Interest Rate Theory and Stochastic Duration

by

Henrik Paulov Hammer

THESIS

for the degree of

Master of Science

Modeling and Data Analysis

(Master i Modellering og Dataanalys)

Faculty of Mathematics and Natural Sciences

University of Oslo

May 2014

Det matematisk- naturvitenskapelige fakultet

Universitetet i Oslo

Abstract

With the new regulations of Basel III and Solvency II there is a necessity to have tools that can measure different types of financial and insurance risk in a portfolio. *Stochastic Duration* is such a measure. This new type of measure, which is for the first time implemented in this thesis, can be used to analyze the sensitivity of complex portfolios of interest rate derivatives with respect to the stochastic fluctuation of the entire term structure of interest rates or the yield surface without assuming as in the classical case (Macaulay duration) flat or piecewise flat interest rates. It is conceivable that this concept will serve as an important tool within risk management and replace the classical Macaulay duration.

Moreover, using the concept of immunization strategies based on stochastic duration we will be able to hedge the expected uncertainty due to the changes in the forward rate in complex bond portfolios.

Keywords: Stochastic Duration, Immunization Strategy, HJM-modeling, Vasicek, Hull-White, CIR.

Acknowledgements

I would like to thank my supervisor Frank Proske for giving me a challenging and interesting thesis. In connection to that I would like to thank both Frank and Ph.D. student Sindre Duedahl for the help and guidance throughout this work.

Specially I would like to thank Michael Graff, Henrik Nyhus, Anders Hovdenes, and Olav Breivik for a lot of interesting conversations and a good student environment, as well as of course all the others I have met during my Master at University of Oslo and Bachelor at University of Agder.

At last, but not at least, I would like to thank my family and friends for their endless support. A special thanks goes to my girlfriend Yaiza who has been a good discussion partner and helped me pushing limits.

Contents

List of Figures

Abbreviations

w.r.t.	**w**ith **r**espect **t**o,
s.t.	**s**uch **t**hat
a.s.	**a**lmost **s**urley
ZCB	**Z**ero **C**oupon **B**ond
EMM	**E**quivalent **M**artingale **M**easure
SDE	**S**tochastic **D**ifferential **E**quation
SPDE	**S**tochastic **P**artial **D**ifferential **E**quation
SSDE	**S**emi-linear **S**tochastic **D**ifferential **E**quation
CIR	**C**ox-**I**ngersoll-**R**oss
HJM	**H**eath-**J**arrow-**M**orton
YTM	**Y**ield-**t**o-**M**aturity
PCA	**P**rincipal **C**omponent **A**nalysis
CG	**C**ylindrical **G**aussian Measure
CBM	**C**ylindrical **B**rownian **M**otion
BM	**B**rownian **M**otion
ONB	**O**rtho**n**ormal **B**asis

Symbols

Ω	Sample Space
$\mathcal{F}$	Filtration
$\mathbb{P}$	Probability Measure
$\mathbb{Q}$	Risk Neutral Measure
$\tilde{\mathbb{P}}$	Centered Forward Rate Measure
$P(t,T)$	ZCB price
$B(t)$	Normalizer
$Y(t,T)$	Yield Curve
$\tilde{}$	Normalized Value
H, K	Separable Hilbert Spaces
$HS(H,K)$	Hilbert-Schmith operators
$\mathcal{V}[0,T]$	Itô Integrable Integrands w.r.t. BM
$\mathcal{L}_{[0,T]}(H,K)$	Itô Integrable Integrands w.r.t. CBM
H_w	Consistent Hilbert Space
$D_{\hat{f}}$	Malliavin Derivative w.r.t. the centered forward curve $\hat{f}$

Chapter 1

Introduction

The main objective in this master thesis is the analysis and implementation of the Stochastic Duration applied to bond portfolios. But in order to work with the Stochastic Duration we need some elementary understanding of Interest Rate Theory.

Chapter 2: Short Rate Models

In this chapter we go through the most elementary tools and thoughts within interest rate theory. The interest is in deriving prices on a ZCB, using different short rate models.

For a more thorough review [1] is recommended.

Chapter 3: A First Look at the HJM-model

Instead of modeling the short rate, an alternative, presented by Heath, Jarrow and Merton, is to model the instantaneous forward rate. With such a model we obtain that the arbitrage free drift only depends on the volatility structure. Later on we extend the model, using the Musiela parametrization. The parametrization exhibits some interesting properties we want to study [2].

Further we aim at discussing a calibration procedure that we later want to use in connection with a practical example. To read more about the subject [3] is recommended.

Chapter 4: Infinite Dimensional Stochastic Analysis

A market observation is that there exists time-to-maturity specific risk. With a possible infinite time-to-maturity we need infinite dimensions of noises. This part is mainly build on [4].

Chapter 5: Generalized HJM framework

Through the generalized HJM model we include the possibility of ZCB's having infinite time-to-maturity. The generalized HJM model is used in the paper [5], where Stochastic Duration is presented and constructed.

Chapter 6: Stochastic Duration

The concept of Stochastic Duration is presented in this chapter. We go through several examples and provide a program for the stochastic duration on a simulated portfolio.

When we have calculated the stochastic duration of a portfolio we can use the immunization strategy to hedge the interest rate risk. Read more about Stochastic Duration in [5] and [6].

Chapter 7: Stochastic Duration an Example

In the last chapter we import data of the US Treasury yield curve and a Future contract on a 2 year Treasury Note. The data are collected from www.quandl.com. We go through a principal component analysis and use the estimated parameters to derive the stochastic duration on a portfolio of a 2 year Treasury Note.

Appendix A: Mathematical Tools

The appendix goes through the most important mathematical tools used in chapter two and three in this thesis.

Chapter 2

Short Rate Models

Our main interest in this chapter is to derive a price on a zero-coupon bond (ZCB).

Definition 2.1 (Zero-Coupon Bond [1])**:**
A zero-coupon bond with maturity date T, also called a T-bond, is a contract which guarantees the holder 1 dollar to be paid on the date T. The price at time t of a bond with maturity date T is denoted by $P(t, T)$.

We are going to treat the ZCB price as a derivative w.r.t. the instantaneous short rate as the underlying process. But, as we will encounter later, we don't necessarily need to provide a dynamic on the short rate (also called overnight rate). We can e.g. use the relation between the instantaneous forward rate and short rate to deduce a dynamic of the short rate given the dynamics on the forward rate. The latter is referred to as the HJM-framework.

In all cases we would like to have a model that creates an arbitrage free price. An arbitrage means that with initial portfolio value at time 0 of zero we, $\mathbb{P}$-a.s., have a portfolio at a later time T that is bigger or equal than zero, with a probability bigger than 0 for the value of the portfolio being bigger than zero. Basically a non-risky portfolio with just the up-side.

Theorem 2.2 (First Fundamental theorem [1])**:**
A market is arbitrage free if there exist a probability measure $\mathbb{Q}(A)$ that is an equivalent Martingale measure (EMM) to $\mathbb{P}(A)$ s.t. the normalized asset price is a Martingale.

A proof is provided in [1]. The second fundamental theorem is about completeness in the market. Given that we find a model of our normalized market that is arbitrage free, then the market is complete iff the Martingale measure is unique[1]. The question is whether our dynamics creates a model which is both arbitrage free and complete. In fact the question is; yes we find a model for an arbitrage free price; but no, the market isn't complete. This leads us to what is called Martingale modeling.

2.1 Zero-Coupon Market

We are going to assume that there exist a "risk less" asset referred to as the Normalizer, and a zero-coupon bond following the assumptions

Assumption 2.3 (Regular Market)**:**
We have the following assumption of a regular market

1. There exist a market for T-bonds for each T
2. $P(t,t) = 1$ for all t (if not there is an arbitrage possibility)
3. For a fixed t, the bond price $P(t,T)$ is differentiable w.r.t. time-of-maturity T.

Definition 2.4 (Normalizer)**:**
The normalizer process is defined as

$$B_t = \exp\Big\{\int_0^t r_s ds\Big\}, \tag{2.1}$$

which is the solution of the SDE

$$\begin{cases} dB(t) = r_t B(t)dt \\ B(0) = 1. \end{cases}$$

[1]This is not the case in the infinite dimensional model.

2.2 Short rate models

The short rate is defined as the limit, $T \to t$, of the instantaneous forward rate.

Definition 2.5 (Instantaneous Short Rate)**:**
Let the instantaneous forward rate with maturity T, contracted at time t be defined as

$$f(t,T) \stackrel{def}{=} -\frac{\partial \log(P(t,T))}{\partial T}, \tag{2.2}$$

then the instantaneous short rate at time t is defined as $f(t,t)$

$$r_t \stackrel{def}{=} f(t,t) = \lim_{T \to t} f(t,T). \tag{2.3}$$

By (2.2) and (2.3) we derive the following relation between the instantaneous short rate, r_t, and the ZCB price, $P(t,T)$,

$$P(t,T) = \exp\Big\{ -\int_t^T r_s ds \Big\}. \tag{2.4}$$

Clearly $P(t,t) = 1$ for all $t \in \mathbb{R}_+$.

Still we haven't chosen the model for r_t. It might be deterministic, but this relies on the future to be certain. Due to liquidity risk, default risk, competitive bond market where the prices are based on supply and demand, and of course company ratings, a better approach is to model r_t on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$. Then by the relation (2.4) we see that the price of $P(t,T)$ is stochastic w.r.t. the underlying process r_t. But what is a fair price?

From a mathematical point of view the fair price is the expected arbitrage free price. Recall the First Fundamental theorem: An arbitrage free price is equivalent with the existence of an Equivalent Martingale measure (EMM) to $\mathbb{P}$ in the normalized market. We define the normalization of the ZCB price (also commonly called the discounted ZCB price).

Definition 2.6 (Normalized ZCB price)**:**

The normalized ZCB price is defined as

$$\tilde{P}(t,T) = \frac{P(t,T)}{B(t)}. \tag{2.5}$$

By the First Fundamental theorem we would like the normalized ZCB price to be a Martingale under the EMM $\mathbb{Q}$. I.e. from the definition of the Martingale we have that

$$\tilde{P}(s,T) = E_{\mathbb{Q}}\big[\tilde{P}(t,T)|\mathcal{F}_s\big], \tag{2.6}$$

for $t \geq s$. This leads us to the price of the ZCB

$$P(s,T) = B(s)E_{\mathbb{Q}}\big[\tilde{P}(t,T)|\mathcal{F}_s\big]. \tag{2.7}$$

Putting in for $\tilde{P}(t,T)$ yields

$$P(s,T) = B(s)E_{\mathbb{Q}}\Big[exp\big\{-\int_0^T r_u\,du\big\}|\mathcal{F}_s\Big].$$

We can in fact derive the SDE of the ZCB. Using the Martingale representation theorem we know that the dynamics of the normalized ZCB price is

$$d\tilde{P}(t,T) = \sigma_t dW_t^{\mathbb{Q}}$$

for some function $\sigma_t \in L_2(\mathbb{Q})$(existence of second moment). If we assume that the Girsanov transform between the Equivalent Martingale measure $\mathbb{Q}$ and the observed probability measure $\mathbb{P}$ was on the form

$$dW_t^{\mathbb{Q}} = -\lambda_t dt + dW_t^{\mathbb{P}},$$

then by Itô's formula we derive the dynamics of the ZCB price as

$$\begin{aligned} dP(t,T) = d(B(t)\tilde{P}(t,T)) &= \tilde{P}(t,T)dB(t) + B(t)d\tilde{P}(t,T) \\ &= P(t,T)r_t dt + P(t,T)\sigma_t dW^{\mathbb{Q}} \\ &= P(t,T)r_t dt + P(t,T)\sigma_t(-\lambda_t dt + dW^{\mathbb{P}}) \\ &= P(t,T)(r_t - \sigma_t\lambda_t)dt + P(t,T)\sigma_t dW^{\mathbb{P}}. \end{aligned}$$

Based on the last equation we are in the "core" of Martingale modeling. The function λ_t isn't unique in the ZCB market. This comes from the fact that we are working within incomplete markets. Therefore it is common to define the models directly via the $\mathbb{Q}$-dynamics.

2.2.1 The Portfolio Setup

We want to model the arbitrage free ZCB price based on a short rate dynamics. Assume that under the objective probability measure $\mathbb{P}$ the dynamics of r_t is the solution of a SDE of the form

$$dr_t = \mu(t,r_t)dt + \sigma(t,r_t)dW_t^{\mathbb{P}}, \tag{2.8}$$

where we recall that the dynamics of the normalizer is

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

The idea is to let the risk free asset $B(t)$ be the benchmark. Then under the EMM $\mathbb{Q}$ the expected return should be equal to the benchmark, $B(t)$. Assume that the price of a ZCB takes the form

$$P(t,T) = F(t,r_t;T), \tag{2.9}$$

where we assume that F is a smooth function of three variables. Since $P(T,T) = 1$ we have the obvious relation that $F(T,r_T;T) = 1$ for all r_T.

We create a portfolio based on the two assets; the Normalizer and the ZCB with the corresponding stock holding, α_{ZCB} and α_B. Then the portfolio value at time t is

$$V_{t;T} = \alpha_B B(t) + \alpha_{ZCB} P(t,T),$$

and by linearity we derive the SDE

$$\begin{aligned} dV_{t;T} &= d[\alpha_B(t)B(t)] + d[\alpha_{ZCB}(t)P(t,T)] \\ \text{(Self-Financing)} &= \alpha_B dB(t) + \alpha_{ZCB} dP(t,T) \end{aligned}$$

when we assume self-financing portfolios. A self-financing portfolio is a portfolio choice where the the stock holding doesn't change during the portfolio time. Hence, $\alpha_B(t) \equiv \alpha_B$.

Let η_B and η_{ZCB} be the weights of the portfolio. I.e.

$$\eta_B(t) = \frac{\alpha_B(t)B(t)}{\alpha_B(t)B(t) + \alpha_{ZCB}(t)P(t,T)}.^{2} \tag{2.10}$$

Then we deduce that the portfolio weights can be written as

$$\alpha_{ZCB} = V_{t;T}\frac{\eta_{ZCB}(t)}{P(t,T)}.$$

Plugging into the portfolio value dynamics we derive that

$$dV_{t;T} = V_{t;T}\big(\frac{\eta_B(t)}{B(t)}dB(t) + \frac{\eta_{ZCB}(t)}{P(t,T)}dP(t,T)\big).$$

Earlier we assumed that $P(t,T) = F(t,r_t;T)$, and from the dynamics of the short rate model we have, using the Itô's formula, that

$$\begin{aligned} dP(t,T) = dF(t,r_t;T) = [F^t(t,r_t;T) + \mu F^r(t,r_t;T) + \frac{1}{2}\sigma^2 F^{rr}(t,r_t;T)]dt \\ + \sigma F^r(t,r_t;T)dW_t^{\mathbb{P}}, \end{aligned}$$

[2]Similar for the ZCB

where e.g. $F^r = \frac{\partial F}{\partial r}$ etc. By plugging in the portfolio process

$$dV_{t;T} = V_{t;T}\big(\frac{\eta_B}{B(t)}B(t)r_t + \frac{\eta_{ZBC}}{F(t,r_t;T)}[F^t(t,r_t;T) + \mu F^r(t,r_t;T) + \frac{1}{2}\sigma^2 F^{rr}(t,r_t;T)]\big)dt$$
$$+V_{t;T}\eta_{ZCB}\sigma F^r(t,r_t;T)dW_t^{\mathbb{P}}.$$

Using Girsanov's theorem,

$$dW_t^{\mathbb{Q}} = -\lambda_t dt + dW_t^{\mathbb{P}},$$

we change our portfolio to be a dynamic under the risk-neutral measure $\mathbb{Q}$

$$dV_{t;T} = V_{t;T}\big(\eta_B r_t + \frac{\eta_{ZBC}}{F(t,r_t;T)}[F^t(t,r_t;T) + (\mu - \sigma\lambda)F^r(t,r_t;T) + \frac{1}{2}\sigma^2 F^{rr}(t,r_t;T)]\big)dt$$
$$+V_{t;T}\eta_{ZCB}\sigma F^r(t,r_t;T)dW_t^{\mathbb{Q}}.$$

Under the risk neutral measure $\mathbb{Q}$, the drift of the ZCB portfolio must be equal to the Benchmark(Normalizer)[3]. The Portfolio process holding just the Benchmark is equivalent to having a portfolio weight of $\eta_B(t) \equiv 1$. By the property $\eta_B(t) + \eta_{ZCB}(t) = 1$ $\eta_{ZCB}(t) \equiv 0$. Hence

$$V_{t;T}\big(\eta_B r_t + \frac{\eta_{ZBC}}{F(t,r_t;T)}[F^t(t,r_t;T) + \mu F^r(t,r_t;T) + \frac{1}{2}\sigma^2 F^{rr}(t,r_t;T)]\big)dt = V_{t;T}r_t dt,$$

which leads to

$$\eta_{ZBC}[F^t(t,r_t;T) + (\mu - \sigma\lambda)F^r(t,r_t;T) + \frac{1}{2}\sigma^2 F^{rr}(t,r_t;T)] + F(t,r_t;T)r_t(\eta_B - 1) = 0.$$

Using the property, $\eta_B + \eta_{ZCB} = 1$, again

$$F^t(t,r_t;T) + (\mu - \sigma\lambda)F^r(t,r_t;T) + \frac{1}{2}\sigma^2 F^{rr}(t,r_t;T) - F^r(t,r_t;T)r_t = 0.$$

Recall the boundary condition $F(T,r_T;T) = 1$. Then we have found what is called the term structure equation.

[3]Under the risk neutral measure the Brownian motion should fluctuate around the path of the Normalizer (the risk free asset)

Proposition 2.7 (Term structure equation)**:**
In an arbitrage free market $F(t, r_t; T)$ will satisfy the term structure equation

$$F^t(t, r_t; T) + (\mu - \sigma\lambda)F^r(t, r_t; T) + \tfrac{1}{2}\sigma^2 F^{rr}(t, r_t; T) - F(t, r_t; T)r_t = 0,$$
$$F(T, r_T; T) = 1.$$

We can in fact generalize this equation to all T-claims, where we have the boundary condition $F(\tau, r_\tau, T) = \Phi(r_\tau)$, for a contract Φ. Here we see what was meant by the view of a ZCB price being the financial derivative w.r.t. the underlying process r_t and the contract $\Phi(r_\tau) = 1$.

We generalize the Proposition (2.7) for all T-claims and apply the Feynman-Kac stochastic representation formula.

Proposition 2.8:
Let a T-claim be contracted as $\Phi(r_\tau)$. Then in an arbitrage free market the price of the contract at time t is

$$p(t; \Phi) = F(t, r(t); T),$$

where the functional [4] *F solves the boundary condition*

$$F^t(t, r_t; T) + (\mu - \sigma\lambda)F^r(t, r_t; T) + \tfrac{1}{2}\sigma^2 F^{rr}(t, r_t; T)] - F(t, r_t; T)r_t = 0,$$
$$F(\tau, r_\tau; T) = \Phi(r_\tau).$$

Further more by the Feynman-Kac stochastic representation, F is solved by

$$F(t, r; T) = B(t)^{-1} E_{\mathbb{Q}}\Big[B(\tau)\Phi(r_\tau)|\mathcal{F}_t\Big],$$

where r_t is $\mathcal{F}_t$-adapted stochastic process with the following $\mathbb{Q}$-dynamics

$$dr_t = [\mu(t, r_t) - \lambda_t \sigma(t, r_t)]dt + \sigma(t, r_t)dW_t^{\mathbb{Q}}.$$

[4] A functional is a function of a function

As long as we have a finite dimension of noises the normalizer (risk free asset) could in fact be an ZCB with maturity $S \neq T$ (see [1]).

Martingale modeling

It is common to define the r_t dynamics under the $\mathbb{Q}$-measure. This is in literature, [1], referred to as Martingale modeling. This means that instead of having an representation of the $\mathbb{Q}$-dynamics in the following way

$$dr_t = [\mu(t, r_t) - \lambda_t \sigma(t, r_t)]dt + \sigma(t, r_t)dW_t^{\mathbb{Q}},$$

we neglect[5] the second term in the drift and define models through the dynamics

$$dr_t = \mu(t, r_t)dt + \sigma(t, r_t)dW_t^{\mathbb{Q}}.$$

2.2.2 Affine Term Structure Models

Affine term structure models are a family of models that have a certain "nice" solution to ZCB prices. All of the models presented in the upcoming section have an Affine Term Structure.

Definition 2.9 (Affine Term Structure Models)**:**
If the solution to the term structure equation (prop. 2.7) $F(t, r_t; T)$ is on the form

$$F(t, r_t; T) = \exp\Big\{A(t, T) - B(t, T)r_t\Big\}, \tag{2.11}$$

where $A(t, T)$ and $B(t, T)$ are deterministic functions, then the model possesses an Affine term structure.

[5] More precisely neglect the procedure of going from $\mathbb{P}$-dynamics to $\mathbb{Q}$-dynamics

Given an affine term structure, the term structure equation[6] have the following form

$$A^t(t,T) - r_t B^t(t,T) - \mu(t,r_t)B(t,T) + \frac{1}{2}\sigma(t,r_t)^2 B(t,T)^2 - r_t = 0. \tag{2.12}$$

In order for the solution to satisfy the boundary value condition in the term structure equation the functions A and B need to satisfy the boundary values

$$\begin{aligned} A(T,T) &= 0, \\ B(T,T) &= 0. \end{aligned}$$

Assume that the drift and the volatility structure have the following form [1],

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

and put them into the term structure equation (2.11). Then

$$A^t(t,T) - r_t B^t(t,T) - (\alpha(t)r_t + \beta(t))B(t,T) + \frac{1}{2}\big(\sqrt{\gamma(t)r_t + \delta(t)}\big)^2 B(t,T)^2 - r_t = 0.$$

Because the term structure equation holds for every r_t we get, by dividing the terms w.r.t. r_t-relations, following two systems to solve

$$\begin{aligned} &A^t(t,T) - \beta(t)B(t,T) + \tfrac{1}{2}\delta(t)B(t,T)^2 = 0, \\ &r_t\big(- B^t(t,T) - \alpha(t)B(t,T) + \tfrac{1}{2}\gamma(t)B(t,T)^2\big) = r_t. \end{aligned}$$

This leads us to the following proposition.

Proposition 2.10 (Affine Term Structure)**:**
Assume that μ and σ are given as

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

[6] Remember that we are using Martingale modeling

then we have a solution of the term structure on the form

$$F(t, r_t; T) = \exp\Big\{A(t,T) - B(t,T)r_t\Big\}, \tag{2.13}$$

where $B(t,T)$ *and* $A(t,T)$ *are solved through the differential equations*

$$\begin{cases} B^t(t,T) + \alpha(t)B(t,T) - \frac{1}{2}\gamma(t)B(t,T)^2 = -1 \\ B(T,T) = 0 \end{cases}$$

and

$$\begin{cases} A^t(t,T) - \beta(t)B(t,T) + \frac{1}{2}\delta(t)B(t,T)^2 = 0 \\ A(T,T) = 0 \end{cases}$$

respectively.

This type of differential equation is commonly referred to as Riccati equations.

2.2.3 Some specific short rate models

We are going to present three short rate models. The three models are well known as the Vasicek, Cox-Ingersoll-Ross and Hull-White model [7];

$$\begin{aligned} \text{Vasicek: } dr_t &= k[\theta - r_t]dt + \sigma dW_t^{\mathbb{Q}}, \\ \text{CIR: } dr_t &= k[\theta - r_t]dt + \sigma\sqrt{r_t}dW_t^{\mathbb{Q}}, \\ \text{Hull-White: } dr_t &= [\theta_t - a_t r_t]dt + \sigma_t dW_t^{\mathbb{Q}}. \end{aligned}$$

All models have their pros and cons. From a mathematical educational point of view, showing the approach rather than finding the dynamics fitted perfectly in the market is important. The Vasicek model deficiency is that it allows, with probability (or the to high probability) bigger than zero of having negative interest rates. But in methodical research the model is very nice because of the simple structure, and we find analytical solution to both ZCB prices and option-prices easily.

The CIR model is a non-negative short rate model. Its deficiency is that it doesn't provide a satisfying noise term, and that it is difficult to deal with, although it provides analytical solutions to the most important derivatives[7].

The Hull-White model is similar (extension) to the Vasicek model with the exception of possible t dependence in the parameters. This improvement provides a consistency relation with today's term structure (or yield curve observed today).

Vasicek Model

As we defined earlier the Vasicek model have the following $\mathbb{Q}$-dynamics,

$$dr_t = k[\theta - r_t]dt + \sigma dW_t^{\mathbb{Q}}. \tag{2.14}$$

We can solve the SDE quite easily by using Itô's formula on the function $g(t,x) = e^{kt}x$, where the underlying process is r_t. We derive

$$\begin{aligned} de^{k\,t}r_t &= k\,e^{k\,t}\,r_t\,dt + e^{k\,t}dr_t \\ \text{(The dynamics)} &= \theta\,e^{k\,t}dt + e^{k\,t}\,\sigma\,dW_t^{\mathbb{Q}}. \end{aligned}$$

Solving this equation over the interval $[s,t]$ yields

$$r_t = e^{-k\,(t-s)}\,r_s + \,\theta\,(1 - e^{-k(t-s)}) + \int_s^t \sigma\,e^{-k\,(t-u)}\,dW_u^{\mathbb{Q}}. \tag{2.15}$$

See Figure 2.1 for an example of a possible trajectory of r_t using exact discretization. From equation (2.15) we see that the only element from the filtration $\mathcal{F}_s$ we need in order to say something about the future value of r_t is r_s. This is the Markov property, which, heuristically, is the property that the future trajectory is only dependent on today's state. To derive the ZCB price we are going to use this useful fact.

There is two ways of deriving a pricing formula. One way is to use the fact that a Vasicek model is within the Affine model framework and solve the term structure equation. This

[7]Recall that the ZCB-price is seen as a derivative w.r.t. the contract $\Phi(r_T)$

```
T = 2; s = 0; r = 0.03; θ = 0.06; k = 0.01; σ = 0.04; δ = 0.01;
sum[δ_, T_, s_, σ_] := Table[σ Exp[-k t] RandomVariate[NormalDistribution[0, Sqrt[δ]]], {t, (T - s)/δ}];
Accsum = Accumulate[sum[δ, T, s, σ]];
shortrate[T_, s_, r_, θ_, k_, δ_] := Table[r Exp[-k t] + θ (1 - Exp[-k t]) + Accsum[[IntegerPart[t/δ]]],
   {t, s + δ, T, δ}];
Interestratecurve = Transpose[{Table[i, {i, 0, T, δ}], Prepend[shortrate[T, s, r, θ, k, δ], r]}];
ListLinePlot[Interestratecurve, ImageSize → Large, PlotRange → {{0, T}, {0, 0.07}}, LabelStyle → Bold,
 PlotStyle → Black, FrameStyle → Thick]
```

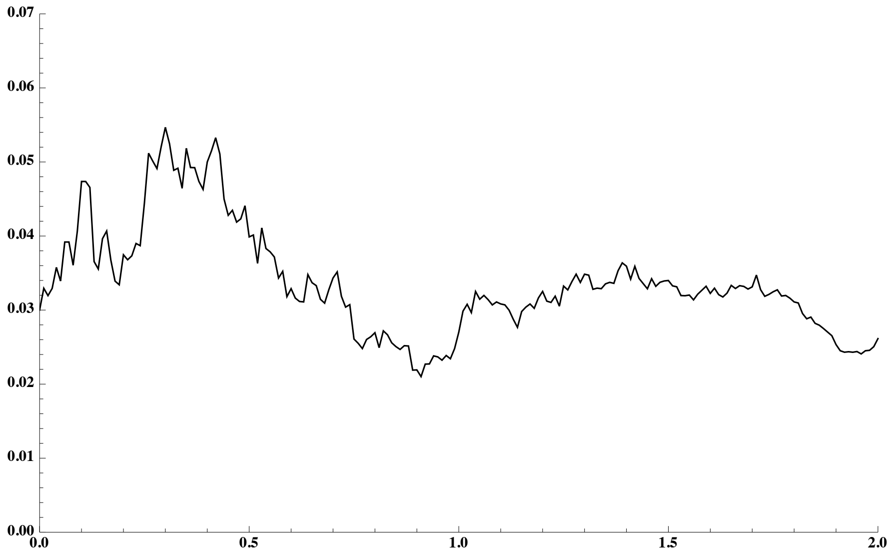

Figure 2.1: *Program for one possible path of the Vasicek model*

approach we use for the CIR model, while for the Vasicek we use the Feynman-Kac representation formula. Let the contract be $\Phi(r_T) = 1$. Then from Proposition 2.8 the ZCB price $P(t, T)$ can be derived by solving the expectation

$$P(t, T) = E_{\mathbb{Q}}\Big[\frac{B_t}{B_T}\, 1 \,|\, \mathcal{F}_t\Big]. \tag{2.16}$$

By putting in for the normalizer, B_t, and the short rate, r_t, we derive the following integrand, using the Markov property

$$P(t, T) = E_{\mathbb{Q}}\Big[\exp\Big\{-\int_t^T [e^{k\,(t-s)}\, r_t + \frac{1}{k}\,\theta\,(1 - e^{k(t-s)}) + \int_t^s \sigma\, e^{k\,(u-s)}\, dW_u^{\mathbb{Q}}]ds\Big\}|r_t\Big].$$

Because of the first part being $\mathcal{F}_t$-measurable and deterministic we simplify the solution to

$$P(t,T) = \exp\Big\{-\int_t^T e^{k\,(t-s)}\, r_t + \frac{1}{k}\,\theta\,(1 - e^{k(t-s)}) ds\Big\} E_{\mathbb{Q}}\Big[\exp\{-\int_t^T \int_t^s \sigma\, e^{k\,(u-s)}\, dW_u^{\mathbb{Q}} ds\}\Big].$$

The first integral is easily solvable. We therefore approach the solution of the second term. To find an solution we need to use the stochastic Fubini theorem and the moment generating function for a Gaussian distributed random variable. Firstly, by the stochastic Fubini theorem, we can rewrite the stochastic-deterministic integral as

$$\begin{aligned} \int_t^T \int_t^s \sigma\, e^{k\,(u-s)}\, dW_u^{\mathbb{Q}} ds &= \int_0^T \int_0^T 1_{[t,T]}(s) 1_{[t,s]}(u) \sigma\, e^{k\,(u-s)}\, dW_u^{\mathbb{Q}}\, ds \\ (S.Fubini) &= \int_0^T \int_0^T 1_{[t,T]}(s) 1_{[t,s]}(u) \sigma\, e^{k\,(u-s)}\, ds\, dW_u^{\mathbb{Q}}, \end{aligned}$$

where we change the integrand in the following matter

$$1_{[t,T]}(s) 1_{[t,s]}(u) = 1_{[t,T]}(s) 1_{[t,s]}(u) 1_{[t,T]}(u) = 1_{[t,T]}(s) 1_{[u,\infty)}(s) 1_{[t,T]}(u).$$

Hence,

$$\int_0^T \int_0^T 1_{[u,T]}(s) 1_{[t,T]}(u) \sigma\, e^{k\,(u-s)}\, ds\, dW_u^{\mathbb{Q}} = \int_t^T \int_u^T \sigma\, e^{k\,(u-s)}\, ds\, dW_u^{\mathbb{Q}}.$$

The next procedure is to use the moment generating function. The moment generating function for a Gaussian distributed r.v. X is given by

$$M_X(t) \overset{def}{=} E[e^{Xt}] = e^{tE[X] + \frac{1}{2}t^2 Var[X]}.$$

Letting $t = 1$ and $X = \int_t^T \int_u^T \sigma\, e^{k\,(u-s)}\, ds\, dW_u^{\mathbb{Q}}$, which clearly is Gaussian distributed due to the definition of the Itô integral and the deterministic integrand, we find the expectation to be zero[8] and variance

$$Var_{\mathbb{Q}}[-\int_t^T \int_u^T \sigma\, e^{k\,(u-s)}\, ds\, dW_u^{\mathbb{Q}}] = \int_t^T \Big(\int_u^T \sigma\, e^{k\,(u-s)}\, ds\Big)^2 du.$$

[8]The expectation of a Itô integral is zero

Rewriting the integral yields

$$2\int_t^T\int_u^T\int_u^v \sigma^2\, e^{k\,(u-s)}e^{k\,(u-v)}\, ds\, dv\, du = -\frac{\sigma^2}{k}B(t,T)^2 - \frac{\sigma^2}{k^2}(B(t,T)-T+t),$$

where $B(t,T) = \frac{1}{k}\left(1-e^{-k(T-t)}\right)$.

Then we have derived the price of the ZCB given the Vasicek model.

$$P(t,T) = \exp\left\{A(t,T) - B(t,T)\, r_t\right\}, \tag{2.17}$$

where

$$A(t,T) = \left(\frac{\theta}{k} - \frac{\sigma^2}{2k^2}\right)(B(t,T)-T+t) - \frac{\sigma^2}{4k}B(t,T)^2.$$

We see that the ZCB price given the Vasicek model have an Affine structure.

Example 2.1:

Let the time-interval; $[t = 0.2, T = 1]$, interest rate at time t; $r_t = 0.03$, the long run interest rate; $\theta = 0.06$, the speed of the mean reversion; $k = 0.01$ and the volatility be; $\sigma = 0.02$. Then by the program in Figure 2.2 we find the price of the ZCB. $P(t,T) = 0.957893$.

```
B[t_, T_] := 1/k (1 - Exp[-k (T - t)])

A[t_, T_] := (θ/k - σ^2/(2 k^2)) (B[t, T] - T + t) - σ^2/(4 k) B[t, T]^2

P[t_, T_, r_] := Exp[A[t, T] - B[t, T] r]

T = 1; t = 0.2; r = 0.03; θ = 0.06; k = 0.01; σ = 0.02;
P[t, T, r]

0.957893
```

Figure 2.2: *Program for finding the ZCB price using the Vasicek model*

Cox-Ingersoll-Ross model

The $\mathbb{Q}$-dynamics of the CIR model is

$$dr_t = k[\theta - r_t]dt + \sigma\sqrt{r_t}dW_t^{\mathbb{Q}}.$$

The model is constructed s.t. the short rate doesn't become negative. This comes from the fact that the CIR model is connected to a squared Ornstein-Uhlenbeck process

$$dX_t = -\alpha X_t dt + \beta dW_t.$$

We show this fact by using Itô formula for $g(t,x) = \sqrt{x}$, with the underlying process $x = r_t$. Then

$$\begin{aligned}
d\sqrt{r(t)} &= \frac{1}{2}\, r(t)^{-\frac{1}{2}}\, dr(t) - \frac{1}{4}\, r(t)^{-\frac{3}{2}}\, (dr(t))^2 \\
&= \frac{1}{2}\, r(t)^{-\frac{1}{2}} \big(\, k[\theta - r(t)]dt + \sigma\sqrt{r(t)}dW_t^{\mathbb{Q}}\big) - \frac{1}{4}\, r(t)^{-\frac{3}{2}}\, \sigma^2 r(t)dt \\
&= \frac{1}{2} r(t)^{-\frac{1}{2}} (k\,\theta - \frac{1}{2}\sigma^2)dt + \frac{1}{2}\big(- k\sqrt{r(t)}dt + \sigma dW_t^{\mathbb{Q}}\big) \\
(k\,\theta = \frac{1}{2}\sigma^2) &= \frac{1}{2}\big(- k\sqrt{r(t)}dt + \sigma dW_t^{\mathbb{Q}}\big).
\end{aligned}$$

This is obvious an Ornstein-Uhlenbeck process for $X_t = \sqrt{r_t}$, $\alpha = \frac{1}{2}k$ and $\beta = \frac{1}{2}\sigma$. Since an Ornstein-Uhlenbeck obvious is Gaussian distributed we have that the CIR model is some form of a non-central χ^2 distributed r.v. .

Using the fact that the CIR model has an affine term structure,

$$\alpha(t) = -k, \beta(t) = k\theta, \gamma(t) = \sigma^2, \delta(t) = 0,$$

we use Proposition 2.10 to derive the ZCB price. Hence we have the following two systems to solve

$$\begin{aligned}
&B^t(t,T) - kB(t,T) - \tfrac{1}{2}\sigma^2 B(t,T)^2 = -1 \\
&B(T,T) = 0,
\end{aligned}$$

and

$$\begin{aligned}
&A^t(t,T) = k\,\theta B(t,T) \\
&A(T,T) = 0.
\end{aligned}$$

In this case we are going to use (Mathematica) to help us. By Appendix B.1.1 we find that

$$B(t,T) = \frac{2(\mathrm{e}^{h(T-t)} - 1)}{2h + (k+h)(\mathrm{e}^{h(T-t)} - 1)}, \tag{2.18}$$

and

$$A(t,T) = \frac{k\theta(h+k)(T-t) + 2\log[2h] - 2\log[h-k+(h+k)\mathrm{e}^{h(T-t)}]}{\sigma^2}. \tag{2.19}$$

In the same manner as the Vasicek section we provide a program for pricing the ZCB via the CIR-model.

Example 2.2:

Here is a program scheme for the price of a ZCB with the CIR-model as the underlying short rate process. Let the time interval; $[t = 0.2, T = 1]$, short rate at time t; $r_t = 0.03$, mean reversion parameter $\theta = 0.06$; speed of convergence; $k = 0.01$, and the volatility be; $\sigma = 0.02$. Then the price of the ZCB is 0.976193 by Figure 2.3.

```
B2[t_, T_, h_] := 2 (Exp[h (T - t)] - 1) / (2 h + (k + h) (Exp[h (T - t)] - 1))
A2[t_, T_, h_] := 2 k θ (h T + k T + 2 Log[e^(-h T) (e^(h T) (h - k) + e^(h T) (h + k))]) / ((h - k) (h + k)) -
   2 k θ ((h + k) t + 2 Log[e^(-h t) (e^(h t) (h - k) + e^(h T) (h + k))]) / ((h - k) (h + k))
P[t_, T_, h_, r_] := Exp[A2[t, T, h] - B2[t, T, h] r]

T = 1; t = 0.2; r = 0.03; θ = 0.06; k = 0.01; σ = 0.02;
P[t, T, Sqrt[k^2 + 2 σ^2], r]
0.976193
```

Figure 2.3: *Program for finding the ZCB price using the CIR model*

Example 2.3:

For the Vasicek model we provided a plot of the trajectory to the short rate using exact discretization. Instead of exact discretization we are going to use an Euler type of stochastic discretization. The approximation improves for smaller Δt. The scheme is shown in Figure 2.4.

Hull-White model

The Hull-White model is an extension of the Vasicek model where the parameters have time dependence. Recall that the $\mathbb{Q}$-dynamics when using the Hull-White model is on the form

$$dr_t = [\theta_t - a_t r_t]dt + \sigma_t dW_t^{\mathbb{Q}}. \tag{2.20}$$

```
Δt = 0.1; r = 0.03; θ = 0.06; k = 0.05; σ = 0.02; m = 1000;
r = Table[r, {i, m}];
Do[
 ΔW = RandomVariate[NormalDistribution[0, Δt]];
 Δr = k (θ - r[[i - 1]]) Δt + (*Exp[-k(Δt i)]*) σ √r[[i - 1]] ΔW;
 r[[i]] = r[[i - 1]] + Δr;
 , {i, 2, m}]
time = Table[i Δt, {i, 0, m - 1}];
rt = Transpose[{time, r}];
ListLinePlot[rt, ImageSize → Large, PlotStyle → Black, FrameStyle → Thick, LabelStyle → Bold,
 PlotRange → {0, 0.08}]
```

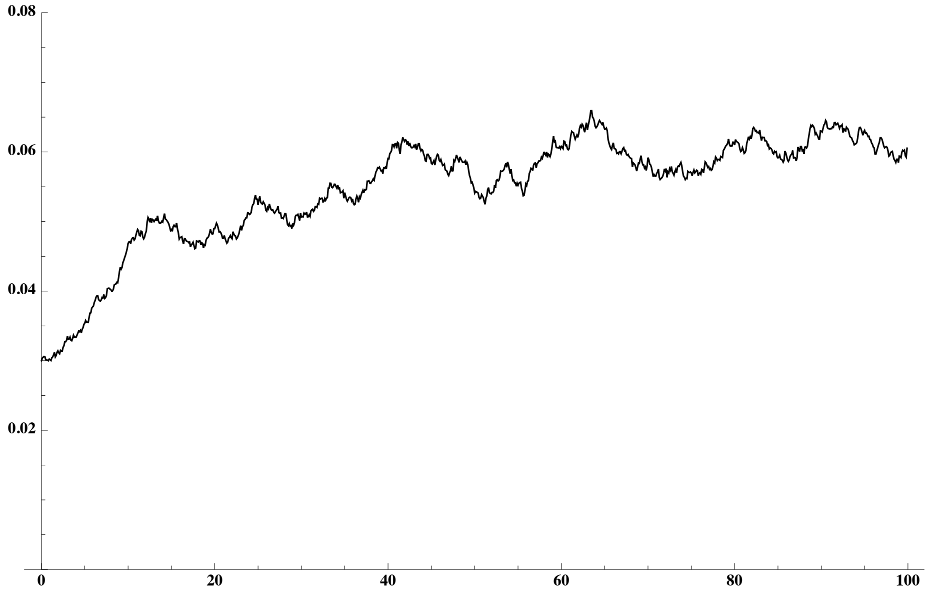

Figure 2.4: *A short-rate trajectory using the CIR model*

The main benefit of this model is that the initial forward (θ_t) and volatility curve (σ_t) can be fitted. But the flexibility has it defiance in no analytical solutions in general [3]. Therefore it is normal to reduce the model to time dependency only on θ_t when investigating the model;

$$dr_t = [\theta_t - a\, r_t]dt + \sigma dW_t^{\mathbb{Q}} \tag{2.21}$$

We easily see that the Hull-White model have affine term structure since we can choose $\alpha(t) = a$, $\beta(t) = \theta_t$, $\gamma(t) = 0$ and $\delta(t) = \sigma$. Solving the first part of the term structure equation yields

$$B(t,T) = \frac{1}{a} + \mathrm{e}^{at}C_1,$$

where, by solving w.r.t. the boundary value $B(T,T) = 0$, we derive following solution

$$B(t,T) = \frac{1}{a}\big(1 - \mathrm{e}^{-a(T-t)}\big).$$

To find $A(t,T)$ we integrate or solve the second term structure equation,

$$A(t,T) = \int_t^T \theta_s \, B(s,T)ds - \frac{1}{2}\sigma^2 \int_t^T B(s,T)^2 ds. \tag{2.22}$$

We want to choose θ_t s.t. it follows the initial forward curve. By the relation,

$$f(0,T) = \frac{\partial \log P(0,T)}{\partial T},$$

and knowing that the Hull-White model has an affine term structure we derive that

$$f(0,T) = \frac{\partial}{\partial T}\big(A(0,T) - r(0)B(0,T)\big). \tag{2.23}$$

Plugging in for $A(0,T)$ and $B(0,T)$

$$\begin{aligned} f(0,T) &= \frac{\partial}{\partial T}\big(\int_0^T \theta_s \, B(s,T)ds - \frac{1}{2}\sigma^2 \int_0^T B(s,T)^2 ds - r(0)\frac{1}{a}\big(1 - \mathrm{e}^{-aT}\big)\big) \\ &= \frac{\partial}{\partial T}\int_0^T \theta_s B(s,T)ds - \frac{1}{2a^2}\sigma^2 \mathrm{e}^{-2aT}\big(\mathrm{e}^{at} - 1\big)^2 - r(0)\mathrm{e}^{-aT} \\ \text{(Appendix B.1.2.1)} &= \int_0^T \theta_s \frac{\partial}{\partial T} B(s,T)ds - \frac{1}{2a^2}\sigma^2 \mathrm{e}^{-2aT}\big(\mathrm{e}^{at} - 1\big)^2 - r(0)\mathrm{e}^{-aT}. \end{aligned}$$

Define

$$\psi(T) \stackrel{def}{=} \int_0^T \frac{\partial}{\partial T}\theta_s B(s,T)ds - r_0 \mathrm{e}^{-aT}$$

and

$$h(T) \stackrel{def}{=} \frac{1}{2a^2}\sigma^2 \mathrm{e}^{-2aT}\big(\mathrm{e}^{at} - 1\big)^2.$$

Then

$$\psi(T) = f(0,T) + h(T).$$

By further calculation in appendix B.1.2.2 we derive that

$$\frac{\partial}{\partial T}\psi(T) = \theta_T - a\psi(T).$$

Putting in for ψ yields that

$$\theta_T = \frac{\partial}{\partial T}\big[(f(0,t) + h(T))\big] + a(f(0,T) + h(T)).$$

In order to calculate θ_T we need to fit the initial forward curve e.g. by a parametric model, and estimate a and σ.

Example 2.4:

We fit the initial forward using the Svensson family. The Svensson family parametric model is defined as [8]

$$f_S(x,z) = z_1 + (z_2 + z_3 x)\mathrm{e}^{-z_4 x} + z_5 x \mathrm{e}^{-z_6 x}. \tag{2.24}$$

Since we observe the yield curves, we use the defined relation

$$Y(t,T) \overset{def}{=} \frac{1}{T-t}\int_t^T f(t,s)ds,$$

and estimate the parameters. We see by Figure 2.5 that we get a fairly good estimate of the parameters. Calculating θ_T, we derive a price for the ZCB given that $a = 0.3$, $\sigma = 0.02$, $t = 0.2$, $T = 2.2$ and the short rate at time 0 is $r_0 = 0.0128$. From Figure 2.6 we see that the price of the ZCB is 0.981973.

```
fS[T_, {z1_, z2_, z3_, z4_, z5_, z6_}] := z1 + (z2 + z3 T) Exp[-z4 T] + z5 T Exp[-z6 T]

YieldCurve[T_, {z1_, z2_, z3_, z4_, z5_, z6_}] :=
  1/T (T z1 + (z2 - e^(-T z4) z2)/z4 + (z3 (1 - e^(-T z4) (1 + T z4)))/z4^2 + (z5 (1 - e^(-T z6) (1 + T z6)))/z6^2)

Y = {1.24, 1.31, 1.32, 1.28, 1.42, 1.73, 2.20, 2.67, 2.91, 3.01} / 100;
TtoM = Table[i , {i, {3/12, 6/12, 9/12, 1, 1 + 2/12, 3 + 2/12, 5 + 2/12, 7 + 2/12, 9 + 2/12, 10}}];

Ymin = NMinimize[Sum[(YieldCurve[TtoM[[i]], {z1, z2, z3, z4, z5, z6}] - Y[[i]])^2, {i, Length[Y]}],
   {z1, z2, z3, z4, z5, z6}];

First[Ymin]

1.06656 × 10^-6

ZEst = {z1, z2, z3, z4, z5, z6} /. Last[Ymin];

Show[Plot[YieldCurve[T, ZEst], {T, 0, 10}, ImageSize → Large, PlotStyle → Black],
 ListPlot[Transpose[{TtoM, Y}], PlotStyle → Black], PlotRange → {0, 0.035}, AxesOrigin → 0,
 LabelStyle → Bold]
```

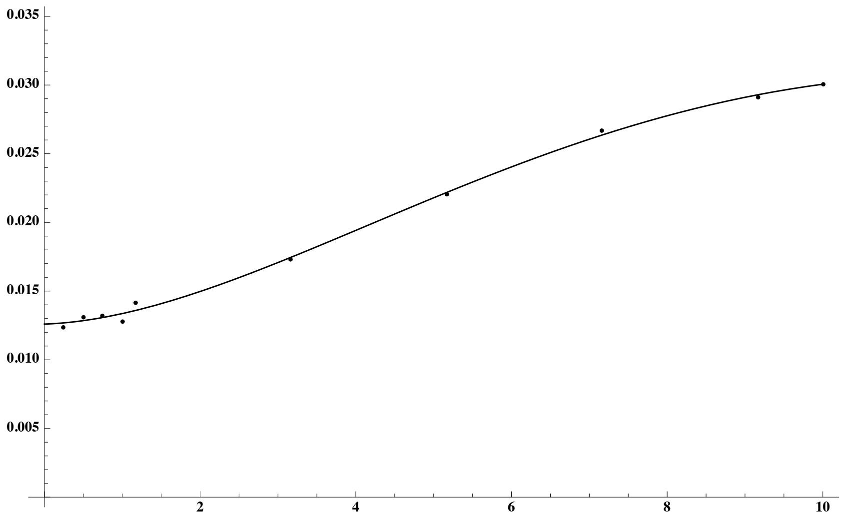

Figure 2.5: *Estimating the Svensson Curve*

```
a = 0.03; σ = 0.02; r = 0.0128;
h[T_] := 1/(2 a^2) σ^2 Exp[-2 a T] (Exp[a T] - 1)^2
θ[T_] := D[fS[T, ZEst] + h[T], T] + a (fS[T, ZEst] + h[T])
B[t_, T_] := 1/a (1 - Exp[-a (T - t)])

A[t_, T_] := Integrate[θ[s] B[s, T], {s, t, T}] - 1/2 σ^2 Integrate[B[s, T]^2, {s, t, T}]
P[t_, T_] := Exp[A[t, T] - r B[t, T]]
P[0.2, 2.2]
0.981973
```

Figure 2.6: *ZCB price using the Hull-White model with the Svensson Family describing the initial forward curve*

Chapter 3

A first look at the HJM-model

The HJM-Model describes the instantaneous forward rate rather than the short rate. This means that the SDE have the following form under the $\mathbb{Q}$-measure,

$$df(t,u) = \alpha(t,u)dt + \sigma(t,u)dW_t^{\mathbb{Q}},$$

or equivalent,

$$f(t,u) = f(0,u) + \int_0^t \alpha(s,u)ds + \int_0^t \sigma(s,u)dW_s^{\mathbb{Q}}.$$

Definition 3.1 (Instantaneous forward rate [1])**:**
The instantaneous forward rate with maturity T, contracted at time t, are defined as

$$f(t,T) \overset{def}{=} -\frac{\partial \log P(t,T)}{\partial T} \tag{3.1}$$

3.1 HJM no-arbitrage drift condition

Using Definition 3.1 the forward rate[1]

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

[1]Instantaneous forward rate

By the First Fundamental theorem we would like the normalized ZCB price, $\tilde{P}(t,T)$, to be a Martingale under the $\mathbb{Q}$-dynamics. Plugging in for the forward rate the normalized ZCB price have the following path

$$\tilde{P}(t,T) = \exp\Big\{ -\big(\int_0^t f(s,s)ds + \int_t^T f(t,u)du \big) \Big\},$$

where $f(s,s) = r_s$ (definition 3.5). Using Fubini and Stochastic Fubini we derive that

$$\begin{aligned}
\int_0^t f(s,s)ds &= \int_0^t f(0,s) + \int_0^s \alpha(v,s)dv + \int_0^s \sigma(v,s)dW_v^{\mathbb{Q}} ds \\
(Fubini) &= \int_0^t f(0,s)ds + \int_0^T \int_0^T 1_{[0,t]}(s)1_{[0,s]}(v)\alpha(v,s)\,ds\,dv \\
&\quad + \int_0^T \int_0^T 1_{[0,t]}(s)1_{[0,v]}(u)\sigma(v,s)dsdW_v^{\mathbb{Q}} \\
&= \int_0^t f(0,s)ds + \int_0^t \int_v^t \alpha(v,s)\,ds\,dv + \int_0^t \int_v^t \sigma(v,s)\,ds\,dW_v^{\mathbb{Q}},
\end{aligned}$$

since

$$1_{[0,t]}(s)1_{[0,s]}(v) = 1_{[0,t]}(v)1_{[0,t]}(s)1_{[v,\infty]}(s) = 1_{[0,t]}(v)1_{[v,t]}(s),$$

and

$$\int_t^T f(t,u)du \overset{fub.}{=} \int_t^T f(0,u)du + \int_0^t \int_t^T \alpha(v,u)dudv + \int_0^t \int_t^T \sigma(v,u)dudW_v^{\mathbb{Q}}.$$

Adding the two integrals together yields

$$\begin{aligned}
\int_t^T f(t,u)du + \int_0^t f(u,u)du &= \int_0^T f(0,u)du \\
&\quad + \int_0^t \Big(\int_t^T \alpha(v,u)du + \int_v^t \alpha(v,u)du \Big) dv \\
&\quad + \int_0^t \Big(\int_t^T \sigma(v,u)du + \int_v^t \sigma(v,u)du \Big) dW_v^{\mathbb{Q}},
\end{aligned}$$

which result in the equation

$$\underbrace{\int_0^T f(0,u)du}_{\text{Constant}} + \int_0^t \int_v^T \alpha(u,v)dudv + \int_0^t \int_v^T \sigma(v,u)dudW_v^{\mathbb{Q}}. \tag{3.2}$$

The first term is known (observed in the market), and the two last terms we define as X_t, yielding the dynamics

$$dX_t = \tilde{\alpha}_T(t,t)dt + \tilde{\sigma}_T(t,t)dW_t^{\mathbb{Q}},$$

where $\tilde{\alpha}$ and $\tilde{\sigma}$ are defined as the integrands in equation (3.2). If we take a look at the dynamics of the normalized ZCB price,

$$\tilde{P}(t,T) = \mathrm{e}^{-(C+X_t)},$$

we find, using Itô's formula, that

$$\begin{aligned}
d\tilde{P}(t,T) &= -\tilde{P}(t,T)dX_t + \frac{1}{2}\tilde{P}(t,T)(dX_t)^2 \\
&= -\tilde{P}(t,T)(\tilde{\alpha}_T(t,t)dt + \tilde{\sigma}_T(t,t)dW_t^{\mathbb{Q}}) + \frac{1}{2}\tilde{P}(t,T)\tilde{\sigma}_T(t,t)^2dt \\
&= \tilde{P}(t,T)\big(\frac{1}{2}\tilde{\sigma}_T(t,t)^2 - \tilde{\alpha}_T(t,t)\big)dt - \tilde{P}(t,T)\tilde{\sigma}_T(t,t)dW_t^{\mathbb{Q}}.
\end{aligned}$$

By the First Fundamental theorem the normalized ZCB price need to be a Martingale since we assume arbitrage free prices. Hence

$$\frac{1}{2}\tilde{\sigma}_T(t,t)^2 - \tilde{\alpha}_T(t,t) = 0,$$

yielding the no arbitrage drift condition,

$$\alpha(t,T) = \sigma(t,T)\int_t^T \sigma(t,u)du, \tag{3.3}$$

when differentiating w.r.t. T on both sides.

Theorem 3.2 (HJM no-arbitrage condition)**:**
If we assume no arbitrage, then by the First Fundamental theorem the risk neutral dynamics of $f(t,u)$ is on the form

$$f(t,u) = f(0,u) + \int_0^t \alpha(s,u)ds + \int_0^t \sigma(s,u)dW_s^{\mathbb{Q}}, \tag{3.4}$$

where

$$\alpha(t,T) = \sigma(t,T)\int_t^T \sigma(t,u)du, \tag{3.5}$$

referred to as the HJM no-arbitrage condition.

3.2 The Musiela Parametrization

Instead of using the parametrization time-of-maturity, T, Musiela proposed to use time-to-maturity, $x = T - t$. This yield a slight difference in the forward rate dynamics. Define

$$f_t(x) \stackrel{def}{=} f(t, t+x) \tag{3.6}$$

Because of the Musiela parametrization we get a t-dependence in the second variable. Let $\frac{\partial}{\partial T}$ stand for differentiating w.r.t. to the second variable. Then formally

$$\begin{aligned} df_t(x) = df(t,t+x) &= df(t;t+x) + \frac{\partial}{\partial T}f(t,t+x)dt \\ &= \alpha(t,t+x)dt + \sigma(t,t+x)dW_t^{\mathbb{Q}} + \frac{\partial}{\partial T}f(t,t+x)dt. \end{aligned}$$

We observe that

$$\begin{aligned} df(t,T) &= df_t(T-t) \\ &= df_t(x) - \frac{\partial}{\partial x}f_t(T-t)dt \\ &= \alpha(t,t+x)dt + \sigma(t,t+x)dW_t^{\mathbb{Q}} + \frac{\partial}{\partial T}f(t,t+x)dt - \frac{\partial}{\partial x}f_t(T-t)dt, \end{aligned}$$

which we know is equal to

$$\alpha(t,t+x)dt + \sigma(t,t+x)dW_t^{\mathbb{Q}}.$$

Hence,

$$\frac{\partial}{\partial T}f(t,t+x)dt = \frac{\partial}{\partial x}f_t(T-t)dt.$$

Proposition 3.3 (The Musiela Equation)**:**
Assume that we have the forward rate dynamics under $\mathbb{Q}$ *given as*

$$df(t,T) = \alpha(t,T)dt + \sigma(t,T)dW_t^{\mathbb{Q}}. \tag{3.7}$$

Then by the Musiela parametrization $f_t(x) = f(t, t+x)$ *we have the following dynamic*

$$df_t(x) = \Big[\big(\frac{\partial}{\partial x} f_t(T-t) + \alpha_t(x)\big)dt + \sigma_t(x)dW_t^{\mathbb{Q}}\Big]_{x=T-t}, \tag{3.8}$$

commonly referred to as the Musiela equation, where

$$\begin{aligned}\sigma_t(x) &= \sigma(t, t+x),\\ \alpha_t(x) &= \sigma_t(x)\int_0^x \sigma_t(u)du.\end{aligned}$$

3.3 Choices of Volatility structure

In the forthcoming we are going to show some choices of volatility structure. We start with the Vasicek and Hull-White because the choices of volatility structure are identical in the non t-dependence case. In fact, then the Vasicek model is a specific choice of the initial forward rate.

3.3.1 Vasicek and Hull-White

Assume that $\sigma(t,T) = \sigma e^{-k(T-t)}$. Then the arbitrage-free drift condition provide the following drift

$$\begin{aligned}\alpha(t,T) &= \sigma(t,T)\int_t^T \sigma(t,s)ds\\ &= \sigma e^{-k(T-t)}\int_t^T \sigma e^{-k(s-t)}ds\\ &= -\sigma e^{-k(T-t)}\frac{1}{k}\Big[\sigma e^{-k(s-t)}\Big]_{s=t}^T\\ &= \frac{\sigma^2}{k}e^{-k(T-t)}\big(1 - e^{-k(T-t)}\big).\end{aligned}$$

Putting into the forward rate yields

$$
\begin{aligned}
f(t,T) &= f(0,T) + \int_0^t \frac{\sigma^2}{k} e^{-k(T-s)}\big(1-e^{-k(T-s)}\big)ds + \int_0^t e^{-k(T-s)}\sigma dW_s \\
\text{(Integrating)} &= f(0,T) + \frac{\sigma^2}{k^2}e^{-kT}\big(\frac{e^{-kT}}{2} - 1 - \frac{e^{2kt-kT}}{2} + e^{kt}\big) + \int_0^t e^{-k(T-s)}\sigma dW_s.
\end{aligned}
$$

Using the relation $r_t = f(t,t)$ we deduce that

$$
\begin{aligned}
r_t &= f(0,T) + \frac{\sigma^2}{2k}\big(e^{-2kt} - 2e{-}kt - 1 + 2\big) + \int_0^t e^{-k(t-s)}dW_s \\
&= f(0,t) + \frac{\sigma^2}{2k^2}\big(1-e^{-kt}\big)^2 + \int_0^t e^{-k(t-s)}\sigma dW_s.
\end{aligned}
$$

Define $\phi(t) \stackrel{def}{=} f(0,t) + \frac{\sigma^2}{2k^2}\big(1-e^{-kt}\big)^2$ and $X_t = \int_0^t \sigma e^{ks} dW_s$. Then

$$
r_t = \phi(t) + e^{-kt}X_t.
$$

Using Itô's formula, with X_t as the underlying process yields

$$
\begin{aligned}
dr_t &= (\phi'(t) - kX_t)dt + e^{-kt}dX_t \\
&= (\phi'(t) - k(r(t) + \phi(t))dt + \sigma dW_t \\
&= k\Big(\frac{\phi'(t) - k\phi(t)}{k} - r(t)\Big) + \sigma dW_t \\
&\stackrel{def}{=} k\big(\theta(t) - r(t)\big)dt + \sigma dW_t.
\end{aligned}
$$

This is the Hull-White model. The Vasicek model is the specific choice of $\theta(t) = \theta$. From earlier deductions we know that, given the specific choice of mean reversion,

$$
r_t = e^{-kt}r_0 + \theta(1-e^{-kt}) + \int_0^t \sigma e^{-k(t-u)}dW_u.
$$

Hence if we choose

$$
f(0,t) = r_0 e^{-kt} + \theta(1-e^{-kt}) - \frac{\sigma^2}{2k^2}\big(1-e^{-kt}\big)^2
$$

we derived the Vasicek model. An another way of deriving the same initial forward rate is provided in appendix B. There we solve the differential equation

$$\theta(t) = \theta$$

w.r.t. the initial forward rate.

3.3.2 CIR

Letting $\sigma(t,T) = e^{-k(T-t)}\sqrt{r(t)}\sigma$ we derive the CIR-model when specifying the initial forward rate structure. Using the arbitrage-free drift condition the drift in the model becomes

$$\begin{aligned}
\alpha(t,T) &= \sigma(t,T)\int_t^T \sigma(t,u)du \\
&= e^{-k(T-t)}\sqrt{r(t)}\sigma\int_0^t e^{-k(u-t)}\sqrt{r(t)}\sigma du \\
&= e^{-k(T-t)}r(t)\sigma^2\frac{1}{k}(1-e^{-k(T-t)}).
\end{aligned}$$

Putting in for the the drift we find the forward rate,

$$f(t,T) = f(0,T) + \int_0^t e^{-k(T-s)}r_s\frac{\sigma^2}{k}\left(1-e^{-k(T-s)}\right)ds + \int_0^t e^{-k(T-s)}\sigma\sqrt{r(s)}dW_s,$$

and the short-rate,

$$r_t = f(0,t) + \int_0^t e^{-k(t-s)}r_s\frac{\sigma^2}{k}\left(1-e^{-k(t-s)}\right)ds + \int_0^t e^{-k(t-s)}\sigma\sqrt{r(s)}dW_s.$$

Defining

$$\phi(t) = f(0,t) + \int_0^t e^{-k(t-s)}r(s)\frac{\sigma^2}{k}\left(1-e^{-k(t-s)}\right)ds$$

and letting

$$X_t = \int_0^t e^{ks}\sigma\sqrt{r(s)}dW_s$$

yields the same procedure as above. By Itô's formula

$$dr(t) = k\Big(\frac{\theta'(t) - k\theta(t)}{k} - r(t)\Big)dt + \sqrt{r(t)}\sigma dW_t.$$

Then solving the differential equation

$$\frac{\theta'(t) - k\theta(t)}{k} = \theta$$

w.r.t. the initial forward rate yields the CIR model.

3.4 Calibration of the forward curve: An introduction

In the market we observe bond prices. The first task is to derive the *yield-to-maturity* (YTM) curve. The YTM is the annual return of holding a bond. To calculate the YTM we need (in parentheses Norway)

- Maturity date
- Settlement date (trade date + 3 working days)
- Bond prices
- Face Value (100)
- Coupon rate
- Coupon interval (Annual)
- Day convention (Actual/365)

Then the YTM is the solution of

$$PV_B = \frac{FV}{(1+YTM)^{\frac{M-1}{I}+\frac{n}{365}}} + \sum_{i=0}^{M-1} \frac{C_i FV}{I(1+YTM)^{\frac{i}{I}+\frac{n}{365}}},$$

where PV_B is the present value of the bond(observed price), FV is the face value, C_i is the coupon rate, I is the coupon interval, M is the number of payments until maturity and n is the number of days until the first payment. When we are solving w.r.t. YTM, numerical approaches is preferable.

When the YTM is calculated we want to find the continuously compounded spot rate defined as

$$Y(t,T) = -\frac{\log[P(t,T)]}{(T-t)},$$

where $P(t,T)$ is the ZCB price. The reason we want to find the continuously compounded spot rate is because of the relation

$$Y(t,T) = \frac{1}{T-t}\int_t^T f(t,s)ds.$$

When we have found the YTM's we can easily calculate the related ZCB price by

$$PV_{ZCB} = \frac{1}{(1+YTM)^{\frac{m}{365}}},$$

where m is the number of days between the settlement date and maturity date. Having the ZCB price we derive the observed continuously compounded spot rate

$$Y^{obs} = -\frac{\log[PV_{ZCB}]}{\frac{m}{365}}.$$

A first time calibration:

Recall that the forward rate is the solution

$$f(t,T) = f(0,T) + \int_0^t \alpha(s,T)ds + \int_0^t \sigma(s,T)dW_s^{\mathbb{Q}},$$

where $\alpha(t,T)$ follow the arbitrage-free drift condition. Because of the arbitrage drift condition we only need to calibrate the initial forward curve $f(0,T)$ and the volatility structure $\sigma(t,T)$. We calibrate the initial forward rate due to today's forward curve, using e.g. smoothing splines, Svensson curve or Nelson-Siegel curve. In Norway the Svensson curve is used [3].

Because of Girsanov's theorem the volatility structure is equal under the objective probability measure $\mathbb{P}$ and under the Equivalent Martingale measure $\mathbb{Q}$. Therefore the volatility structure $\sigma(t,T)$ can be estimated through historical data. We have observed the continuously compounded spot rate

$$Y(t,T) = \frac{1}{T-t}\int_t^T f(t,s)ds.$$

Putting in for the forward rate yields

$$= \frac{1}{T-t}\int_t^T \big(f(0,s) + \int_0^t \alpha(u,s)du + \int_0^t \sigma(u,s)dW_u\big)ds.$$

After calculating the continuously compounded spot rate for the time-to-maturity τ_k we have a data set

$$x_j(\tau_k) = Y(t_j, t_j + \tau_k),$$

for each time of observation $t_1, t_2, \ldots, t_J$. Then it is two ways of estimating the volatility structure. If the time-series $\{x_j(\tau_k)\}_{j=1}^J$ have a small auto-correlation we can use the pure observations. But if there are a clear auto-correlation, a common approach [3] is to estimate the volatility structure w.r.t. the increments

$$\Delta x_j(\tau_k) = x_{j+1}(\tau_k) - x_j(\tau_k),$$

where $t_{j+1} = t_j + \delta$.

If there is a small auto-correlation we see that the volatility structure of the continuously compounded spot rate is

$$Var[Y(t,T)] = \frac{1}{(T-t)^2}Var[\int_0^t \int_t^T \sigma(u,s)dW_u ds].$$

Using stochastic Fubini theorem

$$
\begin{aligned}
&= \frac{1}{(T-t)^2} Var[\int_0^t \int_t^T \sigma(u,s) ds dW_u] \\
&= \frac{1}{(T-t)^2} \int_0^t E\Big[\Big(\int_t^T \sigma(u,s) ds\Big)^2\Big] du,
\end{aligned}
$$

which for a deterministic volatility structure is

$$
\frac{1}{(T-t)^2} \int_0^t \Big(\int_t^T \sigma(u,s) ds\Big)^2 du.
$$

Example 3.1:

Assume the Vasicek/Hull-White volatility structure. Then

$$
\sigma(u,s) = \sigma \mathrm{e}^{-k(s-u)}.
$$

Integrating the inner integral yields

$$
\int_t^T \sigma(u,s) ds = -\frac{1}{k}\sigma\Big(\mathrm{e}^{-k(T-u)} - \mathrm{e}^{-k(t-u)}\Big).
$$

By squaring the inner integral and integrating the outer integral,

$$
\int_0^t \frac{1}{k^2}\sigma^2(\mathrm{e}^{-k(T-u)} - \mathrm{e}^{k(t-u)})^2 du = \frac{\sigma^2}{2k^3}\mathrm{e}^{-2k(T+t)}(\mathrm{e}^{2kt}-1)(\mathrm{e}^{kT}-\mathrm{e}^{kt})^2,
$$

we find the variance of the continuously compounding spot rate that we are estimating,

$$
Var[Y(t,T)] = \frac{1}{(T-t)^2}\frac{\sigma^2}{2k^3}\mathrm{e}^{-2k(T+t)}(\mathrm{e}^{2kt}-1)(\mathrm{e}^{kT}-\mathrm{e}^{kt})^2
$$

Note that, if we have day-to-day observations of the volatility, $t = \frac{1}{365}$.

◇

If there are some auto-correlation we continue estimating the volatility structure based on the increments of the observed continuously compounded spot rate. On increment form the continuously compounded spot rate take the form [1]

$$dY(t,T) = \frac{1}{T-t}\Big(r_t\,dt + \int_t^T \alpha(t,s)ds\,dt + \int_t^T \sigma(t,s)ds\,dW_t\Big).$$

Discretizing the time-steps yields

$$\Delta x_j(\tau) = \frac{1}{\tau}\Big(r_{t_j}\Delta t_j + \int_{t_j}^{t_j+\tau} \alpha(t_j,s)ds\Delta t_j + \int_{t_j}^{t_j+\tau} \sigma(t_j,s)ds\Delta W_{t_j}\Big),$$

where $\Delta t_j = \delta$. We then find that

$$\begin{aligned}
Var[\Delta x_j(\tau)] &= Var\Big[\frac{1}{\tau}\Big(r_{t_j}\Delta t_j + \int_{t_j}^{t_j+\tau} \alpha(t_j,s)ds\Delta t_j + \int_{t_j}^{t_j+\tau} \sigma(t_j,s)ds\Delta W_{t_j}\Big)\Big] \\
&= \frac{1}{\tau^2}\Big(\int_{t_j}^{t_j+\tau} \sigma(t_j,s)ds\Big)^2 Var[\Delta W_{t_j}] \\
&= \frac{1}{\tau^2}\Big(\int_{t_j}^{t_j+\tau} \sigma(t_j,s)ds\Big)^2 \Delta t_j,
\end{aligned}$$

for a deterministic volatility structure.

Estimating procedure using PCA

The estimating procedure of the volatility structure is as follows:

- Use Principal Component Analysis to find an approximation (the important components) of time-to-maturity specific risk
- Choose and fit volatility structures for the important volatility components

A market observation is that there are time-to-maturity specific risk. In the forthcoming chapters we are going to present models of possible infinite time-to-maturity. If we follow the market observation we then have infinite dimension of noise. Clearly we need to reduce the amount of dimensions, and our tool is Principal Component Analysis. Principal Component Analysis within interest rate theory is presented in both [3] and [8]. PCA is based on the spectral decomposition theorem.

Assume that we can decompose the vector into

$$X = \mu + AY = \mu + \sum_{i \geq 1} Y_i a_i,$$

where

$$E[X] = \mu \quad \text{and} \quad Cov[X] = Q = ALA^T,$$

$$E[Y] = 0 \quad \text{and} \quad Cov[Y] = L,$$

and ALA^T is the spectral decomposition with L as the diagonal matrix of eigenvalues, $DiagM\{\lambda_i\}_{i \geq 1}$. Estimating the covariance matrix Q based on K time-to-maturities we have by the spectral decomposition that

$$Q = \sum_{k=1}^{K} \lambda_i a_i a_i^T,$$

where a_i's are the vector elements in matrix A. A good property of the decomposition presented is that the total variance of X is equal to the total variation of Y

$$\sum_{k=1}^{K} Var[X_k] = \sum_{k=1}^{K} Var[Y_k] = tr[L].$$

This means that for a $d \leq K$ we describe

$$\frac{\sum_{k=1}^{d} Var[Y_k]}{tr[L]}$$

amount of the total variation by the d first components. Observing that the d first components describe more than e.g. 99% of the variance we can approximate the covariance matrix

$$Q \approx Q^{approx} = \sum_{i=1}^{d} \lambda_i a_i a_i^T.$$

We estimate the volatility structure of d-th component minimizing e.g. by the least square

$$\sum_{k=1}^{K} \left(\sqrt{\lambda_k a_k a_k^T} - \frac{1}{\tau_k} \int_{t_j}^{t_j + \tau_k} \sigma(t_j, s) ds \right)^2 \to \min_{\sigma}.$$

Hence we approximate the infinite dimension of noise term by

$$df(t,T) = \alpha(t,T)dt + \sigma_1(t,T)dW_t^1 + \sigma_2(t,T)dW_t^2 + \cdots + \sigma_d(t,T)dW_t^d,$$

where W_t^k, for $k = 1, \ldots, d$, are independent Brownian Motion. Then the arbitrage drift condition is

$$\alpha(t,T) = \sum_{k=1}^{d} \sigma_k(t,T) \int_t^T \sigma(t,s)ds.$$

Within interest rate theory three components is usually sufficient for an good approximation of the volatility structure. A code for the Principal Component Analysis for the first example in [8] is presented in Chapter 7.

Chapter 4

Infinite Dimensional Stochastic Analysis

There are two aspects concerning infinite dimensional modeling of the term structure. The first aspect is modeling of the dynamics in time and space with time-to-maturity as the space variable. Then we get an infinite maturity horizon. This problem we already approached under the first view of HJM-modeling through the Musiela parametrization.

A market view is that there exists maturity specific noise. With an infinite maturity horizon we need infinitely many sources of noise. This actually leads to the field of infinite dimensional stochastic analysis.

The finite dimensional stochastic model has the following shortcoming from the viewpoint of a fixed trader. From a complete market[1] point of view we can by a finite dimension of noises perfectly hedge (i.e. replicate) a, e.g., call option on a bond with $x = 5$ years by means of a bond with $x = 30$ years. This contradicts market observations. The risk we don't take into account is the "maturity specific risk". Hence we would like to model the maturity specific risk. The solution is a stochastic partial differential equation with infinite dimensional noise. I.e. the instantaneous forward rate is modeled in the following

[1] We have an incomplete market and make it complete through using internal relations between the same type of bonds(e.g. ZCB with its derivatives and maturities)

way

$$df_t(x) = \left(\frac{\partial}{\partial x} f_t(x) + \alpha_t(x)\right)dt + \sum_{i\in\mathbb{N}} \sigma_t^{(i)}(x)dW_t^{(i)}, \tag{4.1}$$

where $\{W_t^{(i)}\}_{i\geq 1}$ are independent Brownian motions and each represent "maturity specific risk".

4.1 Cylindrical Brownian Motion (CBM)

In this part we want to generalize infinite dimensional stochastic partial differential equations on the form

$$dX(t) = (drift)dt + \sum_{i\in\mathbb{N}} \sigma_t^{(i)}(x)dW_t^{(i)}. \tag{4.2}$$

E.g. choosing $\sigma_t^{(i)} \equiv 1$ the variance

$$\begin{aligned} Var[X(t)] &\geq Var[\int_0^t \sum_{i\in\mathbb{N}} \sigma_s^{(i)}(x)dW_s^{(i)}] \\ &= \sum_{i\in\mathbb{N}} Var[W_t^i] \\ &= \sum_{i\in\mathbb{N}} t = \infty. \end{aligned}$$

Therefore we need to introduce a framework for the study of SPDE's. The solution space for SPDE's is a separable Hilbert space, on which we are going to put additional constraints.

Definition 4.1 (Hilbert Space)**:**
A Hilbert space H is a vector space with an inner product

$$\langle\cdot,\cdot\rangle : H \times H \mapsto \mathbb{R},$$

where the inner product has the following properties

1. $\langle x+y, z\rangle_H = \langle x, z\rangle_H + \langle y, z\rangle_H$ where $x, y, z \in H$ (Linearity 1)

2. $\langle \alpha x, z\rangle_H = \alpha\langle x, z\rangle_H$ where $x, z \in H$ and $\alpha \in \mathbb{R}$ (Linearity 2)

3. $\langle \alpha x, z\rangle_H = \langle \alpha z, x\rangle_H$ where $x, z \in H$,

s.t. H is complete w.r.t. the norm

$$\|x\| \overset{def}{=} \sqrt{\langle x, x\rangle}.$$

We remark that if H is complete then for each Cauchy sequences there exists an $x \in H$ which the sequence converges too. I.e. if $\|x_n - x_m\| \to 0$ is a Cauchy sequence there exists an $x \in H$ s.t. $\|x_n - x\| \to 0$ when $n \to \infty$.

Definition 4.2 (Separable Hilbert Space)**:**
A Hilbert space H is called separable if it exists a dense countable subset $\{y_1, y_2, \dots\}$ of H s.t. for all $\varepsilon > 0$ and all $x \in H$ there exists $y \in \{y_1, y_2, \dots\}$ s.t. $\|y - x\|_H < \varepsilon$.

The choice of Hilbert space becomes clear when we define the Cylindrical Brownian motion. The reason for adding the separability is because we would prefer the Borel σ-algebra[2], $\mathcal{B}(H)$[3], to be equal the σ-algebra generated by "balls". This property is preferable because the σ-algebra generated by the balls simplifies some of the proves. We are not going through those, but rather refer to [8][4].

Theorem 4.3 (ONB Representation theorem)**:**
Let H be a Hilbert space. Then there exists an orthonormal basis (ONB) u_k, $k \geq 1$ of H, i.e.

$$\langle u_i, u_j\rangle_H = \begin{cases} 1, & i = j \\ 0, & i \neq j, \end{cases}$$

s.t. for all $x \in H$ we have the representation

$$x = \sum_{k\in\mathbb{N}} \langle x, u_k\rangle_H u_k.$$

[2]Sometimes referred as σ-field when working with random variable that is an element of $\mathbb{R}$
[3]The Borel σ-algebra is the smallest σ-algebra generated by open balls. A existence of a dense countable subset means that we can create balls that equal the Borel σ algebra
[4]Chapter 3

Furthermore Parseval's equality provides

$$\|x\|_H^2 = \sum_{i\in\mathbb{N}} \langle x, u_i\rangle^2.$$

Before we use the definition of Brownian motion and Theorem (4.3) to define the Cylindrical Brownian motion we define the Cylindrical Gaussian measure.

Definition 4.4 (Cylindrical Gaussian Measure (CG) [4])**:**
Let H be a real separable Hilbert space, then the r.v. $X : H \mapsto L^2(\Omega, \mathcal{F}, \mathbb{P})$ on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a Cylindrical Standard Gaussian if

1. The mapping $X(h)$ is linear; i.e. $X(\alpha k + \beta h) = \alpha X(k) + \beta X(h)$
2. For an arbitrary $h \in H$, $X(h)$ is a Gaussian r.v. with mean zero and variance $\|h\|_H^2$
3. If $h, h' \in H$ are orthogonal, i.e. $\langle h, h'\rangle_H = 0$, then the r.v.s $X(h)$ and $X(h')$ are independent.

We note that by Theorem (4.3), letting $\{u_j\}_{j\geq 1}$ be an orthonormal basis in H and $h \in H$, we can represent $X(h)$ as a $\mathbb{P}$-a.s. convergent series (by kolmogorov three series[5])

$$X(h) = X\big(\sum_{j\in\mathbb{N}} \langle h, u_j\rangle_H u_j\big) = \sum_{j\in\mathbb{N}} \langle h, u_j\rangle_H X(u_j),$$

where in the last equality we used the linearity property in Definition (5.4).

Definition 4.5 (Cylindrical Brownian motion [4])**:**
A family $\{W_t\}_{t\geq 0}$ defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ is called a Cylindrical Brownian motion in a Hilbert space H if

1. For an arbitrary $t \geq 0$ the mapping $W_t : H \mapsto L^2(\Omega, \mathcal{F}, \mathbb{P})$ is linear.
2. For an arbitrary $h \in H$, $W_t(h)$ is an $\mathcal{F}_t$-Brownian motion
   (a) $W_0(h) = 0$ $\mathbb{P}$-a.s.

[5] [4]

(b) For $0 \leq t_1 \leq \cdots \leq t_n = t$ we have that

$$W_{t_1}(h), W_{t_2}(h) - W_{t_1}(h), \ldots, W_t(h) - W_{t_{n-1}}(h)$$

are independent of each other.

(c) For $t \geq s$ we have that

$$W_t(h) - W_s(h) \stackrel{d}{=} W_{t-s}(h)$$

are equally distributed, where $W_{t-s}(h) \sim CG\big[0, (t-s)\|h\|_H^2\big]$

3. For arbitrary $h, h' \in H$ and $t \geq 0$, $E[W_t(h)W_t(h')] = t\langle h, h'\rangle_H$

Because of the linearity we can represent the Cylindrical Brownian motion as a $\mathbb{P}$-a.s. convergent series

$$W_t(h) = \sum_{j\in\mathbb{N}} \langle h, u_j\rangle_H W_t(u_j),$$

where $\{u_j\}_{j\geq 1}$ is an ONB in H and $W_1(u_j)$, for $j \geq 1$, is a sequence of independent standard Gaussian distributed random variables.[6]

In the forthcoming we are going to work with the completed filtration generated by the Cylindrical Brownian motion. I.e.

$$\mathcal{F}_t = \mathcal{N} \cup \sigma\{W_u; 0 \leq u \leq t\},$$

where $\mathcal{N}$ stands for the $\mathbb{P}$-null sets.

4.2 Itô integral w.r.t. Cylindrical Brownian Motion

Our class of function is going to be a subset of Hilbert-Schmidt operators. Therefore we define the Hilbert-Schmidt operators first.

[6] $\|u_j\|_H^2 = 1$

Definition 4.6 (Hilbert-Schmidt operators)**:**

Let H and K be two Hilbert spaces. Then the operator

$$A : H \longrightarrow K$$

with the condition that $\sum_{l \leq 1} \|A(u_l)\|_K^2 < \infty$ for an ONB u_l, $l \geq 1$ in H, is called a Hilbert-Schmidt operator. The class of Hilbert-Schmidt operators from H to K is denoted $HS(H, K)$.

Definition 4.7:

Let $\mathcal{L}(H, K) = \mathcal{L}_{[0,T]}(H, K)$ be the class of functions $f_{(t,\omega)} \in HS(H, K)$, i.e.

$$f_{(t,\omega)} : \mathbb{R}_+ \times \Omega \longrightarrow HS(H, K),$$

with the following properties

1. $(t, w) \mapsto f_{(t,\omega)}$ is $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}$-measurable where $\mathcal{B}$ denotes the Borel σ-algebra and $h \mapsto f_{(\cdot,\cdot)}(h)$ is $\mathcal{B}(HS(H, K))$-measurable
2. $f_{(t,\omega)}$ is $\mathcal{F}_t$-adapted
3. Existence of second moment: $E\big[\int_0^T \|f_t\|_{HS(H,K)}^2 dt\big] < \infty$

In the forthcoming we are going to make sense of the integral

$$\mathcal{I}_C[f](\omega) = \int_0^T f_{(t,\omega)} dW_t, \tag{4.3}$$

where $f_{(t,\omega)} \in \mathcal{L}(H, K)$. Firstly we construct the Itô integral w.r.t. CBM through defining Itô integral w.r.t. CBM on elementary functions[7]. Then expand and show that the elementary function can approximate the functions in $\mathcal{L}(H, K)$. After the procedure we logically define the integral (4.3).

[7]Sometimes called elementary process

An elementary function $\phi \in \mathcal{L}(H, K)$ is defined as

$$\phi_{(t,\omega)} = e_0(w)1_0(t) + \sum_{j=1}^{n} e_j(\omega)1_{(t_j,t_{j+1}]}(t), \tag{4.4}$$

where $0 \leq t_1 \leq \cdots \leq t_n$. We observe that e_j must be $\mathcal{F}_{t_j}$-adapted. For an $h \in H$ we define the Itô integral w.r.t. the CBM in a Hilbert space K[8] for elementary function as

$$(\int_0^t \phi_{(t,\omega)} dW_t)(h) \overset{def}{=} \sum_{j=1}^{n} W_{t_{j+1}\wedge t}[e_j(h)] - W_{t_j \wedge t}[e_j(h)].$$

Using this definition we derive the Itô-isometry for elementary functions.

Lemma 4.8 (Itô-isometry)**:**
If $\phi_{(t,\omega)} \in \mathcal{L}(H, K)$ *is a bounded elementary function then*

$$E\Big[\Big((\int_0^t \phi_s dW_s)[h]\Big)^2\Big] = \int_0^T E[\|\phi_s(h)\|_K^2] ds < \infty \tag{4.5}$$

Sketch of proof. Firstly we use the definition of the Itô integral w.r.t. CBM,

$$E\Big[\Big((\int_0^t \phi_s dW_s)[h]\Big)^2\Big] = E\Big[\Big(\sum_{j=1}^{n} W_{t_{j+1}\wedge t}[e_j(h)] - W_{t_j \wedge t}[e_j(h)]\Big)^2\Big],$$

and divide the summation into two parts,

$$\begin{aligned} &E\Big[\sum_{j=1}^{n} \big(W_{t_{j+1}\wedge t}[e_j(h)] - W_{t_j\wedge t}[e_j(h)]\big)^2\Big] \\ &+ E\Big[\sum_{(j\neq i)=1}^{n} \big(W_{t_{j+1}\wedge t}[e_j(h)] - W_{t_j\wedge t}[e_j(h)]\big)\big(W_{t_{i+1}\wedge t}[e_i(h)] - W_{t_i\wedge t}[e_i(h)]\big)\Big]. \end{aligned}$$

Using the linearity property of expectation, and the property of the CBM we derive the first expectation

$$E\Big[\big(W_{t_{j+1}\wedge t}[e_j(h)] - W_{t_j\wedge t}[e_j(h)]\big)^2\Big] = (t_{j+1} \wedge t - t_j \wedge t) E\big[\|e_j(h)\|_K^2\big].$$

[8] In [4] they have been proper with the definition of CBM and the space that the CBM is defined on. This is why they need to work with adjoint operators

When deriving the second expectation we use the rule of double expectation,

$$E\Big[E\big[\big(W_{t_{j+1}\wedge t}[e_j(h)] - W_{t_j\wedge t}[e_j(h)]\big)\big(W_{t_{i+1}\wedge t}[e_i(h)] - W_{t_i\wedge t}[e_i(h)]\big)|\mathcal{F}_{t_j\vee t_i}\big]\Big].$$

Then $i\wedge j$[9]-term is $\mathcal{F}_{t_j\vee t_i}$-measurable since $i\wedge j+1 \leq j\vee i$ when $i\neq j$, and the $i\vee j$-term is independent of $\mathcal{F}_{t_j\vee t_i}$. Hence

$$E\Big[\underbrace{E\big[\big(W_{t_{j\vee i+1}\wedge t}[e_{j\vee i}(h)] - W_{t_{j\vee i}\wedge t}[e_{j\vee i}(h)]\big)\big]}_{=0 \text{ by CBM}} rest\Big] = 0.$$

Letting $n\to\infty$ we find that

$$\sum_{j=1}^{n}(t_{j+1}\wedge t - t_j\wedge t)E\Big[\|e_j(h)\|_K^2\Big] \longrightarrow \int_0^t E[\|\phi(h)\|_K^2]ds,$$

using linearity properties of the norm. □

The next step is to approximate all functions in $\mathcal{L}(H,K)$ by an bounded elementary function. This is proved through a three steps proof similar to [9].

Proposition 4.9 (The function space)**:**
If $f\in\mathcal{L}(H,K)$, then there exists a sequence of bounded elementary functions ϕ_n, $n\leq 1$ approximating f in $\mathcal{L}(H,K)$, i.e.,

$$\|\phi_n(s) - f_s\|_{\mathcal{L}}(H,K)^2 = E\Big[\int_0^T \|\phi_n(s) - f_s\|^2_{HS(H,K)}dt\Big] \longrightarrow 0$$

as $n\to\infty$.

Proof. [4] □

It is possible to extend the class $\mathcal{L}(H,K)$ to $\mathcal{P}(H,K)$ where the main difference is that instead of assuming existence of second moment we are weakening this property (def.4.7.3) and assume

$$\mathbb{P}\Big[\int_0^T \|f_t\|_H^2 S(H,K)dt < \infty\Big] = 1.$$

[9] $\wedge = inf$, $\vee = sup$

The assumption create a local Itô integral w.r.t Cylindrical Brownian motion, where clearly the relation $\mathcal{L} \subset \mathcal{P}$ holds. In addition we need to work with progressive measurable stochastic process where stopping times is a necessary concept (for a stopped process we know that the second moment exists).

Itô integral w.r.t. independent CBM

Assume that W_t is a CBM in a Hilbert Space K, and let $\{u_l\}_{l\geq 1} \subset K$ be a sequence of ONB in K. Then we have showed that for a $k \in K$

$$W_t(k) = \sum_{l\in\mathbb{N}} \langle k, u_l\rangle_K W_t(u_l).$$

We clearly see that $W_t(u_l)$ is independent and distributed $CG(0, t\langle u_l, u_l\rangle) = CG(0, t)$ which is identical in distribution to a Brownian motion. For simplicity we define $W_t(u_l) = W_t^l$ and refer it as a Brownian motion.

Let $f_t \in \mathcal{L}(H, K)$ and $h \in H$, then we know that there exists a bounded elementary function ϕ_t s.t.

$$\Big(\int_0^t f_s dW_s\Big)(h) \overset{approx}{=} \Big(\int_0^t \phi_s dW_s\Big)(h).$$

From the definition of the Itô integral w.r.t. Cylindrical Brownian motion

$$\Big(\int_0^t \phi_s dW_s\Big)(h) = \sum_{j=1}^{n} W_{t_{j+1}\wedge t}[e_j(h)] - W_{t_j \wedge t}[e_j(h)].$$

Since $e_j(h) \in K$ we use the ONB representation theorem (See 4.3) and rewrite the elementary function for an ONB u_l, $l \geq 1$, as

$$e_j(h) = \sum_{l\in\mathbb{N}} \langle e_j(h), u_l\rangle_K u_l.$$

Using the linearity property of the Cylindrical Brownian motion the sum is

$$
\begin{aligned}
&= \sum_{j=1}^{n} \sum_{l\in\mathbb{N}} \langle e_j(h), u_l\rangle_K (W_{t_{j+1}\wedge t}[u_l] - W_{t_j\wedge t}[u_l]) \\
&= \sum_{l\in\mathbb{N}} \sum_{j=1}^{n} \langle e_j(h), u_l\rangle_K (W^l_{t_{j+1}\wedge t} - W^l_{t_j\wedge t}).
\end{aligned}
$$

Since $\phi_t \in \mathcal{L}(H,K)$ we obviously have that $\langle \phi_t(h), u_l\rangle_K \in \mathcal{V}[0,T]$. Let $\phi^l_t(h) \overset{def}{=} \langle \phi_t(h), u_l\rangle_K$. Then $\phi^l_t(h) \in \mathcal{V}$ is an elementary function since

$$
\begin{aligned}
\phi^l_t = \langle \phi_t(h), u_l\rangle_K &= \langle e_0(h)1_0(t)\rangle_K + \sum_{j=1}^{n} \langle e_j(h)1_{(t_j,t_{j+1}]}, u_l\rangle_K \\
\text{(Inner product prop.)} &= \langle e_0(h)1_0(t), u_l\rangle_K + \sum_{j=1}^{n} \langle e_j(h)1_{(t_j,t_{j+1}]}(t), u_l\rangle_K \\
\text{(Elementary function)} &= \langle e_0(h), u_l\rangle_K 1_0(t) + \sum_{j=1}^{n} \langle e_j(h), u_l\rangle_K 1_{(t_j,t_{j+1}]}(t) \\
&\overset{def}{=} e^l_0(h)1_0(t) + \sum_{j=1}^{n} e^l_j(h)1_{(t_j,t_{j+1}]}.
\end{aligned}
$$

Let $f^l_t \in \mathcal{V}$ be approximated by ϕ^l_t. Then

$$
\begin{aligned}
&= \sum_{l\in\mathbb{N}} \sum_{j=1}^{n} e^l_j(h)\big(W^l_{t_{j+1}\wedge t} - W^l_{t_j\wedge t}\big) \\
\text{(1-dim. Itô integral)} &= \sum_{l\in\mathbb{N}} \int_0^t \phi^l_s(h) dW^l_s \\
\text{(Approx.)} &= \sum_{l\in\mathcal{N}} \int_0^t f^l_s(h) dW^l_s.
\end{aligned}
$$

Hence for an $f_t \in \mathcal{L}(H,K)$, there exists an $f^l_t \in \mathcal{V}[0,T]$ s.t.

$$
\begin{aligned}
\Big(\int_0^t f_t dW_t\Big)(h) &= \sum_{l\in\mathbb{N}} \int_0^t \langle f_t(h), u_l\rangle_K dW^l_s && (4.6) \\
&= \sum_{l\in\mathbb{N}} \int_0^t f^l_s(h) dW^l_s. && (4.7)
\end{aligned}
$$

4.3 Infinite Dimensional Itô Integral w.r.t. CBM

Let H,K be separable Hilbert spaces, $f_t \in \mathcal{L}(H,K)$, W_t a CBM in the Hilbert space K, and $\{g_j\}_{j\geq 1} \subset H$ be an ONB in H. Since $f_t \in \mathcal{L}(H,K)$ there exists an elementary function $\phi_t \in \mathcal{L}(H,k)$ s.t. it approximate f_t. Then

$$\begin{aligned} E\Big[\big(\sum_{i\in\mathbb{N}} \int_0^t \phi_s dW_s(g_i)\big)^2\Big] &\overset{indp}{=} \sum_{i\in\mathbb{N}} E\Big[\big(\int_0^t \phi_s dW_s(g_i)\big)^2\Big] \\ \text{(Multip. with 1)} &= \sum_{i\in\mathbb{N}} E\Big[\big(\int_0^t \phi_s dW_s(g_i)\langle g_i, g_i\rangle_H\big)^2\Big] \\ \text{(Scalar)} &= \sum_{i\in\mathbb{N}} E\Big[\big\langle \int_0^t \phi_s dW_s(g_i)g_i, g_i\big\rangle_H^2\Big]. \end{aligned}$$

On the other hand, by the Itô isometry deduced from the Itô integrals w.r.t. CBM, we know as well that

$$\begin{aligned} \sum_{i\in\mathbb{N}} E\Big[\big(\int_0^t \phi_s dW_s(g_i)\big)^2\Big] &= E\Big[\int_0^t \sum_{i\in\mathbb{N}} \|\phi_s(g_i)\|_K^2 ds\Big] \\ &= E\Big[\int_0^t \|\phi_s\|^2_{HS(H,K)} ds\Big] < \infty \\ &= \|\phi_s\|^2_{\mathcal{L}(H,K)}. \end{aligned}$$

The latter implies that $E\Big[\big\langle \int_0^t \phi_s dW_s(g_i)g_i, g_i\big\rangle_H^2\Big] < \infty$. Using Parseval's equality we have that

$$E\Big[\big\langle \int_0^t \phi_s dW_s(g_i)g_i, g_i\big\rangle_H^2\Big] = E\Big[\big\| \int_0^t \phi_s dW_s(g_i)g_i\big\|_H^2\Big].$$

From this point of view it is suitable to define the infinite dimensional Itô integral w.r.t. CBM for an elementary processes as

$$\int_0^t \phi_s dW_s \overset{def}{=} \sum_{i\in\mathbb{N}} \big(\int_0^t \phi_s dW_s\big)(g_i)g_i$$

Clearly, by the calculations above $\int_0^t \phi_s dW_s \in L^2(\mathbb{P}, H)$. We state the infinite dimensional Itô isometry through the following proposition.

Proposition 4.10 (Infinite Dimensional Itô Isometry)**:**
Let $\phi_t \in \mathcal{L}(H,K)$ be a bounded elementary function, then

$$\|\int_0^t \phi_s dW_s\|_{L^2(\Omega,H)} = \|\phi_s\|_{\mathcal{L}(H,K)}.$$

Proof. The calculations above. □

Recall that a function $f_t \in \mathcal{L}(H,K)$ can be approximated by a bounded elementary function $\phi_t \in \mathcal{L}(H,K)$. Using the definition of the Itô integral w.r.t. CBM we derive that for a bounded elementary function

$$\sum_{i\in\mathbb{N}}(\int_0^t \phi_s dW_s)(g_i)g_i = \sum_{i\in\mathbb{N}}\sum_{j=1}^{n}\big(W_{t_{j+1}\wedge t}[e_j(g_i)] - W_{t_j\wedge t}[e_j(g_i)]\big)g_i$$

We see that the infinite dimensional Itô integral is a linear function. Using this property we can derive the Itô integral w.r.t. CBM. For a function $f_t \in \mathcal{L}(H,K)$ and $h \in H$ we have that

$$\begin{aligned}
\big(\int_0^t f_s dW_s\big)(h) &= \big(\sum_{i\in\mathbb{N}}\big(\int_0^t \phi_s dW_s\big)(g_i)g_i\big)(h)\\
\text{(Linearity)} &= \big\langle \sum_{i\in\mathbb{N}}\big(\int_0^t \phi_s dW_s\big)(g_i)g_i, h\big\rangle_H\\
&= \big\langle \sum_{i\in\mathbb{N}}\sum_{l\in\mathbb{N}}\int_0^t \langle\phi_s(g_i), u_l\rangle_K dW_s^l\, g_i, h\big\rangle_H\\
\text{(Scalar linearity)} &= \big\langle \sum_{l\in\mathbb{N}}\int_0^t \sum_{i\in\mathbb{N}}\langle\phi_s(g_i), u_l\rangle_K g_i dW_s^l, h\big\rangle_H\\
\text{(Adjoint operator)} &= \big\langle \sum_{l\in\mathbb{N}}\int_0^t \sum_{i\in\mathbb{N}}\langle g_i, \phi_s^*(u_l)\rangle_H g_i dW_s^l, h\big\rangle_H\\
\text{(ONB representation thm.)} &= \big\langle \sum_{l\in\mathbb{N}}\int_0^t \phi_s^*(u_l) dW_s^l, h\big\rangle_H\\
\text{(Scalar, Linearity)} &= \sum_{l\in\mathbb{N}}\int_0^t \langle\phi_s^*(u_l), h\rangle_H dW_s^l\\
&= \sum_{l\in\mathbb{N}}\int_0^t \langle\phi(h), u_l\rangle_K dW_s^l.
\end{aligned}$$

4.4 Itô's formula

Here we mainly use Theorem (2.10) in [4], but we work with the Itô integral w.r.t CBM instead of the local version.

Definition 4.11 (Infinite-dimensional Itô process)**:**
Let H and K be separable Hilbert spaces, and $\{W_t\}_{0\geq t\geq T}$ be a K-valued Cylindrical Brownian motion on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\geq t\geq T}, \mathbb{P})$. Then the infinite-dimensional Itô process is a stochastic process X_t on $(\Omega, \mathcal{F}, \{\mathcal{F}\}_{0\geq t\geq T}, \mathbb{P})$ of the form

$$X_t = X_0 + \int_0^t U_{(s,\omega)} ds + \int_0^t V_{(s,\omega)} dW_s,$$

where X_0 is an $\mathcal{F}_0$-measurable H-valued r.v., $V_{(s,\omega)} \in \mathcal{L}(H, K)$ and $U_{(s,\omega)}$ is an H-valued $\mathcal{F}_s$-measurable process s.t.[10]

$$\mathbb{P}\Big[\int_0^T \|U_{(s,\omega)}\|_H ds < 1\Big] = 1.$$

Given the infinite dimensional Itô process we can approach Itô's formula.

Theorem 4.12:
Let X_t, an infinite-dimensional Itô process, be the solution to the SDE

$$dX_t = Udt + VdW_t.$$

Further assume that a function $F : [0, T]\times \mapsto \mathbb{R}$ is continuous and its partial derivatives[11] *F_t, F_x, F_{xx} are continuous and bounded on a bounded subset of $[0, T] \times H$. Then the following, Itô formula, holds*

$$\begin{aligned} F(t, X_t) = F(0, X_0) + &\int_0^t F_s(s, X(s)) + \langle F_x(s, X_s), U(s)\rangle_H + \frac{1}{2} tr[F_{xx}(s, X_s)V(s)V(s)^T]ds \\ &+ \int_0^t \langle F_x(s, X(s)), V(s)dW_s\rangle_H \end{aligned}$$

[10] Bochner integrable
[11] Frechet partial derivative

$\mathbb{P}$-a.s. for all $t \in [0,T]$.

Proof. The proof is a special case of the proof provided in [4]. Note that we are working with $V \in \mathcal{L}(H,K)$, while the generalized proof work with $V \in \mathcal{P}(H,K)$. □

4.5 Martingales and Martingale Representation theorem

Heuristically, if a process consist with the property that the best prediction in the future is today's state, the process has the Martingale property. This property is going to be the main subject when we define Martingales in Hilbert spaces.

Definition 4.13 (Martingales in Hilbert Spaces [4])**:**
Let H be a separable Hilbert Space, measurable w.r.t. its Borel σ-algebra $\mathcal{B}(H)$. Fix $T > 0$ and let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\leq T}, \mathbb{P})$ be a filtered probability space. Let $\{M_t\}_{t\leq T}$ be an H-valued process adapted to the filtration $\{\mathcal{F}_t\}_{t\leq T}$ and $E[\|M_t\|_H] < \infty$. Then M_t is called a Martingale if for any $0 \leq s \leq t$,

$$E[M_t|\mathcal{F}_s] = M_s.$$

Let $\mathcal{F}_t^l = \sigma\{W_s(u_l) : s \leq t\}$ for an ONB $\{u_l\}_{l\geq 1} \subset K$ be the smallest filtration generated by the independent CBMs $W_s(u_l)$. Further define $\mathcal{F}_t^W \overset{def}{=} \cup_{l\geq 1}\mathcal{F}_t^l$. This is the filtration generated by all the CBMs. Using this filtration we are ready to present the Martingale representation theorem w.r.t. CBMs.

Theorem 4.14 (Martingale Representation theorem w.r.t. CBMs)**:**
Let K and H be a separable Hilbert space, W_t a CBM in the Hilbert Space K, and M_t a scalar $\mathcal{F}_t^W$-Martingale s.t. $E[M_0^2] < \infty$ for all $t \geq 0$. Then there exists a unique process $f_t \in \mathcal{L}(H,K)$ s.t.

$$M_t = E[M_0] + \Big(\int_0^t f_t dW_t\Big)(h) \tag{4.8}$$

for an $h \in H$.

Recall that this is equivalent with the existence of $f_t^l(h) \stackrel{def}{=} \langle f_t(h), u_l \rangle$, where $f_t^l \in \mathcal{V}$, s.t.

$$M_t = E[M_0] + \sum_{i \in \mathbb{N}} f_t^l(h) dW_t^l. \tag{4.9}$$

Sketch of proof. Because the linear span[12],

$$\text{span}\Big\{ \mathrm{e}^{\int_0^t h(t)dW_t(u_j) - \frac{1}{2}\int_0^t h(t)dt} : h \in L^2([0,T], \mathbb{R}),\, l = 1,2,3,\dots \Big\},$$

is dense in $L^2(\Omega, \mathcal{F}_t^W, \mathbb{P})$ we can use the Martingale representation theorm

$$M_t = E[M_0] + \sum_{j \in \mathbb{N}} \int_0^t f_s^l(h) dW_s^l,$$

where

$$E[M_t^2] = E[M_0]^2 + E[\int_0^t \|f_s(h)\|_K^2]ds < \infty.$$

The latter is because of the assumption that $f_t \in \mathcal{L}(H,K)$, using the Itô isometry. □

Using the Martingale representation theorem w.r.t. CBM we derive the Martingale representation theorem for Martingales in a Hilbert space H.

Theorem 4.15 (Martingale Representation theorem in a Hilbert space)**:**
Let H and K be separable Hilbert Spaces, W_t a CBM in the Hilbert space K, and M_t an H-valued continuous $\mathcal{F}_t^W$-Martingale s.t. $E[\|M_t\|_H^2] < \infty$ for all $t \geq 0$. Then there exists a unique process $f_t \in \mathcal{L}(H,K)$ s.t.

$$M_t = E[M_0] + \int_0^t f_s \, dW_s.$$

Sketch of proof. Since M_t is an H-valued Martingale s.t. $E[\|M_t\|_H^2] < \infty$, we can choose a ONB $\{g_j\}_{j \geq 1} \subset H$ and represent the Martingale w.r.t CBM as

$$\langle M_t, g_i \rangle_H = E[\langle M_0, g_i \rangle_H] + \sum_{l \in \mathbb{N}} \int_0^t \langle f_s(g_i), u_l \rangle dW_s^l.$$

[12] [4],[9]

Assuming that $E[\|M_t\|_H^2] < \infty$ we have by Parseval's equality that $E[\sum_{j\geq 1}\langle M_t, g_j\rangle_H^2] < \infty$. This implies existence of an inner product and we clearly work in a Hilbert space. Hence we can represent

$$M_t = \sum_{j\in\mathbb{N}} \langle M_t, g_j\rangle g_j.$$

Therefore

$$M_t = \sum_{j\in\mathbb{N}} E[\langle M_0, g_j\rangle_H g_j] + \sum_{j\in\mathbb{N}}\sum_{l\in\mathbb{N}} \int_0^t \langle f_s(g_j), u_l\rangle dW_s^l\, g_j.$$

Since $E[\|M_t\|_H^2] < \infty$ for all t, we use the Parseval's equality again and find that

$$E[\sum_{j\in\mathbb{N}} \langle M_0, g_j\rangle_H g_j] = E[M_0].$$

The stochastic sum we have by definition

$$\begin{aligned}\sum_{j\in\mathbb{N}}\sum_{l\in\mathbb{N}} \int_0^t \langle f_s(g_j), u_l\rangle dW_s^l g_j &= \sum_{j\in\mathbb{N}} \Big(\int_0^t f_s dW_s\Big)(g_j) g_j \\ &= \int_0^t f_s dW_s.\end{aligned}$$

Hence

$$M_t = E[M_0] + \int_0^t f_s dW_s$$

for $f_s \in \mathcal{L}(K, H)$. □

4.6 Girsanov's theorem

To extend Girsanov's theorem for Cylindrical Brownian motion, we mainly extend the Novikov condition.

Theorem 4.16 (Girsanov's theorem w.r.t. CBM)**:**
Let H and K be a separable Hilbert space and W_t be a CBM in the Hilbert space K. Then $W_t(u_l)$, $l \geq 1$ is a sequence of independent Brownian motions for an ONB $\{u_l\}_{l\geq 1} \subset K$.

Further let $q_t \in HS(H,K)$ *s.t.*

$$E\Big[\exp\Big\{\frac{1}{2}\int_0^t \|q_t\|^2_{HS(H,K)}ds\Big\}\Big] < \infty,$$

referred to as the extended Novikov condition. Define the probability measure

$$\mathbb{Q}(A) \overset{def}{=} E[1_A Z_T], A \in \mathcal{F}_T^W,$$

where for an $h \in H$ *the likelihood process*

$$Z_t(h) \overset{def}{=} \exp\Big\{\sum_{l\in\mathbb{N}}\int_0^t q_s^l(h)dW_s^l - \frac{1}{2}\int_0^t \|q_s(h)\|_K^2 ds\Big\}.$$

Then the process

$$\hat{W}_t^l \overset{def}{=} W_t^l - \int_0^t q_s^l(h)ds$$

is an independent Brownian motion w.r.t. the new probability measure $\mathbb{Q}$*.*

Proof. Because of the extended Novikov condition we know that Z_T is a Martingale. Then the rest follows from the finite dimensional Girsanov's theorem. □

4.7 Stochastic Fubini

We state the stochastic Fubini theorem for infinite dimensional Cylindrical Brownian motion. The stochastic Fubini is used in the calculations of finding the arbitrage free drift-condition in the generalized HJM-model.

Theorem 4.17 (Stochastic Fubini theorem)**:**
Let W_t *be an infinite dimensional CBM on a filtered probability space* $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\leq T}, \mathbb{P})$*. Then if*

$$\int_G \|\Phi(\cdot,\cdot,g)\|_{\mathcal{L}(H,K)}\mu(dx) < \infty$$

for a finite measurable space $(G, \mathcal{G}, \mu)$*, where*

$$\Phi : \big([0,T]\times\Omega\times G, \mathcal{B}([0,T])\otimes\mathcal{F}_{t<\leq T}\otimes\mathcal{G}\big) \longmapsto (H, \mathcal{B}(H))$$

is a measurable map, and for every $g \in G$, the process $\Phi(\cdot, \cdot, g)$ is $\{\mathcal{F}_t\}_{t \leq T}$-adapted,

$$\int_G \Big(\int_0^T \Phi(t, \cdot, g) dW_t \Big) \mu(dx) = \int_0^T \Big(\int_G \Phi(t, \cdot, g) \mu(dx) \Big) dW_t.$$

Proof. Theorem (2.8) and Corollary (2.3) in [4].

□

4.8 Stochastic Differential Equations

Let W_t be an infinite dimensional Cylindrical Brownian motion in the separable Hilbert space K on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \leq T}, \mathbb{P})$. We are going to study (semi-linear) stochastic partial differential equation of the form

$$\begin{cases} dX_t = [AX_t + F(t, X_t)]dt + B(t, X_t)dW_t \\ X_0 = x_0, \end{cases}$$

where

$$F : \Omega \times [0, T] \times C([0, T], H) \longmapsto H,$$
$$B : \Omega \times [0, T] \times C([0, T], H) \longmapsto HS(H, K),$$

and A is the generator of a strongly continuous semigroup. Recall that under the Musiela parametrization the left shift provides an extra dt term. The operator in front of $f_t(x)$, $\frac{d}{dx}$ can be shown to be the generator of the left shift operator. But in order to get an solution we need the left shift operator holding the conditions of a strongly continuous semigroup.

Definition 4.18 (Strongly continuous semi-group)**:**
Consider a family of linear functions $S_t : K \mapsto K$, where K is a separable Hilbert space, s.t.

1. $S_0(f) = f$

2. $S_s(S_t(f)) = S_{s+t}(f)$, $s, t \geq 0$, $f \in K$

3. $\|S_t(f) - f\|_K \overset{t \to 0^+}{\longrightarrow} 0, f \in K$

4. $\sup_{\|f\|_K \leq 1} \|S_t(f)\|_K < \infty$, $t \geq 0$

Then S_t,$t \geq 0$ is called a *strongly continuous semi-group.*

A strongly continuous semi-group generate a operator A of the stochastic process X_t. We define the generator in the following way.

Definition 4.19 (Infinitesimal generator)**:**
Let S_t, $t \geq 0$ be a strongly continuous semigroup on K. Define the space $D \subseteq K$ by

$$D \overset{def}{=} \Big\{ f \in K : \lim_{t \to 0^+} \frac{S_t(f) - f}{t} < \infty \ w.r.t. \ \|\cdot\|_K \Big\}.$$

Then the infinitesimal generator of S_t is defined by

$$A(f) \overset{def}{=} \lim_{t \to 0^+} \frac{S_t(f) - f}{t}, \quad f \in D.$$

If we let A be the generator of S_t, then the solution, $X(t)$, of the stochastic partial differential equation

$$\begin{cases} dX_t = [AX_t + F(t, X_t)]dt + B(t, X_t)dW_t \\ X_0 = x_0, \end{cases}$$

given the conditions

$$\mathbb{P}\Big(\int_0^t \|X(t)\|^H dt < \infty \Big) = 1$$
$$\mathbb{P}\Big(\int_0^t \|F(t, X_t)\|_H dt < \infty \Big) = 1$$
$$E\Big[\int_0^t \|B(t, X_t)\|^2_{HS(H,K)} \Big] < \infty,$$

is called a *mild solution* if

$$X(t) = S_t x_0 + \int_0^t S_{t-s} F(s, X_s) ds + \int_0^t S_{t-s} B(s, X_s) dW_s$$

Remark 4.20:

One shows that there is a unique solution to X_t if

$$\|F(t,x) - F(t,y)\|_K + \|B(t,x) - B(t,y)\|_{HS(H,K)} \leq C\|x - y\|_K$$

for all $x, y \in K$ and $t \in [0, T]$.

Chapter 5

Generalized HJM framework

We are going to present the HJM framework which follows the market observation, time-to-maturity specific risk. Because we are assuming possible infinite time-to-maturity, we use the framework of Cylindrical Brownian motion. Still we assume that the ZCB price are given by

$$P(t,T) = \exp\Big\{ - \int_t^T f(t,s)ds \Big\},$$

where $f(t,T)$ is the instantaneous forward rate. In the forthcoming we want to model the forward curves,

$$x \mapsto f_t(x) \stackrel{def}{=} f(t,t+x),$$

using the following Hilbert space of functions $K = H_w$, profoundly specified by [2].

5.1 The Generalized HJM framework

Definition 5.1 (Consistent Hilbert space)**:**
Let $w : [0,\infty) \mapsto (0,\infty)$ be increasing functions s.t.

$$\int_0^\infty \frac{x^2}{w(x)} dx < \infty.$$

Then the space H_w defined as the space of functions $f : [0,\infty) \mapsto \mathbb{R}$ with the properties

1. Absolutely continuous
2. $\int_0^{\infty}(\frac{d}{dx}f(x))^2 w(x)dx < \infty$

is a Hilbert space with the inner product

$$\langle f, g\rangle_{H_w} \stackrel{def}{=} f(0)g(0) + \int_0^{\infty} \frac{\partial}{\partial x}[f(x)]\frac{\partial}{\partial x}[g(x)]w(x)dx$$

for function $f, g \in H_w$.

Properties 5.2 (H_w)**:**

The space H_w has following important properties

1. The linear function (Evaluation functional) $\delta_x : H_w \mapsto \mathbb{R}$ is bounded. I.e.

$$|f(x)| \leq C\|f\|_{H_w} = C\big(f(0)^2 + \int_0^{\infty}(\frac{\partial}{\partial x}[f(x)])^2 w(x)dx\big)^{1/2}$$

   for all $f \in H_w$; $\delta_x[f] = f(x)$

2. The linear function (Integration functional) $\mathcal{J}_x : H_w \mapsto \mathbb{R}$ is bounded. I.e.

$$|\mathcal{J}_x[f]| = |\int_0^x f(s)ds| \leq C\|f\|_K$$

   for all $f \in H_w$; $\mathcal{J}_x[f] = \int_0^x f(s)ds$.

3. The left shift operator $S_t : H_w \mapsto H_w$ defined by

$$(S_t f)[x] = f(t + x)$$

   is a strongly continuous semigroup where the generator of S_t is $A = \frac{d}{dx}$.

4. $f(\infty) \stackrel{def}{=} \lim_{x\to\infty} f(x)$ exists for all $f \in H_w$ since

$$f(\infty) = f(0) + \int_0^{\infty} \frac{d}{dx}[f(x)]dx < \infty$$

5. Consider the subspace $H_w^0 \subseteq H_w$ given by

$$H_w^0 \stackrel{def}{=} \Big\{ f \in H_w : f(\infty) = 0 \Big\}.$$

Define the operation $\star$ by

$$(f \star g) \stackrel{def}{=} f(x) \int_0^x g(s) ds.$$

Then

$$\|f \star g\|_{H_w} \leq C \|f\|_{H_w} \|g\|_{H_w}$$

for all $f, g \in H_w^0$

The properties enable us to set up the following model for the forward rate $f_t(x)$ with time-to-maturity specific risk.

Let H be a separable Hilbert space, and W_t a CBM in H_w on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \leq T}, \mathbb{Q})$. Then we define the infinite dimensional HJM model under the Musiela parametrization as the mild solution of the SPDE

$$df_t = [Af_t + \alpha_t(\cdot, f_t)]dt + \sigma_t(\cdot, f_t) dW_t, \ \ \cdot \in \Omega,$$

where

$$\alpha : [0, \infty) \times \Omega \times H_w \longmapsto H,$$
$$\sigma : [0, \infty) \times \Omega \times H_w \longmapsto HS(H, H_w),$$

and A is the generator of the left shift operator S_t. By property (3), S_t being a strongly continuous semi-group, we know that the forward rate process satisfies

$$f_t = S_t f_0 + \int_0^t S_{t-s} \alpha_s(f_s) ds + \int_0^t S_{t-s} \sigma_s(f_s) dW_s. \tag{5.1}$$

For simplicity let $\alpha_s \overset{def}{=} \alpha_s(f_s)$ and $\sigma_s \overset{def}{=} \sigma_s(f_s)$. Because of the linearity property of the evaluation functional

$$\begin{aligned}\delta_x(f_t) &= \delta_x[S_t f_0] + \int_0^t \delta_x[S_{t-s}\alpha_s]ds + \int_0^t \delta_x[S_{t-s}\sigma_s]dW_s \\ &= \delta_x[S_t f_0 + \int_0^t S_{t-s}\alpha_s ds + \int_0^t S_{t-s}\sigma_s dW_s], \mathbb{Q}\text{-a.s.},\end{aligned}$$

not knowing what happens with the $\mathbb{Q}$-null sets.

Assume that the initial forward rate is an element of H_w. Recall that by the definition of the infinite dimensional HJM model as the mild solution of (6.1)

$$\begin{aligned}\mathbb{Q}\Big(\int_0^t \|f_s\|_H ds < \infty\Big) &= 1, \\ \mathbb{Q}\Big(\int_0^t \|\alpha_s\|_H ds < \infty\Big) &= 1,\end{aligned}$$

and $\sigma_t \in \mathcal{L}(H, H_w)$.

In the forthcoming we are going to work with a bond market consisting of a ZCB, priced

$$P(t,T) = \exp\Big\{ - \mathcal{J}_{T-t}[f_t]\Big\},$$

and a risk free normalizer

$$B(t) = \exp\Big\{ \int_0^t \delta_0(f_s)ds\Big\}.$$

5.2 The Arbitrage-free Drift Condition:

By the First Fundamental theorem the discounted bond prices are Martingales under the risk neutral probability measure $\mathbb{Q}$. Because of the existence of a Girsanov transform, we define the HJM-model directly under the $\mathbb{Q}$-measure[1].

[1] Martingale Modeling

Fix t and time-of-maturity T. Then by the linearity property of the integration functional

$$\begin{aligned}
-\log[P(t,T)] &= \mathcal{J}_{T-t}[S_t f_0] + \int_0^t \mathcal{J}_{T-t}[S_{t-s}\alpha_s]ds + \int_0^t \mathcal{J}_{T-t}[S_{t-s}\sigma_s]dW_s \\
(\mathcal{J}_{T-s} = \mathcal{J}_{T+s} - \mathcal{J}_{T-s}) &= \mathcal{J}_T[f_0] - \mathcal{J}_t[f_0] + \int_0^t \mathcal{J}_{T-s}[\alpha_s] - \mathcal{J}_{t-s}[\alpha_s]ds \\
&\quad + \int_0^t \mathcal{J}_{T-s}[\sigma_s] - \mathcal{J}_{t-s}[\sigma_s]dW_s \\
(\text{Reorganize}) &= (\mathcal{J}_T[f_0] + \int_0^t \mathcal{J}_{T-s}[\alpha_s]ds + \int_0^t \mathcal{J}_{T-s}[\sigma_s]dW_s) \\
&\quad - (\mathcal{J}_t[f_0] + \int_0^t \mathcal{J}_{t-s}[\alpha_s]ds + \int_0^t \mathcal{J}_{t-s}[\sigma_s]dW_s) \\
&\overset{def}{=} I_1 - I_2,
\end{aligned}$$

where we use the relation $\mathcal{J}_{T-t}(S_s) = \mathcal{J}_{T+s} - \mathcal{J}_{t+s}$. Since the integration functional $\mathcal{J}_{T-s}$ by the property (2) is a bounded, deterministic linear operator, $\mathcal{J}_{T-s}(\sigma_s) \in HS(H, H_w)$. Furthermore, knowing that $\sigma_s \in \mathcal{L}(H, H_w)$ implies $I_{T-s}(\sigma_s) \in \mathcal{L}(H, H_w)$.

We take some extra interest in I_2, since we, in fact, can show that it is equal to the risk-free normalizer. Define $u = x + s$, then

$$\begin{aligned}
I_2 &= \mathcal{J}_t[\delta_u f_0] + \int_0^t \mathcal{J}_{t-s}[\delta_u \alpha_s]ds + \int_0^t \mathcal{J}_{t-s}[\delta_u \sigma_s]dW_s \\
(S_s\delta_x = \delta_u) &= \mathcal{J}_t[\delta_u f_0] + \int_0^t \mathcal{J}_{t-s}[S_s\delta_x \alpha_s]ds + \int_0^t \mathcal{J}_{t-s}[S_s\delta_x \sigma_s]dW_s \\
(\text{Left-shift}) &= \mathcal{J}_t[\delta_u f_0] + \int_0^t \mathcal{J}_{[s,t]}[\delta_x \alpha_s]ds + \int_0^t \mathcal{J}_{[s,t]}[\delta_x \sigma_s]dW_s \\
(\text{Indicator}) &= \mathcal{J}_t[\delta_u f_0] + \int_0^t \mathcal{J}_t[\delta_x(\alpha_s 1_{[s,t]})1_{[0,t]}(s)]ds + \int_0^t \mathcal{J}_t[\delta_x(\sigma_s 1_{[s,t]})1_{[0,t]}(s)]dW_s.
\end{aligned}$$

From earlier we know that

$$1_{[0,t]}(s)1_{[s,t]}(x) = 1_{[0,t]}(x)1_{[x,t]}(s).$$

Using Fubini and Stochastic Fubini

$$\begin{aligned}
(Fubini) &= \mathcal{J}_t[\delta_u(f_0)] + \mathcal{J}_t[\int_0^u \delta_x \alpha_s ds] + \mathcal{J}_t[\int_0^u \delta_x \sigma_s dW_s] \\
(x = u - s) &= \mathcal{J}_t\Big[\delta_u(f_0) + \int_0^u \delta_{u-s}\alpha_s ds + \int_0^u \delta_{u-s}\sigma_s dW_S\Big] \\
(\text{Left-shift}) &= \mathcal{J}_t\Big[\delta_0\big(S_u + \int_0^t S_{u-s}\alpha_s ds + \int_0^u S_{u-s}\sigma_s dW_s\big)\Big] \\
&= \mathcal{J}_t[\delta_0(f_u)] = \log[B(t)].
\end{aligned}$$

Then we derive the logarithm of the discounted ZCB price as

$$\begin{aligned}
\log[\tilde{P}(t,T)] &= \log\big[\frac{P(t,T)}{B(t)}\big] = \log[P(t,T)] - \log[B(t)] \\
&= -(I_1 - I_2) - \log[B(t)] \\
&= -I_1 + \log[B(t)] - \log[B(t)] \\
&= -I_1.
\end{aligned}$$

Using Itô's formula on the discounted ZCB price yields

$$\begin{aligned}
\tilde{P}(t,T) = \tilde{P}(0,T) &- \int_0^t \mathcal{J}_{T-s}[\alpha_s]\tilde{P}(s,T)ds - \int_0^t \mathcal{J}_{T-s}(\sigma_s)\tilde{P}(s,T)dW_s \\
&+ \frac{1}{2}\int_0^t \|I_{T-s}(\sigma_s)\|^2_{H_w(\mathcal{L}(H,H_w))}\tilde{P}(s,T)ds.
\end{aligned}$$

By the First Fundamental theorem $\tilde{P}(t,T)$ is a Martingale under the $\mathbb{Q}$-measure. Hence

$$\mathcal{J}_{T-s}(\alpha_s) = \frac{1}{2}\|\mathcal{J}_{T-s}(\sigma_s)\|^2_{H_w}.$$

Differentiating on both sides, and applying property 4. and 5., we find the arbitrage free drift condition

$$\alpha_s = \sigma_s \star \sigma_s = \sigma_s \mathcal{J}_x[\sigma_s].$$

If we in addition assume that

$$\|\sigma_t(f) - \sigma_t(g)\|_{HS(H,H_w)} \leq C\|f - g\|_{H_w}$$

we have a unique, continuous solution to

$$df_t = [Af_t + \alpha_t(f_t)]dt + \sigma_t(f_t)dW_t$$

in the no-arbitrage case [8][2].

An unexpected property of this model is that *the long rates never fall.*

Theorem 5.3 (Long rates never fall)**:**
Assume that we work with the framework presented above and especially consider property 4. . If we define the long rate as

$$\ell_t \stackrel{def}{=} f_t(\infty) = \lim_{x\to\infty} f_t(x),$$

then

$$\ell_s \leq \ell_t$$

for $0 \leq s \leq t$.

Sketch of Proof. The discounted ZCB price $\tilde{P}(t,T)$ under the Musiela parametrization has the following relation

$$\tilde{P}(t,T) = \exp\Big\{ - \int_0^t f_s(0) - \int_0^{T-t} f_t(x)dx\Big\},$$

where $f_t(x) = \delta_x f_t$. Consider

$$\big(\tilde{P}(t,T)\big)^{\frac{1}{T}} = \exp\Big\{ - \frac{1}{T}\big(\int_0^t f_s(0) + \int_0^{T-t} f_t(x)dx\big)\Big\}.$$

Letting $T \to \infty$ then

$$\frac{1}{T}\int_0^t f_s(0)ds \to 0,$$
$$\frac{1}{T}\int_0^{T-t} f_t(x)dx \to \ell_t.$$

[2](Proposition 6.2)

Hence

$$\lim_{T\to\infty} (\tilde{P}(t,T))^{\frac{1}{T}} = \mathrm{e}^{-\ell_t}.$$

Then, by following the proof of [8] using the Hölder's inequality[3],

$$E\big[|X||Y||\mathcal{F}\big] \leq E[|X|^p|\mathcal{F}]^{1/p} E[|Y|^q|\mathcal{F}]^{1/q},$$

for $p, q \geq 1$ and $\frac{1}{p} + \frac{1}{q} = 1$, and Fatou's lemma

$$E[\liminf_{n\to\infty} X_n|\mathcal{F}] \leq \liminf_{n\to\infty} E[X_n|\mathcal{F}],$$

where we remark that $\lim_{n\to\infty} X_n = \liminf_{n\to\infty} X_n = \limsup_{n\to\infty} X_n$ when the limit exist, we derive that

$$\mathrm{e}^{-\ell_s} \geq \mathrm{e}^{-\ell_t}.$$

Hence $\ell_t \geq \ell_s$ for $t \geq s$. □

5.3 Generalized Bond Portfolios

We are going to consider portfolios where the investor is allowed to own bonds of any maturity. In order to define a reasonable notion of trading strategy for portfolios with bonds of possible infinite time-to-maturity we would like to redefine the discounted bond price to follow the Musiela parametrization. For a time-to-maturity x, using properties that we have gone through, we can reparameterize the discounted bond price using a left shift operator following the third property.

Definition 5.4 (Generalized model for the discounted ZCB)**:**
Define

$$\tilde{P}_t(x) \stackrel{def}{=} \tilde{P}(t, t+x),$$

[3] Recall that the expectation is an integral

and let $\sigma_t^l(f,x) \in \mathcal{V}[0,T]$, $l \geq 1$, $f,g \in H_w$, $t,x \geq 0$, be a sequence of real valued processes satisfying

$$\begin{aligned}
&(1) \quad \sum_{l\in\mathbb{N}} \|\sigma_t^l(f,x) - \sigma_t^l(g,x)\|_{H_w}^2 \leq C\|f-g\|_{H_w}^2 \\
&(2) \quad \sigma_t^l(f,0) = 0 \\
&(3) \quad \sum_{l\in\mathbb{N}} |\sigma_t^l(f,x)|^2 \leq C|f(x)|^2.
\end{aligned}$$

Then in a arbitrage free market the generalized discounted bond price, $\tilde{P}_t(x)$, is given by the strictly positive, unique mild solution of the SPDE

$$\tilde{P}_t(x) = \tilde{P}_0(x) + \int_0^t A\tilde{P}_s(x)ds + \sum_{l\in\mathbb{N}} \int_0^t \sigma_s^l(\tilde{P}_s(x),x), dW_s^l,$$

where W_t^l, $l \geq 1$, is a sequence of independent risk neutral CBMs, assuming that $\tilde{P}_0 \in H_w$, and $\delta_x\tilde{P}_t = \tilde{P}_t(x)$.

Through the second condition we assume that the volatility of the discounted bond price vanish at time-to-maturity $x = 0$, while the third condition ensures that the discounted bond price is strictly positive. The latter comes from $\sigma_t^l(x,f)$ given condition 3. is linear w.r.t f[4]. Further the choice of the discounted bond price process being the mild solution of the SPDE above ensures that the non-Musiela parametrization, $\tilde{P}(t,T)$, is a Martingale w.r.t. the risk neutral measure $\mathbb{Q}$.

The next step is to generalize the portfolio process, finishing with generalizing the self-financing trading strategy.

5.3.1 Generalized Portfolio Process

Let $c_1, c_2, \ldots, c_N$ be the stock-holding of bonds with the corresponding time-to-maturity $x_1, x_2, \ldots, x_N$ owned by an investor at time t. Then the portfolio value $V_t(\pi)$ for the

[4] $\sigma_t^l(x,f) = \sigma_t^l(x)f_t(x)$, which is an obvious relation when we look at discounted ZCB deduced earlier.

strategy $\pi = (c_1, c_2, \ldots, c_N)$ at a time t is

$$V_t(\pi) = \sum_{i=1}^{N} c_i P_t(x_i).$$

This is equivalent to

$$V_t(\pi) = (\sum_{i=1}^{N} c_i \delta_{x_i})[P_t],$$

where δ_x is the evaluation functional. By the relation presented it seem reasonable to define the generalized portfolio process for a portfolio strategy π as

$$V_t(\pi) \stackrel{def}{=} \pi_t(P_t),$$

where $\pi_t : H_w \to \mathbb{R}$ is a linear functional satisfying

$$\sup_{\|f\|_{H_w} \leq 1} |\pi_t(f)| < \infty,$$

and $\pi_t(f)$ is $\mathcal{F}_t$-adapted for all f. By the linearity assumption we deduce that

$$V_t(\pi) = \pi_t(P_t) = \langle \pi_t, P_t \rangle_{H_w}.$$

Furthermore the discounted portfolio process

$$\begin{aligned} \tilde{V}_t(\pi) &\stackrel{def}{=} B(t)V_t(\pi) = B(t)\pi_t(P_t) \\ (\text{Linearity}) &= \pi_t(B(t)P_t) \\ &= \pi_t(\tilde{P}_t) \end{aligned}$$

is equal to the portfolio strategy of the discounted ZCB price.

5.3.2 Generalized Self-Financing trading strategy

Under a self-financing strategy the portfolio strategy is decided at time 0 and invariant w.r.t. $t > 0$. Therefore it is reasonable to define the Generalized Self-Financing trading strategy as the following.

Definition 5.5 (Generalized Self-Financing strategy)**:**
Let $\pi_t : H_w \mapsto \mathbb{R}$ be a portfolio strategy

$$V_t(\pi) = \pi_t(P_t)$$

s.t.

$$\sum_{l \in \mathbb{N}} E\big[\int_0^t |\pi_s(\sigma_s^l(\tilde{P}))|^2 ds\big] < \infty.$$

Then $\{\pi_t\}_{t \geq 0}$ is Self-Financing if $\pi_t = \pi_0 \stackrel{def}{=} \pi$ for all $t \geq 0$.

From the definition we deduce that the discounted portfolio value at time t is the solution

$$\tilde{V}_t(\pi) = V_0(\pi) + \sum_{l \in \mathbb{N}} \int_0^t \pi[\sigma_s^l(\tilde{P}_s)] dW_s^l \tag{5.2}$$

for a $V_0 \in \mathbb{R}$, where we note that $V_0 = \tilde{V}_0$. If two portfolios are identical for all t and the corresponding portfolio strategy is self-financing, we can show that the two portfolio strategies must be equal. This leads to the theorem about uniqueness of hedging strategy.

Theorem 5.6 (Uniqueness of hedging strategy)**:**
Assume that we are working with the framework presented. Define

$$\mathcal{H} \stackrel{def}{=} \{g \in H_w : g(0) = 0\}.$$

For two self-financing portfolio strategies π_t^1, π_t^2, $0 \leq t \leq T$, *assuming that*

$$\tilde{V}_T(\pi^1) = \tilde{V}_T(\pi^2),$$

then $\pi^1 = \pi^2$.

Proof. Let π_t^1 and π_t^2 be two self-financing portfolio strategies and define the strategy $\tau_T \stackrel{def}{=} \pi^1 - \pi^2$. Assuming that $\tilde{V}_T(\pi^1) = \tilde{V}_T(\pi^2)$ we have that

$$0 = \tilde{V}_T(\tau) = V_0(\pi^1) - V_0(\pi^2) + \sum_{l \in \mathbb{N}} \int_0^T \tau(\sigma_s^l(\tilde{P}_s)) dW_s^l \tag{5.3}$$

Since we know that the portfolio value is going to be zero, the expectation must be equal to zero. Knowing that the expectation of an Itô integral is equal to zero we deduce that

$$0 = E[\tilde{V}_T(\tau)] = V_0(\pi^1) - V_0(\pi^2).$$

But if $V_0(\pi^1) - V_0(\pi^2) = 0$, then by equation (6.3)

$$\sum_{l\in\mathbb{N}} \int_0^T \tau(\sigma_s^l(\tilde{P}_s))dW_s^l = 0.$$

Because W_s^l, $l \geq 1$, is independent CBM, the Itô isometry yields

$$\sum_{l\in\mathbb{N}} E\Big[\int_0^T [\tau(\sigma_s^l(\tilde{P}_s))]^2 ds\Big] = 0.$$

The only solution is that the integrand is equal to zero, hence

$$\tau(\sigma_s^l(\tilde{P}_s)) = 0.$$

On the other hand we know by the Itô isometry that for a function $g \in \mathcal{H}$ there exists a sequence α_l, $l \geq 1$[5], s.t.

$$\{\sum_{l\in\mathbb{N}} \alpha_l \sigma_t^l(f) : \sum_{l\in\mathbb{N}} \alpha_l^2 < \infty, \alpha_l \in \mathbb{R}, t \leq 0, f \in H_w\} = \{f \in H_w : f(0) = 0\}.$$

Then

$$\begin{aligned}
\tau(g) &= \tau(\sum_{l\in\mathbb{N}} \alpha_l \sigma_s^l(\tilde{P}_s)) \\
\text{(Linearity)} &= \sum_{l\in\mathbb{N}} \alpha_l\, \tau(\sigma_s^l(\tilde{P}_s)) \\
&= \sum_{l\in\mathbb{N}} \alpha_l\, 0 = 0,
\end{aligned}$$

implying that $\tau(g) = 0$ for all $g \in \mathcal{H}$.

[5] Comes from the definition of the Consistent Hilbert Space letting f(0)=0

Let $f \in H_w$ and define $g \stackrel{def}{=} f - f(0) \in \mathcal{H}$. Then

$$\begin{aligned} 0 = \tau(g) &= \tau(f) - \tau(f(0)1) \\ (Linearity) &= \tau(f) - f(0)\tau(1), \end{aligned}$$

implying

$$\tau(f) = f(0)\tau_1.$$

If we put $\tilde{P}_s$ in for f we find that

$$0 = \tilde{V}_s(\tau) = \tau(\tilde{P}_s) = \tilde{P}_s(0)\tau(1).$$

We deduced that $\tau(1) = 0$ which implies that $\tau(f) = 0$ for all $f \in H_w$. Hence $\tau = 0$ and we have proved that $\pi^1 = \pi^2$. $\square$

The bond market presented is incomplete even if there exists a unique risk neutral probability measure, $\mathbb{Q}$. This means that we cannot replicate our possible portfolio. However our bond market is approximately complete in the sense that for all $\varepsilon > 0$ there exists an approximate trading strategy π^ε such that the L^2-distance between the portfolio value and the claim is smaller than ε. Hence the market have under this conditions a unique Equivalent Martingale measure, $\mathbb{Q}$. This we are going to use in the upcoming chapter; Stochastic Duration.

Chapter 6

Stochastic Duration

Duration is a well known concept within interest rate theory [5]. Until now it is based on deterministic interest rates. For a complex bond portfolio with options, swaps, caps, and other interest rate derivatives, the duration deterministic based sensitivity analysis is not satisfying since the classical duration concept (Macaulay duration) requires flat or piecewise flat interest rates for its computation and is only applicable to portfolios composed of rather simple interest rate derivatives as e.g. zero-coupon bonds or swaps. In the forthcoming we are going to introduce a concept called Stochastic Duration, which can be in contrast to the Macaulay duration used to measure the sensitivity of complex bond portfolios with respect to stochastic fluctuations of the entire term structure of interest rates or the yield surface. As a special case of this concept we will present the Stochastic Duration based on the Vasicek model. Later on we will derive a numerical estimate of the stochastic duration and provide an immunization strategy for a portfolio.

6.1 Macaulay duration

For a given portfolio of one ZCB contracted with interest rate r, at time t and maturity T, the present value is the discounted future price[1]

$$pv_t = \mathrm{e}^{-\int_t^T r dy} = \mathrm{e}^{-r(T-t)}.$$

An infinitesimal change in r yields

$$\frac{\partial}{\partial r} pv_t = -(T-t)\mathrm{e}^{-r(T-t)}.$$

Dividing on the present value we derive what is called the Macaulay Duration in continuous time

$$\frac{\frac{\partial}{\partial r} pv_t}{pv_t} = -(T-t).$$

Note that this is the same as taking the derivative of the logarithm of the portfolio. We see that the Macaulay Duration is the time-to-maturity. Therefore the name duration. Because of a linearity property, the duration of the portfolio is the sum of weighted durations of bonds in the portfolio. By this reasoning we can interpret the Macaulay duration as the mean time-to-maturity.

Example 6.1:

Assume that we have two ZCBs in our portfolio, contracted with equal interest rate, but different time-to-maturity, T_1, T_2. Further assume that the stock holding in the ZCBs is α_1 and α_2 respectively. Then the present value of our portfolio is

$$pv = \alpha_1 pv_1 + \alpha_2 pv_2.$$

[1] Recall that by definition the future price of ZCB is equal to one

We derive the Macaulay duration of the portfolio

$$\begin{aligned} d_{1,2}^{Mac} = \frac{d}{d\,r}\log[pv] &= \frac{d}{d\,r}\log[\alpha_1 pv_1 + \alpha_2 pv_2] \\ &= \frac{1}{pv}[-\alpha_1 pv_1 T_1 - \alpha_2 pv_2 T_2] \\ &= \eta_1 d_1 + \eta_2 d_2, \end{aligned}$$

where the portfolio weight η_i is defined as

$$\eta_i = \frac{\alpha_i\, pv_i}{pv}$$

and $d_i = -T_i$, for $i = 1, 2$.

Given that we have an portfolio with the duration $-T_1$, and would like a portfolio with an another, specified, duration, how can we change the portfolio s.t. we have the duration we want? Actually, by simple algebraic operation we find a so called immunization strategy.

Example 6.2:
Assume that we want a duration of $-T_{1,2}$. Then the immunization strategy would be to find the portfolio weight of a ZCB 2 in the market s.t. the portfolio duration is $-T_{1,2}$. By simple algebraic operation we derive the portfolio weight needed

$$\eta_2 = \frac{d_{1,2} - \eta_1 d_1}{d_2}.$$

If the portfolio weight is negative we interpret the result as the portion we need to sell of ZCB 2 in order to immunize our initial portfolio. This means that we go short in ZCB 2.

6.2 Stochastic Duration

In the forthcoming we are going to assume that the forward rate under the Musiela parametrization is the mild solution. That is f_t solves

$$f_t = S_t f_0 + \int_0^t S_{t-s}\alpha_s ds + \int_0^t S_{t-s}\sigma_s dW_s,$$

where σ_s is a deterministic process of Hilbert-Schmidt operators with existing inverse σ_s^{-1} a.e., and where α_s satisfies the HJM-no arbitrage condition (see Chapter 3 and 5). In what follows we may for convenience assume that the risk premium is zero.

The concept of stochastic duration serves as a tool to measure the changes of complex bond portfolios due to changes of the yield curve or forward curve. By the latter equation we see what is affecting the changes in the forward rate. Using Girsanov transform (Section 4.6) we can combine all the changes into a new Brownian Motion

$$d\hat{W}_t = dW_t - \tilde{\alpha}_t dt,$$

under a probability measure $\tilde{\mathbb{P}}$, where

$$\tilde{\alpha}_t = \sigma_t^{-1}(\alpha_t)$$

satisfies the Novikov condition in Theorem (4.16). Recall that the volatility structure doesn't change under the change of measure.

The centered forward curve $\hat{f}_t$ is given by

$$\hat{f}_t = f_t - f_0 = \int_0^t S_{t-s}\sigma_s d\hat{W}_s$$

and a Gaussian random field under $\tilde{\mathbb{P}}$. Then it would be suitable to define the stochastic duration as the derivative w.r.t. $\hat{f}$. Because of stochasticity we cannot use the standard calculus derivative. We need to work within the framework of Malliavin Calculus w.r.t. the centered forward curve $\hat{f}_t$. In order to be able to apply this theory (see [4]) we shall in the sequel assume - also in view of a more general framework w.r.t. α_t - that the SDE

$$dX_t = \alpha_t dt + dW_t$$

has a unique strong solution.

The latter implies in connection with the properties of the left-shift operator S_t and the diffusion coefficient σ_t that the filtrations generated by W_t and $\hat{f}_t$ coincide.

The fundamental theorem of calculus for a Malliavin derivative w.r.t. the Brownian motion, W_t, is

$$D_W \int_0^t u_s dW_s = u_s 1_{[0,t]}(s) + \int_s^t D_W(u_s) dW_s$$

with adequately chosen properties on the functions and functionals [10, Chapter 3]. Using the same theory[2], presented in the paper [5] in the case of strong solutions f_t, we get the same fundamental theorem for the Malliavin derivative w.r.t. the centered forward curve in the risk neutral world (i.e. under $\hat{\mathbb{P}}$):

$$D_{\hat{f}} \int_0^t u_s d\hat{f}_t = u_s 1_{[0,t]}(s) + \int_s^t D_{\hat{f}}(u_s) d\hat{f}_s. \tag{6.1}$$

The latter is called in this master thesis the *Stochastic Duration*. Note the trivial fact that for an deterministic function, u_s, the stochastic duration is

$$D_{\hat{f}} \int_0^t u_s d\hat{f}_t(x) = u_s 1_{[0,t]}(s).$$

We are going to present the theory using the Vasicek model. The volatility function for the Vasicek is deterministic and we use the last property shown. For the CIR model we need to take into account that it is stochastic w.r.t. the short rate.

Definition 6.1 (Stochastic Duration [5])**:**
Let F be a square integrable functional of the forward curve $\hat{f}$ w.r.t. the $\tilde{\mathbb{P}}$-measure. Assume that F is Malliavin differentiable w.r.t. $\hat{f}$. Then the stochastic duration of F is the stochastic process

$$D_{\hat{f}}\, F \in L^2(\Omega, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}; H).$$

In the next example we need an auxiliary result.

Lemma 6.2 (Chain Rule)**:**
Let F be Malliavin differentiable random variable w.r.t. $\hat{f}$. Further suppose that $g : \mathbb{R} \mapsto \mathbb{R}$ is continuously differentiable with bounded derivative. Then $g(F)$ is Malliavin differentiable w.r.t $\hat{f}$ and

$$D_{\hat{f}}\, g(F) = g'(F)\, D_{\hat{f}} F,$$

[2] Itô-Wiener Chaos Expansion, Skorohod Integral, Malliavin Calculus

where g' is the derivative of g.

Proof. [5] □

Example 6.3:

Assume that we are going to buy a ZCB contracted at time t. What is the stochastic duration of buying a future on the ZCB. Let the price of the ZCB at time t (future) be denoted by $P_t(x)$. By the definition of the instantaneous forward rates we know that the price of a ZCB at time t is

$$P_t(x) = \exp\Big\{ -\int_0^x f_t(y)dy \Big\}.$$

First by applying the chain rule,

$$D_{\hat{f}} P_t(x) = P_t(x) D_{\hat{f}} \int_0^x f_t(y)dy,$$

and then finding the stochastic duration of the integral,

$$\begin{aligned} D_{\hat{f}} \int_0^x f_t(y)dy &= \int_0^x D_{\hat{f}} f_t(y)dy \\ &= \int_0^x D_{\hat{f}}(f_0(y) + \hat{f}_t(y))dy \\ &= \int_0^x (0 + 1_{[0,t]}(s))dy \\ &= x\, 1_{[0,t]}(s), \end{aligned}$$

we derive the stochastic duration of the future ZCB price $P_t(x)$

$$D_{\hat{f}} P_t(x) = x\, 1_{[0,t]}(s) P_t(x).$$

Remark that $P_t(x)$ is stochastic up to time t where the price is settled. This means that the stochastic duration is stochastic. Therefore it would be suitable to talk about expected stochastic duration. We then see that

$$E[D_{\hat{f}} P_t(x)] = x\, 1_{[0,t]}(s) E[P_t(x)]$$

◇

6.2.1 Generalized Portfolio

The stochastic duration of a portfolio of stochastic interest rate derivatives; like bond options or swaps, is the main subject for this section. Let us assume that the bond portfolio value at time τ is a square integrable functional of our forward curve. Then we know (see [5]) that this portfolio value has a chaos decomposition w.r.t. $\hat{f}_t$, which actually can be interpreted as a Taylor expansion (on some locally convex space) in infinite dimensions.

In what follows we aim at studying the first order approximation of the chaos decomposition, that is $I_0(f_0) + I_1(f_1)$, where $I_0(f_0) \in \mathbb{R}$ and

$$I_1(f_1) = \int_0^\tau g_s^\tau(S_{\tau-s}\sigma_s)d\hat{W}_s$$

for a deterministic process g_s^τ of (continuous) linear functionals on our Hilbert space. So using Girsanov's theorem we see that the portfolio value is approximately

$$I_0(f_0) + \int_0^\tau g_s^\tau(S_{\tau-s}\alpha_s)ds + \int_0^\tau g_s^\tau(S_{\tau-s}\sigma_s)dW_s$$

under the original probability measure.

The main underlying process of the portfolio is the forward curve. Therefore we may assume that our portfolio take the form

$$Z_x(\tau) = z_v + \int_v^\tau g_s^\tau(S_{\tau-s}\alpha_s)ds + \int_v^\tau g_s^\tau(S_{\tau-s}\sigma_s)dW_s.$$

We are watching the portfolio at time v, where x is the set of time-to-maturities in the portfolio.

We want to work with the Vasicek model by using implicitly a 1-dimensional Brownian motion as an approximation for W_s. The (approximated) stochastic duration of our portfolio is given by

$$\begin{aligned} D_{\hat{f}} Z_x(\tau) &= D_{\hat{f}}(I_0(f) + I_1(f_1)) \\ &= D_{\hat{f}}(I_1(f_1)) \\ &= f_1, \end{aligned}$$

where

$$\begin{aligned} f_1(t,x) &= E[\hat{f}_t(x) I_1(f_1)] \\ &= \int_0^t (S_{t-s}\sigma_s)(x)\, g_s^\tau(S_{\tau-s}\sigma_s) ds \end{aligned}$$

by the construction of the stochastic integral, $I_1(f_1)$ (see [4]).

Approximating the g function

The function is presented in the paper [6]

We interpret the g-function as the total volatility structure of the portfolio. This structure we approximate with series of step functions. Assume that

$$g_s^\tau(\cdot) = \sum_{n=1}^{N} a_n(\tau) b_n(s,\cdot),$$

where $a_i(\tau)$ approximate the τ-dependence in the g-function and $b_i(s,\cdot)$ evaluates the variable function within g and approximate the process up to τ.

The a-function is assumed to be an integral with step functions as integrands. I.e.

$$a_n(\tau) = \int_0^\tau h_n(s) ds,$$

where

$$h_n(s) = \sum_{n=1}^{M} \gamma_{n,m} 1_{(t_{m-1},t_m]}(s).$$

The b-function is divided into an evaluation function and a time dependent function

$$b_n(s,\cdot) = \sum_{k=1}^{K} b_{n,k}(s)\delta_{x_k}(\cdot),$$

where $b_{n,k}(s)$ is approximated by the step function

$$b_{n,k}(s) = \sum_{m=1}^{M} \beta_{n,k,m} 1_{(t_{m-1},t_m]}(s).$$

Putting in for g in the stochastic part of the portfolio process yields

$$\begin{aligned}
\int_0^t g_s^\tau(S_{\tau-s}\sigma_s)dW_s &= \sum_{n=1}^{N} \int_0^\tau a_n(\tau)b_n(s, S_{\tau-s}\sigma_s)dW_s \\
&= \sum_{n=1}^{N} a_n(\tau) \int_0^\tau b_n(s, S_{\tau-s}\sigma_s)dW_s \\
&= \sum_{n=1}^{N} \int_0^\tau h_n(s)ds \int_0^\tau b_n(s, S_{\tau-s}\sigma_s)dW_s \\
&= \sum_{n=1}^{N} \int_0^\tau b_i(s, S_{\tau-s}\sigma_s)h_n(s)dW_s + \int_0^\tau \int_0^s b_n(u, S_{\tau-u}\sigma_u(\cdot))dW_u\, h_n(s)ds.
\end{aligned}$$

This means that our portfolio is decomposed as

$$Z(\tau) = drift + \sum_{n=1}^{N} \int_0^\tau b_i(s, S_{\tau-s}\sigma_s)h_n(s)dW_s.$$

We also assume here that $Z_v = Z_v(\tau)$ is absolutely continuous w.r.t τ. If we define

$$\theta(s) \stackrel{def}{=} \sum_{n=1}^{N} b_i(s, S_{\tau-s}\sigma_s)h_n(s),$$

we see that the quadratic variation of the portfolio is

$$QV_\tau \stackrel{def}{=} [Z,Z]_\tau = \int_0^\tau (\theta(s))^2 ds.$$

By elementary computation the quadratic variation at τ_i is

$$
\begin{aligned}
QV^g_{\tau_i} = & \sum_{m=1}^{M-1} \sum_{n,l=1}^{N} \sum_{k,j=1}^{K} \beta_{n,k,m}\beta_{j,l,m}\gamma_{m,n}\gamma_{m,l}\, \delta_{x_j} \frac{\sigma_s}{C_\sigma} \delta_{x_k} \frac{\sigma_s}{C_\sigma} \Big(\lambda_{\tau_i,t_m}[\sigma_s] - \lambda_{\tau_i,t_{m-1}}[\sigma_s] \Big) \\
& + \sum_{n,l=1}^{N} \sum_{k,j=1}^{K} \beta_{n,k,m}\beta_{j,l,m}\gamma_{m,n}\gamma_{m,l}\, \delta_{x_j} \frac{\sigma_s}{C_\sigma} \delta_{x_k} \frac{\sigma_s}{C_\sigma} \Big(\lambda_{\tau_i,\tau_i}[\sigma_s] - \lambda_{\tau_i,t_{M-1}}[\sigma_s] \Big),
\end{aligned}
$$

where

$$
\lambda_{\tau,t}[\sigma_s] = \int_0^t (\delta_{\tau-s}\sigma_s)^2 ds,
$$

and C_σ are the constants w.r.t. x in the volatility function. Note that

- M is the number of discretized time-steps approximating the time of the observations
- K is the number of different time-to-maturities in the portfolio
- N is the number of different types of a- and b-functions

At the times τ_i we observe the portfolio values $Z^{obs}_{t,T}(\tau_i)$ and estimate the quadratic variation

$$
QV^{obs}_{\tau_i} = \sum_{j=1}^{i} \Big(Z^{obs}_{t,T}(\tau_j) - Z^{obs}_{t,T}(\tau_{j-1}) \Big)^2 .
$$

The idea is to estimate the $\beta_{n,k,m}, \gamma_{n,m}$'s by minimizing the least square of the quadratic variation function and the observed quadratic variation. Assume that we have more than $N^3 + N^2$ observations, then we minimize

$$
\sum_{i=1}^{\geq N^3+N^2} \Big(QV^g_{\tau_i} - QV^{obs}_{\tau_i} \Big)^2 \longrightarrow \min_{\beta,\gamma} .
$$

We end this section with an simulation example.

Example 6.4:

For simplicity we assume that the volatility structure follows the Vasicek/Hull-White model

$$
\sigma_s(x,f) = \sigma \exp\{-ax\}.
$$

Then we derive that

$$\begin{aligned}\lambda_{\tau,t}[\sigma_s] &= \int_0^t \big(\sigma \exp\{-a(\tau - s)\}\big)^2 ds \\ &= \frac{\sigma^2}{2a}(\exp\{-2a(\tau - t)\} - \exp\{-2a(\tau)\}).\end{aligned}$$

By the telescope we obtain

$$\begin{aligned} QV^g_{\tau_i} &= \sum_{m=1}^{M-1}\sum_{n,l=1}^{N}\sum_{k,j=1}^{K} \beta_{n,k,m}\beta_{j,l,m}\gamma_{m,n}\gamma_{m,l}\, e^{-ax_j}e^{-ax_k}\left(\frac{\sigma^2}{2a}e^{-2a(\tau_i-t_m)} - \frac{\sigma^2}{2a}e^{-2a(\tau_i-t_{m-1})}\right) \\ &\quad + \sum_{n,l=1}^{N}\sum_{k,j=1}^{K} \beta_{n,k,m}\beta_{j,l,m}\gamma_{m,n}\gamma_{m,l}\, e^{-ax_j}e^{-ax_k}\left[\frac{\sigma^2}{2a}e^{-2a(\tau_i-\tau_i)} - \frac{\sigma^2}{2a}e^{-2a(\tau_i-t_{M-1})}\right] \\ &= \sum_{m=1}^{M-1}\sum_{n,l=1}^{N}\sum_{k,j=1}^{K} \beta_{n,k,m}\beta_{j,l,m}\gamma_{m,n}\gamma_{m,l}\, \frac{\sigma^2}{2a}e^{-a(x_j+x_k+2\tau_i)}\left(e^{2at_m} - e^{2at_{m-1}}\right) \\ &\quad + \sum_{n,l=1}^{N}\sum_{k,j=1}^{K} \beta_{n,k,m}\beta_{j,l,m}\gamma_{m,n}\gamma_{m,l}\, \frac{\sigma^2}{2a}e^{-a(x_j+x_k+2\tau_i)}\left[e^{2a\tau_i} - e^{2at_{M-1}}\right]. \end{aligned}$$

In this example we are also looking at the simulated portfolio. The main interest is in constructing a reasonable program for calculating the stochastic duration. The program will serve as a tool for understanding the "non-analysis" part of the concept.

Initial Values

We need the following initial values: **the standard deviation of the portfolio**; $\sigma port$, **the Vasicek parameters**; σc and ac, **the number of functions**; $N1$, **the number of time-to-maturity** in the portfolio; K, **and the number of time-discretization**; M. The number of time-discretization is divided into the integer time length L and the number of time steps between each integer, δ, yielding the time set t. This results in the **parameters**; β and γ, which must be estimated. In order to estimate the parameters we need to **observe** at least $M \times L \times N1 = \text{Length}[EstPar]$ of the quadratic variation of the portfolio. **The observation times** are τc. The last initial value need is the **time-to-maturities** in the portfolio; x. For contracts that have a maturity before the ZCB portfolio starts (futures, options) the x's are constants, while e.g. in the second hand market, the x's vary due to t.

```
N1 = 1; K = 2; δ = 5; L = 3; M = L δ; (*Creates the amount of parameters; here we can change*)
σport = 0.08; σc = 0.08; ac = 0.08; Icap = 40;
(* Initial conditions: σc and ac is known from the Vasicek model,
Icap is the initial capital (portfolio value*)
β = Table[beta[i, j, l], {i, N1}, {j, M}, {l, K}];
γ = Table[gamma[i, l], {i, N1}, {l, M}];
EstPar = EstPar = DeleteCases[DeleteCases[Flatten[Join[β, γ]], _Integer], _Real];
t = Table[i/δ, {i, 0, M}]; (*Time*)
τc = Table[i/(δ Length[EstPar]) M, {i, 0, Length[EstPar]}];
x = RandomVariate[UniformDistribution[{L, L + 5}], K] (*Table[10,{j,K1}];*)
{6.02923, 7.69865}
```

Simulating the Portfolio

When we simulate the portfolio many choices of structures can be made. The program provides a simple Brownian motion with the volatility structure equal to the Vasicek model. Recall that

$$\begin{aligned} Var\left[\int_0^\tau \sigma \mathrm{e}^{-ax} dW_s\right] &= \sigma^2 \mathrm{e}^{-2ax} Var[W_\tau] \\ &= \sigma^2 \mathrm{e}^{-2ax} \tau \end{aligned}$$

When we have simulated the portfolio we easily calculate the observed quadratic variation.

```
SimPortfolio =
  Icap + Accumulate[Table[RandomVariate[NormalDistribution[0, σport Exp[-ac (τc[[i + 1]] - τc[[i]])]]],
     {i, 1, Length[EstPar]}]];

SimPortfolio = Prepend[SimPortfolio, Icap];

QVobs = Accumulate[Table[(SimPortfolio[[j + 1]] - SimPortfolio[[j]])^2, {j, 1, Length[SimPortfolio] - 1}]];
```

The Quadratic Variation Function

We define the volatility structure (Vasicek) and the quadratic variation function.

```
σ[s_, σc_, ac_] := σc Exp[-ac s] (*Volatility Function*)

θ[τ_, σc_, ac_, β_, γ_, N_, K_, M_, t_, x_] :=
 Sum[ Sum[γ[[n]][[m]] γ[[ν]][[μ]] (Min[t[[m + 1]], τ] - Min[t[[m]], τ]) (Min[t[[μ + 1]], τ] - Min[t[[μ]], τ]),
    {m, M}, {μ, M}]
   Sum[
    Integrate[Sum[β[[n]][[m]][[k]] β[[ν]][[m]][[k]] β[[n]][[m]][[κ]] β[[ν]][[m]][[κ]]
       σ[τ - s + x[[k]], σc, ac] σ[τ - s + x[[κ]], σc, ac] , {k, K}, {κ, K}],
     {s, Min[t[[m]], τ], Min[t[[m + 1]], τ]}], {m, M}]
  , {n, N}, {ν, N}] (*Quadratic Variation Function*)
```

Minimization

A simple strategy is to minimize the distance between $QVobs$ and QVg through a least square. Note that the least square distance isn't a good estimate on how good the estimation is due to possible stochastic volatility functions in the portfolio. In this example the distance is rather small

```
sum = Sum[(θ[j/M, σc, ac, β, γ, N1, K, M, t, x] - QVobs[[j]])^2, {j, Length[QVobs]}];
Mβγ = NMinimize[sum, EstPar]; (*FindMinimum,NMinimize,Minimize*)
First[Mβγ]
βest = β /. Last[Mβγ];
γest = γ /. Last[Mβγ];

8.30575 × 10^-7

EstQVobs = Table[θ[j/(δ L), σc, ac, βest, γest, N1, K, M, t, x], {j, Length[QVobs]}];
```

In Figure 6.1 we see a plot of the quadratic variation function and the observed quadratic variations.

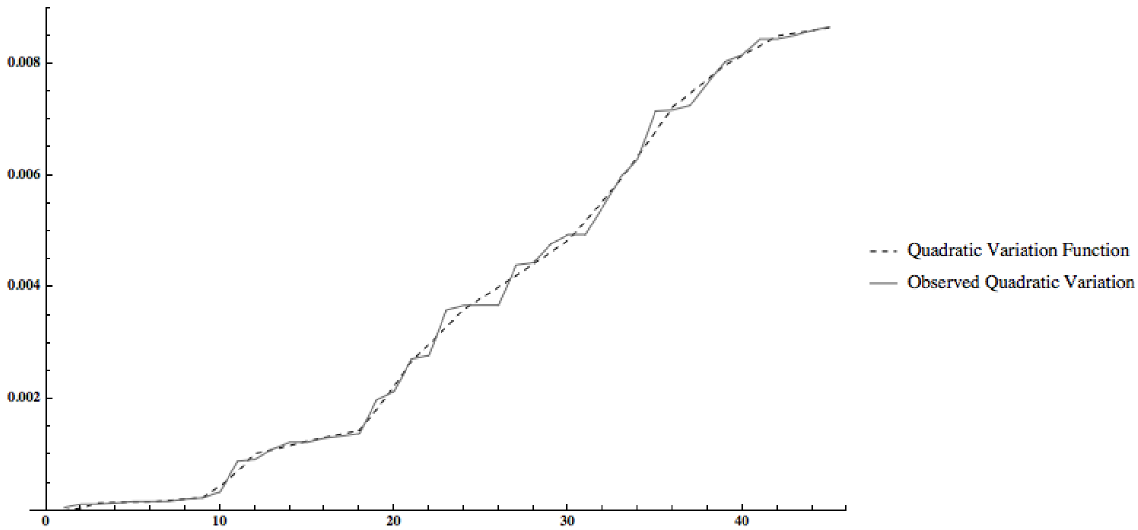

Figure 6.1: *QVobs vs. QV-function*

The g-function

Before we calculate the stochastic duration we take a look at the g-function that estimates the volatility function of the portfolio. The changes aren't radical in this case, but that is due to choices of initial values.

```
a[τ_, n_, M_, γ_] := Sum[γ[[n]][[m]] (Min[t[[m + 1]], τ] - Min[t[[m]], τ]), {m, M}]

b[τ_, s_, n_, {σc_, ac_}, x_, M_, K_, β_] :=
 Sum[β[[n]][[m]][[k]] Boole[t[[m]] ≤ s < Min[t[[m + 1]], τ]] σ[x + τ - s, σc, ac], {k, K}, {m, M}]
 (*σ[x[[k]]+τ-s,σc,ac]*)

g[τ_, s_, {σc_, ac_}, x_, M_, K_, N_, β_, γ_] :=
 Sum[a[τ, n, M, γ] b[τ, s, n, {σc, ac}, x, M, K, β], {n, 1, N}]

t221 = Table[{20 τ , 20 s, g[τ, s, {σc, ac}, Mean[x], M, K, N1, βest, γest]}, {τ, 0, 3, 30/365},
   {s, 0, 3, 30/365}];
```

By Figure 6.2 the function g have stochastic changes as s increases, while for the observation time there are changes due to the step function.

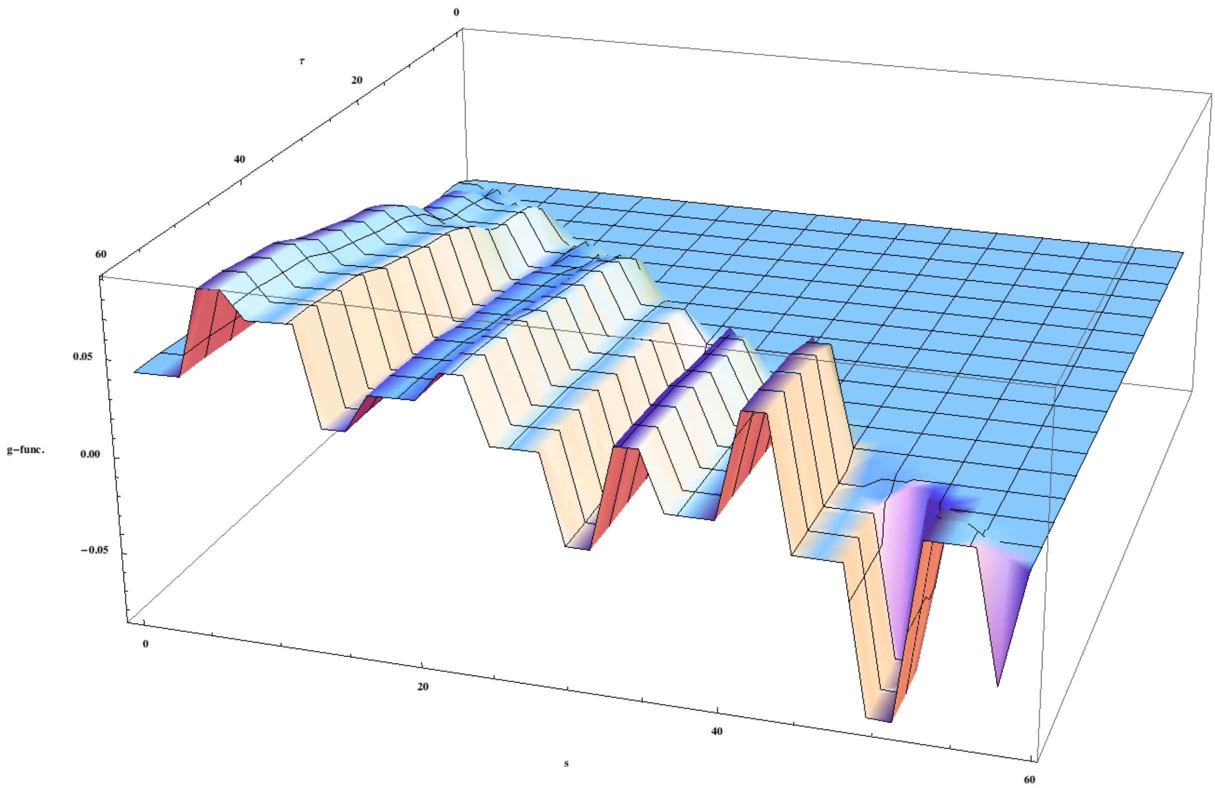

Figure 6.2: *The g-function*

Stochastic Duration

Recall that the stochastic duration of the portfolio is

$$D_{\hat{f}} Z_x(\tau) = f_1,$$

where

$$f_1(t,x) = \int_0^t (S_{t-s}\sigma_s)(x) g_s^{\tau}(S_{\tau-s}\sigma_s) ds.$$

Therefore

```
f1[v_, y_] := NIntegrate[σ[y + v - s, σc, ac] g[τc[[-1]], s, {σc, ac}, y, M, K, N1, βest, γest], {s, 0, v}]

t222 = Table[{20 v , y, f1[v, y]}, {v, 30/365, 3, 30/365}, {y, 30/365, 3, 30/365}];
```

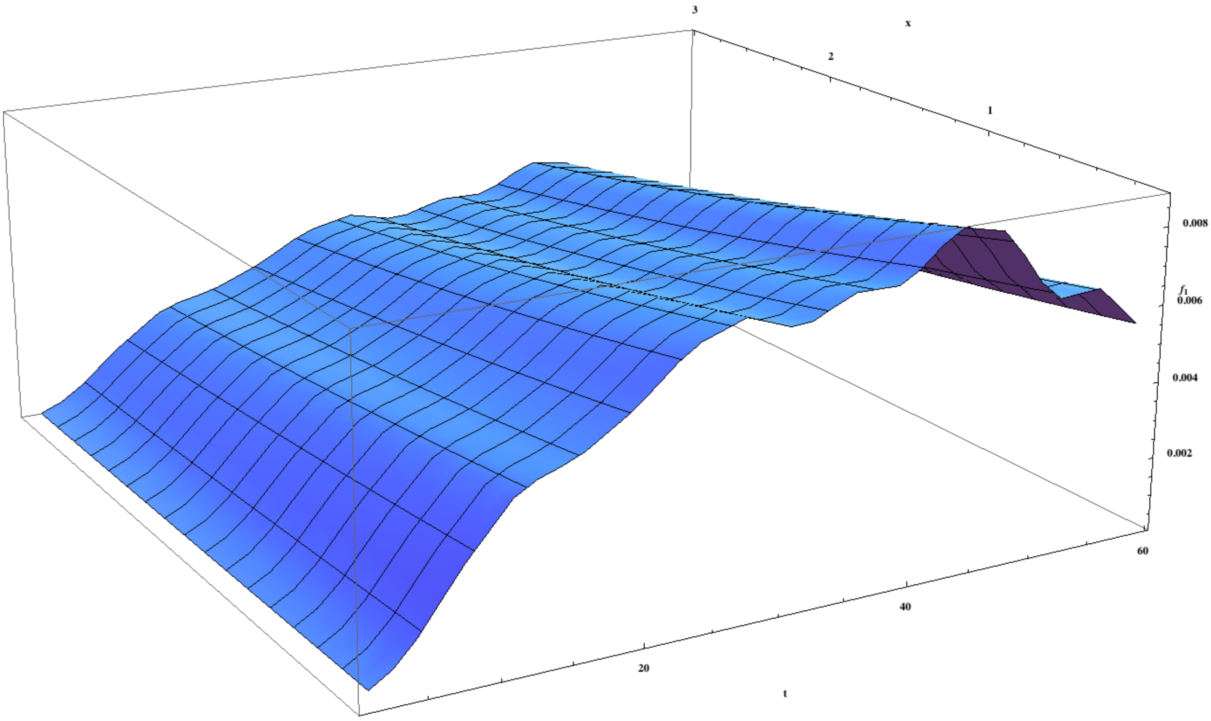

Figure 6.3: *The Stochastic Duration*

◇

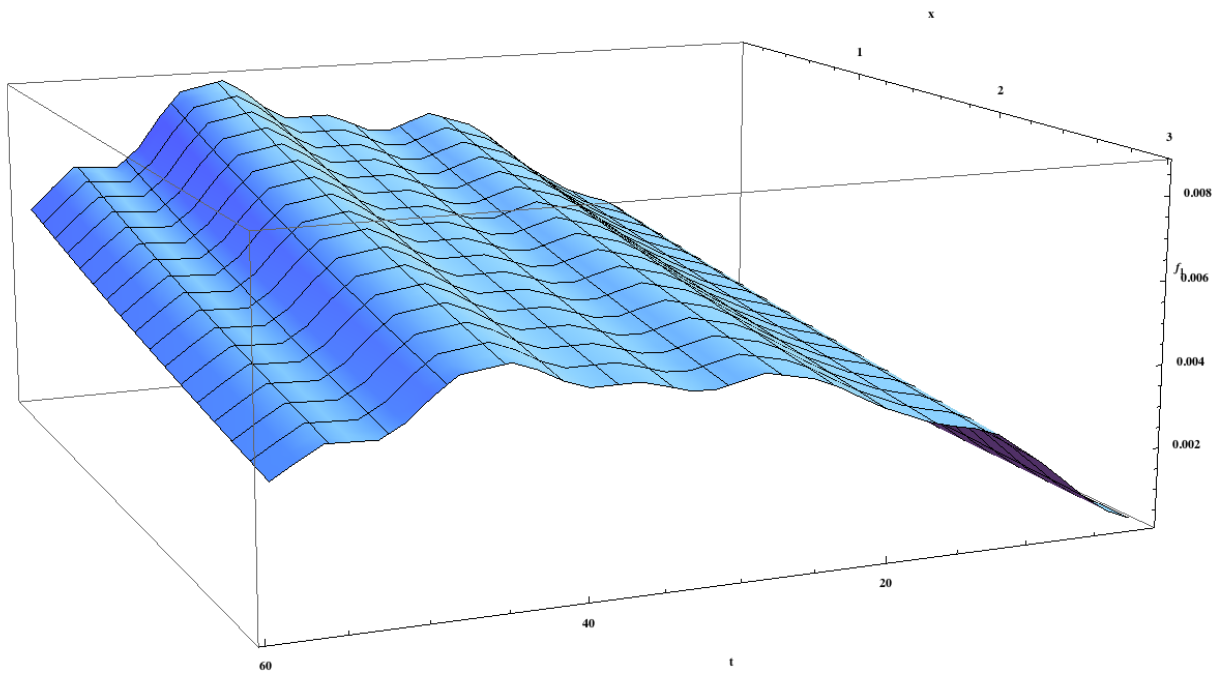

Figure 6.4: *The Stochastic Duration*

6.3 Immunization Strategy

The immunization strategy is based on the paper [11], where the main idea is to minimize the stochastic duration in time and space (time-to-maturity). Assume that we have a portfolio $V_x(\tau)$ with it's estimated stochastic duration

$$D_{\hat{f}} V_x(\tau).$$

Then we want to minimize our risk using the hedging strategy H_t with the stochastic duration

$$D_f H_x(\tau).$$

The latter stochastic duration is "most likely" stochastic and it will be sensible to minimize the expected stochastic duration. Let, e.g., the hedging strategy be of α_{ZCB} shares of a Future and α_{call} shares of a call option on a ZCB. Then the immunization strategy would be

$$E\Big[\int_0^x \int_0^t (D_{\hat{f}} V_t - D_{\hat{f}} H_t^\alpha)^2 \, ds \, dy\Big] \longrightarrow \min_\alpha$$

w.r.t. α_{ZCB} and α_{call} (under some boundary conditions).

The stochastic duration on the future is already derived in an example. We therefore derive the stochastic duration on a call option of a ZCB

$$\big(P(t,T) - K\big)^+.$$

Example 6.5:
Let $F = \big(P(t,T) - K\big)^+$ then

$$F = 1_{[K,\infty)}\big[P(t,T)\big](P(t,T) - K).$$

Using Example 2.2 in [5] we have that

$$\begin{aligned} D_{\hat{f}}F &= 1_{[K,\infty)}\big[P(t,T)\big]D_{\hat{f}}(P(t,T) - K) \\ &= 1_{[K,\infty)}\big[P(t,T)\big](-(T-t)P(t,T) - 0) \\ &= -1_{[K,\infty)}\big[P(t,T)\big](T-t)P(t,T). \end{aligned}$$

$\diamond$

Chapter 7

Stochastic Duration an Example

As a major example we are going to calculate the *stochastic duration* for a Future portfolio on a two year Treasury Note[1]. The data are collected from www.quandle.com, which refer further to Chicago Mercantile Market and US Treasury From this website the following data collection was used; (Treasury Yield Curve Rates) and (2 Year Treasury Note Futures, March 2013, TUH2013, CBOT).

Since we are working with a possible infinite dimensional noise, we need to reduce the dimensions. Using PCA[2] to reduce the dimension, we estimate a Vasicek model for the total covariance of the important components. This is a strong simplification since a better approach would be to estimate a volatility structure for each of the important components. But for the purpose here it is well enough.

Furthermore we use the program presented in Chapter 6, using the parameters that are estimated by means of the PCA procedure.

[1]A Note is equal a Bond with time-to-maturity of 2-10 years [8]

[2]presented in Chapter 3

7.1 Principal Component Analysis

As earlier noticed we are going to use a PCA procedure. Firstly we need to import the (Treasury Yield Curve Rates) from 01.10.93 until 31.07.2001[3]. An issue with this data set is that the numbers are interpolations of observations s.t. we have observation for the time-to-maturities $(0.25, 0.5, 1, 2, 3, 5, 7, 10, 20, 30)$. This means that there will be some variation in the data set from the real observations.

```
dataimport =
  Import[
   "http://www.quandl.com/api/v1/datasets/USTREASURY/YIELD.csv?&trim_start=1993-10-01&trim_end=2001-07-31&
     sort_order=desc", "Data"];

TtM = Table[i, {i, {0.25, 0.5, 1, 2, 3, 5, 7, 10, 20, 30}}];

Wdata = Table[dataimport[[i]][[j]] / 100, {i, 3, Length[dataimport]}, {j, {3, 4, 5, 6, 7, 8, 9, 10, 11, 12}}];
Length[Wdata]

1960
```

We have 1960 observations that is provided in percentage. Hence we need to divide by 100.

With the new data set we estimate the covariance, and use the spectral decomposition theorem.

```
CovarianceWd = Covariance[Wdata];
{u, w, v} = SingularValueDecomposition[CovarianceWd];

u.w.Transpose[v] == CovarianceWd

True
```

By Figure 7.1 we see that we can approximate the covariance matrix by the three first components, (99.7%).

Therefore we approximate the covariance by the covariance of the three first components. Comparing the diagonal on the approximated covariance and the estimated covariance we see that the difference is small.

[3]The reason why this dates are used is because of an example in [8]

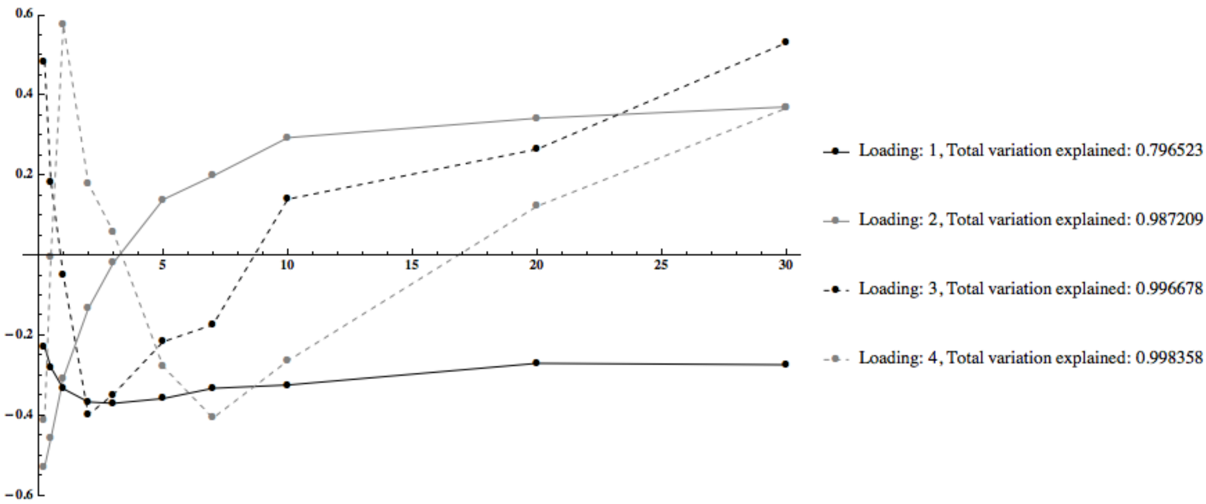

Figure 7.1: *The important components of the Yield Curve*

```
Sum[Diagonal[w][[i]] Outer[Times, Transpose[v][[i]], Transpose[v][[i]]], {i, Length[Diagonal[w]]}] ==
 CovarianceWd

EstCov = Sum[Diagonal[w][[i]] Outer[Times, Transpose[v][[i]], Transpose[v][[i]]], {i, 3}];

Diagonal[EstCov]
Diagonal[CovarianceWd]

True

{0.0000559628, 0.000059144, 0.0000613263, 0.0000644055, 0.0000635025,
 0.0000612088, 0.0000551678, 0.0000578492, 0.0000467664, 0.0000510921}

{0.0000562036, 0.0000593128, 0.0000617063, 0.000064502, 0.0000635427,
 0.0000613593, 0.0000553911, 0.0000580185, 0.0000470086, 0.0000513019}
```

Performing PCA on this data set, [8], means that we are assuming the auto-correlation to be small. Because of a simplified example, we don't put concerns on this issue. We recall from Example 3.1 that we need to estimate the parameters using the relation

$$Var[Y(t,T)] = \frac{1}{(T-t)^2}\frac{\sigma^2}{2k^3}\mathrm{e}^{-2k(T+t)}(\mathrm{e}^{2kt}-1)(\mathrm{e}^{kT}-\mathrm{e}^{kt})^2,$$

where, since we observe day-to-day changes in the volatility structure, $t = \frac{1}{365}$. Using least square we estimate the parameters in the Vasicek model to be $\sigma = 0.1495$ and $k = 0.00816$.

```
VarVasicek[T_, t_, σ_, k_] := 1/(T - t)^2 σ^2/(2 k^3) Exp[-2 k (T + t)] (Exp[2 k t] - 1) (Exp[k T] - Exp[k t])^2

t = 1/365; TtM;

minσk = NMinimize[Sum[(VarVasicek[TtM[[i]] + t, t, σ, k] - Diagonal[EstCov][[i]])^2, {i, Length[TtM]}], {σ, k}]
estσk = {σ, k} /. Last[minσk];

{1.13759×10^-10, {σ → 0.149521, k → 0.00815519}}
```

Using the estimated parameters we are ready to calculate the Stochastic Duration of an portfolio of (2 Year Treasury Note Futures, March 2013, TUH2013, CBOT).

7.2 Stochastic Duration of the Portfolio

Clearly finding the Stochastic Duration for a future TUH2013, isn't comparable with the estimation of the volatility structure for the period (1993-2001). But the procedure is the same as we show in this chapter.

Firstly we import the data. The data of interest are the settlement data and the date of the transaction.

```
portdata =
  Import[
   "http://www.quandl.com/api/v1/datasets/CME/TUH2013.csv?&auth_token=LqWy2oLARYBxWdmoFX25&trim_start=2011-12-
     30&trim_end=2013-03-28&sort_order=desc", "Data"];

portdata[[1]]
{Date, Open, High, Low, Last, Change, Settle, Volume, Prev. Day Open Interest}

Wport = portdata[[2 ;; 116]];
m = Length[Wport];

settle = Transpose[Wport][[7]];

date = Transpose[Wport][[1]];
```

Initial Values

We use the initial values derived from the PCA procedure, and divide the time into three main time steps with an discretization of five per time step. Since we are working with a portfolio of only 2 year Treasury Notes the time-to-maturity is $x = \{2\}$.

```
K = 1; δ = 5; M = 15; N1 = 1; σc = estσk[[1]]; ac = estσk[[2]];
β = Table[beta[i, j, l], {i, N1}, {j, M}, {l, K}];
γ = Table[gamma[i, l], {i, N1}, {l, M}];
EstPar = EstPar = DeleteCases[DeleteCases[Flatten[Join[β, γ]], _Integer], _Real];
t = Table[i/δ, {i, 0, M}];
x = {2};
```

The Portfolio

We want to do some transaction on each of the days. The total stock holding is decided by a Uniform$(20, 40)$, which means that we at least have 20 stocks and at most have 40 stocks in the Future TUH2013. Using the settlement values as the value of one stock in the Future, we derive the Quadratic Variation

```
SimPortfolio = Table[RandomVariate[UniformDistribution[{20, 40}]] settle[[-i]], {i, 60}];

QVobs = Accumulate[Table[(SimPortfolio[[j + 1]] - SimPortfolio[[j]])^2, {j, 1, Length[SimPortfolio] - 1}]];
```

Minimization

The next step in the procedure is to minimize the Quadratic Variation function using the program in chapter 6.

```
σ[s_, σc_, ac_] := σc Exp[-ac s]

θ[τ_, σc_, ac_, β_, γ_, N_, K_, M_, t_, x_] :=
 Sum[ Sum[γ[[n]][[m]] γ[[ν]][[μ]] (Min[t[[m + 1]], τ] - Min[t[[m]], τ]) (Min[t[[μ + 1]], τ] - Min[t[[μ]], τ]),
    {m, M}, {μ, M}]
   Sum[
    Integrate[Sum[β[[n]][[m]][[k]] β[[ν]][[m]][[k]] β[[n]][[m]][[κ]] β[[ν]][[m]][[κ]]
       σ[τ - s + x[[k]], σc, ac] σ[τ - s + x[[κ]], σc, ac] , {k, K}, {κ, K}],
     {s, Min[t[[m]], τ], Min[t[[m + 1]], τ]}], {m, M}]
  , {n, N}, {ν, N}]


sum = Sum[(θ[j/20, σc, ac, β, γ, N1, K, M, t, x] - QVobs[[j]])^2, {j, Length[QVobs]}];
Mβγ = FindMinimum[sum, EstPar];
First[Mβγ]
βest = β /. Last[Mβγ];
γest = γ /. Last[Mβγ];

FindMinimum::cvmit : Failed to converge to the requested accuracy or precision within 100 iterations. >>

2.25028 × 10^13

EstQVobs = Table[θ[j/20, σc, ac, βest, γest, N1, K, M, t, x], {j, Length[QVobs]}];
```

Here we get a hugh least square estimate, 2.25×10^{13}, but we see by Figure 7.2, that the fit is rather good.

The g-function

As in Chapter 6 we define the g-function through the estimated a- and b-function.

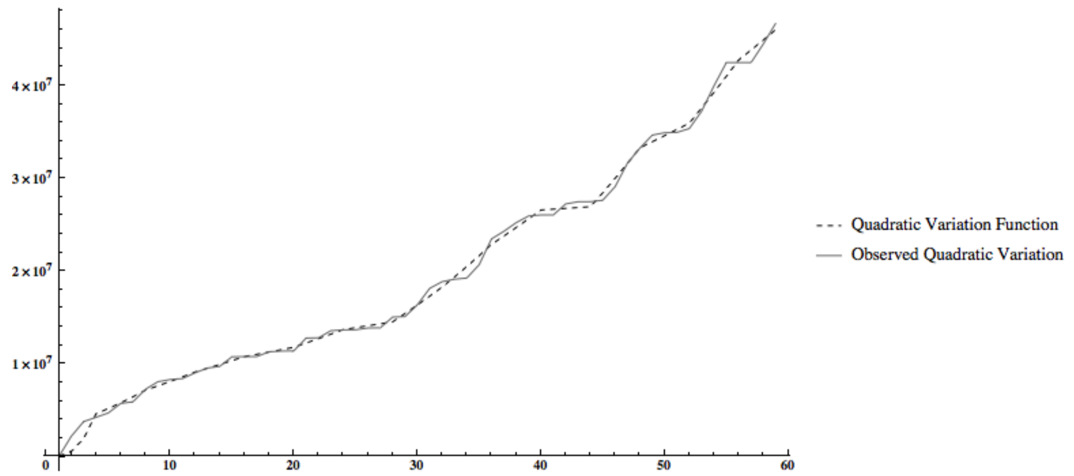

Figure 7.2: *The Quadratic Variation function vs. the observed Quadratic Variation*

```
a[τ_, n_, M_, γ_] := Sum[γ[[n]][[m]] (Min[t[[m + 1]], τ] - Min[t[[m]], τ]), {m, M}]

b[τ_, s_, n_, {σc_, ac_}, x_, M_, K_, β_] :=
 Sum[β[[n]][[m]][[k]] Boole[t[[m]] < s ≤ Min[t[[m + 1]], τ]] σ[x + τ - s, σc, ac], {k, K}, {m, M}]

g[τ_, s_, {σc_, ac_}, x_, M_, K_, N_, β_, γ_] := Sum[a[τ, n, M, γ] b[τ, s, n, {σc, ac}, x, M, K, β], {n, 1, N}]

t221 = Table[{20 τ , 20 s, g[τ, s, {σc, ac}, 2, M, K, N1, βest, γest]}, {τ, 0, 3, 30/365}, {s, 0, 3, 30/365}];
```

By Figure 7.3 and Figure 7.4 we see the estimated path of the g-function. Recall that the g-function describes the portfolio volatility structure.

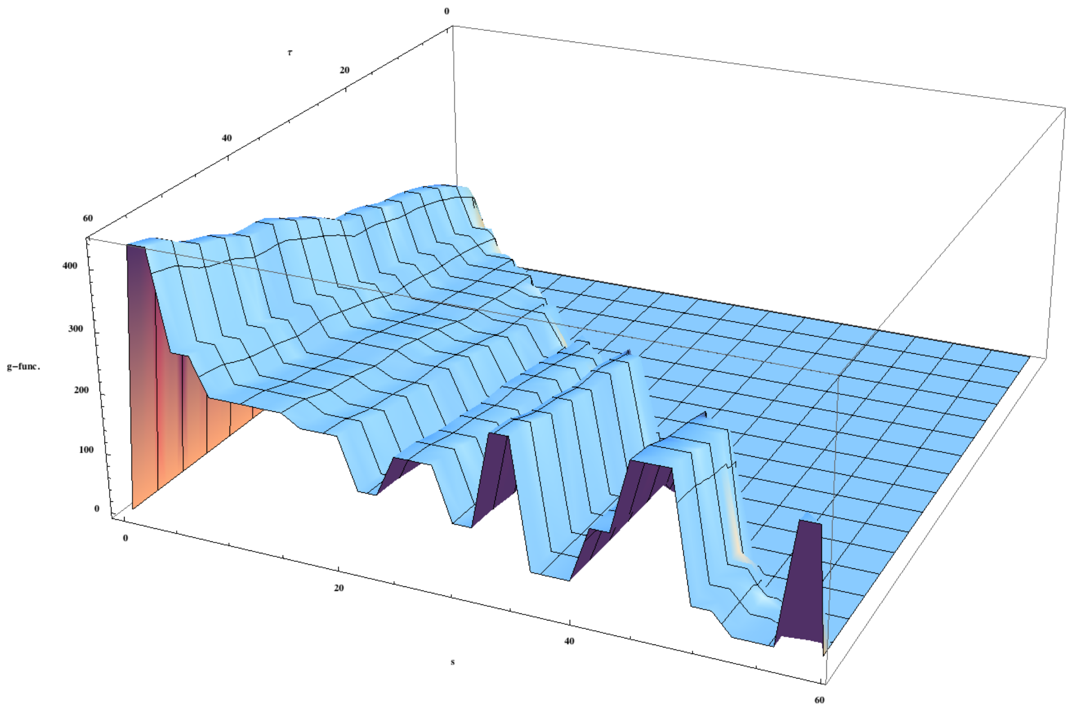

Figure 7.3: *The volatility structure of the portfolio, g*

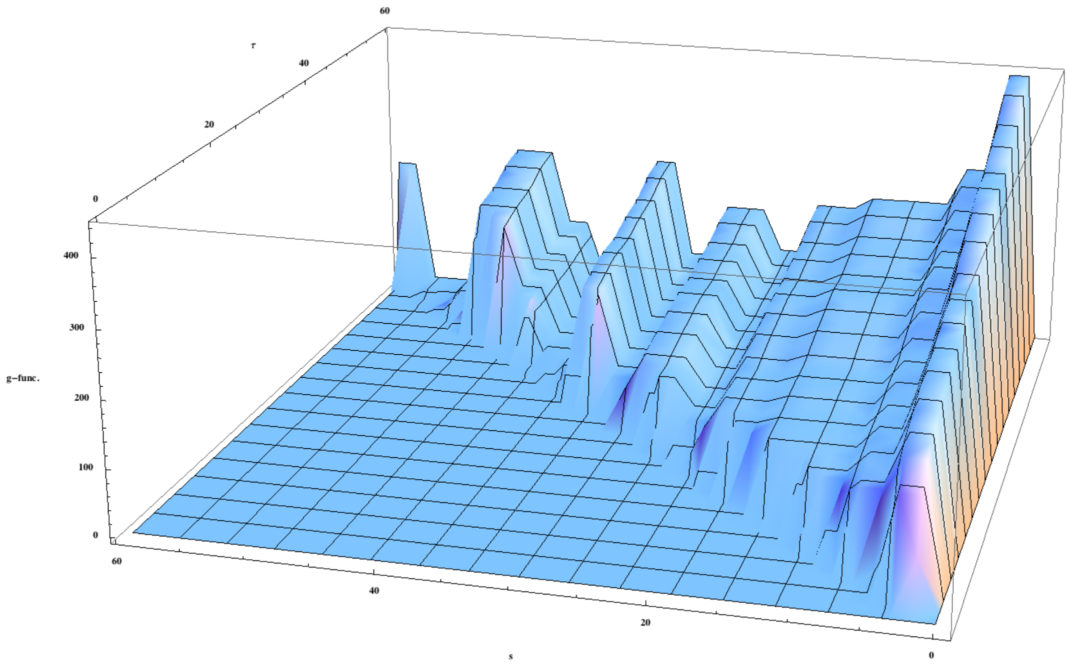

Figure 7.4: *The volatility structure of the portfolio, g*

Stochastic Duration

Using the volatility structure of the portfolio, we find the *stochastic duration*. Recall that the stochastic duration describes the changes in the portfolio value due to changes in the forward curve.

```
f1[v_, y_] := NIntegrate[σ[y + v - s, σc, ac] g[t[[-1]], s, {σc, ac}, y, M, K, N1, βest, γest], {s, 0, v}]

t222 = Table[{20 v , y, f1[v, y]}, {v, 30/365, 3, 30/365}, {y, 30/365, 3, 30/365}];
```

We find the Stochastic Duration in time and space.

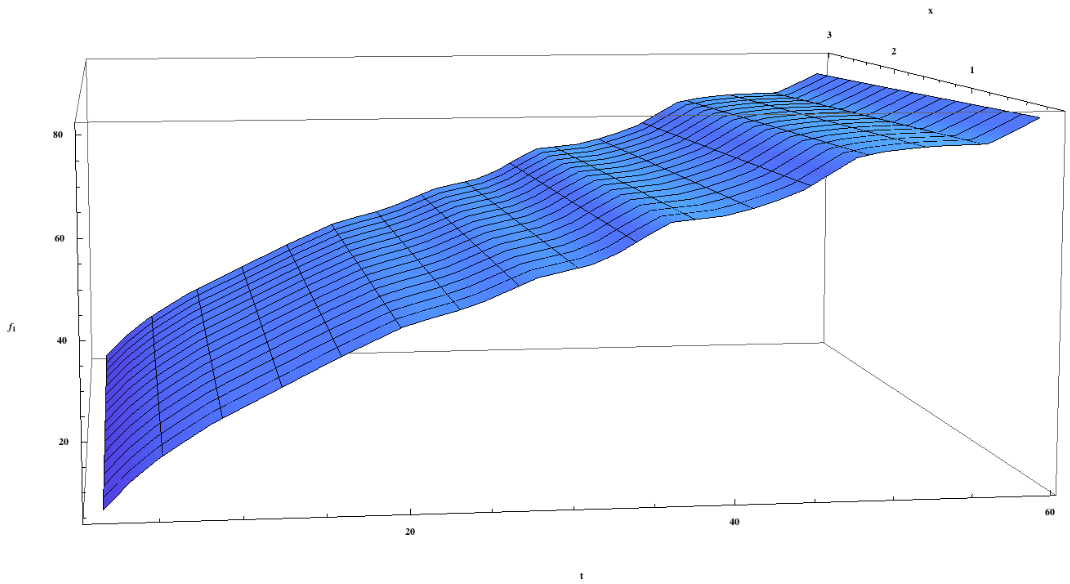

Figure 7.5: *The Stochastic Duration*

Appendix A

Mathematical Tools

The short rate r is not deterministic. This means that we can't with exact mathematical analysis predict the path of r, which brings us into the field of probability theory. A way to model r is to define it as an r.v. on the probability space

$$(\Omega, \mathcal{F}, \mathbb{P}),$$

where Ω denotes the sample space, $\mathcal{F}$ denotes the σ-algebra on the sample space, and $\mathbb{P}$ the probability measure(distribution).

Definition A.1 (Sample space)**:**
A set $\Omega \neq \emptyset$ representing the collection of all possible outcomes of a random experiment is called sample space.

Knowing the possible outcomes of r, Ω, we would like to figure out which family of outcomes that are reasonable to put a probability measure on. The definition of σ-algebra ensures some good properties of the family.

Definition A.2 (σ-algebra)**:**
A family $\mathcal{F}$ of subsets of Ω is called σ-algebra on Ω if

1. $\emptyset \in \mathcal{F}$

2. $A \in \mathcal{F} \implies A^C \in \mathcal{F}$ [1]

3. $A_1, A_2, \dots \in \mathcal{F} \implies \cup_{i \geq 1} \in \mathcal{F}$

The last element of the probability space, the probability measure, tells us to which degree the different measurable outcomes occur. There are several such measure (χ^2, Log-Gaussian, Pareto), but the most common within financial theory is the Gaussian distribution. With the probability measure we get a set of working tools; moments, characteristic function (the Fourier transform of an r.v.), etc.

Definition A.3 (Probability measure)**:**
A function

$$\mathbb{P} : \mathcal{F} \mapsto [0, 1]$$

is called a probability measure on $(\Omega, \mathcal{F})$, if

1. $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(\Omega) = 1$

2. $A_1, \dots, A_n, \dots \in \mathcal{F}$ with $A_i \bigcap A_j = \emptyset$ [2] $\implies \mathbb{P}(\bigcup_{i \geq 1} A_i) = \sum_{i \geq 1} \mathbb{P}(A_i)$ [3]

Heuristically, for an event $A \in \mathcal{F}$, $\mathbb{P}(A)$ is the probability that event A occurs.

Properties A.4 (Properties of the probability measure)**:**
We have following properties of the probability measure

1. $A \subseteq B \implies \mathbb{P}(A) \leq \mathbb{P}(B)$

2. $\mathbb{P}(\bigcup_{i \geq 1} A_i) \leq \sum_{i \geq 1} \mathbb{P}(A_i)$

3. $\lim_{i \to \infty} \mathbb{P}(A_i) = \mathbb{P}(A)$ if $A = \bigcup_{i \geq 1} A_i$

In some cases an event has probability zero. This sets are referred to as the $\mathbb{P}$-null sets[4].

---

[1] $A^C \stackrel{def}{=} \Omega - A$
[2] A_i and A_j is disjoint
[3] In the literature this is called σ-additivity
[4] Of course given that we are working with the probability measure $\mathbb{P}$

Definition A.5 ($\mathbb{P}$-null sets)**:**
An event $A \in \mathcal{F}$ s.t. $\mathbb{P}(A) = 0$ is called a $\mathbb{P}$-null set.

On the other side we may have some events that have probability one. In this case the event have the $\mathbb{P}$-a.s. property.

Definition A.6 ($\mathbb{P}$-almost surely)**:**
We say that an event holds $\mathbb{P}$-almost surely (a.s.) if there exist a $\mathbb{P}$-null set $N \in \mathcal{F}$ s.t. the event holds for all $\omega \in N^C = \Omega - N$.

A.1 Random Variable

Working with the outcomes isn't always easy, especially if the events aren't numbers. For example could we categorize sections in companies into $\{1, 2, 3, 4, 5, ...\}$ instead of the HR-section, Actuary-section, etc. By doing so we can define expectations and other statistical tools that relay on number theory. For this type of purpose we need a mapping function, and the mapping function is commonly referred to as the random variable or stochastic variable.

Definition A.7 (Random Variable)**:**
Let X be the function

$$X : \Omega \mapsto E,$$

where $E \subseteq \mathbb{R}$. Then X is called a $(E, \mathcal{E})$-random variable (r.v.) on $(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathcal{E}$ is the measurability of X s.t., for an $A \in \mathcal{E}$, the inverse image of the r.v. $X^{-1}(A) \in \mathcal{F}$.

Note that the Ω might be equal to E, like it is for dices. A special case of E is real valued random variables. Then

$$X : \Omega \longmapsto \mathbb{R}$$

is a r.v. on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ if the set

$$A \stackrel{def}{=} \{\omega \in \Omega : X(\omega) < x\}$$

is an event for all $x \in \mathbb{R}$. I.e. $A \in \mathcal{F}$. If the random variable have two or more dimension we call it a random vector (r.v).

Definition A.8 (Random Vector)**:**
A function

$$X : \Omega \mapsto \mathbb{R}^n$$

with

$$X(\omega) = (X_1(\omega), X_2(\omega), \ldots, X_n(\omega))^T$$

is called a random vector, if $X_1, \ldots, X_n$ are r.v.'s.

A.2 Expectation

The most common statistic is the expectation which is defined as the Lebesgue integral w.r.t. the probability measure $\mathbb{P}$. The reason why we are working with Lebesgue integral is because the random variable might be discrete. In discrete case the Riemann integral is equal to zero due to it's definition.

Definition A.9 (Expectation w.r.t. $\mathbb{P}$)**:**
Let $X \geq 0$ be a positive r.v. Then the expectation (integral) is defined as

$$E[X] \stackrel{def}{=} \int_\Omega X(\omega)\mathbb{P}(d\omega).$$

For a general X we define

$$E[X] \stackrel{def}{=} \int_\Omega X(\omega)\mathbb{P}(d\omega) \stackrel{def}{=} E[X^+] - E[X^-],$$

where $X^+ = \max(x, 0)$ and $X^- = \max(-x, 0)$. If $E[X^+], E[X^-] < \infty$ then X is called $\mathbb{P}$-integrable.

Properties A.10 (Properties of the expectation)**:**
Let X and Y be two r.v. and $\alpha, \beta \in \mathbb{R}$[5]

[5] $\mathbb{R}$ is an example of a field

1. $E[\alpha X + \beta Y] = \alpha E[X] + \beta E[Y]$ (Linearity)

2. If $X \leq Y$ $\mathbb{P}$-a.s. then $E[X] \leq E[Y]$

A.3 Conditional probability and expectation

In many cases there would be interesting to understand what properties a random variable have given knowledge. E.g. for a stochastic process, defined later, it is interesting to understand the process given knowledge about the process up to a certain time s. The field is conditional probability and expectation.

Let X be a r.v. on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$, and $A \subseteq \mathcal{A}$. Then the conditional probability,

$$\mathbb{P}(X|A) \stackrel{def}{=} \frac{\mathbb{P}(1_A(\omega)\, X(\omega))}{E[1_A(\omega)]},$$

where $1_A(\omega)$ is the indicator function

$$1_A(\omega) = \begin{cases} 1 & \text{if } \omega \in \text{A} \\ 0 & \text{else,} \end{cases}$$

can be interpreted as the probability of X knowing that $\omega \in A$, and $1_{\mathcal{A}} = \Omega$. In the same way the conditional expectation,

$$E[X|A] \stackrel{def}{=} \frac{E[1_A(\omega)\, X(\omega)]}{E[1_A(\omega)]}$$

can be interpreted as the expectation of X given that $\omega \in A$. A more formal definition of the conditional expectation is provided.

Definition A.11 (Conditional expectation w.r.t. $\mathcal{A}$)**:**
Let X be a r.v. s.t. $E[|X|] < \infty$. Then the expected value of X given $\mathcal{A}$ is the unique r.v. Y s.t.

$$E[1_A X] = E[1_A Y]$$

for all $A \in \mathcal{A}$, and $\{Y \leq q\} \in \mathcal{A}$ for all $q \in \mathbb{R}$. The r.v. Y is denoted by $E[X|\mathcal{A}]$.

Properties A.12 (Conditional Expectation)**:**
Let X,Y be two r.v., $\alpha, \beta \in \mathbb{R}$ and $\mathcal{A}, \mathcal{B}$ be two σ-algebras. Then

1. $E[\alpha X + \beta Y|\mathcal{A}] = \alpha E[X|\mathcal{A}] + \beta E[Y|\mathcal{B}]$ (Linearity)
2. $E[E[X|\mathcal{A}]] = E[X]$ (Rule of double expectation)
3. $E[X|\mathcal{A}] = X$, if X is independent of $\mathcal{A}$
4. $E[X|\mathcal{A}] = E[X]$, if X is a r.v. on $(\Omega, \mathcal{A}, \mathbb{P})$
5. $E[X|\mathcal{A}] = E[E[X|\mathcal{B}]|\mathcal{A}]$ if $\mathcal{A} \subseteq \mathcal{B}$

A.4 Stochastic processes

The short rate change by time. Therefore the most general way to define the short rate is as a stochastic process r_t. But in order to define a stochastic process we need to know the measurability for each time t. The measurability for each time t is controlled by the filtration.

Definition A.13 (Filtration)**:**
Let $\{\mathcal{F}_t\}_{0 \leq t \leq T}$ be a family of σ-algebras on $(\Omega, \mathcal{F}, \mathbb{P})$ s.t.

$$\mathcal{F}_{t_1} \subset \mathcal{F}_{t_2} \subset \ldots | \subset \mathcal{F}$$

for all $0 \leq t_1 < t_2 \leq ... \leq T$. Then $\{\mathcal{F}_t\}_{0 \leq t \leq T}$ is called the filtration on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

The smallest necessary information from the path is called a Markov process. A Markov process have the property that the forthcoming trajectory only depends on today's state. This means that we can reduce the filtration $\mathcal{F}_t$ to today's observation r_t.

Definition A.14 (Stochastic Process)**:**

Let $\mathcal{T} \neq \emptyset$ be the parameter space. Then the collection

$$\{r_t\}_{t\in\mathcal{T}}$$

of random variables is called Stochastic Process on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$.

Example A.1 (Parameter Space)**:**

E.g. could the parameter space be $\mathcal{T} = [0, T]$ or $\mathcal{T} = \{1, 2, \ldots, n\}$.

Definition A.15 (The Markov Property; 1-dim.)**:**

Assume that a one dimensional stochastic process r_t starts at x. Then r_t has the Markov family property if

$$E^x[f(r_{t+h})|\mathcal{F}_t] = E^{r_t}[f(r_h)],$$

for all bounded and Borel-measurable functions f.

A.5 Brownian motion and Itô Integral

One of the most used stochastic processes is called Brownian motion. This is a process that fluctuates due to a Gaussian distribution. With the properties of independent, stationary increments and stochastic continuity[6], Brownian motion is an element in a broader class of stochastic processes. Namely the Levy Processes.

Definition A.16 (Brownian motion)**:**

A stochastic process $(W_t)_{0\leq t\leq T}$ on the probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\leq t\leq T}, \mathbb{P})$ is called a Brownian motion if the process has the following properties

1. $W_0 = 0$ $\mathbb{P}$-a.s.

2. Independent increments: i.e. for $0 \leq t_1 \leq \cdots \leq t_n$

$$W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_{t_n} - W_{t_{n-1}}$$

[6]Stochastic continuity is a condition that preserves jumps at fixed time. But in fact Brownian motion W_t has a continuous version which means no jumps. Hence it follow the Stochastic continuity condition.

are independent of each other

3. Stationary increments: i.e. $W_t - W_s$ is equally distributed as W_{t-s}

$$W_t - W_s \overset{d}{=} W_{t-s}$$

4. Gaussian distributed[7]: $W_{t-s} \sim Gaussian(0, t-s)$,

where the filtration $\mathcal{F}_t$ is the smallest σ-algebra containing the path of the process W_t up to time t

$$\mathcal{F}_t \overset{def}{=} \sigma(W_u; 0 \leq u \leq t)$$

Because of the independence[8] and stationary[9] property, the Brownian motion is a process without memory. Hence a Brownian motion is a Markov process.

The stochastic integral w.r.t. a Brownian motion, W_t, also called an Itô integral, is defined on the set of processes, $f \in \mathcal{V}[0,T]$ as

$$\mathcal{I}[f] = \int_0^t f_s dW_s,$$

where the set of processes $\mathcal{V}[0,T]$ have the following requirements: A function $f \in \mathcal{V}[0,T]$ is the mapping

$$f : \Omega \times [0,T] \longmapsto \mathbb{R},$$

where

- f is measurable w.r.t. $\mathcal{F} \otimes \mathcal{B}([0,T])$
- f_t is $\mathcal{F}_t$-adapted
- $E\big[\int_0^T f_s^2 ds\big] < \infty$

[7] The Gaussian distribution is the same distribution as the well known Normal distribution

[8] Independent on the past

[9] The past doesn't change the distributional properties

The last requirement ensure the existence of the Itô-isometry

$$E\Big[\Big(\int_0^T f_s dW_s\Big)^2\Big] = E\Big[\int_0^T f_s^2 ds\Big]$$

Properties A.17 (Itô integral)**:**

The stochastic integral w.r.t. a Brownian motion have the following properties

1. (Linearity) $\int_0^T \alpha f_s + \beta g_s dW_s = \alpha \int_0^T f_s dW_s + \beta \int_0^T g_s dW_s$
2. $E\big[\int_0^T f_s dW_s\big] = 0$
3. (Itô Isometry) $Var\big[(\int_0^T f_s dW_s)^2\big] = E\big[\int_0^T f_s^2 ds\big]$
4. There exists a continuous version of $Y_t \overset{def}{=} \int_0^t f_s dW_s$

By property 4. we may assume that the process Y_t is continuous.

A.6 Itô's Lemma

Assume that X_t is an Itô process, i.e. that

$$X_t = X_0 + \int_0^t h_s ds + \int_0^t f_s dW_s,$$

where $f_s \in \mathcal{V}[0,T]$ and $t \in [0,T]$, and h_s is $\mathcal{F}_t$-adapted with the property

$$E[\int_0^t |h_s| ds] < \infty.$$

Then, which form does the function,

$$g(t, X_t),$$

take? The answer is Itô's Lemma.

Theorem A.18 (Itô's Lemma)**:**
Let X_t be an Itô process and $g \in C^2([0,\infty] \times \mathbb{R})$[10]. Then the function $g(t,x)$ is an Itô process for $x = X_t$ as the underlying process, where

$$g(t,X_t) = g(0,X_0) + \int_0^t \frac{\partial}{\partial t} g(s,X_s)ds + \int_0^t \frac{\partial}{\partial x} g(s,X_s)dW_s + \frac{1}{2}\int_0^t \frac{\partial^2}{\partial x^2} g(s,X_s)(dX_s)^2,$$

using that $(dW_t)^2 = dt$ and $dW_t\,dt = dt\,dW_t = dt\,dt = 0$.

Note that the stochastic differential equation (SDE) of the Itô process is

$$dX_t = h_t\,ds + f_t\,dW_t.$$

A.7 Stochastic Differential Equations

Assume we have a SDE on the form

$$dX_t = h(t,X_t)dt + f(t,X_t)dW_s.$$

Then the solution is defined as the process X_t, in the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$, satisfying the equation

$$X_t = x_0 + \int_0^t h(s,X_s)ds + \int_0^t f(s,X_s)dW_s,$$

where $h : [0,T] \times \mathbb{R} \to \mathbb{R}$ and $f : [0,T] \times \mathbb{R} \to \mathbb{R}$ are given. If

$$|h(t,x)| + |f(t,x)| \leq C(1 + |x|)$$

and

$$|h(t,x) - h(t,y)| + |f(t,x) - f(t,y)| \leq D|x - y|$$

[10]For a $g \in C^2([0,\infty) \times \mathbb{R})$, the function is twice differentiable on $(0,\infty) \times \mathbb{R}$

for all $x, y \in \mathbb{R}$ and $t \in [0, T]$, the solution X_t exist and is unique. Knowing that a solution is unique makes us being able to do certain transformations in order to derive an appealing solution. See the Vasicek model present earlier.

A.8 Important Financial Tools

One crucial modeling assumption is that the discounted financial market is following an Martingale. It's an assumption of fair price. Heuristically the Martingale property is the model assumption that the best predicted value in the future is today's value.

Definition A.19:
Let X_t be a stochastic process on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\leq 0}, \mathbb{P})$. Further assume that $E[|X_t|] < \infty$ for all $t \in [0, T]$ and

$$E[X_t|\mathcal{F}_s] = X_s$$

for all $s \in [0, t]$. Then X_t is called a $\mathcal{F}_t$-Martingale.

For a stochastic process following the path of a stochastic integral w.r.t. a Brownian motion the Martingale representation theorem is an important tool in understanding the process under arbitrage free (fair) prices.

Theorem A.20 (Martingale Representation theorem)**:**
Let W_t be an Brownian motion, and assume that $M_t \in L^2(\mathbb{P})$ is a $\mathcal{F}_t$-Martingale. Then there exists a unique stochastic process $f \in \mathcal{V}[0, t]$ for all $t \geq 0$ s.t.

$$M_t = E[M_0] + \int_0^t f_s \, dW_s.$$

Note that assuming $M_t \in L^2(\mathbb{P})$ is the same as assuming existence of the second moment. This condition can be relaxed. Then we are working with local Martingales, where the main difference is that we need to work with stopping times, knowing that for a stopped process the second moment exist.

In most of the cases the financial process have a drift under the objective probability measure $\mathbb{P}$. But a necessity in financial modeling is that the discounted financial asset is priced due to a Martingale. We therefore need a change of measure. Using Girsanov transform we replace the objective probability measure by an appropriate chosen probability measure $\mathbb{Q}$. The transform will remain a Brownian motion and we choose this measure s.t. the process is a Martingale.

Theorem A.21 (Girsanov's theorem)**:**
Let $\tilde{W}_t$ be an Itô process on the form

$$\tilde{W}_t = W_t - \int_0^t h_s ds.$$

We define the likelihood process as

$$Z_t(h_t) = \exp\Big\{ \int_0^t h_s dW_s - \frac{1}{2}\int_0^t h_s^2 ds\Big\}$$

for $t \in [0, T]$. Further assume that h_t satisfies the Novikov condition

$$E\Big[\exp\Big\{\frac{1}{2}\int_0^t h_t^2 ds\Big\}\Big] < \infty$$

and that the Girsanov transformation $\mathbb{Q}$ of the measure $\mathbb{P}$ is defined by the probability measure

$$\mathbb{Q}(A) \overset{def}{=} E_{\mathbb{P}}[1_A Z_T].$$

Then $\tilde{W}_t$ is a Brownian motion under $\mathbb{Q}$.

Generally, the Brownian motion under the "risk neutral" measure $\mathbb{Q}$ will be denoted $W_t^{\mathbb{Q}}$.

Under the risk neutral measure we are going to derive several partial differential equation of the solution to the arbitrage free price of ZCB (also applied in deriving arbitrage free price for an option on the underlying process of a ZCB). An auxiliary tool is the Feynman-Kac proposition.

Proposition A.22 (Feynman-Kac)**:**
Let X_t be the solution of the stochastic differential equation

$$dX_t = h(t, X_t)dt + f(t, X_t)dW_t^{\mathbb{Q}},$$

where $X_0 = x_0$. Then for an risk free interest rate r the solution $F(t,x)$ of the boundary value problem

$$\begin{aligned}&\frac{\partial F}{\partial t}(t,x) + h(t,x)\frac{\partial F}{\partial x}(t,x) + \frac{1}{2}f(t,x)^2\frac{\partial^2}{\partial x^2}(t,x) - rF(t,x) = 0\\ &F(T,x) = \phi(x)\end{aligned}$$

is

$$F(t,x) = \mathrm{e}^{-r(T-t)}E_{\mathbb{Q}}[\phi(X_T)|\mathcal{F}_t],$$

given that the second moment exist for the integrand in the stochastic part of $F(t, X_t)$.

The integrability condition ensure that the stochastic integral is equal to zero. Feynman-Kac is used in the portfolio setup, Chapter 2.

The last theorem is very often used in this thesis. The Stochastic Fubini theorem serve in the same way as the Fubini theorem. It's a change of inner integrands.

Theorem A.23 (Stochastic Fubini theorem)**:**
Consider a stochastic process $f(u,s) \in \mathcal{V}[0,T]$. Then

$$\int_0^T\int_0^T f(u,s)dW_u ds = \int_0^T\int_0^T f(u,s)dsdW_s.$$

Proof. Theorem (6.2) in [3], where the assumption of $f(u,s) \in \mathcal{V}[0,T]$ is a special case of the proof. □

For a more thorough review I refer to [9] and [1].

Appendix B

B.1 Calculations

B.1.1 Deriving the CIR zero-coupon price

First we find the function $B(t,T)$. We solve the first differential equation

```
DSolve[D[B[t, T], t] - k B[t, T] - 1/2 σ^2 B[t, T]^2 == -1, B[t, T], {t, T},
 GeneratedParameters → (Subscript[c, #] &)]
```

$$\left\{\left\{B[t,T] \to \frac{-k+\sqrt{-k^2-2\sigma^2}\,\mathrm{Tan}\left[\frac{1}{2}\left(t\sqrt{-k^2-2\sigma^2}+2\sqrt{-k^2-2\sigma^2}\,c_1[T]\right)\right]}{\sigma^2}\right\}\right\}$$

w.r.t to the boundary condition $B(T,T) = 0$

```
Solve[Sqrt[-k^2 - 2 σ^2] Tan[1/2 (T Sqrt[-k^2 - 2 σ^2] + 2 Sqrt[-k^2 - 2 σ^2] c1)] == k, c1]
```

$$\left\{\left\{c_1 \to \mathrm{ConditionalExpression}\left[\frac{1}{2}\left(-T+\frac{2\left(\mathrm{ArcTan}\left[\frac{k}{\sqrt{-k^2-2\sigma^2}}\right]+\pi\,C[1]\right)}{\sqrt{-k^2-2\sigma^2}}\right),\ C[1]\in \mathrm{Integers}\right]\right\}\right\}$$

Since this applies for all $C[1]$, element in the integers, we choose the constant equal to zero. By transforming the equation to exponential functions and simplifying it by using that

$$\sqrt{-k^2-2\sigma^2} = \mathrm{i}\sqrt{k^2+2\sigma^2} = \mathrm{i}\,h\,^{1},$$

[1] $\mathrm{i} \overset{def}{=} \sqrt{-1}$

we end up with the following function for $B(t,T)$,

$$B(t,T)=\frac{2(\mathrm{e}^{h(T-t)}-1)}{2h+(k+h)(\mathrm{e}^{h(T-t)}-1)}. \tag{B.1}$$

The next thing would be to check if we in fact have the correct answer. The analytical tools of Mathematica provid a satisfying answer.

```
B[t_, T_] := (2 (Exp[(T - t) Sqrt[k^2 + 2 σ^2]] - 1)) /
   (2 Sqrt[k^2 + 2 σ^2] + (k + Sqrt[k^2 + 2 σ^2]) (Exp[(T - t) Sqrt[k^2 + 2 σ^2]] - 1))

FullSimplify[D[B[t, T], t] - k B[t, T] - 1/2 σ^2 B[t, T]^2]

-1

B[T, T]

0
```

The second differential equation we solve in the same way or just integrate $k\theta B(t,T)$ over t. We derive the following simplified solution

$$A(t,T)=\frac{k\theta(h+k)(T-t)+2\log[2h]-2\log[h-k+(h+k)\mathrm{e}^{h(T-t)}]}{\sigma^2}, \tag{B.2}$$

where we still have defined $h=\sqrt{k^2+2\sigma^2}$. By rearrenging and using the same definition of an affine term structure as [7] the result is equal.

B.1.2 Calculations; the Hull-White model

B.1.2.1

Recall that $B(s,T) = \frac{1}{a}\big(1 - \mathrm{e}^{-a(T-s)}\big)$. Then

$$\begin{aligned}
\frac{\partial}{\partial T}\int_0^T \theta_s B(s,T)ds &= \frac{\partial}{\partial T}\int_0^T \theta_s \frac{1}{a}\big(1 - \mathrm{e}^{-a(T-s)}\big)ds \\
\text{(Linearity)} &= \frac{\partial}{\partial T}\Big[\int_0^T \theta_s \frac{1}{a} ds - \mathrm{e}^{-aT}\int_0^T \frac{\theta_s}{a}\mathrm{e}^{-as}ds\Big] \\
&= \theta_T \frac{1}{a} - \frac{\partial}{\partial T}\Big[\mathrm{e}^{-aT}\int_0^T \frac{\theta_s}{a}\mathrm{e}^{as}ds\Big] \\
\text{(Product rule)} &= \frac{\theta_T}{a} + a\mathrm{e}^{-aT}\int_0^T \frac{\theta_s}{a}\mathrm{e}^{-as}ds - \mathrm{e}^{-aT}\frac{1}{a}\theta_T \mathrm{e}^{aT} \\
&= \int_0^T \theta_s \mathrm{e}^{-a(T-s)}ds \\
&= \int_0^T \frac{\partial}{\partial T}\theta_s \frac{1}{a}\big(1 - \mathrm{e}^{-a(T-s)}\big)ds \\
&= \int_0^T \frac{\partial}{\partial T}\theta_s B(s,T)ds.
\end{aligned}$$

B.1.2.2

Recall that

$$\psi(T) = f(0,T) + h(T),$$

where

$$\psi(T) \overset{def}{=} \int_0^T \frac{\partial}{\partial T}\theta_s B(s,T)ds - r_0 \mathrm{e}^{-aT}$$

and

$$h(T) \overset{def}{=} \frac{1}{2a^2}\sigma^2 \mathrm{e}^{-2aT}(\mathrm{e}^{aT} - 1)^2.$$

From Appendix B.1.2.1

$$\psi(T) = \int_0^T \theta_s \mathrm{e}^{-a(T-s)}ds.$$

Then

$$\begin{aligned}
\frac{\partial}{\partial T}\psi(T) &= \frac{\partial}{\partial T}(\mathrm{e}^{-aT}\int_0^T \theta_s \mathrm{e}^{as}ds + r_0 e^{-aT}) \\
(\text{Product rule}) &= -a\mathrm{e}^{-aT}\int_0^T \theta_s \mathrm{e}^{as}ds + \mathrm{e}^{-aT}\theta_T \mathrm{e}^{aT} + ar_0\mathrm{e}^{-aT} \\
&= \theta_T - a\psi(T)
\end{aligned}$$

B.1.3 Deriving the Vasicek model by means of the forward rate

An another way of deducing the same initial forward rate structure is by solving the differential equation

$$\frac{\phi'(t) + k\phi(t)}{k} = \theta$$

w.r.t. the initial forward rate $f(0,t)$

```
θ[t_] := f[t] + σ^2/(2 k^2) (1 - Exp[-k t])^2

DSolve[D[θ[t], t] + k θ[t] / k == ϕ, f[t], t]

{{f[t] → (e^(-k t) (-(e^(-k t) σ^2)/k - (e^(k t) (σ^2 - 2 k^2 ϕ))/k))/(2 k) + e^(-k t) C[1]}}
```

Simplifying the expression

$$f(0,T) = \theta - \frac{\sigma^2(1+e^{-2T})}{2k^2} + e^{-kt}C_1$$

and solving the initial condition $f(0,0) = r_0$,

```
Solve[(e^(-k 0) (-(e^(-k 0) σ^2)/k - (e^(k 0) (σ^2 - 2 k^2 ϕ))/k))/(2 k) + e^(-k 0) c1 == r, c1]

{{c1 → (k^2 r + σ^2 - k^2 ϕ)/k^2}}
```

we derive the same specified initial forward rate

$$f(0,t) = r_0 e^{-kt} + \theta(1-e^{-kt}) - \frac{\sigma^2}{2k^2}\left(1-e^{-kt}\right)^2.$$

Remark B.1:

In the program it was necessary to split between the constant and function. Therefore in the program $\phi(t) = \theta(t)$ and $\theta = \phi$.

Bibliography

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

[2] Damir Filipovic. *Consistency problems for Heath-Jarrow-Morton interest rate models.* Number 1760. Springer, 2001.

[3] Damir Filipovic. Term-structure models: A graduate course, 2009.

[4] Leszek Gawarecki and Vidyadhar Mandrekar. *Stochastic Differential Equations in Infinite Dimensions: With Applications to Stochastic Partial Differential Equations.* Springer, 2010.

[5] Paul C Kettler, Frank Proske, and Mark Rubtsov. Sensitivity with respect to the yield curve: duration in a stochastic setting. In *Inspired by Finance*, pages 363–385. Springer, 2014.

[6] Sindre Duedahl. Stochastic yield curve duration and portfolio immunization strategies. *Working paper*, 2014.

[7] Damiano Brigo and Fabio Mercurio. *Interest rate models-theory and practice: with smile, inflation and credit.* Springer, second edition, 2007.

[8] René Carmona and Michael R Tehranchi. *Interest Rate Models: an Infinite Dimensional Stochastic Analysis Perspective.* Springer, 2007.

[9] Bernt Øksendal. *Stochastic differential equations.* Springer, sixth edition, 2003.

[10] Giulia Di Nunno, Bernt Øksendal, and Frank Proske. *Malliavin calculus for Lévy processes with applications to finance.* Springer, 2008.

[11] John Hull and Alan White. The optimal hedging of interest rate sensitive securities. *University of Toronto*, 1994.