HELSINGIN YLIOPISTO — HELSINGFORS UNIVERSITET — UNIVERSITY OF HELSINKI

Tiedekunta/Osasto — Fakultet/Sektion — Faculty	Laitos — Institution — Department
Faculty of Science	Department of Mathematics and Statistics

Tekijä — Författare — Author
Teodor Duevski

Työn nimi — Arbetets titel — Title

Modeling the term structure of zero-coupon bonds

Oppiaine — Läroämne — Subject
Theoretical and Computational Methods

Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
Master's thesis	May 2019	71 p.

Tiivistelmä — Referat — Abstract

In this thesis we model the term structure of zero-coupon bonds. Firstly, in the static setting by norm optimization Hilbert space techniques and starting from a set of benchmark fixed income instruments, we obtain a closed from expression for a smooth discount curve. Moving on to the dynamic setting, we describe the stochastic modeling of the fixed income market. Finally, we introduce the Heath-Jarrow-Morton (HJM) methodology. We derive the evolution of zero-coupon bond prices implied by the HJM methodology and prove the HJM drift condition for non arbitrage pricing in the fixed income market under a dynamic setting.

Knowing the current discount curve is crucial for pricing and hedging fixed income securities as it is a basic input to the HJM valuation methodology. Starting from the non arbitrage prices of a set of benchmark fixed income instruments, we find a smooth discount curve which perfectly reproduces the current market quotes by minimizing a suitably defined norm related to the flatness of the forward curve. The regularity of the discount curve estimated makes it suitable for use as an input in the HJM methodlogy.

This thesis includes a self-contained introduction to the mathematical modeling of the most commonly traded fixed income securities. In addition, we present the mathematical background necessary for modeling the fixed income market in a dynamic setting. Some familiarity with analysis, basic probability theory and functional analysis is assumed.

Avainsanat — Nyckelord — Keywords
discount curve, forward curve, term-structure estimation, zero-coupon bond, the HJM model

Säilytyspaikka — Förvaringsställe — Where deposited
Kumpulan tiedekirjasto

Muita tietoja — Övriga uppgifter — Additional information

University of Helsinki

Department of Mathematics and Statistics

Master's thesis

Modeling the term structure of zero-coupon bonds

Author:
Teodor Duevski

Supervisors:
Prof. Jani Lukkarinen
Dr. Jaakko Lehtomaa

Acknowledgements

Above all, I would like to express my gratitude to my master's thesis supervisor Prof. Jani Lukkarinen for providing me with this opportunity and for his guidance during the thesis writing process. His valuable input not only as a supervisor, but also as a mentor during the past two years helped me tremendously in guiding my research and study interests. I would also like to thank my second supervisor Dr. Jaakko Lehtomaa for his thoughtful suggestions which were of immense help in shaping the structure of this thesis.

Contents

To Helsinki, my family and VSO in that order.

Chapter 1

Introduction

Any investment under which the borrower or the issuer is obliged to make payments of a fixed income at some given predetermined dates in the future is termed a fixed income security. These securities are traded on the so called fixed income markets or more commonly known interest markets. The fixed income market in its broadest sense includes any interest rate sensitive claim. Due to the fact that changing interest rates constitute a major risk for banks, insurance companies and other financial institutions modeling the term structure of interest rates is of fundamental importance [1]. The prices of fixed income securities are usually expressed in terms of a fundamental bond unit called a zero-coupon bond. In financial modeling it is often assumed that at the current date t one can observe an initial term structure of zero-coupon bond prices for a continuum of maturities called the discount curve.

Starting with the seminal work by Vasiček [2], the main goal of the mathematical modeling of the fixed income market has been twofold:

1. To consistently price all default free zero-coupon bonds of varying maturities. That is, to obtain a smooth discount curve for all maturities given a finite number of actively traded fixed income securities.

2. To price all interest rate sensitive contingent claims in an arbitrage free way at a future date given the prices of the zero-coupon bonds.

Problem 1 is fundamentally deterministic in nature. We fix a time t as the current date and given the t time prices of a set of benchmark fixed income instruments with varying maturities, we try to estimate a smooth discount curve that gives the t prices of zero-coupon bonds for any maturity. Since the discount curve is an infinite dimensional object, to address Problem 1 one needs an interpolation method in order to complete the information obtained from the finite number of fixed income instruments. One possible approach taken in the early literature, is to impose a particular parametric functional form for the entire or parts of the discount curve and calibrate the parameters by minimizing the pricing error between the function and the market quotes [3, 4]. These parametric methods are more suitable for studies where the general shape of the discount curve is more important than the exact values. In practise they have been used in monetary policy modeling and recommendations by central banks [4]. The functional parametric forms can however be too restrictive for financial institutions involved in trading, since they do not perfectly reproduce the market quotes of the observed actively traded fixed income instruments.

An alternative approach taken in the literature is to suitably define a norm which is related to the smoothness and goodness of fit of the curve. In order to obtain maximal smoothness and goodness of fit, the norm is then minimized subject to a matching constraint. Often the minimization is done on a transformation of the discount curve, called the forward curve with the assumption that forward rates over short time distances should not vary substantially during economic periods of low volatility [5, 6]. Starting with the work by Lorimier and Delbean, it has been a common practise to directly minimize a norm related to the discount curve starting from a set of zero-coupon bonds as benchmark instruments whose price is to be matched [7]. Zero-coupon bonds are however rarely traded in practise (US Treasury Bills are the most common example), so one often needs to rely on financial data from more actively traded fixed income instruments such as LIBOR rates, interest rate swaps and interest rate futures. This has been the approach taken in a recent work by Filipović, whereby a suitably defined norm in a Hilbert space is minimized given a matching constraint for pricing of actively traded fixed income instruments [8].

Problem 2 is stochastic in nature. Given a solution to Problem 1 i.e. a smooth discount curve at time t we try to impose a model for the evolution of the zero-coupon prices or some transformation thereof. The modeling is done on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathcal{P})$, carrying a d-dimensional Brownian motion $W = (W_1, W_2, \dots, W_d)^T$. Once that model is in place, i.e. we have the prices for zero-coupon bonds at any time in the future over any maturity date, we price any contingent fixed income claim based on the evolution of the zero-coupon bond prices. Starting with Vasiček, the main approach taken in the early literature has been to impose the Itó dynamics of the short rate [2]. The Vasiček model assumes that the short rate is a mean reverting stochastic process with constant drift and volatility parameter under a martingale measure $\mathcal{Q}$ which is equivalent to the objective probability measure $\mathcal{P}$ [2]. This model leads to an explicit closed form solution for the zero-coupon bond prices, however it cannot be fitted to an initial yield and volatility curve [1]. Building on the initial work done by Vasiček there have been numerous extensions proposed for modeling the short rate. The most widely known and used in practice are the Hull-White and the Ho-Lee models [9, 10]. Similarly as the Vasiček model, these models specify the dynamics of the short rate under the equivalent martingale measure $\mathcal{Q}$ and one is often faced with a parameter estimation problem. To calibrate to data, the parameter estimation problem is solved via a so called inversion of the yield curve method, whereby the $\mathcal{Q}$ evolution parameters are chosen in such a way as to best "match" the solution of the term structure equation [11]. An alternative approach currently known as the HJM methodology, developed in the seminal work by Heath, Jarrow and Morton circumvents the problem of inversion of the yield curve by imposing a model for the evolution of the forward curve directly under the objective probability measure $\mathcal{P}$ [12]. This model leads to stochastic spot rates processes with multiple stochastic factors influencing the term structure, thus making all the previously considered short rate models a special case of this general framework [12].

In this thesis, we address the two problems described above. Firstly, by using optimization in Hilbert space methods and building on the work by Filipović we estimate a smooth discount curve given a set of benchmark instruments. Then, we move to the stochastic setting and use this curve as an input in the Heath-Jarrow-Morton methodology [8, 12].

In Chapter 2, we briefly introduce the notion of an interest rate and describe the most commonly traded fixed income securities. Under a deterministic setting, we derive formulas for the arbitrage free prices of the fixed income securities that are used as a benchmark for estimating the discount curve. In particular, we focus on the arbitrage free expressions of forward rates, simple rates, coupon bonds and interest rate swaps. In Chapter 3, we describe and complement the method devised by Filipović in order to estimate a smooth discount curve given the fixed set of benchmark fixed instruments introduced in Chapter 2 [8]. The most important results of this chapter are Theorems 3.6 and 3.7, which give unique closed form solutions for the discount curve under the conditions that the short rate is minimized or exogenously specified respectively.

In Chapter 4, we move to the stochastic setting and collect the most important tools from stochastic analysis used in the modeling of the fixed income market. The main tools used in the subsequent chapters are stated here, namely the Radon-Nikodym theorem, the Girsanov's change of measure theorem, Itó's theorem, the Martingale representation theorem and Fubini's theorem for stochastic integrals. These results are not stated nor proved in their most general form. Instead the regularity assumptions imposed are often relaxed and geared towards the applications of the theorems in subsequent chapters.

In Chapter 5, using the toolbox from Chapter 4, we describe the modeling of the equities and fixed income markets. We highlight the similarities and the differences between the two types of markets and we introduce the short rate modeling approach to the fixed income market by using the Ho-Lee model as a particular example.

In Chapter 6, we describe the Heath-Jarrow-Morton methodology for the dynamics of the forward curve. Using first fundamental theorem of asset pricing introduced in Chapter 5 and the toolkit from Chapter 4 from the HJM dynamics of the forward curve we derive the dynamics of the zero-coupon bond prices and the short rate dynamics. We end the chapter by deriving the famous HJM drift condition, which specifies the dynamics of the forward curve under a martingale equivalent measure $\mathcal{Q}$. We take the HJM drift condition to be the no arbitrage condition for the fixed income market. In Chapter 7 we summarize and present possible directions for further research.

Chapter 2

The economics of interest rates and related contracts

In this chapter, we explore the no arbitrage pricing of some frequently traded fixed income securities. After a brief introduction to time preference theory of interest, we focus our attention on defining the main fixed income contracts that are going to be used throughout the thesis. We price these contracts in an arbitrage free way using zero-coupon bonds. In this and the following chapter we consider a deterministic setting and we fix $t \in \mathbf{R}_+$ as today's date. We end the chapter by defining the notions of duration and convexity as measures for the sensitivity of the bond price to the shift in the yield curve.

2.1 Interest as an economic phenomenon

2.1.1 Time preference theory

In the most general sense interest can be described as a value differential between present goods and future goods of the same quantity and quality. In his magnum opus Capital and Interest, Bohm-Bawerk noticed that ”present goods have in general a greater subjective value than future goods of equal quantity and quality” [13]. Thus, the interest phenomenon in the most general sense is grounded on the simple fact that present consumption is preferred to future consumption.

Definition 2.1 (Time preference)**.** We are going to define time preference as the current relative valuation placed on receiving a good today compared to receiving the same good at a future date.

A positive time preference for a given good indicates that the good is valued more today than at some future time. The longer the time preference for a given good, the higher is the valuation for the good today compared to some future time period. Due to the necessity for present consumption, the fact that human beings tend to underestimate future needs the time preference is in most cases positive i.e. given a choice between present or future goods of the same quantity and quality human beings choose the former. This does not imply that human beings must consume through every action so that their time preference would be the only factor determining their action. Rather, it means that in order to survive human beings must at some point prefer shorter production processes to longer ones - even though the longer ones tend to be more productive. Ludwig von Mises provides a classic example illustrating the above mentioned phenomenon [14]. Consider an individual faced with the following three fishing production alternatives: the first one leads to catching of one fish at the end of one hour production process, the second one leads to catching of ten fish at the end of an eight-hour production process and the third one to catching of hundred fish at the end of a full work week production process. In light of the figures given, if one disregards the need for survival through time it is expected that one would choose the last alternative, which is the most productive. However, the need of survival might prompt the individual to choose any of the three offered alternatives, based on his or her own time preference. For instance, suppose one chooses the ten fish alternative. This means that the need to survival prompted the individual to choose a less productive shorter process over a more productive longer process (the hundred fish alternative); however, the additional benefit of choosing the ten fish process (nine fish in seven hours) was large enough to overcome the individual's time preference for choosing the shortest alternative. The argument presented here is an adaptation of Mises' argument taken from [15].

Thus, in a rudimentary market economy, where simple trade of goods is allowed between individuals, the interest rate is determined by the time preference of each individual comprising the economy. At a given point in time, individuals with a shorter time preference are willing to supply goods (a bundle of which in a modern market economy is given in terms of money prices) i.e. lend to other individuals, whereas individuals with a longer time preference will in general demand consumption goods. It is important to note that the time preference theory does not imply that we can divide individuals in a market economy into lenders and borrowers, instead it merely asserts that the subjective difference between the degree to which one prefers present to future benefits (given all else equal) determines the interest rate in a simple economy where trade in terms of goods between individuals is allowed.

2.1.2 Supply and demand theory of interest

The subjective time preference theory explains the nature of the interest phenomenon. It is a useful conceptual tool aiding in understanding the rate of interest in a simple market economy where individuals are allowed to trade present and future goods. In the modern economy, interest is defined as price of future monetary unit in terms of today's monetary unit and according to standard macroeconomic theory it is determined by the supply and demand for credit. The supply and demand schedules are not in general determined by the subjective time preferences of each individual. Instead, they are mostly influenced by institutional, governmental, bank and central bank policies on the supply side and the funding needs of corporations, governments and individuals on the demand side. In equilibrium the interest rate is given by the intersection of the supply and demand schedules for credit.

2.1.3 The fixed income market

The need for economic growth and the turn to more roundabout and longer methods of production has lead to the development of highly complex and mutually dependent credit markets each of which generates a separate interest rate usually termed as the yield. These credit markets are termed fixed income or interest rate markets. Any investment under which the borrower or the issuer is obliged to make payments of a fixed income at some given future predetermined dates is termed a fixed income security. These securities are traded on the fixed income markets. The simplest type of an interest contract is described by a so called zero-coupon bond, which is a contract that pays its holder an amount of 1 at the end of the lending period. The end of the lending period is usually termed the contract maturity date. The prices of other fixed income securities, usually called fixed income derivatives are often mathematically expressed in terms of a collection of zero-coupon bond prices for different maturities. In this chapter we price some of the most important fixed income derivatives, such as the simple rate, the forward rate agreement and the interest rate swap. We price these contracts under the so called no arbitrage assumption which roughly states that an investor starting from zero initial capital at time t cannot make a guaranteed risk free profit at some at some later date (a precise definition of the no arbitrage principle is provided in Subsection 5.1.3).

2.2 Simple rate and yield

Definition 2.2 (Money interest)**.** Money interest refers to the rent paid by a borrower of money to the lender over a period of time. At time t, the lender pays an amount of

1 to the borrower and at time T the lender receives an amount $1 + R$. R is defined to be the money interest. We call t the lending time and T the maturity date of the loan. Throughout this the thesis we will assume that $R \geq 0$.

Interest is usually expressed in annualized interest rates and the following two definitions will be used throughout this thesis.

Definition 2.3 (Simple rate)**.** The simple rate $L(t, T)$ is defined so that it satisfies the following relation:

$$1 + R = 1 + (T - t)L(t, T). \tag{2.1}$$

Definition 2.4 (Yield)**.** The continuously compounded rate or yield $y(t, T)$ is defined so that it satisfies the following relation:

$$1 + R = e^{(T-t)y(t,T)}. \tag{2.2}$$

Definition 2.5 (Term structure)**.** The term structure of interest rates prevailing at t is a mapping $T \mapsto L(t, T)$ for fixed t. The term structure of yields is called a yield curve and it is a mapping $T \mapsto y(t, T)$. Throughout the thesis we will assume that both the term structure of interests and the yield curve are smooth curves for all $T \geq t$, where t the current time is held fixed.

Definition 2.6 (Time series of term structures)**.** Time series of term structures is a mapping $(t, T) \mapsto L(t, T)$ for each $t \in \mathbf{R}_+$ and $T \geq t$.

Lemma 2.7. *Given a fixed* $t \in \mathbf{R}_+$ *the following relations hold:*

(i) $L(t, T) \geq y(t, T)$ *for all* $T \geq t$.

(ii) $\lim_{T \to t} L(t, T) = \lim_{T \to t} y(t, T)$.

Proof.

(i) Using equations (2.2) and (2.1) we see that:

$$L(t, T) = \frac{e^{(T-t)y(t,T)} - 1}{T - t}.$$

Thus, the inequality follows from the convexity of the exponential function.

(ii) Using the relation derived in *i)* and L'Hôpital's rule we obtain:

$$\begin{aligned}\lim_{T\to t} L(t,T) &= \lim_{T\to t} \frac{e^{(T-t)y(t,T)}-1}{T-t} \\ &= \lim_{T\to t} \left(y(t,T)e^{(T-t)y(t,T)} + \frac{\partial y}{\partial T}(t,T)(T-t)e^{(T-t)y(t,T)} \right) \\ &= \lim_{T\to t} y(t,T).\end{aligned}$$

□

2.3 The Short Rate and the Money Market Account

Definition 2.8 (Short rate)**.** We define

$$\lim_{T\to t} L(t,T) \equiv r(t)$$

and call $r(t)$ the short rate.

Remark 2.9. Using conclusion *ii)* in Lemma 2.7 we notice

$$\lim_{T\to t} y(t,T) = r(t).$$

The short rate is the interest rate an investor earns on a loan over the short period $[t, t+dt]$.

Definition 2.10 (Money market account)**.** Continuously reinvesting at the short rate gives us the money market account which is defined to satisfy the following relation:

$$B(t+dt) = B(t)(1+r(t)dt). \tag{2.3}$$

Rewriting (2.3) in differential notation we have:

$$dB(t) = B(t)r(t)dt. \tag{2.4}$$

Imposing $B(0) = 1$, the solution of (2.4) is given by:

$$B(t) = e^{\int_0^t r(s)ds}. \tag{2.5}$$

As can be seen from (2.3), the money market account is a risk-free asset because its return at $t+dt$ is already known at time t. Hence, the short rate is the risk-free rate of return over an infinitesimal period $[t, t+dt]$.

2.4 Zero-coupon bonds

A bond is a securitized (tradable) form of a loan. The simplest type of a tradable bond contract is the so called zero-coupon bond.

Definition 2.11 (Zero-coupon bond)**.** A zero-coupon bond with maturity T is a contract that pays its holder an amount of 1 at T. We denote its price at time t with maturity T by $P(t,T)$.

Intuitively, a zero-coupon bond specifies the time value at time t of a monetary unit which has a value of 1 at T. It is also called a T-bond.

Definition 2.12 (Discount curve)**.** If we fix t and vary the maturity date T, we get the term structure of the zero-coupon bond prices prevailing at t called the discount curve. Similarly as in the case of the yield curve, we are going to assume that the discount curve is smooth in T.

For simplicity, throughout this chapter we will assume that the following hold:

1. There is no credit risk, which means that $P(T,T) = 1$.
2. There exists a frictionless market for all zero-coupon bonds of all maturities.

In practise Assumption 2 implies that all costs and restraints associated with trading transactions are non-existent and an investor can always purchase, sell, or short sell a zero-coupon bond of any maturity. Assumption 1 implies that the borrower (issuer of the bond) will not default and repay the full price of the zero-coupon bond at the maturity date T.

In reality these assumptions are not always satisfied, since zero-coupon bonds are not traded for all maturities and $P(T,T)$ might be smaller than 1. This can occur if there is some positive probability that the issuer of the bond defaults. Even though not always supported in practise, these two assumptions are a starting point in the mathematical analysis of the fixed income market. They are of vital importance throughout the first two chapters of the thesis where the deterministic fixed income market (t fixed) is analyzed. In particular, Assumption 2 will allow us to express the non arbitrage prices of other more complex fixed income instruments in terms of zero-coupon bond prices in this deterministic setting.

From no arbitrage requirements the price of the zero-coupon bond at time t with maturity date T can be expressed in terms of the simple rate (2.1) and the yield (2.2) via the following relation:

$$P(t,T) = \frac{1}{1+(T-t)L(t,T)} = e^{-(T-t)y(t,T)}. \tag{2.6}$$

Using expression (2.6) and the definition of the short rate

$$\lim_{T\to t} y(t,T) = r(t),$$

we immediately get the following relation between the short rate and the derivative of the discount curve:

$$r(t) = -\frac{\partial P}{\partial T}(t,t). \tag{2.7}$$

2.5 Forward and futures rates

Definition 2.13 (Forward rate agreement)**.** A forward rate agreement (FRA) is a fixed income contract which is specified at time t, by a future period $[T_0, T_1]$, with length denoted by $\delta = T_1 - T_0$, a fixed rate K and a notional N.

The term notional in the definition above refers to a contractually agreed upon time t value of the FRA. The forward rate agreement allows the holder to fix a rate K at time t for a loan of an amount N over a future period $[T_0, T_1]$. This allows the holder to take an loan at time T_0 with a predetermined rate K known at time t and thus hedge against unexpected changes in the unknown interest rate over a period $[T_0, T_1]$.

Lemma 2.14. *The non arbitrage rate K of the FRA is called the simple forward rate. It is denoted by $F(t, T_0, T_1)$ and is given in terms of zero-coupon bond prices as:*

$$F(t,T_0,T_1) = \frac{1}{\delta}\left(\frac{P(t,T_0)}{P(t,T_1)} - 1\right). \tag{2.8}$$

Proof. The payoff for the FRA holder at time T_1 is given by $N\delta(L(T_0,T_1) - K)$. We rewrite the payoff using expression (2.6) as:

$$N\left(\frac{1}{P(T_0,T_1)} - 1 - \delta K\right). \tag{2.9}$$

Now, we calculate the present value of expression (2.9) using zero-coupon bonds. First, we note that the time T_0 value of the $\frac{1}{P(T_0,T_1)}$ term is equal to 1. Therefore, its time t value under no arbitrage must be the price of a zero-coupon bond which matures at T_0, given by $P(t,T_0)$. Since the term $1 + \delta K$ is deterministic (known at t), we can discount it directly via a T_1-bond. Thus, for the present t time value of the FRA we obtain:

$$N\left(P(t,T_0) - P(t,T_1) - \delta K P(t,T_1)\right). \tag{2.10}$$

Since there is a counter-party to the FRA contract (the FRA seller), the present value of the contract has to be zero. Equating (2.10) to 0 and solving for K we obtain (2.8) which proves the claim. □

Corollary 2.15. *At $t = T_0$ the value of the simple forward rate given by (2.8) is equal to $L(T_0, T_1)$.*

Proof. Immediate consequence of equations (2.8) and (2.6) using the no credit risk assumption, $P(T_0, T_0) = 1$. □

Definition 2.16 (Instantaneous forward rate)**.** For forward rate agreements with infinitesimally small lending periods we define the so called instantaneous forward rate, denoted by $f(t,T)$ as

$$\lim_{T_1 \to T_0 \equiv T} F(t, T_0, T_1) = f(t,T).$$

Definition 2.17 (Forward curve)**.** If we fix t and vary the maturity date T, we get the term structure of instantaneous forward rates prevailing at t which we call the forward curve.

The relations between the forward curve, discount curve and the yield curve derived in Lemma 2.18 be used frequently throughout the text.

Lemma 2.18.

(i) $P(t,T) = e^{-\int_t^T f(t,u)du}$.

(ii) $y(t,T) = \frac{1}{T-t}\int_t^T f(t,u)du$.

Proof.

(i) First we rewrite the expression for the simple forward rate, (2.8), as:

$$F(t,T_0,T_1) = -\frac{1}{P(t,T_1)}\left(\frac{P(t,T_1) - P(t,T_0)}{T_1 - T_0}\right). \tag{2.11}$$

Taking the limit $T_1 \to T_0 \equiv T$ on both sides in (2.11), we notice that $f(t,T) = -\frac{\partial \log P(t,T)}{\partial T}$ which proves the claim.

(ii) Follows immediately from *i)* and (2.6).

□

Corollary 2.19. *The short rate and the forward rate are related by $r(t) = f(t,t)$.*

Proof. From result *i)* in Lemma 2.18 we notice that $f(t,t) = -\frac{1}{P(t,t)}\frac{\partial P}{\partial T}(t,t)$. The conclusion now follows form the no credit risk assumption $P(t,t) = 1$ and (2.7). □

Lemma 2.20. *The second order derivative of the discount curve is given by:*

$$\frac{\partial^2 P}{\partial T^2}(t,T) = \lim_{h\to 0} \frac{P(t,T)}{h}\left(F(t,T-h,T) - F(t,T,T+h)\right). \tag{2.12}$$

Proof. Using equation (2.8) we obtain the following expression for the second derivative of the discount curve:

$$\begin{aligned}\frac{\partial^2 P}{\partial T^2}(t,T) &= \lim_{h\to 0} \frac{P(t,T-h) - 2P(t,T) + P(t,T+h)}{h^2} \\ &= \lim_{h\to 0} \frac{P(t,T)}{h}\left(F(t,T-h,T) + \frac{1}{h}\left(\frac{1}{1+hF(t,T,T+h)} - 1\right)\right).\end{aligned} \tag{2.13}$$

Now, define:

$$g(h) \equiv \frac{1}{1+hF(t,T,T+h)}.$$

We now calculate:

$$g'(h) = -(1+hF(t,T,T+h))^{-2}\left(F(t,T,T+h) + h\frac{\partial F}{\partial T_1}(t,T,T+h)\right). \tag{2.14}$$

Now, since $F(t,T,T) = 0$ as can be seen from (2.8) using (2.14) we obtain $g'(0) = 0$. Explicitly calculating again:

$$\begin{aligned}g''(h) = &- (1+hF(t,T,T+h))^{-2}\left(2\frac{\partial F}{\partial T_1}(t,T,T+h) + h\frac{\partial^2 F}{\partial T_1^2}(t,T,T+h)\right) \\ &+ (1+hF(t,T,T+h))^{-3}\left(F(t,T,T+h) + h\frac{\partial F}{\partial T_1}(t,T,T+h)\right)^2\end{aligned}$$

Thus, $g''(0) = -2\frac{\partial F}{\partial T_1}(t,T,T)$. Using Taylor expansion for small h we obtain:

$$g(h) = g(0) + hg'(0) + \frac{1}{2}h^2 g''(0) + O(h^3) = 1 - h^2\frac{\partial F}{\partial T_1}(t,T,T) + O(h^3). \tag{2.15}$$

Plugging in (2.15) in (2.13) we see that the $O(h^3)$ term becomes $O(h)$ and thus is equal to zero when h goes to zero and we obtain:

$$\begin{aligned}\frac{\partial^2 P}{\partial T^2}(t,T) &= \lim_{h\to 0}\frac{P(t,T)}{h}\left(F(t,T-h,T) - h\frac{\partial F}{\partial T_1}(t,T,T)\right)\\ &= \lim_{h\to 0}\frac{P(t,T)}{h}\left(F(t,T-h,T) - F(t,T,T+h) + F(t,T,T)\right),\end{aligned}$$

which proves the claim. □

According to (2.12), the second order derivative of the discount curve depends on the differences in the forward rates between two consecutive infinitesimal lending periods. Relation (2.12) will be used as an economic justification for defining the minimization norm in Chapter 3.

Definition 2.21 (Interest rate future)**.** An interest rate futures contract, similarly to the forward rate agreement allows its holder to manage the exposure to the simple interest rate $L(T_0, T_1)$ prevailing over a future period $[T_0, T_1]$. Unlike the forward rate agreement the interest rate future is daily marked to market, so that the payoff at time t is zero.

The price of the futures contract is quoted as:

$$P_{ft}(t, T_0, T_1) = 100(1 - R_{ft}(t, T_0, T_1)),$$

where $R_{ft}(t, T_0, T_1)$ is the futures rate. Marking to market means that at time $t + dt$ there is a cash flow to the holder of the futures contract given by the difference in the futures price at times $t + dt$ and t, expressed as $P_{ft}(t + dt, T_0, T_1) - P_{ft}(t, T_0, T_1)$.

Interest rate futures are more commonly traded than forward rate agreements. However, In general there is no closed-form formula for the futures rate. Therefore, in order to price interest rate futures for simplicity we take the futures rate to be equal to the simple forward rate for a forward contract over the same period.

2.6 Coupon bonds and interest rate swaps

Definition 2.22 (Fixed coupon bond)**.** A fixed coupon bond is a financial contract specified by a collection of coupon dates $T_1, T_2, \ldots, T_n$. At each date T_i the holder of the contract receives a fixed (known at time t) payment coupon with value c_i for all $i \in \{1, 2, \ldots, n\}$. A principle value of amount N is paid to the holder at the maturity date T_n.

The price of the fixed coupon bond at time $t \leq T_n$, $c(t)$ can be expressed as a sum of zero-coupon bonds $P(t, T_i)$ with maturity dates T_i as follows:

$$c(t) = \sum_{i=1}^{n} P(t, T_i) c_i \mathbb{1}_{t<T_i} + P(t, T_n)N. \tag{2.16}$$

Using (2.6), we may rewrite the price of coupon bond in terns of the yield as follows:

$$c(t) = \sum_{i=1}^{n} e^{-(T_i-t)y(t,T_i)} c_i \mathbb{1}_{t<T_i} + e^{-(T_n-t)y(t,T_n)} N. \tag{2.17}$$

Definition 2.23 (Floating rate note)**.** A floating rate note is specified by a collection of dates $T_0, T_1, \dots, T_n$, where T_0 is called the first reset date and T_n is the maturity date. The note pays floating rate coupon payments at all dates starting with T_1 given by $c_i = (T_i - T_{i-1})L(T_{i-1}, T_i)N$ and a notional N at the maturity date T_n.

Unlike the coupons in the fixed coupon bond, the coupons payed by the floating rate note are not deterministic. Their value depends on the simple interest rate $L(T_{i-1}, T_i)$, which is not know at the contract specifying date t.

Theorem 2.24. *The arbitrage free price of the floating rate note at time $t \leq T_0$ is $P_{FN}(t) = NP(t, T_0)$.*

Proof. To prove the theorem we construct a self financing portfolio with the same cash payments as the floating rate note. Assume that at time $t \leq T_0$ we have an amount of $NP(t, T_0)$ and we use that amount to buy N, $P(t, T_0)$ zero-coupon bonds. The bonds mature at T_0 and thus we have a net cash flow of N at T_0. We use the obtained cash flow N at T_0 to buy $\frac{N}{P(T_0,T_1)}$, $P(T_0, T_1)$ zero-coupon bonds; thus at T_1 we receive $\frac{N}{P(T_0,T_1)} = N\left(\frac{1}{P(T_0,T_1)} - 1\right) + N = c_1 + N$ (where we have used relation (2.6)). We keep the coupon c_1 and reinvest the notional N into $P(T_1, T_2)$ bonds. Repeating this investment process at each T_i we receive $c_i + N$. Thus, we have succeeded in constructing a self financing portfolio which replicates the floating rate payments. Absence of arbitrage implies that a self financing portfolio yielding the same cash flows as a traded security (in this particular case the floating rate note) must have initial price equal to $P_{FN}(t)$. $\square$

Definition 2.25 (Interest rate swap)**.** An interest rate swap is an exchange of fixed and floating coupon payments. It is specified by a collection of dates $T_0, T_1, \cdots, T_n$, a fixed rate K and a notional N. For simplicity in notation we are going to assume that the time periods have equal length $\delta = T_i - T_{i-i}$ for all $i \in \{1, 2, \dots, n\}$. At time T_i the holder pays fixed δKN and receives floating $\delta NL(T_{i-1}, T_i)$.

Thus, by definition the value of the interest rate swap V_p to the holder is equal to the value of the floating rate note less the price of the fixed coupon bond. Using the result from Theorem 2.24 and equation (2.16) the price of the interest rate swap can be written as:

$$V_p(t) = N(P(t,T_0) - P(t,T_n) - \delta K \sum_{i=1}^{n} P(t,T_i)).$$

Definition 2.26 (Forward swap rate)**.** The value of K which renders the payoff of the interest rate swap zero is called the forward swap rate for $t \leq T_0$ or the spot swap rate for $t = T_0$. It is denoted by R_{swap} and its value is given by:

$$R_{swap} = \frac{P(t,T_0) - P(t,T_n)}{\delta \sum_{i=1}^{n} P(t,T_i)}. \tag{2.18}$$

2.7 Duration and Convexity

We next define two important first and second order sensitivity measures for bond portfolio risk management. Duration and convexity measure how bond prices change under parallel shift of the yield curve.

Definition 2.27 (Duration)**.** We define the duration D as the relative first order sensitivity of the bond price $p(y)$, under a parallel shift of the yield curve from $y(t,T)$ to $y(t,T)+s$, where $s \in \mathbf{R}$ is a small fixed parameter. Thus, D can be expressed as:

$$D = -\frac{1}{p(y)} \frac{d}{ds} p(y;s)\bigg|_{s=0}, \tag{2.19}$$

where $p(y;s)$ denotes the bond price for the shifted yield curve $y(t,T)+s$ and $\frac{d}{ds}p(y;s)\big|_{s=0}$ denotes the Gateaux functional derivative of the bond price. Using result *ii* from Lemma 2.18, we notice that a parallel shift of s in the yield curve is equivalent to a shift of s in the forward curve. Equivalently the duration can be defined as the relative first order sensitivity of the bond price under a parallel shift of the forward curve.

As a concrete example, consider a fixed coupon bond with a price given by (2.17). For simplicity of notation assume $t = 0$ and that for all coupon date payments we have $T_i > 0$. Then according to (2.17), the price at of the bond p at time zero is a function of the yield curve at time $t = 0$, which we are going to denote by $y(T) \equiv y(0,T)$. We further denote $y_i \equiv y(T_i)$ and we assume that the principal of the bond N is contained in the coupon payment c_n. Then from (2.17), the price of the bond is given by a functional transformation of the yield curve:

$$p = \sum_{i=1}^{N} e^{-T_i y_i} c_i.$$

To calculate the duration consider a parallel shift of the yield curve $y(T) \mapsto y(T) + s$ for some fixed $s \in \mathbf{R}$. Then according to (2.19) the duration of the bond is given by:

$$D = -\frac{1}{p}\frac{d}{ds}\left(\sum_{i=1}^{N} e^{-T_i(y_i+s)} c_i\right)\Bigg|_{s=0} = \frac{\sum_{i=1}^{N} T_i e^{-T_i y_i} c_i}{\sum_{i=1}^{N} e^{-T_i y_i} c_i}.$$

The aim of duration hedging is to immunize a portfolio with respect to small parallel shifts in the forward curve, $f \equiv f(t,T)$. Specifically if $\Pi(f;s)$ is the value of the portfolio to be hedged as a function of s, $H(f;s)$ the value of the hedging instrument as a function of s the aim is to find the amount a of $H(f;s)$ to be held such that the following requirement is satisfied:

$$\frac{d}{ds}(\Pi(f;s) + aH(f;s))\Bigg|_{s=0} = 0$$

Using the definition of duration it is easy to see that the solution is given by:

$$a = \frac{-D_\Pi \Pi(f)}{D_H H(f)},$$

where D_Π and D_H denote the duration of the portfolio and the hedging instrument respectively and $\Pi(f)$ and $H(f)$ denote the amounts held before the shift in the forward curve.

Definition 2.28 (Convexity)**.** We define the convexity C as the relative second order sensitivity of the bond price $p(y)$, under a parallel shift of the yield curve from $y(t,T)$ to $y(t,T) + s$, where s is a fixed parameter. Thus, C is given by:

$$C = \frac{1}{p(y)}\frac{d^2}{ds^2}p(y;s)\Bigg|_{s=0}$$

Chapter 3

Smooth discount curve estimation

In financial modeling, it is often assumed that the discount curve of zero-coupon bond prices can be observed for a continuum of maturities. This discount curve can be used further on as an exogenously given initial condition in stochastic models for modeling the time series of terms structures and pricing of various fixed income instruments. In practise zero-coupon bonds are rarely traded and one has to derive the implied zero-coupon bond prices from the non arbitrage prices of fixed income instruments which are more actively traded. Such examples include forward rates, simple rates, coupon bonds, and interest rate swaps. In particular, we are going to use the non arbitrage price expressions for these contracts as defined in Equations (2.8), (2.6), (2.16) and (2.18), to derive the zero-coupon bond prices at the starting time, for finitely many maturity dates. Once we have a large enough price vector of zero-coupon bonds we face an interpolation problem of finding a discount curve which is smooth and matches the derived prices. The method presented in this chapter is based on [8].

3.1 Problem statement

For notational simplicity, we set the starting date to be $t = 0$ and we denote by $q = (q_1, q_2, q_3, \ldots, q_n)^T$ the vector of observed prices of n fixed income instruments at time $t = 0$. We further denote by $0 \leq T_1 \leq T_2 \leq \cdots \leq T_N$ the union of all cash-flow dates of these instruments and by $C = (c_{ij})$ the corresponding $n \times N$ cash-flow matrix, where c_{ij} specifies the cash-flow at time T_j for instrument i. If instrument i does not have a cash-flow at time T_j we set $c_{ij} = 0$. All dates which are relevant for the valuation of the instrument are included in the cash-flow matrix; thus we allow for both positive and negative values in the the matrix entries c_{ij}. We are going to assume that the matrix C has full rank. If this were not the case we would be including redundant fixed income instruments in our matrix C since they can be replicated by a linear combination of other instruments and do not impose additional requirements on the discount curve. For instance, using formula (2.16), it is easy to see that two coupon bonds with the same characteristics but different notional values impose the same constraint on the discount curve (assuming a non arbitrage setting in which the formula is valid). Since, the vector of observed prices is obtained by discounting the cash-flow matrix we may write the linear system:

$$Cd = q, \tag{3.1}$$

where $d = (P(T_1), P(T_2), P(T_3), \ldots, P(T_N))^T$ is a collection of prices for zero-coupon bonds with maturities T_i. Thus, C and q are known and the aim is to solve (3.1) for d.

In order to simplify the notation we suppress the current time $t = 0$ in the arguments and we write $P(T_i)$ instead of $P(0, T_i)$. If C were an invertible square matrix the unique solution to (3.1) would be simply given by $d = C^{-1}q$. In practise however, there are usually a lot more cash-flow dates than fixed income instruments available at a given date ($N \gg n$) so the linear system is under-determined and there exist many vectors d which satisfy the relation (3.1). Hence, there are two problems to be solved:

1. Which admissible discount vector should be chosen?
2. Once the admissible discount vector is chosen how do we find the discount curve; that is how do we interpolate for $P(T)$, where $T \in (0, T_N)$?

3.1.1 Market examples

In the framework of the linear system defined above we first note that since the discount curve has to start at face value i.e. $P(0) = 1$, we need to impose $T_1 = 0$, $c_{11} = 1$ and $c_{1j} = c_{i1} = 0$ for all $i \neq j$. The definition of the other entries of the matrix C will depend on the type of benchmark instruments available at the current date. In particular, we have the following simple examples:

1. Treasury bills: Treasury bills are short term debt obligations backed by the U.S. treasury department with maturity less than 1 year. Since the treasury bills do not pay coupons they are the market prototype for zero-coupon bonds. Thus a treasury bill i with maturity T_j will enter the matrix C by setting $c_{ij} = 1$ and $c_{ik} = 0$ for all $k \neq j$.
2. LIBOR rates: LIBOR is the average inter-bank interest rate at which selections of banks operating in the London currency market are prepared to lend to one

another. The LIBOR is calculated daily as the average of the estimates offered by the collection of banks and has maturities ranging from overnight to one year. The LIBOR rate is the market prototype for the simple rate as defined in (2.1). Thus according to (2.6) a LIBOR rate i enters the framework above as instruments with price 1 today i.e. $q_i = 1$ and a cash-flow $1+T_iL(T_i)$ at time T_i. Furthermore, as we can see from the definition of the short rate, $\lim_{T\to t} L(t,T) \equiv r(t)$, the overnight LIBOR rate can be used as an approximation for the current short rate.

3. Interest rate futures: Any type of interest rate future such as the Treasury bills futures or the Eurodollar futures in this framework are approximated by forward rate agreements (see discussion at the end of Section 2.5). In accordance with relation (2.8) a simple forward rate $F(T_{i-1}, T_i)$ will enter as an instrument with price 0 today, a cash-flow of -1 at time T_{i-1} and a cash-flow of $1+(T_i - T_{i-1})F(T_{i-1}, T_i)$ at time T_i.

Thus, the advantage of using the representation given by (3.1) lies in the fact that we can use many readily available prices for various fixed income contracts as inputs. The reliability of the method is however questionable, since we have to assume that the current market prices of these instruments match their mathematically derived values. In particular, we have to assume that at any given point in time there is no arbitrage on the market.

3.2 Optimal discount curve in Hilbert space

Instead of searching for an admissible discount vector d, and then continuously interpolating to find the discount curve (which can for example be done by using polynomial splines) we use an alternative approach by converting the problem into an optimisation problem in infinite dimensional Hilbert space given the matching constraint. First, we briefly review some of the mathematical tools for optimization in functional spaces.

3.2.1 Optimization of Functionals in Hilbert space

The following subsection introduces the notion of the Fréchet derivative and states the necessary conditions for applying Lagrangian methods to optimization problems in functional analysis. For a more complete discussion the reader is referred to Chapter 4 in [16] and Section 43.8 in [17].

Definition 3.1 (Fréchet derivative)**.** Let V and W be normed vector spaces, and $U \subset V$ be an open subset of V. A map $f : U \to W$ is called Fréchet differentiable at $x \in U$ if there exists a bounded linear operator $A : V \to W$ such that:

$$\lim_{\|h\|\to 0} \frac{\|f(x+h) - f(x) - Ah\|_{\mathbf{W}}}{\|h\|_{\mathbf{V}}} = 0.$$

If such an operator A exists it is unique and we will denote the Fréchet derivative of the map f at x by $f'(x) \equiv A$. Denote by $B(V, W)$ the set of all bounded linear operators $V \to W$. A map f that is Fréchet differentiable at every point in U is said to be C^1 if the function $f' : U \to B(V, W)$, $x \mapsto f'(x)$ is continuous.

Theorem 3.2 (Lagrange multiplier rule)**.** *Let V and W be normed vector spaces and $u \in V$. Suppose $f : U(u) \subseteq V \to \mathbf{R}$ and $g : U(u) \subseteq V \to W$ are C^1 in $U(u)$, where $U(u)$ is an open neighbourhood of u. Let u be a solution to the following minimization problem:*

$$\begin{aligned} &\min f(u) \\ &s.t.\ g(u) = 0, \end{aligned} \tag{3.2}$$

where $g'(u) \in B(V, W)$ is surjective. Then there exists a functional $\lambda^* : W \to R$ such that the following holds:

$$f'(u) + \lambda^* g'(u) = 0. \tag{3.3}$$

Proof. See proof of Proposition 1 in Section 4.14 in [16]. □

In the special case that is going to be used throughout this chapter we have $W = \mathbf{R}^n$ and $g(u) = \big(g_1(u), g_2(u), \dots, g_n(u)\big)$. Thus the surjectivity of $g'(u)$ is equivalent to the fact that for every $w = (w_1, w_2, \dots, w_n)$ the system $g_j'(u)h = w_j$ for $j = 1, 2, \dots, n$ has a solution $h \in V$. Then condition (3.3) is equivalent to:

$$f'(u) + \sum_{j=1}^{n} \lambda_j^* g_j'(u) = 0,$$

where $\lambda^* = \big(\lambda_1^*, \lambda_2^*, \dots, \lambda_n^*\big) \in \mathbf{R}^n$. Suppose now that f and g satisfy the requirements of Theorem 3.2 and take $W = \mathbf{R}^n$. We define the Lagrangian for the optimization problem (3.2), $\mathcal{L}(u, \lambda) = f(u) + \lambda \cdot g(u)$. It is easy to see that $\mathcal{L}(u, \lambda + s) - \mathcal{L}(u, \lambda) = s \cdot g(u)$ for all $s \in \mathbf{R}^n$. Thus, the optimization constraint $g(u) = 0$ is equivalent to the fact that Fréchet derivative of the Lagrangian with respect to $\mathbf{R}^n$ vanishes since $s \in \mathbf{R}^n$ was arbitrary. The Fréchet derivative of the Lagrangian with respect to V satisfies $D_u\mathcal{L}(u, \lambda) = f'(u) + \lambda g'(u)$ thus if $\lambda = \lambda^*$ and u is a solution to the optimization problem, then the Fréchet derivative of the Lagrangian with respect to V vanishes. Thus, if u is a solution to the optimization problem (3.2) then the Fréchet derivatives of the Lagrangian with respect to V and $\mathbf{R}^n$ vanish. For a discussion on the sufficient conditions see Section 43.8 in [17].

In our further work whenever faced with an optimization problem we are going to assume that all the sufficient and necessary conditions for applying the Lagrange multiplier rule stated in this chapter hold.

3.2.2 Hilbert space of discount curves

Let us define the vector space of discount curves $\mathbf{H}$ to consist of all real functions $P\colon [0, T_{max}] \to \mathbf{R}$ with absolutely continuous first derivatives. Then, P'' exists a.e. and $P'(T) - P'(0) = \int_0^T P''(y)dy$ for all $0 \leq T \leq T_{max}$[18]. In addition we require that P'' is square integrable i.e. $\int_0^{T_{max}} P''(x)^2 dx < \infty$.

We require absolute continuity of the first derivative in order to guarantee the existence and continuity of the forward curve, which as explained in the previous chapter is given by $f(T) = (-\log P(T))'$. T_{max} is chosen to be bigger than the largest maturity date of the benchmark fixed-income instruments available (i.e. $T_{max} > T_N$). The next step is to choose a suitable norm in the vector space $\mathbf{H}$ which will be later on minimized. There are various choices that can be made; however, economic intuition sets up some general guidelines which ought to be followed.

As we saw in the previous section, in particular by (2.12), the second derivative of the discount curve is determined by the differences in the forward rate between two consecutive infinitesimal periods. Economic intuition suggests that the forward rate locked in today should not vary over two consecutive infinitesimal periods. This requirement can be achieved by minimizing the second order derivative of the discount curve in some proper sense. As an initial simplification we are going to impose that the short rate be minimized. As can be seen from relation (2.7), this can be achieved by minimizing the first order derivative of the discount curve at 0.

Hence, starting from the premises that the instantaneous short rate should be minimized and the flatness of the forward curve we may define the vector space $\mathbf{H}$ scalar product as follows. For any two elements $\alpha \in \mathbf{H}$ and $\beta \in \mathbf{H}$ we have:

$$\langle \alpha, \beta \rangle_{\mathbf{H}} = \alpha(0)\beta(0) + \alpha'(0)\beta'(0) + \int_0^{T_{max}} \alpha''(y)\beta''(y)dy.$$

In particular, for the norm of the discount curve, $P \in \mathbf{H}$, we obtain:

$$\|P\|_{\mathbf{H}}^2 = \langle P, P \rangle_{\mathbf{H}} = P(0)^2 + P'(0)^2 + \int_0^{T_{max}} P''(y)^2 dy \tag{3.4}$$

Lemma 3.3. $\mathbf{H} = \{P\colon [0, T_{max}] \to \mathbf{R} | P'$ *is absolutely continuous and* $\int_0^{T_{max}} P''(x)^2 dx < \infty\}$ *is a real Hilbert space under the scalar product* $\langle \alpha, \beta \rangle_{\mathbf{H}} = \alpha(0)\beta(0) + \alpha'(0)\beta'(0) + \int_0^{T_{max}} \alpha''(y)\beta''(y)dy$ *for all* $\alpha \in \mathbf{H}$ *and* $\beta \in \mathbf{H}$.

Proof. From its definition we can immediately see that the scalar product is symmetric and linear in both arguments and that $\langle \alpha, \alpha \rangle_{\mathbf{H}} = 0$ if an only if $\alpha = 0$ a.e. It remains to show completeness under the norm induced from the inner product. Let $(P_n)_{n\geq 1}$ be a Cauchy sequence in $\mathbf{H}$. That is for any $\epsilon > 0$ there exists a positive integer N such that for all integers $n, m > N$ we have $\|P_n - P_m\|^2_{\mathbf{H}} < \epsilon$. Thus, we obtain:

$$\|P_n - P_m\|^2_{\mathbf{H}} = (P_m(0) - P_n(0))^2 + (P'_m(0) - P'_n(0))^2 + \int_0^{T_{max}} (P''_m(y) - P''_n(y))^2 dy < \epsilon \quad (3.5)$$

From (3.5), we see that $(P_n(0))_{n\geq 1}$ is a Cauchy sequence in $\mathbf{R}$, hence it converges to some $a \in \mathbf{R}$. Similarly, $(P'_n(0))_{n\geq 1}$ is Cauchy in $\mathbf{R}$ converging to some value $b \in \mathbf{R}$. The last term in the sum in (3.5) implies that $(P'')_{n\geq 1}$ is Cauchy in $L^2[0, T_{max}]$ thus it converges to some $f \in L^2[0, T_{max}]$. First, we note that by Cauchy - Schwartz inequality we have:

$$\left(\int_0^{T_{max}} |f(y)|\, dy \right)^2 \leq T_{max} \int_0^{T_{max}} f(y)^2 dy < \infty,$$

so $f \in L[0, T_{max}]$. Now define $u(x) \equiv b + \int_0^x f(y)dy$ for all $x \in [0, T_{max}]$. Then u is absolutely continuous and $u(0) = b$ (see Lemma 3.70 in [19]). Next we define $P(x) \equiv a + \int_0^x u(y)dy$. Then $P(0) = a$, $P'(x) = u(x)$, thus $P'(x)$ is absolutely continuous. Furthermore, $P''(x) = u'(x) = f(x)$ so clearly $P'' \in \mathbf{H}$. Explicitly, we compute:

$$\begin{aligned} \|P_n - P\|^2_{\mathbf{H}} &= (P_n(0) - P(0))^2 + (P(0) - P'_n(0))^2 + \int_0^{T_{max}} (P(y) - P''_n(y))^2 dy \\ &= (P_n(0) - a)^2 + (b - P'_n(0))^2 + \int_0^{T_{max}} (f(y) - P''_n(y))^2 dy \end{aligned} \quad (3.6)$$

Taking the limit as $n \to \infty$ in (3.6) we obtain $P_n \to P$ and since $P \in \mathbf{H}$ we have that $\mathbf{H}$ is indeed a Hilbert space. □

For each time $T \in [0, T_{max}]$, we define a linear functional $\Phi_T : \mathbf{H} \to \mathbf{R}$, which evaluates the discount curve at T, so $\Phi_T(P) = P(T)$. Using the fundamental theorem of calculus and integration by parts we have:

$$\Phi_T(P) = P(T) = P(0) + \int_0^T P'(y)dy = P(0) + TP'(0) + \int_0^T (T - y)P''(y)dy. \quad (3.7)$$

Using the triangle inequality, equation (3.7) and the fact that $T_{max} \geq T$ we obtain:

$$\begin{aligned}|\Phi_T(P)|| &\leq |P(0)| + |TP'(0)| + \int_0^T (T-y)|P''(y)|dy \\ &\leq |P(0)| + |T_{max}P'(0)| + \int_0^{T_{max}} (T_{max}-y)|P''(y)|dy.\end{aligned}$$

Finally, using the Cauchy-Schwartz inequality we have:

$$\begin{aligned}|\Phi_T(P)| &\leq |P(0)| + |T_{max}P'(0)| + \int_0^{T_{max}} (T_{max}-y)|P''(y)|dy \\ &\leq \sqrt{1+T_{max}^2+T_{max}^3}\|P\|_{\mathbf{H}}.\end{aligned}$$

Thus, $\Phi_T(P)$ is a bounded linear functional. Now, according to the Riesz representation theorem for each T there exists a unique element ϕ_T such that for all $P \in \mathbf{H}$ we have:

$$\langle \phi_T, P \rangle_{\mathbf{H}} = \Phi_T(P)$$

The following lemma gives us an explicit form for the Riesz element.

Lemma 3.4. *Denote $m = \min(x, T)$. The linear functional Φ_T on $\mathbf{H}$ can be uniquely represented by the element $\phi_T \in \mathbf{H}$ given by*

$$\phi_T(x) = 1 - \frac{1}{6}m^3 + \frac{1}{2}xT(2+m).$$

Proof. Using integration by parts and the fundamental theorem of calculus we first rewrite $P(T)$ as follows:

$$P(T) = P(0) + \int_0^T P'(y)dy = P(0) + TP'(0) + \int_0^T (T-y)P''(y)dy \tag{3.8}$$

Next from equations (3.4) and (3.8) we obtain the following conditions for ϕ_T:

$$\begin{aligned}\phi_T(0) &= 1 \\ \phi_T'(0) &= T \\ \phi_T''(x) &= (T-x)\mathbb{1}_{x\in[0,T]}.\end{aligned}$$

Integrating the last condition we have:

$$\phi_T'(x) = T + \int_0^x (T-y)\mathbb{1}_{y\in[0,T]}dy. \tag{3.9}$$

When $x \geq T$, the integral in (3.9) does not depend on x due to the indicator function. Thus, it is enough to integrate to m and we obtain:

$$\phi'_T(x) = T - \frac{m^2}{2} + Tm. \tag{3.10}$$

For the final form of the element we have the expression:

$$\phi_T(x) = 1 + \int_0^x (T - \frac{m'^2}{2} + Tm')dy, \tag{3.11}$$

where now $m' = \min(y, T)$. Splitting the integral limits in expression (3.11), we have:

$$\int_0^x (T - \frac{m'^2}{2} + Tm')dy = \int_0^m (T - \frac{m'^2}{2} + Tm')dy + \int_m^x (T - \frac{m'^2}{2} + Tm')dy \tag{3.12}$$

Noticing that the $m' = y$ inside the first integral on the left hand side of expression (3.12) and $m' = T$, if $T < x$ in the second integral, we obtain the following expression:

$$\phi_T(x) = 1 - \frac{1}{6}m^3 + \frac{T}{2}m^2 - \frac{T^2}{2}m + xT(1 + \frac{T}{2}). \tag{3.13}$$

Treating the cases when $T > x$ and $T < x$ in expression (3.13) separately proves the claim. □

From (3.10) we have:

$$\phi'_T(x) = -\frac{m^2}{2} + T + Tm. \tag{3.14}$$

Thus, we also obtain:

$$\phi''_T(x) = (T - x)\mathbb{1}_{x \in [0,T]}.$$

Next define a linear map $M\colon \mathbf{H} \to \mathbf{R}^n$ via the following:

$$MP = C(\Phi_{T_1}(P), \Phi_{T_2}(P), \ldots, \Phi_{T_N}(P))^T,$$

where C is the cash-flow matrix.

Lemma 3.5. *The adjoint operator of M, $M^*\colon \mathbf{R}^n \to \mathbf{H}$ exists and is defined via $(M^*z)(x) = \sum_{i=1}^N \phi_{T_i}(x) C_i^T \cdot z$.*

Proof. Since M is bounded and linear a unique adjoint operator exists defined to satisfy the following:

$$\langle MP, z \rangle_{\mathbf{R}^n} = \langle P, M^*z \rangle_{\mathbf{H}},$$

for all $z \in \mathbf{R}^n$ and for all $P \in \mathbf{H}$. Using the definition of M and Riesz representation of the functional Φ_T explicitly we have:

$$\begin{aligned}
\langle MP, z\rangle_{\mathbf{R}^n} &= \langle C(\Phi_{T_1}(P), \Phi_{T_2}(P), \cdots, \Phi_{T_N}(P))^T, z\rangle_{\mathbf{R}^n} \\
&= \langle C(\langle \phi_{T_1}, P\rangle_{\mathbf{H}}, \langle \phi_{T_2}, P\rangle_{\mathbf{H}}, \cdots, \langle \phi_{T_N}, P\rangle_{\mathbf{H}})^T, z\rangle_{\mathbf{R}^n} \\
&= \langle (\langle \phi_{T_1}, P\rangle_{\mathbf{H}}, \langle \phi_{T_2}, P\rangle_{\mathbf{H}}, \cdots, \langle \phi_{T_N}, P\rangle_{\mathbf{H}})^T, C^T z\rangle_{\mathbf{R}^n} \\
&= \sum_{i=1}^{N} \langle \phi_{T_i}, P\rangle_{\mathbf{H}} (C^T z)_i = \langle P, \sum_{i=1}^{N} \phi_{T_i} (C^T z)_i\rangle_{\mathbf{H}},
\end{aligned}$$

where in the last equality we have used the linearity of the inner product in $\mathbf{H}$. The claim now follows from the definition and the uniqueness of the adjoint. □

3.2.3 Smooth discount curve

Using the definition of the linear map M and the definition of the norm in the Hilbert space $\mathbf{H}$, we may write the problem of finding a smooth discount curve to match the prices of the benchmark fixed instruments as an optimization problem in an infinite dimensional space. Namely, our goal is to solve the following problem:

$$\begin{aligned}
&\min_{P \in \mathbf{H}} \frac{1}{2} \|P\|^2_{\mathbf{H}} \\
&s.t.\ MP = q
\end{aligned} \tag{3.15}$$

Theorem 3.6. *There exists a unique closed form solution of the optimization problem* (3.15) *and it is given by* $P^*(T) = p \cdot \phi(T)$. *Here* $p = C^T (CSC^T)^{-1} q$, $\phi(T) = (\phi_{T_1}(T), \phi_{T_2}(T) \cdots \phi_{T_N}(T))^T$ *and* S *is a positive definite matrix with components* $S_{ij} = \langle \phi_{T_i}, \phi_{T_j}\rangle_{\mathbf{H}}$.

Proof. The Lagrangian for the optimization problem (3.15) is given by:

$$\begin{aligned}
\mathcal{L}(P, \lambda) &= \frac{1}{2}\|P\|^2_{\mathbf{H}} + \lambda \cdot (MP - q) \\
&= \frac{1}{2}\|P\|^2_{\mathbf{H}} + \langle \lambda, MP\rangle_{\mathbf{R^n}} - \langle \lambda, q\rangle_{\mathbf{R^n}} \\
&= \frac{1}{2}\|P\|^2_{\mathbf{H}} + \langle M^* \lambda, P\rangle_{\mathbf{H}} - \langle \lambda, q\rangle_{\mathbf{R^n}},
\end{aligned}$$

where to obtain the last equality we have used the properties of the adjoint operator. Explicitly computing, we have:

$$\begin{aligned}
\mathcal{L}(P + h, \lambda) - \mathcal{L}(P, \lambda) &= \langle P, h\rangle_{\mathbf{H}} + \frac{1}{2}\langle h, h\rangle_{\mathbf{H}} + \langle M^* \lambda, h\rangle_{\mathbf{H}} \\
&= \langle P + M^* \lambda, h\rangle_{\mathbf{H}} + \frac{1}{2}\langle h, h\rangle_{\mathbf{H}}.
\end{aligned}$$

Thus, the Fréchet derivative of the Lagrangian in H is represented by $A_{\mathbf{H}}h = \langle P + M^*\lambda, h\rangle_{\mathbf{H}}$. Similarly, we compute:

$$\mathcal{L}(P, \lambda + s) - \mathcal{L}(P, \lambda) = \langle MP - q, s\rangle_{\mathbf{R^n}}$$

Hence, the Fréchet derivative of the Lagrangian in $\mathbf{R}^n$ is represented by $A_{\mathbf{R}^n}s = \langle MP - q, s\rangle_{\mathbf{R}^n}$. Since at optimum the Fréchet derivatives are equal to zero, we must have $\langle P + M^*\lambda, h\rangle_{\mathbf{H}} = 0$ and $\langle MP - q, s\rangle_{\mathbf{R}^n} = 0$ for all $h \in \mathbf{H}$ and for all $s \in \mathbf{R}^n$. Hence, we obtain the following conditions of the optimal discount curve P^*:

$$\begin{aligned} &P^* + M^*\lambda^* = 0 \\ &MP^* - q = 0. \end{aligned}$$

From the first optimally condition we obtain $P^* = -M^*\lambda^*$. Plugging into the second condition we have $-MM^*\lambda^* = q$.

Now, $MM^*\colon \mathbf{R}^n \to \mathbf{R}^n$ is a linear map. Let $z \in \mathbf{R}^n$ be arbitrary. Then, using the definitions of M and M^* and Lemma 3.5 we have:

$$\begin{aligned} MM^*z &= M(\sum_{i=1}^{N} \phi_{T_i} C_i^T \cdot z) \\ &= C\left(\Phi_{T_1}(\sum_{i=1}^{N} \phi_{T_i} C_i^T \cdot z), \Phi_{T_2}(\sum_{i=1}^{N} \phi_{T_i} C_i^T \cdot z), \ldots,\ \Phi_{T_N}(\sum_{i=1}^{N} \phi_{T_i} C_i^T \cdot z)\right)^T \\ &= C\left(\langle \phi_{T_1}, \sum_{i=1}^{N} \phi_{T_i} C_i^T \cdot z\rangle_{\mathbf{H}}, \langle \phi_{T_2}, \sum_{i=1}^{N} \phi_{T_i} C_i^T \cdot z\rangle_{\mathbf{H}}, \ldots, \langle \phi_{T_N}, \sum_{i=1}^{N} \phi_{T_i} C_i^T \cdot z\rangle_{\mathbf{H}}\right)^T. \end{aligned}$$

From the last relation we notice that the map MM^* can be represented by the matrix CSC^T, where S is a $N \times N$ symmetric matrix with entries $S_{ij} = \langle \phi_{T_i}, \phi_{T_j}\rangle_{\mathbf{H}}$. Using the Riesz representation theorem we have, $S_{ij} = \langle \phi_{T_i}, \phi_{T_j}\rangle_{\mathbf{H}} = \Phi_{T_i}(\phi_{T_j}) = \phi_{T_j}(T_i) = \phi_{T_i}(T_j)$. Hence, we have an explicit form for the matrix S entries:

$$S_{ij} = 1 - \frac{1}{6}\min(T_i, T_j)^3 + T_iT_j + \frac{1}{2}T_iT_j\min(T_i, T_j)$$

Since we have assumed that C has full rank, the matrix CSC^T is invertible if S is positive definite. Using the explicit form for the S matrix entries, we will choose the maturity dates T_k such that the matrix S is positive definite. By the construction of the matrix S, the matrix CSC^T is invertible. Therefore, $-\lambda^* = (CSC^T)^{-1}q$ and for the the optimal discount curve we obtain $P^* = M^*(CAC^T)^{-1}q$ which proves the claim. $\square$

Hence, we have constructed an explicit form of the discount curve that exactly replicates the prices q of all fixed income instruments available on the market at time $t = 0$ and minimizes the norm defined by (3.4). The corresponding forward curve can be calculated using result i from Lemma 2.18. Explicitly, we obtain:

$$f^*(T) = -\frac{\partial \log P^*(T)}{\partial T} = -\frac{p \cdot \phi'(T)}{p \cdot \phi(T)},$$

where $\phi'(T) = \big(\phi'_{T_1}(T), \phi'_{T_2}(T), \cdots \phi'_{T_N}(T)\big)$ and is computed using (3.14).

3.2.4 Exogenously specified short rate

As was explained in the previous section, the term $P'(0)^2$ leads to a minimization of the instantaneous short rate. In this section we fix an exogenously specified short rate $r \in \mathbf{R}$.

Define the linear functional $\Psi(P) = P'(0)$. Trivially, from the definition of the norm (3.4), Ψ is bounded. Furthermore, it is clear that $\psi(x) = x$ satisfies $\langle \psi, P \rangle_{\mathbf{H}} = \Psi(P)$ and by uniqueness it is the Rietz representation of the functional Ψ. Now, we add $\Psi(P) = -r$ as an additional constraint to (3.15). We have the following optimization problem:

$$\begin{aligned} &\min_{P \in \mathbf{H}} \frac{1}{2} \|P\|^2_{\mathbf{H}} \\ &s.t.\ MP = q \\ &s.t.\ \Psi(P) = -r \end{aligned} \tag{3.16}$$

To solve (3.16) we are going to first reduce it to an optimization problem of the form given by (3.15). To that end, define the modified price vector $\tilde{q} = (q, -r)^T$ and the modified cash flow matrix $\tilde{C} = (\tilde{c}_{ij})$ to be a block diagonal $(n+1) \times (N+1)$ matrix such that $\tilde{c}_{ij} = c_{ij}$ for $i \in \{1, \ldots, n\}$, $j \in \{1, \ldots, N\}$, $\tilde{c}_{(n+1)(N+1)} = 1$ and $\tilde{c}_{ij} = 0$ otherwise. Similarly as in the previous subsection define the linear map $\tilde{M} : \mathbf{H} \to \mathbf{R}^{n+1}$ such that $\tilde{M}P = \tilde{C}(\Phi_{T_1}(P), \Phi_{T_2}(P), \cdots, \Phi_{T_N}(P), \Psi(P))^T$. Therefore, we can rewrite the optimization problem (3.16) in the following form:

$$\begin{aligned} &\min_{P \in \mathbf{H}} \frac{1}{2} \|P\|^2_{\mathbf{H}} \\ &s.t. \tilde{M}P = \tilde{q}. \end{aligned} \tag{3.17}$$

Theorem 3.7. *There exists a unique closed form solution to the optimization problem 3.16 given by* $\tilde{P}^*(T) = \tilde{p} \cdot \tilde{\phi}(T)$, *where* $\tilde{p} = \tilde{C}^T(\tilde{C}\tilde{S}\tilde{C}^T)^{-1}\tilde{q}$, $\tilde{\phi}(T) = (\phi(T), \psi(T))^T$

and $\tilde{S}$ is a positive definite $(N+1)\times(N+1)$ matrix with components $\tilde{S}_{ij} = S_{ij}$ for $i, j \in \{1, 2 \ldots N\}$, $\tilde{S}_{i(N+1)} = T_i$ for $i < N+1$ and $\tilde{S}_{(N+1)(N+1)} = 1$.

Proof. Since we have reduced (3.16) to the optimization problem given by (3.17) the proof is analogous to the proof of Theorem 3.6. □

3.3 Discount curve sensitivities

The discount curve derived in Theorem 3.6 $P^*(T; q, C)$, depends on the prices and cash flows of the instruments used through the price vector q and the cash-flow matrix C. Thus, for hedging purposes it is important to examine the sensitivity of the optimal curve with respect to changes in q and changes in C. In particular, we have the following two lemmas:

Lemma 3.8. *The directional derivative $D_q P^* \cdot v \in \mathbf{H}$ of the optimal discount curve along a vector $v \in \mathbf{R}^n$ is given by $(D_q P^* \cdot v)(T) = c(v)^T \cdot \phi(T)$ where $c(v) = C^T(CSC^T)^{-1}v$*

Proof. Using the expression for the optimal discount curve in we have:

$$P^*(T) = \sum_{i=1}^{N} \left(C^T(CSC^T)^{-1}q\right)_i \phi_{T_i}(T) = \sum_{i=1}^{N}\sum_{j=1}^{n}(C^T(CSC^T)^{-1})_{ij} q_j \phi_{T_i}(T)$$

Thus, using the last expression:

$$\frac{\partial P^*}{\partial q_k}(T) = \sum_{i=1}^{N}(C^T(CSC^T)^{-1})_{ik}\phi_{T_i}(T)$$

Finally, we obtain:

$$(D_q P^* \cdot v)(T) = \sum_{k=1}^{n} v_k \frac{\partial P^*}{\partial q_k}(T) = \sum_{i=1}^{N}\sum_{k=1}^{n}(C^T(CSC^T)^{-1})_{ik} v_k \phi_{T_i}(T),$$

which proves the claim. □

Lemma 3.9. *The directional derivative $D_C P^* \cdot m \in \mathbf{H}$ of the optimal discount curve along a matrix $m \in \mathbf{R}^{n\times N}$ is given by $(D_C P^* \cdot m)(T) = f(m)^T \cdot \phi(T)$ where $f(m) = \left(m^T - C^T(CSC^T)^{-1}\left(CSm^T + mSC^T\right)\right)(CSC^T)^{-1}q$.*

Proof. Using the identity that for an invertible matrix K one has $(K^{-1})' = -K^{-1}K'K^{-1}$ we explicitly compute:

$$\begin{aligned}\frac{\partial P^*}{\partial C_{ij}}(T) &= -q^T(CSC^T)^{-1}\frac{\partial CSC^T}{\partial C_{ij}}(CSC^T)^{-1}C\phi(T) + p^T(CSC^T)^{-1}I_{ij}\phi(T)\\ &= q^T(CSC^T)^{-1}\big(I_{ij} - (CSI_{ji} + I_{ji}SC^T)(CSC^T)^{-1}C\big)\phi(T).\end{aligned}$$

Using the previous expression and the fact that $\sum_{i=1}^{N}\sum_{j=1}^{n} m_{ij}I_{ij} = m$ we have:

$$(D_C P^* \cdot m)(T) = \sum_{i=1}^{N}\sum_{j=1}^{n} m_{ij}\frac{\partial P^*}{\partial C_{ij}}(T) = f(m)^T\phi(T),$$

which proves the claim. □

The discount curve sensitivity can be used to hedge a portfolio of securities against changes in the discount curve. To illustrate this in practise suppose that we have a portfolio consisting entirely of coupon bonds each of which has a price p_i. For instance, suppose that the bond portfolio generates a cash flow of c_k at time τ_k, where the index k runs from 1 to K. If we denote the value of the portfolio today by V we have:

$$\frac{\partial V}{\partial p_i} = \sum_{k=1}^{K} c_k \frac{\partial P^*}{\partial p_i}(\tau_k)$$

Hence, the change in the value of the portfolio can be written using Lemma 3.8:

$$\Delta V = \sum_{i=1}^{n}\frac{\partial V}{\partial p_i}\Delta p_i = \sum_{i=1}^{n}\sum_{k=1}^{K} c_k\frac{\partial P^*}{\partial p_i}(\tau_k)\Delta p_i = \sum_{i=1}^{n}\sum_{k=1}^{K} c_k D_q P^* \cdot e_i(\tau_k)\Delta p_i$$

Therefore, by obtaining an amount of $-s_i$, where $s_i = \sum_{k=1}^{K} c_k D_q P^* \cdot e_i(\tau_k)$ of instrument i we can hedge the portfolio against changes in the prices of the benchmark instruments. Since hedging against changes in the individual prices of the instruments in the portfolio is not necessarily consistent with the movement of the discount curve an alternative approach is to hedge against shifts in the interest rate that are deemed most likely. For instance, an approach often used in practise is to consider functional shifts $\mu_j(T)$ $j = 1, \cdots, J$ to the forward curve $f^*(T)$. The sensitivity of the portfolio $V = V(f^*)$ is expressed via the following Gateaux functional derivative:

$$\left.\frac{dV(f^* + \epsilon s_j)}{d\epsilon}\right|_{\epsilon=0}$$

The choice $J = 1$ and $s_1(T) = 1$ is the standard duration hedge and can be treated in the same manner as stated in Section 2.7.

3.4 Optimal market quotes

Thus far it had been assumed that market quotes enter in the price vector q and the cash-flow matrix C without any error. In practise, we observe a price spread, determined by the so called bid-ask range. If we denote the bid and ask prices by q_b and q_a we have the following constraint $q_b \leq q \leq q_a$, where the inequality is component-wise. Thus, given the optimal discount curve for a given price vector as determined by Theorem 3.6, we can find the optimal admissible price that produces the smoothest curve.

Lemma 3.10. *The norm of the optimal discount curve P^* is given by $\|P^*\|_{\mathbf{H}}^2 = q^T \cdot (CSC^T)^{-1}q$.*

Proof. Using Theorem 3.6 and the definition of an adjoint operator we obtain:

$$\begin{aligned}
\|P^*\|_{\mathbf{H}}^2 &= \langle P^*, P^* \rangle_{\mathbf{H}} \\
&= \langle M^*(CSC^T)^{-1}q, M^*(CSC^T)^{-1}q \rangle_{\mathbf{H}} \\
&= \langle MM^*(CSC^T)^{-1}q, (CSC^T)^{-1}q \rangle_{\mathbf{R}^n} \\
&= \langle (CSC^T)(CSC^T)^{-1}q, (CSC^T)^{-1}q \rangle_{\mathbf{R}^n} \\
&= \langle q, (CSC^T)^{-1}q \rangle_{\mathbf{R}^n},
\end{aligned}$$

which proves the claim. □

Thus, we have to solve the following convex quadratic programming problem:

$$\begin{aligned}
&\min_{q \in \mathbf{R}^n} q^T \cdot (CSC^T)^{-1}q \\
&s.t.\ q_b \leq q \leq q_a
\end{aligned}$$

Now, if the market quotes enter through C as is the case with swaps and forward rates and if the bid ask ranges of the instruments are denoted by α_b and α_a we need to solve the following problem.

$$\begin{aligned}
&\min_{\alpha \in \mathbf{R}^n} q^T \cdot (C(\alpha)SC^T(\alpha))^{-1}q \\
&s.t.\ \alpha_b \leq \alpha \leq \alpha_a
\end{aligned} \tag{3.18}$$

Unlike the previous case, Problem (3.18) is not necessarily a convex programming problem, however with the help of the next lemma it can be solved using gradient based optimization algorithms.

Lemma 3.11. *The partial derivative of the norm of the optimal discount curve P^* with respect to an instrument quote α_i is given by:*

$$\frac{\partial \|P^*\|^2_{\mathbf{H}}}{\partial \alpha_i} = -2q^T \cdot \left((CSC^T)^{-1} \frac{\partial C}{\partial \alpha_i} SC^T (CSC^T)^{-1} q \right).$$

Proof. The result is an immediate consequence of Lemma 3.10 following a similar calculation as in Lemma 3.9. □

Chapter 4

Stochastic Calculus

In this chapter we collect the basic results from probability theory and stochastic analysis which are going to be used in modeling the fixed income market. We start by proving the Radon-Nikodym and Bayes' theorems. Moving on, we introduce the concept of a stochastic basis, which is a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathcal{P})$ carrying a d-dimensional Brownian motion $W = (W_1, W_2, \ldots, W_d)^T$. We move on to the Itó's theorem, The Girsanov change of measure theorem and the Martingale representation theorem. We finish by presenting the Fubini's theorem for stochastic integrals. The mathematical results in this chapter are not stated in their most general form. Instead the theorems are presented in a form suitable for their application in mathematical finance. In particular, geared towards the modeling of the equities and the fixed income markets.

4.1 Radon-Nikodym Theorem and Equivalent measures

Definition 4.1 (Absolutely continuous measures)**.** Let $(\Omega, \mathcal{F})$ be a measurable space and $\mathcal{P}$ and $\mathcal{Q}$ two measures on $(\Omega, \mathcal{F})$. We say that $\mathcal{Q}$ is absolutely continuous with respect to $\mathcal{P}$, written $\mathcal{Q} << \mathcal{P}$ if the following holds: For all $A \in \mathcal{F}$ we have that if $\mathcal{P}(A) = 0$, then $\mathcal{Q}(A) = 0$.

Definition 4.2 (Equivalent measures)**.** Let $(\Omega, \mathcal{F})$ be a measurable space and $\mathcal{P}$ and $\mathcal{Q}$ two measures on $(\Omega, \mathcal{F})$. If $\mathcal{Q}$ is absolutely continuous with respect to $\mathcal{Q}$ and $\mathcal{P}$ is absolutely continuous with respect to $\mathcal{P}$ we say that $\mathcal{Q}$ and $\mathcal{P}$ are equivalent measures. We denote $\mathcal{P} \sim \mathcal{Q}$.

If $\mathcal{P}$ and $\mathcal{Q}$ are probability measures on $(\Omega, \mathcal{F})$ we have the following simple result.

Lemma 4.3. *Two probability measures $\mathcal{P}$ and $\mathcal{Q}$ on a measurable space $(\Omega, \mathcal{F})$ are equivalent if and only if $\mathcal{P}(A) = 1 \iff \mathcal{Q}(A) = 1$ for all $A \in \mathcal{F}$.*

Proof. Follows immediately form the definition of a probability measure and the definition of equivalent measures. □

Remark 4.4. It follows from the previous lemma that if $\mathcal{Q}$ and $\mathcal{P}$ are two equivalent probability measures we have $\mathcal{P}(A) > 0 \iff \mathcal{Q}(A) > 0$ for all $A \in \mathcal{F}$.

Theorem 4.5 (Radon-Nikodym theorem)**.** *Let $(\Omega, \mathcal{F})$ be a measurable space and $\mathcal{P}$ and $\mathcal{Q}$ two finite, positive measures on $(\Omega, \mathcal{F})$ such that $\mathcal{Q}$ is absolutely continuous with respect to $\mathcal{P}$. Then there exists a non-negative function $L : \Omega \mapsto \mathbf{R}$ such that L is $\mathcal{F}$-measurable, $\int_\Omega L(x)d\mathcal{P}(x) < \infty$ and $\mathcal{Q}(A) = \int_A L(x)d\mathcal{P}(x)$ for all $A \in \mathcal{F}$. We call L the Radon-Nikodym derivative of $\mathcal{Q}$ with respect to $\mathcal{P}$. The Radon-Nikodym derivative is uniquely determined $\mathcal{P} -$ a.e. and we write $L(x) = \frac{d\mathcal{Q}(x)}{d\mathcal{P}(x)}$.*

Proof. We define a new measure on $(\Omega, \mathcal{F})$, $\lambda(A) = \mathcal{P}(A) + \mathcal{Q}(A)$ for all measurable events $A \in \mathcal{F}$. We trivially observe that by its definition λ is a finite positive measure on $(\Omega, \mathcal{F})$ and any function $g \in L^2(\lambda)$ it holds that $g \in L^2(\mathcal{P})$ and $g \in L^2(\mathcal{Q})$. Define the linear functional $\Phi : L^2(\lambda) \mapsto R$ via the following:

$$\Phi(g) = \int_\Omega g(x)d\mathcal{Q}(x)$$

By the triangle inequality and the Cauchy-Schwartz inequality we have:

$$|\Phi(g)| = |\int_\Omega g(x)d\mathcal{Q}(x)| \leq \int_\Omega |g(x)|d\mathcal{Q}(x) \leq \sqrt{\lambda(\Omega)}\|g\|_{L^2(\lambda)}.$$

Thus, by the Riesz representation theorem there exists a unique $h \in L^2(\lambda)$ such that for all $g \in L^2(\lambda)$ we have:

$$\Phi(g) = \langle h, g\rangle_{L^2(\lambda)}.$$

The last relation implies that for all $g \in L^2(\lambda)$:

$$\int_\Omega gd\mathcal{Q} = \int_\Omega hgd\lambda. \tag{4.1}$$

In particular, letting $g = I_A$ for an arbitrary $A \in \mathcal{F}$ we obtain:

$$\mathcal{Q}(A) = \int_A d\mathcal{Q} = \int_A hd\lambda.$$

Since by definition, $\lambda(A) \geq \mathcal{Q}(A)$ for all $A \in \mathcal{F}$, using the previous relation we have:

$$\mathcal{Q}(A) = \int_A d\mathcal{Q} = \int_A hd\lambda \leq \lambda(A). \tag{4.2}$$

Next, we define the following sets:

$$D_+ = \{x \in \Omega \mid h(x) > 1\},$$

$$D_- = \{x \in \Omega \mid h(x) < 0\},$$

$$D_0 = \{x \in \Omega \mid h(x) = 1\}.$$

Since h is a measurable function, the sets defined above are measurable. In particular, assuming that $\lambda(D_+) > 0$ or $\lambda(D_-) > 0$, from relation (4.2) we have:

$$0 < \lambda(D_+) < \int_{D_+} h d\lambda \leq \lambda(D_+)$$

or

$$0 \leq \mathcal{Q}(D_-) = \int_{D_-} h d\lambda < 0.$$

Hence, $\lambda(D_+) = \lambda(D_-) = 0$, which implies $\mathcal{P}(D_+) = \mathcal{P}(D_-) = 0$ and $\mathcal{Q}(D_+) = \mathcal{Q}(D_-) = 0$. Using the definition of λ, we rewrite relation (4.1) as:

$$\int_\Omega g(1-h) d\mathcal{Q} = \int_\Omega g h d\mathcal{P}. \tag{4.3}$$

Again letting, $g = I_A$ for an arbitrary $A \in \mathcal{F}$ we obtain:

$$\int_A (1-h) d\mathcal{Q} = \int_A h d\mathcal{P}.$$

Letting $A = D_0$ we thus obtain $\mathcal{P}(D_0) = 0$ and from absolute continuity $\mathcal{Q}(D_0) = 0$. Hence we have, $0 \leq h < 1$ $\mathcal{P}$-a.e. Now let $g = \frac{1}{1-h} I_A$. Using (4.3), we finally obtain:

$$\mathcal{Q}(A) = \int_A \frac{h}{1-h} d\mathcal{P}.$$

Setting $L = \frac{h}{1-h}$ proves the claim, since L is $\mathcal{F}$ measurable, $\mathcal{P}$ almost surely positive and is uniquely determined by the Riesz representation. $\square$

If $\mathcal{Q}$ and $\mathcal{P}$ are probability measures, $\mathcal{Q}$ absolutely continuous with respect to $\mathcal{P}$ then the Radon-Nikodym derivative $\frac{d\mathcal{Q}}{d\mathcal{P}} : \Omega \mapsto \mathbf{R}_+$ satisfies:

$$\int_\Omega \frac{d\mathcal{Q}}{d\mathcal{P}} d\mathcal{P} = \mathcal{Q}(\Omega) = 1.$$

Thus in terms of expected values for any random variable $X \in L^1(\mathcal{Q})$ we have:

$$\mathbb{E}^{\mathcal{Q}}(X) = \mathbb{E}^{\mathcal{P}}(\frac{d\mathcal{Q}}{d\mathcal{P}} X), \tag{4.4}$$

where $\mathbb{E}^{\mathcal{Q}}$ and $\mathbb{E}^{\mathcal{P}}$ denote the expected values with respect to $\mathcal{Q}$ and $\mathcal{P}$ respectively.

4.2 Bayes' theorem for conditional expectations

Definition 4.6 (Conditional expectation)**.** Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability space and X a random variable in $L^1(\mathcal{P})$. Let $\mathcal{G}$ be a sigma-algebra on Ω such that $\mathcal{G} \subseteq \mathcal{F}$. Let Z be a random variable satisfying the following properties:

(i) Z is $\mathcal{G}$ measurable.

(ii) For every $G \in \mathcal{G}$ it holds that $\int_G Z(x) d\mathcal{P}(x) = \int_G X(x) d\mathcal{P}(x)$.

We call Z the conditional expectation of X given the sigma-algebra $\mathcal{G}$ and we denote $Z = \mathbb{E}(X|\mathcal{G})$.

As a simple application of the Radon-Nikodym theorem we prove the existence of a conditional expectation.

Lemma 4.7. *Given the setting of Definition 4.6 there exists a random variable Z satisfying conditions i and ii. Furthermore, Z is $\mathcal{P}$-a.s. unique.*

Proof. Since $\mathcal{G} \subseteq \mathcal{F}$, $\mathcal{P}$ is a measure on $\mathcal{G}$. Define a new measure on $(\Omega, \mathcal{G})$ as follows:

$$\nu(G) = \int_G X(x) d\mathcal{P}(x),$$

for all $G \in \mathcal{G}$. Then trivially $\nu << \mathcal{P}$. Applying the Radon-Nikodym theorem, 4.5, we see that the Radon-Nikodym derivative $Z = \frac{d\nu}{d\mathcal{P}}$ is $\mathcal{G}$ measurable $\mathcal{P}$-a.s. unique and satisfies $\nu(G) = \int_G Z d\mathcal{P}$ for all $G \in \mathcal{G}$ which proves the claim. $\square$

Lemma 4.8. *Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability space, $\mathcal{Q}$ a probability measure on $(\Omega, \mathcal{F})$ such that $\mathcal{Q} << \mathcal{P}$. Let furthermore $\mathcal{G}$ be a sigma algebra on Ω such that $\mathcal{G} \subseteq \mathcal{F}$. Then the Radon-Nikodym derivatives with respect to $\mathcal{G}$, $\frac{d\mathcal{Q}}{d\mathcal{P}}\big|_{\mathcal{G}}$ and $\mathcal{F}$, $\frac{d\mathcal{Q}}{d\mathcal{P}}\big|_{\mathcal{F}}$ are related by $\frac{d\mathcal{Q}}{d\mathcal{P}}\big|_{\mathcal{G}} = \mathbb{E}^{\mathcal{P}}\left(\frac{d\mathcal{Q}}{d\mathcal{P}}\big|_{\mathcal{F}}|\mathcal{G}\right)$*

Proof. We begin by noting that by the definition of conditional expectation $\mathbb{E}^{\mathcal{P}}(\frac{d\mathcal{Q}}{d\mathcal{P}}\big|_{\mathcal{F}}|\mathcal{G})$ is $\mathcal{G}$ measurable. Now for any set $G \in \mathcal{G}$ it holds that $G \in \mathcal{F}$. So from the Radon-Nikodym theorem and the definition of conditional expectation we have:

$$\mathcal{Q}(G) = \int_G d\mathcal{Q} = \int_G \frac{d\mathcal{Q}}{d\mathcal{P}}\bigg|_{\mathcal{F}} d\mathcal{P} = \int_G \mathbb{E}^{\mathcal{P}} \left(\frac{d\mathcal{Q}}{d\mathcal{P}}|_{\mathcal{F}} \bigg| \mathcal{G} \right) d\mathcal{P}.$$

The result now follows from the uniqueness of the Radon-Nikodym derivative. $\square$

Theorem 4.9 (Bayes' theorem)**.** *Assume X is a random variable on a probability space $(\Omega, \mathcal{F}, \mathcal{P})$ and let $\mathcal{Q}$ be an absolutely continuous probability measure with respect to $\mathcal{P}$ on $(\Omega, \mathcal{F})$ with Radon-Nikodym derivative $\frac{d\mathcal{Q}}{d\mathcal{P}}$ on $\mathcal{F}$. Assume that $X \in L^1(\Omega, \mathcal{F}, \mathcal{Q})$ and let $\mathcal{G}$ be a sigma algebra such that $\mathcal{G} \subseteq \mathcal{F}$. Then the following holds $\mathcal{Q} - a.s.$*

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

Proof. Denote $A = \{\omega \in \Omega | \mathbb{E}^{\mathcal{P}}(\frac{d\mathcal{Q}}{d\mathcal{P}}|\mathcal{G})(\omega) = 0\}$. From the $\mathcal{G}$-measurablility of the conditional expectation we notice that the set A is $\mathcal{G}$-measurable. We thus have:

$$\mathcal{Q}(A) = \int_A d\mathcal{Q} = \int_A \frac{d\mathcal{Q}}{d\mathcal{P}} d\mathcal{P} = \int_A \mathbb{E}^{\mathcal{P}}(\frac{d\mathcal{Q}}{d\mathcal{P}}|\mathcal{G}) d\mathcal{P} = 0$$

Thus $\mathbb{E}^{\mathcal{P}}(\frac{d\mathcal{Q}}{d\mathcal{P}}|\mathcal{G}) \neq 0$ $\mathcal{Q}$-a.e. It is now enough to prove the following holds $\mathcal{P}$-a.e. on $\mathcal{G}$:

$$\mathbb{E}^{\mathcal{Q}}(X|\mathcal{G})\mathbb{E}^{\mathcal{P}}(\frac{d\mathcal{Q}}{d\mathcal{P}}|\mathcal{G}) = \mathbb{E}^{\mathcal{P}}(\frac{d\mathcal{Q}}{d\mathcal{P}}X|\mathcal{G}). \tag{4.5}$$

If (4.5) holds the claim of the theorem follows from the absolute continuity of $\mathcal{Q}$ with respect to $\mathcal{P}$ and the fact that $\mathcal{Q}(A) = 0$. Let $G \in \mathcal{G}$. Then using Lemma 4.8 and the Radon-Nikodym theorem we have:

$$\int_G \mathbb{E}^{\mathcal{Q}}(X|\mathcal{G})\mathbb{E}^{\mathcal{P}}(\frac{d\mathcal{Q}}{d\mathcal{P}}|\mathcal{G})d\mathcal{P} = \int_G \frac{d\mathcal{Q}}{d\mathcal{P}}\bigg|_{\mathcal{G}} \mathbb{E}^{\mathcal{Q}}(X|\mathcal{G})d\mathcal{P} = \int_G \mathbb{E}^{\mathcal{Q}}(X|\mathcal{G})d\mathcal{Q} = \int_G X d\mathcal{Q}.$$

Furthermore:

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

Since the integrals with respect to $\mathcal{P}$ are equal, (4.5) holds $\mathcal{P}$-a.s. on $\mathcal{G}$ which proves the claim. □

4.3 The general setting

Definition 4.10 (Filtration)**.** Given a measurable space $(\Omega, \mathcal{F})$, a filtration is a sequence of sigma-algebras, $(\mathcal{F}_t)_{t\geq 0}$, where, $\mathcal{F}_t \subseteq \mathcal{F}$ for all $t \geq 0$ and $\mathcal{F}_s \subseteq \mathcal{F}_t$ whenever $s \leq t$.

Definition 4.11 (Filtered probability space)**.** A filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathcal{P})$ is a probability space equipped with the filtration $(\mathcal{F}_t)_{t\geq 0}$ of its sigma-algebra $\mathcal{F}$. We will assume that the probability space is complete and right continuous i.e. $\mathcal{F}_0$ contains all the $\mathcal{P}$-null sets and $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$ for all $t \geq 0$ respectively.

Definition 4.12 (Stochastic basis)**.** A stochastic basis is a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathcal{P})$ carrying a d-dimensional Brownian motion $W = (W_1, W_2, \ldots W_d)^T$

which is adapted to the filtration (see definition below). We will call $\mathcal{P}$ the objective probability measure of the stochastic basis.

Definition 4.13 (Adapted process). A stochastic process $X(\omega, t)$ is called adapted if for all $t \geq 0$ the mapping $\omega \mapsto X(\omega, t)$ is $\mathcal{F}_t$ measurable.

Definition 4.14 (Martingale process). A stochastic process $X = X(\omega, t)$ is an $(\mathcal{F}_t)_{t\geq 0}$ martingale if the following conditions are satisfied:

(i) X is $(\mathcal{F}_t)_{t\geq 0}$ adapted.

(ii) For every $t \geq 0$, $X(t) \in L^1(\Omega, \mathcal{F}, \mathcal{P})$.

(iii) For every s and t such that $0 \leq s \leq t$ it holds that $X(s) = \mathbb{E}^{\mathcal{P}}(X(t)|\mathcal{F}_s)$.

Let $\mathcal{Q}$ be an equivalent probability measure to $\mathcal{P}$ on $(\Omega, \mathcal{F})$. We denote by $D(t) = \frac{d\mathcal{Q}}{d\mathcal{P}}\big|_{\mathcal{F}_t}$ the Radon-Nikodym density process. We have the following simple observations.

Lemma 4.15. *The Radon-Nikodym density process $D(t)$ is an $(\mathcal{F}_t)_{t\geq 0}$ martingale.*

Proof.

(i) By the Radon-Nikodym theorem, 4.5, $D(t)$ is $(\mathcal{F}_t)_{t\geq 0}$ adapted.

(ii) By the Radon-Nikodym theorem $D(t) \in L^1(\Omega, \mathcal{F}, \mathcal{P})$.

(iii) Follows from lemma 4.8 since $s \leq t$ implies $\mathcal{F}_s \subseteq \mathcal{F}_t$.

□

Assuming that $\mathcal{F}_0$ is the trivial sigma algebra i.e. $\mathcal{F}_0 = \{\Omega, \emptyset\}$ and using Lemma 4.8 we have the following observation:

$$D(0) = \frac{d\mathcal{Q}}{d\mathcal{P}}\Big|_{\mathcal{F}_0} = \mathbb{E}^{\mathcal{P}}\left(\frac{d\mathcal{Q}}{d\mathcal{P}}|\mathcal{F}_0\right) = \mathbb{E}^{\mathcal{P}}\left(\frac{d\mathcal{Q}}{d\mathcal{P}}\right) = 1. \tag{4.6}$$

Lemma 4.16. *Let $\mathcal{Q}$ be an equivalent probability measure to $\mathcal{P}$ on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathcal{P})$. An adapted process X is a $\mathcal{Q}$ martingale if and only if the process DX is a $\mathcal{P}$ martingale.*

Proof. Since X is adapted, DX is adapted. Furthermore, from equation (4.4), we see that $X \in L^1(\Omega, \mathcal{F}, \mathcal{Q})$ if and only if $DX \in L^1(\Omega, \mathcal{F}, \mathcal{P})$. From Bayes' theorem 4.9 we have:

$$\mathbb{E}^{\mathcal{Q}}(X(t)|\mathcal{F}_s) = \frac{\mathbb{E}^{\mathcal{P}}(D(t)X(t)|\mathcal{F}_s)}{\mathbb{E}^{\mathcal{P}}(D(t)|\mathcal{F}_s)}.$$

Using the martingale property of the Radon-Nikodym process we obtain:

$$\mathbb{E}^{\mathcal{Q}}(X(t)|\mathcal{F}_s) = \frac{\mathbb{E}^{\mathcal{P}}(D(t)X(t)|\mathcal{F}_s)}{D(s)},$$

which proves the claim. □

Definition 4.17 (Progressive process)**.** We call a stochastic process $X(\omega, t)$ progressive if the mapping $(\omega, s) \mapsto X(\omega, s)$, where $s \in [0.t]$ is $\mathcal{F}_t \otimes \mathbf{B}[0, t]$ measurable for all $t \geq 0$. Here, $\mathbf{B}[0, t]$ denotes the Borel sigma algebra on the interval $[0, t]$. We denote by $\mathcal{A}_T$ the sigma algebra generated by all progressive processes on $\Omega \times [0, T]$. We denote by $\mathcal{A}$ the sigma algebra generated by all progressive processes on $\Omega \times [0, \infty)$.

Definition 4.18 ($\mathcal{L}^2(W)$)**.** We denote by $\mathcal{L}^2(W)$, the set of all $\mathbf{R}^d$ valued progressive processes $h = (h_1, h_2, \ldots, h_d)$ that satisfy $\mathbb{E}(\int_0^\infty \|h(s)\|^2 ds) < \infty$.

Definition 4.19 ($\mathcal{L}(W)$)**.** We denote by $\mathcal{L}(W)$, the set of all $\mathbf{R}^d$ valued progressive processes $h = (h_1, h_2, \ldots, h_d)$ that satisfy $\int_0^t \|h(s)\|^2 ds < \infty$ for all $t > 0$.

For every $h \in \mathcal{L}(W)$ we are going to define the stochastic integral with respect to W as follows:

$$(h \bullet W)_t = \sum_{j=1}^{d} \int_0^t h_j(s) dW_j(s), \tag{4.7}$$

where the one-dimensional stochastic integrals are defined in the usual sense (see for example chapter 6 in [20]). If the t subscript is omitted the integration is assumed to be from 0 to ∞.

Lemma 4.20. *If $h \in \mathcal{L}^2(W)$ then $(h \bullet W)$ is a martingale and the following property called Itó isometry holds $\mathbb{E}((h \bullet W)^2) = \mathbb{E}(\int_0^\infty \|h(s)\|^2 ds)$.*

Proof. For the case when h is simple see Proposition 4.4. in [11]. The general claim is proven in Propositions 2.7 and 2.10 in [21] chapter IV. □

4.3.1 One-dimensional Itó process

An Itó process $X(\omega, t)$ is a process given by the following form:

$$X(t) = X(0) + \int_0^t a(s)ds + (\rho \bullet W)_t, \tag{4.8}$$

here $\rho \in \mathcal{L}(W)$ and a is a progressive process satisfying $\int_0^t |a(s)| ds$ for all $t > 0$. $X(0)$ is an $\mathcal{F}_0$ (in particular we will choose $\mathcal{F}_0 = \{\Omega, \emptyset\}$) measurable scalar random variable, independent of the Brownian motion. We often write (4.8) in differential notation as follows:

$$dX = adt + \rho dW$$

For an Itó process satisfying (4.8) we define the following sets:

$$\mathcal{L}(X) = \{h \text{ progressive} \mid \int_0^t |h(s)a(s)|^2 ds < \infty \text{ for all } t > 0, h\rho \in \mathcal{L}(W)\},$$

$$\mathcal{L}^2(X) = \{h \text{ progressive} \,|\mathbb{E}(\int_0^t |h(s)a(s)|^2 ds) < \infty, h\rho \in \mathcal{L}^2(W)\}.$$

If $h \in \mathcal{L}(X)$ we can define the stochastic integral of h with respect to X as follows:

$$\int_0^t h(s)dX(s) = \int_0^t h(s)a(s)ds + (\rho h \bullet W)_t, \tag{4.9}$$

Let Y is another Itó process given by:

$$Y(t) = Y(0) + \int_0^t b(s)ds + (\sigma \bullet W)_t.$$

We define the quadratic covariation process of X and Y as follows:

$$\langle X, Y\rangle_t = \int_0^t \rho(s) \cdot \sigma(s)ds. \tag{4.10}$$

Since ρ and σ are real valued progressive processes the quadratic variation is symmetric and bilinear.

Lemma 4.21. *Let X, Y and Z be Itó processes as defined by (4.8), and let $a \in \mathbf{R}$ and $b \in \mathbf{R}$ be two constants. Then, the quadratic covariation process satisfies*

(i) $\langle X, Y\rangle_t = \langle Y, X\rangle_t$.

(ii) $\langle aX + bY, Z\rangle_t = a\langle X, Z\rangle_t + b\langle Y, Z\rangle_t$.

Proof. Follows form the definition of quadratic variation (4.10) and the properties of the inner product in $\mathbf{R}^d$. □

In differential notation we write:

$$d\langle X, Y\rangle_t = \rho(t) \cdot \sigma(t)dt.$$

4.3.2 The multidimensional Itó process and Itó formula

Definition 4.22 (n-dimensional Itó process)**.** A process $X = (X_1, X_2, \dots X_n)$ is called an n-dimensional Itó process if each component X_i is an Itó process.

We denote by $\mathcal{L}^2(X)$ ($\mathcal{L}(X)$) the set of progressive processes $h = (h_1, h_2, \dots h_n)$ such that $h_i \in \mathcal{L}^2(X_i)$ ($h_i \in \mathcal{L}(X_i)$). We define the stochastic integral of h with respect to X coordinate-wise:

$$(h \bullet X)_t = \sum_{i=1}^{n} \int_0^t h_i(s) dX_i(s),$$

where the coordinate-wise integrals are defined as in (4.9).

Theorem 4.23 (Itó formula)**.** *Let $f\colon \mathbf{R}^n \times \mathbf{R}_+ \to \mathbf{R}$ be a function such that $f \in \mathcal{C}^{2,1}(\mathbf{R}^n \times \mathbf{R}_+)$. If X is an n-dimensional Itó process then, $f(X(t), t)$ is an Itó process and it holds:*

$$\begin{aligned} f(X(t), t) =& f(X(0), 0) + \int_0^t \frac{\partial f}{\partial s}(X(s), s) ds + \sum_{i=1}^{n} \int_0^t \frac{\partial f}{\partial x_i}(X(s), s) dX_i(s) \\ &+ \frac{1}{2} \sum_{i,j=1}^{n} \int_0^t \frac{\partial^2 f}{\partial x_j \partial x_i}(X(s), s) d\langle X_i, X_j \rangle_s. \end{aligned}$$

Proof. Proven in Theorem 3.3 [21] in chapter IV. □

Corollary 4.24 (Integration by parts formula)**.** *If X and Y are two Itó processes then we have the following integration by parts formula*

$$X(t)Y(t) = X(0)Y(0) + \int_0^t X(s) dY(s) + \int_0^t Y(s) dX(s) + \langle X, Y \rangle_t. \tag{4.11}$$

Proof. Follows by applying Itó formula to $f(x, y) = xy$ □

We often write Itó formula more compactly in differential notation as follows:

$$df = \frac{\partial f}{\partial t} dt + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dX_i + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i} d\langle X_i, X_j \rangle_t. \tag{4.12}$$

In particular, if the dynamics of the one dimensional Itó process X is given by:

$$dX = a dt + \rho dW,$$

where W is a one dimensional Brownian motion and $f = f(x,t) \in \mathcal{C}^{2,1}(\mathbf{R} \times \mathbf{R}_+)$, then (4.12) reduces to the well known formula:

$$df = \left(\frac{\partial f}{\partial t} + \rho^2 \frac{1}{2} \frac{\partial^2 f}{\partial x^2} + a \frac{\partial f}{\partial x} \right) dt + \rho \frac{\partial f}{\partial x} dW$$

4.3.3 Stochastic exponential

Definition 4.25 (Stochastic exponential)**.** For an Itó process X we define the stochastic exponential of X, denoted by $\mathcal{E}_t(X)$ as $\mathcal{E}_t(X) \equiv e^{X(t) - \frac{1}{2}\langle X,X \rangle_t}$.

We have the following lemma which will be used on numerous occasions.

Lemma 4.26. *Let X and Y be two Itó processes. Then, the stochastic exponential satisfies the following properties:*

(i) *$\mathcal{E}$ is a positive Itó process and the unique solution to the stochastic differential equation $dU = UdX$, with $U(0) = e^{X(0)}$.*

(ii) *$\mathcal{E}(X)$ is a continuous martingale if X is a martingale.*

(iii) *$\mathcal{E}(X)\mathcal{E}(Y) = \mathcal{E}(X+Y)e^{\langle X,Y \rangle}$.*

(iv) *$\mathcal{E}(0) = 1$.*

(v) *$\mathcal{E}^{-1}(X) \equiv \frac{1}{\mathcal{E}(X)} = \mathcal{E}(-X)e^{\langle X,X \rangle}$.*

Proof.

(i) Positivity follows from the properties of the exponential function and the fact that $\mathcal{E}(X)$ is an Itó process follows from Theorem 4.23. Applying Itó formula in differential notation, (4.12), to $\mathcal{E}_t(X) = e^{X(t) - \frac{1}{2}\langle X,X \rangle_t}$ we have:

$$d\mathcal{E} = -\frac{1}{2}\mathcal{E} d\langle X,X \rangle + \mathcal{E} dX + \frac{1}{2}\mathcal{E} d\langle X,X \rangle. \tag{4.13}$$

Thus, $\mathcal{E}(X)$ solves the differential equation. Furthermore, from the definition of the stochastic exponential $\mathcal{E}_0(X) = e^{X(0)}$. Thus $\mathcal{E}(X)$ solves the differential equation. Applying Itó formula to $e^{-X(t) + \frac{1}{2}\langle X,X \rangle_t}$ similarly as in (4.13) we see that $d\frac{1}{\mathcal{E}} = -\frac{1}{\mathcal{E}} dX$. Now let V be another solution to the stochastic differential equation, namely $dV = VdX$, with $V(0) = e^{X(0)}$. Applying (4.11) we have

$$d\frac{V}{\mathcal{E}} = -\frac{V}{\mathcal{E}} dX + \frac{V}{\mathcal{E}} dX = 0$$

Now by uniqueness of representation of Itó processes we have $V = \mathcal{E}$ which proves the claim.

(ii) Proven in Lemma 4.2 in [1].

(iii) Explicitly calculating and using Lemma 4.21 we have:

$$\mathcal{E}(X)\mathcal{E}(Y) = e^{X+Y-\frac{1}{2}(\langle X,X\rangle+\langle Y,Y\rangle+2\langle X,Y\rangle)+\langle X,Y\rangle} = \mathcal{E}(X+Y)e^{\langle X,Y\rangle}.$$

(iv) $\mathcal{E}(0) = e^0 = 1$.

(v) The result follows by applying the result in *iii* with $Y = -X$ and using *iv*.

□

As seen from Lemma 4.26 the stochastic exponential is a prototype of a positive Itó process. As we will see in the next chapter, it is often used in financial applications in modeling positive asset prices.

Suppose $\gamma \in \mathcal{L}(W)$. In the next theorem we adopt the following notation:

$$\mathcal{E}_\infty(\gamma \bullet W) \equiv e^{(\gamma\bullet W)_T - \frac{1}{2}\int_0^T \gamma(s)\cdot\gamma(s)ds},$$

where T is some finite time horizon $T \geq t$ after which the stochastic processes defined above stop.

4.3.4 Girsanov's change of Measure theorem

Theorem 4.27 (Girsanov's Change of Measure theorem)**.** *Let $\gamma \in \mathcal{L}(W)$ be such that $\mathcal{E}(\gamma \bullet W)$ is a uniformly integrable martingale with $\mathcal{E}_\infty(\gamma \bullet W) > 0$. Then, the Radon-Nikodym derivative $\frac{d\mathcal{Q}}{d\mathcal{P}} = \mathcal{E}_\infty(\gamma \bullet W)$ defines an equivalent probability measure $\mathcal{Q}$ to $\mathcal{P}$ and the process $\bar{W}(t) = W(t) - \int_0^t \gamma(s)ds$ is a $\mathcal{Q}$ Brownian motion.*

Proof. The claim is proven in Theorem 1.12 in chapter VIII in [21]. □

Corollary 4.28. *The $\mathcal{F}_t$ conditional Radon-Nikodym derivative under the conditions of the Girsanov's change of measure theorem is given by $\frac{d\mathcal{Q}}{d\mathcal{P}}\big|_{\mathcal{F}_t} = \mathcal{E}_t(\gamma \bullet W)$ for all $t \geq 0$.*

Proof. Using Lemma 4.8 we have:

$$\frac{d\mathcal{Q}}{d\mathcal{P}}\bigg|_{\mathcal{F}_t} = \mathbb{E}^{\mathcal{P}}(\frac{d\mathcal{Q}}{d\mathcal{P}}|\mathcal{F}_t).$$

Now applying Girsanov's theorem we obtain:

$$\left.\frac{d\mathcal{Q}}{d\mathcal{P}}\right|_{\mathcal{F}_t} = \mathbb{E}^{\mathcal{P}}(\mathcal{E}_\infty(\gamma \bullet W)|\mathcal{F}_t).$$

Since the stochastic integral $(\gamma \bullet W)_t$ is a martingale by conclusion *ii* in Lemma 4.26 we obtain:

$$\left.\frac{d\mathcal{Q}}{d\mathcal{P}}\right|_{\mathcal{F}_t} = \mathbb{E}^{\mathcal{P}}(\mathcal{E}_\infty(\gamma \bullet W)|\mathcal{F}_t) = \mathcal{E}_t(\gamma \bullet W).$$

□

A sufficient condition for the uniform integrability of the stochastic exponential in Girsanov's change of measure theorem is provided by the following result.

Theorem 4.29 (Novakov's condition)**.** *If* $\mathbb{E}\left(e^{\frac{1}{2}\int_0^\infty \gamma(s)\cdot\gamma(s)ds}\right) < \infty$ *then* $\mathcal{E}(\gamma \bullet W)$ *is a uniformly integrable martingale with* $\mathcal{E}_\infty(\gamma \bullet W) > 0$.

Proof. The claim follows from Proposition 1.15 in [21] chapter VIII and Proposition 1.26 in [21] chapter IV. □

4.3.5 Martingale representation theorem

Theorem 4.30 (Martingale representation theorem)**.** *Assume that the filtration* $(\mathcal{F}_t)_{t\geq 0}$ *is generated by the Brownian motion* W*. Then every* $\mathcal{P}$ *martingale* $M(t)$ *has a continuous modification and there exists* $\psi \in \mathcal{L}^2(W)$ *such that,* $M(t) = M(0) + \int_0^t \psi(s)dW(s)$.

Proof. The claim is proven in Theorem 3.5 in [21] chapter V. □

Corollary 4.31. *Under the assumptions of theorem 4.30 for every* $\mathcal{Q}$ *probability measure equivalent to* $\mathcal{P}$ *there exists a* $\gamma \in \mathcal{L}(W)$ *that satisfies the conditions in Girsanov's theorem, thus* $D(t) = \frac{d\mathcal{Q}}{d\mathcal{P}}\big|_{\mathcal{F}_t} = \mathcal{E}_t(\gamma \bullet W)$ *for all* $t \geq 0$.

Proof. From Lemma 4.15 the Radon-Nikodym density process is $\mathcal{P}$ martingale. Furthermore, using a similar argument as in the proof of Bayes' theorem 4.9, we conclude $D(t) > 0$ $\mathcal{Q}$-a.s and since $\mathcal{Q}$ and $\mathcal{P}$ are equivalent we also have $D(t) > 0$ $\mathcal{P}$ - a.s. Now, applying the Martingale representation theorem on $D(t)$ there exists a $\psi \in \mathcal{L}^2(W)$ such that $D(t) = D(0) + \int_0^t \psi(s)dW(s)$. Now using (4.6) we can write:

$$D(t) = D(0) + \int_0^t \psi(s)dW(s) = 1 + \int_0^t \psi(s)dW(s)$$

Multiplying and dividing by $D(s)$ inside the integral we obtain:

$$D(t) = 1 + \int_0^t D(s)\frac{\psi(s)}{D(s)}dW(s).$$

Define $\gamma(t) = \frac{\psi(t)}{D(t)}$. Now by Radon-Nikodym theorem and the fact that $\psi \in \mathcal{L}^2(W)$ the Novakov condition is satisfied. Thus, we have:

$$D(t) = 1 + \int_0^t D(s)\gamma(s)dW(s).$$

Conclusions i and iv in Lemma 4.26 complete the proof. □

Suppose now that $X(t) = X(0)+\int_0^t a(s)ds+\int_0^t \sigma(s)dW(s)$ is an Itó process. From lemma 4.20 we see that if $\sigma \in \mathcal{L}^2(W)$ then the integral term $\int_0^t \sigma(s)dW(s)$ is a martingale. Conversely, given that the filtration is generated by the Brownian motion using the martingale representation theorem we may conclude that any Itó process which is a martingale can be written in an integral form $X(t) = X(0) + \int_0^t \psi(s)dW(s)$ for some $\psi \in \mathcal{L}^2(W)$. We succinctly state this result in the next lemma.

Lemma 4.32. *Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathcal{P})$ be a probability space such that the filtration $(\mathcal{F}_t)_{t\geq 0}$ is generated by the Brownian motion W. Let $X(t) = X(0) + \int_0^t a(s)ds + \int_0^t \sigma(s)dW(s)$ be an Ito process, such that $\sigma \in \mathcal{L}^2(W)$. Then X is a $\mathcal{P}$ martingale if and only if $a = 0$ $\mathcal{P} \otimes dt$ a.s.*

Proof. Follows directly from Lemma 4.20 and the martingale representation theorem. □

Lemma 4.33 (Martingale representation under the Girsanov's change of measure)**.** *Let $\mathcal{Q}$ be an equivalent measure to $\mathcal{P}$ such that the conditions of Girsanov's change of measure theorem hold. Suppose furthermore that the filtration $(\mathcal{F}_t)_{t\geq 0}$ is generated by the Brownian motion W. Denote by $\bar{W}(t) = W(t) - \int_0^t \gamma(s)ds$ the $\mathcal{Q}$ Brownian motion and let $Y(t)$ be a $\mathcal{Q}$ martingale. Then there exists an adapted process $\bar{\psi}(t)$ such that $Y(t) = Y(0) + \int_0^t \bar{\psi}(s)d\bar{W}(s)$.*

Proof. By Corollary 4.28, the Radon-Nikodym derivative for the $\mathcal{Q}$ Girsanov measure is given by $D(t) = \frac{d\mathcal{Q}}{d\mathcal{P}}\big|_{\mathcal{F}_t} = \mathcal{E}_t(\gamma \bullet W)$. Now since $Y(t)$ is a $\mathcal{Q}$ martingale by Lemma 4.16 $Y(t)D(t)$ is a $\mathcal{P}$ martingale. By the Martingale Representation theorem $Y(t)D(t)$ has a continuous modification and there exists a $\psi \in \mathcal{L}^2(W)$ such that:

$$Y(t)D(t) = Y(0) + \int_0^t \psi(s)dW(s), \tag{4.14}$$

where we have used the fact that $D(0) = 1$. Now applying the Itó formula with $f(x) = \frac{1}{x}$, where $x = D(t)$ and using the fact that $D(t) = \mathcal{E}_t(\gamma \bullet W)$ we obtain:

$$d(\frac{1}{D(t)}) = \frac{1}{D(t)}\gamma(t)dW(t) + \frac{1}{D(t)}\gamma(t) \cdot \gamma(t)dt. \tag{4.15}$$

Since $D(t) > 0$ $\mathcal{P}$ a.s. we may write $dY(t) = d(Y(t)D(t)\frac{1}{D(t)})$. Now using (4.11) we have:

$$dY(t) = d(Y(t)D(t)\frac{1}{D(t)}) = Y(t)D(t)d(\frac{1}{D(t)}) + \frac{1}{D(t)}d(Y(t)D(t)) + d\langle\frac{1}{D}, YD\rangle_t. \tag{4.16}$$

Inserting (4.14) and (4.15) into (4.16) and noticing that $d\langle\frac{1}{D}, YD\rangle_t = \frac{1}{D(t)}\gamma(t)\psi(t)dt$ we obtain:

$$dY(t) = (Y(t)\gamma(t) \cdot \gamma(t) - \frac{1}{D(t)}\psi(t) \cdot \gamma(t))dt + (\frac{1}{D(t)}\psi(t) - Y(t)\gamma(t))dW(t). \tag{4.17}$$

Using (4.17) and the fact that $dW(t) = d\bar{W}(t) + \gamma(t)dt$ we obtain:

$$dY(t) = (\frac{1}{D(t)}\psi(t) - Y(t)\gamma(t))d\bar{W}(t).$$

Setting $\bar{\psi}(t) = \frac{1}{D(t)}\psi(t) - Y(t)\gamma(t)$ proves the claim. □

4.3.6 Fubini's theorem for stochastic integrals

We finish this background chapter on stochastic analysis by stating the stochastic version of Fubini's theorem and an important corollary.

Theorem 4.34 (Fubini theorem for stochastic integrals)**.** *Consider an $\mathbf{R}^d$ valued stochastic process $\phi(\omega, s, t)$ with two indices such that $0 \le t, s \le T$ for some $T > 0$ satisfying the following assumptions:*

(i) ϕ *is* $\mathcal{A}_T \otimes \mathcal{B}[0, t]$ *measurable.*

(ii) $\sup_{t,s} \|\phi(s, t)\| < \infty$ *for all* $\omega \in \Omega$.

Then $\lambda(t) = \int_0^T \phi(t, s)ds \in \mathcal{L}(W)$ *and there exists a* $\mathcal{F}_T \otimes \mathcal{B}[0, t]$ *measurable modification* $\psi(s)$ *of* $\int_0^T \phi(t, s)dW(t)$ *such that* $\int_0^T \psi^2(s)ds < \infty$. *Moreover:*

$$\int_0^T \left(\int_0^T \phi(t, s)dW(t)\right) ds = \int_0^T \left(\int_0^T \phi(t, s)ds\right) dW(t)$$

Proof. The claim follows from Theorem 6.2 and Theorem 6.3 in [1]. □

Corollary 4.35. *Let $\phi = \phi(\omega, s, t)$ be a stochastic process with two indices satisfying the assumptions in Fubini's theorem for stochastic integrals. Then the process*

$$u(s) = \int_0^s \phi(t,s)dW(t),$$

where $s \in [0,T]$, has a progressive modification $\pi(s)$, such that $\int_0^T \pi^2(s)ds < \infty$. Whenever a progressive modification exists we will use the same letter to denote the process and its progressive modification.

Proof. See proof of Corollary 6.3 in [1]. $\square$

Chapter 5

Stochastic modeling of the equities and the fixed income markets

In our further work we take the stochastic setting as defined in Section 4.3. Specifically, we take a stochastic basis $(\omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathcal{P})$ carrying a d-dimensional Brownian motion $W = (W_1, W_2, \dots W_d)^T$. Furthermore, we assume that the filtration $(\mathcal{F}_t)_{t\geq 0}$ is generated by the Brownian motion, so that the Martingale representation theorem 4.30 holds. Firstly we briefly outline the main pillars of financial modeling of the equities market proving one direction of the first fundamental theorem of asset pricing and then we turn our attention to the stochastic modeling of the fixed income market. In particular, we focus on stochastic models for the dynamics of the short rate.

5.1 The equities market

5.1.1 Financial markets

Definition 5.1 (Financial market of equities)**.** A financial market of equities is a collection of random processes $S = (B, S_1, \dots, S_n)$ with a risk-free asset, the money market account, B, as defined in 2.10, satisfying $dB = Brdt$, with $B(0) = 1$, and n risky assets S_i whose price processes satisfy $dS_i = S_i(\mu_i dt + \sigma_i dW)$ with $S_i(0) > 0$.

The short rate r, the appreciation rates μ_i and the volatility vectors $\sigma_i = (\sigma_{i1}, \sigma_{i2}, \dots, \sigma_{id})$ are assumed to form progressive processes. Furthermore, we assume $\sigma_{ij} \in \mathcal{L}^2(W)$ for all indices $i \in \{1, 2 \ldots, n\}$ and $j \in \{1, 2 \ldots, d\}$. Thus,

$$X_0(t) = \int_0^t r(s)ds \tag{5.1}$$

and

$$X_i(t) = \int_0^t \mu_i(s)ds + \int_0^t \sigma_i(s)dW(s) \tag{5.2}$$

are well defined Itó processes. Hence, from Lemma 4.26, we may rewrite the price processes of the financial assets S_i as:

$$S_i(t) = S_i(0)\mathcal{E}_t(X_i).$$

Definition 5.2 (Portfolio)**.** A portfolio $\phi = (\phi_0, \phi_1, \ldots, \phi_n)$ is any $\mathbf{R}^{n+1}$ valued progressive process.

Definition 5.3 (Value process)**.** For a given portfolio ϕ we define its value process V as $V = \phi \cdot S = \sum_{i=0}^n \phi_i S_i$.

Definition 5.4 (Self financing portfolio)**.** A portfolio ϕ is called self-financing for the price vector S if $\phi \in \mathcal{L}(S)$ and any trading gains or losses over any period of time are solely due to value changes in the underlying assets. Mathematically $dV = \phi \cdot dS = \sum_{i=0}^n \phi_i dS_i$.

5.1.2 Numeraires and the preservation of the self financing property

Often we pick an asset S_j for some $j \leq n$ and express the values of the other assets in terms of the asset S_j. We call the asset S_j a *numeraire*. In particular, we define the discounted asset price vector $\bar{S} = \frac{S}{S_j}$ and the discounted value process $\bar{V} = \frac{V}{S_j}$. Here, the division is understood component-wise i.e. $\bar{S} = (\frac{B}{S_j}, \frac{S_1}{S_j}, \ldots, 1\mathbb{1}_{i=j}, \ldots \frac{S_n}{S_j})$ and similarly for the discounted value process. Applying Itó formula in differential notation (4.12) to $f(x) = \frac{1}{x}$ we obtain:

$$d(\frac{1}{S_d}) = -\frac{1}{S_d}(\mu_d - \sigma_d \cdot \sigma_d)dt - \frac{1}{S_d}\sigma_d dW. \tag{5.3}$$

From (5.3) and the integration by parts formula given by (4.11), we obtain that the dynamics of $\bar{S}_i$ for $i \neq d$ satisfies:

$$d\bar{S}_i = \bar{S}_i(\mu_i - \mu_d + \sigma_d \cdot \sigma_d)dt + \bar{S}_i(\sigma_i - \sigma_d)dW. \tag{5.4}$$

For $i = d$, we have $d\bar{S}_d = 0$.

Lemma 5.5. *Let ϕ be a portfolio such that $\phi \in \mathcal{L}(S) \cap \mathcal{L}(\bar{S})$. Then ϕ is self financing for S if and only if it is self financing for $\bar{S}$.*

Proof. Suppose ϕ is self financing for S. Then using the properties of quadratic variation, in particular Lemma 4.21, (4.11) and the definition of a self financing portfolio we have:

$$\begin{aligned}
d\bar{V} &= d\left(\frac{V}{S_d}\right) \\
&= \frac{1}{S_d}dV + Vd\left(\frac{1}{S_d}\right) + d\langle V, \frac{1}{S_d}\rangle \\
&= \sum_{i=0}^{n}\phi_i\frac{1}{S_d}dS_i + \sum_{i=0}^{n}\phi_i S_i d(\frac{1}{S_d}) + d\langle\sum_{i=0}^{n}\phi_i S_i, \frac{1}{S_d}\rangle \\
&= \sum_{i=0}^{n}\phi_i d\bar{S}_i - \sum_{i=0}^{n}\phi_i\langle S_i, \frac{1}{S_d}\rangle + d\langle\sum_{i=0}^{n}\phi_i S_i, \frac{1}{S_d}\rangle \\
&= \sum_{i=0}^{n}\phi_i d\bar{S}_i
\end{aligned}$$

Thus ϕ is self financing for $\bar{S}$. The other direction follows similarly. □

In the future we fix the money market account B as a numeraire. Hence, from (5.4) we obtain the dynamics for $\bar{S}$ as:

$$d\bar{S}_i = \bar{S}_i(\mu_i - r)dt + \bar{S}_i\sigma_i dW. \tag{5.5}$$

5.1.3 First fundamental theorem of asset pricing

Definition 5.6 (Arbitrage portfolio)**.** An arbitrage portfolio is a self financing portfolio ϕ whose value process satisfies $V(0) = 0$, $V(T) \geq 0$ and $\mathcal{P}(V(0) > 0) > 0$ for some $T > 0$.

Definition 5.7 (Arbitrage freedom)**.** If no arbitrage portfolio exists for any $T > 0$ the market model is termed arbitrage free.

Definition 5.8 (Equivalent martingale measure)**.** An equivalent martingale measure $\mathcal{Q}$, abbreviated EMM, is a probability measure on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0})$ equivalent to $\mathcal{P}$ such that the discounted price processes $\bar{S}_i$ are $\mathcal{Q}$ martingales for all indices i.

Lemma 5.9. *Suppose $\mathcal{Q}$ is an EMM of the form given by Girsanov's theorem. Then γ satisfies $-\sigma_i \cdot \gamma = \mu_i - r$ for all $i \geq 1$ $\mathcal{Q} \otimes dt$ a.s. Conversely, if γ is a solution of $-\sigma_i \cdot \gamma = \mu_i - r$ that satisfies the conditions of Girsanov's theorem then $\mathcal{Q}$ given by Girsanov's theorem is in fact an EMM.*

Proof. Let $\mathcal{Q}$ be an EMM given by Girsanov's theorem. Then, using (5.5), we obtain the discount price dynamics in terms of the $\mathcal{Q}$ Brownian motion:

$$d\bar{S}_i = \bar{S}_i(\mu_i - r + \sigma_i \cdot \gamma)dt + \bar{S}_i \sigma_i d\bar{W}$$

Now, using Lemma 4.32 it is clear that $-\sigma_i \cdot \gamma = \mu_i - r$ for all $i \geq 1$ $\mathcal{Q} \otimes dt$ a.s. The converse statement is obvious. $\square$

The factor $-\gamma$ is often termed the market price of risk. On the right hand side of $-\sigma_i \cdot \gamma = \mu_i - r$ we have the excess return of the risky asset i, whereas on the left hand side we have a linear combination of the volatilities of asset i which intuitively specify the risk of investing in asset i multiplied by factor loading $-\gamma$. Since γ is the same for all risky assets i it specifies the cumulative market risk that the investor takes on.

Definition 5.10 (Admissible strategy)**.** A self financing strategy ϕ is admissible if there exists an EMM $\mathcal{Q}$ such that the discounted value process $\bar{V}$ is a $\mathcal{Q}$ martingale.

It turns out that the existence of an EMM rules out arbitrage. This is one direction of the first fundamental theorem of Asset pricing.

Theorem 5.11 (First fundamental theorem of asset pricing)**.** *Suppose there exists an EMM $\mathcal{Q}$. Then there exists no admissible arbitrage strategy.*

Proof. Let $\bar{V}$ be a discounted value process for some admissible strategy ϕ, with $\bar{V}(0) = 0$ and $\bar{V}(T) \geq 0$. Since V is a $\mathcal{Q}$ martingale for some EMM $\mathcal{Q}$, we have:

$$0 \leq \mathbb{E}^{\mathcal{Q}}(\bar{V}(T)) = \bar{V}(0) = 0$$

Thus $V(T) = 0$ $\mathcal{Q}$ a.s. and since $\mathcal{Q} \sim \mathcal{P}$, $V(T) = 0$ $\mathcal{P}$ a.s. which proves the claim $\square$

It turns out that the absence of arbitrage among admissible strategies in continuous time finance is not sufficient for the existence of an EMM in the most general sense. A version of the reverse statement of the first fundamental theorem of asset pricing under the conditions of no-free lunch under vanishing risk (NFLVR) (a form of asymptotic arbitrage) is proved in [22] and [23]. In this thesis we take the absence of arbitrage and the existence of an EMM as essentially equivalent assuming that the NFLVR assumptions given in [23] on the price vector S are satisfied.

5.2 The fixed income market

Unlike the financial market of equities described in the previous section, the fixed income market consists of infinitely many financial assets the *zero-coupon bonds*, whose price we denote by $P(t,T)$. Our fixed income market at each time t is going to consist of infinitely many financial assets $P(t,T)$ for each maturity date T. The zero-coupon bond price $P(t,T)$ is thus a stochastic object with two variables. For a fixed t, $P(t,T)$ becomes a function of the maturity date T, called the discount curve (see Definition 2.12). For a fixed maturity T, $P(t,T)$ is going to be a scalar stochastic process, which gives the prices of bonds with the same maturity date at different times t. To simplify the mathematical modeling of the bond market and in order to price the different fixed income derivatives, such as LIBOR rates, swap rates and interest rate futures introduced in Chapter 2. we will impose the following regularity assumptions on the zero-coupon bond prices:

1. At any time t there exists a frictionless market for all zero-coupon bonds with any maturity $T \geq t$.
2. There is no credit risk, so the the relation $P(t,t) = 1$ holds for all t.
3. For each fixed t, the bond price $P(t,T)$ is differentiable with respect to the time of maturity T.

The intuitive meaning behind the first two assumptions was outlined in Section 2.4. The third assumption is technical in nature and we use it to express the arbitrage free prices of the different fixed income derivatives as explained in Chapter 2. In Chapter 3 we have described a method of how one can derive a smooth discount curve given the prices of certain fixed income derivatives. Here we move on to the stochastic setting and we ask the following question: What is a reasonable stochastic model of the fixed income market and how do we obtain a model $P(t,T)$ as a stochastic process? Thus our goal is to model an arbitrage free family of zero-coupon bond price processes $\{P(\cdot,T); T \geq 0\}$.

While modeling the equities asset market, we imposed that their price processes satisfy the Itó dynamics of the form given by (5.1) and (5.2). The most common approach in the early fixed income literature is to exogenously specify the Itó dynamics of the risk free short rate $r(t)$ and then derive the implied dynamics of the zero-coupon bond price process $P(t,T)$. This is the approach we explore below. An alternative approach which is based in specifying the dynamics of the forward curve directly developed by Heath, Jarrow and Morton will be described thoroughly in the next chapter [12].

5.2.1 Short rate models

Since the short rate at time t in Definition 2.8 is the risk-free interest rate one can earn on a loan over the interval $[t, t+dt]$, the zero-coupon bond prices $P(t,T)$ are going to depend on the short rate dynamics over the interval $[t,T]$. Thus, one possible starting point in modeling the fixed income market is to specify the Itó dynamics of the short rate exogenously. Namely, we have the following assumptions:

1. The short rates follow an Itó process:

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

   where $b(t)$ and $\sigma(t)$ are are assumed to be progressive processes such that the Itó process for the short rate and the money market account $B(t)$ given by (2.5) are well defined.

2. There is no arbitrage on the market.

Analogously to the market for equities, we will take the existence of an equivalent martingale measure $\mathcal{Q}$ to the objective probability measure $\mathcal{P}$ such that the discounted price process $\frac{P(t,T)}{B(t)}$ is a $\mathcal{Q}$ martingale to be equivalent with the no arbitrage assumption. In particular, we are going to assume that there exists an equivalent martingale measure of the form given by Girsanov's theorem. Namely, $D(t) = \frac{d\mathcal{Q}}{d\mathcal{P}}|_{\mathcal{F}_t} = \mathcal{E}_t(\gamma \bullet W)$. We will denote by $\bar{W}(t) = W(t) - \int_0^t \gamma(s)ds$ the Girsanov transformed $\mathcal{Q}$ Brownian motion. Using the first assumption for the dynamics of the short rate it is easy to see that under the Girsanov $\mathcal{Q}$ Brownian motion the short rates satisfy:

$$dr(t) = (b(t) + \sigma(t)\cdot\gamma(t))dt + \sigma(t)d\bar{W} \tag{5.6}$$

A concrete example for a short rate model satisfying the general structure of equation (5.6) is the Ho-Lee model where the volatility is taken to be constant $\sigma(t) = \sigma$ under the equivalent probability measure $\mathcal{Q}$ [10]. Mathematically, the Ho-Lee model is given by:

$$dr(t) = b(t)dt + \sigma d\bar{W}.$$

In Chapter 6 we will see that the Ho-Lee model is just a special case of the simplest Heath-Jarrow-Morton model for the forward curve, where the volatility is assumed to be constant.

Moving back to the general setting we observe that since the discounted value process is a $\mathcal{Q}$ martingale, we have:

$$\frac{P(t,T)}{B(t)} = \mathbb{E}^{\mathcal{Q}}\left(\frac{P(T,T)}{B(T)}|\mathcal{F}_t\right). \tag{5.7}$$

Now, using the definition of the money market account, the no credit risk assumption $P(T,T)=1$ and the F_t - measurability of $B(t)$ from (5.7) we obtain:

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

Lemma 5.12. *For any $T>0$ there exists a progressive $\mathbf{R}^n$ valued process $v(t,T)$, $t\leq T$ such that*

$$dP(t,T) = P(t,T)r(t)dt + P(t,T)v(t,T)d\bar{W}(t)$$

Proof. Since we have assumed that $\frac{P(t,T)}{B(t)}$ is a $\mathcal{Q}$ martingale under $\mathcal{Q}$ which is the Girsanov equivalent measure, using Lemma 4.33 we may conclude that there exists an adapted process $\bar{\Psi}(t,T)$ for any $T>0$ such that:

$$\frac{P(t,T)}{B(t)} = P(0,T) + \int_0^t \bar{\psi}(s,T)d\bar{W}(s).$$

Now, using (4.11) and the dynamics of the money market account $dB(t) = B(t)r(t)dt$ we have:

$$\begin{aligned} dP(t,T) &= d\left(\frac{P(t,T)}{B(t)}B(t)\right) \\ &= B(t)d\left(\frac{P(t,T)}{B(t)}\right) + \frac{P(t,T)}{B(t)}dB(t) \\ &= P(t,T)r(t)dt + B(t)\bar{\psi}(t,T)d\bar{W}(t) \\ &= P(t,T)r(t)dt + P(t,T)v(t,T)d\bar{W}(t), \end{aligned}$$

where $v(t,T) \equiv \frac{B(t)\bar{\psi}(t,T)}{P(t,T)}$ which proves the claim. □

An immediate consequence of Lemma 5.12 is that the zero-coupon bond prices under the objective probability measure $\mathcal{P}$ satisfy the following dynamics:

$$\frac{dP(t,T)}{P(t,T)} = \left(r(t) - v(t,T)\cdot\gamma(t)\right)dt + v(t,T)dW(t).$$

Similarly as we saw by Lemma 5.9 in the financial market for equities, $-\gamma$ again is the market price of risk for the fixed income market specifying the excess instantaneous return over the risk free rate $r(t)$ in units of volatility $v(t,T)$.

Chapter 6

Heath-Jarrow-Morton methodology

6.1 The Heath-Jarrow-Morton model

Consider the general setting of a financial market as defined in Section 4.3. Assume we are given an $\mathbf{R}$ valued stochastic process $\alpha = \alpha(\omega, t, T)$ and an $\mathbf{R}^d$ valued stochastic process, $\sigma = (\sigma_1(\omega, t, T), \sigma_2(\omega, t, T), \ldots \sigma_d(\omega, t, T)$, with two indices such that the following conditions hold:

(a) α and σ are $\mathcal{A} \otimes \mathcal{B}$ measurable.

(b) $\int_0^T \int_0^T |\alpha(s,t)| ds dt < \infty$.

(c) $\sup_{s,t \leq T} \|\sigma(s,t)\| < \infty$ for all T and for all $\omega \in \Omega$.

For a given integrable initial forward curve $T\colon \ \mapsto f(0,T)$ we assume that the forward rate process $f(\cdot, T)$ satisfies the following Itó dynamics for $t \leq T$:

$$f(t,T) = f(0,T) + \int_0^t \alpha(s,T) ds + \int_0^t \sigma(s,T) dW(s). \tag{6.1}$$

In (6.1) the date of maturity T, is held fixed and the forward curve is a stochastic process in t for each maturity date T. The integral with respect to the d-dimensional Brownian motion is defined via (4.7). We notice that the assumptions (a), (b) and (c), are sufficient to guarantee the well-definiteness of the integrals. In particular $\sigma \in \mathcal{L}(W)$. Furthermore, assumptions (a) and (c) on σ, are sufficient conditions for the application of the Fubini's theorem for stochastic integrals 4.34, whereas assumption (a) and (b) on

α, provide sufficient conditions for the classical Fubini's theorem. Since the forward rate dynamics is specified before the maturity date T i.e. for $t \leq T$ without loss of generality we set $\alpha(s,t) = 0$ and $\sigma(s,t) = 0$ whenever $s > t$.

Remark 6.1. We note using Corollary 2.19 that the short rate under the HJM model is given by:

$$r(u) = f(u,u) = f(0,u) + \int_0^u \alpha(s,u)ds + \int_0^u \sigma(s,u)dW(s) \tag{6.2}$$

Now σ satisfies the assumptions in Corollary 4.35, hence the process $\int_0^u \sigma(s,u)dW(s)$ has a progressive modification such that the condition given in Corollary 4.35 is satisfied. Furthermore, assumption (b) and the fact that initial forward curve is integrable, guarantee (using the progressive modification of the $\int_0^u \sigma(s,u)dW(s)$ process) that $r(u)$ has a progressive modification such that $\int_0^t |r(u)|du < \infty$ for all $t > 0$. This in particular allows us to define the money market account which is given by (2.5).

6.2 Zero-coupon bond prices dynamics

The dynamics for the forward curve imposes the following for the dynamics of the zero-coupon bond prices.

Lemma 6.2. *For every maturity T, the zero-coupon bond price follows an Itó process of the form:*

$$P(t,T) = P(0,T) + \int_0^t P(s,T)(r(s) + b(s,T))ds + \int_0^t P(s,T)v(s,T)dW(s).$$

Here, $v(s,T) = -\int_s^T \sigma(s,u)du$ is the zero-coupon bond volatility and $b(s,T)$ is defined as:

$$b(s,T) = -\int_s^T \alpha(s,u)du + \frac{1}{2}\| \ v(s,T)\|^2.$$

Proof. Using the first result from Lemma 2.18 we compute explicitly:

$$\begin{aligned} \log P(t,T) &= -\int_t^T f(t,u)du \\ &= -\int_t^T f(0,u)du - \int_t^T \left(\int_0^t \alpha(s,u)ds\right) du \\ &- \int_t^T \left(\int_0^t \sigma(s,u)dW(s)\right) du \end{aligned} \tag{6.3}$$

In (6.3) we use the Fubini's theorem in the first integral and Fubini's theorem for stochastic integrals on the second integral to obtain:

$$\begin{aligned}\log P(t,T) = &-\int_t^T f(0,u)du - \int_0^t \left(\int_t^T \alpha(s,u)du\right) ds \\ &- \int_0^t \left(\int_t^T \sigma(s,u)du\right) dW(s).\end{aligned} \tag{6.4}$$

We rewrite the integral limits as follows:

$$-\int_t^T f(0,u)du = -\int_0^T f(0,u)du + \int_0^t f(0,u)du \tag{6.5}$$

Similarly, for the other two terms in (6.4) we have:

$$-\int_0^t \left(\int_t^T \alpha(s,u)du\right) ds = -\int_0^t \left(\int_s^T \alpha(s,u)du\right) ds + \int_0^t \left(\int_s^t \alpha(s,u)du\right) ds. \tag{6.6}$$

$$\begin{aligned}-\int_0^t \left(\int_t^T \sigma(s,u)du\right) dW(s) = &-\int_0^t \left(\int_s^T \sigma(s,u)du\right) dW(s) \\ &+ \int_0^t \left(\int_s^t \sigma(s,u)du\right) dW(s).\end{aligned} \tag{6.7}$$

Using the fact that $\alpha(s,u) = 0$ whenever $u < s$ we have:

$$\begin{aligned}\int_0^t \left(\int_s^t \alpha(s,u)du\right) ds &= \int_0^t \left(\int_0^s \alpha(s,u)du\right) ds + \int_0^t \left(\int_s^t \alpha(s,u)du\right) ds \\ &= \int_0^t \left(\int_0^t \alpha(s,u)du\right) ds\end{aligned} \tag{6.8}$$

Now applying Fubini's theorem we have:

$$\int_0^t \left(\int_0^t \alpha(s,u)du\right) ds = \int_0^t \left(\int_0^t \alpha(s,u)ds\right) du$$

Again using the fact that $\alpha(s,u) = 0$ whenever $u < s$ we firstly conclude :

$$\int_0^t \left(\int_u^t \alpha(s,u)ds\right) du = 0.$$

Using the last two relations and (6.8) we obtain:

$$\int_0^t \left(\int_s^t \alpha(s,u)du\right) ds = \int_0^t \left(\int_0^u \alpha(s,u)ds\right) du. \tag{6.9}$$

Arguing similarly and applying Fubini's theorem for stochastic integrals we have:

$$\int_0^t \left(\int_s^t \sigma(s,u)du\right) dW(s) = \int_0^t \left(\int_0^u \sigma(s,u)dW(s)\right) du \tag{6.10}$$

Plugging in expression (6.9) in (6.6) we obtain:

$$-\int_0^t \left(\int_t^T \alpha(s,u)du\right) ds = -\int_0^t \left(\int_s^T \alpha(s,u)du\right) ds + \int_0^t \left(\int_0^u \alpha(s,u)ds\right) du \quad (6.11)$$

Similarly, plugging in (6.10) into (6.7) we have:

$$\begin{aligned} -\int_0^t \left(\int_t^T \sigma(s,u)dW(u)\right) ds = &-\int_0^t \left(\int_s^T \sigma(s,u)dW(u)\right) ds \\ &+ \int_0^t \left(\int_0^u \sigma(s,u)dW(s)\right) du \end{aligned} \quad (6.12)$$

Now, we use expressions (6.5), (6.11) and (6.12) to rewrite (6.4) as follows:

$$\begin{aligned} \log P(t,T) = &-\int_0^T f(0,u)du - \int_0^t \left(\int_s^T \alpha(s,u)du\right) ds \\ &- \int_0^t \left(\int_s^T \sigma(s,u)dW(u)\right) ds \\ &+ \int_0^t \left(f(0,u) + \int_0^u \alpha(s,u)ds + \int_0^u \sigma(s,u)dW(s)\right) du \end{aligned} \quad (6.13)$$

Using (6.2) we notice that the integral in the second line reduces to $\int_0^t r(u)du$. Identifying the $b(s,T)$ and $v(s,T)$ terms in the first line of (6.13) and collecting like terms we obtain:

$$\log P(t,T) = \log P(0,T) + \int_0^t \left(r(s) + b(s,T) - \frac{1}{2}\| \, v(s,t)\|^2\right) ds + \int_0^t v(s,T)dW(s)$$

To deduce the zero-coupon bond dynamics we apply Itó formula to $f(X(t)) = e^{X(t)}$, with $X(t) = \log P(t,T)$. In particular, using the notation from Theorem 4.23, we have $f(X(0)) = P(0,T)$, $\langle X, X\rangle_s = \| \, v(s,T)\|^2$, $\frac{\partial f}{\partial x}(X(s)) = P(s,T)$ and $\frac{\partial^2 f}{\partial x^2}(X(s)) = P(s,T)$. Thus, applying Itó theorem in the integral form yields:

$$\begin{aligned} P(t,T) = P(0,T) &+ \int_0^t P(s,T)\left(r(s) + b(s,T) - \frac{1}{2}\| \, v(s,T)\|^2\right) ds \\ &+ \int_0^t P(s,T)v(s,T)dW(s) + \frac{1}{2}\int_0^t P(s,T)\| \, v(s,T)\|^2 ds \end{aligned} \quad (6.14)$$

Simplifying (6.14) proves the claim. □

We recall that the money market account $B(t)$ is the risk-free asset which gives us the return from continuously investing and reinvesting at the short rate. We have the following result regarding the discounted zero-coupon bond price dynamics $\frac{P(t,T)}{B(t)}$.

Corollary 6.3. *For all $t \leq T$ the discounted zero-coupon bond follows a dynamics given by:*

$$\frac{P(t,T)}{B(t)} = P(0,T) + \int_0^t \frac{P(s,T)}{B(s)} b(s,T)ds + \int_0^t \frac{P(s,T)}{B(s)} v(s,T)dW(s).$$

Proof. Recall that the return of the money market account is defined via Remark 6.1 and is given by:

$$B(t) = e^{\int_0^t r(s)ds}.$$

Define $f(X(t),t) = X(t)e^{-\int_0^t r(s)ds}$. We apply Itó theorem 4.23 to f with $X(t) = P(t,T)$. Explicitly computing, we have $f(X(0),0) = X(0) = P(0,T)$. Furthermore,

$$\frac{\partial f}{\partial s}(X(s),s) = -X(s)e^{-\int_0^s r(u)du} r(s) = -\frac{P(s,T)}{B(s)} r(s).$$

Similarly, $\frac{\partial f}{\partial x}(X(s),s) = \frac{1}{B(s)}$ and $\frac{\partial^2 f}{\partial x^2}(X(s),s) = 0$. Thus, according to Itó theorem with the dynamics of the zero-coupon bond prices derived in Lemma 6.2 we obtain:

$$\begin{aligned}\frac{P(t,T)}{B(t)} &= P(0,T) - \int_0^t \frac{P(s,T)}{B(s)} r(s)ds + \int_0^t \frac{1}{B(s)} P(s,T)(r(s)+b(s,T))ds \\ &\quad + \int_0^t \frac{1}{B(s)} P(s,T)v(s,T)dW(s).\end{aligned} \tag{6.15}$$

Simplifying (6.15) yields the result. □

6.3 Short rate dynamics

The Heath-Jarrow-Morton model for the dynamics of the forward curve as presented in Section 6.1 specifies the dynamics that the short rate needs to follow. To derive the short rate dynamics given the dynamics of the forward rate we impose the following regularity assumptions:

(a) α and σ are $\mathcal{A} \otimes \mathcal{B}$ measurable.

(b) $\int_0^T \int_0^T |\alpha(s,t)| ds dt < \infty$.

(c) $\sup_{s,t \leq T} \|\sigma(s,t)\| < \infty$ for all T and for all $\omega \in \Omega$.

(d) $f(0,T)$, $\alpha(t,T)$ and $\sigma(t,T)$ are differentiable in T.

(e) $\int_0^T |\frac{\partial f}{\partial u}(0,u)| du < \infty$.

(f) $\frac{\partial \alpha}{\partial T}(t,T)$ and $\frac{\partial \sigma}{\partial T}(t,T)$ are $\mathcal{A} \otimes \mathcal{B}$ measurable.

(g) $\int_0^T \int_0^T |\frac{\partial \alpha}{\partial T}(s,t)| ds dt < \infty$.

(h) $\sup_{s,t\leq T} \|\frac{\partial\sigma}{\partial t}(s,t)\| < \infty$ for all T and for all $\omega \in \Omega$.

Lemma 6.4. *Under the assumptions a) - h) the short rate dynamics is an Itó process of the form:*

$$r(t) = r(0) + \int_0^t \epsilon(u)du + \int_0^t \sigma(u,u)dW(u),$$

where

$$\epsilon(u) = \alpha(u,u) + \frac{\partial f}{\partial u}(0,u) + \int_0^u \frac{\partial\alpha}{\partial u}(s,u)ds + \int_0^u \frac{\partial\sigma}{\partial u}(s,u)dW(s).$$

Proof. First note that

$$r(t) = f(t,t) = f(0,t) + \int_0^t \alpha(s,t)ds + \int_0^t \sigma(s,t)dW(s). \tag{6.16}$$

Using the fundamental theorem of calculus we write:

$$f(0,t) = f(0,0) + \int_0^t \frac{\partial f}{\partial u}(0,u)du = r(0) + \int_0^t \frac{\partial f}{\partial u}(0,u)du \tag{6.17}$$

Next, we have:

$$\begin{aligned}\int_0^t \alpha(s,t)ds &= \int_0^t \alpha(s,s)ds + \int_0^t (\alpha(s,t) - \alpha(s,s))\, ds \\ &= \int_0^t \alpha(s,s)ds + \int_0^t \left(\int_s^t \frac{\partial\alpha}{\partial u}(s,u)du\right) ds\end{aligned} \tag{6.18}$$

Since for $s > u$ we have $\alpha(s,u) = 0$ we have:

$$\int_0^s \frac{\partial\alpha}{\partial u}(s,u)du = 0$$

Using the last relation, we can rewrite (6.18) and obtain:

$$\int_0^t \alpha(s,t)ds = \int_0^t \alpha(s,s)ds + \int_0^t \left(\int_0^t \frac{\partial\alpha}{\partial u}(s,u)du\right) ds \tag{6.19}$$

Now we use Fubini's theorem on the last term in (6.19) to obtain:

$$\int_0^t \alpha(s,t)ds = \int_0^t \alpha(s,s)ds + \int_0^t \left(\int_0^t \frac{\partial\alpha}{\partial u}(s,u)ds\right) du. \tag{6.20}$$

Similarly, noticing that:

$$\int_u^t \frac{\partial\alpha}{\partial u}(s,u)ds = 0$$

Using the previous relation and (6.20) we finally obtain:

$$\int_0^t \alpha(s,t)ds = \int_0^t \alpha(s,s)ds + \int_0^t \left(\int_0^u \frac{\partial\alpha}{\partial u}(s,u)ds\right) du \tag{6.21}$$

Treating the integral with respect to the forward curve volatility in a similar manner (now applying the Fubini's theorem for stochastic integrals, which can be done based on the extra assumptions) we have the following:

$$\begin{aligned}\int_0^t \sigma(s,t)dW(s) &= \int_0^t \sigma(s,s)dW(s) + \int_0^t (\sigma(s,t) - \sigma(s,s))\,dW(s) \\ &= \int_0^t \sigma(s,s)dW(s) + \int_0^t \left(\int_s^t \frac{\partial\sigma}{\partial u}(s,u)du\right) dW(s) \\ &= \int_0^t \sigma(s,s)dW(s) + \int_0^t \left(\int_0^u \frac{\partial\sigma}{\partial u}(s,u)dW(s)\right) du.\end{aligned} \tag{6.22}$$

Plugging in (6.22), (6.21) and (6.17) into (6.16) completes the proof. □

6.4 Arbitrage Pricing under HJM

In view of the first fundamental theorem of asset pricing as presented in the previous section let $\mathcal{Q} \sim \mathcal{P}$ be an equivalent probability measure of the form given by Girsanov's theorem. By $\bar{W}$ we denote the $\mathcal{Q}$ transformed Brownian motion. Analogously to the asset pricing model, we will call $\mathcal{Q}$ an equivalent martingale measure (EMM) for the bond market if the discounted zero-coupon bond price process $\frac{P(t,T)}{B(t)}$ is a $\mathcal{Q}$ martingale for $t \leq T$ for all maturities T. In the first fundamental theorem of asset pricing we have seen that the existence of an EMM rules out arbitrage opportunities over admissible strategies. Thus, we have the following sufficient condition for no arbitrage in the bond market, called the HJM drift condition. We are going to assume that the regularity assumptions (a) - (h) hold.

Theorem 6.5 (HJM drift condition)**.** *Suppose $b(s,T)$ and $v(s,T)$ are defined as in Lemma 6.2. Moreover suppose $\mathcal{Q} \sim \mathcal{P}$ is a measure equivalent to $\mathcal{P}$ given by Girsanov's theorem. If $\mathcal{Q}$ is an EMM for the bond market then $b(t,T) = -v(t,T)\cdot\gamma(t)$. In this case the $\mathcal{Q}$ dynamics of the forward rate is given by $f(t,T) = f(0,T) + \int_0^t d(s,T)ds + \int_0^t \sigma(s,T)d\bar{W}(s)$, where the HJM drift $d(s,T)$ is given by $d(s,T) = \sigma(s,T)\cdot\int_s^T \sigma(s,u)du$.*

Proof. Writing Corollary 6.3 in differential form we have:

$$d\frac{P(t,T)}{B(t)} = \frac{P(t,T)}{B(t)}b(t,T)dt + \frac{P(t,T)}{B(t)}v(t,T)dW(t).$$

Hence, for the $\mathcal{Q}$ transformed dynamics we obtain:

$$d\frac{P(t,T)}{B(t)} = \frac{P(t,T)}{B(t)}(b(t,T) + v(t,T)\cdot\gamma(t))dt + \frac{P(t,T)}{B(t)}v(t,T)d\bar{W}(t) \tag{6.23}$$

Since $\mathcal{Q}$ is an EMM, from (6.23) we must have $b(t,T) = -v(t,T)\cdot\gamma(t)$ $\mathcal{Q}\otimes dt$ a.s. which implies $b(t,T) = -v(t,T)\cdot\gamma(t)$, $\mathcal{P}\otimes dt$ a.s. Differentiating both sides of the last relation with respect to T and solving for $\alpha(t,T)$ we obtain:

$$\alpha(t,T) = d(t,T) - \sigma(t,T)\cdot\gamma(t). \tag{6.24}$$

Plugging in equation (6.24) into (6.1) and transforming to the $\bar{W}$ Brownian motion proves the claim. □

Imposing additional assumptions, we have the converse statement which states sufficient conditions for $\mathcal{Q}$ to be an EMM.

Theorem 6.6. *Suppose $b(s,T)$ and $v(s,T)$ are defined as in Lemma 6.2. Suppose $\mathcal{Q}\sim\mathcal{P}$ is a measure equivalent to $\mathcal{P}$ given by Girsanov's theorem. Let $b(t,T) = -v(t,T)\cdot\gamma(t)$, $\mathcal{P}\otimes dt$ a.s. and suppose $v(t,T)$ satisfies the Novakov condition i.e. $\mathbb{E}^{\mathcal{Q}}\left(e^{\frac{1}{2}\int_0^t v(s,T)\cdot v(s,T)ds}\right) < \infty$. Then $\mathcal{Q}$ is an EMM.*

Proof. Since $b(t,T) = -v(t,T)\cdot\gamma(t)$ the discounted bond price dynamics under $\mathcal{Q}$ satisfies:

$$d\frac{P(t,T)}{B(t)} = \frac{P(t,T)}{B(t)}v(t,T)d\bar{W}(t).$$

Thus, from Lemma 4.26 we may write it in integral form as:

$$\frac{P(t,T)}{B(t)} = P(0,T)\mathcal{E}_t(v(\cdot,T)\bullet\bar{W}).$$

From Theorem 4.29 and the last relation we obtain that $\frac{P(t,T)}{B(t)}$ is a $\mathcal{Q}$ martingale. Therefore, $\mathcal{Q}$ is in fact an EMM for the bond market. □

As a simple application of the HJM drift condition let $\mathcal{Q}$ be an EMM and consider the simplest HJM model where the volatility is constant i.e. $\sigma(t,T) = \sigma$ for all $T > 0$. Then the HJM drift is given by $d(s,T) = \sigma\cdot\sigma(T-s)$ and for the forward rate we obtain:

$$f(t,T) = f(0,T) + \sigma\cdot\sigma t(T - \frac{t}{2}) + \sigma\bar{W}(t).$$

Hence, for the short rate $r(t)$ we have:

$$r(t) = f(t,t) = f(0,t) + \frac{\sigma\cdot\sigma t^2}{2} + \sigma\bar{W}(t),$$

which is a special case of the Ho-Lee model presented in (5.6), with $b(t) = \sigma\cdot\sigma t$.

Chapter 7

Concluding remarks

The discovery of the subsequently called Brownian motion by the botanist Robert Brown in 1827 sparked a revolution in the physical sciences. At the same time, it served as a starting point in the mathematical modeling of financial markets. Eminently, one of the first mathematical descriptions of Brownian motion was applied by Louis Bachelier in 1900 to study stock and option pricing [24]. Since the early work of Bachelier, the field of mathematical finance has grown tremendously, owning largely due to the development of new mathematical tools suitable for modeling and pricing increasingly complex financial contracts. Most of the early work in mathematical finance, has been devoted to modeling the stock (equities) market. Starting with the work by Vasiček, the mathematical finance field has made a significant progress in modeling the fixed income market [2].

Throughout this thesis, we have explored two approaches to addressing arguably the two most important problems in the mathematical modeling of the fixed income market. Using functional analysis and optimization in Hilbert space methods in Chapter 3 we have presented a method to estimate a smooth discount curve given a set of benchmark actively traded fixed income securities [8]. The method was developed in a recent work by Filipović and Willems and relies on the non arbitrage prices of actively traded fixed income securities in order to estimate the discount curve of zero-coupon bond prices traded at the current date [8]. The discount curve obtained via this method has maximal smoothness, reproduces the market quotes perfectly and is given in a closed form (see Theorem 3.6). This makes the method easily implementable in practise. The method presented however, relies on norm minimization techniques in a Hilbert functional space. Therefore, the shape of the discount curve is highly sensitive to the a priori specification of the Hilbert space norm. The choice for the norm is based on the economic argument that the forward rates over two infinitesimal periods should not vary a lot. Whether this is reasonable assumption should be empirically validated in further work. In addition, the method derives the discount curve based on actively traded fixed income quotes by assuming that their market prices match the mathematically derived non arbitrage prices for all maturity dates T; thus assuming that the market is arbitrage free. Although, this is a common practise in mathematical finance, the non arbitrage assumption certainly affect the practical applicability of the method especially among financial institutions involved in trading on the fixed income markets.

In Chapter 5 we have highlighted the differences and similarities between the fixed income and the equities market. Unlike the equities market, the fixed income market consists of infinitely many financial assets. Hence, its mathematical modeling requires some separate techniques, a survey of which is presented in Chapter 4.

In Chapter 6 we have presented the famous Heath-Jarrow-Morton (HJM) valuation methodology, whereby the dynamics of the forward curve is directly specified. From the dynamics of the forward curve, we have derived the dynamics of the zero-coupon bond prices in Lemma 6.2. As can be seen from Lemma 6.2, the discount curve, $P(0,T)$, is a basic input to the HJM valuation methodology. Since the discount curves in Theorems 3.6 and 3.7 are smooth, they satisfy the regularity requirements in the HJM methodology. Thus, combining the method presented in Chapter 3 with the HJM methodology, in particular, Theorems 3.6 and 3.7 and Lemma 6.2, it is possible to numerically obtain the full term structure of zero-coupon bonds for any maturity date T at any future date $t \leq T$. The numerical implementation is left for future research. Another question of fundamental importance which could be addressed in further research is whether the discount curve derived by the method presented in Chapter 3 is dynamically consistent (i.e. leads to arbitrage free evolution) with the HJM valuation methodology [25, 26].

Bibliography

[1] Damir Filipovic. *Term-Structure Models. A Graduate Course.* Springer, 2009.

[2] Oldrich Vasicek. An equilibrium characterization of the term structure. *Journal of financial economics*, 5(2):177–188, 1977.

[3] Charles R Nelson and Andrew F Siegel. Parsimonious modeling of yield curves. *Journal of business*, pages 473–489, 1987.

[4] Lars EO Svensson. Estimating and interpreting forward interest rates: Sweden 1992-1994. Technical report, National bureau of economic research, 1994.

[5] Leif Andersen. Discount curve construction with tension splines. *Review of Derivatives Research*, 10(3):227–267, 2007.

[6] Ken Adams. Smooth interpolation of zero curves. *Algo Research Quarterly*, 4(1/2): 11–22, 2001.

[7] Freddy Delbaen and Sabine Lorimier. Estimation of the yield curve and the forward rate curve starting from a finite number of observations. *Insurance: Mathematics and Economics*, 11(4):259–269, 1992.

[8] Damir Filipovic and Sander Willems. Exact smooth term-structure estimation. *SIAM Journal on Financial Mathematics*, 9(3):907–929, 2018.

[9] John Hull and Alan White. Pricing interest-rate-derivative securities. *The Review of Financial Studies*, 3(4):573–592, 1990.

[10] Thomas SY Ho and Sang-Bin Lee. Term structure movements and pricing interest rate contingent claims. *the Journal of Finance*, 41(5):1011–1029, 1986.

[11] Tomas Björk. *Arbitrage theory in continuous time.* Oxford university press, 2009.

[12] David Heath, Robert Jarrow, and Andrew Morton. Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica: Journal of the Econometric Society*, pages 77–105, 1992.

[13] Eugen von Böhm-Bawerk. Capital and interest. 3 vols. *Trans. George D. Huncke and Hans F. Sennholz. South Holland, Ill.: Libertarian Press*, 1959.

[14] Ludwig Von Mises. *Human action*. Ludwig von Mises Institute, 1966.

[15] Jörg Guido Hülsmann. A theory of interest. *Quarterly Journal of Austrian Economics*, 5(4):77–110, 2002.

[16] Eberhard Zeidler. *Applied functional analysis: main principles and their applications*, volume 109. Springer Science & Business Media, 2012.

[17] Eberhard Zeidler. *Nonlinear functional analysis and its applications: III: variational methods and optimization*. Springer Science & Business Media, 2013.

[18] Walter Rudin. *Real and complex analysis*. Tata McGraw-Hill Education, 2006.

[19] Holopainen Ilkka. *Lecture notes in Real Analysis I*. University of Helsinki, 2017.

[20] Carlos A Braumann. *Introduction to Stochastic Differential Equations with Applications to Modelling in Biology and Finance*. Wiley, 2019.

[21] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293. Springer Science & Business Media, 2013.

[22] Freddy Delbaen and Walter Schachermayer. The fundamental theorem of asset pricing for unbounded stochastic processes. *Mathematische annalen*, 312(2):215–250, 1998.

[23] Freddy Delbaen and Walter Schachermayer. A general version of the fundamental theorem of asset pricing. *Mathematische annalen*, 300(1):463–520, 1994.

[24] Louis Bachelier. *Louis Bachelier's theory of speculation: the origins of modern finance*. Princeton University Press, 2011.

[25] Robert A Jarrow. Forward rate curve smoothing. *Annu. Rev. Financ. Econ.*, 6(1): 443–458, 2014.

[26] Christian Gourieroux and Alain Monfort. Linear-price term structure models. *Journal of Empirical Finance*, 24:24–41, 2013.