formula only separately. However, as we have already mention in Chapter 4, all is not lost, since in LMM models we can derive a formula for approximated swaption prices.

The most known LMM model is the *BGM model* [3], which comes from the acronym of the names of its authors, Brace, Gatarek and Musiela and assumes a simple lognormal evolution of forward LIBOR rates. Since its appearance, several extensions have been proposed, which vary in assumptions of special local volatility functions and/or stochastic volatility processes.

First, we define a discrete time structure and the corresponding notation, then using the martingale property of forward LIBOR rates, derive their dynamics under different measures. We briefly introduce the pioneer work of Brace, Gatarek and Musiela and present one of the BGM model's extension, the LMM with a stochastic volatility process. At the end of this Chapter, we also introduce a framework of a concrete model, called the LMMPlus model, that is used in practice for the valuation of life insurance liabilities.

5.1 Preparation

The first part of this Chapter is mainly based on [14], since we find the derivation of forward LIBOR dynamics coherent and straightforward. Consequently, the notation is adjusted. We define the following time structure:

$$0 = T_0 < T_1 < \cdots < T_{N-1} < T_N,$$
$$\tau_n = T_{n+1} - T_n,$$

where $n = 0, 1, \dots, N-1$ and $N \in \mathbb{N}$ is predefined. Note, this defines a finite set of zero-coupon bonds $P(t, T_n)$ for $\{n : t < T_n \leq T_N\}$. Therefore, we can rewrite the definition of the simply-compounded forward interest rate (or forward LIBOR rate) – like we did in Chapter 4 – as

$$L_n(t) = F(t, T_n, T_{n+1}) = \frac{1}{\tau_n}\left(\frac{P(t, T_n)}{P(t, T_{n+1})} - 1\right).$$

As we can see, it is easy to derive the zero-coupon bond price recursively. However, by t moving forward, zero-coupon bonds begin to successively expire, hence it is reasonable to introduce an index function $q(t)$ satisfying

$$T_{q(t)-1} \leq t < T_{q(t)}. \tag{13}$$

By that, we can identify the first zero-coupon bond that has not expired by time t. Now we can get back to the previous idea and recursively deduce the price of a zero-coupon bond for any $t < T_n$. From the definition of the forward LIBOR rate it is clear, that

$$P(t, T_{n+1}) = \frac{P(t, T_n)}{1 + \tau_n L_n(t)}.$$

Hence, after applying this relation recursively, the zero-coupon bond price is given by

$$P(t, T_n) = P\big(t, T_{q(t)}\big) \prod_{i=q(t)}^{n-1} \frac{1}{1 + \tau_i L_i(t)}. \tag{14}$$

We can also define the discrete-time equivalent $B(t)$ of the bank account $\beta(t)$ introduced in Chapter 1, which represents the time t value of investing one unit of currency at time 0. While $\beta(t)$ accrues continuously at the short-rate, $B(t)$ grows by reinvesting the initial one unit of currency at each time instant T_n for the next time instant T_{n+1}. We know from Chapter 2, that

$$L_n(T_n) = F(T_n, T_n, T_{n+1}) = L(T_n, T_{n+1}).$$

Hence, the value of a simply-compounded money market account at time $T_{q(t)}$ is given by

$$B(T_{q(t)}) = \prod_{i=0}^{q(t)-1} \left(1 + \tau_i L_i(T_i)\right),$$

while its time t "equivalent" by

$$B(t) = P(t, T_{q(t)}) \prod_{i=0}^{q(t)-1} \left(1 + \tau_i L_i(T_i)\right). \tag{15}$$

Note, that $B(t)$ can also be a numeraire, hence in the following we introduce the *spot measure* $\mathbb{Q}^B$ as the probability measure defined by the money market account.

Definition 17 Spot Measure *The spot measure $\mathbb{Q}^B$ is a measure induced by the numeraire $B(t)$, under which $V(t)/B(t)$ is a martingale. Hence,*

$$V(t) = B(t)\mathbb{E}_t^B \left[\frac{V(T)}{B(T)}\right],$$

where $\mathbb{E}_t^B$ is the conditional expectation under the spot measure $\mathbb{Q}^B$ and corresponding to the filtration $\mathcal{F}_t$.

5.2 Forward LIBOR Rate Dynamics

Remember the index function $q(t)$ defined by Equation (13) and consider the set of forward LIBOR rates at time t, fixed at time instants with index $n \geq q(t)$, i.e. $L_{q(t)}(t), L_{q(t)+1}(t), \dots, L_{N-1}(t)$.

In Chapter 3, we proved that – considering the discrete-time structure and the adjusted notation – $L_n(t)$ is a martingale under the T_{n+1}-forward measure $\mathbb{Q}^{n+1} \equiv \mathbb{Q}^{T_{n+1}}$. Therefore, assume a set of driftless dynamics for $n \geq q(t)$:

$$dL_n(t) = \sigma_n(t)^\top dW^{n+1}(t), \tag{16}$$

where $W^{n+1}(t)$ is the m-dimensional Wiener process under the T_{n+1}-forward measure $\mathbb{Q}^{n+1}$ and $\sigma_n(t)$ is a m-dimensional process, i.e.

$$W^{n+1}(t) = [W^{n+1,1}(t), \dots, W^{n+1,m}(t)]_{t\geq 0}^\top,$$
$$\sigma_n(t) = [\sigma_n{}^1(t), \dots, \sigma_n{}^m(t)\,]_{t\geq 0}^\top.$$

For the moment, we do not provide further details regarding the process $\sigma_n(t)$. Now we simply assume it corresponds to the conditions of this framework.

We derive the dynamics of $L_n(t)$ under the T_n-forward measure $\mathbb{Q}^n$. As we will see, by moving from the T_{n+1}-forward measure to the adjacent forward measure we obtain a recursive relation, that will help us to represent the forward-rate dynamics under the terminal T_N-forward measure.

Lemma 4 [14] *Assume that $L_n(t)$ is given by Equation (16). The dynamics of $L_n(t)$ under the T_n-forward measure $\mathbb{Q}^{T_n}$ are given by*

$$dL_n(t) = \sigma_n(t)^\top \left(\frac{\tau_n \sigma_n(t)}{1 + \tau_n L_n(t)} dt + dW^n(t) \right),$$

where $W^n(t)$ is the m-dimensional Wiener process under the T_n-forward measure $\mathbb{Q}^n$.

Proof. From the result of the Change of Numeraire Theorem in Chapter 3 we know, that the Radon-Nikodym derivative defining measure $\mathbb{Q}^n$ is given by

$$M_T = \frac{d\mathbb{Q}^n}{d\mathbb{Q}^{n+1}} = \frac{P(T, T_n)}{P(0, T_n)} \frac{P(0, T_{n+1})}{P(T, T_{n+1})}.$$

Considering the numeraire $P(t, T_{n+1})$, one can see that the stochastic process M_t is a martingale under the T_{n+1}-forward measure, since

$$\mathbb{E}_t^{n+1}[M_T] = \mathbb{E}_t^{n+1}\left[\frac{P(T, T_n)}{P(0, T_n)} \frac{P(0, T_{n+1})}{P(T, T_{n+1})}\right] = \frac{P(t, T_n)}{P(0, T_n)} \frac{P(0, T_{n+1})}{P(t, T_{n+1})} = M_t$$

for $t \le T$.

Note, that $P(t, T_n)/P(t, T_{n+1})$ can be expressed in terms of the forward LIBOR rate, so that M_t can be rewritten as

$$M_t = \frac{P(0, T_{n+1})}{P(0, T_n)} \big(1 + \tau_n L_n(t)\big).$$

For a given n the only t-dependent value in this process is the forward rate $L_n(t)$, hence the derivation of M_t is as simple as

$$dM_t = \frac{P(0, T_{n+1})}{P(0, T_n)} \tau_n dL_n(t) = \frac{P(0, T_{n+1})}{P(0, T_n)} \tau_n \sigma_n(t)^\top dW^{n+1}(t).$$

If we divide the previous equation by M_t, we get

$$\frac{dM_t}{M_t} = \gamma_n(t)^\top dW^{n+1}(t),$$

where

$$\gamma_n(t) = \frac{\tau_n \sigma_n(t)}{1 + \tau_n L_n(t)}.$$

It is clear, that

$$\frac{dM_t}{M_t} = d\big(ln(M_t)\big),$$

hence using the Itô's Lemma (see Appendix B), with $g(t, M_t) = ln(M_t)$ we obtain

$$M_t = \exp\left(\int_0^t \gamma_n(s)^\top dW^{n+1}(s) - \frac{1}{2}\int_0^t \gamma_n(s)^\top \gamma_n(s)\, ds\right).$$

Assuming $\gamma_n(t)$ is locally bounded, i.e. for every $t > 0$ there exists a positive constant $C(t)$ such that

$$\gamma_n(t)^\top \gamma_n(t) \le C(t),$$

$\gamma_n(t)$ satisfies the Novikov's condition (see [17]).

By that, we can apply Girsanov's Theorem I (see Appendix B), where we set

$$W(t) = W^{n+1}(t),$$
$$a(t, \omega) = -\gamma_n(t),$$
$$Y(t) = W^n(t),$$

so that

$$dW^n(t) = dW^{n+1}(t) - \frac{\tau_n \sigma_n(t)}{1 + \tau_n L_n(t)} dt,$$

and $W^n(t)$ is a Wiener process under the T_n-forward measure $\mathbb{Q}^n$. Substituting

$$dW^{n+1}(t) = dW^n(t) + \frac{\tau_n \sigma_n(t)}{1 + \tau_n L_n(t)} dt \tag{17}$$

into Equation (16) we obtain

$$dL_n(t) = \sigma_n(t)^\top \left(\frac{\tau_n \sigma_n(t)}{1 + \tau_n L_n(t)} dt + dW^n(t)\right),$$

i.e. the dynamics of the forward LIBOR rate under the T_n-forward measure $\mathbb{Q}^n$.

■

Note, that the previous Lemma can be used iteratively. Therefore, in the next Lemma we prove that the dynamics of $L_n(t)$ for any $n \ge q(t)$ can be represented under the terminal measure $\mathbb{Q}^N \equiv \mathbb{Q}^{T_N}$.

Lemma 5 [14] *Assume that* $L_n(t)$ *is given by Equation (16). The dynamics of* $L_n(t)$ *under the* T_N*-forward measure* $\mathbb{Q}^{T_N}$ *are given by*

$$dL_n(t) = \sigma_n(t)^\top \left(-\sum_{j=n+1}^{N-1} \frac{\tau_j \sigma_j(t)}{1+\tau_j L_j(t)} dt + dW^N(t) \right), \tag{18}$$

where $W^N(t)$ is the Wiener process under the T_N-forward measure $\mathbb{Q}^{T_N}$.

Proof. Consider the result of *Lemma 4*, which must hold for any $q(t) \le n \le N-1$. Hence, applying Equation (17) recursively we obtain dynamics of the Wiener process under the terminal measure $\mathbb{Q}^{T_N}$ as

$$\begin{aligned} dW^N(t) &= dW^{N-1}(t) + \frac{\tau_{N-1}\sigma_{N-1}(t)}{1+\tau_{N-1}L_{N-1}(t)} dt \\ &= dW^{N-2}(t) + \sum_{j=N-2}^{N-1} \frac{\tau_j \sigma_j(t)}{1+\tau_j L_j(t)} dt \\ &\vdots \\ &= dW^{n+1}(t) + \sum_{j=n+1}^{N-1} \frac{\tau_j \sigma_j(t)}{1+\tau_j L_j(t)} dt. \end{aligned}$$

Again, $dW^{n+1}(t)$ can be substituted into Equation (16), which gives us the desired dynamics under the T_N-forward measure $\mathbb{Q}^{T_N}$.

■

In the following Lemma we show that the dynamics of forward LIBOR rates can also be represented under the spot measure $\mathbb{Q}^B$ using a very similar approach. As we discuss later, considering the dynamics in $\mathbb{Q}^B$ have an advantage over using the dynamics in the terminal measure, since its drift term changes as time t moves.

Lemma 6 *Assume that $L_n(t)$ is given by Equation (16). The dynamics of $L_n(t)$ under the spot measure $\mathbb{Q}^B$ are given by*

$$dL_n(t) = \sigma_n(t)^\top \left(\sum_{j=q(t)}^{n} \frac{\tau_j \sigma_j(t)}{1+\tau_j L_j(t)} dt + dW^B(t) \right), \tag{19}$$

where $W^B(t)$ is the Wiener process under the spot measure $\mathbb{Q}^B$.

Proof. Recall Equations (14) and (15) from Section 5.1 and proceed as follows. Clearly, for $0 \le t \le T_{n+1}$ the normalized bond price

$$\frac{P(t,T_{n+1})}{B(t)} = \frac{P\big(t,T_{q(t)}\big)\prod_{i=q(t)}^{n}\frac{1}{1+\tau_i L_i(t)}}{P\big(t,T_{q(t)}\big)\prod_{i=0}^{q(t)-1}\big(1+\tau_i L_i(T_i)\big)}$$
$$= \left(\prod_{i=0}^{q(t)-1}\big(1+\tau_i L_i(T_i)\big)^{-1}\right)\left(\prod_{i=q(t)}^{n}\big(1+\tau_i L_i(t)\big)^{-1}\right)$$

is the function of forward LIBOR rates only.

Since the money-market account is a numeraire, by using the Change of Numeraire Theorem we can ensure that the corresponding spot measure $\mathbb{Q}^B$ is equivalent to the T_{n+1}-forward measure $\mathbb{Q}^{T_{n+1}}$. Therefore, $P(t,T_{n+1})/B(t)$ is a martingale under $\mathbb{Q}^B$. Let us define a stochastic process $N_n(t)$ as

$$N_n(t) = \frac{B(0)}{P(0,T_{n+1})}\frac{P(t,T_{n+1})}{B(t)}.$$

Assuming N_t satisfies the desired conditions, according to the Martingale Representation Theorem presented (see Appendix B) there exists a unique stochastic process $\eta_n(t)$, such that $\eta_n \in \mathcal{V}^n(0,t)$ (also Appendix B) for all $t \geq 0$ and

$$N_n(t) = N_0 + \int_0^t \eta_n(t)^\top\, dW^B(t),$$

where $N_0 = 1$. Therefore, by applying the Itô's Lemma where $g(t,N_t) = ln\big(N_n(t)\big)$, the stochastic process $N_n(t)$ can be written as

$$N_n(t) = exp\left(\int_0^t \frac{1}{N_n(s)}\eta_n(s)^\top dW^B(s) - \frac{1}{2}\int_0^t \frac{1}{N_n(s)^2}\eta_n(s)^\top\eta_n(s)\ dW^B(s)\right),$$

where, $\frac{1}{N_t}$ is locally bounded (see [18]). Hence,

$$\nu_n(t) = \frac{1}{N_t}\eta_n(t)$$

satisfies the Novikov's condition, and

$$dln\big(N_n(t)\big) = \nu_n(t)^\top dW^B(t) - \frac{1}{2}\nu_n(t)^\top\nu_n(t)dt. \tag{20}$$

We also know from the normalized bond price, that the previous derivation can be expressed as

$$d\ln\big(N_n(t)\big) = d\ln\left\{\frac{B(0)}{P(0,T_{n+1})}\left(\prod_{i=0}^{q(t)-1}\big(1+\tau_i L_i(T_i)\big)^{-1}\right)\left(\prod_{i=q(t)}^{n}\big(1+\tau_i L_i(t)\big)^{-1}\right)\right\},$$

where

$$\frac{B(0)}{P(0,T_{n+1})}\left(\prod_{i=0}^{q(t)-1}\big(1+\tau_i L_i(T_i)\big)^{-1}\right)$$

is not time dependent. Hence,

$$d\ln\big(N_n(t)\big) = -\sum_{i=q(t)}^{n} d\ln\big(1+\tau_i L_i(t)\big). \tag{21}$$

Previously, we could see the fact – which comes directly from Girsanov's theorems (see Appendix B) – that the diffusion coefficient $DC(\cdot)$ of a stochastic process remains unchanged after an equivalent measure change.

Remark 2 *The diffusion coefficient $DC(\cdot)$ of a process defined as*

$$dX_t = (\dots)dt + v_t dW_t,$$

is the vector v_t.

Therefore, since $DC\big(L_i(t)\big)$ under any measure is $\sigma_i(t)$ for $i = q(t), \dots, n-1$, by applying the Itô's Lemma on $\ln\big(1+\tau_i L_i(t)\big)$ we get

$$DC\left(\ln\big(1+\tau_i L_i(t)\big)\right) = \frac{\tau_i \sigma_i(t)}{1+\tau_i L_i(t)},$$

while

$$DC\left(d\ln\big(N_n(t)\big)\right) = -\sum_{i=q(t)}^{n}\frac{\tau_i \sigma_i(t)}{1+\tau_i L_i(t)}.$$

Clearly, we can equate the diffusion coefficients of Equations (20) and (21), so that we get

$$v_n(t) = -\sum_{i=q(t)}^{n}\frac{\tau_i \sigma_i(t)}{1+\tau_i L_i(t)}.$$

Finally, we can derive dynamics of $L_n(t)$ as

$$dL_n(t) = \frac{1}{\tau_n} d\left(\frac{P(t,T_n)}{P(t,T_{n+1})}\right) = \frac{1}{\tau_n} d\left(\frac{\frac{P(t,T_n)}{B(t)}}{\frac{P(t,T_{n+1})}{B(t)}}\right)$$

$$= \frac{1}{\tau_n} d\left(\frac{\frac{P(0,T_n)}{B(0)} N_{n-1}(t)}{\frac{P(0,T_{n+1})}{B(0)} N_n(t)}\right) = \frac{1}{\tau_n} \frac{P(0,T_n)}{P(0,T_{n+1})} d\left(\frac{N_{n-1}(t)}{N_n(t)}\right).$$

For $d(N_{t-1}/N_t)$ we can apply the multiplication rule of the Itô's Lemma (see Appendix B), so that

$$d\left(\frac{N_{n-1}(t)}{N_n(t)}\right) = \frac{N_{n-1}(t)}{N_n(t)} \big(\nu_{n-1}(t) - \nu_n(t)\big)^\top \big(-\nu_n(t)dt + dW^B(t)\big),$$

which results

$$dL_n(t) = \frac{1}{\tau_n} \frac{P(t,T_n)}{P(t,T_{n+1})} \left(\frac{\tau_n \sigma_n(t)}{1+\tau_n L_n(t)}\right)^\top \left(\sum_{i=q(t)}^{n} \frac{\tau_i \sigma_i(t)}{1+\tau_i L_i(t)} dt + dW^B(t)\right)$$

$$= \sigma_n(t)^\top \left(\sum_{i=q(t)}^{n} \frac{\tau_i \sigma_i(t)}{1+\tau_i L_i(t)} dt + dW^B(t)\right).$$

■

Summarizing this section, in the following box we provide a collective definition of the forward LIBOR rate dynamics under all the probability measures discussed earlier.

The forward LIBOR rate ($L_n(t)$) dynamics under the T_{n+1}-forward measure $\mathbb{Q}^{T_{n+1}}$ are given by the formula:

$$dL_n(t) = \sigma_n(t)^\top dW^{n+1}(t), \qquad n = 0, 1, \dots, N-1.$$

Under the terminal forward measure $\mathbb{Q}^{T_N}$:

$$dL_n(t) = \sigma_n(t)^\top \left(-\sum_{j=n+1}^{N-1} \frac{\tau_i \sigma_j(t)}{1+\tau_j L_j(t)} dt + dW^N(t)\right), \qquad n = 0, 2, \dots, N-2.$$

Under the spot measure $\mathbb{Q}^B$:

$$dL_n(t) = \sigma_n(t)^\top \left(\sum_{i=q(t)}^{n} \frac{\tau_i \sigma_i(t)}{1+\tau_i L_i(t)} dt + dW^B(t)\right), \qquad n = 0, 1, \dots, N-1.$$

Take a look at the dynamics in the terminal and the spot measure, and note the difference between them:

- Under the terminal measure, the number of terms in the drift summation is fixed. Hence, a forward rate with smaller lower index is probably more biased than a forward rate with higher lower index.
- Under the spot measure, the number of terms in the drift summation is decreasing as time increases.

Also, the discretized cash account is intuitively close to the well-known risk-neutral continuous cash account, therefore we choose dynamics under the spot measure $\mathbb{Q}^B$.

5.3 The Choice of $\sigma_n(t)$

Until now, we have not provided assumptions regarding the diffusion process $\sigma_n(t)$. In the previous Section we considered it to be a "well-behaved" process, so that the results were non-explosive.

However, to build a proper model, one must be specific about $\sigma_n(t)$. In general, we assume that the stochastic diffusion process $\sigma_n(t)$ is defined as

$$\sigma_n(t) = \lambda_n(t)\varphi\big(L_n(t)\big), \tag{22}$$

where $\lambda_n(t)$ is a bounded deterministic m-dimensional function and $\varphi : \mathbb{R} \longrightarrow \mathbb{R}$ is a *local volatility* process of $L_n(t)$.

Local volatility functions are very tractable and widely used to find a parametric distribution that is flexible enough to consistently price quoted options in a market. Although there are many functions proposed (see [1]), most of them lead only to skews in the implied volatility. Some more flexible functions allow for smile-shaped volatilities, but each local volatility function has its own limitations. Hence, stochastic volatility processes (see Section 5.5.1) are used instead, or in addition. Obviously not all choices of $\varphi(\cdot)$ are allowed – the existence and the uniqueness of the solution must be ensured.

We assume $L_n(0)$ being non-negative for all n, set $\varphi(0) = 0$ and consider a local volatility function $\varphi(\cdot)$ that satisfies *the regularity conditions* [8]:

- φ is *Lipschitz continuous*
- φ satisfies the *growth condition*

$$\varphi^2(x) \leq K(1 + x^2), \qquad \forall x \in \mathbb{R}$$

  where K is a positive constant.

This ensures "non-explosive, pathwise unique solutions" [8] of the stochastic differential equation for forward LIBOR rates under all forward measures. The previous assumptions and proved results are presented in *Interest Rate Modeling* [7], [8] by Piterbarg & Andersen in form of a lemma.

There are some standard parametrizations for φ, such as $\varphi\big(L_n(t)\big) = L_n(t)$ (the log-normal specification) or $\varphi\big(L_n(t)\big) = bL_n(t) + a$ (the displaced log-normal specification). The former – used in the BGM model – implies flat implied volatilities and satisfies the regularity conditions. However, for the latter that is considered by the LMMPlus model, we need to impose additional restrictions. Since the displaced local volatility function $\varphi\big(L_n(t)\big) = bL_n(t) + a$ clearly violates $\varphi(0) = 0$, we assume $bL_n(0) + a > 0$. In addition, we must prevent $1 + \tau_n L_n(t)$ in the denominator of the drift becoming zero, therefore the following also must hold:

$$\frac{a}{b} < \frac{1}{\tau_n}, \qquad n = 1, 2, \dots, N-1.$$

Again, detailed explanation and the proof of the results, i.e. the existence of non-explosive unique solutions of the stochastic differential equation can be found in [8].

Another possible parametrization proposed by Piterbarg & Andresen [8] is the specification of the displaced log-normality as $\varphi\big(L_n(t)\big) = (1-b)L_n(0) + bL_n(t)$. By that, the constant component is different for each forward LIBOR rate and we must require

$$\frac{(1-b)}{b} < \frac{1}{L_n(0)\tau_n}.$$

In case of stochastic volatility models – like LMMPlus model – the separation of the diffusion process still holds, however it is extended by a stochastic volatility process, for which we provide further details in Section 5.5.1.

5.4 Brace-Gatarek-Musiela (BGM) Model

The BGM model is also called the log-normal forward LIBOR rate model, because by choosing the skew function as $\varphi(x) = x$ we get

$$dL_n(t) = L_n(t)\lambda_n(t)^\top dW^{n+1}(t),$$

where $\lambda_n(t) : \mathbb{R} \longrightarrow \mathbb{R}^m$ is a deterministic function and $W^{n+1}(t)$ is an m-dimensional Wiener process under the T_{n+1}-forward measure.

5.5 LIBOR Market Model with a Stochastic Volatility Process

In this Section, we present an extension of the standard BGM model, for which we introduce the stochastic volatility process (SVP) and discuss its inconveniences associated with instantaneous correlations between Wiener processes of the SVP and forward LIBOR rates dynamics (see Section 5.5.1). Then, in Section 5.5.2 we guide the reader through the algorithm of the rank reduction, including the discussion of the instantaneous correlation between forward LIBOR rates; the construction of an instantaneous correlation matrix from the sample covariance matrix or using a parametric form; and applying the PCA to reduce the number of factors, i.e. the dimension of the Wiener process W^B. We end this Section by introducing the LMMPlus model.

5.5.1 The Stochastic Volatility Process

The main drawback of the BGM model is that it implies a flat volatility structure. In markets, however, we can observe the non-monotonic behavior of volatility structure with smiles and skews. As we have discussed earlier, some local volatility functions are able to capture such a behavior, but as we also said, they have some limitations. Therefore, in order to capture the (realistic) stochastic behavior of volatility and adapt market smiles and skews, we introduce the stochastic volatility process. While existence itself of an SVP generates a smile-shaped structure, skew-shaped volatilities can be captured only by assuming for example a non-zero correlation between the Wiener processes of the SVP and the forward rate dynamics, assuming a displaced diffusion or a non-linear local volatility function φ. As discussed by Brigo & Mercurio [1], at least one of these requirements should be fulfilled.

Considering the LMM model extended by an SVP, the forward LIBOR rate dynamics in the spot measure $\mathbb{Q}^B$ are given as

$$dL_n(t) = \sqrt{z(t)}\varphi\big(L_n(t)\big)\lambda_n(t)^\top \left(\sqrt{z(t)}\mu_n(t)dt + dW^B(t)\right),$$

where we replaced $\sigma_n(t)$ according to Equation (22),

$$\mu_n(t) = \sum_{i=q(t)}^{n} \frac{\tau_i \varphi\big(L_i(t)\big)\lambda_i(t)}{1 + \tau_i L_i(t)}$$

and $z(t)$ is a *mean-reverting process* – more precisely the Cox-Ingersoll-Ross (CIR) process – with dynamics

$$dz(t) = \kappa\big(\theta - z(t)\big)dt + \epsilon\sqrt{z(t)}dZ^B(t),$$

where parameters κ, θ and ϵ are positive constants, and $Z^B(t)$ is a Wiener process in the spot measure $\mathbb{Q}^B$ that is in general correlated to $W^B(t)$.

This clean form, is not observable under other probability measures, due to the non-zero correlation between the Wiener processes of dynamics, which is quite inconvenient, for example when pricing swaptions (see Chapter 6). In the following we show how the z-process changes when we move into the T_{n+1}-forward measure $\mathbb{Q}^{n+1}$, i.e. the measure, where forward LIBOR rates are martingales. From Section 5.2 we know, that the dynamics in $\mathbb{Q}^{n+1}$ are given as

$$dL_n(t) = \sqrt{z(t)}\varphi\big(L_n(t)\big)\lambda_n(t)^\top dW^{n+1}(t),$$

where $W^{n+1}(t)$ is a Wiener process under $\mathbb{Q}^{n+1}$.
Therefore, by applying Lemmas from Section 5.2 the following must hold:

$$dW^{n+1}(t) = \sqrt{z(t)}\mu_n(t)dt + dW^B(t).$$

In general, we define

$$a(t) = \frac{d\langle Z^B(t), W^B(t)\rangle}{dt},$$

where $\langle Z^B(t), W^B(t)\rangle$ is the *quadratic covariation process* (see Appendix B) given by

$$\mathrm{d}\langle Z^B(t), W^B(t)\rangle = dZ^B(t)dW^B(t).$$

Then, we can write (see [8])

$$dZ^B(t) = a(t)^\top dW^B(t) + \sqrt{1 - \|a(t)\|^2}d\widehat{W}(t),$$

where $\widehat{W}(t)$ is a Wiener process independent from $W^B(t)$. Moving to the T_{n+1}-forward measure $\mathbb{Q}^{n+1}$ by substituting $dW^B(t)$ we get

$$dZ^B(t) = dZ^{n+1}(t) - \sqrt{z(t)}a(t)^\top \mu_n(t)dt,$$

where

$$dZ^{n+1}(t) = a(t)^\top dW^{n+1}(t) + \sqrt{1 - \|a(t)\|^2}d\widehat{W}(t).$$

Hence, the process $z(t)$ under the T_{n+1}-forward measure $\mathbb{Q}^{n+1}$ is given by the stochastic differential equation

$$dz(t) = \kappa\left(\theta - z(t)\left(1 + \frac{\epsilon}{\kappa}\mu_n(t)^\top a(t)\right)\right)dt + \epsilon\sqrt{z(t)}dZ^{n+1}(t),$$

where $Z^{n+1}(t)$ is a Wiener process under $\mathbb{Q}^{n+1}$.

Obviously, as discussed by Piterbarg & Andersen in *Interest Rates Modeling* [8], the change in the drift makes it harder to deal with the dynamics when implementing measure change techniques. Therefore, it is common to assume independent Wiener processes, i.e. simply set $a(t) = 0$, so that the SVP dynamics become the same regardless of which measure we choose. However, the framework of Piterbarg & Andersen considers a special local volatility function

$$\varphi\big(L_n(t)\big) = (1-b)L_n(0) + bL_n(t),$$

such that the volatility skew can be controlled with parameter b. As we will see in next Chapters, the LMMPlus model assumes a simple displaced-diffusion (see Section 5.6), in which case a non-zero correlation is recommended.

5.5.2 Instantaneous Correlations and Rank Reduction

In one-factor models for interest rates – such as the vanilla short-rate models – various points on the forward curve are perfectly correlated. Although this type of forward curves – where the points move co-monotonically – is common in markets, some more complex types of curve changes can also be observed. Therefore, in LMM models we control these correlations by using m-dimensional Wiener processes.

Using the same covariance notation as before, we have for any n and k

$$d\langle L_n, L_k\rangle(t) = z(t)\varphi\big(L_n(t)\big)\varphi\big(L_k(t)\big)\lambda_n(t)^\top\lambda_k(t)dt.$$

Hence, the correlation between $L_n(t)$ and $L_k(t)$ is given by

$$\mathrm{Corr}\big(dL_n(t), dL_k(t)\big) = \frac{d\langle L_n, L_k\rangle(t)}{\sqrt{d\langle L_n, L_n\rangle(t)\, d\langle L_k, L_k\rangle(t)}} = \frac{\lambda_n(t)^\top\lambda_k(t)}{\|\lambda_n(t)\|\|\lambda_k(t)\|}.$$

Obviously, when we have only one Wiener process, i.e. $m = 1$, the correlation $\mathrm{Corr}\big(dL_n(t), dL_k(t)\big)(t) = 1$ for any n and k. As the number of the Wiener processes m increases, the ability to capture more complex correlations improves. However, it also increases the complexity of our model. Therefore, it is our ambition to find an "optimal" dimension of Wiener processes. Once we get the empirical data, we can find the smallest m

that is high enough to sufficiently replicate the sample variance-covariance matrix of empirical data by using the tools of the *principal component analysis* (PCA) (see Appendix C).

Remark 3 *Recall the notation* $d\langle L_n, L_k\rangle(t)$*, satisfying*

$$d\langle L_n, L_k\rangle(t) = dL_n(t)dL_k(t).$$

Hence,

$$d\langle L_n, L_k\rangle(t) = z(t)\varphi\big(L_n(t)\big)\varphi\big(L_k(t)\big)\lambda_n(t)^\top\lambda_k(t)\Big[z(t)\mu_n(t)^\top\mu_k(t)(dt)^2 + \sqrt{z(t)}\mu_n(t)^\top dW^B(t) + \sqrt{z(t)}\mu_k(t)^\top dW^B(t) + \big(dW^B(t)\big)^2\Big].$$

Note, that $\big(dW^B(t)\big)^2 = dt$ *and all the other terms in the expression are of order* $\mathcal{O}\big(dt^{3/2}\big)$ *or higher. According to Piterbarg & Andersen* [7], *those terms can be neglected for small* dt*, so that*

$$d\langle L_n, L_k\rangle(t) = z(t)\varphi\big(L_n(t)\big)\varphi\big(L_k(t)\big)\lambda_n(t)^\top\lambda_k(t)dt.$$

5.5.2.1 Sample Correlation Matrix

Consider now the possession of some market data. In the following, we introduce a way to obtain empirical instantaneous correlations and present a possible parametric formulation for them.

Assume the *Musiela parametrization,* i.e. instead of a fixed *time of maturity* a fixed *time to maturity*. Hence, for a fixed value of τ we can define "sliding" forward rates $l(t, x)$ with tenor x as

$$l(t, x) = L(t, t + x, t + x + \tau).$$

Let $x_1, \dots, x_{N_x}$ be a given set of tenors and $t_0, \dots, t_{N_t}$ a given set of calendar times, so that we can define the $N_x \times N_t$ matrix O with elements

$$O_{i,j} = \frac{l\big(t_j, x_i\big) - l\big(t_{j-1}, x_i\big)}{\sqrt{t_j - t_{j-1}}}, \qquad i = 1, \dots, N_x, \qquad j = 1, \dots, N_t,$$

where $\sqrt{t_j - t_{j-1}}$ is the annualizing factor. Hence, by constructing the forward curve from market observable derivatives contracts (see techniques presented in Chapter 6 of [7]) we can obtain a $N_x \times N_x$ variance-covariance matrix defined as

$$C = \frac{OO'}{N_t}.$$

To fit empirical market data, we need to find a number m of Wiener processes that is high enough to sufficiently replicate the variance-covariance matrix C. A concrete example of the usage of PCA is presented in [8], where Piterbarg & Andersen state, that in general three or four dimensions are sufficient, indeed the loss of variance can be small even for $m = 2$. However, in some cases, when a derivative heavily depends on the correlation between forward LIBOR rate with tenors close to each other, higher number of dimensions is required.

Now, if we introduce the diagonal matrix

$$c = \begin{pmatrix} \sqrt{C_{1,1}} & 0 & \cdots & 0 \\ 0 & \sqrt{C_{2,2}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sqrt{C_{N_x,N_x}} \end{pmatrix},$$

then the empirical forward LIBOR rate correlation matrix R, i.e. the matrix of sample estimates $R_{i,j}$ of instantaneous correlation between forward rates $l(t, x_i)$ and $l(t, x_j)$ $i, j = 1, \dots, N_x$ (when time-homogeneity of correlations is assumed) is given as

$$R = c^{-1}Cc^{-1}.$$

In general, we expect matrix R with elements $R_{i,i} \geq 0$ and $R_{i,i} = 1$ to be real, symmetric and positive definite. Since for larger tenors changes in adjacent forward rates are more correlated, we expect the sub-diagonals to be increasing, i.e. $i \longmapsto R_{i+p,i}$ must be increasing for a fixed p. It should also hold that correlations between $l(\cdot, x_i)$ and $l(\cdot, x_j)$ decline in $|x_i - x_j|$. As discussed in [8], the decline is usually steep for small $|x_i - x_j|$, but it has near-flat asymptote for tenors far from each other. Their conclusion from the observed data was that the decay rate of correlations and the level of the asymptote depend also on $\min\{x, y\}$ – as the decay rate is decreasing and the asymptote level is increasing with $\min\{x, y\}$.

In practice, matrix R is often relatively noisy, in sense that it does not satisfy our expectations, i.e. might contain non-intuitive entries. To mention one of the possible reasons: it is known that correlations tend to change over time. Also, an $N_x \times N_x$ correlation matrix is characterized by $N_x(N_x - 1)/2$ entries, which is often too high for practical purposes. Hence, it is common practice to work with a parametric form based on less parameters, which also smooths the correlation matrix. Some well-known parametric formulations can be found in [1] or [8], however, for the following parametric form, we refer the reader to [14], where it is given as

$$Corr\left(dL_i(t), dL_j(t)\right) = q\left(T_i - t, T_j - t\right),$$

where

$$q(x,y) = \rho_\infty + (1-\rho_\infty)\exp\left(-|x-y|\left(b\exp\left(a(min\{x,y\})\right)\right)\right),$$

$-1 \leq \rho_\infty \leq 1$, $b \geq 0$ and $a \in \mathbb{R}$. We can use least-squares optimization against the empirical correlation to find the optimal parameters $\xi^* = (a^*, b^*, \rho_\infty^*)^\top$:

$$\xi^* = \underset{\xi}{\operatorname{argmin}}\left(tr\left(\left(R - R_2(\xi)\right)\left(R - R_2(\xi)\right)^\top\right)\right),$$

where R is the empirical correlation matrix and $R_2(\xi)$ is the correlation matrix generated by ξ. From the value of the minimalized function at ξ^* we get the average absolute correlation error.

In some cases, parametric forms can generate matrices R, that are not positive definite. In practice, it is not common, but possible that R has some negative eigenvalues. Then the correlation matrix has to be "repaired". For further details, we refer the reader to [8].

5.5.2.2 Correlation PCA

We briefly introduce two possible ways to use PC analysis on an empirical correlation matrix. Both of them were presented in the book of Piterbarg & Andersen [8].

Consider a p-dimensional vector Y, where $Y_i \sim N(0,1)$ for $i = 1, \dots, p$. Define a positive definite correlation matrix R for Y, given as

$$R = \mathbb{E}(YY^\top),$$

and let hold the following approximation

$$Y \approx DX,$$

where X is a m-dimensional vector of independent normal random variables, $m \leq p$ and D is a $(p \times m)$-dimensional matrix. Note, that DX is a vector of variables with zero means and unit variance, therefore $DD^\top$ can be interpreted as a correlation matrix. We require $DD^\top$ to have ones on its diagonal and must find an optimal D^* that minimizes the function

$$h(D;R) = tr\left((R - DD^\top)(R - DD^\top)^\top\right),$$

or in formulas

$$D^* = \underset{D}{\operatorname{argmin}}\, h(D;R), \qquad v(D) = \mathbf{1},$$

where $v(D)$ is a p-dimensional vector of the diagonal elements of $DD^\top$ and $\mathbf{1}$ is a p-dimensional vector of ones.

The following technique is proposed by Piterbarg & Andersen, and its proof can be found in their book called *Interest Rate Modeling* [8]. Consider a p-dimensional vector μ and D_μ given as the "unconstrained optimum" (see [8])

$$D_\mu = \underset{D}{\operatorname{argmin}}\, h\big(D; R + \operatorname{diag}(\mu)\big).$$

Let μ^* be the solution of

$$v\big(D_\mu\big) - \mathbf{1} = \mathbf{0}, \tag{23}$$

then $D^* = D_\mu^*$. Hence, instead of looking for an optimal D we can solve the p-dimensional root-search problem, i.e. Equation (23). Note, that for a fixed μ we can easily find D_μ using PCA (see Appendix C), where $R + \operatorname{diag}(\mu)$ is considered to be the target correlation matrix.

In the end, this algorithm returns a correlation matrix $D^*(D^*)^\top$ that has reduced rank m. There are several other algorithms for finding an optimal rank-reduced correlation matrix (see [8]). In the following, we introduce an alternative, heuristic method.

Another approach to find the rank-reduced correlation matrix can also be found in *Interest Rate Modeling* [8], called the *"Poor Man's Correlation PCA"*.

We suppose that PCA from Appendix C can be directly applied to the correlation matrix and compute the $p \times p$ matrix

$$R_m = E_m \Lambda_m (E_m)^\top,$$

where Λ_m is the $m \times m$ diagonal matrix of the m largest eigenvalues of R and E_m is the $p \times m$ matrix of eigenvector corresponding to the eigenvalues.

We need to ensure that R_m has a diagonal of ones, therefore consider the following approximation:

$$R \approx r_m^{-1} R_m r_m^{-1},$$

where

$$r_m = \begin{pmatrix} \sqrt{(R_m)_{1,1}} & 0 & \dots & 0 \\ 0 & \sqrt{(R_m)_{2,2}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sqrt{(R_m)_{p,p}} \end{pmatrix}.$$

Note, that D is therefore given as

$$D = r_m^{-1} E_m \sqrt{\Lambda_m}.$$

Although this algorithm is easier to apply, according to Piterbarg & Andersen, the difference between the two approximations is small only if R_m is close to having ones on its diagonal. In most cases, when large, complex correlation matrices are considered, the approximations highly differ.

5.6 The LMMPlus Model

The LMMPlus model introduced in [19] is an extension of the BGM model, that is used in practice for the valuation of life insurance liabilities by the Zurich Insurance Company, Bratislava. It is an LMM model, that considers a displaced lognormal formulation of the local volatility function together with the previously introduced stochastic volatility process (SVP). Hence, it has some very useful properties:

- with a displacement, the model does not only correspond to the current/post-crisis interest rate environment, but gives us a significant freedom to control the shape of the distribution.
- knowing the volatility skew, the model predicts a realistic evolution of the implied volatility and replicates forward rate dynamics more realistically. This may be substantial, for example when considering a path-dependent option.

However, it must also be said, that due to an SVP the model becomes more complex and run times increase, too.

In the LMMPlus model we consider a local volatility process of $L_n(t)$, that is given by the function $\varphi(L_n(t)) = L_n(t) + \delta$. Therefore, the dynamics of the forward LIBOR rate and the SVP in the spot measure $\mathbb{Q}^B$ for $n = 0, 1, \dots, N-1$ are

$$dL_n(t) = \sqrt{z(t)}(L_n(t) + \delta)\lambda_n(t)^\top \left(\sqrt{z(t)}\mu_n(t)dt + dW^B(t)\right)$$

$$dz(t) = \kappa\big(\theta - z(t)\big)dt + \epsilon\sqrt{z(t)}dZ^B(t),$$

where

$$\mu_n(t) = \sum_{i=q(t)}^{n} \frac{\tau_i(L_i(t) + \delta)\lambda_i(t)}{1 + \tau_i L_i(t)},$$

κ, θ and ϵ are positive constants, and the displacement factor δ must satisfy $\delta < \tau_i{}^{-1}$.

In the LMMPlus model the correlation – discussed in Section 5.5.1 – between the Wiener processes of the forward LIBOR rate dynamics and the z-process dynamics is *not* zero. Instead, we define a constant instantaneous correlation given by the formula

$$dZ^B(t) = \boldsymbol{\rho_{WZ}}^\top dW^B(t) + \sqrt{1-\rho^2}d\widehat{W}(t), \tag{24}$$

where $\boldsymbol{\rho_{WZ}}$ is an m-dimensional vector of constants $\frac{\rho}{\sqrt{m}}$ and also (for future reference) consider another instantaneous correlation $\rho_i(t)$ defined by equation

$$\rho_i(t) = \frac{\langle dZ^B(t), dY_i^B(t)\rangle}{dt}, \tag{25}$$

where $Y_i^B(t)$, given as

$$dY_i^B(t) = \frac{\lambda_i(t)}{\|\lambda_i(t)\|}\,dW^B(t),$$

is Gaussian with mean 0 and quadratic variation t, identifying it as a Wiener process (see [8]).

Note, that in this form the LMMPlus is barely a model, rather than a framework. In order to create a working model that is used in practice, we refer the reader to Chapter 7, where we complete it by presenting the theory of the LMM model calibration and define a specific structure of the deterministic component $\lambda_n(t)$ assumed by the LMMPlus model.

5.7 The Piterbarg & Andersen Model

Although, we have not specified it yet, in previous Sections we have already discussed some assumptions of another model introduced in [8]. Now we would like to present the concrete model proposed by Piterbarg & Andersen, who defined an LMM model with displaced diffusion and a stochastic volatility process in measure $\mathbb{Q}^B$ as

$$dL_n(t) = \sqrt{z(t)}\big((1-b)L_n(0) + bL_n(t)\big)\lambda_n(t)^\top\left(\sqrt{z(t)}\mu_n(t)dt + dW^B(t)\right)$$

$$dz(t) = \kappa\big(z_0 - z(t)\big)dt + \epsilon\sqrt{z(t)}dZ^B(t),$$

where

$$\mu_n(t) = \sum_{i=q(t)}^{n} \frac{\tau_i\big((1-b)L_n(0) + bL_n(t)\big)\lambda_i(t)}{1+\tau_i L_i(t)},$$

$z_0 = z(0) = 1$ and the correlation between $W^B(t)$ and $Z^B(t)$ is considered to be zero. Since $d\langle W^B(t), Z^B(t)\rangle = 0$, dynamics of the SVP remains unchanged also in other probability measures. Hence, dynamics in the T^{n+1}-forward measure are much simpler, in formulas

$$dL_n(t) = \sqrt{z(t)}\big((1-b)L_n(0) + bL_n(t)\big)\lambda_n(t)^\top dW^{n+1}(t)$$

$$dz(t) = \kappa\big(z_0 - z(t)\big)dt + \epsilon\sqrt{z(t)}dZ^{n+1}(t).$$

As we can see, the model itself does not differ extremely from the LMMPlus model. The main differences are between the local volatility functions considered by each model and the assumed values of correlations between the Wiener processes of forward rates and the SVP. This results that in the Piterbarg & Andersen model slopes at ATM strikes are controlled by the parameter b instead of non-zero correlations [1]. This model is studied in details and well-summarized in [14], where they also discuss the case of time-dependent parameters.

6 Swaption Pricing in the LMMPlus Model

In this Chapter, we derive the pricing formula for swaptions in the LMMPlus model, meaning that we consider a non-zero correlation between the d-dimensional Wiener process W^B and the Wiener process Z^B of SVP dynamics. This Chapter is mainly based on the work of Wu & Zhang (see [20]), who proposed a pricing formula for caplets using the same approach.

As we will see, the price of a swaption cannot be obtained analytically, therefore a number of approximations are made in the valuation. First, we express the price in terms of a *moment generating function*, then derive the partial differential equation satisfied by the moment generating function and eventually compute the moment generating function, which is then used to price swaptions. Before moving on to the pricing itself, we should point out, that by the term "swaption" used in this Chapter, we refer solely to the payer swaption, since the price of a receiver swaption can be derived later using the *put-call parity* identity.

Recall the definitions and results from Chapter 4 and consider T_k is the expiry time of a swaption, so that if the swaption is exercised, the cashflows take place at times $T_{k+1}, T_{k+2}, \dots, T_m$. Hence, the annuity and swap rates are given as

$$A_{k,m}(t) = \sum_{j=k}^{m-1} \tau_j P\left(t, T_{j+1}\right), \qquad S_{k,m}(t) = \frac{P(t,T_k) - P(t,T_m)}{A_{k,m}(t)}.$$

We also know:

- the swaption price at the expiry time is

  $$V_{swaption}(T_k) = A_{k,m}(T_k)\max\left\{S_{k,m}(T_k) - K, 0\right\},$$

  where K is the strike rate.
- $S_{k,m}(t)$ is a martingale in measure $\mathbb{Q}^{k,m}$, in formulas

  $$\mathbb{E}_t^A\left[S_{k,m}(T_k)\right] = S_{k,m}(t).$$
- the swaption price at time t can be expressed as

  $$V_{swaption}(t) = A_{k,m}(t)\mathbb{E}_t^{k,m}\left[\max\left\{S_{k,m}(T_k) - K, 0\right\}\right].$$

6.1 Introducing the Moment Generating Function

Note, that the expectation term in the swaption price can be expressed as

$$\mathbb{E}_t^A\left[S_{k,m}(T_k)\mathbf{1}_{S_{k,m}(T_k)>K}\right] - K\mathbb{E}_t^A\left[\mathbf{1}_{S_{k,m}(T_k)>K}\right]$$

or after some adjustments

$$S_{k,m}(t)\left(\mathbb{E}_t^{k,m}\left[e^{ln\left(\frac{S_{k,m}(T_k)}{S_{k,m}(t)}\right)}\mathbf{1}_{S_{k,m}(T_k)>K}\right]-\frac{K}{S_{k,m}(t)}\mathbb{E}_t^{A}\left[\mathbf{1}_{S_{k,m}(T_k)>K}\right]\right),$$

where

$$\mathbf{1}_{S_{k,m}(T_k)>K}=\begin{cases}1, & S_{k,m}(T_k)>K\\ 0, & \text{otherwise}\end{cases}.$$

If we define the random variable $X=ln\left(\frac{S_{k,m}(T_k)}{S_{k,m}(t)}\right)$ and the constant $x=ln\left(\frac{K}{S_{k,m}(t)}\right)$, then the price of a swaption can be rewritten, so that

$$V_{swaption}(t)=A_{k,m}(t)S_{k,m}(t)(\mathbb{E}_t^{A}[e^X\mathbf{1}_{X>x}]-e^x\mathbb{E}_t^{A}[\mathbf{1}_{X>x}]).$$

In terms of the moment generating function defined for a random variable Y as

$$M_Y(t)\equiv M(t)=\mathbb{E}[e^{tY}],$$

we can express the expectations $\mathbb{E}_t^{k,m}[\mathbf{1}_{X>x}]$ and $\mathbb{E}_t^{k,m}[e^X\mathbf{1}_{X>x}]$ as

$$\mathbb{E}_t^{k,m}[\mathbf{1}_{X>x}]=\frac{1}{2}+\frac{1}{\pi}\int_0^\infty\frac{Im\left(e^{-iux}M_X(iu)\right)}{u}du,$$

$$\mathbb{E}_t^{k,m}[e^X\mathbf{1}_{X>x}]=\frac{1}{2}+\frac{1}{\pi}\int_0^\infty\frac{Im\left(e^{-iux}M_X(1+iu)\right)}{u}du.$$

Due to the fact, that they can be demonstrated with similar arguments, we introduce the derivation only of the first equation (see Appendix D). These relations can be described by a more general proposition, which – together with its detailed proof – can be found in [21].

Since expectations are finally replaced, the price of a swaption in terms of the moment generation function is given by the formula

$$V_{swaption}(t)=A_{k,m}(t)\left[\frac{S_{k,m}(t)-K}{2}+\cdots\right.$$
$$\left.+\frac{1}{\pi}\int_0^\infty\left(S_{k,m}(t)\frac{Im\left(e^{-iux}M_X(1+iu)\right)}{u}-K\frac{Im\left(e^{-iux}M_X(iu)\right)}{u}\right)du\right].$$

Note, as we have already mentioned in Section 5.6, the LMMPlus model considers displaced forward LIBOR rates. Therefore, swap rates become displaced, too. However, this does not change the payoff and the replacement of rates by $L_n(t) + \delta$ and $S_{k,m}(t) + \delta$ leads to the same swaption pricing formula, where $X = ln\left(\frac{S_{k,m}(T_k)+\delta}{S_{k,m}(t)+\delta}\right)$ and $ln\left(\frac{K+\delta}{S_{k,m}(t)+\delta}\right)$.

As we can see, this formula contains two improper integrals with moment generating functions of the log of the forward swap rate. In order to use it in practice, we must somehow approximate the moment generating function M_X. For that, we must derive the swap rate dynamics in measure $\mathbb{Q}^{k,m}$, which is done in the following section and solve the *Kolmogorov backward equation* (see Section 6.3) satisfied by the moment generating function of the log-forward swap rate.

6.2 Swap Rate and SVP Dynamics in Measure $\mathbb{Q}^{k,m}$

In order to derive the dynamics of the forward swap rate, recall the definition of the forward swap rate introduced in Chapter 4 as

$$S_{k,m}(t) = \frac{\sum_{i=k}^{m-1} \tau_i P(t, T_{i+1}) L_i(t)}{A_{k,m}(t)} = \sum_{i=k}^{m-1} \alpha_i(t) L_i(t),$$

where $A_{k,m}(t)$ is the annuity rate defined earlier, and

$$\alpha_i(t) = \frac{\tau_i P(t, T_{i+1})}{A_{k,m}(t)}.$$

In the following, we need to find the dynamics of the forward swap rate and the corresponding stochastic volatility process in measure $\mathbb{Q}^{k,m}$. Since all the tools and techniques of measure change are introduced in earlier Chapters, we take liberties with skipping some details and walk through the derivation of the measure change in a more careless manner – meaning that some technical or computational intermediary steps are left as an “exercise” to the reader.

As we can see, the forward swap rate $S_{k,m}(t)$ is a function of forward LIBOR rates $L_k(t), L_{k+1}(t), \dots, L_{m-1}(t)$. In Chapter 4 we also showed, that it is a martingale in measure $\mathbb{Q}^{k,m}$. Hence, by applying the Itô’s Lemma we get that

$$dS_{k,m}(t) = \sqrt{z(t)} \sum_{i=k}^{m-1} \frac{\partial S_{k,m}(t)}{\partial L_i(t)} (L_i(t) + \delta) \lambda_i(t)^\top dW^{k,m}(t),$$

or in other terms

$$dS_{k,m}(t) = \sqrt{z(t)}\big(S_{k,m}(t) + \delta\big) \sum_{i=k}^{m-1} \omega_i(t)\lambda_i(t)^\top dW^{k,m}(t),$$

where $\omega_i(t)$ are the stochastic weights defined as

$$\omega_i(t) = \left(\frac{L_i(t) + \delta}{S_{k,m}(t) + \delta}\right) \times \frac{\partial S_{k,m}(t)}{\partial L_i(t)}.$$

Random weights depending on the entire forward curve prohibit the analytical treatment of these dynamics. Therefore, to get around this complication they can be approximated by freezing them in time 0, so that

$$\omega_i(t) \approx \omega_i(0). \tag{26}$$

Remember that we expressed the swap rate as a weighted sum of forward rates $L_k(t), L_{k+1}(t), \dots, L_{m-1}(t)$. We can expect that these weights $\alpha_i(t)$ in the definition of the swap rate vary little over time, therefore $\partial S_{k,m}(t)/\partial L_i(t)$ must be near-constant quantities. Although, it can depend on the chosen function $\varphi(\cdot)$, in cases when "forward curves are reasonably flat and the forward curve movements are predominantly parallel" [8], we can assume that $(L_i(t) + \delta)/\big(S_{k,m}(t) + \delta\big)$ are also close to constant. Hence, in the following we consider the swap rate dynamics given by the approximation

$$dS_{k,m}(t) \approx \sqrt{z(t)}\big(S_{k,m}(t) + \delta\big) \sum_{i=k}^{m-1} \omega_i(0)\lambda_i(t)^\top dW^{k,m}(t).$$

For the dynamics of stochastic volatility process, recall Section 5.5.1 where we defined the dynamics in measure $\mathbb{Q}^B$ by the equation

$$dz(t) = \kappa\big(\theta - z(t)\big)dt + \epsilon\sqrt{z(t)}dZ^B(t),$$

where $Z^B(t)$ is a Wiener process in measure $\mathbb{Q}^B$. The desired dynamics of the stochastic volatility process in measure $\mathbb{Q}^{k,m}$ can be expressed through techniques of measure change used in Section 5.2.

Define the stochastic process M_t satisfying

$$M_T = \left.\frac{d\mathbb{Q}^{k,m}}{d\mathbb{Q}^B}\right|_{\mathcal{F}_T} = \frac{B(0)}{A_{k,m}(0)} \frac{A_{k,m}(T)}{B(T)},$$

where the annuity rate $A_{k,m}(t)$ and the simply-compounded money market account $B(t)$ are the numeraires of the corresponding measures. Since it is a martingale in measure $\mathbb{Q}^B$, we get

$$M_t = \frac{1}{A_{k,m}(0)} \frac{A_{k,m}(t)}{B(t)}.$$

We assume the stochastic process M_t is "well-behaved", meaning that The Martingale Representation Theorem is applicable and M_t can be expressed in form of an exponential martingale (see Section 5.2), such that

$$d\ln(M_t) = v(t)^\top dW^{k,m}(t) - \frac{1}{2}v(t)^\top v(t)dt.$$

We would like to find the diffusion coefficient $DC[\cdot]$ of the log of the stochastic process M_t, so we can eventually apply the Girsanov's Theorem. Clearly,

$$d\ln(M_t) = d\left[\ln\left(A_{k,m}(t)\right) - \ln\big(B(t)\big)\right] = d\left[\ln\left(A_{k,m}(t)\right) - \ln\left(P\big(t, T_{q(t)}\big)\right)\right],$$

therefore the formula for the desired diffusion coefficient is given as

$$DC[\ln(M_t)] = DC\left[\ln\left(A_{k,m}(t)\right)\right] - DC\left[\ln\left(P\big(t, T_{q(t)}\big)\right)\right].$$

The first term can be obtained from

$$d\ln\left(A_{k,m}(t)\right) = \frac{dA_{k,m}(t)}{A_{k,m}(t)} = \frac{\sum_{j=k}^{m-1} \tau_j dP\big(t, T_{j+1}\big)}{A_{k,m}(t)} = \sum_{j=k}^{m-1} \frac{\tau_j\, P\big(t, T_{j+1}\big)}{A_{k,m}(t)} d\ln\left(P\big(t, T_{j+1}\big)\right)$$

$$= \cdots + \sum_{j=k}^{m-1} \alpha_j(t) DC\left[\ln\left(P\big(t, T_{j+1}\big)\right)\right] dZ^{k,m}(t).$$

If we change the notation, so that $\sigma_A = DC\left[\ln\left(A_{k,m}(t)\right)\right]$ and $\sigma_{j+1}^P = DC\left[\ln\left(P\big(t, T_{j+1}\big)\right)\right]$ for $j = k, k+1, \dots, m-1$, the diffusion coefficient $DC[\ln(M_t)]$ can be given by the formula

$$DC[\ln(N_t)] = \sigma_A - \sigma_{q(t)}^P,$$

where $\sigma_A = \sum_{j=k}^{m-1} \alpha_j(t) \sigma_{j+1}^P$ and $\sum_{j=k}^{m-1} \alpha_j(t) = 1$. Hence,

$$DC[\ln(M_t)] = \sum_{j=k}^{m-1} \alpha_j(t)\big(\, \sigma_{j+1}^P - \sigma_{q(t)}^P\big),$$

Obviously, the difference of the diffusion coefficients can be expressed easily by using Equation (14), i.e. the definition of the zero-coupon bond price, so that

$$\sigma_{j+1}^P - \sigma_{q(t)}^P = DC\left[\ln\left(\frac{P\big(t, T_{j+1}\big)}{P\big(t, T_{q(t)}\big)}\right)\right] = DC\left[\sum_{i=q(t)}^{j} \ln\big(1 + \tau_i L_i(t)\big)\right].$$

This term could look familiar, since in Section 5.2 we have already showed that

$$DC\big[\ln\big(1 + \tau_i L_i(t)\big)\big] = \frac{\tau_i(L_i(t) + \delta)\sqrt{z(t)}\lambda_i(t)}{1 + \tau_i L_i(t)}.$$

Therefore, the difference $\sigma^P_{j+1} - \sigma^P_{q(t)}$ can be substituted by the term

$$\sum_{i=q(t)}^{j} \frac{\tau_i (L_i(t) + \delta)\sqrt{z(t)}\lambda_i(t)}{1 + \tau_i L_i(t)},$$

by which we have finally got the diffusion process

$$v(t) = \sigma_A - \sigma^P_{q(t)} = \sum_{j=k}^{m-1} \left(\alpha_j(t) \sum_{i=q(t)}^{j} \frac{\tau_i (L_i(t) + \delta)\sqrt{z(t)}\lambda_i(t)}{1 + \tau_i L_i(t)} \right),$$

and can apply the Girsanov's Lemma to define $W^A(t)$ by the formula

$$dW^{k,m}(t) = dW^B(t) + \sum_{j=k}^{m-1} \alpha_j(t)\sqrt{z(t)} \sum_{i=q(t)}^{j} \frac{\tau_i (L_i(t) + \delta)\lambda_i(t)}{1 + \tau_i L_i(t)} dt.$$

However, we must consider the correlations between the numeric Wiener process $Z^B(t)$ and the d-dimensional vector $W^B(t)$ of Wiener processes. Recall Equations (24) and (25) from Section 5.6, where we also defined the d-dimensional vector $\boldsymbol{\rho_{ZW}} = \left(\frac{\rho}{\sqrt{d}}, \quad \dots, \quad \frac{\rho}{\sqrt{d}}\right)^\top$. Then $Z^B(t)$ is defined as

$$dZ^B(t) = \boldsymbol{\rho_{ZW}}^\top dW^B(t) + \sqrt{1-\rho^2} d\widehat{W}(t)$$

where $\widehat{W}(t)$ is a scalar Wiener process independent from $W^B(t)$, and the correlation $\rho_i(t)$ as

$$\rho_i(t) = \frac{\langle dZ^B(t), \left(\frac{\lambda_i(t)}{\|\lambda_i(t)\|}\right)^\top dW^B(t) \rangle}{dt}.$$

Using $dW^B(t)$ obtained from the Girsanov's Theorem we can write

$$dZ^B(t) = \boldsymbol{\rho_{WZ}}^\top \left(dW^{k,m}(t) - \sum_{j=k}^{m-1} \alpha_j(t)\sqrt{z(t)} \sum_{i=q(t)}^{j} \frac{\tau_i (L_i(t) + \delta)\lambda_i(t)}{1 + \tau_i L_i(t)} dt \right)$$
$$+ \sqrt{1-\rho^2} d\widehat{W}(t).$$

Note, that

$$\boldsymbol{\rho_{ZW}}^\top \lambda_i(t) \ = \rho_i(t)\|\lambda_i(t)\|,$$

hence the formula can be rewritten as

$$dZ^B(t) = dZ^{k,m}(t) - \sum_{j=k}^{m-1} \alpha_j(t)\sqrt{z(t)}\xi_j(t)\, dt,$$

where

$$dZ^{k,m}(t) = \boldsymbol{\rho_{ZW}}^\top dW^{k,m}(t) + \sqrt{1-\rho^2} d\widehat{W}(t),$$

$$\xi_j(t) = \sum_{i=q(t)}^{j} \frac{\tau_i(L_i(t)+\delta)\rho_i(t)\|\lambda_i(t)\|}{1+\tau_i L_i(t)}.$$

Finally, the desired dynamics of $z(t)$ in measure $\mathbb{Q}^{k,m}$ are given by the formula

$$dz(t) = \kappa\left(\theta - \tilde{\xi}(t)z(t)\right)dt + \epsilon\sqrt{z(t)}dZ^{k,m}(t),$$

where

$$\tilde{\xi}(t) = 1 + \frac{\epsilon}{\kappa}\sum_{j=k}^{m-1} \alpha_j(t)\xi_j(t).$$

Note, that $\xi_j(t)$ depends on the forward LIBOR rates $L_i(t)$, which – similarly to the weights $\omega_i(t)$ discussed earlier in the swap rate dynamics – prevents us valuating swaptions analytically. However, as it is discussed by Wu & Zhang in [20], we can freeze forward LIBOR rates in time 0, so that $\xi_j(t)$ is approximated as

$$\xi_j(t) \approx \sum_{i=q(t)}^{j} \frac{\tau_i(L_i(0)+\delta)\rho_i(t)\|\lambda_i(t)\|}{1+\tau_i L_i(0)}.$$

As mentioned earlier, we also expect that the coefficients $\alpha_j(t)$ vary little over time. Therefore, the last approximation we use is

$$\alpha_j(t) \approx \alpha_j(0). \tag{27}$$

Hence, the approximated dynamics of $z(t)$ are given as

$$dz(t) \approx \kappa\left(\theta - \tilde{\xi}(t)z(t)\right)dt + \epsilon\sqrt{z(t)}dZ^{k,m}(t), \tag{28}$$

where

$$\tilde{\xi}(t) \approx 1 + \frac{\epsilon}{\kappa}\sum_{j=k}^{m-1} \alpha_j(0)\xi_j(t).$$

Results of this Section are summarized in the box below. In the following, we move towards solving the final value problem for the moment generating function.

Using Equations (26) and (27) approximated dynamics of the forward swap rate and the corresponding stochastic volatility process in the LMMPlus model are given by formulas

$$dS_{k,m}(t) = \sqrt{z(t)}\big(S_{k,m}(t) + \delta\big)\tilde{\lambda}_{k,m}(t)^{\top} dW^{k,m}(t)$$

$$dz(t) = \kappa\left(\theta - \tilde{\xi}(t)z(t)\right)dt + \epsilon\sqrt{z(t)}dZ^{k,m}(t),$$

where the instantaneous correlation between the d-dimensional $W^{k,m}(t)$ and the numeric $Z^{k,m}(t)$ is given by the formula

$$dZ^{k,m}(t) = \frac{\rho}{\sqrt{d}}\left(dW^{(k,m),1}(t) + \cdots + dW^{(k,m),d}(t)\right) + \sqrt{1-\rho^2}\,d\widehat{W}(t)$$

and

$$\tilde{\lambda}_{k,m}(t) \approx \sum_{i=k}^{m-1} \omega_i(0)\lambda_i(t), \qquad \omega_i(t) = \left(\frac{L_i(t) + \delta}{S_{k,m}(t) + \delta}\right) \times \frac{\partial S_{k,m}(t)}{\partial L_i(t)}$$

$$\tilde{\xi}(t) \approx 1 + \frac{\epsilon}{\kappa}\sum_{j=k}^{m-1} \alpha_j(0)\xi_j(t), \qquad \xi_j(t) \approx \sum_{i=q(t)}^{j} \frac{\tau_i(L_i(0)+\delta)\rho_i(t)\|\lambda_i(t)\|}{1+\tau_i L_i(0)}.$$

6.3 Final Value Problem for the Moment Generating Function

Define the expectation of function of a process $u(X_t, t)$ that represents the moment generating function of X_T as

$$u(X_t, t) = \mathbb{E}_t[e^{lX_T}],$$

where $\mathbb{E}_t$ is the expectation with the filtration $\mathcal{F}_t$. We can easily show that $u(X_t, t)$ is a martingale: for $t > s$

$$\mathbb{E}_s[u(X_t, t)] = \mathbb{E}_s\Big[\mathbb{E}_t[e^{lX_T}]\Big] = \mathbb{E}_s[e^{lX_T}] = u(X_s, t),$$

therefore it must be driftless. Let X_T be the log of the function $\varphi\left(S_{k,m}(t)\right) = S_{k,m}(t) + \delta$ at expiry. The moment generating function

$$\phi(X_t, z(t), t, l) = \mathbb{E}_t[e^{lX_T}]$$

must satisfy the Kolmogorov backward equation (see [20] and [22])

$$\frac{\partial \phi}{\partial t} + \mu_X \frac{\partial \phi}{\partial x} + \mu_z \frac{\partial \phi}{\partial z} + \frac{1}{2}\sigma_X^2 \frac{\partial^2 \phi}{\partial x^2} + \rho_{Xz}\sigma_X\sigma_z \frac{\partial^2 \phi}{\partial x \partial z} + \frac{1}{2}\sigma_z^2 \frac{\partial^2 \phi}{\partial z^2} = 0,$$

where $\mu_X, \mu_z, \sigma_X, \sigma_z$ are the drifts and volatilities of the processes X_t and $z(t)$ and ρ_{Xz} is their correlation.

Applying the Itô's Lemma we get

$$dX_t = d\ln\left(\varphi\left(S_{k,m}(t)\right)\right) = -\frac{1}{2}z(t)\left\|\tilde{\lambda}_{k,m}(t)\right\|^2 dt + \sqrt{z(t)}\tilde{\lambda}_{k,m}(t)^\top dW^{k,m}(t),$$

which together with the dynamics of $z(t)$ derived in the previous Section give us the desired drifts and volatilities as

$$\mu_X = -\frac{z(t)}{2}\left\|\tilde{\lambda}_{k,m}(t)\right\|^2$$

$$\mu_z = \kappa\left(\theta - \tilde{\xi}(t)z(t)\right)$$

$$\sigma_X^2 = z(t)\left\|\tilde{\lambda}_{k,m}(t)\right\|^2$$

$$\sigma_z^2 = \epsilon^2 z(t),$$

and the following equation:

$$\rho_{Xz}\sigma_X\sigma_z = \epsilon z(t)\sum_{i=k}^{m-1}\omega_i(0)\|\lambda_i(t)\|\rho_i(t),$$

because

$$\rho_{Xz} = \frac{\langle \frac{\tilde{\lambda}_{k,m}}{\left\|\tilde{\lambda}_{k,m}\right\|}^\top dXdz\rangle(t)}{\sqrt{\langle \frac{\tilde{\lambda}_{k,m}(t)^\top}{\left\|\tilde{\lambda}_{k,m}(t)\right\|} dX(t)\rangle^2 \langle dz(t)\rangle^2}} = \frac{\sigma_X\sigma_z \frac{\tilde{\lambda}_{k,m}(t)^\top}{\left\|\tilde{\lambda}_{k,m}(t)\right\|} dW^B(t)dZ^B(t)}{\sigma_X\sigma_z}$$

$$= \frac{\tilde{\lambda}_{k,m}(t)^\top}{\left\|\tilde{\lambda}_{k,m}(t)\right\|} dW^B(t)dZ^B(t) = \frac{\sum_{i=k}^{m-1}\omega_i(0)\lambda_i(t)^\top dW^B(t)dZ^B(t)}{\left\|\tilde{\lambda}_{k,m}(t)\right\|}$$

$$= \frac{\sum_{i=k}^{m-1}\omega_i(0)\|\lambda_i(t)\|\rho_i(t)}{\left\|\tilde{\lambda}_{k,m}(t)\right\|}.$$

Also note, that the moment generating function satisfies the final condition, i.e.

$$\phi(X_T, z(T), T, l) = \mathbb{E}_T[e^{lX_T}] = e^{lX_T}.$$

In general, when solving the Kolmogorov backward equation we look for solutions of the following form:

$$\phi(X, z, t, l) = e^{A(\tau,l)+B(\tau,l)z+lx},$$

where $\tau = T - t$, and functions A and B are obtainable analytically for constant coefficients (see [20] or [23]). Substituting this functional form in Kolmogorov's partial differential equation and using

$$\frac{\partial\phi}{\partial\tau} = -\phi\left(\frac{\partial A}{\partial\tau} + z\frac{\partial B}{\partial\tau}\right), \qquad \frac{\partial\phi}{\partial x} = \phi l, \qquad \frac{\partial\phi}{\partial z} = \phi B, \qquad \frac{\partial^2\phi}{\partial x^2} = \phi l^2, \qquad \frac{\partial^2\phi}{\partial z^2} = \phi B^2$$

$$\frac{\partial^2 \phi}{\partial x \partial z} = \phi B l,$$

gives us two ordinary differential equations

$$\frac{dA}{d\tau} = \kappa\theta B,$$

$$\frac{dB}{d\tau} = b_2 B^2 + b_1 B + b_0,$$

where

$$b_2 = \frac{1}{2}\epsilon^2, \qquad b_1 = \epsilon l \sum_{i=k}^{m-1} \omega_i(0)\|\lambda_i(\tau)\|\rho_i(\tau) - \kappa\tilde{\xi}(\tau), \qquad b_0 = \frac{1}{2}\left\|\tilde{\lambda}_{k,m}(\tau)\right\|^2 (l^2 - l).$$

The second ordinary differential equation is a Riccatti equation, for which an analytical solution is available in case of constant coefficients. However, if we assume coefficients constants in a time interval (τ_j, τ_{j+1}), i.e. piece-wise constants, the analytical solutions can be extended through recursion.

This system of two ordinary differential equations with subject to the initial conditions

$$A(0) = A_0, \qquad B(0) = B_0,$$

can be solved by transforming the Riccatti equation into a linear first order equation (see *Remark 4*) to get B and integrating the first differential equation afterwards to get A, so that

$$A(\tau) = A_0 + \kappa\theta\left(\frac{-b_1 + d_b}{2b_2}\tau - \frac{1}{b_2}\ln\left(\frac{1 - he^{d\tau}}{1 - h}\right)\right),$$

$$B(\tau) = B_0 + \left(\frac{-b_1 + d_b}{2b_2} - B_0\right)\left(\frac{1 - e^{d\tau}}{1 - he^{d\tau}}\right),$$

where

$$d_b = \sqrt{b_1^2 - 4b_2 b_0}, \qquad h = \frac{2B_0 b_2 + b_1 - d_b}{2B_0 b_2 + b_1 + d_b}.$$

Finally, considering the case of piece-wise constant coefficients, let

$$A_0 = A(\tau_j, l),$$

$$B_0 = B(\tau_j, l)$$

and replace τ with $\tau - \tau_j$, so that we can arrive at final solutions

$$A(\tau, l) = A(\tau_j, l) + \kappa\theta\left(B_+(\tau - \tau_j) - \frac{2}{\epsilon^2}\ln\left(\frac{1 - h_j e^{d_b(\tau - \tau_j)}}{1 - h_j}\right)\right),$$

$$B(\tau,l) = B(\tau_j,l) + \left(B_+ - B(\tau_j,l)\right)\left(\frac{1-e^{d_b(\tau-\tau_j)}}{1-h_j e^{d_b(\tau-\tau_j)}}\right),$$

where $B_+ = \frac{-b_1+d_b}{2b_2}$ is the particular solution of the Riccatti equation (see *Remark 4*) and

$$h_j = \frac{2b_2 B(\tau_j,l) + b_1 - d_b}{2b_2 B(\tau_j,l) + b_1 + d_b}.$$

Hence, we have found the approximate moment generating function, which can be substituted for $M_X(t)$ in the swaption price formula derived in Section 6.1. In the following Remark we present a possible method to transform the Ricatti differential equation into an integrable form. For further steps of the derivation of $A(\tau,l)$ and $B(\tau,l)$ see Appendix E.

Remark 4 *For the (Riccatti) differential equation*

$$\frac{dB}{d\tau} = b_2 B^2 + b_1 B + b_0,$$

we can look for solutions in form

$$B = Y_1 + Y_2.$$

Setting $d_b = \sqrt{b_1^2 - 4b_2 b_0}$ *the particular solution* Y_1 *is given as*

$$Y_1 = \frac{-b_1 \pm d_b}{2b_2},$$

from which – without loss of generality – we can take the one with the "+" sign (see [20]*) and denote it by* B_+*. Since the general solution satisfies*

$$Y_2 = B - B_+,$$

the following must hold for its derivative:

$$\frac{dY_2}{d\tau} = \frac{d(Y_2 + B_+)}{d\tau} = b_2(Y_2 + B_+)^2 + b_1(Y_2 + B_+) + b_0.$$

Using the fact, that the particular solution satisfies $b_2 {B_+}^2 + b_1 B_+ + b_0 = 0$*, we get a new equation*

$$\frac{dY_2}{d\tau} = b_2 Y_2^2 + d_b Y_2,$$

which is from the class of Bernoulli equations, and therefore has an explicit solution. We can simply use the substitution $u = Y_2^{-1}$ *and convert it into the linear differential equation*

$$\frac{du}{d\tau} + d_b u = -b_2,$$

that allows integration.

7 Calibration

In order to create a useful model from the framework discussed in Chapter 5, we must define the structure of $\lambda_n(t)$. Once it is done, the model can be calibrated to market-observable prices of chosen *calibration instruments*, so that we can find optimal parameters in our model. Our aim in this Chapter is to present the issues of the LMM model calibration and propose two possible calibration methods:

- a grid-based process proposed by Piterbarg & Andersen [8],
- a specific approach introduced in the LMMPlus model calibration methodology.

As we will see, these methods are very different. While the first one considers exogenously given parameters for the stochastic volatility dynamics and the second one defines the correlation of forward LIBOR rates through a parametric function. First we present the grid-based construction of the time-to-maturity component $\lambda_n(t)$ in the Piterbarg & Andersen model, then we show the process considered by the LMMPlus model, which results an economically interpretable framework that can be easily applied and communicated.

7.1 Construction of $\boldsymbol{\lambda_n(t)}$

The only thing we have not discussed in Chapter 5 is the choice of deterministic factor volatility components, i.e. the m-dimensional vectors $\lambda_k(t)$ for $k = 1, \dots, N-1$, which – as we showed in Section 5.5.2 – are closely related to correlations between forward LIBOR rates and the volatility of their processes. In this Section, first we would like to guide the reader through the steps of the calibration proposed in [8], which also underlies the framework of another work (see [14]) that we often refer to. Then we introduce another approach of choosing the time-to-maturity component proposed in the LMMPlus model calibration methodology.

7.1.1 The Piterbarg & Andersen Approach

Now we introduce a general approach proposed by Piterbarg & Andersen. Note, in this method we suppose the stochastic volatility dynamics are given exogenously, as trader's input. Then according to Piterbarg & Andersen [8], we can assume that the time-to-maturity component is expressed by the following funcions:

$$\lambda_k(t) = h(t, T_k - t), \qquad \|\lambda_k(t)\| = g(t, T_k - t),$$

where $h : \mathbb{R}_+^2 \longrightarrow \mathbb{R}^m$ and $g : \mathbb{R}_+^2 \longrightarrow \mathbb{R}$. We consider the parametrization of g using the Rebonato function, i.e. the following parametric form:

$$g(t,x) = g(x) = (a + bx)e^{-cx} + d,$$

where $a, b, c, d \in \mathbb{R}_+$. As we will see in the next Section, this also coincides with the assumptions of the LMMPlus model. Note, this specification is not t-dependent – it only depends on the time to maturity $(T_k - t)$. Although in this thesis we keep ourselves to this parametrization, we should mention that in order to include dependency on time t, function g can be also expressed as a multiplication of two functions depending on the time and the time to maturity separately (see [8]).

It is common to consider function $\|\lambda_k(t)\|$ to be a piecewise constant in time t, with discontinuities at $T_n,\ n = 1, \dots, N-1$,

$$\|\lambda_k(t)\| = \sum_{n=1}^{k} 1_{\{T_n - 1 \le t \le T_n\}} \|\lambda_{n,k}\| = \sum_{n=1}^{k} 1_{\{q(t)=n\}} \|\lambda_{n,k}\|,$$

where $\|\lambda_{n,k}\|$ is constructed from a $(N_t \times N_x)$ -dimensional matrix G using a two-dimensional interpolation method, so that for a rectangular grid of time and tenors $\{t_i\} \times \{x_j\}$

$$G_{i,j} = (a + bx_j)e^{-cx_j} + d,$$

where $i = 1, \dots, N_t, j = 1, \dots, N_x$.

If we suppose a simple linear interpolation done separately in both dimensions, then for $1 \le n \le k \le N-1$

$$\|\lambda_{n,k}\| = w_{++}G_{i,j} + w_{+-}G_{i,j-1} + w_{-+}G_{i-1,j} + w_{--}G_{i-1,j-1},$$

where

$$w_{++} = \frac{(T_{n-1} - t_{i-1})(\tau_{n,k} - x_{j-1})}{(t_i - t_{i-1})(x_j - x_{j-1})}, \qquad w_{+-} = \frac{(T_{n-1} - t_{i-1})(x_j - \tau_{n,k})}{(t_i - t_{i-1})(x_j - x_{j-1})},$$

$$w_{-+} = \frac{(t_i - T_{n-1})(\tau_{n,k} - x_{j-1})}{(t_i - t_{i-1})(x_j - x_{j-1})}, \qquad w_{--} = \frac{(t_i - T_{n-1})(x_j - \tau_{n,k})}{(t_i - t_{i-1})(x_j - x_{j-1})}.$$

Now, define the $(N-n) \times (N-n)$ instantaneous correlation matrix $R(T_n)$ for a fixed T_n as

$$R(T_n) = \mathrm{Corr}\left(dL_i(T_n), dL_j(T_n)\right), \qquad i,j = n, \dots, N-1,$$

which can be computed, for instance, using the parametric form presented in Section 5.5.2.1 and consider a diagonal volatility matrix $c(T_n)$ with elements

$$(c(T_n))_{i,j} = \begin{cases} \|\lambda_{n,n+j-1}\|, & i = j \\ 0, & i \neq j \end{cases}.$$

Hence, an instantaneous covariance matrix $C(T_n)$ can be computed as

$$C(T_n) = c(T_n)R(T_n)c(T_n).$$

If we let $H(T_n)$ be an $(N-n) \times m$ matrix with elements $(H(T_n))_{j,i} = h_i(T_n, T_{n+j-1} - T_n)$ for $j = 1, \dots, N-n$ and $i = 1, \dots, m$, then it should follow that

$$C(T_n) = H(T_n)H(T_n)^\top.$$

Therefore, we have two representations of the covariance matrix, so that we can write the equation

$$H(T_n)H(T_n)^\top = c(T_n)R(T_n)c(T_n), \tag{29}$$

from which we can construct $H(T_n)$, i.e. the set of vectors $h(T_n, T_{n+j-1} - T_n)$ for the whole grid, and using that, the full set of $\lambda_k(t)$.

In general, Equation (29) does not have a solution. As it is stated by Piterbarg & Andersen (see [8]), while normally the left-hand side does not have a full rank, the right-hand side does. Therefore, the PCA decomposition (see Appendix C) should be applied. Although, it can be performed in many ways, we introduce the “Correlation PCA” [8], which is strongly recommended by Piterbarg & Andersen (see Remark below) and discussed earlier in Section 5.5.2.2.

By applying the PCA we have

$$R(T_n) = D(T_n)D(T_n)^\top,$$

where the $(N-n) \times m$ matrix D can be found by following the steps of Section 5.5.2.2. Hence, we can replace $R(T_n)$ in Equation (29) and define $H(T_n)$ as

$$H(T_n) = c(T_n)D(T_n).$$

Remark 5 *Note, that the previous decomposition is independent from the volatility matrix $c(T_n)$. Hence, updates made on guesses of the elements of matrix G do not affect the correlation PCA. For a more detailed justification, we refer the reader to* [8].

7.1.2 Construction of $\lambda_n(t)$ in the LMMPlus Model

In the LMMPlus model, a concrete form of $\lambda_n(t)$ is defined through the Rebonato function scaling a set of factor loadings, that need to be estimated using the PCA.

The m-dimensional time-to-maturity component $\lambda_n(t)$ of the forward LIBOR rate volatility is given as

$$\lambda_n(t) = \begin{cases} 0, & x \leq 0 \\ \beta_n(x)g(x), & x \in (0, T_n], \\ \beta(T_n)g(T_n), & x > T_n \end{cases}$$

where $x = T_n - t$, $g(x)$ is the exponential Rebonato function defined earlier with parameters $a, b, c, d \in \mathbb{R}$ and $\beta_n(x)$ is an m-dimensional vector – also called the factor loading – with elements β_n^i, $i = 1, \dots, m$, that are constant over the interval $x \in (T_{n-1}, T_n]$. For a better understanding, the elements can be defined using a different approach such as

$$\beta^i(x) = \begin{cases} \beta_1^i, & x \in (0, T_1] \\ \beta_2^i, & x \in (T_{n-1}, T_n] \\ & \dots \\ \beta_n^i, & x \in (T_{n-1}, T_n] \\ & \dots \\ \beta_N^i, & x \in (T_{N-1}, T_N] \end{cases}, \qquad i = 1, \dots, m.$$

Now, considering this parametrization, if we recall the forward LIBOR rate dynamics (see Equation (19) in Chapter 5), then the component containing the Wiener process for $x \in (0, T_n]$ is

$$\sigma_n(t)^\top dW^B = \varphi\big(L_n(t)\big)\lambda_n(t)^\top dW^B = g(x)\varphi\big(L_n(t)\big)\beta_n(x)^\top dW^B.$$

Therefore β_n^i can be thought as an exposure of the forward LIBOR rate $L_n(t)$ to the i-th component of the m-dimensional Wiener process W^B. Hence, we refer to this set of vectors as factor loadings.

We can show that the choice of this parametric form gives us a very simple representation of the forward LIBOR rate correlations. Indeed, it is given via a small number m of factor loadings. However, to conform the framework presented earlier, first we need $\lambda_n(t)$ to satisfy the assumption

$$\|\lambda_n(t)\| = g(x),$$

so that the factor loadings do not have an impact on the forward LIBOR rate volatility. This can be simply achieved by normalizing the factor loadings, such that

$$\|\beta_n(x)\|^2 = \big(\beta_n^1(x)\big)^2 + \dots + \big(\beta_n^m(x)\big)^2 = 1.$$

Now, recall Section 5.5.2, where we derived the instantaneous correlation between two forward LIBOR rates as

$$\mathrm{Corr}\big(dL_n(t), dL_k(t)\big) = \frac{\lambda_n(t)^\top \lambda_k(t)}{\|\lambda_n(t)\| \|\lambda_k(t)\|}.$$

Then, using the parametrization presented earlier, the instantaneous correlation between $L_n(t)$ and $L_k(t)$ is given by

$$\mathrm{Corr}\big(dL_n(t), dL_k(t)\big) = \beta^1_{q(T_n-t)}\beta^1_{q(T_k-t)} + \cdots + \beta^m_{q(T_n-t)}\beta^m_{q(T_k-t)},$$

meaning that factor loadings describe the correlation between movements of two points on the zero-coupon curve. Using PCA, we can choose the most efficient set of factor loadings, that describes the LMMPlus model correlations and captures our target correlations.

According to the LMMPlus model [24], first two factors from the PCA can explain approximately 90% of the variance, meaning that by applying the algorithm discussed in Appendix C, we decided to set the Wiener process dimension $m = 2$. The first factor influences each point on the curve along the same direction, i.e. describes the parallel shift, while the second one can be interpreted as the "tilt" factor, which affects the short and long end of the curve oppositely, in different degrees. Hence, we refer to the second factor as the short term one (corresponding to the short rate volatility), since it has greater effect on the short term yields. Similarly, the first factor is associated with the long term volatility, i.e. we refer to it as the long term factor.

In the case, when we insist on the interpretation of factor one and two as the short and long term factors respectively, another approach of using the PCA can be considered. Instead of deriving loading factors from the target correlation matrix discussed in Section 5.5.2.1, we can construct another correlation matrix from a simpler model, the extended *two-factor Black-Karasinki (BK) model*, for which we refer the reader to [25]. At this point, we would like to underline, that doing this results a less accurate fit to target correlations, however, the LMMPlus model favors the fact of obtaining loading factors, whose influence on rates is easy to communicate through wider business. Note, that we can consider short rates equivalent to spot rates of up to (approximately) one year maturity, due to the negligible effect of the long rate factor at these maturities. and long rate equivalent to the spot rate at the maturity, where the short rate factor is zero (for example, 10 years).

Using the BK method, we can model yearly forward rates, for example, to thirty years [24] and derive a correlation matrix. As we said, the second factor has higher weighting on the short end of the curve and lower on longer maturities. Hence, if we apply PCA similar to "Poor Man's" (see Section 5.5.2.2) on the correlation matrix derived from the BK model, then we can choose the short rate factor candidate $F^1(n)$ as the second principal component (eigenvector) scaled by the second eigenvalue. Note, that (after constructing the first factor loading) the second factor loading can be simply determined using the equation

$$(\beta_n^1)^2 + (\beta_n^2)^2 = 1.$$

According to the LMMPlus model methodology, the first factor loadings can be modeled as Rebonato parametric forms minimizing the objective function

$$OF = \sum_{n=1}^{29} \left(\beta_n^1 - F^1(n)\right)^2,$$

such that a minimum is found satisfying

$$\beta_n^1 = (a + bn)e^{-cn} + d$$

and

$$\beta_n^2 = \sqrt{1 - (\beta_n^1)^2}.$$

Hence, we can determine all the Rebonato parameters and also loading factors β_n^1 and β_n^2 for $n = 1, \dots, 29$.

An advantage of this approach is that we can extrapolate factor loadings, for example, to 120 years using the same optimized Rebonato function or interpolate them for semi-annually or monthly steps, so that it allows smooth and complete correlations. Note, that we have also got factors one and two as we required, driving short and long term structure of interest rates respectively.

We would like to draw attention to the notation of the parameters in the factor loadings' Rebonato form, that can be misleading: they do not correspond to those in function g introduced earlier in this Section.

7.1.3 The Rebonato Function

Recall the time-to-maturity component $\lambda_n(t)$ discussed earlier in this Chapter and defined by scaling the (principal component) factor loading by a Rebonato function

$$g(x) = (a + bx)e^{-cx} + d,$$

where $a, b, c, d \in \mathbb{R}_+$. For a better understanding, at the end of this Section we dedicate a few words to the nature of the calibration parameters emerged from the Rebonato function. For that, we also refer the reader to Figure 4, a visual representation of the Rebonato parameters.

Since $\lim_{x \to \infty}(a + bx)e^{-cx} = 0$, parameter d is the long-term value of time-to-maturity component of the forward LIBOR rate volatility. Also, if $x = 0$, the value of the function is equal to $(a + d)$, which can be interpreted as the initial value of the forward LIBOR rate volatility. The parameters b and c influence the behavior of the function, while c is also largely responsible for the curvature. These two parameters, together with a determine whether a hump occurs and if it does, its location is given by $(b - ac)/bc$.

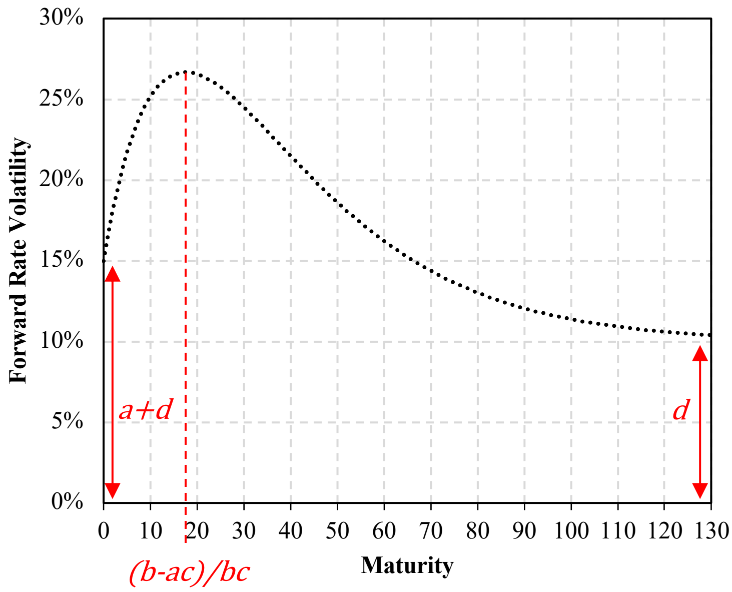

Figure 4 *A Representation of The Rebonato Function's Parameters*

7.2 Data Acquisition and Pre-Processing

The first stage of the calibration process is the data acquisition and pre-processing – one must be specific about the input parameters, market data and a set of target prices before calibrating the model. Therefore, in this Section we discuss some possible ways of obtaining a continuous initial yield curve $L_n(0)$, the choice of calibration instruments and propose a concrete value of the displacement factor based on expert judgements from the internal documents of Zurich Insurance Company.

7.2.1 Initial Yield Curve

One of the inputs that is assumed by all LMM models is the initial yield curve $L_n(0)$, which must be recovered somehow from market data. For that, first we have to define the risk-free reference rate, which can be derived for example from government bonds. The problem is, that in general not all the required rates are directly observable on market. Therefore, we must construct a curve that matches the market rates where we want them and use an estimation method to get a complete set of rates.

Instead of deriving a whole interpolation method, we refer the reader to *Interpolation Methods For Curve Construction* [26] written by Hagan & West, where the authors present and compare a wide range of yield curve interpolation algorithms. One can also use simple theoretical term structure methods, such as Vasicek [27]. However, in case of theoretical methods the yield curve is determined to take a shape from basic categories, therefore using an empirical method is recommended, for example the *monotone convex spline interpolation* proposed by Hagan & West and applied in [6]. Piterbarg & Andersen [7] also dedicated a whole chapter to the theory of constructing yield curves.

Two other alternative interpolation methods are the Smith-Wilson method (used in the LMMPlus model [24]) and the Nelson-Siegel method (see for example [28] and [29]), which are used in practice for both interpolation and extrapolation, i.e. when the long-term forward rate needs to be determined.

In the following, we introduce one of the simpler methods, the Nelson-Siegel model, which was modified by Svensson (see [28]), who added two more parameters to the original model redefining the formula for the instantaneous forward rate as

$$f(t,T) = \beta_0 + \beta_1 e^{-\frac{\tau(t,T)}{\varrho_1}} + \beta_2 \frac{\tau(t,T)}{\varrho_1} e^{-\frac{\tau(t,T)}{\varrho_1}} + \beta_3 \frac{\tau(t,T)}{\varrho_2} e^{-\frac{\tau(t,T)}{\varrho_2}},$$

where $\beta_0, \beta_1, \beta_2, \beta_3, \varrho_1, \varrho_2$ are estimated using the least-squares method. It is a model that is often used in practice and can properly fit long maturities. Using the definitions from Chapter 2, we can transform the equation, so that for the implied spot rate we get

$$R(t,T) = \beta_0 + \beta_1 \left(\frac{1 - e^{-\frac{\tau(t,T)}{\varrho_1}}}{\frac{\tau(t,T)}{\varrho_1}} \right) + \beta_2 \left(\frac{1 - e^{-\frac{\tau(t,T)}{\varrho_1}}}{\frac{\tau(t,T)}{\varrho_1}} - e^{-\frac{\tau(t,T)}{\varrho_1}} \right)$$
$$+ \beta_3 \left(\frac{1 - e^{-\frac{\tau(t,T)}{\varrho_2}}}{\frac{\tau(t,T)}{\varrho_2}} - e^{-\frac{\tau(t,T)}{\varrho_2}} \right).$$

Obviously, the Nelson-Siegel-Svensson method can be easily applied, however, as we can see on the following Figure – where we used NSS model to construct an estimated yield curve, based only on a few German government bonds from April 13, 2020 – it is far from being perfect. Deviations from the market are inevitable because of the least-squares minimization when searching parameters, meaning that the model cannot exactly replicate the yield curve. Therefore, when all the market points are required to be included in the yield term structure, a spline approach mentioned earlier is more convenient.

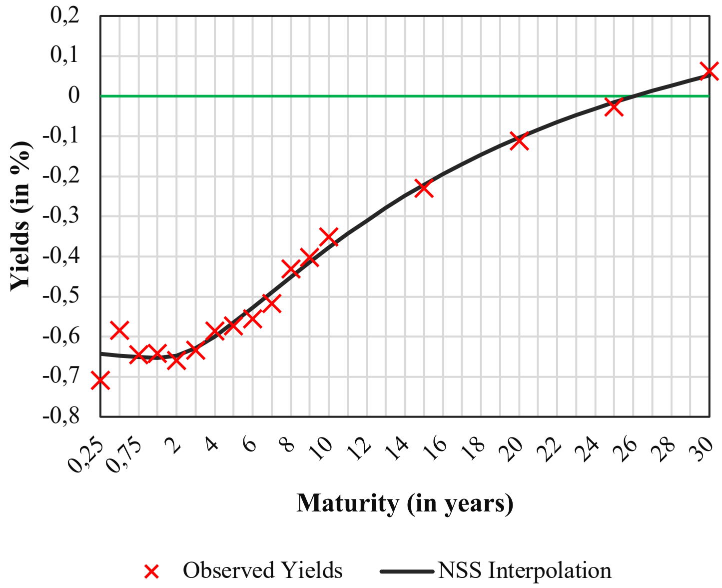

***Figure 5** Spot Rates Estimated by The NSS Method*

7.2.2 Calibration Instruments

In a standard LMM model calibration a set of swaptions and caps should be chosen with observable prices on the market, which would serve as calibration target prices. To define precisely which swaptions and caps should be included in the calibration, we refer the reader to [8], where the authors compare two opposing approaches: *global* and *local*.

In short, the global approach prescribes calibrating the model to a large set of options, including swaptions and caps, which results a large set of consistently priced instruments. Considering the "grid-based" calibration presented in Section 7.1.1, it is recommended to choose ATM swaptions (and also caps) with swaption maturities and swap tenors that coincide with points in our grid. Choosing this method is reasonable for example when the LMM model is intended to be used on exotic securities.

In the local approach, also called the *parsimonious,* one must carefully choose a small number of swaptions and caps and focus on "specification of smooth and realistic term structures of forward rate volatilities" [8]. Hence, it usually involves strong time-homogeneity assumptions on the time-to-maturity component $\lambda_n(t)$. Both have some disadvantages: while in the global approach, the calibration can result forward rate volatilities that are extremely non-stationary, in the local approach the mispricing of certain options is inevitable, which is obviously problematic in cases, when a model is used for pricing complex instruments. Another unanswered question arises, whether we should include both swaption and cap markets or only either of them. In general, choosing the global approach and including both markets (with weights according to their relative importance) is the most reasonable decision, since it is more generally applicable and also many interest rate securities depend on LIBOR rates and swap rates simultaneously. However, as stated in [8], the followers of the local approach argue, that the model should be calibrated only to one of the markets, according to properties of the security to be priced.

Consider now the time-to-maturity component $\lambda_n(t)$ from Section 7.1.2 expressed in terms of factor loadings given in the form of a Rebonato function. As we can see, a time-homogeneity assumption is made in the LMMPlus model, corresponding with the theory of the parsimonious approach mentioned above and the idea proposed by Piterbarg & Andersen [8], that "the evolution of the volatility structure should be as close to being time-homogeneous as possible".

The LMMPlus model follows the parsimonious approach and is calibrated purely to swaption prices or swaption implied volatilities. Obviously, its calibration is not grid-based, since the time-to-maturity component is described by factor loadings. Therefore, the set of target swaptions is chosen in a different way. According to its methodology, we can choose a volatility surface built for multiple swaption maturities (times to expiry dates) and strikes with a specific swap tenor (for example 10 years), and also an ATM surface, which is the mapping of ATM swaption implied volatilities across multiple maturities and swap tenors. By that, we get two slices of the *swaption implied volatility cube* (an object showing how swaption prices vary along dimensions of the swaption maturity, the swaption strike and the swap tenor), that have a common intersection at ten year tenor. According to the description of the LMMPlus model, it is recommended to fit to at least a 10×10 ATM surface to maintain the stability of the calibration. We return to this topic in Section 7.3, where the concrete calibration algorithm of the LMMPlus model is presented.

7.2.3 Displacement Factor δ

Recall the definition of the local volatility function introduced in Section 5.6 as

$$\varphi\big(L_n(t)\big) = L_n(t) + \delta.$$

We mentioned that a forward rate displacement parameter δ is used to shift down the evolution of forward rates, so that negative forward rates do not cause inconveniencies in the model. In fact, $-\delta$ becomes the lower bound of forward LIBOR rates and influences the distribution of them. It also affects the quality of the model, since forward rates below $-\delta$ cannot be modelled and similarly, for strikes below $-\delta$ a close fit cannot be achieved. Therefore, the forward displacement parameter δ must be chosen wisely. A possible way is to include it in the optimization process when calibrating the model, however in the methodology of the LMMPlus model [24] it is recommended to consider it as an input parameter.

Note, that in order to include a wide range of swaption maturities, strikes and swap terms with many negative strikes, a relatively high displacement value should be considered. That also generates a less asymmetrical distribution with less/no exploding rates. In the LMMPlus model, the displacement factor is set to $\delta = 45\%$. Empirical tests validating the choice of δ – that is out of the scope of this thesis – are covered in the internal document of Zurich Insurance Company [24].

7.3 Calibration Algorithm

Consider a given time structure and suppose, that we have managed to find an optimal dimension of the Wiener process W^B. Now, we continue with the framework introduced in Section 7.1.1 and discuss the calibration algorithm proposed by Piterbarg & Andersen [8]. As we said earlier, this method supposes that parameters of the stochastic volatility dynamics and the local volatility function are specified by the user, i.e. given exogenously.

Once the calibration instruments are chosen, the discrepancies between the model and the quoted market prices must be minimized. For that, in general we assume a chosen set of N_S swaptions $V_{swaption,1}, V_{swaption,2}, \dots, V_{swaption,N_S}$ and N_C caps $V_{cap,1}, V_{cap,2}, \dots, V_{cap,N_C}$ and introduce a *calibration objective function* $\mathcal{I}$ (see [8]) given as

$$\mathcal{J}(G) = \frac{\omega_S}{N_S}\sum_{i=1}^{N_S}\left(\bar{V}_{swaption,i}(G) - \hat{V}_{swaption,i}\right)^2 + \frac{\omega_C}{N_C}\sum_{i=1}^{N_C}\left(\bar{V}_{cap,i}(G) - \hat{V}_{cap,i}\right)^2$$
$$+ \frac{\omega_{\partial t}}{N_x N_t}\sum_{i=1}^{N_t}\sum_{j=1}^{N_x}\left(\frac{\partial G_{i,j}}{\partial t_i}\right)^2 + \frac{\omega_{\partial x}}{N_x N_t}\sum_{i=1}^{N_t}\sum_{j=1}^{N_x}\left(\frac{\partial G_{i,j}}{\partial x_j}\right)^2 + \frac{\omega_{\partial t^2}}{N_x N_t}\sum_{i=1}^{N_t}\sum_{j=1}^{N_x}\left(\frac{\partial^2 G_{i,j}}{\partial t_i^2}\right)^2$$
$$+ \frac{\omega_{\partial x^2}}{N_x N_t}\sum_{i=1}^{N_t}\sum_{j=1}^{N_x}\left(\frac{\partial^2 G_{i,j}}{\partial x_j^2}\right)^2,$$

where $\omega_S, \omega_C, \omega_{\partial t}, \omega_{\partial x}, \omega_{\partial t^2}, \omega_{\partial x^2} \in \mathbb{R}_+$ are exogenously specified weights, $\hat{V}$ denotes the quoted market price of swaptions/caps and $\bar{V}(G)$ denotes their generated prices as functions of G introduced in Section 7.1.1. Obviously, first two terms measure the accuracy of fitting to target prices. The other terms can be interpreted this way:

- the third term – the mean-squared average of the derivatives of G – penalizes the volatility functions that are too variable through time
- the fifth term – the mean-squared average of the second derivatives of G with respect to time – controls the smoothness of volatilities (through time) and penalizes discrepancies from a linear evolution
- the remaining two terms are similar to third and fifth, however, they measure constancy and smoothness in the "time-to-maturity direction".

Since swaption prices can vary too much over a long period of time, unitary weighting scheme can make extremely high-priced swaptions overvalued in the objective function. On the contrary, implied volatilities are more flat, therefore in practice a commonly used approach is to apply the error function to them. Market prices $\hat{V}_{swaption,i}$, $\hat{V}_{caps,i}$ can be converted into a constant implied volatilities $\widehat{IV}_{S_i}, \widehat{IV}_{C_i}$. Then, if we denote $\overline{IV}_{S_i}(G), \overline{IV}_{C_i}(G)$ as corresponding model implied volatilities, we can write a new objective function as

$$\mathcal{J}(G) = \frac{\omega_S}{N_S}\sum_{i=1}^{N_S}\left(\overline{IV}_{S_i}(G) - \widehat{IV}_{S_i}\right)^2 + \frac{\omega_C}{N_C}\sum_{i=1}^{N_C}\left(\overline{IV}_{C_i}(G) - \widehat{IV}_{C_i}\right)^2 + \cdots.$$

A possible way – suggested by Piterbarg & Andersen – to refine the calibration is to consider different weights for each swaption or cap instead of common weights ω_S, ω_C.

By introducing the objective function we can finally move to the calibration algorithm proposed by Piterbarg & Andersen [8]. Assuming a given time structure, a selected time and tenor grid $\{t_i\} \times \{x_j\}$, a chosen number m of factors (Wiener process dimension), a correlation matrix R and a set of calibration swaptions and caps (together with a set of weights in the calibration objective function), we can proceed by following steps:

1. From a guessed matrix G construct the full grid $\|\lambda_{n,k}\|$ for all LIBOR indices $k = 1, \dots, N-1$ and expiry indices $n = 1, \dots, k$ (see Section 7.1.1)
2. Compute the matrix $H(T_n)$ for each $n = 1, \dots, N-1$ and volatility loadings $\lambda_k(T_n)$ from $\|\lambda_{n,k}\|$ (see Section 7.1.1)
3. Given $\lambda_k(t)$ for all $k = 1, \dots, N-1$ compute model swaption and cap prices.
4. Find the value of the objective function $\mathcal{I}(G)$.
5. Update G and repeat previous steps until the function $\mathcal{I}(G)$ is minimized.

For some numerical optimizers that can be used for calibration purposes, we refer the reader to the book of Piterbarg & Andersen, where a couple of useful algorithms and methods (available also in some numerical packages) are listed.

As we mentioned earlier, following the parsimonious approach of choosing calibration instruments, the LMMPlus model is calibrated to swaptions, while caps are excluded. Therefore, in this thesis we did not discuss cap pricing. Due to the exclusion of target caps and the non-grid-based construction of the time-to-maturity component $\lambda_n(t)$, the calibration objective function used in the LMMPlus model is "cleaner" relative to the one proposed by Piterbarg & Andersen. Meaning that it is defined only by the first term, the mean-squared error of implied volatilities. To end this Chapter, in the following we introduce the objective function and the calibration methodology used in the LMMPlus model.

7.3.1 Calibration of the LMMPlus Model

The LMMPlus model, introduced in Section 5.6 with time-to-maturity constructed through factor loadings as shown in Section 7.1.2, contains only 8 parameters that need to be found:

- A set of parameters $a, b, c, d \in \mathbb{R}_+$ in the Rebonato function
$$g(x) = (a + bx)e^{-cx} + d.$$
- Positive constants κ, θ and ϵ of the stochastic volatility process (SVP), given by
$$dz(t) = \kappa\big(\theta - z(t)\big)dt + \epsilon\sqrt{z(t)}dZ^B(t).$$

- The constant correlation ρ between the Wiener processes of the SVP and the forward LIBOR rate dynamics, satisfying

$$dZ^B(t) = \frac{\rho}{\sqrt{2}} \sum_{i=1}^{2} dW^{B,i}(t) + \sqrt{1-\rho^2} d\widehat{W}(t).$$

Clearly, the first set of parameters controls the time-to-maturity component $\lambda_n(t)$, whereas the remaining four affect the SVP, the correlation of its shock component with the forward rate shock and allow us fitting to the target skew surface. Recall, that we have chosen two slices of the implied volatility cube (see Section 7.2.2), the ATM surface and a volatility surface through maturities and strikes with a 10-year swap tenor. While the away-from-the-money surface serves to find the SVP parameters, the ATM surface drives the parameters of the Rebonato function.

Similarly to the calibration method introduced earlier, our aim is to optimize the calibration objective function that measures the accuracy of fitting model outputs to target implied volatilities. The algorithm repeatedly chooses optimization parameters, combines them with input parameters and produces a new set of modeled swaption prices until they sufficiently resemble market swaption prices. As we mentioned earlier, the objective function in this case is simpler compared to the general methodology discussed earlier as we do not have to include penalizations for the variability and regularity of the volatility structure.

If we assume a given set of N_S swaptions with market implied volatilities $\widehat{IV}_1, \widehat{IV}_i, \dots, \widehat{IV}_{N_S}$ and denote the set of optimization parameters as

$$\psi = (a, b, c, d, \kappa, \theta, \epsilon, \rho),$$

then the calibration objective function in the LMMPlus model is given by

$$\phi(\psi) = \sum_{\substack{i=1 \\ \omega_i \neq 0}}^{N_S} \frac{\omega_i}{W} \left(\widehat{IV}_i - \overline{IV}_i(\psi) \right)^2,$$

where ω_i are the user specified weights satisfying

$$W = \sum_{i=1}^{N_S} \omega_i,$$

and $\overline{IV}_i(\psi)$ are the model implied volatilities. Note, all parameters are bounded: first seven parameters must be positive, whereas the correlation parameter $\rho \in [-1, 1]$.

Conclusion

In this thesis we examined the LIBOR Market Models modeling the evolution of forward LIBOR rates, which are directly observable on markets. Our aim was to provide a comprehensive but comprehensible guide, which makes understandable the broad issue of building up an advanced interest rate model and deriving a pricing formula using measure changes. We focused on deriving the forward LIBOR rate dynamics in several probability measures and introducing some advanced LMM models together with their calibration methods to market asset prices. We described two models in details: one proposed by Piterbarg & Andersen and the LMMPlus model, used in practice by the Zurich Insurance Company.

First, in Chapter 1 we discussed the necessary assumptions of an arbitrage-free economy and presented the risk-neutral probability measure. In Chapter 2, we introduced the definition of the zero-coupon bond together with some spot interest rates and described the Forward Rate Agreement to define forward rates. Then in Chapter 3, we briefly discussed the Change of Numeraire Theorem and used it to define forward measures induced by zero-coupon bonds as numeraires. We also showed the martingale property of forward rates under these measures. These Chapters were mainly based on the book *Interest Rate Models - Theory and Practice with Smile, Inflation and Credit* [1] written by Brigo & Mercurio. Chapter 4 was dedicated to swaps and swaptions. We presented the vanilla interest rate swap and derived its price using its payoff. Then we described the swaption as an option contract allowing the owner to enter a swap contract. Again, we derived its price from its payoff, however, a pricing formula using the approach proposed by Black [2] was presented, too.

In Chapter 5, we introduced the discrete-time equivalents of numeraires presented in Chapter 2 and derived the evolution of forward LIBOR rates under measures induced by them using the Change of Numeraire Theorem and some other techniques of stochastic calculus. Then we discussed some possible formulations of the forward LIBOR rate evolution: shortly mentioned the basic lognormal formulation proposed by Brace, Gatarek and Musiela [3] known as the BGM model and introduced displaced linear local volatility functions. We presented two single-factor displacements proposed by Piterbarg & Andersen and assumed by the LMMPlus model, separately. As we said, the assumption of displaced rates allowed working with negative interest rates and offered the control over implied volatility skews. We extended both models by introducing a stochastic volatility process used to a more realistic modeling of volatilities and allowing volatility smiles. We also examined the issue of obtaining an empirical correlation matrix and performing principal component analysis to reduce the number of factors of uncertainty. Then, due to differences between the two models considered in this thesis, at the end of this Chapter we presented their concrete frameworks separately.

In Chapter 6, we derived the pricing formula for swaptions following the approach proposed by Wu & Zhang [20] and originally used to pricing caplets. Since there did not exist an exact analytical solution for swaption prices under the assumptions of LIBOR Market Models, we derived an approximate formula expressed in terms of moment generating functions.

In Chapter 7, we provided the calibration methodology for both the Piterbarg & Andersen and the LMMPlus model. We discussed model inputs and presented two ways of constructing the deterministic time-to-maturity component of volatility: a grid-based method proposed by Piterbarg & Andersen and a more heuristic, economically interpretable method assumed by the LMMPlus model. At the end of this thesis we also introduced objective calibration functions considered by both models, whose optimization led us to complete models calibrated to market prices of target assets.

This thesis is a theory-oriented guide to understand the issue of LIBOR Market Models and build up a framework that can relatively accurately price swaptions. We provided the methodology of two models, which in fact differ in many assumptions. The one proposed by Piterbarg & Andersen considers a forward rate evolution independent from the stochastic volatility process, while in the LMMPlus model the correlation between them is given by a constant, which needs to be found during the calibration process. Also, the grid-based calibration method does not involve the optimization of parameters of the stochastic volatility process, but instead considers them as input parameters given exogenously. On the other hand, as we said the LMMPlus follows a heuristic and more easily applicable calibration method based on expert judgments, while sacrificing modeling accuracy to ease the economical comprehension and numerical solutions.

Therefore, a possible continuation of this work would be a detailed comparison of their performance. Due to differences described above, implied volatility skews and smiles are controlled by different parameters. Hence, an examination of the effect of parameters on implied volatilities in both cases would be reasonable. Also, overcoming the technical difficulties of converting the swaption pricing formula to a form that can be solved numerically, is not trivial. For example, improper integrals included in the pricing formula must be adjusted, for instance by mapping the infinite interval to a finite one or using the approximation mentioned in *Libor Market Model with Stochastic Volatility* [14] written by Seonmi Lee.

In Chapter 7, we discussed some possible interpolation methods used for constructing a continuous initial yield curve. However, as we saw, the presented Nelson-Siegel-Svensson reproduced the curve poorly, especially for the short end. We therefore recommend a detailed examination of choosing the proper interpolation method.

Finally, the most obvious continuation of our thesis would be a computer-based practical implementation of the theoretical methodology. It would be an interesting exercise to apply both algorithms on concrete, real-life data, so that one would be able to compare their accuracy and runtime. Eventually, calibrated models could be used for simulations to obtain Monte Carlo estimates of swaption prices. Finally, for a fully functioning model a validation process should be also prepared, in order to ensure that the market/target swaption prices agree with the Monte Carlo estimates.

Bibliography

[1] D. Brigo and F. Mercurio, Interest Rate Models - Theory and Practice with Smile, Inflation and Credit, 2nd ed., Springer, 2007.

[2] F. Black, "The pricing of commodity contracts," *Journal of Financial Economics,* vol. 3, no. 1-2, pp. 167-179, 1976.

[3] A. Brace, D. Gatarek and M. Musiela, "The Market Model of Interest Rate Dynamics," *Mathematical Finance,* vol. 7, no. 2, pp. 127-155, 1997.

[4] M. Vojtek, "Stochastické modely úrokovej miery a ich kalibrácia na trhové dáta," FMFI UK, Bratislava, 2001.

[5] J. Guliš, "LIBOR model oceňovania časovej štruktúry úrokových mier a jeho kalibrácia," Bratislava, 2013.

[6] N. Šelić, "The LIBOR Market Model," University of the Witwatersrand, Johannesburg, 2006.

[7] V. Piterbarg and L. Andersen, Interest Rate Modeling: Volume I: Foundations and Vanilla Models, New York: Atlantic Financial Press, 2010.

[8] V. Piterbarg and L. Andersen, Interest Rate Modeling: Volume II: Term Structure Models, New York: Atlantic Financial Press, 2010.

[9] J. M. Harrison and D. M. Kreps, "Martingales and Arbitrage in Multiperiod Securities Markets," *Journal of Economic Theory,* vol. 20, no. 3, pp. 381-408, 1979.

[10] J. M. Harrison and S. R. Pliska, "Martingales and stochastic integrals in the theory of continuous trading," *Stochastic Processes and their Applications,* vol. 11, no. 3, pp. 215-260, 1981.

[11] J. M. Harrison and S. R. Pliska, *A stochastic calculus model of continuous trading: Complete markets,* vol. 15, no. 3, pp. 313-316, 1983.

[12] H. Geman, N. El Karoui and J.-C. Rochet, "Changes of Numéraire, Changes of Probability Measure and Option Pricing," *Journal of Applied Probability,* vol. 32, no. 2, pp. 443-458, Jun. 1995.

[13] J. C. Hull, Options, Futures and Other Derivatives, Upper Saddle River: Pearson, 2015.

[14] S. Lee, "Libor Market Model with Stochastic Volatility," University of Amsterdam, Amsterdam, 2012.

[15] F. Black and M. Scholes, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy,* vol. 81, no. 3, pp. 637-654, 1973.

[16] R. C. Merton, "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *Journal of Finance,* vol. 29, no. 2, pp. 449-470, 1974.

[17] N. Ikeda and S. Watanabe, Stochastic Diferrential Equations and Diffusion Processes, Tokyo: North Holland/Kodansha, 1992.

[18] P. Protter, Stochastic Integration and Differential Equations, Berlin: Springer-Verlag Berlin Heidelberg, 2004.

[19] Internal Document of Zurich Insurance Company: *Module IX – Life Valuation Models*, 2019.

[20] L. Wu and F. Zhang, "Fast Swaption Pricing Under a Market Model With Stochastic Volatility," Department of Mathematics, University of Science and Technology Clear Water Bay, Hong Kong, 2005.

[21] D. Duffie, J. Pan and K. Singleton, "Transform Analyisis And Asset Pricing For Affine Jump-Diffusions," *Econometrica,* vol. 68, no. 6, pp. 1343-136, 2000.

[22] B. Øksendal, Stochastic Differential Equations, Fifth, Ed., New York: Springer-Verlag Berlin Heidelberg, 2003.

[23] S. Heston, "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options," *The Review of Financial Studies,* vol. 6, no. 2, pp. 327-343, 1993.

[24] Internal Document of Zurich Insurance Company: *LMM+: Derivation, Implementation and MC Calibration Methodology*, 2017.

[25] C. Turfus, "Analytic Option Prices for the Black-Karasinski Short Rate Model," 2019. [Online]. Available: https://www.researchgate.net/publication/313514008_Analytic_Option_Prices_for_the_Black-Karasinski_Short_Rate_Model. [Accessed 18 4 2020].

[26] P. Hagan and G. West, "Interpolation For Curve Construction," *Applied Mathematical Finance,* vol. 13, no. 2, pp. 89-129, 2006.

[27] O. Vašíček, "An equilibrium characterization of the term structure," *Journal of Financial Economics,* vol. 5, no. 2, pp. 177-188, 1977.

[28] M.-M. Magurean, "Extrapolation Methods of the Term Structure of Interest Rates under Solvency II," Sapienza Università di Roma, Rome, 2016.

[29] H. Hladíková and J. Radová, "Term Structure Modelling by Using Nelson-Siegel Model," *European Financial and Accounting Journal,* vol. 7, no. 2, pp. 36-55, 2012.

[30] J. Hunter, "Measure Theory," 2011. [Online]. Available: https://www.math.ucdavis.edu/~hunter/m206/measure_notes.pdf. [Accessed 29 4 2020].

[31] P. Chen, "Modellin the Stochastic Correlation," KTH Royal Institute of Technology, Stockholm, 2016.

Appendix A – Measure Theory

In this Appendix, we introduce some properties of probability measures we refer to throughout this thesis. We provide definitions of equivalent measures and the Radon-Nikodym derivative with the fundamental background from measure theory.

Consider a measurable space $(X, \mathcal{A})$, consisting of a set X and a σ-algebra $\mathcal{A}$ and the extended set of real numbers $\overline{\mathbb{R}} = \{-\infty\} \cup \mathbb{R} \cup \{\infty\}$. If a property holds for all $x \in X \backslash N$, where N is a set of measure zero we say it holds almost everywhere, a.e. for short.

A signed measure ν on X is a function $\nu : \mathcal{A} \longrightarrow \overline{\mathbb{R}}$ such that:

- $\nu(\emptyset) = 0$;
- ν attains at most on of the values $\infty, -\infty$;
- if $\{A_i \in \mathcal{A} : i \in \mathbb{N}\}$ is a disjoint collection of measurable sets, then

$$\nu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \nu(A_i).$$

Singular Measures [30]

Measures μ and ν on space $(X, \mathcal{A})$ are singular, written

$$\mu \perp \nu,$$

if there exist sets M, $N \in \mathcal{A}$ such that $M \cap N = \emptyset$, $M \cup N = X$ and $\mu(M) = 0$, $\nu(N) = 0$.

Absolute Continuity and Equivalence of Measures [30]

Consider the signed measure ν and a measure μ on $(X, \mathcal{A})$. Then ν is absolutely continuous with respect to μ, written

$$\nu \ll \mu,$$

if $\nu(A) = 0$ for every $A \in \mathcal{A}$ such that $\mu(A) = 0$.

If the measures are mutually continuous, they are called equivalent.

Lebesgue-Radon-Nikodym Theorem [30]

Let ν be a σ-finite signed measure and μ a σ-finite measure on space $(X, \mathcal{A})$. Then there exist unique σ-finite signed measures ν_a, ν_s such that

$$\nu = \nu_a + \nu_s, \qquad \text{where } \nu_a \ll \mu \text{ and } \nu_s \perp \mu.$$

Moreover, there exists a measurable function $f : X \longrightarrow \overline{\mathbb{R}}$, uniquely defined up to μ-a.e. equivalence, such that

$$\nu_a(A) = \int_A f d\mu$$

for every $A \in \mathcal{A}$, where the integral is well-defined as an extended real number. Function

$$f = \frac{d\nu}{d\mu}$$

is called the Radon-Nikodym derivative of ν with respect to μ.

Appendix B – Stochastic Calculus

In this Appendix, we provide the framework, i.e. definitions, lemmas and theorems used in our thesis. We leave out the absolute basics, like the definition of a stochastic process or the Wiener process. For a complete overview of the stochastic calculus theory we refer the reader, for example, to the book *Stochastic Differential Equations* [22] written by Bernt Øksendal.

Martingales [22]

A filtration on $(\Omega, \mathcal{F}, \mathbb{P})$ is a family $\mathcal{M} = \{\mathcal{M}_t\}_{t\geq 0}$ of σ-algebras $\mathcal{M}_t \subset \mathcal{F}$ such that

$$0 \leq s < t \implies \mathcal{M}_s \subset \mathcal{M}_t.$$

A stochastic process $\{M_t\}_{t\geq 0}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ is called a martingale with respect to a filtration $\{\mathcal{M}_t\}_{t\geq 0}$ if

1. M_t is $\mathcal{M}_t$-measurable for all t,
2. $\mathbb{E}[|M_t|] < \infty$ for all t,
3. $\mathbb{E}[M_s|M_t] = M_t$ for all $s \geq t$.

Itô Process and the Itô's Lemma [22]

Let W_t be an m-dimensional Wiener process under a non-specified probability measure. An Itô process is an n-dimensional stochastic process $X_t = (X_t^1, X_t^2, \dots, X_t^n)'$ on $(\Omega, \mathcal{F}, \mathbb{P})$ given by

$$dX_t = u(t,\omega)dt + v(t,\omega)dW_t,$$

where for process $u : \mathbb{R} \times \Omega \longrightarrow \mathbb{R}^n$ satisfies

$$\int_0^t |u(s,\omega)|ds < \infty, \qquad 0 \leq t \leq T,$$

and – by defining $|v(s,\omega)|^2 = tr\{v(s,\omega)v(s,\omega)'\}$ – process $v : \mathbb{R} \times \Omega \longrightarrow \mathbb{R}^{n\times m}$ satisfies

$$\int_0^t |v(s,\omega)|^2 ds < \infty, \qquad 0 \leq t \leq T.$$

Let $g(t,x)$ be a C^2 map $[0,\infty) \times \mathbb{R}^n \longrightarrow \mathbb{R}^p$. Then the process $Y(t,\omega) = g(t,X_t)$ is also an *Itô process whose component number k, Y^k is given by the formula*

$$dY^k = \frac{\partial g_k}{\partial t}(t,X)dt + \sum_i \frac{\partial g_k}{\partial x^i}(t,X)dX^i + \frac{1}{2}\sum_{i,j} \frac{\partial^2 g_k}{\partial x^i \partial x^j}(t,X)dX^i dX^j.$$

Proof. We refer the reader to [22].

■

Multiplication Rule of the Itô's Lemma

If X_t and Y_t are two stochastic processes defined as

$$dX_t = u_X(t,\omega)dt + v_X(t,\omega)dW_t,$$
$$dY_t = u_Y(t,\omega)dt + v_Y(t,\omega)dB_t,$$

then

$$d(X_t Y_t) = X_t dY_t + Y_t dX_t + dX_t dY_t.$$

Proof. The formula can be obtained by a direct application of the Itô's Lemma on the stochastic process $X_t Y_t$.

■

Quadratic Covariation of Stochastic Processes [31]

Consider a partition Π of the interval $\langle 0, t\rangle$ as $\Pi : 0 = t_0 < t_1 < \cdots < t_n = t$. If X is a stochastic process, its quadratic variation, i.e. a pathwise measurement of its variation is defined as

$$\langle X\rangle_t = \lim_{\|\Pi\|\to 0} \sum_{i=1}^{n-1} \left(X_{t_{i+1}} - X_{t_i}\right)^2,$$

where $\|\Pi\| = \max_{1\leq i\leq n-1}\{t_{i+1} - t_i\}$. Hence, if we consider another stochastic process Y, the quadratic covariation process between X and Y is defined as

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

If X is an Itô process defined as

$$dX_t = u_X(t,\omega)dt + v_X(t,\omega)dW_t,$$

their quadratic variation can be computed as

$$\langle X\rangle_t = \langle \int_0^{\cdot} v_X(s,\omega)dW_s, \int_0^{\cdot} v_X(s,\omega)dW_s\,\rangle_t = \int_0^t v_X^2(s,\omega)d\langle W,W\rangle_s = \int_0^t v_X^2(s,\omega)ds,$$

where we used that $d\langle W, W\rangle_t = dt$ (see [31]).

Again, if Y is another Itô process – driven by another Wiener process – defined as

$$dY_t = u_Y(t,\omega)dt + v_Y(t,\omega)dB_t$$

then the quadratic covariation between X and Y is given as

$$\langle X,Y\rangle_t = \langle \int_0^{\cdot} v_X(s,\omega)dW_s \,, \int_0^{\cdot} v_Y(s,\omega)dB_s \,\rangle_t = \int_0^t v_X(s,\omega)v_Y(s,\omega)d\langle W,B\rangle_s,$$

or in differential form

$$d\langle X,Y\rangle_t = v_X(t,\omega)v_Y(t,\omega)d\langle W,B\rangle_t.$$

Also, we can write

$$dX_t dY_t = d\langle X,Y\rangle_t = v_X(t,\omega)v_Y(t,\omega)d\langle W,B\rangle_t.$$

Some Classes of Functions and the Itô Integral [22]

Let $\mathcal{V} = \mathcal{V}(S,T)$ be the class of functions

$$f(t,\omega): [0,\infty) \times \Omega \longrightarrow \mathbb{R}$$

such that

1) $(t,\omega) \longrightarrow f(t,\omega)$ *is $\mathcal{B} \times \mathcal{F}$-measurable, where $\mathcal{B}$ defines a Borel σ-algebra on* $[0,\infty)$.
2) $f(t,\omega)$ *is $\mathcal{F}_t$-adapted.*
3) $\mathbb{E}\left[\int_S^T f^2(t,\omega)dt\right] < \infty.$

Hence, we can define the Itô integral of a function $f \in \mathcal{V}$ as

$$\int_S^T f(t,\omega)dW_t,$$

where W_t is a 1-dimensional Wiener process. For more details regarding Itô integrals, we refer the reader to [22].

The previous idea can be extended into a larger class of integrands f. The second condition can be rewritten as

2') *There exists an increasing family of σ-algebras $\mathcal{H}_t$; $t \geq 0$ such that*
   i) *W_t is a martingale with respect to $\mathcal{H}_t$,*
   ii) *f_t is $\mathcal{H}_t$-adapted.*

Let $W_t = (W_t^1, \dots, W_t^n)$ be an n-dimensional Wiener process and $\mathcal{F}_t^{(n)}$ be a σ-algebra generated by $W(t)$. The set of matrices $v = \left[v(t,\omega)_{ij}\right] \in \mathbb{R}^{m\times n}$ where each v_{ij} satisfies conditions 1), 2') and 3) w.r.t. a filtration $\mathcal{H} = \{\mathcal{H}_t\}_{t\geq 0}$ we denote by $\mathcal{V}_{\mathcal{H}}^{m\times n}(S,T)$.
For $v \in \mathcal{V}_{\mathcal{H}}^{m\times n}(S,T)$ we can define

$$\int_S^T v dW_t,$$

which is an $m \times 1$ matrix, whose i-th component is given by the following sum of 1-dimensional Itô integrals

$$\sum_{j=1}^{n} \int_S^T v_{ij}(t,\omega) dW_t^j.$$

We also introduce another class of functions. First, we extend the previous condition 3) by weakening it to

3') $P\left(\mathbb{E}\left[\int_S^T f^2(t,\omega)dt\right] < \infty\right) = 1.$

Hence, by $\mathcal{W}_{\mathcal{H}}(S,T)$ we denote the class of functions $f(t,\omega)$ satisfying 1), 2') and 3'). Similarly to the previous class, in matrix case we use $\mathcal{W}_{\mathcal{H}}^{m\times n}(S,T)$.

Martingale Representation Theorem [22]

Let $W_t = (W_t^1, \dots, W_t^n)$ be an n-dimensional Wiener process. Suppose M_t is an $\mathcal{F}_t^{(n)}$-martingale under the measure $\mathbb{P}$ and that $M_t \in L^2(\mathbb{P})$ for all $t \geq 0$. Then there exists a unique stochastic process $g(s,\omega)$ such that $g \in \mathcal{V}^n(0,t)$ for all $t \geq 0$ and

$$M_t(\omega) = \mathbb{E}[M_0] + \int_0^t g(s,\omega) dW_s, \qquad \forall t \geq 0.$$

Proof. We refer the reader to [22].

■

Girsanov's Theorem I [22]

Let $Y(t) \in \mathbb{R}^n$ be an Itô process of the form

$$d\,Y(t) = a(t,\omega)dt + dW_t, \qquad t \leq T, \qquad Y(0) = 0,$$

where $T \leq \infty$ is a given constant and W_t is an n-dimensional Wiener process. Put

$$M_t = exp\left(-\int_0^t a(s,\omega)dW_s - \frac{1}{2}\int_0^t a^2(s,\omega)\ ds\right), \qquad t \leq T.$$

Assume that $a(s,\omega)$ satisfies the Novikov's condition

$$\mathbb{E}\left[exp\left(\frac{1}{2}\int_0^t a^2(s,\omega)ds\right)\right] < \infty,$$

where $\mathbb{E} = \mathbb{E}^{\mathbb{P}}$ is the expectation under the measure $\mathbb{P}$. Define the measure $\mathbb{Q}$ on $\left(\Omega, \mathcal{F}_T^{(n)}\right)$ by

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

Then $Y(t)$ is an n-dimensional Wiener process under the measure $\mathbb{Q}$, for $t \leq T$.

Proof. We refer the reader to [22]**.**

■

Girsanov's Theorem II [22]

Let $Y(t) \in \mathbb{R}^n$ be an Itô process of the form

$$dY(t) = \beta(t,\omega)dt + \theta(t,\omega)dW_t, \qquad t \leq T,$$

where $W_t \in \mathbb{R}^m$, $\beta(t,\omega) \in \mathbb{R}^n$ and $\theta(t,\omega) \in \mathbb{R}^{n\times m}$. Suppose there exists a process $u(t,\omega) \in \mathcal{W}_{\mathcal{H}}^m$ and $\alpha(t,\omega) \in \mathcal{W}_{\mathcal{H}}^n$ such that

$$\theta(t,\omega)\, u(t,\omega) = \beta(t,\omega) - \alpha(t,\omega)$$

and assume that $u(t,\omega)$ satisfies the Novikov's condition

$$\mathbb{E}\left[exp\left(\frac{1}{2}\int_0^t u^2(s,\omega)ds\right)\right] < \infty.$$

Put

$$M_t = exp\left(-\int_0^t u(s,\omega)dW_s - \frac{1}{2}\int_0^t u^2(s,\omega)\ ds\right), \qquad t \leq T.$$

and

$$d\mathbb{Q} = M_T d\mathbb{P} \qquad\qquad on \left(\Omega, \mathcal{F}_T^{(n)}\right).$$

Then

$$\widehat{W_t} = \int_0^t u(s,\omega)ds + W_t, \qquad t \leq T$$

is a Wiener process under the measure $\mathbb{Q}$ and in terms of $\widehat{W_t}$ the process $Y(t)$ has the stochastic integral representation

$$dY(t) = \alpha(t,\omega)dt + \theta(t,\omega)d\widehat{W_t}.$$

Proof. We refer the reader to [22]**.**

■

Appendix C – Principal Components Analysis (PCA)

In this Appendix we follow the steps of Piterbarg & Andersen presented in *Interest Rate Modeling: Volume I* [7].

Consider a p-dimensional random variable $Z \sim N(0, \Sigma)$, where Σ has full rank, i.e. is positive definite. Let us define the approximation of Z as

$$Z \approx DX,$$

where X is an r-dimensional vector of independent normal random variables, $r \leq p$ and D is a $(p \times r)$-dimensional matrix. By an optimal approximation, we mean L^2 closeness of the covariance matrix $DD^\top$ to Σ. Hence, we define the optimal D^* as the matrix that minimizes

$$f(D) = tr\big((\Sigma - DD^\top)(\Sigma - DD^\top)^\top\big),$$

where $tr(A)$ is the trace of a matrix A.
It can be shown [7], that

$$D^* = E_r\sqrt{\Lambda_r},$$

where Λ_r is an $r \times r$ diagonal matrix of the r largest eigenvalues of Σ, and E_r is a $p \times r$ matrix of p-dimensional eigenvectors, each corresponding to an eigenvalue in Λ_r.
Now, we can write the approximation of Z as

$$Z \approx \tilde{Z} \coloneqq E_r\sqrt{\Lambda_r}X = \sqrt{\lambda_1}e_1X_1 + \sqrt{\lambda_2}e_2X_2 + \cdots + \sqrt{\lambda_r}e_rX_r,$$

where e_i is the i-th column of E_r and λ_i is the i-th among the eigenvalues sorted in decreasing order of magnitude. The vector e_i is called the i-th principal component of Z, and the random variable $\sqrt{\lambda_i}X_i$ the i-th principal factor.

Since $tr\big(Cov(Z,Z)\big) = \sum_{i=1}^{p}\lambda_i$ and $tr\left(Cov\big(\tilde{Z},\tilde{Z}\big)\right) = \sum_{i=1}^{r}\lambda_i$, by fraction

$$\frac{\sum_{i=1}^{r}\lambda_i}{\sum_{i=1}^{p}\lambda_i}$$

we can explain the loss of the total variance.

Appendix D – Deriving the Equation from Section 6.1

We derive the equality

$$\mathbb{E}_t^A[\mathbf{1}_{X>x}] = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \frac{Im\left(e^{-iux}M_X(iu)\right)}{u}du$$

introduced in Section 6.1, using the following properties:

$$\mathrm{Im}(z) = \frac{z-\bar{z}}{2i}$$

$$M_X(z) = \mathbb{E}[e^{zX}]$$

$$\overline{M_X(z) =} M_X(\bar{z})$$

$$\int_0^\infty \frac{sin(x)}{x}dx = \frac{\pi}{2}.$$

The equality follows from

$$\frac{1}{\pi}\int_0^\infty \frac{\mathrm{Im}\left(e^{-iux}M_X(iu)\right)}{u}du = \frac{1}{\pi}\int_0^\infty \frac{e^{-iux}M_X(iu) - e^{iux}M_X(-iu)}{u}du$$

$$= \frac{1}{\pi}\int_0^\infty \frac{e^{-iux}\int_{-\infty}^{+\infty} e^{iuy}f(y)dy - e^{iux}\int_{-\infty}^{+\infty} e^{-iuy}f(y)dy}{2iu}du$$

$$= \frac{1}{\pi}\int_{-\infty}^{+\infty}\int_0^\infty \frac{e^{-iu(x-y)} - e^{iu(x-y)}}{2iu}du\, f(y)dy$$

$$= \frac{1}{\pi}\int_{-\infty}^{+\infty}\int_0^\infty \frac{\sin\left(u(x-y)\right)}{u}du\, f(y)dy = \frac{1}{2}\int_{-\infty}^{+\infty} \mathrm{sign}(x-y)f(y)dy$$

$$= \frac{1}{2}\left(\int_x^{+\infty} f(y)dy - \int_{-\infty}^{x} f(y)dy\right)$$

$$= \frac{1}{2}\left(\int_x^{+\infty} f(y)dy - \left(\int_{-\infty}^{\infty} f(y)dy - \int_x^{+\infty} f(y)dy\right)\right) = \int_x^{+\infty} f(y)dy - \frac{1}{2}$$

$$= \mathbb{E}_t^A[\mathbf{1}_{X>x}] - \frac{1}{2}.$$

Appendix E – Results from Section 6.3

Recall differential equations

$$\frac{dA}{d\tau} = \kappa\theta B,$$

$$\frac{dB}{d\tau} = b_2 B^2 + b_1 B + b_0,$$

where we considered the decomposition of B into a particular and a general solution, written $B = Y_1 + Y_2$. In *Remark 4*, we found the particular solution Y_1 and transformed the second differential equation into an integrable linear differential equation

$$\frac{du}{d\tau} + d_B = -b_2,$$

where $u = Y_2^{-1}$. Since both sides can be integrated, the solution for function u is

$$u = e^{d_b \tau} c - \frac{b_2}{d_B}.$$

Hence,

$$Y_2 = \frac{1}{e^{d_b \tau} c - \frac{b_2}{d_B}},$$

and the constant c can be found by using the relation $Y_2(0) = B(0) - Y_1$. Since the value of B at time 0 is B_0 and for Y_1 we chose the particular solution with the "+" sign, constant c must in fact satisfy $Y_2(0) = B_0 - B_+$. Therefore,

$$Y_2 = \frac{1}{e^{d_b \tau}\left(\frac{2b_2}{2b_2 B_0 + b_1 - d_b} + \frac{b_2}{d_B}\right) - \frac{b_2}{d_B}}.$$

By rearranging the fraction we get the general solution Y_2 expressed as

$$Y_2 = \frac{d_B}{b_2} \frac{h e^{d_b \tau}}{1 - h e^{d_b \tau}},$$

where

$$h = \frac{2b_2 B_0 + b_1 - d_b}{2b_2 B_0 + b_1 + d_b}.$$

We have managed to find both the particular and the general solution, so we can derive $B(\tau)$ as

$$B(\tau) = Y_1 + Y_2 = B_+ + \frac{d_B}{b_2} \frac{h e^{d_b \tau}}{1 - h e^{d_b \tau}} = \frac{-b_1 + d_B}{2b_2} + \frac{d_B}{b_2} \frac{h e^{d_b \tau}}{1 - h e^{d_b \tau}},$$

from which we can get to the final form

$$B(\tau) = B_0 + \frac{(-2b_2B_0 - b_1 + d_b)}{2b_2}\left(\frac{1 - e^{d_b\tau}}{1 - he^{d_b\tau}}\right).$$

The definition of function $A(\tau)$ can be obtained by a simple integration, such that

$$A(\tau) = A_0 + \kappa\theta \int_0^\tau B(s)ds = A_0 + \kappa\theta B_0\tau + \frac{(-2b_2B_0 - b_1 + d_b)}{2b_2} \int_0^\tau \frac{1 - e^{d_bs}}{1 - he^{d_bs}} ds,$$

where

$$\int_0^\tau \frac{1 - e^{d_bs}}{1 - he^{d_bs}} ds = \tau - \int_1^{e^{d_b\tau}} \frac{1 - h}{1 - hx} dx = \tau - \frac{1}{d_B}\left(\frac{h - 1}{h}\right) \ln\left(\frac{1 - he^{d_b\tau}}{1 - h}\right).$$

Therefore, the final form of $A(\tau)$ is given as

$$A(\tau) = A_0 + \frac{\kappa\theta}{2b_2}\left((-b_1 + d_B)\tau - 2\ln\left(\frac{1 - he^{d_b\tau}}{1 - h}\right)\right).$$