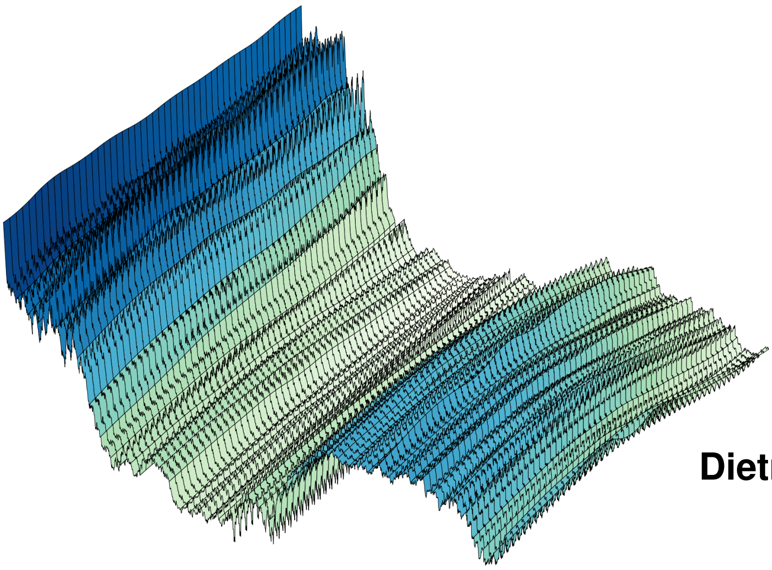
KU LEUVEN

FACULTY OF SCIENCE

A Path Integral & Field Theory Formalism for Interest Rates and Zero Coupon Bond Options

A Physicist's Approach



Dietmar DEJAEGHER

Supervisor: Prof. Ewald Janssens
Institute for Quantum Solid-State Physics,
KU Leuven

Co-supervisor: Frank De Jonghe, PhD
Financial Quantitative & Analytics Services,
EY

Thesis presented in
fulfillment of the requirements
for the degree of Master of Science
in Physics

Academic year 2022-2023

Abstract

Many models attempting to describe interest rates' dynamics assume an exact correlation between rates for different maturities. However, such a trivial correlation structure is not observed in market data and allows for troublesome arbitrage situations. The aim of this master's thesis is to incorporate a non-trivial correlation structure between the instantaneous forward interest rates at different maturities by means of a stochastic field. To achieve this, the most relevant concepts in finance and physics are introduced in such a way that reveals their significant interconnectedness. Two of the most well-known interest rate models, *Vasicek* and *Heath-Jarrow-Morton* (HJM), are described and translated to a path integral formulation. Closed-form pricing formulas for the zero coupon bond and a European call option on the zero coupon bond are rederived together with the no-arbitrage condition on the drift term of the HJM model. This description eases the transition to the field theory extension in which the stochastic nature of the instantaneous forward interest rates is governed by a statistical field. This is achieved by allowing for fluctuations in the *maturity direction* as well as in the *present time direction* into a one-factor HJM model. The no-arbitrage condition on the drift and pricing formula for the zero coupon bond option are rederived in a formalism that is familiar from quantum field theory. The theoretical results are tested against real market data allowing for a comparison of the models. A calibration to caplet prices shows that the field theory has almost double the accuracy of a one-factor HJM model in terms of the *root mean square error* (RMSE). The field theory model can be further improved by, for example, including higher order terms in the field's Lagrangian. Other possible outlooks are discussed as well.

Acknowledgments

At the start of my journey as a physics' student at KU Leuven I would have never thought my master's thesis would have anything to do with (theoretical) finance. I have always had the pleasure and privilege to study that what I find utmost interesting and that was no different when writing this thesis. Nevertheless, it was not always a walk in the park and I definitely encountered many difficulties along the way. Therefore, I would like to thank all the people that have helped me conquer these struggles and everyone that made this final work possible.

First of all, I want to express my eternal gratitude to my parents for their everlasting support throughout my studies, their endless patience (as I have not always been a role model student) and their numerous attempts in trying to understand me (obscurely) explaining something new I learned and thought was worth sharing.

Next, I would very much like to thank my supervisors professor Ewald Janssens and Dr Frank De Jonghe for taking the time to oversee my progress and help me to stay focused on what was most relevant. I also want to thank them for providing me with the much appreciated feedback and suggestions which guided me in trying to create a coherent story between the different subjects.

I would also like to thank KBC Bank for providing me with the necessary data for this thesis. This data was essential in obtaining my numerical results. Furthermore, I want to thank Ivo De Saedeleer (Head of Quants at KBC Bank) and the rest of the entire quant team at KBC Bank for giving me the opportunity to join the team during a year long internship, for making me feel welcome and helping me with my every question. I learned so much and gained a great amount of hands-on experience. I also want to specially thank Dr Yves De Wulf (interest rate quant at KBC Bank) for the many clear explanations on fascinating topics concerning interest rates and so much more. This greatly helped me forming an idea of how experts in the field of interest rates look at the bigger picture.

Last but certainly not least, I want to express my profound gratitude and immense appreciation to Eva for always supporting and comforting me with her own thesis-stress.

To all these people and everyone I undoubtedly forgot, I sincerely want to say: thank you!

Summary in Layman's Terms

The past couple of years, one was inundated with news on falling (and rising) stock prices, energy crises, seemingly ever-increasing inflation numbers, imminent financial recessions, the debt ceiling of the United States of America and Central Banks upping interest rates. A study of the stock market's and interest rates' behavior might thus be more relevant than ever. In this master's thesis, we focus on the latter. We explain two ways of modeling interest rates frequently used by practitioners and indicate their relation to popular theories in physics. Most of these models are based on *diffusion*, a special type of random process. The idea of diffusion is not something spectacularly complicated or abstract. In fact, we encounter examples of diffusion in our every day lives when we pour milk into our cup of coffee, when steam from a hot shower spreads out in the bathroom or when smoke of a blown-out candle dissipates. Diffusion-based interest rates models can be used to calculate the prices of certain financial products. However, they have one major flaw: they fail to account for the relationship, or *correlation*, seen in the market between the rates on loans over different periods of time.

In this thesis work, we included this correlation by deploying a particularly fruitful theoretical formalism from physics: *quantum field theory*. The mathematical tools developed for (and by) quantum field theory proved extremely useful for including this missing relationship in a parsimonious manner. To make the transition to this new model as smooth as possible, we first introduce a *path integral*, which is just a fancy way of saying that we take into account every possible path that interest rates can take as they change from one value to another. These path integrals allow us to "disguise" certain interest rate models such that they are more easily treated with the tools from quantum field theory. Using these tools, we can again calculate the price of those financial products such that the theoretical results can be compared.

Finally, we check how well all these models can predict prices of interest rate products that are traded in the market by looking at the numerical prices. This shows how even the simplest version of a model built upon these "quantum field theory tools" already has an incredible accuracy in matching market prices. We also suggest possible future research topics that could further deepen the foundations of this way of dealing with stochastic interest rates.

Beknopte Samenvatting

De voorbije paar jaar zijn nieuwsfeiten over dalende (en stijgende) aandelen koersen, hoge inflatie cijfers, dreigende financiële recessies, het schuldenplafond van de Verenigde Staten van Amerika en de Centrale Banken die de rente verhogen moeilijk weg te denken. Een onderzoek over het gedrag van de aandelenmarkt en de rente is dus misschien relevanter dan ooit tevoren. In deze master thesis focussen we op het tweede. We leggen twee manieren voor het modelleren van rentes, die frequent gebruikt worden, uit en duiden hun verband met populaire theorieën uit de fysica. De meeste van deze modellen zijn gebaseerd op een specifiek type willekeurig proces, genaamd *diffusie*. Het concept van diffusie is niet iets ongelofelijk ingewikkeld of abstract. Integendeel, we komen er regelmatig voorbeelden van tegen in ons dagelijks leven zoals wanneer, bijvoorbeeld, melk in een kop koffie wordt gegoten, de badkamer wordt gevuld met stoom tijdens een warme douche of rook verspreidt na het uitblazen van een kaars. Deze op-diffusie-gebaseerde rente modellen kunnen gebruikt worden om specifieke financiële instrumenten te prijzen. Ze hebben echter een grote tekortkoming: ze slagen er niet in de relatie, of *correlatie*, die te zien is in de markt, tussen rentes op leningen over verschillende tijdsperiodes in rekening te brengen.

In deze thesis verwerken we die correlatie door een buitengewoon nuttig theoretisch formalisme uit de fysica in te zetten. Dit formalisme is dat van *kwantumveldentheorie*. Het wiskundige gereedschap dat is voortgekomen uit kwantumvelden theorie is bijzonder handig gebleken wanneer deze ontbrekende relatie wordt geïntegreerd. Om de transitie naar dit nieuwe model zo vlot mogelijk te laten verlopen introduceren we eerst een *padintegraal*, wat een stoere manier is om te zeggen dat we rekening houden met ieder pad dat de rente kan volgen wanneer het van de ene waarde naar de andere verandert. Met deze padintegraal kunnen we de gebruikelijke rentemodellen “vermommen” zodat ze gemakkelijker bruikbaar zijn voor het gereedschap van kwantumveldentheorie. We kunnen dan met dit gereedschap de prijzen van diezelfde financiële instrumenten berekenen om zo de theoretische resultaten te vergelijken.

Tenslotte zullen we nagaan hoe goed de prijzen die deze modellen voorspellen overeenkomen met marktprijzen van financiële producten die afhankelijk zijn van de rente. Hiermee tonen we aan dat zelfs het meest eenvoudige model, opgebouwd met dat “kwantumvelden-gereedschap”, al met ongelofelijke nauwkeurigheid prijzen kan berekenen van rente producten die overeenkomen met markprijzen. Verder stellen we nog mogelijke onderwerpen voor toekomstig onderzoek voor, om zo de funderingen van deze kijk op het werken met stochastische rentes te kunnen uitdiepen.

List of Figures

Symbols

$\mathcal{S}$	action
$\mathcal{A}(t,x)$	statistical field
$S(t)$	asset price process
$C(t,S;K,s)$	price at time t of a European call option on an underlying asset S with strike K and expiry s
$B(t_1,t_2)$	risk-free cash bank account
$\mathcal{P}_{n\|m}$	conditional probability density
$\mathcal{B}(t,s)$	coupon bond price
$\Phi(d)$	standard normal cumulative distribution function
$D(t_1,t_2)$	discount factor for discounting from time t_2 to t_1
Δ_0	region $\left\{(x,t)\in\mathbb{R}^2 \middle\| t_0\le t\le x \text{ and } t_0\le x\le t_*\right\}$
$\delta(x)$	Dirac delta
$\alpha(t,x)$	risk-neutral drift
$\eta(t)$	Gaussian noise as "derivative" of Wiener process
$\mathbb{E}$, $\langle\ \rangle$	expectation value
$\mathcal{F}_t$	filtration
$\mathcal{F}$	region $\left\{(x,t)\in\mathbb{R}^2 \middle\| t_i\le t\le t_f \text{ and } t\le x\le t+T_{FR}\right\}$
$\mathcal{D}[.]$	functional integration measure
$F(t;s_1,s_2)$	forward rate
$\mathcal{P}_n$	hierarchy of probability densities
$\mathbb{I}$	identity operator
$f(t,x)$	instantaneous forward rate
i	imaginary unit
$\mathcal{L}$	Lagrangian
ω_i	market state
Z	partition function
$\mathcal{Z}$	partition functional
$V^A(t,\omega_i)$	value of portfolio A at time t for market state ω_i
$\mathcal{P}$	probability distribution
$G(x,x';t)$	propagator
$P(t,S;K,s)$	price at time t of a European put option on an underlying asset S with strike K and expiry s
μ	rigidity parameter
$r(t)$	risk-free rate
$\mathcal{R}$	region $\left\{(x,t)\in\mathbb{R}^2 \middle\| t_0\le t\le t_* \text{ and } t_*\le x\le s\right\}$
J	source function
$\Theta(t)$	step function
K	strike value of an option
$\mathcal{T}$	region $\Delta_0\oplus\mathcal{R}$

T_{FR}	maximum duration for instantaneous forward interest rates found in the market
$\mathcal{W}$	transition probability per unit time
σ	volatility
$W(t)$	Wiener process
$R(t,T)$	yield to maturity at time t over a duration T
$\mathrm{ZCB}C(t;K,t_*,s)$	price at time t of a European call option on a zero coupon bond $P(t,s)$ with strike K and expiry t_*
$\mathrm{ZCB}P(t;K,t_*,s)$	price at time t of a European put option on zero coupon bond $P(t,s)$ with strike K and expiry t_*
$P(t,s)$	price at time t of a zero coupon bond with maturity s

Acronyms

ATM	*at the money*
CKE	*Chapman-Kolmogorov equation*
CMS	*constant maturity swap*
DCF	*discount factor*
FPE	*Fokker-Planck equation*
HJM	*Heath-Jarrow-Morton*
ITM	*in the money*
O-U	*Ornstein-Uhlenbeck*
OTM	*out the money*
PDE	*partial differential equation*
QFT	*quantum field theory*
QM	*quantum mechanics*
RHS	*right-hand-side*
RMSE	*root mean square error*
SDE	*stochastic differential equation*
ZCB	*zero coupon bond*

Contents

Chapter 1

Introduction

It is in the nature of scientists to apply their expertise to problems in other fields. Such interdisciplinary research has led to numerous discoveries, new insights and even technological advancements. A few examples are the discovery of the structure of the DNA, where the developed techniques of X-ray diffraction played a crucial role [41], or our increased understanding of climate change as a consequence of the work on complex physical systems, which was awarded the 2021 Nobel prize in physics [28]. This feature of scientific curiosity and the urge to deploy your own skill set to seemingly unrelated problems also holds for physicists. One of such fields where physicists (and mathematicians) have been thriving in specifically is the field of (mathematical) finance. This can be seen from the many widely used financial models developed by physicists and mathematicians.

The most prominent example, although not introduced by physicists, is undoubtedly the Black-Scholes-Merton model, commonly known as the Black-Scholes model, named after its conceivers, Fischer Black (applied mathematician), Myron Scholes (financial economist) and Robert C. Merton (economist). The model was first introduced[1] in 1973 [11] and resulted in a Nobel prize in economics in 1997 for Scholes and Merton. Due to his death in 1995, Fischer Black was, sadly, ineligible to be awarded the prestigious prize, however, his contributions were not unacknowledged, as the model name suggests. The Black-Scholes model attempts to describe the dynamics of stock prices in financial markets given certain assumptions. The stock price process in this model very much resembles a diffusion process. In fact, the Black-Scholes equation, which is a *partial differential equation* (PDE) within this model, is, up to a transformation, completely equivalent to the heat diffusion equation [21]. The Black-Scholes equation is the PDE within this model that governs the evolution of the price of a *financial derivative*[2]. Its similarity to the heat equation makes it fairly straightforward to solve, given certain boundary conditions. The model is still, to this day, commonly used in finance to obtain (quick estimates for) option prices. We will discuss the Black-Scholes model and its properties in more detail later on in section 2.1.4.

A second well-known example of a mathematical model used for modeling stock price behavior and valuating financial derivatives are the Lévy processes, named after French mathematician Paul Lévy. This "newer" way of modeling financial markets, with respect to the Black-Scholes model, was proposed in the late 1980s and early 1990s and has shown better consistency with

[1] It was actually Louis Bachelier who introduced the random walk as a description for stock price movements in 1900. However, he did so in a slightly different context [7].

[2] A financial derivative, in short, is a product that *derives* its value from the price of some underlying asset. We kindly ask the reader not to worry about any financial terminology that remains unclear at this moment. Chapter 2 contains detailed explanations of the most important terms and is included after the introduction. Here, we will merely try to sketch the relevance of this thesis.

market data when it comes to option pricing. Furthermore, the increasing interest and activity of strong mathematical minds in the field of finance has led to exciting progress in the study of Lévy processes with various new insights in applied modeling as a result [50]. An extensive study of Lévy processes in finance is beyond the scope of this thesis.

A third example is the Vasicek interest rate model. It was named after Czech mathematician Oldřich Vašíček and first introduced in 1977. Vašíček aimed to model the *short rate* by having its dynamics driven by some source of *market risk*. It was the first interest rate model to include *mean-reversion* and has been quite useful in the pricing of interest rate derivatives [56]. Due to its simplicity, the model did, however, have some undesirable features, such as time-independent parameters and relatively high probabilities for large negative interest rates. The latter was first resolved by John Cox, Jonathan Ingersoll and Stephen Ross in the so called Cox-Ingersoll-Ross model in 1985 [23]. The former problem was solved by the (perhaps even more) famous Hull-White model, after John Hull and Alan White [35]. The Hull-White model is sometimes called the "Extended Vasicek Model" since it incorporates a time-dependence into various parameters that are constant in the "regular" Vasicek model. Even today, the Hull-White model is still very popular. Although the Vasicek model has some (serious) shortcomings, it is nevertheless instructive and will be studied in more detail in section 2.3.1.

There are many more such cases of papers and books considering connections between physics and finance or economics [13, 53, 57, 61]. In fact, a whole new research field has been developed in recent years, named *econophysics* [54]. With this and the example of the Vasicek model in mind, we arrive at the reason as to why this and previous work on interest rate models makes sense: namely the fact that interest rates are not constant. Otherwise, there would be no need for models describing interest rate dynamics. But before we start to discuss the non-constant nature of interest rates, we need to talk about the *time-value of money*.

Due to various economic, social and political reasons, such as, for example, inflation, financial crises or currency devaluations, the true value of money is time-dependent. A euro today is worth less (depending on global economic conditions, worth more) than a euro one year from now. If one were to know by exactly how much the value of one euro will be depreciated over the upcoming year, one would have the opportunity to use this edge over his competitors and earn some money, as explained in the following thought experiment.
Consider for example the following extreme situation where Bob knows for certain that €100 today will be worth only the equivalent of €50 next year. This means that you could, for example, buy 10 expensive apples of €10 a piece today, but only 5 apples next year because they would then cost a current equivalent of €20 a piece. Suppose also that Alice, who is completely unaware of this enormous inflation, is prepared to give out a loan of €100 to Bob, but she expects Bob to pay her €110 back next year. Bob could then take out a loan from Alice, buy 10 apples today for €100 and consequently sell 6 apples next year[3] for €120. Bob would then have €120 and 4 apples, but he still owes Alice €110. After the loan is settled, Bob is left with €10 and 4 apples (each now worth €20). But Bob started out with no money. This is what is called an *arbitrage opportunity* [51].

Because of such arbitrage opportunities, it is of vital importance to financial institutions to measure or predict this deterioration in money value as accurately as possible. They ideally want to know the exact present value by *discounting* the future value of a sum of money, or, vice-versa, want to know the expected future value by *compounding* the present-day value.

If we were to encompass all the possible information regarding the devaluation of money into

[3] It is important for the reader to understand that this is a thought experiment and apples that preserve in good condition for one year actually do not exist.

a single quantity[4], we would have a simple (constant) *risk-free* interest rate. The fact that it is "risk-free" means that the loan is never defaulted and the money is always paid back [51]. For example, the interest rate on a cash account at big banks or on government bonds can be considered "risk-free". This is however a theoretical (but useful) idea as nothing is ever entirely free of risk (see for example the bankruptcy of Bear Stearns and Lehman Brothers). Nevertheless, let us denote this risk-free rate by r and consider a fixed deposit of an amount N made at a certain time $t \geq 0$. Assume the deposit accrues at the rate r such that, at a future time $s \geq t$, the value of the deposit has increased to

$$N(1+r)^T,$$

where $T = s - t$. In this example we gained a small amount of rN over the course of first year. The second year, we would gain $N(1+r)r$. The third year $N(1+r)^2 r$, and so on. This is called *discrete compounding* on an annual basis however one could also accrue the interest on a monthly, weekly or even daily basis. In the case we compound ε times per year, the deposit grows to

$$N\left(1+\frac{r}{\varepsilon}\right)^{\varepsilon T}.$$

Almost always and throughout this thesis as well, one assumes *continuous compounding*, i.e. we take the limit $\varepsilon \to \infty$ such that the deposit has the value

$$Ne^{rT} \tag{1.1}$$

at future time $s = t + T$. We will call the factor e^{rT} the *risk-free cash bank account* (or just cash bank account) as it represents the value one notional will have accumulated to on a cash bank account over the period $T = s - t$. The factor e^{-rT} on the other hand is a theoretical quantity called a *discount factor* (DCF) and was devised to describe the evolution of the time-value of money (or the time-value for any financial asset). We will later see, in section 2.2.3, that discounting is not always done with the same quantity and that different types of quantities with which we discount certain prices or values are called *numeraires* [12]. We will denote the cash bank account over a time period $[t_1, t_2]$ as $B(t_1, t_2)$. Note that $\forall\ 0 \leq t_1 \leq t_2 :\ 0 < B(t_1, t_2)$ and $B(t, t) = 1$ for all times t (this holds for any interest rate[5]). The DCF used for discounting over the time interval $[t_1, t_2]$ will be denoted $D(t_1, t_2)$. Since the DCF is the inverse of the cash bank account, we can interpret $D(t_1, t_2)$ as the amount of cash we must deposit at time t_1 on such a cash bank account to have 1 notional at future time t_2, hence $D(t, t) = 1 = B(t, t)$, for all times t.

The set of all DCFs at a certain time t for different *maturities* s (naturally, T is called the "time to maturity", sometimes also named the *tenor*) comprise a *discount curve* $s \mapsto D(t, s)$. Such a curve is often called a *term structure of DCFs*. One can see in the market that this term structure changes over time. Even more so, its evolution behaves stochastically. For a time-dependent (stochastic) interest rate $r(t)$, the DCFs (and cash bank account) would take the form

$$D(t_1, t_2) = B^{-1}(t_1, t_2) = e^{-\int_{t_1}^{t_2} r(t)\mathrm{d}t}.$$

Since the interest rate $r(t)$ is stochastic, both the DCF and cash bank account are also stochastic [15]. If one wants to obtain a specific value, an expectation value is needed. DCFs will be discussed in more depth in section 2.2.1. It is this rate $r(t)$, generally called the *spot rate*, that Vašíček aims to model. The different types of interest rates will be discussed in section 2.2.4. The Vasicek (and other) spot rate model is unfortunately only reliable for modeling the initial

[4]We will later see that this quantity has a lot more complexity to it.

[5]Note that under the dubious assumption of positive interest rates, one can also assume $\forall t_1 \leq t_2 : 1 \leq B(t_1, t_2)$.

or short-time curve. In fact, spot rate models only provide information about one point on the entire discount curve [5]. We will later see how David Heath, Robert Jarrow and Andrew Morton attempt to model the *entire* discount curve by investigating a more "fundamental" quantity, named *instantaneous forward rates*, denoted $f(t,s)$. However, it will become clear that, in this model, the forward rates at different maturities s are exactly correlated, i.e. $\forall\ t_1, t_2 \leq s_1, s_2 : \langle f(t_1, s_s) f(t_2, s_2)\rangle_c = \delta(t_1 - t_2)$ where $\langle .\ ,\ .\rangle_c$ denotes the connected-correlation function, and δ is the Dirac-delta. This exact correlation, however, is not something that is observed in the market data. On the contrary, we will show that the (empirical) correlation between instantaneous forward rates at a certain time t is rapidly decreasing across different tenors.

Our aim is to model the behavior of the DCFs (and by extension the discount curve) by considering these as "observables" arising from the dynamics of more the fundamental curve (and later a field). To achieve this, we will deploy concepts and formalisms we know and love from physics such as, Lagrangians, path integrals (both Wiener- and Feynman-), stochastic fields,... This will allow us to construct a formalism for the description of interest rates from a physics perspective and incorporate a non-trivial correlation between the forward rates in a parsimonious manner. To make sure that the context of this thesis as a whole is clear for both physicists and economists, the layout will be the following.

We start with a concise discussion of the core concepts in both finance (chapter 2) and physics (chapter 3) required to have a good understanding of the main ideas of this thesis. Thereafter, these concepts can be combined to translate the stochastic processes for interest rates to a path integral formalism in chapter 4. This path integral formulation allows for a relatively smooth transition to the construction of a field theory for forward interest rates in chapter 5. This field theory is then tested against real market data and compared with conventional models in chapter 6. A short summary of the most important results and insights of this thesis is given together with a brief discussion of further prospects in chapter 7.

Chapter 2

Finance for Physicists

This chapter will go over the most important financial concepts that are relevant for acquiring a decent understanding of the main ideas of this thesis. The explanations are aimed at physicists or mathematicians with a strong analytical foundation but with limited to no knowledge of finance. This also means we will try to make pertinent connections with physics. First, some of the most basic and fundamental concepts in finance together with the famous Black-Scholes model, that was mentioned in the introduction, are discussed. Second, we talk about the *zero coupon bond* (ZCB) and the different types of interest rates. Lastly, we describe the two main interest rate models relevant for this thesis and some their most important characteristics.

2.1 Basics & Black-Scholes

2.1.1 European Options

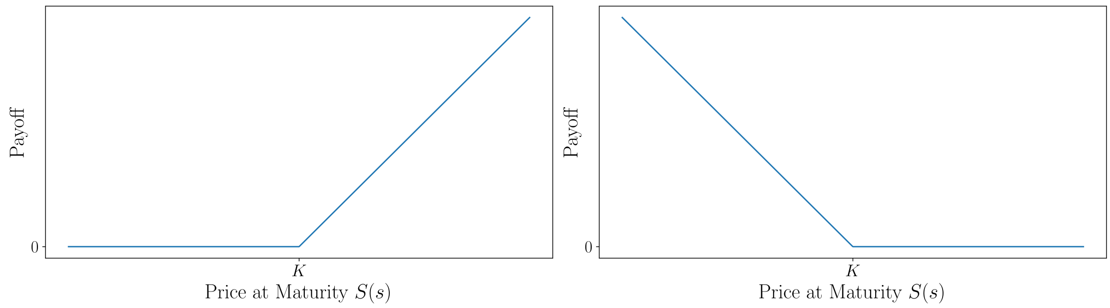
Financial derivatives, as already stated in the introduction, are financial products whose value is derived from the value of an underlying asset. The underlying assets are mostly stocks, but they could also be bonds, interest rates or even crude oil or corn. Derivatives play a big role in the global financial markets as they had a gross market value of around USD 18.6 trillion at the end of 2013, as estimated by the Bank for International Settlements [51]. Because of this sizable amount, it is only reasonable to discuss one of the most basic derivatives that are being traded: the *European call and put options.* The European call option is a financial contract that allows the holder to buy the underlying asset, S, from the issuer at a predetermined future time s, the *maturity* (or sometimes the *expiry*), for a predetermined price K, the *strike price* or strike for short. Note that the holder of the call option has the *right* and not the *obligation* to buy the underlying asset at maturity. We will denote the price of such a call option on an underlying S with maturity time s and strike K at time $t \leq s$ as $C(t, S; s, K)$. It is important to note that $S = S(t)$ also has an implicit time-dependence. The value of the underlying is always considered at the same time instance as is given as argument in the option price unless explicitly stated otherwise. The put option on the other hand allows the holder of the contract to sell the underlying asset to the issuer. Again, the holder has the right and not the obligation to do so. The price at time t, of a put option with expiry s and strike K is denoted $P(t, S; K, s)$. Since the holder has the right and not the obligation to buy/sell the underlying asset, or to *exercise* the option, we can safely assume he only does so when it would be beneficial. In other words, the call/put option will only be exercised when the price of the underlying at maturity $S(s)$ is higher/lower than the predetermined strike price K. The payoff at maturity for the call option is thus given as

$$C(s, S; K, s) = \max[S(s) - K, 0] \equiv [S(s) - K]^+, \tag{2.1}$$

while the payoff at maturity for the put option is given as

$$P(s, S; K, s) = \max[K - S(s), 0] \equiv [K - S(s)]^+. \tag{2.2}$$

A graphical representation of both payoff functions can be seen in fig. 2.1. When the price S of the underlying of a call option is below the strike K, we say the call option is *out the money* (OTM). On the other hand, when the price of the underlying is above the strike, the call option is *in the money* (ITM). When the price is equal to the strike, the call option is said to be *at the money* (ATM). Naturally, for the put option it is the other way around. The put option is OTM when the price of the underlying is above the strike, ITM when the price is below the strike and ATM when both are equal.



(a) (b)

Figure 2.1: Graphical representation of the payoff functions of a European call (a) and European put (b) with strike K in function of the price of the underlying at maturity $S(s)$.

We will later see that, under the market assumption of no-arbitrage, there exists a special relation between the price of a European call and European put option with equal maturity and strike on the same underlying asset, called *put-call parity* [12]. The put-call parity relation for options with a maturity s and strike K on an non-dividend paying asset S is given at a time $t \leq s$ as

$$C(t, S; K, s) - P(t, S; K, s) = S(t) - Ke^{-r(s-t)}, \tag{2.3}$$

where r is taken to be the so called risk-free interest rate (more on this in section 2.2.4). This relation allows one to calculate the price of a put option, given the price of the corresponding call option and vice versa.

2.1.2 No-Arbitrage & Other Market Assumptions

Just as is the case for (almost all) physical theories, certain assumptions are made to make a financial model tractable. A first assumption is that markets are *frictionless*. This means that trading (i.e. buying and/or selling) involves no transaction costs and there are no differences between bid- and ask prices, just one market price. There are also no taxes, transaction delays and interest rates are the same for lending as for borrowing [59]. A second assumption imposed on the market is that of an *efficient* market. In an efficient market, asset prices contain and reveal all available market information and thus every market participant has equal access to that information. This would imply that assets and derivatives are only traded at their “fair price” [34]. A third assumption is that the market is *complete*, meaning any derivative can be replicated by trading specific amounts of the underlying [59]. In section 2.1.3 we will see a very important consequence of this market assumption. A fourth assumption is that market participants are completely rational. One example of completely rational behavior is only exercising an option when it is ITM. A fifth market assumption is the assumption of no arbitrage.

No-Arbitrage

The assumption of no arbitrage loosely states that "one can not make money from nothing" or "there is no such thing as a free lunch". More formally, if we denote the value of a given portfolio A, i.e. a collection of assets, at time t as $V^A(t)$, then, for some future time $s \geq t$, the portfolio value $V^A(s)$ is a stochastic variable. We call portfolio A an *arbitrage portfolio* if, $V^A(t) \leq 0$ and there exists a future time $s \geq t$ such that $V^A(s) \geq 0$ with a non-zero probability for $V^A(s) > 0$. The no-arbitrage assumption then states that no such arbitrage portfolios exist.

Under the assumption of no-arbitrage, we can prove the *monotonicity theorem*, which in turn allows us to easily derive the put-call parity relation [12].

Monotonicity theorem. Under the no-arbitrage assumption, if two portfolios A and B have values $V^A(s,\omega_i) \geq V^B(s,\omega_i)$ at a time s for a certain market state ω_i, then $V^A(t) \geq V^B(t)$ for all times $t \leq s$, independent of the market condition. Additionally, if $V^A(s,\omega_j) > V^B(s,\omega_j)$ then $V^A(t) > V^B(t),\ \forall t < s$.

Proof. Consider a new portfolio $C = A - B$. Then $V^C(s,\omega_i) \geq 0$. Suppose there exist a time $t < s$ for which $V^C(t) = -\varepsilon < 0$. The portfolio C plus ε amount of cash is an arbitrage portfolio as it has zero value at time t but $V^C(s,\omega_i) + \varepsilon \geq \varepsilon > 0$. Therefore

$$V^C(t) \geq 0 \implies V^A(t) \geq V^B(t),\ \forall t < s.$$

If, on the other hand, $V^C(s,\omega_i) > 0$, then we require $V^C(t) > 0$ for $t < s$, otherwise C is an arbitrage portfolio. □

A natural consequence of the monotonicity theorem is the "*law of one price*", which states that, if there exist a time s for which $V^A(s,\omega_i) = V^B(s,\omega_i)$, then $V^A(t) = V^B(t)$ for all times t.

We are now ready to proof the put-call parity relation given in eq. (2.3). The proof uses a clever construction of different portfolios to arrive at the desired result. This is a frequently reoccurring method for proving theorems in mathematical finance.

Put-Call Parity. Under the assumption of no arbitrage, the European put and call with equal characteristics, i.e. same underlying S, maturity s and strike K, obey the relation given in eq. (2.3).

Proof. Consider a portfolio A containing one European call on an underlying S with strike K and maturity s plus Ke^{-rs} worth of cash. Also consider a second portfolio B containing one European put with the same characteristics as the call of portfolio A and one unit of the underlying S. At the maturity time s, the value of the portfolios are

$$V^A(s) = [S(s) - K]^+ + Ke^{-rs}e^{rs} = \begin{cases} S(s) & \text{if } S(s) > K \\ K & \text{if } S(s) \leq K \end{cases} = [K - S(s)]^+ + S(s) = V^B(s).$$

The portfolios have equal values at time $s \geq t$, which means that, by the monotonicity theorem they have equal value at all times t, i.e.

$$C(t, S; K, s) + Ke^{-r(s-t)} = P(t, S; K, s) + S(t).$$

This is exactly equal to eq. (2.3). $\square$

These market assumptions, especially the assumption of no-arbitrage, will allow us to neglect most inconveniences when introducing different models and deriving pricing formulas. No-arbitrage, for example, is closely related to the next concept: risk-neutral pricing of derivatives.

2.1.3 Risk-Neutral Pricing

To discuss the risk-neutral pricing we must first introduce the concepts of *martingales* and *martingale measures*. We call a probability measure Q a martingale measure for a stochastic process $X = \{X_i | i = 0, 1, \ldots, n\}$ if, for all $0 \leq i \leq n$, X_i obeys

- $\forall i \in \{0, 1, \ldots, n\}: \ \mathbb{E}^Q[|X_i|] < \infty$
- $\mathbb{E}^Q[X_n | X_0, X_1, \ldots, X_{n-1}] = X_{n-1}$.

We say X_i is a Q-martingale [51]. By using the identity for iterated expectation values

$$\mathbb{E}[X_n | X_0] = \mathbb{E}[\mathbb{E}[X_n | X_{n-1}] | X_0], \tag{2.4}$$

the martingale conditions yields

$$\mathbb{E}^Q[X_j | X_i] = X_i, \ \forall 0 \leq i \leq j \leq n. \tag{2.5}$$

When applying the definition of a martingale to the value-process of a portfolio it represents a "fair game", i.e. the value of the portfolio on day j is its value on day $i \leq j$ [12].

A simple example of a martingale process would be to consider a coin-flipping game in which we start with €1 and each time we land on heads, we gain €1, while each time we land on tails, we lose €1. The probability for either outcomes after each coin flip is equal to 1/2 and independent of any previous coin flips. This means that, if we denote our value after n coin flips by X_n, we have

$$\begin{aligned}
\mathbb{E}[X_n | X_0, X_1, \ldots, X_{n-1}] &= \mathbb{E}[X_n | X_{n-1}] \\
&= \mathbb{E}[X_{n-1} + 1 | X_{n-1}] + \mathbb{E}\left[X_{n-1} - 1 \middle| X_{n-1} \right] \\
&= (X_{n-1} + 1)\mathcal{P}(\text{heads} | X_{n-1}) + (X_{n-1} - 1)\mathcal{P}(\text{tails} | X_{n-1}) \\
&= (X_{n-1} + 1)\frac{1}{2} + (X_{n-1} - 1)\frac{1}{2} \\
&= X_{n-1}.
\end{aligned}$$

We see that the conditional expectation for our wallet to be worth X_n after n coin flips is equal to its value after $n - 1$ flips and thus obeys the martingale property stated above.

Now, there is a *fundamental theorem of finance* or *fundamental theorem of asset pricing* [5, 12, 51] which states that in a complete arbitrage-free market, there exists a probability measure to which the evolution of the price of financial derivatives are a martingale. More generally, a complete arbitrage-free market is equivalent to the statement that for any given asset S, there exists a probability measure Q, such that ratio $\frac{S(t)}{\mathcal{N}(t)}$, for a numeraire $\mathcal{N}(t)$, are Q-martingales. Probability measure Q is called the *risk-neutral (martingale) measure* with respect to the numeraire $\mathcal{N}(t)$. This theorem implies that, for a derivative V on an assets S with maturity s, its price at time $t \leq s$ reads

$$\frac{V(t,s)}{\mathcal{N}(t)} = \mathbb{E}^Q \left[\frac{V(s,s)}{\mathcal{N}(s)} \middle| S(t) = S_t \right]. \tag{2.6}$$

This equation is the called the *risk-neutral pricing formula*. Popular choices for the numeraire $\mathcal{N}(t)$ are the DCF, $D(t,s)$ or a ZCB, the price of which will be denoted $P(t,s)$ [12]. We will introduce these in section 2.2.2.

For now, we can see that, for a cash bank account as numeraire, i.e. $\mathcal{N}(t) = B(0,t) = e^{\int_0^t \mathrm{d}t' r(t')}$, we find

$$V(t,s) = \mathbb{E}^Q \left[e^{-\int_t^s \mathrm{d}t' r(t')} V(s,s) \middle| S(t) = S_t \right], \tag{2.7}$$

where we place the cash bank account $B(0,t)$ at present time t inside the since its value is a known quantity at time t.

We can thus conclude that the assumption of a complete and arbitrage-free market is essential for the risk-neutral pricing of financial instruments. The combination of these two market assumptions imply the existence of a "risk-neutral (martingale) measure" under which the discounted price-process of a financial instrument is a martingale. The risk-neutral price is then defined as the expectation value of the discounted future price under this risk-neutral martingale measure.

2.1.4 Black-Scholes Equation

As already mentioned in the introduction, the Black-Scholes option pricing model is one of the most well-known (but also one of the most basic) models in mathematical finance. Because of this reason, the model is very well researched and there are at least a dozen different ways of deriving the Black-Scholes PDE [58]. We will only consider the original derivation by means of a simple *hedging* argument. The concept of hedging loosely translates to "eliminating every risk", hence we will construct a portfolio whose change is completely risk-free.

In the Black-Scholes model, we assume the risk-free interest rate to be constant and the underlying of a derivative, S, to follow a *geometric Brownian motion*:

$$\mathrm{d}S = \mu S \,\mathrm{d}t + \sigma S \,\mathrm{d}W(t) , \tag{2.8}$$

where μ is a drift-term, σ is the *volatility* and $W(t)$ represents a *Wiener process* (more on Wiener processes and Brownian motions in section 3.1.2). Moreover, under the risk-neutral measure, the drift term μ in eq. (2.8) is equal to the risk-free interest rate r (in the case of a dividend paying underlying, one has $\mu = r - d$, where d are the continuously payed dividends).
We now construct a portfolio A that consists of one call option[1] with price $C(t)$ at time t (we leave the option's characteristics implicit as they do not impact our derivation) and a loan of Δ amount of the underlying. The value of the portfolio at any time t (before the option's maturity) is given by

$$V^A(t) = C(t) - \Delta S(t).$$

We now wonder how the portfolio value changes over an infinitesimal time interval $[t, t + \mathrm{d}t]$:

$$\mathrm{d}V^A = \mathrm{d}C - \Delta \,\mathrm{d}S$$

[1] Note that instead of a call option, this could be any other derivative and the derivation would not change.

From Itô's Lemma[2] this becomes

$$\mathrm{d}V^A = \frac{\partial C}{\partial t}\,\mathrm{d}t + \frac{\partial C}{\partial S}\,\mathrm{d}S + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2}\,\mathrm{d}t - \Delta\,\mathrm{d}S\,.$$

If we now choose $\Delta = \frac{\partial C}{\partial S}$ we are left with

$$\mathrm{d}V^A = \left(\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2}\right)\mathrm{d}t\,. \tag{2.9}$$

This equation shows that the change in the portfolio's value is entirely deterministic as it does not depend on the stochastic Wiener process W. Hence it is risk-free. By no-arbitrage, the portfolio's value should thus grow at the same rate as the value V^A in cash under the risk-free rate, that is

$$\mathrm{d}V^A = rV^A\,\mathrm{d}t\,. \tag{2.10}$$

Combining eq. (2.9) and eq. (2.10) yields

$$\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + rS\frac{\partial C}{\partial S} - rC = 0. \tag{2.11}$$

This is the famous Black-Scholes equation. One can obtain the call option price by solving the PDE with boundary condition

$$C(t = s, S; K, s) = [S(s) - K]^+. \tag{2.12}$$

Black-Scholes as Heat-Diffusion

It is interesting to note that one can transform eq. (2.11) to a simple heat equation [21]. This can be achieved by defining $T \equiv s - t$ and $u \equiv \ln\left(\frac{S}{K}\right)$. The option value $C(t, S) \equiv C(t, S; K, s)$ can then be rewritten in function of T and u as $C(t, S) \equiv K\tilde{C}(T, u)$. If then also define

$$g(u, T) = \tilde{C}(T, u)e^{-au-bT}, \tag{2.13}$$

where

$$a \equiv -\frac{1}{\sigma^2}\left(r - \frac{\sigma^2}{2}\right)$$
$$b \equiv -\frac{1}{2\sigma}\left(r + \frac{\sigma}{2}\right)^2.$$

Plugging eq. (2.13) into eq. (2.11) indeed yields the famous heat equation in a straightforward manner

$$\frac{\partial g(u, T)}{\partial T} = \frac{\sigma^2}{2}\frac{\partial^2 g(u, T)}{\partial u^2}, \tag{2.14}$$

[2] As we will not include a discussion on Itô calculus in this thesis, we will simply mention Itô's Lemma here and refer to appropriate literature [31]: Let $g : \mathbb{R} \to \mathbb{R}$ be a twice continuously differentiable function and $X(t)$ a stochastic process on $[0, s]$ with

$$\mathrm{d}X(t) = \mu(X(t), t)\,\mathrm{d}t + \sigma(X(t), t)\,\mathrm{d}W(t)$$

for a Wiener process $W(t)$. Then the following holds

$$\mathrm{d}g(X) = g'(X(t))\,\mathrm{d}X(t) + \frac{1}{2}g''(X(t))\sigma(X(t), t)^2\,\mathrm{d}t\,.$$

Naturally, the solution of the Black-Scholes equation can be expressed in terms of the solution to the heat equation. If we consider the general boundary condition $g(u, T=0) = h(u)$, the solution to eq. (2.14) is given as

$$g(u,T) = \int_0^u G(u,u';T)g(u')\,\mathrm{d}u'\,,$$

where $G(u,u';T)$ is the *propagator*. An explicit solution for $g(u,T)$ thus allows us to transform back to a solution for $C(t,S)$ via

$$\begin{aligned} C(t,S) &= Kg(u,T)e^{au+bT} \\ &= K^{1-a}S^a g\left(\ln\left(\frac{S}{K}\right), s-t\right)e^{b(s-t)}. \end{aligned}$$

For a European call option, subject to the boundary condition given in eq. (2.12), the solution famously reads [21]

$$C(t,S;K,s) = S\Phi(d_+) - Ke^{-r(s-t)}\Phi(d_-), \tag{2.15}$$

where

$$d_\pm = \frac{\ln\left(\frac{S}{K}\right) + \left(r \pm \frac{\sigma^2}{2}\right)(s-t)}{\sigma\sqrt{s-t}} \tag{2.16}$$

and

$$\Phi(d) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d} e^{-\frac{x^2}{2}}\,\mathrm{d}x\,. \tag{2.17}$$

We displayed the price of a few European call option prices for different times to expiry $T = s-t$ in fig. 2.2

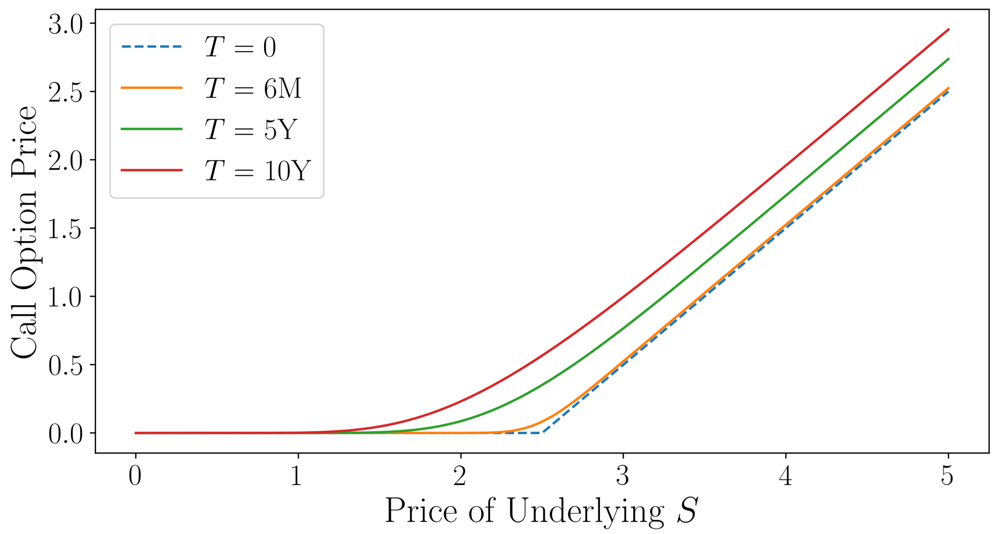


Figure 2.2: Graphical representation of European call option prices in function of the price of the underlying at maturity for different times to maturity T. The parameter values are $K = 2.5$, $r = 0.02$, $\sigma = 0.1$ year$^{\frac{1}{2}}$. The blue dashed line shows the option's payoff at maturity.

2.1.5 Feynman-Kac Formula

The Feynman-Kac Formula, named after physicist Richard Feynman and mathematician Mark Kac, reveals a connection between PDEs and a corresponding stochastic process [60]. It allows one to solve certain PDEs by simulating the random paths of a stochastic process and giving a stochastic representation of the solutions for these PDEs. The formula states that, for a PDE of the form

$$\frac{\partial\varphi(x,t)}{\partial t}+\frac{1}{2}v^2(x,t)\frac{\partial^2\varphi(x,t)}{\partial x^2}+\mu(x,t)\frac{\partial\varphi(x,t)}{\partial x}-V(x,t)\varphi(x,t)=0 \tag{2.18}$$

with boundary condition $\varphi(x,t=t_f)=g(x_f)$ and known functions $\mu(x,t), v(x,t), V(x,t), g(x)$, the solution $\varphi(x,t)$ can be written as an expectation value

$$\varphi(x,t)=\mathbb{E}\left[e^{-\int_{t_0}^{t_f}V(X,t')\mathrm{d}t'}g(X)\,\middle|\, X=x\right], \tag{2.19}$$

where X is a stochastic process

$$\mathrm{d}X=\mu(x,t)\,\mathrm{d}t+v(x,t)\,\mathrm{d}W(t) \tag{2.20}$$

for a Wiener process $W(t)$ (under the same measure as the expectation value) and initial conditions $X(0)=x$.

When taking, for example, $\varphi=C$, the price of a European call option on an underlying $x=S$ and comparing eq. (2.11) with eq. (2.18), we can identify

$$\begin{aligned}\mu(S,t)&=rS\\ v(S,t)&=\sigma S\\ V(S,t)&=r\\ g(S)&=[S(s)-K]^+,\end{aligned}$$

where we set $t_f=s$, the option's maturity. We also know that, under the risk-neutral measure Q, the price-process of the underlying follows the geometric Brownian motion given by eq. (2.8) with $\mu=r$ (not to be confused with the function $\mu(t,x)$ in eq. (2.18)). By the Feynman-Kac formula eq. (2.19), we then find that

$$\boxed{C(t,S;K,s)=\mathbb{E}^Q\left[e^{r(s-t_0)}(S_s-K)^+\,\middle|\, S_t=S(t)\right]}, \tag{2.21}$$

where we introduced the stochastic price-process S_t, which obeys eq. (2.8) for a Wiener process under the risk-neutral measure Q. This result is directly related to the concept of risk-neutral valuation from section 2.1.3. This can be seen by comparing eq. (2.7) in the case where the asset is a European call with eq. (2.21). It allows us to price derivatives by considering *all* possible random paths the underlying can follow (under the risk-neutral measure), between begin- and endpoint $S(0)$ and $S(s)$ respectively, and taking the expectation value of the discounted payoff. This enables the use of Monte Carlo simulations to approximate derivative prices [60] and also naturally leads to the concept of *path integrals* and how these might be usefull in constructing a formalism of derivative valuation. We will see more on path integrals in chapter 3.

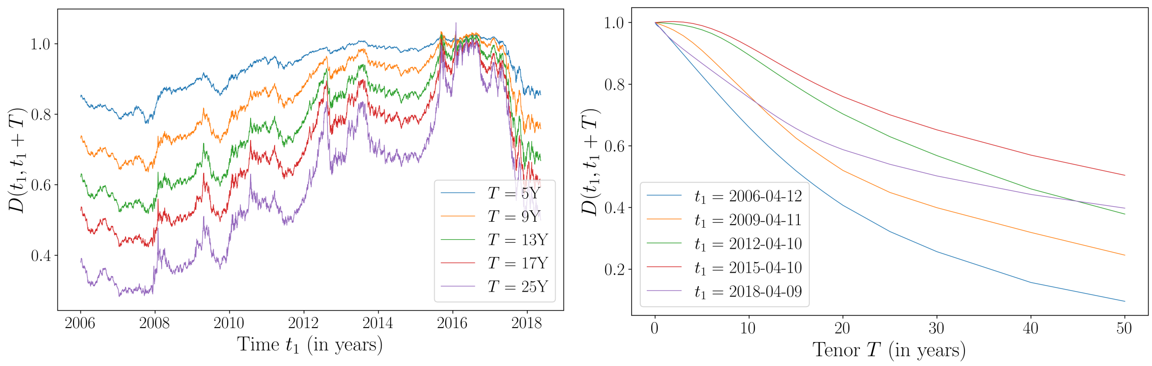
2.2 Zero Coupon Bonds & Interest Rates

2.2.1 Discount Factors

As already discussed briefly in the introduction, DCFs are mathematical instruments used for discounting future values of financial instruments to their current value. We will always assume continuous discounting such that the DCF for the time interval $[t_1, t_2]$ reads

$$D(t_1, t_2) = e^{-\int_{t_1}^{t_2} r(t)\mathrm{d}t} \tag{2.22}$$

A DCF can also be seen as a representation of the "cost of borrowing", i.e. for a hypothetical loan of €1 at time t_1 over a time period $t_2 - t_1$ you would have pay back a total amount of $B(t_1, t_2) = D(t_1, t_2)^{-1}$. We will often refer to the "end time" t_2 of the DCF as its maturity and the time period $t_2 - t_1$ is called the tenor. Note that the tenor of a DCF does not represent an instance in time but rather a duration. DCFs can have equal tenors but totally different values. When the interest rate r is constant, eq. (2.22) reduces to $D(t_1, t_2) = e^{-r(t_2 - t_1)}$, however, this is almost never the case and interest rates follow the interest rate market that fluctuates stochastically [8]. This stochastic behavior can clearly be seen from historical market data by keeping $T \equiv t_2 - t_1$ fixed and letting t_1 vary. The evolution of DCFs through time t_1 for 5 curves with tenors ranging from $T = 5Y$ to $T = 25Y$ are shown in fig. 2.3a. On the other hand, if we



(a) (b)

Figure 2.3: Plots of historical DCFs as a function of time t_1 keeping $T = t_2 - t_1$ fixed on each individual path (a) and as a function of $T = t_2 - t_1$ keeping t_1 fixed on each individual curve (b).

keep t_1 fixed and let $T = t_2 - t_1$ vary, we get what was previously called a discount curve. Such discount curves can be seen as a function of their tenor $T = t_2 - t_1$ in fig. 2.3b for various dates[3] t_1. Each of the discount curves displayed corresponds to a different point in time, i.e. a different date t_1. The curve thus evolves stochastically through time. Note that the DCFs have in fact been larger than 1 for a significant period of time. This is because interest rates where indeed negative for a while.

We will consider all the different DCFs as observables arising from the dynamics of an underlying fundamental rate curve. Each point on this rate curve will turn out to follow certain stochastic dynamics (see section 2.3). From what we see in fig. 2.3a, it safe to assume that these observables are (highly[4]) correlated over different tenors (at least in the short term) [8].
It is clear that, on the contrary to stock prices, DCFs are functions of two variables. Hence, they comprise a 2-dimensional surface laying in a 3-dimensional space. A graphical representation of this surface can be seen in fig. 2.4. Note how it can clearly be seen in figs. 2.3b and 2.4 that, for $T = s - t = 0$, all DCFs have value equal to 1.

Figures 2.3a and 2.3b are thus 2-dimensional sections of this surface. It is only natural to explore the idea of extending the fundamental rate curve to a stochastic "rate field" that underlies the DCF dynamics.

[3]The dates for this and other figures is 12th of April in different years. This exact day can differ due to weekend days.
[4]But not exactly!

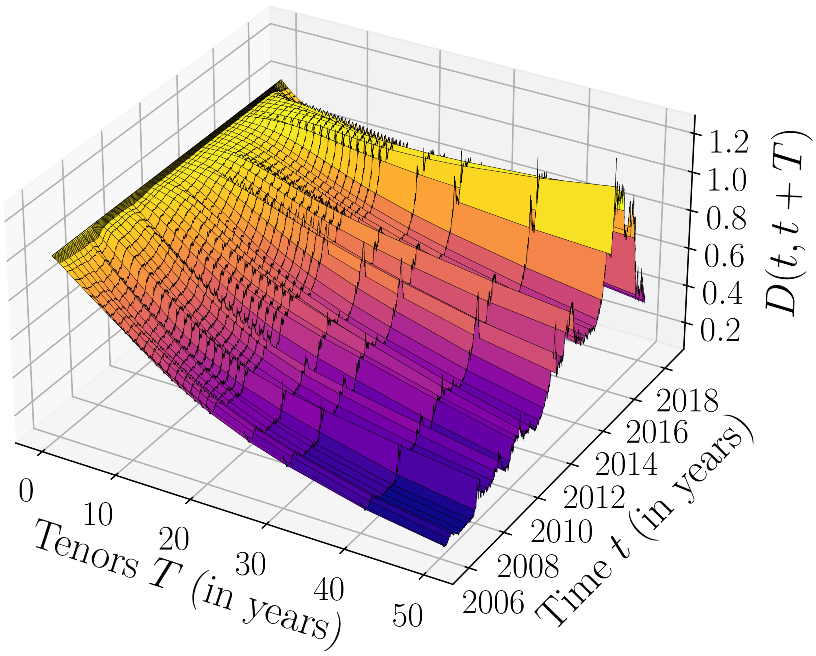


Figure 2.4: Graphical representation of the historical "discount surface".

2.2.2 Zero Coupon Bonds

The DCF is very closely related to a theoretical financial product named the ZCB. It is a theoretical instrument in the sense that it is not actually traded in the market but acts as a very useful tool in developing and describing models or calculations. It is called a *zero coupon* bond since there are no in-between payments over the course of the bond's lifetime but only one payment at maturity. The reason a ZCB is so similar to its corresponding DCF is that the value of a ZCB at its maturity time s is equal to €1. By the monotonicity theorem, the value of the ZCB at time $t \leq s$ is equal to the value of €1 discounted from time s to t [27]. The price of a ZCB at time t with maturity $s \geq t$ will be denoted $P(t, s)$. Furthermore, to simplify our lives we will make the following assumptions concerning ZCBs:

- ZCBs are traded in a frictionless nature for all maturities $s > 0$
- $\forall s > 0 : P(s, s) = 1$
- $P(t, s)$ is differentiable in s.

This final condition ensures that the curve $s \mapsto P(t, s)$, or *term-structure of ZCBs*, is smooth.

Now, since the payout of a ZCB is equal to 1, its risk-neutral price at time $t \leq s$ reads

$$P(t, s) = \mathbb{E}^Q \left[e^{-\int_t^s r(t')\mathrm{d}t'} \middle| \mathcal{F}_t \right], \tag{2.23}$$

where $\mathcal{F}_t$ is called a *filtration* and can loosely be interpreted as all the available (relevant) market data up to time t [8]. This is one of two possible definitions for the ZCB we will consider. Here, the ZCB is defined in terms of the spot rate $r(t')$. We will see in section 2.2.4 that it can also be defined in relation to the instantaneous forward interest rate. Note how this reduces to a regular DCF $D(t, s)$ in the case of deterministic interest rates. Moreover, at any certain *present* time t_0, the ZCB price $P(t_0, s)$ is a known deterministic quantity[5] (since it can be determined from the market), while the value of the DCF $D(t_0, s)$ is stochastic as it depends on the future evolution of $r(t)$ between t_0 and s [15].

Due to its simplicity and widely spread use in mathematical finance and interest rate models, we will often fall back to the ZCB and a European call option on a ZCB to check the analytical tractability of a model or to compare different approaches to certain derivations.

2.2.3 Numeraires

We mentioned earlier that one could use different quantities aside from DCFs for "discounting" future values and that such quantities are named "numeraires". In fact, a numeraire is defined as *any* positive non-dividend-paying asset [15]. This means that, given a numeraire $\mathcal{N}$ and a probability measure $Q^{\mathcal{N}}$, such that, for any asset V_1, the ratio $\frac{V_1}{\mathcal{N}}$ is a $Q^{\mathcal{N}}$-martingale (such as in eq. (2.6)), then, for an arbitrary numeraire $\mathcal{U}$, there exists another probability measure $Q^{\mathcal{U}}$ such that, for any asset V_2, the ratio $\frac{V_2}{\mathcal{U}}$ is a $Q^{\mathcal{U}}$-martingale, i.e.

$$\frac{V_2(t)}{\mathcal{U}(t)} = \mathbb{E}^{Q^{\mathcal{U}}}\left[\frac{V_2(s)}{\mathcal{U}(s)}\middle|\mathcal{F}_t\right] \quad \text{for } 0 \leq t \leq s. \tag{2.24}$$

This is quite the mouthful but it can loosely translated to: "For any numeraire, there exists a corresponding probability measure such that the risk-neutral pricing formula hold". It is important to keep in mind that all the market assumptions explained in section 2.1.2 are necessary for this to hold.

The two most commonly used numeraires are the cash bank account (or, equivalently the DCF) and the ZCB. In the case of the former, the risk-neutral price for a derivative V reads

$$\begin{aligned} V(t_0) &= B(t_0, t_0)\mathbb{E}\left[\frac{V(s)}{B(t_0, s)}\middle|\mathcal{F}_{t_0}\right] \\ &= \mathbb{E}\left[e^{-\int_{t_0}^{s} r(t)\mathrm{d}t}V(s)\middle|\mathcal{F}_{t_0}\right] \\ &= \mathbb{E}\left[D(t_0, s)V(s)\middle|\mathcal{F}_{t_0}\right], \end{aligned} \tag{2.25}$$

where we used the fact that $B(t_0, t_0) = 1$ and left the corresponding probability measure implicit. In the case of the latter, the risk-neutral pricing formula for a derivative with maturity $s \geq t_0$ becomes

$$\begin{aligned} V(t_0) &= P(t_0, s)\mathbb{E}\left[\frac{V(s)}{P(s, s)}\middle|\mathcal{F}_{t_0}\right] \\ &= P(t_0, s)\mathbb{E}\left[V(s)\middle|\mathcal{F}_{t_0}\right], \end{aligned} \tag{2.26}$$

since $P(s, s) = 1$ for all s. Note the that big difference between eq. (2.25) and eq. (2.26) is that in the case of the ZCB as numeraire, the discounting is done *outside* the expectation value. The use of the "correct" numeraire can thus greatly simplify the valuation procedure [34].

[5] This does however not eliminate the fact that $t \mapsto P(t, s)$ is a stochastic process. The price of the ZCB is only known with certainty at the *present* and *maturity*.

2.2.4 Interest Rates

We have already mentioned a few different types of interest rates such as spot rate, forward rate and instantaneous forward rate. All these rates specify the "cost of borrowing" in a slightly alternate fashion. In this section we will briefly discuss the relevant types of interest rates, starting with one that was not referred to before: the *yield to maturity*.

Yield to Maturity

The yield to maturity, $R(t,T)$, indicates the *internal* rate of return at time t on a ZCB of maturity $s = t + T$ and is given as

$$R(t,T) = -\frac{1}{T}\ln P(t,s). \tag{2.27}$$

It might be of note that definitions vary among literature [18, 56]. We will assume the definition as given by Vašíček in [56]. $R(t,T)$ is named the "internal" rate of return as it is the value that must be plugged into eq. (2.23) to obtain the traded ZCB price. The yield to maturity represents the average rate over the time period $[t,s]$ at which an investment at time t in the ZCB $P(t,s)$ will accumulate [18].

Forward Rate

The forward rate shows a rate of interest between future times s_1 and s_2, fixed at an earlier time $t \le s_1 \le s_2$ and is defined as [18]:

$$F(t;s_1,s_2) = \frac{1}{s_2 - s_1}\ln\frac{P(t,s_1)}{P(t,s_2)}. \tag{2.28}$$

It is easy to see that, if $s_1 = t$ the forward rate $F(t;t,s_2) = R(t,s_2 - t)$. The forward rate also allows us to define the most important type of interest rate for the transition to a field theory formalism: the instantaneous forward interest rate

Instantaneous Forward Rate

The instantaneous forward rate gives the rate of interest on a loan of infinitesimal duration over the interval $[s, s + \mathrm{d}s]$, as seen from time $t \le s$ and is defined as

$$f(t,s) = \lim_{s_2 \to s} F(t;s,s_2) = -\frac{\partial}{\partial s}\ln P(t,s) \tag{2.29}$$

or, equivalently

$$P(t,s) = e^{-\int_t^s f(t,x)\mathrm{d}x}. \tag{2.30}$$

This definition relates the instantaneous forward rates to the ZCB price. Note that $t \mapsto P(t,s)$ is still a stochastic process for every value of s. The concept of an instantaneous forward interest rate is entirely artificial but it does however function very well as a tool to make models and price calculations a lot more convenient [18]. Just as the DCFs, the instantaneous forward rate is also a function of two variables. It can thus be represented as a 2-dimensional surface and will therefore play a crucial role in the path integral- and field theory description of interest rates. For convenience, we will often use the term "forward (interest) rate" when actually talking about the instantaneous forward interest rate. Hopefully this is clear from context as not to confuse it with the actual forward rate $F(t;s_1,s_2)$.

Spot or Risk-Free Rate

The spot rate or risk-free rate is the interest rate used to calculate DCFs. What we actually mean by "risk-free rate" is an *instantaneous* risk-free rate

$$\begin{aligned} r(t) &= \lim_{T \to 0} R(t,T) = R(t,0) \\ &= \lim_{s \to t} f(t,s) = f(t,t). \end{aligned}$$

It describes the (continuously compounded) rate of interest on a so called risk-free bank account. It is also sometimes called the *short rate* and it is this rate that is modeled in the short-rate models such as Vašíček and Hull-White [35, 56].

2.3 Interest Rate Models

Generally, interest rate models can be categorized in two families. On the one hand, there are the *short-rate models* that aim to model the risk-free rate but only yield reliable results over a short time period. These models can, in principle, model an initial forward rate curve but cannot produce viable evolutions of the forward rate curve as they do not maintain the assumption of no-arbitrage [5]. Also, often they do not have sufficient flexibility to be accurately calibrated to such an initial forward rate curve [27]. However, most short-rate models are diffusion-like and relatively simple which makes them analytically tractable [27]. On the other hand, there are the *term-structure models*. These attempt to model the entire term structure of (forward) interest rates and can give decent results over longer time periods. They take an initial forward rate curve $f(t_0, s)$, which is to be determined from the market, as opposed to trying to derive this initial curve from a short-rate model. This approach enables us to incorporate (almost) all market information on the forward rates that is available from this initial curve. Term-structure models are able to yield an arbitrage-free evolution of the plugged-in initial forward rate curve[6] [5].

In this chapter, we will only discuss the two most relevant interest rate models for this thesis: the Vasicek model and the HJM framework, but there are many more commonly used (short-) interest rate models such as Black-Derman-Toy, Black-Karasinski, Cox-Ingersol-Ross, Ho-Lee, Hull-White,... [9, 10, 23, 33, 35].

2.3.1 Vasicek's Model

The Vasicek model is a short-rate model and thus aims to model the spot interest rate. It does so by arguing the validity of some assumptions on the behavior of the spot rate. We will, under these assumptions, derive a PDE for the ZCB price and, in a specific case, its closed-form pricing formula. This will be done following the original paper by Vašíček [56]. The first assumption states that $r(t)$ follows a stochastic process and is a continuous function of time t. Furthermore, it is assumed that the future values of the spot rate, given its current value, are independent of the process in the past, i.e. $r(t)$ follows a *Markov process* (more on Markov processes in section 3.1.1). A second assumption states that the price of a ZCB depends only on the different values of the short rate over the lifetime of the bond. Finally, all the market assumptions explained in section 2.1.2 are in effect. To summarize,

[6] Note that this does not imply that this evolution is a correct one. It only implies that the evolution that is produced is effectively free of arbitrage.

A.1 $r(t)$ follows a diffusion process in the sense that

$$\mathrm{d}r(t) = g(r,t)\,\mathrm{d}t + \rho(r,t)\,\mathrm{d}W(t) \tag{2.31}$$

for a Wiener process $W(t)$. The functions $g(r,t), \rho^2(r,t)$ are the instantaneous drift and variance respectively.

A.2 $P(t,s)$ is determined by solely by the evolution of the short rate process over the lifetime of the ZCB; $\{r(\tau)|t \le \tau \le s\}$.

A.3 The market is efficient, frictionless and arbitrage free.

By A.1 and A.2, one can conclude that, at every time t, $P(t,s)$ is a function of of the value of the short rate at that time, $r(t)$, since, due to the Markovian property, the evolution of the spot rate for times beyond time t only depends on the present rate value, $r(t)$. Most of the time, we will leave this dependence implicit and just write $P(t,s)$.

Using now A.1 and $P(t,s) = P(t,s,r(t))$, one can show that, by Itô's differentiation,

$$\frac{\mathrm{d}P}{P} = \mu(t,s)\,\mathrm{d}t - \nu(t,s)\,\mathrm{d}W(t)\,, \tag{2.32}$$

with

$$\mu(t,s) = \frac{1}{P(t,s,r)}\left(\frac{\partial}{\partial t} + g\frac{\partial}{\partial r} + \frac{\rho^2}{2}\frac{\partial^2}{\partial r^2}\right)P(t,s,r) \tag{2.33}$$

$$\nu(t,s) = -\frac{\rho}{P(t,s,r)}\frac{\partial P(t,s,r)}{\partial r}. \tag{2.34}$$

Note that, since the bond value P is driven by a single stochastic process $W(t)$, bonds with different maturities s are correlated perfectly. This is an undesirable feature as we will see in section 6.2 that this is not actually the case for real market data and bonds for different maturities do have a non-trivial correlation.

We will now derive a partial differential equation for the ZCB price P by means of portfolio construction. Consider a portfolio consisting of a sold amount N_1 of ZCBs with maturity s_1 and a bought amount N_2 of ZCBs with maturity s_2. The total portfolio value is given by $N = N_2 - N_1$ and is governed by eq. (2.32).

Take these amounts $N_{1,2}$ as follows

$$N_i = \frac{N\nu(t,s_j)}{\nu(t,s_i) - \nu(t,s_j)} \quad \text{for } i,j \in \{1,2\} \text{ and } i \ne j. \tag{2.35}$$

From eq. (2.32) and Itô's Lemma, it follows that

$$\mathrm{d}N = [N_2\mu(t,s_2) - N_1\mu(t,s_2)]\,\mathrm{d}t + [N_2\nu(t,s_2) - N_1\nu(t,s_1)]\,\mathrm{d}W(t)\,. \tag{2.36}$$

Plugging in the specified amounts from eq. (2.35) yields

$$\mathrm{d}N = N\frac{\mu(t,s_2)\nu(t,s_1) - \mu(t,s_1)\nu(t,s_2)}{\nu(t,s_1) - \nu(t,s_2)}\,\mathrm{d}t\,. \tag{2.37}$$

Such a portfolio would effectively be riskless as the stochastic element $\mathrm{d}W$ is absent and by the no-arbitrage assumption in A.3 one finds for arbitrary s_1, s_2

$$\frac{\mu(t,s_2)\nu(t,s_1) - \mu(t,s_1)\nu(t,s_2)}{\nu(t,s_1) - \nu(t,s_2)} = r(t),$$

or, equivalently

$$\frac{\mu(t,s_1)-r(t)}{\nu(t,s_1)}=\frac{\mu(t,s_2)-r(t)}{\nu(t,s_2)}.$$

This ratio is thus independent from the choices for $s_{1,2}$. Introduce the function $\lambda(t,r)$ as

$$\lambda(t,r)=\frac{\mu(t,s,r)-r}{\nu(t,s,r)}\quad \textit{for } s\geq t, \tag{2.38}$$

where $r=r(t)$ denotes the current spot rate value. The quantity $\lambda(t,r)$ is called the *market price of risk*. It specifies the increase in expected rate of return on a ZCB per unit of risk. Using eq. (2.38) and substituting μ and ν from eqs. (2.33) and (2.34) yields

$$\frac{\partial P}{\partial t}+(g+\rho\lambda)\frac{\partial P}{\partial r}+\frac{1}{2}\rho^2\frac{\partial^2 P}{\partial r^2}-rP=0 \tag{2.39}$$

subject to boundary condition $P(s,s,r)=1$.

A Specific Case

Vašíček proposes the specific case for which $\lambda(t,r)=\lambda$ is a constant and $r(t)$ follows an *Ornstein-Uhlenbeck* (O-U) process (more on O-U processes in section 3.1.5):

$$\mathrm{d}r(t)=k[\theta-r(t)]\,\mathrm{d}t+\rho\,\mathrm{d}W(t)\,, \tag{2.40}$$

where θ is the mean-reversion level, k is the velocity of the mean-reversion and ρ is the "instantaneous volatility" of the process, all of which are taken constant. Important to note is that Vašíček never claimed that the O-U process is the best description of the short rate behavior and it only served as an educative example.

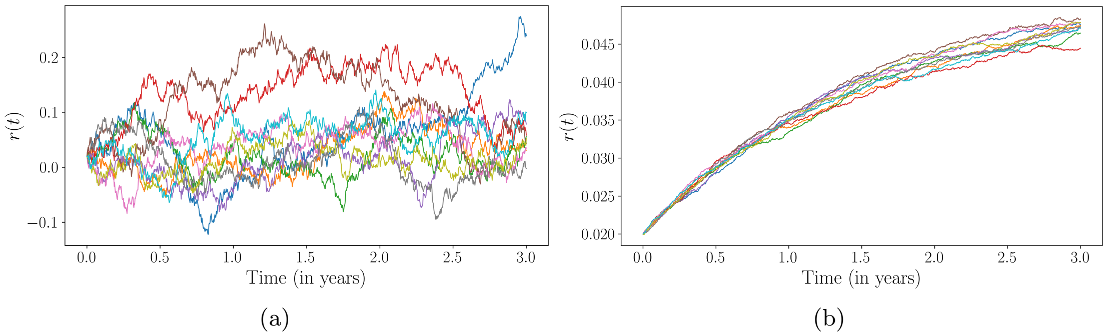


Figure 2.5: Simulation of 10 spot rate processes following the Vašíček dynamics as in eq. (2.40) with parameter values $k=0.7/\text{year},\ \theta=0.05,\ \rho=0.1\ \text{year}^{\frac{1}{2}}$ (a) and $k=0.7/\text{year},\ \theta=0.05,\ \rho=0.001\ \text{year}^{\frac{1}{2}}$ (b).

In figs. 2.5 and 2.5b, we displayed 10 simulations of a high and low volatility Vasicek process following the dynamics of eq. (2.40). We can clearly see that the processes to tend to the mean-reversion level $\theta=0.05$. Of course, this mean-reversion is more apparent in the low volatility regime of fig. 2.5b than in the high volatility regime of fig. 2.5a. Furthermore, it is important to note that, as can be seen in fig. 2.5a, interest rates can become (highly) negative under Vašíček 's model. Unfortunately, this is no longer a purely theoretical feature (to some extent) and some countries actually already implemented interest rates below zero [43]. We will also encounter negative interest rates in the market data discussed in section 6.2.

It will later be shown that the conditions expectation value and variance for r following an O-U process are given respectively as

$$\mathbb{E}[r(s)|r(t)] = \theta\left[1 - e^{-k(s-t)}\right] + r(t)e^{-k(s-t)} \tag{2.41}$$

$$\mathrm{Var}[r(s)|r(t)] = \frac{\rho^2}{2k}\left[1 - e^{-k(s-t)}\right] \tag{2.42}$$

for $s \geq t$. Under the assumption that the sport rate follows an O-U process, a solution to eq. (2.39) has the form

$$P(t,s,r) = \exp\left\{\left(\frac{1 - e^{-k(s-t)}}{k}\right)[R - r] - (s-t)R - \frac{\rho^2}{4k}\left(\frac{1 - e^{-k(s-t)}}{k}\right)^2\right\} \tag{2.43}$$

with $s \geq t$ and with

$$R = \theta - \frac{\rho\lambda}{k} - \frac{\rho^2}{2k^2}. \tag{2.44}$$

Plugging eq. (2.43) into eqs. (2.33) and (2.34), one finds

$$\mu(t,s) = r(t) + \frac{\rho\lambda}{k^2}\left(1 - e^{-k(s-t)}\right)$$
$$\nu(t,s) = \frac{\rho}{k}.$$

Taking $\lambda = 0$, the bond price reduces to the well-documented result

$$\boxed{P(t,s) = A(t,s)e^{-\tilde{B}(t,s)r(t)},} \tag{2.45}$$

with

$$A(t,s) = \exp\left\{\left(\theta - \frac{\rho^2}{2k}\right)[\tilde{B}(t,s) - (s-t)] - \frac{\rho^2}{4k}\tilde{B}(t,s)^2\right\} \tag{2.46}$$

$$\tilde{B}(t,s) = \frac{1 - e^{-k(s-t)}}{k}, \tag{2.47}$$

where $\tilde{B}(t,s)$ is just a definition to simplify notation and has no direct relation with the cash bank account $B(t,s)$. We plotted the price of a ZCB according to Vašíček 's formula eq. (2.45) in function of its time to maturity $T = s - t$ as seen from time $t = 0$. The result can be seen in fig. 2.6.

2.3.2 Heath-Jarrow-Morton Framework

As stated earlier, short-rate models still have quite some inadequacies, such as the inability of producing an arbitrage-free evolution of the term-structure. The approach as suggested by Heath, Jarrow and Morton tries to work around this issue by introducing a framework that models the (instantaneous) forward rate curve as a whole instead of modeling the short-rate and deriving the forward rate from that. By taking the instantaneous forward rates as the fundamental underlying quantities, they managed to derive a way of describing an arbitrage-free stochastic evolution of the entire forward rate curve [15]. An exciting consequence of this approach is that the framework also encompasses (almost) all spot rate models [59].

Just as in the previous section, we will derive some of the most relevant results from the original paper by Heath, Jarrow and Morton [32] but we will not go into much depth as far as the

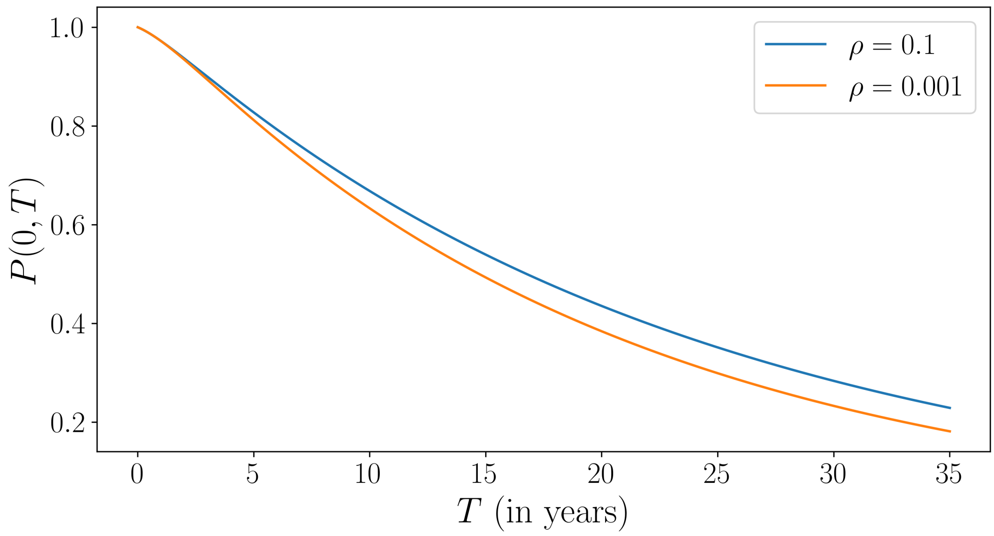


Figure 2.6: Price of a ZCB $P(t=0,s,r(0))$ plotted in function of its time to maturity time T with parameter values $r_0 = 0.02, k = 0.7/\text{year},\ \theta = 0.05,\ \rho = 0.1\ \text{year}^{\frac{1}{2}}$ and $\rho = 0.001\ \text{year}^{\frac{1}{2}}$.

mathematical complexities are concerned. We gladly refer to the following rigorous literature for the meticulous reader: [15, 27, 37, 48].

The instantaneous forward interest rate, $f(t,s)$, is assumed to follow the dynamics given by

$$\mathrm{d}f(t,s) = \alpha(t,s)\,\mathrm{d}t + \sigma(t,s)\,\mathrm{d}W(t)\,, \tag{2.48}$$

with a general solution of the form

$$f(t,s) = f(0,s) + \int_0^t \mathrm{d}t'\,\alpha(t',s)\,\mathrm{d}t' + \int_0^t \mathrm{d}t'\,\sigma(t',s)\,\mathrm{d}W(t')\,, \tag{2.49}$$

where $t \le s$ and $W(t)$ is a Wiener process. In the above stochastic process, the random fluctuations of the entire forward rate curve, starting from the initial curve $s \mapsto f(0,s)$, are governed by the (single) Wiener process $W(t)$. This implies that all forward interest rates are perfectly correlated because they all depend on the same underlying stochastic factor [59], i.e. every forward rate $f(t,s)$ has the same stochastic "shock" at every time t, independent of the future time s. We call the term $\alpha(t,s)$ the *drift velocity* and $\sigma(t,s)$ the *deterministic*[7] *volatility function* or *volatility structure*. Models with a deterministic volatility structure are called *Gaussian models* and we will restrict ourselves to this case [42]. One can include multiple independent Brownian motions to drive the fluctuations by replacing the second term on the *right-hand-side* (RHS) of eq. (2.48) by a sum

$$\mathrm{d}f(t,s) = \alpha(t,s)\,\mathrm{d}t + \sum_{i=1}^{K} \sigma_i(t,s)\,\mathrm{d}W_i(t)\,. \tag{2.50}$$

This is referred to as a K-factor model. A K-factor model would provide a work-around for these (undesirable) trivial correlations in the one-factor model. However, we will explain in chapter 5 that even in the limit of $K \to \infty$, there are still certain shortcomings. We will restrict ourselves to one-factor models, i.e. $K = 1$. We will later extend the HJM model to a model with independent fluctuations for every future time s as well as for every time t.

Note that, by comparing eqs. (2.40) and (2.48), one can translate the Vasicek model to the HJM framework by taking $\sigma(t,s) = \rho e^{-k(s-t)}$ [8].

[7] Volatility can also be modeled to be stochastic but this will not be discussed here.

No Arbitrage Condition

We will now show that, following [59], under the no-arbitrage assumption, the drift function $\alpha(t,s)$ is completely fixed by the volatility $\sigma(t,s)$. Assume the ZCB prices $P(t,s)$ still obey eq. (2.39) and note that, for $t=s$, one has

$$0 = \mu(t,t)\,\mathrm{d}t + \nu(t,t)\,\mathrm{d}W(t)\,. \tag{2.51}$$

As this is the case for all values of t, this process is deterministic and we require $\nu(t,t)=0$ for all t.

Now differentiate eq. (2.29) with respect to t and substituting in eq. (2.48). Following Itô calculus, one finds

$$\mathrm{d}f(t,s) = \frac{\partial}{\partial s}\left[\frac{1}{2}\nu(t,s)^2 - \mu(t,s)\right]\mathrm{d}t - \frac{\partial}{\partial s}\nu(t,s)\,\mathrm{d}W(t)\,. \tag{2.52}$$

Reminding ourselves that spot-rate $r(t)=f(t,t)$ and that, in the absence of arbitrage and under the risk-neutral measure, $\mu(t,s)=r(t)$, we can compare eq. (2.39) and (2.48) to find

$$\sigma(t,s) = -\frac{\partial}{\partial s}\nu(t,s) \tag{2.53}$$

and

$$\alpha(t,s) = \frac{\partial}{\partial s}\left[\frac{1}{2}\nu(t,s)^2 - r(t)\right] = \nu(t,s)\frac{\partial}{\partial s}\nu(t,s). \tag{2.54}$$

Combining eqs. (2.53) and (2.54) with the fact that $\nu(t,t)=0$ for all values of t, we can conclude that

$$\boxed{\alpha(t,s) = \sigma(t,s)\int_t^s \sigma(t,x)\,\mathrm{d}x\,.} \tag{2.55}$$

This relationship between the risk-neutral drift and the volatility results from the assumption of no arbitrage.

As a first example, consider now the case for which volatility is a positive constant, i.e. $\sigma(t,s) = \sigma > 0$. We then immediately find from eq. (2.55) that

$$\alpha(t,s) = \sigma^2(s-t), \tag{2.56}$$

for all $t\in[0,s]$.

As a second example, we take the volatility to be the Vasicek case [8]:

$$\sigma(t,s) = \rho e^{-k(s-t)}, \tag{2.57}$$

for positive constants ρ and k. The drift function is then found to be

$$\begin{aligned}\alpha(t,s) &= \rho e^{-k(s-t)}\int_t^s \rho e^{-k(x-t)}\,\mathrm{d}x \\ &= \frac{\rho^2}{k}e^{-k(s-t)}\left[1 - e^{-k(s-t)}\right].\end{aligned}$$

ZCB Option

We mentioned earlier that HJM models with a deterministic volatility structure are referred to as Gaussian models. We also said that we would restrict ourselves to these types of models and this is not without reason. The big advantage of such models is that they allow for a closed form solution for the option price on a ZCB. In the case of a deterministic volatility structure $\sigma(t,s)$ both the forward- and the spot rate are normally distributed under the risk-neutral measure. The ZCB option price is then log-normally distributed (also under the risk-neutral measure) and this leads to a “Black-Scholes”-like pricing formula [42].

We will derive the pricing formula for a European call option on a ZCB $P(t,s)$ with maturity s. Consider the price at time t_0 for call option with strike K and expiry at time[8] t_* such that $0 \leq t_0 \leq t_* \leq s$ and denote its price at time t as $\mathrm{ZCBC}(t;K,t_*,s)$. At expiry t_*, we have

$$\mathrm{ZCBC}(t_*;K,t_*,s) = [P(t_*,s) - K]^+. \tag{2.58}$$

Risk-neutral pricing then implies

$$\begin{aligned}\mathrm{ZCBC}(t_0;K,t_*,s) &= \mathbb{E}\left[e^{-\int_{t_0}^{t_*} r(t)\mathrm{d}t}\mathrm{ZCBC}(t_*;K,t_*,s)\middle| \mathcal{F}_{t_0}\right], \\ &= \mathbb{E}\left[D(t_0,t_*)\ \mathrm{ZCBC}(t_*;K,t_*,s)\middle| \mathcal{F}_{t_0}\right].\end{aligned} \tag{2.59}$$

We saw earlier in eq. (2.32) that $P(t,s)$ is log-normally distributed and thus we can use log-normal models to calculate the expectation value in eq. (2.59). To do this we first consider a more general approach discussed in [8]. Suppose an asset S and DCF $D(t_0,s)$ are jointly log-normally distributed under the risk-neutral martingale measure. We use the following notation for the variances

$$\begin{aligned}\mathrm{Var}[\ln S] &= \sigma_1^2 s \\ \mathrm{Var}[\ln D(t_0,s)] &= \sigma_2^2 s\end{aligned}$$

and their correlation

$$\mathcal{C}[\ln S, \ln D(t_0,s)] = \rho.$$

The price at time $t = t_0$ for purchasing the asset at time $s \geq t_0$ is given as

$$F = \frac{\mathbb{E}[D(t_0,s)S(s)]}{\mathbb{E}[D(t_0,s)]}. \tag{2.60}$$

Since both $S(s)$ and $D(t_0,s)$ are log-normally distributed, these can be written as

$$S(s) = \mathbb{E}[S(s)]e^{\alpha_1 W_1 - \alpha_1^2/2} \qquad \text{with } \alpha_1^2 \equiv \sigma_1^2 s, \tag{2.61}$$

$$D(t_0,s) = \mathbb{E}[D(t_0,s)]e^{\alpha_2(\rho W_2 + \overline{\rho} W_1) - \alpha_2^2/2} \qquad \text{with } \alpha_2^2 \equiv \sigma_2^2 s, \tag{2.62}$$

where $\overline{\rho} = \sqrt{1-\rho^2}$, $W_1, W_2 \sim N(0,1)$ and W_1 is independent of W_2. This means

$$\mathbb{E}[D(t_0,s)S(s)] = \mathbb{E}[S(s)]\mathbb{E}[D(t_0,s)]\exp\left[\frac{1}{2}(\alpha_1 + \rho\alpha_2)^2 + \frac{1}{2}\alpha_2^2\overline{\rho}^2 - \frac{1}{2}\alpha_1^2 - \frac{1}{2}\alpha_2^2\right] \tag{2.63}$$

[8] When talking about options on bonds, we call the maturity of the option, i.e. the time at which the option is exercised, the *expiry* and the time at which the bond matures the *maturity*.

$$= \mathbb{E}[S(s)]\mathbb{E}[D(t_0, s)]e^{\rho\alpha_1\alpha_2}, \tag{2.64}$$

such that, from eq. (2.60)

$$F = \mathbb{E}[S(s)]e^{\rho\sigma_1\sigma_2 s} \tag{2.65}$$

and

$$S(s) = Fe^{\alpha_1 W_2 - \alpha_1^2/2 - \rho\alpha_1\alpha_2}. \tag{2.66}$$

We now know the price at time $t = t_0$ of a call on the asset S with strike K and expiry s reads

$$\begin{aligned} C(t_0, S; K, s) &= \mathbb{E}\left[D(t_0, s)\ [S(s) - K]^+ \middle| S(t_0) = S_0\right] \\ &= \mathbb{E}[D(t_0, s)]\mathbb{E}\left[e^{\rho W_2 - \rho^2\alpha_2^2/2}\ [S(s) - K]^+ \middle| S(t_0) = S_0\right]. \end{aligned}$$

This can be rewritten as

$$C(t_0, S; K, s) = \mathbb{E}[D(t_0, s)]\mathbb{E}\left[Fe^{(\alpha_1+\rho\alpha_2)W_2 - \frac{1}{2}(\alpha_1+\rho\alpha_2)^2} - Ke^{\rho\alpha_2 W_2 - \frac{1}{2}\rho^2\alpha_2^2}\middle| W_2 \geq -w, S(t_0) = S_0\right],$$

for

$$w = \frac{1}{\alpha_1}\left(\ln\frac{F}{K} - \frac{1}{2}\alpha_1^2 - \rho\alpha_1\alpha_2\right).$$

From probability theory, we have $\forall y, z \in \mathbb{R} : \mathbb{E}\left[e^{yW_2 - y^2/2}\middle| W_2 \geq -w\right] = \Phi(y + w)$, the call price becomes

$$\begin{aligned} C(t_0, S; K, s) &= \mathbb{E}[D(t_0, s)]\left\{F\Phi[(\alpha_1 + \rho\alpha_2) + w] - K\Phi\left(\frac{\ln\frac{F}{K} - \frac{\alpha_1^2}{2}}{\alpha_1}\right)\right\} && (2.67) \\ &= \mathbb{E}[D(t_0, s)]\left[F\Phi\left(\frac{\ln\frac{F}{K} + \frac{1}{2}\sigma_1^2 s}{\sigma_1\sqrt{s}}\right) - K\Phi\left(\frac{\ln\frac{F}{K} - \frac{1}{2}\sigma_1^2 s}{\sigma_1\sqrt{s}}\right)\right]. && (2.68) \end{aligned}$$

Note that eq. (2.68) holds for all log-normally distributed assets and is extremely similar to eq. (2.15). Since the ZCB $P(t, s)$ is log-normally distributed by eq. (2.39) we can use eq. (2.68) to find the price at time t_0 of a call option with strike K and expiry t_* on $P(t, s)$, such that $t_0 \leq t_* \leq s$. In the case of a constant volatility structure σ, the variance is given as [42]

$$\operatorname{Var} \ln P(t_0, s) = \sigma^2(s - t_*)^2(t_* - t_0). \tag{2.69}$$

Furthermore, for a ZCB, one has $\mathbb{E}[D(t_0, t_*)] = P(t_0, t_*)$ and $F = \frac{P(t_0,s)}{P(t_0,t_*)}$. Hence, the price of the call option reads

$$\boxed{\mathrm{ZCBC}(t_0; K, t_*, s) = P(t_0, s)\Phi(d_+) - KP(t_0, t_*)\Phi(d_-),} \tag{2.70}$$

where

$$d_\pm = \frac{\ln\left[\frac{P(t_0,s)}{P(t_0,t_*)K}\right] \pm \frac{1}{2}\sigma^2(s - t_*)^2(t_* - t_0)}{\sigma(s - t_*)\sqrt{t_* - t_0}}.$$

Chapter 3

Physics for Economists

The purpose of this chapter is to give the reader with a strong background in finance and economics a (very) brief introduction to the main ideas of physics that underlie the core objective of this thesis. We hope to sketch the journey that goes from stochastic processes in (classical) statistical mechanics all the way to (Feynman) path integrals in *quantum field theory* (QFT). This journey will start with some statistical mechanics where we discuss Markov processes, Langevin equations and how a path integral is constructed in this context. Thereafter, we generalize the main equations for stochastic processes to statistical field theory. To conclude, we go into more detail about some QFT. We explain how the path integral was first constructed in *quantum mechanics* (QM) and how to evaluate Gaussian (functional) integrals. Finally, we discuss the free scalar field and the harmonic oscillator in the field theory formalism.

3.1 Some Statistical Mechanics

Statistical mechanics is the incredibly interesting study of many-particle systems in which all particles interact, collide and bounce off of each other. The movement of a single particle in such a system is near impossible to predict, however their collective behavior may be described by assuming that the movement of an individual particle follows a *stochastic process*. In a stochastic process, the evolution of the position of a particle is subject to *fluctuations* and may take the form

$$\frac{\mathrm{d}x}{\mathrm{d}t} = g(x,t) + \rho(x,t)\eta(t). \tag{3.1}$$

This is a *stochastic differential equation* (SDE) of the Langevin type[1] (more on Langevin equations in section 3.1.2) and $\eta(t)$ is the *noise* [60]. We will restrict ourselves to the simplest and most common type of noise which is *Gaussian white noise* with the following mean and correlation respectively,

$$\langle \eta(t) \rangle = 0 \tag{3.2}$$

$$\langle \eta(t)\eta(t') \rangle = \delta(t-t'). \tag{3.3}$$

It must be stated that $\eta(t)$ is not a well-defined stochastic process. It can, however, be more or less considered as the derivative of a Wiener process $W(t)$, i.e. $\mathrm{d}W(t) = \eta(t)\,\mathrm{d}t$, despite that such a derivative does not exist [55]. A random process $x(t)$ as given in eq. (3.1) is entirely specified by its *hierarchy of probability densities* [60]. The probability of finding the value $x(t_i)$

[1] It might be instructive to point at resemblance with eq. (2.40) for $g(x,t) = k(\theta - x)$ and $\rho(x,t) = \rho$.

to be within the infinitesimal interval $[x_i, x_i + \mathrm{d}x_i]$ for $i = 1, \ldots, n$ (for $n \in \mathbb{N}_0$) reads

$$\mathcal{P}_n(x_1, t_1; \ldots; x_n, t_n)\,\mathrm{d}x_1 \ldots \mathrm{d}x_n\,. \tag{3.4}$$

This hierarchy obeys the following properties

- $\mathcal{P}_n \geq 0$
- $\mathcal{P}_n$ is invariant under the interchange of pairs (x_i, t_i) and (x_j, t_j) for $i, j = 1, \ldots, n$ and $i \neq j$
- $\int \mathcal{P}_n\,\mathrm{d}x_n = \mathcal{P}_{n-1}$ and $\int \mathcal{P}_1\,\mathrm{d}x_1 = 1$.

From this hierarchy of probability densities one may construct the *conditional probability density* $\mathcal{P}_{n|m}$ which gives the probability of having $x(t_i) = x_i$ for $i = m+1, \ldots, m+n$ given $x(t_j) = x_j$ for $j = 1, \ldots, m$. The conditional probability is defined as

$$\mathcal{P}_{n|m}(x_{m+1}, t_{m+1}; \ldots; x_{m+n}, t_{m+n}|x_1, t_1; \ldots; x_m, t_m) = \frac{\mathcal{P}_{n+m}(x_1, t_1; \ldots; x_{n+m}, t_{n+m})}{\mathcal{P}_m(x_1, t_1; \ldots; x_m, t_m)}. \tag{3.5}$$

We will focus on stochastic processes for which eq. (3.4) obeys the very specific (and useful) *Markovian property*.

3.1.1 Markov Processes

We call $\mathcal{P}_{1|1}(x_2, t_2|x_1, t_1)$ the *transition probability (density)* as it represents the probability that $x(t_2)$ takes value x_2 given $x(t_1) = x_1$. The Markov property states that, for any time sequence $t_1 < \cdots < t_n$, one has

$$\mathcal{P}_{1|n-1}(x_n, t_n|x_1, t_1; \ldots; x_{n-1}, t_{n-1}) = \mathcal{P}_{1|1}(x_n, t_n|x_{n-1}, t_{n-1}). \tag{3.6}$$

The transition probability density at time t_n, given $x(t_{n-1}) = x_{n-1}$, is independent from the values of x at earlier times $t_i < t_{n-1}$ [55]. A Markov process is thus completely determined by $\mathcal{P}_1(x_1, t_1)$ and $\mathcal{P}_{1|1}(x_2, t_1|x_1, t_1)$ since

$$\begin{aligned}\mathcal{P}_n(x_1, t_1; \ldots; x_n, t_n) &= \mathcal{P}_1(x_1, t_1)\mathcal{P}_{n-1|1}(x_2, t_2; \ldots; x_n, t_n) \\ &= \mathcal{P}_1(x_1, t_1)\mathcal{P}_{1|1}(x_2, t_2|x_1, t_1) \ldots \mathcal{P}_{1|1}(x_n, t_n|x_{n-1}, t_{n-1}).\end{aligned} \tag{3.7}$$

In the case of $n = 3$, we find by the properties of the hierarchy of probability densities and eq. (3.7)

$$\begin{aligned}\int \mathrm{d}x_2\, \mathcal{P}_3(x_1, t_1; x_2; t_2; x_3; t_3) &= \mathcal{P}_2(x_1, t_1; x_3, t_3) \\ &= \int \mathrm{d}x_2\, \mathcal{P}_1(x_1, t_1)\mathcal{P}_{1|1}(x_3, t_3|x_2, t_2)\mathcal{P}_{1|1}(x_2, t_2|x_1, t_1).\end{aligned}$$

Hence

$$\mathcal{P}_{1|1}(x_3, t_3|x_1, t_1) = \int \mathrm{d}x_2\, \mathcal{P}_{1|1}(x_3, t_3|x_2, t_2)\mathcal{P}_{1|1}(x_2, t_2|x_1, t_1) \tag{3.8}$$

This equation is called the *Chapman-Kolmogorov equation* (CKE). The usefulness of this equation is rather limited, however it can be cast into an equivalent and more convenient form: the *Master Equation*. For the derivation we refer to [55, 60] and we just posit the equation here:

$$\frac{\partial}{\partial t}\mathcal{P}_{1|1}(x, t|x_0, t_0) = \int \left[\mathcal{W}(x|x')\mathcal{P}_{1|1}(x', t'|x_0, t_0) - \mathcal{W}(x'|x)\mathcal{P}_{1|1}(x, t|x_0, t_0)\right] \mathrm{d}x'\,, \tag{3.9}$$

where $\mathcal{W}(x|x')$ is the transition probability for going from x' to x per unit time. The CKE has been cast into a differential form: the Master equation.

3.1.2 Brownian Motion & Langevin Equation

The Langevin approach offers a concrete and clear picture of the effects of fluctuations in macroscopic systems [55]. We will apply this approach to the well-known case of the *Brownian particle*. Consider a test (or Brownian) particle submerged in a fluid consisting of many randomly moving smaller particles. The equations of motion of the Brownian particle according to Newton's law is

$$m\dot{v} = F(t) + \mathfrak{F}(t), \tag{3.10}$$

where $\dot{v} = \frac{\mathrm{d}v}{\mathrm{d}t}$, m is the mass of the test particle, $F(t)$ is the force as a result of some external field and $\mathfrak{F}(t)$ is the force due to the many random collisions between the Brownian particle and fluid particles. We assume the test particle is much heavier than the fluid particles. This way, the velocity of the Brownian particle v is always in a kind of equilibrium. The random fluctuations in v are the result of a type of *friction* that slows the test particle down and of the numerous collisions with the fluid particles [60]. Since we assumed the test particle to be much heavier than the fluid particle, the particles of the surrounding fluid slow down a lot faster and can be considered to be in a constant state of equilibrium. We can thus write

$$\frac{\mathfrak{F}(t)}{m} = -\gamma v + \xi(t), \tag{3.11}$$

where γ is a friction coefficient and ξ is the so called *Langevin force* (or Gaussian white noise) obeying eqs. (3.2) and (3.3) with a small alteration for the latter: $\langle \xi(t)\xi(t')\rangle = \Omega\delta(t-t')$ for some positive constant Ω. Taking this all into consideration and in the absence of an external field, i.e.$F(t) = 0$, eq. (3.10) takes the form

$$\dot{v} = -\gamma v + \xi(t). \tag{3.12}$$

This is the Langevin equation for a Brownian particle [60].

Langevin equations can also be considered for stochastic processes aside from macroscopic systems. In general, a Langevin equation is a SDE that clearly shows the influence of both a deterministic- or drift term and a stochastic- or diffusion term on the evolution of a stochastic Markovian process [25].

3.1.3 Fokker-Planck Equation

There exists an intricate relation between the Langevin equation for a stochastic process $x(t)$ and the time-evolution of its corresponding transition probability $\mathcal{P}_{1|1}(x,t|x_0,t_0)$. The equation describing the latter is called a *Fokker-Planck equation* (FPE). A derivation of the FPE from the Langevin equation is given in [60] and starts from a general Langevin type equation as in eq. (3.1) where the noise term η still has the properties as given in eqs. (3.2) and (3.3). The main idea of the derivation is to evaluate the *conditional average*, i.e. the average with respect to the conditional probability density $\mathcal{P}_{1|1}$, of an arbitrary function $R(x)$. This can be done using the CKE in eq. (3.8). A Taylor expansion of $R(x)$ and rearrangement of terms ultimately yields

$$\frac{\partial}{\partial t}\mathcal{P}_{1|1}(x,t|y,t') = -\frac{\partial}{\partial x}\left[g(x,t)\mathcal{P}_{1|1}(x,t|y,t')\right] + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[\rho(x,t)^2\mathcal{P}_{1|1}(x,t|y,t')\right] \tag{3.13}$$

$$= \hat{H}_{\mathrm{FP}}\mathcal{P}_{1|1}(x,t|y,t'), \tag{3.14}$$

where the operator $\hat{H}_{\mathrm{FP}}$ is the Fokker-Planck Hamiltonian, an analog to a quantum Hamiltonian (more in section 3.3) [62]. The Fokker-Planck Hamiltonian in this case reads

$$\hat{H}_{\mathrm{FP}} = \frac{1}{2}\frac{\partial}{\partial x}\left[\frac{\partial}{\partial x}\rho(x,t)^2 - 2g(x,t)\right]. \tag{3.15}$$

Equation (3.14) is naturally called the FPE for $\mathcal{P}_{1|1}(x,t|y,t')$ corresponding to the stochastic process $x(t)$ obeying

$$\mathrm{d}x(t) = g(x(t),t)\,\mathrm{d}t + \rho(x(t),t)\eta(t)$$

given $x(t') = y$ for $t' < t$. If $g(x,t)$ is a linear function of x, e.g $g(x,t) = g_0 + g_1 x$, and $\rho(x,t)$ a constant, we call eq. (3.14) a *linear* FPE. Furthermore, if $g_1 < 0$, the (stationary) solution $\mathcal{P}_1(x)$ is a Gaussian distribution [55].

3.1.4 Path Integral for Markov Processes

We are now almost equipped to derive the path integral for a general Markov process as in eq. (3.1), but first we obtain an expression for the *Wiener path integral*. We start by constructing a partition of the time interval $[t_0, t_f]$ of N sub-intervals, each of size $(t_f - t_0)/N$. The probability that the process takes a value in $[a,b]$ (with $a < b$) at time $t > t_0$, given respective initial and final conditions $x(t_0)x_0$ and $x(t_f) = x_f$, reads

$$\mathcal{P}(a \le x \le b, t) = \int_a^b \mathrm{d}x\, \mathcal{P}_{1|1}(x,t|x_0,t_0). \tag{3.16}$$

The conditional probability density $\mathcal{P}_{1|1}(x,t|x',t')$ is sometimes also called the *(time) propagator* [60]. Since $\mathcal{P}_{1|1}(x,t|x',t')$, corresponding to the Langevin equation, obeys the CKE, we reiterate this to find the probability that $x(t)$, starting at $x(t_0) = x_0$, takes a value in $[a_j, b_j]$ at time $t_j = t_0 + j(t_f - t_0)/N$ to end in $x(t_f) = x_f$

$$\int_{a_1}^{b_1} \dots \int_{a_{N-1}}^{b_{N-1}} \mathrm{d}x_1 \dots \mathrm{d}x_{N-1}\, \mathcal{P}_{1|1}(x_f,t_f|x_{N-1},t_{N-1}) \dots \mathcal{P}_{1|1}(x_1,t_1|x_0,t_0). \tag{3.17}$$

This equation can be thought of as the integration over all possible paths that the stochastic process $x(t)$ could follow between $x(t_0) = x_0$ and $x(t_f) = x_f \equiv x_N$. A graphical representation of such paths can be seen in fig. 3.1. Now, by taking the limits $N \to \infty$ and $b_j - a_j \to 0$, for all $j = 1, \dots, N-1$, the continuous trajectory becomes more and more accurate [60].

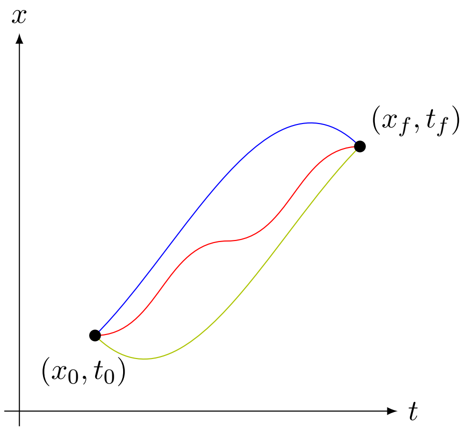


Figure 3.1: Graphical representation of three possible random paths between (x_0, t_0) and (x_f, t_f).

For a Wiener process $W(t)$, one has [60]

$$\mathcal{P}_{1|1}(W_j,t_j|W_{j-1},t_{j-1}) = \frac{1}{\sqrt{2\pi\Omega(t_j - t_{j-1})}} \exp\left[-\frac{(W_j - W_{j-1})^2}{2\Omega(t_j - t_{j-1})}\right], \tag{3.18}$$

where Ω is a positive constant such that $\langle\eta(t)\eta(t')\rangle = \Omega\delta(t-t')$. Substituting eq. (3.18) into eq. (3.17) yields the probability of following a single given path, which reads

$$\prod_{j=1}^{N} \frac{\mathrm{d}W_j}{\sqrt{2\pi\Omega\delta t}} \exp\left[-\frac{1}{2\Omega\delta t}\sum_{j=1}^{N}(W_j - W_{j-1})^2\right], \tag{3.19}$$

where $\delta t = t_j - t_{j-1}$ for all $j = 1,\ldots,N$. In taking the limit $N \to \infty$ and $\delta t \to 0$, the exponential in eq. (3.19) becomes

$$\exp\left[-\frac{1}{2\Omega}\sum_{j=1}^{N}\left(\frac{W_j - W_{j-1}}{\delta t}\right)^2 \delta t\right] \to e^{-\frac{1}{2\Omega}\int_{t_0}^{t_f}\left(\frac{\mathrm{d}W}{\mathrm{d}t}\right)^2 \mathrm{d}t} \tag{3.20}$$

and the product

$$\prod_{j=1}^{N} \frac{\mathrm{d}W_j}{\sqrt{2\pi\Omega\delta t}} \to \mathcal{D}[W] \tag{3.21}$$

which is called the *Wiener (integration) measure* [60]. The transition probability for the Wiener process is thus written as the *Wiener path integral*

$$\mathcal{P}_{1|1}(W,t|W_0,t_0) = \int \mathcal{D}[W] e^{-\frac{1}{2\Omega}\int_{t_0}^{t_f}\left(\frac{\mathrm{d}W}{\mathrm{d}\tau}\right)^2 \mathrm{d}\tau}, \tag{3.22}$$

where τ is just an integration variable. It is also important to note that, since we can consider the noise term $\eta(t)$ as a derivative of the Wiener process, eq. (3.22) can be written as

$$\mathcal{P}_{1|1}(\eta,t|\eta_0,t_0) = \int \mathcal{D}[\eta] e^{-\frac{1}{2\Omega}\int_{t_0}^{t_f}\eta(\tau)^2 \mathrm{d}\tau}, \tag{3.23}$$

with

$$\lim_{\substack{N\to\infty \\ \delta t\to 0}} \prod_{j=1}^{N} \sqrt{\frac{\delta t}{2\pi\Omega}} \frac{\mathrm{d}W_j}{\delta t} = \mathcal{D}[\eta]. \tag{3.24}$$

Now that all this is established, we can roughly go over the derivation for the Wiener path integral for the general Markov process. For a detailed explanation we again refer to [60]. The main idea is to discretize the Langevin equation eq. (3.1) and use eq. (3.20) to obtain the corresponding probability in x-space. For simplicity, we will consider $\rho(x,t) = \rho$, a positive constant. This is done by finding the Jacobian, $\mathcal{J}$, for the transformation $\{W_j\} \to \{x_j\}$ and using $\mathcal{P}(\{x_j\}) = \mathcal{J}\mathcal{P}(\{W_j\})$. In doing so, we can obtain a "discretized form" of the conditional probability.

$$\mathcal{P}_{1|1}(x,t|x_0,t_0) = \left(\frac{1}{2\pi\Omega\delta t}\right)^{\frac{N}{2}} \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \mathrm{d}W_1\ldots\mathrm{d}W_N\,\delta(x_f - x_N)\exp\left[-\frac{1}{2\Omega\delta t}\sum_j (W_j - W_{j-1})^2\right]. \tag{3.25}$$

Taking the continuum limit of $N \to \infty$ and $\delta t \to 0$ yields

$$\mathcal{P}_{1|1}(x,t|x_0,t_0) = \int \mathcal{D}[x] e^{\mathcal{S}[x(t)]}, \tag{3.26}$$

where

$$\mathcal{S}[x(t)] = \int_{t_0}^{t} \mathrm{d}t'\,\mathcal{L}[x(t'),\dot{x}(t')]\,\mathrm{d}t' \tag{3.27}$$

and

$$\mathcal{L}[\dot{x}(t'), x(t')] = -\frac{1}{2\Omega\rho^2}\left[\dot{x}(t') - g(x(t'), t')\right]^2. \tag{3.28}$$

We call eqs. (3.27) and (3.28) the *(stochastic) action* and *(stochastic Lagrangian)* respectively. The path integration measure in (3.26) is formally defined as

$$\int \mathcal{D}[x] = \lim_{\substack{N\to\infty \\ \delta t\to 0}} \prod_{j=1}^{N} \int_{-\infty}^{\infty} \frac{\mathrm{d}W_j}{\sqrt{2\pi\Omega\delta t}}. \tag{3.29}$$

This Lagrangian is related to the *Onsager-Machlup function* which, for a stochastic Markovian process as eq. (3.1), reads [44]

$$\mathcal{L}[\dot{x}(t), x(t)] = \frac{1}{2}\left[\frac{\dot{x}(t) - g(x(t), t)}{\rho(x(t), t)}\right]^2. \tag{3.30}$$

We successfully derived a path integral representation for the transition probability density corresponding to a (more or less) general Langevin equation.

Another approach to arrive at eq. (3.26) is by means of *stochastic quantization* [5, 20]. In stochastic quantization, we define the transition probability as a path integral over both the stochastic process $x(t)$ and the noise $\eta(t)$. However, we know how these two processes are related, namely via the Langevin equation in eq. (3.1). Therefore, we impose this relation through a Dirac delta such that the path integral can be written as

$$\begin{aligned}\mathcal{P}_{1|1}(x, t|x_0, t_0) &= \int \mathcal{D}[x]\mathcal{D}[\eta] \prod_{t'=t_0}^{t} \delta\left[\dot{x}(t') - g(x, t') - \rho(x, t')\eta(t')\right] e^{-\frac{1}{2}\int_{t_0}^{t}\mathrm{d}t'\eta(t')^2\mathrm{d}t'} \\ &= \int \mathcal{D}[x] e^{\mathcal{S}[x]},\end{aligned}$$

with

$$\mathcal{S}[x] = -\frac{1}{2}\int_{t_0}^{t}\mathrm{d}t'\left[\frac{\dot{x} - g(x, t')}{\rho(x, t')}\right]^2 \mathrm{d}t'. \tag{3.31}$$

The Lagrangian is thus given as

$$\mathcal{L}[\dot{x}, x] = -\frac{1}{2}\left[\frac{\dot{x} - g(x, t')}{\rho(x, t')}\right]^2, \tag{3.32}$$

which is again the Onsager-Machlup function seen in eq. (3.30). This procedure of stochastic quantization can also be used for *generating functions* and *-functionals*, to which we will come back to in sections 3.2 and 3.3.

From the path integral expression in eq. (3.26), the associated FPE can be rederived, this way confirming the form of eq. (3.26) is indeed correct [60].

3.1.5 Ornstein-Uhlenbeck Process

Using the path integral formulation of a general Langevin equation derived in the previous section, it is instructive to consider a well known example: the O-U process. The O-U process is an interesting process to study as it is the *only stationary, Gaussian* and *Markovian* process. It is stationary in the sense that $\mathcal{P}_1$ is time-independent and its transition probability only depends on the time-difference between the two "states", i.e. $\mathcal{P}_{1|1}(x_2, t_2|x_1, t_1) = \mathcal{T}_\tau(x_2|x_1)$ for $\tau = t_2 - t_1$. It is a Gaussian process since all of its hierarchy of probability densities $\mathcal{P}_n$ are Gaussian distributions and it is a Markov process as it obeys the Markov property given in eq. (3.7) [55]. This means that the Markov process determined by a linear FPE is the O-U process.

Consider thus a stochastic process of the form eq. (3.1) where $g(x,t) = k[\theta - x(t)]$ and $\rho(x,t) = \rho$ with $k, \theta, \rho > 0$. According to eq. (3.1), this process reads

$$\frac{\mathrm{d}x}{\mathrm{d}t} = k[\theta - x(t)] + \rho\eta(t), \tag{3.33}$$

where we now again take $\Omega = 1$. Associated to this process is a linear FPE

$$\frac{\partial}{\partial t}\mathcal{P}_{1|1}(x,t|x_0,t_0) = -\frac{\partial}{\partial x}\left[k(\theta - x)\mathcal{P}_{1|1}(x,t|x_0,t_0)\right] + \frac{1}{2}\rho^2\frac{\partial^2}{\partial x^2}\mathcal{P}_{1|1}(x,t|x_0,t_0) \tag{3.34}$$

to which the stationary solution is an O-U process. By eq. (3.27) and eq. (3.28), the corresponding action and Lagrangian read

$$\mathcal{S}[x(t)] = -\frac{1}{2\rho^2}\int_{t_0}^{t}\mathrm{d}t'\left[\dot{x}(t) - k(\theta - x)\right]^2 \tag{3.35}$$

$$\mathcal{L}[x(t), \dot{x}(t)] = -\frac{1}{2\rho^2}\left[\dot{x}(t) - k(\theta - x)\right]^2 \tag{3.36}$$

respectively. A general solution procedure to such path integrals is given in [60]. In the case of the O-U process the solution is given as [55, 60]

$$\mathcal{P}_{1|1}(x_f, t_f|x_0, t_0) = \left(\frac{k}{\pi\rho^2(1 - e^{-k\tau})}\right)^{\frac{1}{2}}\exp\left[-\frac{k}{\rho^2}\frac{\left[x_f - \theta\left(1 - e^{-k\tau}\right) - x_0 e^{-k\tau}\right]^2}{(1 - e^{-k\tau})}\right], \tag{3.37}$$

where $\tau = t_f - t_0$. This is a Gaussian distribution with conditional mean and variance given by eqs. (2.41) and (2.42) respectively.

3.2 Some Statistical Field Theory

Generalizations for Langevin equations to Langevin field equations were first proposed to describe critical phenomena, phase transitions and their dynamics close to equilibrium. This approach has proven to be fruitful for a great number of problems. This is due to field- and path integral formalism which enables the use of standard methods developed in QFT. This section briefly discusses the main concepts that will be useful to us for what is to come.

The equations derived in section 3.1.4 are readily generalized to a field theory formalism [38, 62]. Such generalization of eq. (3.1) to a Langevin field equation governing the stochastic evolution of a one-compontent scalar field $\varphi(t,x)$ is given as

$$\frac{\partial\varphi(t,x)}{\partial t} = \alpha(\varphi,t,x) + \sigma(\varphi,t,x)\mathcal{A}(t,x) \tag{3.38}$$

for some functionals $\alpha(\varphi,t,x), \sigma(\varphi,t,x)$ of φ and with initial conditions $\varphi(0,x) = \varphi_0(x)$. Note that now x denotes a d-dimensional space vector. The stochastic noise field $\mathcal{A}(t,x)$ has a Gaussian white noise characterization

$$\langle\mathcal{A}(t,x)\rangle = 0 \quad \text{and} \quad \langle\mathcal{A}(t,x)\mathcal{A}(t',x')\rangle = \delta^{(d)}(x - x')\delta(t - t'), \tag{3.39}$$

where $\delta^{(d)}(x)$ is the d-dimensional Dirac delta function. We will later see in section 3.3.3 that stochastic (or quantum) fields do not always have such a "trivial" two-point correlation function. An analogous derivation as described in section 3.1.3 yields the field theory extension of the FPE in eq. (3.14)

$$\frac{\partial}{\partial t}\mathcal{P}_{1|1}(\varphi, t|\varphi_0, t_0) = \mathcal{H}_{\mathrm{FP}}\mathcal{P}_{1|1}(\varphi, t|\varphi_0, t_0), \tag{3.40}$$

where

$$\mathcal{H}_{\mathrm{FP}} = \frac{1}{2}\int \mathrm{d}^d x\, \frac{\delta}{\delta\varphi(x)}\left[\frac{\delta}{\delta\varphi(x)}\sigma(\varphi, t, x)^2 - 2\alpha(\varphi, t, x)\right] \tag{3.41}$$

is the field theory generalization of eq. (3.15). Here, we introduced the functional differentiation which is entirely defined by

$$\frac{\delta\varphi(x)}{\delta\varphi(y)} = \delta^{(d)}(x-y). \tag{3.42}$$

To make this formalism workable with other QFT-methods, one usually constructs a path integral representation of the (time-dependent) correlation functions for the stochastic field φ. These correlation functions are obtained through a generating functional $\mathcal{Z}[J]$ which is given as

$$\mathcal{Z}[J] = \left\langle \exp\left[\int \mathrm{d}^d x\, \mathrm{d}t\, \varphi(t,x)J(t,x)\right]\right\rangle_{\mathcal{A}}, \tag{3.43}$$

where $\langle\ \rangle_{\mathcal{A}}$ denotes the average over the statistical field $\mathcal{A}$ and some *source function* $J(t,x)$. This can be written as a functional integral

$$= \int \mathcal{D}[\mathcal{A}] \exp\left[-\int \mathrm{d}^d x\, \mathrm{d}t\left(\frac{1}{2}\mathcal{A}(t,x)^2 - \varphi(t,x)J(t,x)\right)\right]. \tag{3.44}$$

This is sometimes also called the *partition functional* with the special case $Z \equiv \mathcal{Z}[0]$ the *partition function*. Of course, one must impose eq. (3.38). This can be achieved by denoting

$$E[\varphi, \mathcal{A}; t, x] \equiv \frac{\partial\varphi(t,x)}{\partial t} - \alpha(\varphi, t, x) - \sigma(\varphi, t, x)\mathcal{A}(t,x) = 0 \tag{3.45}$$

and inserting the identity

$$\int \mathcal{D}[\varphi] \det M \prod_{(t,x)} \delta(E[\varphi, \mathcal{A}; t, x]) = 1 \tag{3.46}$$

into eq. (3.44) with

$$M(t, x; t', x') = \frac{\delta E[\varphi, \mathcal{A}; t, x]}{\delta\varphi(t', x')}. \tag{3.47}$$

Suppose now, α and σ are independent of the field φ, i.e. $\alpha(\varphi, t, x) = \alpha(t,x)$ and $\sigma(\varphi, t, x) = \sigma(t,x)$, it is then shown in [62] that the generating functional $\mathcal{Z}[J]$ is formally given as

$$\mathcal{Z}[J] = \int \mathcal{D}[\varphi] \exp\left[\mathcal{S}[\varphi] + \int \mathrm{d}^d x\, \mathrm{d}t\, J(t,x)\varphi(t,x)\right] \tag{3.48}$$

with

$$\mathcal{S}[\varphi] = -\frac{1}{2}\int \mathrm{d}^d x\, \mathrm{d}t\left[\frac{\frac{\partial\varphi(t,x)}{\partial t} - \alpha(t,x)}{\sigma(t,x)}\right]^2, \tag{3.49}$$

where eq. (3.49) is a clear generalization of eq. (3.27) and its integrand can be considered as a type of "free field Lagrangian". It must be stated that in eq. (3.48), we assume the normalization

$Z = 1$. We will later see that this is not always the case but easily resolved by including an extra factor Z^{-1} in eq. (3.48). Also note that, for $J(t,x) = 0$, eq. (3.48) reduces to some sort of field theory extension for eq. (3.26):

$$\mathcal{Z}[0] = \int \mathcal{D}[\varphi] e^{\mathcal{S}}[\varphi], \tag{3.50}$$

with $\mathcal{S}[\varphi]$ given by eq. (3.49).

The n-point correlation function for the φ-field is then obtained from eq. (3.48) as follows

$$\begin{aligned}\langle \varphi(t_1,x_1)\dots\varphi(t_n,x_n)\rangle &= \int \mathcal{D}[\varphi]\ \varphi(t_1,x_1)\dots\varphi(t_n,x_n) e^{\mathcal{S}[\varphi]} \\ &= \frac{\delta}{\delta\varphi(t_1,x_1)}\dots\frac{\delta}{\delta\varphi(t_n,x_n)}\ \mathcal{Z}[J]\bigg|_{J=0}.\end{aligned} \tag{3.51}$$

We have shown how one can calculate correlation functions for a stochastic field, starting from its Langevin field equation. The concept of action, generating functional and correlation functions will return quite frequently in further chapters. The reader must however be aware that this section on statistical field theory is (extremely) far from complete. For a more in-depth discussion on statistical fields and their applications we kindly refer to [20, 38, 62].

3.3 Some Quantum Field Theory

In this chapter, we go over the core concepts in QFT that will help us develop a path integral- and field theory formalism for interest rates. We will start by introducing the path integral in one-particle QM following the approach in [45]. Thereafter, we briefly explain how to evaluate some Gaussian integrals. To conclude this section, we concisely discuss two simple (but useful) examples of quantum fields: the free scalar field and the harmonic oscillator.

3.3.1 Path Integral in Quantum Mechanics

We will describe a system with a *classical* Hamiltonian

$$H(x,p) = \frac{p^2}{2m} + V(x) \tag{3.52}$$

for position x and momentum p, both of which can be vectors. Equation (3.52) corresponds to the quantum mechanical Hamiltonian

$$\hat{H} = H(\hat{x},\hat{p}), \tag{3.53}$$

where $\hat{x}, \hat{p}$ are the position and momentum operator respectively, satisfying the famous Heisenberg's uncertainty principle $[\hat{x},\hat{p}] = \mathrm{i}\mathbb{I}$ with $\mathbb{I}$ the identity operator and the $\hbar = 1$ convention. The time-dependent Schrödinger equation

$$\frac{\partial}{\partial t}\left|\psi(t)\right\rangle = \hat{H}\left|\psi(t)\right\rangle \tag{3.54}$$

describes the time evolution of the states' *ket* $|\psi(t)\rangle$. The solution to eq. (3.54) reads

$$|\psi(t)\rangle = e^{-\mathrm{i}\hat{H}t}\left|\psi(0)\right\rangle \tag{3.55}$$

since $\hat{H}$ has no explicit time-dependence.

One can also consider position states $|x,t\rangle$, which obey

$$\hat{x}\,|x,t\rangle = x\,|x,t\rangle \tag{3.56}$$

for a real eigenvalue x and

$$\langle x,t|x',t\rangle = \delta(x-x'). \tag{3.57}$$

As the operators $\hat{x},\hat{p}$ are time-independent, the states $\{|x\rangle\}$ are time-independent as well. They can therefore be used to construct a basis for any other state [45]. This way, we can define a *wave function*

$$\psi(x,t) = \langle x|\psi(t)\rangle\,. \tag{3.58}$$

Acting on the wave function is the quantum mechanical Hamiltonian operator with $\hat{p} = -\mathrm{i}\frac{\partial}{\partial x}$. Now, by using eqs. (3.55) and (3.58), we can write

$$\psi(x,t) = \langle x|\, e^{-\mathrm{i}\hat{H}t}\,|\psi(0)\rangle\,. \tag{3.59}$$

For a complete set of states $\{|x_0\rangle\}$ satisfying the *completeness relation*

$$\int \mathrm{d}x_0\,|x_0\rangle\,\langle x_0| = \mathbb{I}, \tag{3.60}$$

eq. (3.59) can be rewritten as

$$\begin{aligned}\psi(x,t) &= \int \mathrm{d}x_0\,\langle x|\,e^{-\mathrm{i}\hat{H}t}\,|x_0\rangle\,\langle x_0|\psi(0)\rangle \\ &= \int \mathrm{d}x_0\,K(x,x_0;t)\psi(x_0,0),\end{aligned} \tag{3.61}$$

where

$$K(x,x_0;t) = \langle x|\,e^{-\mathrm{i}\hat{H}t}\,|x_0\rangle \tag{3.62}$$

is the *transition amplitude* (or sometimes *kernel*) for the particle to travel from position x_0 at time $t=0$ to position x at time t. It is the position representation of the Schrödinger time-evolution operator and a quantum mechanical analog to the transition probability discussed in section 3.1 [47]. The path integral formalism gives an expression for the amplitude $K(x,x_0;t)$ in terms of all the possible paths the particle can take between $x(0)=x_0$ and $x(t)=x$ [45]. Therefore, $K(x,x_0;t)$ will take a very different form.

The Feynman path integral in quantum theory, named after famous Nobel prize winning physicist Richard Feynman, was first introduced in his own doctoral thesis in 1942 at Princeton University. His new approach to describing QM (and later QFT) stems from the *principle of least action* which was used to generalize the concept of the Lagrangian and action from classical to quantum mechanics [26].

To obtain a path integral formulation of the transition amplitude in eq. (3.62), we consider the evolution of the wave function over the time interval $[0,s]$ such that $x(0)=x_0$ and $x(s)=x$. We then divide $[0,s]$ into $N-1$ sub-intervals of equal length $\delta t = s/N$. This allows us to write

$$e^{-\mathrm{i}\hat{H}s} = e^{-\mathrm{i}\hat{H}(s-t_{N-1})}\ldots e^{-\mathrm{i}\hat{H}t_1}. \tag{3.63}$$

If, at each instance t_j for $j=1,\ldots,N-1$, we introduce a complete set of states $\{|x_j\rangle\}$ each of them obeying the completeness relation eq. (3.60), we can rewrite eq. (3.62) is a similar manner as used to arrive at eq. (3.61)

$$K(x,x_0;s) = \prod_{j=1}^{N-1}\left(\int \mathrm{d}x_j\,\langle x_{j+1}|\,e^{-\mathrm{i}\hat{H}(t_{j+1}-t_j)}\,|x_j\rangle\right)\langle x_1|\,e^{-\mathrm{i}\hat{H}t_1}\,|x_0\rangle\,, \tag{3.64}$$

where $x_N = x$. The integration is carried out over all the possible values of $x(t)$ for $t \in \{t_1, \ldots, t_{N-1}\}$ given $x(0) = x_0$ and $x(s) = x$. A graphical representation of such a path can be seen in fig. 3.2.

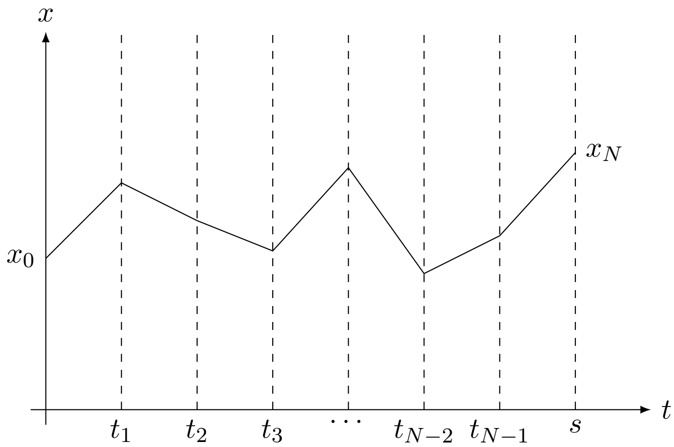


Figure 3.2: Graphical representation of a possible particle's trajectory over the time interval $[0, s]$.

As we will be interested in the continuum limit $N \to \infty$ and $\delta t \to 0$, we can make the approximation [45]

$$e^{-\mathrm{i}\hat{H}\delta t} \approx e^{-\mathrm{i}\frac{\hat{p}^2}{2m}\delta t} e^{-\mathrm{i}V(\hat{x})\delta t}. \tag{3.65}$$

Plugging eq. (3.65) into eq. (3.64) and using the fact that all the states $|x_j\rangle$ are eigenstates to the operator $\hat{x}$ allows us to write [45]

$$K(x, x_0; s) = \prod_{j=1}^{N-1} \int \mathrm{d}x_j \, e^{-\mathrm{i}V(x_j)\delta t} \langle x_{j+1}| \, e^{-\mathrm{i}\frac{\hat{p}^2}{2m}\delta t} \, |x_j\rangle \, . \tag{3.66}$$

We can find a more appropriate expression for $\langle x_{j+1}| \, e^{-\mathrm{i}\frac{\hat{p}}{2m}\delta t} \, |x_j\rangle$ by using a complete set of momentum states $\{|p\rangle\}$ in the same fashion as we used the position states $\{|q\rangle\}$. Such a complete set of momentum states obeys in an analogous manner eqs. (3.56) and (3.60) (except for an extra factor $1/2\pi$ in eq. (3.60)). Using the completeness relation for the momentum states and the fact that $\langle p|x\rangle = e^{-\mathrm{i}p\cdot x}$ we find, for $0 \leq j \leq N-1$,

$$\begin{aligned} \langle x_{j+1}| \, e^{-\mathrm{i}\frac{\hat{p}^2}{2m}\delta t} \, |x_j\rangle &= \int \frac{\mathrm{d}p_j}{2\pi} \langle x_{j+1}| \, e^{-\mathrm{i}\frac{\hat{p}^2}{2m}\delta t} \, |p_j\rangle \langle p_j|x_j\rangle \\ &= \int \frac{\mathrm{d}p_j}{2\pi} e^{-\mathrm{i}\frac{p_j^2}{2m}\delta t} e^{ip_j\cdot(x_{j+1}-x_j)}. \end{aligned} \tag{3.67}$$

This is a Gaussian-type integral and will be discussed in much more detail in section 3.3.2. For now, we just state the result

$$\langle x_{j+1}| \, e^{-\mathrm{i}\frac{\hat{p}^2}{2m}\delta t} \, |x_j\rangle = \sqrt{\frac{m}{2\pi i\delta t}} e^{\mathrm{i}\frac{m}{2}\left(\frac{x_{j+1}-x_j}{\delta t}\right)^2 \delta t}. \tag{3.68}$$

Equation (3.66) then becomes

$$K(x, x_0, s) = \left(\frac{m}{2\pi\mathrm{i}\delta t}\right)^{\frac{N}{2}} \prod_{j=1}^{N-1} \int \mathrm{d}q_j \exp\left[\mathrm{i}\sum_{j=0}^{N-1} \frac{m}{2}\left(\frac{x_{j+1}-x_j}{\delta t}\right)^2 \delta t - \mathrm{i}\sum_{j=0}^{N-1} V(x_j)\delta t\right]. \tag{3.69}$$

The continuum limit $N \to \infty$ and $\delta t \to 0$ the yields the final result

$$K(x, x_0; s) = \langle x| \, e^{-\mathrm{i}\hat{H}s} \, |x_0\rangle = \int \mathcal{D}[x] e^{i\mathcal{S}[x]}, \tag{3.70}$$

where the action is

$$\mathcal{S}[x] = \int_0^s \mathrm{d}t \, \mathcal{L}[x, \dot{x}] \tag{3.71}$$

with Lagrangian

$$\mathcal{L}[x, \dot{x}] = \frac{1}{2} m\dot{x}^2 - V(x) \tag{3.72}$$

and integration measure

$$\int \mathcal{D}[x] = \lim_{\substack{N\to\infty \\ \delta t \to 0}} \left(\frac{m}{2\pi \mathrm{i}\delta t}\right)^{\frac{N}{2}} \prod_{j=1}^{N-1} \int \mathrm{d}q_j \, . \tag{3.73}$$

The path integral formulation of QM is an equivalent description that "sums" over a continuous range of possible paths between x_0 and x over a time period $[0, s]$ (as in the derivation above), with a weight $e^{\mathrm{i}\mathcal{S}}$ associated to each path [45].

Notice how these equations are very similar to eqs. (3.26) to (3.28) and (3.49) from sections 3.1.4 and 3.2. For the Wiener path integral we found

$$\mathcal{P}_{1|1}(x, s|x_0, t_0) = \int \mathcal{D}[x] e^{\mathcal{S}[x]},$$

in statistical field theory

$$\mathcal{Z}[J = 0] = \int \mathcal{D}[\varphi] e^{\mathcal{S}[\varphi]}$$

and for the Feynman path integral in QM

$$K(x, x_0; s, t_0) = \int \mathcal{D}[x] e^{\mathrm{i}\mathcal{S}[x]},$$

where each expression can be considered as some sort of "probability distribution" and in all cases the action is defined in the same way:

$$\mathcal{S} = \int_{t_0}^{s} \mathrm{d}t \, \mathcal{L},$$

leaving the dependence of $\mathcal{S}$ and $\mathcal{L}$, either on a stochastic process x, stochastic field φ or position state x implicit.

We can however perform a transformation such that the Feynman path integral is of exactly the same form as the Wiener integral. This transformation is a *Euclidean rotation* that analytically continues the integrals to imaginary time, i.e.

$$t \to t' = -\mathrm{i}\tau \quad \text{with } \tau \in \mathbb{R}. \tag{3.74}$$

The Lagrangian in eq. (3.72) becomes

$$\mathcal{L}\left[\frac{\mathrm{d}x(\tau)}{\mathrm{d}\tau}, x(\tau)\right] = -\frac{1}{2} m \left[\frac{\mathrm{d}x(\tau)}{\mathrm{d}\tau}\right]^2 - V(x) \tag{3.75}$$

which means the action in eq. (3.71) reads

$$\mathcal{S}[x(\tau)] = -\mathrm{i}\int_0^{\tau_s} \mathrm{d}\tau\, \mathcal{L}\left[\frac{\mathrm{d}x(\tau)}{\mathrm{d}\tau}, x(\tau)\right], \tag{3.76}$$

where $\tau_s = -\mathrm{i}s$. Plugging eqs. (3.75) and (3.76) in to eq. (3.70) yields

$$K(x, x_0; s) = \int \mathcal{D}[x] e^{\mathcal{S}[x(\tau)]}. \tag{3.77}$$

The Euclidean transformation to imaginary time causes the quantum probability amplitudes to appear as a (continuous) sum similar to those from statistical field theory [36]. This allows for a transition from problems in QFT to questions in (classical) statistical field theory and vice versa.

It was actually Kac's discovery that the derivation above can be reproduced using the operator $e^{-\hat{H}t}$ instead of $e^{-\mathrm{i}\hat{H}t}$ to arrive at the Wiener path integral from section 3.1.4 [19]. He showed that this real exponent can, in a rigorous manner, be mathematically justified in relation to the Wiener path integral and that it has applications in statistical physics and stochastic processes. For this reason, eqs. (3.26), (3.50) and (3.70) are often also named *Feynman-Kac equations*.

3.3.2 Gaussian Integrals

In eq. (3.67), we encountered a Gaussian integral and, without mentioning, we already encounter quite some Gaussian *functional* integrals. As it turns out, these types of integrals are very common in field theories. Hence, it makes sense to explain their evaluation procedures. The methods we discuss here are based on the approach described in [45].

We start by briefly sketching the evaluation of the generic integral, which is of the form

$$\int \mathrm{d}x\, e^{-\frac{1}{2}\lambda x^2} = \sqrt{\frac{2\pi}{\lambda}}, \tag{3.78}$$

for $\lambda > 0$. This integral can be evaluated by making the transformation $x \to x' = \sqrt{\frac{2}{\lambda}}x$ and considering the square of this integral. A new transformation to polar coordinates allows a fairly straightforward evaluation of the integral yielding π (up to a factor $\frac{2}{\lambda}$ from the initial transformation). Taking the square root ultimately yields the result in eq. (3.78).

This can relatively easily be extended to the n-dimensional case in which we consider, for $x \in \mathbb{R}^n$

$$Z(A) = \int \mathrm{d}^n x\, e^{-\frac{1}{2}x\cdot Ax}, \tag{3.79}$$

where A is a $n \times n$ symmetric, positive definite matrix and $x \cdot y = x^T y$ denotes the standard scalar product for $x, y \in \mathbb{R}^n$ with x^T the transpose of x. This type of integral corresponds to a free field theory in a way and is the equivalent of a *partition function* [45]. For such a matrix A there exists an orthogonal matrix U (i.e. $UU^T = \mathbb{I}$) such that $UAU^T = \mathrm{diag}(\lambda_1 \ldots \lambda_n)$ with $\lambda_i > 0, \forall i = 1, \ldots, n$. By making the transformation $x \to x' = Ux$, one finds, due to the orthogonality property of U, that

$$Z(A) = \int \mathrm{d}^n x'\, e^{-\frac{1}{2}x'\cdot UAU^T x'} = \prod_{j=1}^{n} \int \mathrm{d}x'_j\, e^{-\frac{1}{2}\lambda_j x'^2_j},$$

where we used $x' \cdot UAU^T x' = \sum \lambda_j x'^2_j$. This is just the product of n generic integrals to which we known the answer, hence

$$Z(A) = \frac{(2\pi)^{\frac{n}{2}}}{\sqrt{\det A}}. \tag{3.80}$$

If, now for example we have a linear term

$$Z(A,b) = \int \mathrm{d}^n x \, e^{-\frac{1}{2}x\cdot Ax + b\cdot x}, \tag{3.81}$$

with $b \in \mathbb{R}^n$, we can readily make the substitution $x' = x - A^{-1}b$ such that the exponent becomes $-\frac{1}{2}x' \cdot Ax' + \frac{1}{2}b \cdot A^{-1}b$, yielding

$$Z(A,b) = e^{\frac{1}{2}b\cdot A^{-1}b} \frac{(2\pi)^{\frac{n}{2}}}{\sqrt{\det A}}. \tag{3.82}$$

This neat little trick is called *completing the square* and will be used quite frequently. Also, one can see how eq. (3.81) is the finite dimensional equivalent of a *partition functional*.

For an arbitrary function[2] $f(x)$, we can define the expectation value as

$$\langle f(x) \rangle = \frac{1}{Z(A, b=0)} \int \mathrm{d}^n x \, f(x) e^{-\frac{1}{2}x\cdot Ax}, \tag{3.83}$$

where the pre-factor $Z(A,0)^{-1}$ functions as a normalization such that $\langle 1 \rangle = 1$. In particular we see

$$\langle x_i \rangle = \frac{1}{Z(A,0)} \frac{\partial}{\partial b_i} Z(A,b)|_{b=0} = 0 \tag{3.84}$$

$$\langle x_i x_j \rangle = \frac{1}{Z(A,0)} \frac{\partial}{\partial b_i} \frac{\partial}{\partial b_j} Z(A,b)|_{b=0} = A^{-1}_{ij}, \tag{3.85}$$

which is the finite dimensional counterpart to eq. (3.51).

An extensive discussion of *non-Gaussian integrals* will not be handled here, however we do briefly mention them as they correspond to *interacting field theories* [45]. The basic integral in this case is given as

$$Z = \frac{1}{Z(A,0)} \int \mathrm{d}^n x \, e^{-\frac{1}{2}x\cdot Ax + b\cdot x - V(x)}, \tag{3.86}$$

where again the pre-factor ensures normalization and $V(x)$ is real function, bounded by below with $V(0) = 0$. These non-Gaussian integrals can only be evaluated by means of *perturbation expansion*. Every term in such a perturbation expansions gets associated a diagram called *Feynman diagrams*. For a more detailed explanation on perturbation expansions and Feynman diagrams we refer to [20, 24, 45, 47, 62].

When extending these methods to path integrals for (quantum) fields, the vectors become fields, the matrices become operators and the partition functions become partition functionals. We hope to demonstrate this by studying two of the most well-known cases: the free scalar field and the harmonic oscillator.

3.3.3 Free Scalar Field

Not dwelling to much on metrics and Einstein-summation convention, the Lagrangian (density) for a free scalar field $\varphi(x)$, with $x = (t, \vec{x})$, $\vec{x} \in \mathbb{R}^d$, reads

$$\mathcal{L}[\dot{\varphi}, \varphi] = \frac{1}{2}\dot{\varphi} - \frac{1}{2}(\nabla\varphi)^2 - \frac{1}{2}m\varphi^2, \tag{3.87}$$

[2] The function f has nothing to do with the instantaneous forward interest rate at this moment.

where in this case $\dot{\varphi} = \frac{\partial \varphi}{\partial t}$. By using integration by parts with φ vanishing at the boundaries, the action can be given the form

$$\mathcal{S}[\varphi] = -\frac{1}{2}\int \mathrm{d}^d x\, \varphi(x)\triangle\varphi(x), \tag{3.88}$$

where

$$\triangle \equiv \frac{\partial^2}{\partial t^2} - \nabla^2 + m^2, \tag{3.89}$$

is named the *Klein-Gordon operator*. The field theory can be defined through a partition functional with source function $J(x)$

$$\mathcal{Z}[J] = \int \mathcal{D}[\varphi] e^{i\mathcal{S}[\varphi]+i\int \mathrm{d}^d x J(x)\varphi(x)}. \tag{3.90}$$

Note that eq. (3.90) is the infinite-dimensional analog of eq. (3.81). We thus hope to rewrite eq. (3.90) by means of an infinite-dimensional analog for completing the square. This is achieved as follows:

Start by defining $\tilde{\varphi}(x) = \varphi(x) - \triangle^{-1}J(x)$, where we leave the expression for $\triangle^{-1}$ undisclosed for now. We then find

$$\begin{aligned}
\mathcal{S}[\tilde{\varphi}] &= -\frac{1}{2}\int \mathrm{d}^d x \left[\varphi(x) - \triangle^{-1}J(x)\right] \triangle \left[\varphi(x) - \triangle^{-1}J(x)\right] \\
&= -\frac{1}{2}\int \mathrm{d}^d x \left[\varphi(x)\triangle\varphi(x) - 2\varphi(x)J(x) + J(x)\triangle^{-1}J(x)\right] \\
&= \mathcal{S}[\varphi] + \int \mathrm{d}^d x\, \tilde{\varphi}(x)J(x) - \frac{1}{2}\int \mathrm{d}^d x\, J(x)\triangle^{-1}J(x).
\end{aligned}$$

Hence, eq. (3.90) can be written as

$$\mathcal{Z}[J] = Z e^{\frac{1}{2}\mathrm{i}\int \mathrm{d}^d x J(x)\triangle^{-1}J(x)}, \tag{3.91}$$

where again $Z = \mathcal{Z}[0]$. Note that eq. (3.91) is an infinite-dimensional generalization of eq. (3.82). We now only need to find $\triangle^{-1}$. This is equivalent to solving the equation

$$-\triangle G(x-y) = \delta^d(x-y), \tag{3.92}$$

such that $G(x)$ is the *Green's function* of the operator $\triangle$ [45]. $G(x)$ is often called the propagator for the field theory. Equation (3.92) can be solved for G by means of Fourier transformations and complex integration (for a detailed derivation see [45]). One ultimately finds

$$\mathrm{i}G(x) = \int \frac{\mathrm{d}^{p-1}p}{(2\pi)^{d-1}} \frac{1}{2\sqrt{m^2+\vec{p}^2}} e^{\mathrm{i}\vec{p}\cdot\vec{x}} \left[\Theta(t)e^{-\mathrm{i}\sqrt{m^2+\vec{p}^2}t} + \Theta(-t)e^{\mathrm{i}\sqrt{m^2+\vec{p}^2}t}\right], \tag{3.93}$$

where $\Theta(t)$ is the stepfunction defined as

$$\Theta(t) = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0. \end{cases} \tag{3.94}$$

The generating functional $\mathcal{Z}[J]$ can thus be rewritten in the form

$$\mathcal{Z}[J] = Z e^{-\frac{1}{2}\int \mathrm{d}^d x \mathrm{d}^d y J(x)\mathrm{i}G(x-y)J(y)} \tag{3.95}$$

equivalent to eq. (3.91).

Note that often one either chooses the normalization $\mathcal{Z}[0] = 1$ or one defines the generating functional in eq. (3.90) with a normalization factor $\mathcal{Z}[0]^{-1}$ in front such that eq. (3.91) takes the form

$$\mathcal{Z}[J] = e^{-\frac{1}{2}\int \mathrm{d}^d x \mathrm{d}^d y J(x) \mathrm{i} G(x-y) J(y)}. \tag{3.96}$$

The n-point correlation functions can be determined from the generating functional in eq. (3.95) as

$$\begin{aligned}\langle \varphi(x_1) \dots \varphi(x_n) \rangle &= \int \mathcal{D}[\varphi] \varphi(x_1) \dots \varphi(x_n) e^{\mathrm{i}\mathcal{S}[\varphi]} \\ &= (-\mathrm{i})^n \frac{\delta}{\delta J(x_1)} \cdots \frac{\delta}{\delta J(x_n)} \mathcal{Z}[J] \bigg|_{J=0}. \end{aligned} \tag{3.97}$$

The n-point correlation function is thus, up to the imaginary unit, of the exact same form as eq. (3.51). When taking a closer look at the two-point correlation function, we see

$$\langle \varphi(x_1)\varphi(x_2) \rangle = \mathrm{i}G(x_1 - x_2). \tag{3.98}$$

which shows the difference with the two-point correlation function of the Gaussian field in eq. (3.39) where (aside from the imaginary unit) the propagator is just the Dirac-delta. This show the non-trivial correlation between different (space-time) points of the quantum field.

With the path integral formalism one can thus define the entirety of the field theory from the generating functional $\mathcal{Z}[J]$, the correct propagator G and the correlation functions.

3.3.4 Harmonic Oscillator

In this section we will discuss probably the most important physical example for the context of this thesis: the (quantum) harmonic oscillator. The importance of this example will become clear once we construct the field Lagrangian for the forward interest rates in chapter 5. The main goal here is to derive an explicit expression for the propagator of the harmonic oscillator. We will do so for special boundary conditions that allow for an easy transition to the field of forward interest rates. The approach we follow is in less detail given in [5].

The harmonic oscillator can be considered as a trivial $(0 + 1)$ field theory, i.e. only a time-component. We will consider the system in Euclidean time, which means that the Euclidean Lagrangian for a particle of mass m and position $x(t)$, moving through a potential $V(x) = m\omega^2 x^2/2$ in the presence of an external force $J(t)$ reads

$$\mathcal{L}[\dot{x}, x] = -\frac{1}{2} m\dot{x}(t)^2 - \frac{1}{2} m\omega^2 x(t)^2 + J(t)x(t) \tag{3.99}$$

The external force $J(t)$ will act as the source function in the generating functional. The action over the time interval $[t_i, t_f]$ is

$$\mathcal{S}[x] = \int_{t_i}^{t_f} \mathrm{d}t \, \mathcal{L}(\dot{x}, x) \tag{3.100}$$

For the context of forward interest rates (which will be discussed in chapter 5), it is in our interest to consider the case where the initial and final positions, $x(t_i)$ and $x(t_f)$, are random. The boundary conditions would then not fix $x(t_i)$ and $x(t_f)$ to some position x_i and x_f respectively but rather, with some mean-reversion process in mind, we impose that

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t}\bigg|_{t=t_i} = \frac{\mathrm{d}x(t)}{\mathrm{d}t}\bigg|_{t=t_f} = 0. \tag{3.101}$$

This way, $x(t_i)$ and $x(t_f)$ can take any possible value but they do not change at the boundaries. That is, we assume the mean reversion level has been reached (up to some small fluctuations) at the final time. Such boundary conditions are named *Neumann* boundary conditions, after famous mathematician John Von Neumann. Using these boundary conditions, the action can be rewritten by means of integration by parts

$$\mathcal{S}[x] = \frac{m}{2}\int_{t_i}^{t_f} \mathrm{d}t \left[x(t)\frac{\mathrm{d}^2 x(t)}{\mathrm{d}t^2} - \omega^2 x(t)^2\right] + \int_{t_i}^{t_f} \mathrm{d}t\, J(t)x(t) \tag{3.102}$$

$$= -\frac{m}{2}\int_{t_i}^{t_f} \mathrm{d}t\, x(t)\left[-\frac{\mathrm{d}^2}{\mathrm{d}t^2} + \omega^2\right] x(t) + \int_{t_i}^{t_f} \mathrm{d}t\, J(t)x(t). \tag{3.103}$$

Let us again define $\triangle \equiv m\omega^2\left[1 - \frac{1}{\omega^2}\frac{\mathrm{d}^2}{\mathrm{d}t^2}\right]$ such that the generating functional takes the form of eq. (3.95)

$$\mathcal{Z}[J, t_i, t_f] = \int \mathcal{D}[x] e^{\mathcal{S}[x]} = \exp\left[\frac{1}{2}\int_{t_i}^{t_f} \mathrm{d}t\,\mathrm{d}t'\, J(t)G(t, t'; t_i, t_f)J(t')\right], \tag{3.104}$$

where $G(t, t'; t_i, t_f) = \triangle^{-1}$ is again the propagator for which we have now included the boundaries t_i, t_f. Furthermore, by definition of $\triangle$, we have, as in eq. (3.92),

$$\triangle G(t, t'; t_i, t_f) = \delta(t - t'). \tag{3.105}$$

From this we will now derive an explicit expression for the propagator $G(t, t'; t_i, t_f)$.

The path integral in eq. (3.104) is evaluated over all possible paths $x(t)$ that satisfy the specified boundary conditions eq. (3.101). All these functions can be written as a Fourier expansion

$$x(t) = a_0 + \sum_{n=1}^{\infty} a_n \cos\left(n\pi\frac{t - t_i}{T}\right), \tag{3.106}$$

where $T \equiv t_f - t_i$. The path integration measure then is

$$\int \mathcal{D}[x] = \mathcal{N}\prod_{n=0}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}a_n \tag{3.107}$$

for some normalization constant $\mathcal{N}$. Plugging eq. (3.106) into eq. (3.100) yields

$$\begin{aligned}
\mathcal{S}[x] - \int_{t_i}^{t_f} \mathrm{d}t\, J(t)x(t) &= -\frac{m}{2}\int_{t_i}^{t_f} \mathrm{d}t \left[a_0 + \sum_{n=1}^{\infty} a_n \cos\left(n\pi\frac{t - t_i}{T}\right)\right]\left[\omega^2 - \frac{\mathrm{d}^2}{\mathrm{d}t^2}\right]\left[a_0 + \sum_{k=1}^{\infty} a_k \cos\left(k\pi\frac{t - t_i}{T}\right)\right] \\
&= -\frac{m}{2}\int_{t_i}^{t_f} \mathrm{d}t \left[a_0 + \sum_{n=1}^{\infty} a_n \cos\left(n\pi\frac{t - t_i}{T}\right)\right]\left[\sum_{k=1}^{\infty} a_k \left(\frac{k\pi}{T}\right)^2 \cos\left(k\pi\frac{t - t_i}{T}\right)\right. \\
&\quad \left. + \omega^2 a_0 + \omega^2\sum_{k=1}^{\infty} a_k \cos\left(k\pi\frac{t - t_i}{T}\right)\right] \\
&= -\frac{m}{2}\int_{t_i}^{t_f} \mathrm{d}t \left[\omega^2 a_0^2 + \sum_{n,k=1}^{\infty} a_n a_k \left(\frac{k\pi}{T}\right)^2 \cos\left(n\pi\frac{t - t_i}{T}\right)\cos\left(k\pi\frac{t - t_i}{T}\right)\right. \\
&\quad \left. + \omega^2 \sum_{n,k=1}^{\infty} a_n a_k \cos\left(n\pi\frac{t - t_i}{T}\right)\cos\left(k\pi\frac{t - t_i}{T}\right)\right],
\end{aligned}$$

where we used the fact that $\int_{t_i}^{t_f} \mathrm{d}t \cos\left(n\pi \frac{t-t_i}{T}\right) = 0,\ \forall n \in \mathbb{Z}$. Now using the orthogonality relation for the cosine, i.e. $\int_{t_i}^{t_f} \mathrm{d}t \cos\left(n\pi \frac{t-t_i}{T}\right) \cos\left(k\pi \frac{t-t_i}{T}\right) = \frac{T}{2}\delta_{n,k},\ \forall n, k \in \mathbb{Z}$, and carrying out the integrals gives

$$\begin{aligned}\mathcal{S}[x] - \int_{t_i}^{t_f} \mathrm{d}t\, J(t)x(t) &= -\frac{m}{2}\left[\omega^2 a_0^2 T + \sum_{n=1}^{\infty} a_n^2 \left(\frac{n\pi}{T}\right)^2 \frac{T}{2} + \omega^2 \sum_{n=1}^{\infty} a_n^2 \frac{T}{2}\right] \\ &= -\frac{m}{2}\omega^2 T\left[a_0^2 + \frac{1}{2}\sum_{n=1}^{\infty} a_n^2\left(1 + \frac{n^2\pi^2}{\omega^2 T^2}\right)\right].\end{aligned}$$

If we now introduce $\kappa_0 = m\omega^2 T$, $\kappa_n = \frac{1}{2}m\omega^2 T\left(1 + \frac{n^2\pi^2}{\omega^2T^2}\right)$, $\forall n \in \mathbb{N}_0$ and $J_n = \int_{t_i}^{t_f} \mathrm{d}t\, J(t)\cos\left(n\pi\frac{t-t_i}{T}\right)$, the action can ultimately be rewritten as

$$\mathcal{S}[a] = -\frac{1}{2}\sum_{n=0}^{\infty}\left[\kappa_n a_n^2 + 2J_n a_n\right], \tag{3.108}$$

where

$$\int_{t_i}^{t_f} \mathrm{d}t\, J(t)x(t) = \sum_{n=0}^{\infty} J_n a_n,$$

is placed on the RHS of eq. (3.108). We see that all the Gaussian integrals over the variables a_n are now decoupled such that eq. (3.104) takes the form

$$\mathcal{Z}[J, t_i, t_f] = \mathcal{N}\prod_{n=0}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}a_n\, e^{-\frac{1}{2}\kappa_n a_n^2 + J_n a_n}.$$

These integrals can all be evaluated by completing the square in the exponent such that, by eq. (3.82),

$$\mathcal{Z}[J, t_i, t_f] = e^{\frac{1}{2}\sum_{n=0}^{\infty}\frac{J_n^2}{\kappa_n}}.$$

Plugging in the expressions for J_n and κ_n allows us to factor out the source functions $J(t)$

$$\mathcal{Z}[J, t_i, t_f] = \exp\left[\frac{1}{2}\int_{t_i}^{t_f} \mathrm{d}t\,\mathrm{d}t'\, J(t)\cos\left(n\pi\frac{t-t_i}{T}\right)\frac{1}{\kappa_n}\cos\left(n\pi\frac{t'-t_i}{T}\right)J(t')\right]. \tag{3.109}$$

Identifying eq. (3.104) with eq. (3.109) shows that

$$\begin{aligned}G(t, t'; t_i, t_f) &= \sum_{n=0}^{\infty}\cos\left(n\pi\frac{t-t_i}{T}\right)\frac{1}{\kappa_n}\cos\left(n\pi\frac{t'-t_i}{T}\right) \\ &= \frac{1}{m\omega^2 T}\left[1 + 2\sum_{n=1}^{\infty}\cos\left(n\pi\frac{t-t_i}{T}\right)\left(\frac{1}{1+\frac{n^2\pi^2}{\omega^2T^2}}\right)\cos\left(n\pi\frac{t'-t_i}{T}\right)\right].\end{aligned} \tag{3.110}$$

The sum can be simplified by introduction $\theta = t - t_i$ and $\theta' = t' - t_i$ and using the trigonometric identity $\cos(\theta \pm \theta') = \cos(\theta)\cos(\theta') \pm \sin(\theta)\sin(\theta')$:

$$2\sum_{n=1}^{\infty}\frac{\cos\left(\frac{n\pi}{T}\theta\right)\cos\left(\frac{n\pi}{T}\theta\right)}{1+\frac{n^2\pi^2}{\omega^2T^2}} = \frac{\omega^2T^2}{\pi^2}\sum_{n=1}^{\infty}\frac{\cos\left(n\pi\frac{\theta+\theta'}{T}\right) + \cos\left(n\pi\frac{\theta-\theta'}{T}\right)}{\frac{\omega^2T^2}{\pi^2} + n^2}. \tag{3.111}$$

The identity[3] [5]

$$\sum_{n=1}^{\infty}\frac{\cos(n\beta)}{a^2+n^2}=\frac{\pi}{2a}\frac{\cosh[(\pi-|\beta|)a]}{\sinh(\pi a)}-\frac{1}{2a^2}$$

for $\beta \in [0, 2\pi]$, allows us to write

$$2\sum_{n=1}^{\infty}\frac{\cos\left(\frac{n\pi}{T}\theta\right)\cos\left(\frac{n\pi}{T}\theta\right)}{1+\frac{n^2\pi^2}{\omega^2T^2}}=\frac{\omega T}{2}\frac{\cosh[\omega(T-|\theta+\theta'|)]+\cosh[\omega(T-|\theta-\theta'|)]}{\sinh(\omega T)}-1. \tag{3.112}$$

By combining this with eq. (3.110), the final expression for the propagator becomes

$$G(t,t';t_i,t_f)=\frac{\cosh[\omega(t_f-t_i-t-t'+2t_i)]+\cosh[\omega(t_f-t_i-|t-t'|)]}{2m\omega\sinh[\omega(t_f-t_i)]}. \tag{3.113}$$

The propagator in eq. (3.113), together with the generating functional eq. (3.104) define the field theory for the harmonic oscillator. Note that eq. (3.113) is a symmetric function of $t,\ t'$. We will encounter this propagator quite frequently (see chapter 5) as it turns out to be equivalent to the propagator of the (simplest) field theory of the forward interest rates.

[3] This identity can easily be validated by plotting both sides as a function of $\beta \in [0, \pi]$.

Chapter 4

Path Integrals for Interest Rates

Now that the link between a stochastic Markov process and its path integral description has been established we can deploy this formalism to the two interest rate models that were discussed in section 2.3: the Vasicek- and HJM model. It might be important to note that this chapter will be a mere reformulation of work already done and nothing "new" is introduced. However, this reformulation is advantageous for a smooth transition into the field theory of interest rates in chapter 5. This is due to the relation between the path integral description of statistical physics and that of QFT, which we touched upon in section 3.3.1. The following two sections will be computationally heavy, as are many problems in a path integral formalism. We start with the simplest of the two interest rate models, which is the Vasicek model.

4.1 The Vasicek Path Integral

We first remind the reader about the dynamics which the short-rate obeys in the Vasicek model, namely those in eq. (2.40);

$$\frac{\mathrm{d}r}{\mathrm{d}t} = k[\theta - r(t)] + \rho\eta(t), \tag{4.1}$$

where $\eta(t)$ is the regular white noise characterized by eqs. (3.2) and (3.3). This is an O-U process which means the action is given by eq. (3.35), but the stochastic process is now denoted $r(t)$:

$$\mathcal{S}[r; t_0, s] = -\frac{1}{2\rho^2}\int_{t_0}^{s} \mathrm{d}t \left[\frac{\mathrm{d}r}{\mathrm{d}t} - k(\theta - r)\right]^2, \tag{4.2}$$

where we now specified the boundaries of the action. Since we ultimately would like to rederive the closed-form pricing formula for the ZCB in the Vasicek model as given in eqs. (2.45) to (2.47), we will be required to evaluate expectation values. Therefore we must construct a sensible (transition) probability distribution and to do this the boundary conditions for $r(t)$ must be specified. At the initial time t_0, the spot rate is fixed by the value observed in the market r_0. At the final time s, the spot rate can take any value. As before, in eq. (3.101), we only impose that $\frac{\mathrm{d}r}{\mathrm{d}t}\big|_{t=s} = 0$. Taking these boundary conditions into account, we can construct a probability distribution by first defining

$$Z = \int \mathcal{D}[r] e^{\mathcal{S}[r;t_0,s]}, \tag{4.3}$$

with integration measure denoted as

$$\int \mathcal{D}[r] = \prod_{t=t_0}^{s} \int_{-\infty}^{\infty} \mathrm{d}r(t)\,. \tag{4.4}$$

Note that Z is essentially defined as $\mathcal{Z}[0]$ and will function as a normalization constant. The probability distribution then uses the action in eq. (4.2)

$$\mathcal{P}_{1|1}(r, s \mid r_0, t_0) = \frac{1}{Z} \int \mathcal{D}[r] e^{\mathcal{S}[r;t_0,s]} \tag{4.5}$$

with boundary conditions

$$r(t_0) = r_0; \quad \left.\frac{\mathrm{d}r}{\mathrm{d}t}\right|_{t=s} = 0. \tag{4.6}$$

We can use the probability distribution in eq. (4.5) to evaluate the expectation value for the ZCB price.

4.1.1 Path Integral for the ZCB

The price at time t_0 of a ZCB with maturity s is given as

$$P(t_0, s) = \mathbb{E}\left[e^{-\int_{t_0}^{s} r(t)\mathrm{d}t} \middle| r(t_0) = r_0 \right]. \tag{4.7}$$

Using eq. (4.5), this can be cast in the following form

$$P(t_0, s) = \frac{1}{Z} \int \mathcal{D}[r] e^{\mathcal{S}} e^{-\int_{t_0}^{s} r(t)\mathrm{d}t}, \tag{4.8}$$

where $\mathcal{S} = \mathcal{S}[r; t_0, s]$ for further convenience. The ZCB price as given in eq. (2.45) can be rederived by evaluating the path integral in eq. (4.8). The derivation is based on [5, 46]. We start by introducing a new quantity $\tilde{\mathcal{S}} = \mathcal{S} - r$, such that

$$P(t_0, s) = \frac{1}{Z} \int \mathcal{D}[r] e^{\tilde{\mathcal{S}}}. \tag{4.9}$$

Performing the change of variables $r(t) = x(t) + \theta$ leaves the integration measure eq. (4.4) invariant and

$$\tilde{\mathcal{S}} = -\frac{1}{2\rho^2} \int_{t_0}^{s} \mathrm{d}t \left[\frac{\mathrm{d}x}{\mathrm{d}t} + kx \right]^2 - \int_{t_0}^{s} \mathrm{d}t\,(x + \theta). \tag{4.10}$$

Define now another new variable $v(t) = \frac{\mathrm{d}x}{\mathrm{d}t} + kx(t)$. Solving this for $x(t)$, we find

$$\begin{aligned} x(t) &= x(t_0) e^{-k(t-t_0)} + e^{-kt} \int_{t_0}^{t} \mathrm{d}t'\, v(t') e^{kt'} \\ &= (r_0 + \theta) e^{-k(t-t_0)} + e^{-kt} \int_{t_0}^{t} \mathrm{d}t'\, v(t') e^{kt'}, \end{aligned}$$

such that

$$\begin{aligned} \int_{t_0}^{s} \mathrm{d}t\, x(t) &= \int_{t_0}^{s} \mathrm{d}t \left[(r_0 + \theta) e^{-k(t-t_0)} + e^{-kt} \int_{t_0}^{t} \mathrm{d}t'\, v(t') e^{kt'} \right] \\ &= \tilde{B}(t_0, s)(r_0 + \theta) + \int_{t_0}^{s} \mathrm{d}t\, \tilde{B}(t, s) v(t), \end{aligned} \tag{4.11}$$

where we introduced $\tilde{B}(t, s)$ in the same way as in eq. (2.47). Plugging eqs. (4.10) and (4.11) and the expression for $v(t)$ back into eq. (4.8) allows us to write

$$P(t_0, s) = \frac{1}{Z} \int \mathcal{D}[r] \exp\left\{ -\frac{1}{2\rho^2} \int_{t_0}^{s} \mathrm{d}t \left[\frac{\mathrm{d}r}{\mathrm{d}t} - k(\theta - r) \right]^2 - \int_{t_0}^{s} \mathrm{d}t\, r \right\}$$

$$
\begin{aligned}
&= \frac{1}{Z}\int \mathcal{D}[x]\exp\left\{-\frac{1}{2\rho^2}\int_{t_0}^{s}\mathrm{d}t\left[\frac{\mathrm{d}x}{\mathrm{d}t}+kx\right]^2-\int_{t_0}^{s}\mathrm{d}t\,(x+\theta)\right\}\\
&= \frac{1}{Z}\int \mathcal{D}[v]\exp\left\{-\frac{1}{2\rho^2}\int \mathrm{d}\tau\, v(t)^2-\tilde{B}(t_0,s)(r_0+\theta)-\int_{t_0}^{s}\mathrm{d}t\,\tilde{B}(t,s)v(t)-\theta(s-t_0)\right\}\\
&= \exp\left[\tilde{B}(t_0,s)r_0-\theta(s-t_0)-\theta\tilde{B}(t_0,s)\right]\frac{1}{Z}\int \mathcal{D}[v]e^{-\frac{1}{2\rho^2}\int_{t_0}^{s}\mathrm{d}t\left[v^2+2\rho^2\tilde{B}(t,s)v\right]}.
\end{aligned} \tag{4.12}
$$

Note that by changing the variables from $x(t)$ to $v(t)$, the path integration measure changes with an extra factor $\mathcal{J}=\det\left(\frac{\mathrm{d}}{\mathrm{d}t}+k\right)$, the Jacobian. However, one can make the same change of integration variables in the partition function Z such that this factor cancels out. Completing the square yields

$$
\begin{aligned}
P(t_0,s) &= \exp\left[\tilde{B}(t_0,s)r_0-\theta(s-t_0)-\theta\tilde{B}(t_0,s)\right]\exp\left\{\frac{\rho^2}{2}\int_{t_0}^{s}\mathrm{d}t\,\tilde{B}(t,s)^2\right\}\\
&= \exp\left[\tilde{B}(t_0,s)r_0-\theta(s-t_0)-\theta\tilde{B}(t_0,s)\right]\\
&\quad\times\exp\left\{\frac{\rho^2}{2k^2}\left[(s-t_0)-2\frac{1-e^{-k(s-t_0)}}{k}+\frac{1-e^{-ka(s-t_0)}}{2k}\right]\right\}.
\end{aligned}
$$

After some rewriting and collecting terms, this can be written in the desired form

$$
\begin{aligned}
P(t_0,s) &= \exp\left\{\left(\theta-\frac{\rho^2}{2k^2}\right)\left[\tilde{B}(t_0,s)-(s-t_0)\right]-\frac{\rho^2}{4k}\tilde{B}(t_0,s)^2\right\}e^{-\tilde{B}(t_0,s)r_0}\\
&= A(t_0,s)e^{-\tilde{B}(t_0,s)r_0},
\end{aligned} \tag{4.13}
$$

which is exactly what we would expect from eq. (2.45).

We have thus successfully translated the Vasicek short-rate model to a path integral formulation in which we rederived the ZCB price. The use of this formalism can readily be extended to more complicated short-rate models (e.g. Hull-White, Ho-Lee) and even to the HJM framework. This is demonstrated in the next section.

4.2 The HJM Path Integral

In this section we will develop the path integral formulation for the (one-factor) HJM framework. We check the no-arbitrage condition for the model and derive a closed-form solution for the price of a ZCB option, following the approach suggested in [5]. We try to explain the derivations with as much detail as necessary.

In a similar fashion as for the Vasicek model, we can setup a path integral description for the HJM. The HJM model describes the dynamics of the (instantaneous) forward interest rates as given in eq. (2.48), or, equivalently

$$
\frac{\partial f(t,x)}{\partial t}=\alpha(t,x)+\sigma(t,x)\eta(t), \tag{4.14}
$$

where $0\le t_0\le t\le x\le s$ and $\eta(t)$ is again a white noise. This time, we obtain a probability measure for $\eta(t)$ over the time interval $[t_1,t_2]$ by discretizing said interval in N equal sub-intervals of length δt. This way, the variable t takes on values $n\delta t$ for $n=1,\dots N$ and $\eta(t)\to\eta(n)$. We can now construct a probability distribution analogously to eq. (3.23) as

$$
\mathcal{P}_{1|1}(\eta_1,t_1|\eta_2,t_2)=\prod_{n=1}^{N}\sqrt{\frac{\delta t}{2\pi}}\int_{-\infty}^{\infty}\mathrm{d}\eta(n)\,e^{-\frac{1}{2}\eta(n)^2\delta t}. \tag{4.15}
$$

When taking the continuum limit $\delta t \to 0$ and $N \to \infty$, we find that

$$\int \mathrm{d}\eta \to \int \mathcal{D}[\eta] \tag{4.16}$$

$$\mathcal{P}_{1|1}(\eta_1, t_1|\eta_2, t_2) \to \int \mathcal{D}[\eta] e^{\mathcal{S}}, \tag{4.17}$$

where

$$\int \mathrm{d}\eta = \prod_{n=1}^{N} \sqrt{\frac{\delta t}{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}\eta(n) \tag{4.18}$$

and

$$\mathcal{S} \equiv \mathcal{S}[\eta; t_1, t_2] = -\frac{1}{2} \int_{t_1}^{t_2} \eta(t)^2 \, \mathrm{d}t \tag{4.19}$$

We have thus constructed a probability measure for the noise $\eta(t)$. Note that here no extra normalization factor Z^{-1} is required as in the Vasicek case.

4.2.1 No Arbitrage Condition, Again

As we saw in section 2.3.2, the HJM framework is only free of arbitrage if a certain relation between the volatility structure $\sigma(t, x)$ and drift term $\alpha(t, x)$ holds. We will now how this relationship comes about in the path integral formulation.

Consider a ZCB with maturity s and price $P(t_0, s)$ at time $t_0 \leq s$ and $P(t_*, s)$ at time $t_0 \leq t_* \leq s$. In order for this ZCB to satisfy risk-neutral pricing under the no arbitrage assumption, as given in eq. (2.6), we require the expectation value of the price at time t_*, discounted by the spot rate to be equal to the price at time t_0, i.e.

$$P(t_0, s) = \mathbb{E}_{[t_0, t_*]} \left[e^{-\int_{t_0}^{t_*} r(t) \mathrm{d}t} P(t_*, s) \middle| \mathcal{F}_{t_0} \right], \tag{4.20}$$

where we have taken the expectation value over the time interval $[t_0, t_*]$. Using the probability measure eq. (4.17), this can be written as

$$P(t_0, s) = \int \mathcal{D}[\eta] e^{-\int_{t_0}^{t_*} r(t) \mathrm{d}t} P(t_*, s) e^{\mathcal{S}[\eta; t_0, t_*]}. \tag{4.21}$$

Using the relation between the ZCB and the forward interest rates in eq. (2.30), this becomes

$$= \int \mathcal{D}[\eta] e^{-X} e^{\mathcal{S}[\eta; t_0, t_*]}, \tag{4.22}$$

where we defined

$$X \equiv \int_{t_0}^{t_*} r(t) \, \mathrm{d}t + \int_{t_*}^{s} f(t_*, x) \, \mathrm{d}x \,. \tag{4.23}$$

Reminding ourselves of the general solutions for the HJM dynamics of the forward rates, given in eq. (2.49), and using $r(t) = f(t, t)$ yields

$$\begin{aligned} X = &\int_{t_0}^{t_*} \mathrm{d}x \left[f(t_0, x) + \int_{t_0}^{x} \alpha(t', x) \, \mathrm{d}t' + \int_{t_0}^{x} \sigma(t', x) \eta(t') \, \mathrm{d}t' \right] \\ &+ \int_{t_*}^{s} \mathrm{d}x \left[f(t_0, x) + \int_{t_0}^{t_*} \alpha(t', x) \, \mathrm{d}t' + \int_{t_0}^{t_*} \sigma(t', x) \eta(t') \, \mathrm{d}t' \right]. \end{aligned}$$

Notice now how we can write

$$\int_{t_0}^{t_*} \mathrm{d}x \int_{t_0}^{x} \mathrm{d}t' \, \alpha(t',x) + \int_{t_*}^{s} \mathrm{d}x \int_{t_0}^{t_*} \mathrm{d}t' \, \alpha(t',x) = \int_{t_0}^{t_*} \mathrm{d}t' \left[\int_{t'}^{t_*} \mathrm{d}x \, \alpha(t',x) + \int_{t_*}^{s} \mathrm{d}x \, \alpha(t',x) \right]$$
$$= \int_{t_0}^{t_*} \mathrm{d}t' \int_{t'}^{s} \mathrm{d}x \, \alpha(t',x).$$

An equivalent reasoning can be applied to the integrals of $\sigma(t',x)\eta(t')$. These integrals can be represented somewhat more graphically, as shown in fig. 4.1, by introducing the following regions of the (x,t)-plane

$$\Delta_0 = \left\{ (x,t) \in \mathbb{R}^2 \middle| t_0 \le t \le x \text{ and } t_0 \le x \le t_* \right\} \tag{4.24}$$
$$\mathcal{R} = \left\{ (x,t) \in \mathbb{R}^2 \middle| t_0 \le t \le t_* \text{ and } t_* \le x \le s \right\} \tag{4.25}$$
$$\mathcal{T} = \Delta_0 \oplus \mathcal{R}. \tag{4.26}$$

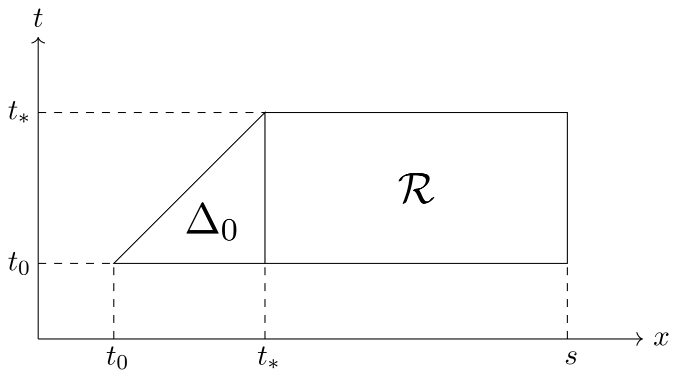


Figure 4.1: Graphical representation of the relevant areas Δ_0 and $\mathcal{R}$ in the (x,t)-plane.

Hence,

$$X = \int_{t_0}^{s} f(t_0,x) \, \mathrm{d}x + \int_{\mathcal{T}} \left[\alpha(t,x) + \sigma(t,x)\eta(t) \right], \tag{4.27}$$

where denoted

$$\int_{\mathcal{T}} = \int_{t_0}^{t_*} \mathrm{d}t' \int_{t'}^{s} \mathrm{d}x \, .$$

Plugging eq. (4.27) back into eq. (4.22) yields:

$$P(t_0,s) = \int \mathcal{D}[\eta] \exp\left[- \int_{t_0}^{s} f(t_0,x) \, \mathrm{d}x - \int_{\mathcal{T}} \alpha(t,x) - \int_{\mathcal{T}} \sigma(t,x)\eta(t) \right] e^{\mathcal{S}[\eta,t_0,t_*]}$$
$$= P(t_0,s) e^{-\int_{\mathcal{T}} \alpha(t,x)} \int \mathcal{D}[\eta] e^{-\int_{\mathcal{T}} \sigma(t,x)\eta(t)} e^{\mathcal{S}[\eta;t_0,t_*]}, \tag{4.28}$$

where we again used eq. (2.30) to rewrite the exponential as the price of a ZCB.

To aid us in the evaluation of the path integral in eq. (4.28) we invoke the moment generating functional, given as

$$\mathcal{Z}[J;t_1,t_2] = \int \mathcal{D}[\eta] e^{\mathcal{S}[\eta;t_1,t_2]} e^{\int_{t_1}^{t_2} J(t)\eta(t) \mathrm{d}t}, \tag{4.29}$$

for some source function $J(t)$. We also saw in section 3.3.3 that, by means of completing the square[1] with the substitution $\tilde{\eta}(t) = \eta(t) - J(t)$, the exponential in the integrand can be rewritten

[1] Notice that before we had a factor $\triangle^{-1}$ which is now just a factor 1.

as

$$\mathcal{S}[\eta; t_1, t_2] + \int_{t_1}^{t_2} J(t)\eta(t)\,\mathrm{d}t = \mathcal{S}[\tilde{\eta}, t_1, t_2] + \frac{1}{2}J(t)^2.$$

Since the probability distribution in eq. (4.17) is normalized, we have

$$\mathcal{Z}[J=0; t_1, t_2] = \int \mathcal{D}[\eta] e^{\mathcal{S}[\eta, t_1, t_2]} = \int \mathcal{D}[\tilde{\eta}] e^{\mathcal{S}[\tilde{\eta}, t_1, t_2]} = 1 \tag{4.30}$$

and thus

$$\mathcal{Z}[J; t_1, t_2] = e^{\frac{1}{2}\int_{t_1}^{t_2} J(t)^2 \mathrm{d}t}. \tag{4.31}$$

Looking back at eq. (4.28), we can use eq. (4.31) and define the integrated volatility $\Sigma(t) \equiv -\int_t^s \mathrm{d}x\,\sigma(t,x)$ as the source function to write

$$\begin{aligned} e^{\int_{\mathcal{T}} \alpha(t,x)} &= \int \mathcal{D}[\eta] e^{\mathcal{S}[\eta, t_0, t_*]} e^{\int_{t_0}^{t_*} \mathrm{d}t \eta(t) \Sigma(t)} \\ &= \mathcal{Z}\left[\Sigma(t); t_0, t_*\right] \\ &= e^{\frac{1}{2}\int_{t_0}^{t_*} \Sigma(t)^2 \mathrm{d}t}. \end{aligned} \tag{4.32}$$

Hence

$$\begin{aligned} \int_{\mathcal{T}} \alpha(t,x) &= \frac{1}{2}\int_{t_0}^{t_*} \left[\int_t^s \mathrm{d}x\,\sigma(t,x)\right]^2 \mathrm{d}t \\ \int_t^s \mathrm{d}x\,\alpha(t,x) &= \frac{1}{2}\left[\int_t^s \mathrm{d}x\,\sigma(t,x)\right]^2 \end{aligned}$$

or, equivalently,

$$\alpha(t,x) = \sigma(t,x)\int_t^x \mathrm{d}y\,\sigma(t,y). \tag{4.33}$$

Note that eqs. (2.55) and (4.33) yield exactly the same condition on the risk-free drift $\alpha(t,x)$. We have thus successfully rederived the no-arbitrage condition on drift term $\alpha(t,x)$ in the path integral formulation of the HJM model.

4.2.2 ZCB Option, Again

Now that we have established that the no-arbitrage condition for the HJM model can be rederived in the path integral formulation, it is only natural to ask whether the ZCB option can be rederived as well. Let us therefore consider the same call option as in section 2.3.2, i.e. a European call option of strike K and expiry t_* on a ZCB of maturity $s \geq t_*$. We saw that the payoff at expiry time t_* was given by

$$\mathrm{ZCB}C(t_*; K, t_*, s) = \left[P(t_*, s) - K\right]^+, \tag{4.34}$$

which means that the risk-neutral price at time $t_0 \leq t_*$ is equal to

$$\mathrm{ZCB}C(t_0; K, t_*, s) = \mathbb{E}_{[t_0,t_*]}\left[e^{-\int_{t_0}^{t_*} \mathrm{d}t r(t)}(P(t_*, T) - K)^+ \middle| \mathcal{F}_{t_0}\right]. \tag{4.35}$$

We will first mold the payoff in eq. (4.34) into a form that fits our path integral description. Using eq. (2.30) and the Dirac delta, the payoff can be written as

$$\left[P(t_*, s) - K\right]^+ = \int_{-\infty}^{\infty} \mathrm{d}G\,\delta\left(G + \int_{t_*}^s \mathrm{d}x\,f(t_*, x)\right)\left(e^G - K\right)^+,$$

where we introduced G as an extra integration variable. Writing the Dirac delta as the integral over a complex exponential, we find

$$= \int_{-\infty}^{\infty} \mathrm{d}G \int_{-\infty}^{\infty} \frac{\mathrm{d}p}{2\pi} e^{\mathrm{i}p\left[G+\int_{t_*}^{s} \mathrm{d}x f(t_*,x)\right]} \left(e^G - K\right)^+ . \tag{4.36}$$

Plugging this form of the payoff back into the expectation value in eq. (4.35), we can pull out the integration over G and p. This means, the risk-neutral price at time $t_0 \leq t_*$ becomes

$$\mathrm{ZCBC}(t_0; K, t_*, s) = \int_{-\infty}^{\infty} \mathrm{d}G\, \Psi(G, t_*, T)(e^G - K)^+ \tag{4.37}$$

with

$$\Psi(G, t_*, s) = \int_{-\infty}^{\infty} \frac{\mathrm{d}p}{2\pi} \mathbb{E}_{[t_0,t_*]} \left[e^{-\int_{t_0}^{t_*} \mathrm{d}t f(t,t)} e^{\mathrm{i}p\left[G+\int_{t_*}^{s} \mathrm{d}x f(t_*,x)\right]} \right],$$

where we used $r(t) = f(t,t)$. Finally rewriting the expectation value using the probability measure eq. (4.17) yields

$$\Psi(G, t_*, s) = \int \mathcal{D}[\eta] \int_{-\infty}^{\infty} \frac{\mathrm{d}p}{2\pi} e^{\mathcal{S}[\eta;t_0,t_*]} e^{-\int_{t_0}^{t_*} \mathrm{d}t f(t,t)} e^{\mathrm{i}pG} e^{\mathrm{i}p \int_{t_*}^{s} \mathrm{d}x f(t_*,x)} . \tag{4.38}$$

Plugging this back into eq. (4.35) gives

$$\mathrm{ZCBC}(t_0; K, t_*, s) = \int \mathcal{D}[\eta] \int_{-\infty}^{\infty} \mathrm{d}G \frac{\mathrm{d}p}{2\pi} e^{\mathcal{S}[\eta;t_0,t_*]} e^{-\int_{t_0}^{t_*} \mathrm{d}t f(t,t)} e^{\mathrm{i}pG} e^{\mathrm{i}p \int_{t_*}^{s} \mathrm{d}x f(t_*,x)} \left(e^G - K\right)^+ .$$

This is the price of the European call option on a ZCB translated to the path integral formulation. The challenge lies now in evaluating this path integral. We can start by simplifying the expression for $\Psi(G, t_*, s)$ in eq. (4.38). This can be done from an inspection of the exponentials and applying eq. (2.49), which shows

$$\begin{aligned} \int_{t_*}^{s} \mathrm{d}x\, f(t_*, x) &= \int_{t_*}^{s} \left[f(t_0, x) + \int_{t_0}^{t_*} \mathrm{d}t\, \alpha(t,x) + \int_{t_0}^{t_*} \mathrm{d}t\, \sigma(t,x)\eta(t) \right] \\ &= \int_{t_*}^{s} \mathrm{d}x\, f(t_0, x) + \int_{\mathcal{R}} \alpha(t,x) + \int_{\mathcal{R}} \sigma(t,x)\eta(t), \end{aligned}$$

such that we can define

$$\Lambda_0 \equiv G + \int_{t_*}^{s} \mathrm{d}x\, f(t_0, x) + \int_{\mathcal{R}} \alpha(t,x), \tag{4.39}$$

where $\int_{\mathcal{R}}$ again denotes an integration over the entire region $\mathcal{R}$ defined in eq. (4.25). This way, the expression for $\Psi(G, t_*, s)$ in eq. (4.38) can be rewritten as

$$\Psi(G, t_*, s) = \int \mathcal{D}[\eta] \int_{-\infty}^{\infty} \frac{\mathrm{d}p}{2\pi} e^{\mathcal{S}[\eta;t_0,t_*]} e^{-\int_{t_0}^{t_*} \mathrm{d}t f(t,t)} e^{\mathrm{i}p\Lambda_0} e^{\mathrm{i}p \int_{\mathcal{R}} \sigma(t,x)\eta(t)} . \tag{4.40}$$

In a similar fashion, we can write

$$\begin{aligned} \int_{t_0}^{t_*} \mathrm{d}x\, f(x, x) &= \int_{t_0}^{t_*} \mathrm{d}x \left[f(t_0, x) + \int_{t_0}^{x} \mathrm{d}t\, \alpha(t,x) + \int_{t_0}^{x} \mathrm{d}t\, \sigma(t,x)\eta(t) \right] \\ &= \int_{t_0}^{t_*} \mathrm{d}x\, f(t_0, x) + \int_{\Delta_0} \alpha(t,x) + \int_{\Delta_0} \sigma(t,x)\eta(t). \end{aligned}$$

Plugging this back into eq. (4.40) yields

$$\Psi(G,t_*,s) = \int \mathcal{D}[\eta] \int_{-\infty}^{\infty} \frac{\mathrm{d}p}{2\pi} e^{\mathcal{S}[\eta;t_0,t_*]} e^{-\int_{t_0}^{t_*}\mathrm{d}x f(t_0,x)} e^{-\int_{\Delta_0}\alpha(t,x) - \int_{\Delta_0}\sigma(t,x)\eta(t)} e^{\mathrm{i}p\Lambda_0} e^{\mathrm{i}p\int_{\mathcal{R}}\sigma(t,x)\eta(t)} \tag{4.41}$$

$$= P(t_0,t_*) \int_{-\infty}^{\infty} \frac{\mathrm{d}p}{2\pi} e^{\Lambda} e^{\Lambda_0}, \tag{4.42}$$

where we again used eq. (2.30) to write $P(t_0,t_*)$ as a pre-factor and defined

$$e^{\Lambda} \equiv e^{-\int_{\Delta_0}\alpha(t,x)} \int \mathcal{D}[\eta] e^{\mathcal{S}[\eta;t_0,t_*]} e^{-\int_{\Delta_0}\sigma(t,x)\eta(t) + \mathrm{i}p\int_{\mathcal{R}}\sigma(t,x)\eta(t)}. \tag{4.43}$$

Note that from eq. (4.43) it becomes clear that price of the call option depends on the interactions between the two domains, Δ_0 and $\mathcal{R}$, which were defined in eq. (4.24) and eq. (4.25) respectively. We will later (section 5.4) find that in the field theory formulation, when using a different numeraire, this derivation considerably simplifies and only the region $\mathcal{R}$ will be required. But for now, the goal is is to obtain an expression for Λ. This can be done by using eq. (4.31), but first of all we have to rearrange some terms in a more convenient form:

$$\begin{aligned} -\int_{\Delta_0}\sigma(t,x)\eta(t) + \mathrm{i}p\int_{\mathcal{R}}\sigma(t,x)\eta(t) &= -\int_{t_0}^{t_*}\mathrm{d}t\int_{t}^{t_*}\mathrm{d}x\,\sigma(t,x)\eta(t) + \mathrm{i}p\int_{t_0}^{t_*}\mathrm{d}t\int_{t_*}^{s}\mathrm{d}x\,\sigma(t,x)\eta(t) \\ &= \int_{t_0}^{t_*}\mathrm{d}t\,\eta(t)\left[-\int_{t}^{t_*}\mathrm{d}x\,\sigma(t,x) + \mathrm{i}p\int_{t_*}^{s}\mathrm{d}x\,\sigma(,x)\right]. \end{aligned}$$

By inspection, we can identify the factor in the square-brackets as some source function. Using eq. (4.31), this allows us to write down

$$\int \mathcal{D}[\eta] e^{\mathcal{S}[W,t_0,t_*]} e^{-\int_{\Delta_0}\sigma(t,x)\eta(t) + \mathrm{i}p\int_{\mathcal{R}}\sigma(t,x)\eta(t)} = \exp\left\{\frac{1}{2}\int_{t_0}^{t_*}\mathrm{d}t\left[-\int_{t}^{t_*}\mathrm{d}x\,\sigma(t,x) + \mathrm{i}p\int_{t_*}^{s}\mathrm{d}x\,\sigma(t,x)\right]^2\right\}. \tag{4.44}$$

Using the condition on the risk-neutral drift given in eq. (4.33), one can show that, after some serious rewriting [5]

$$\Lambda = -\frac{1}{2}q^2p^2, \tag{4.45}$$

where we defined

$$q^2 \equiv \int_{t_0}^{t_*}\mathrm{d}t\left[\int_{t_*}^{s}\mathrm{d}x\,\sigma(t,x)\right]^2 \tag{4.46}$$

$$= 2\int_{\mathcal{R}}\alpha(t,x). \tag{4.47}$$

Equation (4.47) will not be proven here but we will derive this identity in section 5.5. Plugging this expression for Λ back into eq. (4.42) yields

$$\Psi(G,t_*,s) = P(t_0,t_*)\int_{-\infty}^{\infty}\frac{\mathrm{d}p}{2\pi} e^{-\frac{q^2}{2}p^2} e^{\mathrm{i}p\Lambda_0}.$$

By completing the square, one finds

$$\Psi(G,t_*,s) = P(t_0,t_*)\frac{e^{-\frac{\Lambda_0^2}{2q^2}}}{\sqrt{2\pi q^2}}. \tag{4.48}$$

Now that we have successfully derived an expression for $\Psi(G, t_*, s)$, which can be plugged into eq. (4.37) to find

$$\begin{aligned}\mathrm{ZCBC}(t_0; K, t_*, s) &= \int_{-\infty}^{\infty} \mathrm{d}G\, P(t_0, t_*) \frac{e^{-\frac{\Lambda_0^2}{2q^2}}}{\sqrt{2\pi q^2}} (e^G - K)^+ \\ &= \frac{P(t_0, t_*)}{\sqrt{2\pi q^2}} \int_{-\infty}^{\infty} (e^G - K)^+ \exp\left\{ -\frac{1}{2q^2} \left[G + \int_{t_*}^{s} \mathrm{d}x\, f(t_0, x) + \int_{\mathcal{R}} \alpha(t, x) \right]^2 \right\},\end{aligned}$$

where we used the original expression for Λ_0 as defined in eq. (4.39). This can be written as

$$\mathrm{ZCBC}(t_0; K, t_*, s) = \frac{P(t_0, t_*)}{\sqrt{2\pi q^2}} \left[\int_{\ln K}^{\infty} \mathrm{d}G\, e^{-\frac{1}{2q^2}(G+\lambda)^2 + G} - K \int_{\ln K}^{\infty} \mathrm{d}G\, e^{-\frac{1}{2q^2}(G+\lambda)^2} \right],$$

where we introduced $\lambda \equiv q^2/2 + \int_{t_*}^{s} \mathrm{d}x\, f(t_0, x)$ for simplicity's sake. Then, by means of completing the square in the first integral, this becomes

$$\mathrm{ZCBC}(t_0; K, t_*, s) = \frac{P(t_0, t_*)}{\sqrt{2\pi q^2}} \left[e^{\frac{q^2}{2} - \lambda} \int_{\ln K}^{\infty} \mathrm{d}G\, e^{-\frac{1}{2q^2}(G+\lambda-q^2)^2} - K \int_{\ln K}^{\infty} \mathrm{d}G\, e^{-\frac{1}{2q^2}(G+\lambda)^2} \right].$$

We can now perform a change of variables in which we set $G_1 = (G + \lambda - q^2)$ and $G_2 = (G + \lambda)/q$ such that

$$\mathrm{ZCBC}(t_0; K, t_*, s) = \frac{P(t_0, t_*)}{\sqrt{2\pi q^2}} \left[e^{\frac{q^2}{2} - \lambda} q \int_{d_1}^{\infty} \mathrm{d}G_1\, e^{-\frac{G_1^2}{2}} - Kq \int_{d_2}^{\infty} \mathrm{d}G_2\, e^{-\frac{G_2^2}{2}} \right], \tag{4.49}$$

where

$$d_1 = \frac{\ln K + \lambda - q^2}{q} \quad \text{and} \quad d_2 = \frac{\ln K + \lambda}{q}. \tag{4.50}$$

Resubstiting $\lambda = q^2/2 + \int_{t_*}^{s} \mathrm{d}x\, f(t_0, x)$ yields

$$\mathrm{ZCBC}(t_0; K, t_*, s) = P(t_0, t_*) \left[e^{-\int_{t_*}^{s} \mathrm{d}x f(t_0, x)} \Phi(-d_1) - K\Phi(-d_2) \right]. \tag{4.51}$$

By now defining

$$F(t_0, t_*, s) = F \equiv e^{-\int_{t_*}^{s} \mathrm{d}x f(t_0, x)} = \frac{P(t_0, s)}{P(t_0, t_*)},$$

one finds the following expressions for integration boundaries d_1 and d_2:

$$\begin{aligned} d_1 &= \frac{1}{q}\left[\ln\frac{K}{F} - \frac{q^2}{2} \right] \\ d_2 &= \frac{1}{q}\left[\ln\frac{K}{F} + \frac{q^2}{2} \right], \end{aligned}$$

such that we can define

$$d_{\pm} = \frac{1}{q}\left[\ln\frac{F}{K} \pm \frac{q^2}{2} \right], \tag{4.52}$$

where $d_1 = -d_+$ and $d_2 = -d_-$.

This ultimately yields the following price, at time t_0, for an option with strike K and expiry t_* on a ZCB of maturity s:

$$\boxed{\mathrm{ZCBC}(t_0; K, t_*, s) = P(t_0, t_*) \left[F\Phi(d_+) - K\Phi(d_-) \right].} \tag{4.53}$$

One can now easily check that for constant (and positive) volatility $\sigma(t,x)=\sigma$, eq. (4.46) yields the following integration boundaries

$$d_\pm = \frac{\ln\frac{F}{K} \pm \frac{\sigma^2}{2}(T-t*)^2(t_*-t_0)}{\sigma(T-t_*)\sqrt{t_*-t_0}}$$

and thus eq. (4.53) reduces to exactly the same price as documented in section 2.3.2 in eq. (2.70).

To conclude, we again point out that the derivation above is in fact nothing more than a mere translation to the path integral formulation for SDEs, however, in the "classical" derivation of the ZCB call option price, as discussed in section 2.3.2, we started from the assumption of constant volatility[2], while in the "path-integral" derivation no assumptions on the form of $\sigma(t,x)$ were made, except for it being deterministic. This already shows one important strength that the path-integral approach may provide: the volatility structure can be any well-behaved deterministic function and can be plugged into eq. (4.53) directly. We will see that a further extension to a field theory of interest rates provides even more useful tools and functionalities.

[2]One could also have done the derivation for any other (deterministic) form of the volatility structure. Nevertheless, one always should make an assumption on the volatility.

Chapter 5

Field Theory for Interest Rates

With everything that is developed up until now, we are well equipped to take our formulation to the next step and explore how a field theory formalism for the forward interest rates may be constructed. This is quite the interesting approach since, in contrast to stocks which are described by only one *degree of freedom*, the stochastic interest rate curve requires an infinite number of degrees of freedom to be described. The statistical- and quantum field theories discussed in sections 3.2 and 3.3 provide the tools to accurately describe the dynamics of such multi-dimensional objects. However, one might ask whether such a field theory description can be omitted by employing an *multi-factor* HJM model. After all, we have only been discussing the *one-factor* version of the model.

First of all, for any specified volatility structure, there does in general not exist a possible construction of an N-factor model (for $N \in \mathbb{N}_0$) that can consistently capture the entire term structure. This means that the model's parameters would need to be continuously recalibrated [29]. Furthermore, the total amount of such parameters that require continuous recalibration increases with N, or, even more so, increases as $n \times N$, where n is the number of parameters used to specify the volatility structure. We will see that in the (simplest) field theory model, only one additional parameters is required to translate the one-factor model to a field theory model. This parameter will describe the *rigidity* of the forward rate curve and will later turn out to influence the correlation between forward rates of different maturities.

Secondly, finite-factor models allow for unreasonable hedging possibilities. For example, in a three-factor HJM model, one could theoretically hedge a bond with a time to maturity of $T = 30$ years by constructing an appropriate portfolio[1] consisting of bonds with $T = 1, 2$ and 3 months. This way one could eliminate the majority of the interest rate risk exposure, which is of course not possible in reality [29].

Thirdly, multi-factor models have difficulties incorporating non-trivial correlations between forward rates of different maturities [29]. As was mentioned earlier, in the one-factor model for example, all points on the curve $s \mapsto f(t, s)$ are exactly correlated. It will be shown that, when deploying the mathematical tools from QFT, the non-trivial correlation structure can embodied in a parsimonious manner which allows for relatively straightforward analytical studies as well as the development of calibration procedures (more on the model calibration in chapter 6).

The term structure of interest rates has been studied as a stochastic string in [16, 17, 49, 52]. The theory of (Gaussian) random fields was used to describe the forward interest rates and introduce a correlation between the forward rates across different maturities [22, 29, 39, 40]. However, none of these references utilized the many practical tools QFT has to offer. The approach that will

[1] Note that this portfolio would not merely be constructed by buying all such bonds but rather be a complicated combination of buying and selling specific amounts of certain bonds.

be maintained here is one with the path integrals from QFT in mind and will be based on the framework established in [5]. All the derivations are, when necessary, worked out in great detail.

The (one-factor) HJM, as given in eq. (4.14), will be extended to a two-dimensional field theory by including extra independent fluctuations for the forward rates at each different maturity. We will restrict ourselves to the simplest case, which is the free field, analogous to the field discussed in section 3.3.3. However, we will see that there is an intricate connection with the harmonic oscillator from section 3.3.4 as well.

In this chapter, the Lagrangian and action for the theory are constructed in a qualitative and phenomenological manner. This allows for the definition of the moment generating functional and the calculation of the correlation functions. The corresponding propagator to the theory is shown to be equivalent to that of the harmonic oscillator. The "classical" limit of the theory is discussed. The condition required for the theory to be free of arbitrage is derived by means of path integration. To conclude, we study the change of a numeraire, as explained in section 2.2.3, in the context of the field theory, which will enable us to derive the price of a ZCB option for one last time in a fairly straightforward way.

5.1 The Field Lagrangian & Action

In our earlier discussion of the HJM model, in section 2.3.2, the stochastic nature of the forward interest rates at time t was governed by a random *shock* $\eta(t)$ which was applied to the entire curve $x \mapsto f(t,x)$. As stated before, we now aim to generalize this evolution of the forward rate curve by letting $f(t,x)$ fluctuate in both x and t independently. That is, at any point in time t, there are infinitely many maturities (or future times) x and they all behave in a stochastic and (highly but not exactly) correlated manner. So, for any instant t, an infinite amount of independent variables are needed to correctly describe the curve's random behavior.

In short, we will consider $f(t,x)$ to be an independent stochastic variable for all x and t and keep both x and t continuous for the sake of notational simplicity.

For any moment in time t, we find in the market the forward interest rates for a certain duration T_{FR}. The forward interest rates at this moment in time t, i.e. $f(t,x)$, exist for all future times x for which $t \leq x \leq t + T_{FR}$. To make this all somewhat more concrete, we will consider the forward rates' evolution starting from an initial time t_i to a final future time t_f. As per definition of the forward interest rates, $t \leq x$ is always true. The field $f(t,x)$ is thus defined on the domain[2] $\mathcal{F}$ given by

$$\mathcal{F} = \left\{ (x,t) \in \mathbb{R}^2 \middle| t_i \leq t \leq t_f \text{ and } t \leq x \leq t + T_{FR} \right\}. \tag{5.1}$$

A graphical representation of $\mathcal{F}$ can be seen in fig. 5.1. The integration variables $\eta(t)$ for the path integral formulation of the HJM model in section 4.2 only took values for $t \in [t_0, t_*]$, the integration variables for the field theory, being $f(t,x)$, can take values for all $(x,t) \in \mathcal{F}$.

In order for us to develop a field theory description of the forward interest rates, we must construct a Lagrangian $\mathcal{L}$. This Lagrangian must contain a kinetic term, which expresses the time evolution as t changes, and a rigidity term, which constraints the shape of the forward rates in the future time x-direction. Given to form of the regular SDE that governs the forward rate dynamics in the HJM model, in eq. (4.14), and the result in eq. (3.49) from section 3.2, a well-educated guess for the kinetic term of the Lagrangian might be

$$\mathcal{L}_{\text{kin}}[f] = -\frac{1}{2}\left(\frac{\frac{\partial f}{\partial t} - \alpha}{\sigma}\right)^2. \tag{5.2}$$

[2] Here $\mathcal{F}$ denotes a region in the (x,t)-plane and has nothing to do with the aforementioned filtration $\mathcal{F}_t$.

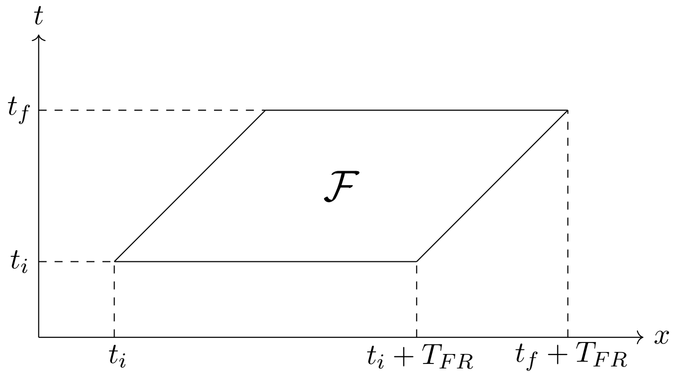


Figure 5.1: Graphical representation of the domain $\mathcal{F}$.

Note that the dependence on (t, x) is left implicit in eq. (5.2) to ease notation. As for the rigidity term, we look for a possible form in the field theory description of a stochastic string. However a simple term of the form $\left(\frac{\partial f}{\partial x}\right)^2$, to mimic the "tension term" in a string system can not be used as it would violate the no-arbitrage condition for the field theory [5, 14]. The simplest form for the rigidity term that can be used to constrain the possible fluctuations in the x-directions while also maintaining the no-arbitrage condition is of the form $\left(\frac{\partial^2 f}{\partial t \partial x}\right)^2$ [14], such that the rigidity terms reads

$$\mathcal{L}_{\text{rig}}[f] = -\frac{1}{2\mu^2}\left[\frac{\partial}{\partial x}\left(\frac{\frac{\partial f}{\partial t} - \alpha}{\sigma}\right)\right]^2, \tag{5.3}$$

where we introduced a new parameter μ, which has units year^{-1}. The rigidity of the forward rates is measured as $1/\mu^2$ and determines the strength of the fluctuations in the x-direction. We will later show that in the limit $\mu \to 0$, i.e. infinite rigidity, the field theory reduces to the one-factor HJM model. For now, using eqs. (5.2) and (5.3), we can write down the simplest form of the Lagrangian

$$\mathcal{L}[f] = -\frac{1}{2}\left(\frac{\frac{\partial f}{\partial t} - \alpha}{\sigma}\right)^2 - \frac{1}{2\mu^2}\left[\frac{\partial}{\partial x}\left(\frac{\frac{\partial f}{\partial t} - \alpha}{\sigma}\right)\right]^2, \tag{5.4}$$

where f, α and σ all are implicitly dependent on (t, x) and $f(t, x) \in \mathbb{R} \cup \{\pm\infty\}$. The action corresponding to this Lagrangian is

$$\mathcal{S}[f] = \int_{t_i}^{t_f} \mathrm{d}t \int_{t}^{t+T_{FR}} \mathrm{d}x\, \mathcal{L}[f] = \int_{\mathcal{F}} \mathcal{L}[f], \tag{5.5}$$

where the integral on the RHS of the final equation again indicates an integration over the entire (finite) domain $\mathcal{F}$. Note that here, wedo no longer specify the integration boundaries of the action. From now on, these will be clear from context.

Of course, to complete the construction of this field theory, we must specify the boundary conditions. We let the boundaries in the t-direction, i.e. $t = t_i$ and $t = t_f$, be specified by an initial and final forward rate curve, i.e. $f(t_i, x)$ and $f(t_f, x)$ with $t_i \leq x \leq t_i + T_{FR}$ and $t_f \leq x \leq t_f + T_{FR}$ respectively.

As for the boundaries in the x-direction, we will require the surface terms of the action to vanish. More explicitly, we impose

$$\frac{\partial}{\partial x}\left(\frac{\frac{\partial f}{\partial t} - \alpha}{\sigma}\right) = 0 \tag{5.6}$$

for $t_i \leq t \leq t_f$ and $x = t$ or $x = t + T_{FR}$.

Using the boundary condition in eq. (5.6), we can perform an integration by parts over x for the second term of the Lagrangian to find

$$\begin{aligned}
\mathcal{S}[f] &= -\frac{1}{2}\int_{t_i}^{t_f} \mathrm{d}t \int_{t}^{t+T_{FR}} \mathrm{d}x \left\{ \left(\frac{\frac{\partial f}{\partial t} - \alpha}{\sigma}\right)^2 + \frac{1}{\mu^2}\left[\frac{\partial}{\partial x}\left(\frac{\frac{\partial f}{\partial t} - \alpha}{\sigma}\right)\right]^2 \right\} \\
&= -\frac{1}{2}\int_{t_i}^{t_f} \mathrm{d}t \left\{ \int_{t}^{t+T_{FR}} \left(\frac{\frac{\partial f}{\partial t} - \alpha}{\sigma}\right)^2 \mathrm{d}x + \frac{1}{\mu^2}\int_{t}^{t+T_{FR}} \left[\frac{\partial}{\partial x}\left(\frac{\frac{\partial f}{\partial t} - \alpha}{\sigma}\right)\right]^2 \mathrm{d}x \right\} \\
&= -\frac{1}{2}\int_{t_i}^{t_f} \mathrm{d}t \left\{ \int_{t}^{t+T_{FR}} \left(\frac{\frac{\partial f}{\partial t} - \alpha}{\sigma}\right)^2 \mathrm{d}x + \frac{1}{\mu^2}\left[\left(\frac{\frac{\partial f}{\partial t} - \alpha}{\sigma}\right)\frac{\partial}{\partial x}\left(\frac{\frac{\partial f}{\partial t} - \alpha}{\sigma}\right)\Bigg|_{x=t}^{x=t+T_{FR}} \right.\right. \\
&\qquad \left.\left. - \int_{t}^{t+T_{FR}} \left(\frac{\frac{\partial f}{\partial t} - \alpha}{\sigma}\right)\frac{\partial^2}{\partial x^2}\left(\frac{\frac{\partial f}{\partial t} - \alpha}{\sigma}\right)\mathrm{d}x \right]\right\} \\
&= -\frac{1}{2}\int_{t_i}^{t_f} \mathrm{d}t \int_{t}^{t+T_{FR}} \mathrm{d}x \left(\frac{\frac{\partial f}{\partial t} - \alpha}{\sigma}\right)\left(1 - \frac{1}{\mu^2}\frac{\partial^2}{\partial x^2}\right)\left(\frac{\frac{\partial f}{\partial t} - \alpha}{\sigma}\right).
\end{aligned} \tag{5.7}$$

We can define the field theory entirely by the partition functional

$$Z \equiv \mathcal{Z}[0] = \int \mathcal{D}[f] e^{\mathcal{S}[f]}, \tag{5.8}$$

where

$$\int \mathcal{D}[f] = \prod_{(x,t)\in\mathcal{F}} \int_{-\infty}^{\infty} \mathrm{d}f(t,x)\,. \tag{5.9}$$

Note that here we have again denoted $Z = \mathcal{Z}[0]$, where $\mathcal{Z}[J]$ is of course the moment generating functional. Since we will soon introduce a convenient change of variables for the field theory and will mostly be working with the newly obtained field, we do not discuss the generating functional and correlation functions here. The partition function in eq. (5.8) represents a sum over all possible configurations of $f(t,x)$ with $(x,t) \in \mathcal{F}$, each of which is *weighted* by $e^{\mathcal{S}[f]}/Z$.

In principle, one can define forward rates for arbitrary far future times, i.e. $t_f \to \infty$. That is, the time at which a certain contract for a given forward rate is disclosed can be taken to be the present or any moment in the future. The region $\mathcal{F}$ in which the field theory is defined is then readily extended to the (semi-)infinite region shown in fig. 5.2.

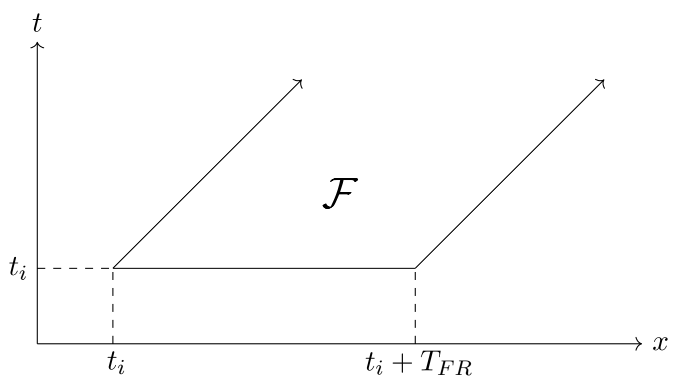


Figure 5.2: Graphical representation of the domain $\mathcal{F}$ in the limit $t_f \to \infty$.

When we look back at how the path integral for the HJM model was treated in section 4.2, we did so by considering the probability measure for the Gaussian noise $\eta(t)$. For computational (and notational) purposes, we want to introduce a stochastic field $\mathcal{A}(t,x)$ that will govern the forward rate dynamics as

$$\frac{\partial f(t,x)}{\partial t} = \alpha(t,x) + \sigma(t,x)\mathcal{A}(t,x). \tag{5.10}$$

This implies

$$f(t,x) = f(t_i,x) + \int_{t_i}^{t} \mathrm{d}t'\, \alpha(t',x) + \int_{t_i}^{t} \mathrm{d}t'\, \sigma(t',x)\mathcal{A}(t',x). \tag{5.11}$$

The field $\mathcal{A}(t,x)$ can be considered as the *drift-less velocity field* for the forward interest rates $f(t,x)$ [5]. From now on, most derivations, such as the propagator, correlation functions,... will be performed for the "velocity" field $\mathcal{A}(t,x)$. To do this we must consider a change of variables which, based on eq. (5.11), reads

$$F : C^\infty(\mathbb{R}^2) \to C^\infty(\mathbb{R}^2) : f(t,x) \mapsto f(t_i,x) + \int_{t_i}^{t} \mathrm{d}t'\, \alpha(t',x) + \int_{t_i}^{t} \mathrm{d}t'\, \sigma(t',x)\mathcal{A}(t',x) \tag{5.12}$$

which has a Jacobian of the form

$$\mathcal{J} = \frac{\delta F[f(t,x)]}{\delta f(t',x)} = \text{constant}.$$

Hence, we find that, up to a constant,

$$\int \mathcal{D}[f] \to \int \mathcal{D}[\mathcal{A}],$$

and so the Lagrangian and action, in terms of the field $\mathcal{A}$, become respectively

$$\boxed{\begin{aligned} \mathcal{L}[\mathcal{A}] &= -\frac{1}{2}\left[\mathcal{A}(t,x)^2 + \frac{1}{\mu^2}\left(\frac{\partial \mathcal{A}(t,x)}{\partial x}\right)^2\right] && (5.13)\\ \mathcal{S}[\mathcal{A}] &= -\frac{1}{2}\int_{\mathcal{F}}\left[\mathcal{A}(t,x)^2 + \frac{1}{\mu^2}\left(\frac{\partial \mathcal{A}(t,x)}{\partial x}\right)^2\right], && (5.14)\end{aligned}}$$

with similar boundary conditions to eq. (5.6)

$$\left.\frac{\partial \mathcal{A}(t,x)}{\partial x}\right|_{x=t} = \left.\frac{\partial \mathcal{A}(t,x)}{\partial x}\right|_{x=t+T_{FR}} = 0. \tag{5.15}$$

The transformation in eq. (5.12) uncovered a symmetry for the theory. We namely see that, for another transformation $\sigma(t,x) \to \zeta(t,x)\sigma(t,x)$, a corresponding change $\mathcal{A}(t,x) \to \zeta(t,x)^{-1}\mathcal{A}(t,x)$, leaves eq. (5.11) invariant. In other words, the field theory for $\mathcal{A}$ with volatility structure σ is completely equivalent to the field theory for $\mathcal{B} \equiv \zeta^{-1}\mathcal{A}$ with volatility structure $\zeta\sigma$. This scaling symmetry will later allow us to normalize the two-point correlation function without any issues. It also implies that the volatility structure has no invariant significance and one needs to fix the symmetry before $\sigma(t,x)$ can be uniquely determined from market data [5].

The boundary conditions in eq. (5.15) allow the field to take all possible values at the boundary and also allow us to rewrite the action (5.14) in a similar form as eq. (3.100). The derivation is completely analogous as the derivation that led us to eq. (5.7). We find

$$\mathcal{S}[\mathcal{A}] = -\frac{1}{2}\int_{\mathcal{F}} \mathcal{A}(t,x)\left(1 - \frac{1}{\mu^2}\frac{\partial^2}{\partial x^2}\right)\mathcal{A}(t,x). \tag{5.16}$$

The moment generating functional and partition function for this theory are thus given as

$$\mathcal{Z}[J] = \frac{1}{Z}\int \mathcal{D}[\mathcal{A}] e^{\mathcal{S}[\mathcal{A}]} e^{\int_{\mathcal{F}} J(t,x)\mathcal{A}(t,x)} \tag{5.17}$$

$$Z = \mathcal{Z}[0] = \int \mathcal{D}[\mathcal{A}] e^{\mathcal{S}[\mathcal{A}]}, \tag{5.18}$$

respectively. If we again define $\triangle \equiv 1 - \frac{1}{\mu^2}\frac{\partial^2}{\partial x^2}$, an equivalent derivation as shown in section 3.3.3 yields for eq. (5.17)

$$\mathcal{Z}[J] = e^{\frac{1}{2}\int_{\mathcal{F}} J(t,x)\triangle^{-1}J(t,x)} \tag{5.19}$$

or, equivalently,

$$= e^{\frac{1}{2}\int_{t_i}^{\infty}\mathrm{d}t\int_t^{t+T_{FR}}\mathrm{d}x\mathrm{d}x' J(t,x)G(x,x';t,t+T_{FR})J(t,x')}, \tag{5.20}$$

where we again denoted the Green's function of $\triangle$ as $G(x,x';t,t+T_{FR})$, which is also the propagator of the theory.

It might also be important to note that the action in eq. (5.16) has no derivative with respect to time t and thus the field $\mathcal{A}(t,x)$ can seen as a system in the x-direction at every point in time t which is equivalent to the harmonic oscillator studied in section 3.3.4 [5]. Indeed, if we fix time t and set $\mu = \frac{1}{\sqrt{m}} = \omega$, we find that eqs. (3.99) and (5.13) are completely equivalent (if one also chooses correct integration boundaries).

5.2 The Propagator

Before we go straight to the propagator for the theory, we remind ourselves that the n-point correlation functions can all be calculated from the moment generating functional in eq. (5.20). This is done in a similar fashion as in eq. (3.97):

$$\langle \mathcal{A}(t_1,x_1)\ldots\mathcal{A}(t_n,x_n)\rangle = \frac{\delta}{\delta J(t_1,x_1)}\cdots\frac{\delta}{\delta J(t_n,x_n)}\mathcal{Z}[J]\bigg|_{J=0}. \tag{5.21}$$

When specifically looking at the two-point correlation function, one of the most frequently encountered correlation functions in finance [5], we find

$$\begin{aligned}\langle \mathcal{A}(t,x)\mathcal{A}(t',x')\rangle &= \frac{1}{Z}\int \mathcal{D}[\mathcal{A}]\ \mathcal{A}(t,x)\mathcal{A}(t',x')e^{\mathcal{S}[\mathcal{A}]} \\ &= \frac{\delta^2}{\delta J(t,x)\delta J(t',x')}\mathcal{Z}[J]\bigg|_{J=0} \\ &= \delta(t-t')G(x,x';t,t+T_{FR})\end{aligned} \tag{5.22}$$

where $\delta(t-t')$ was factored out for further convenience. The propagator $G(x,x';t,t+T_{FR})$ thus measures the influence of a fluctuations of the field $\mathcal{A}$ at a point (t,x) on the fluctuations at another point (t',x').

As stated earlier, the field theory of forward interest rates is, in a way, equivalent to that of a harmonic oscillator. This means that, from eq. (3.113), we can immediately write down the propagator for the field $\mathcal{A}$ by substituting $m = 1/\mu^2$ and $\omega = \mu$, which yields

$$\boxed{G(x,x';t,t+T_{FR}) = \mu\frac{\cosh[\mu(T_{FR}-|x-x'|)] + \cosh[\mu(T_{FR}-(x+x'-2t))]}{2\sinh\mu T_{FR}}.} \tag{5.23}$$

Note that the propagator can be rewritten as

$$G(\theta,\theta';T_{FR}) = \mu\frac{\cosh[\mu(T_{FR}-|\theta-\theta'|)]+\cosh[\mu(T_{FR}-(\theta+\theta'))]}{2\sinh\mu T_{FR}}, \tag{5.24}$$

where $\theta = x-t,\ \theta' = x'-t$. The propagator is thus independent of the time t and only depends on the tenor, or time to maturity, θ and θ'. This means that, if the volatility structure also only depends on $x-t$ (as is the case for the Vasicek model, see eq. (2.57)), none of the forward interest rate's properties and characteristics depend explicitly on time t but only depend on their respective tenors: $\theta = x-t$ and $\theta' = x'-t$. A visualization of the propagator[3], unnormalized and normalized, for $\mu = 0.0566$/year can be seen in fig. 5.3.

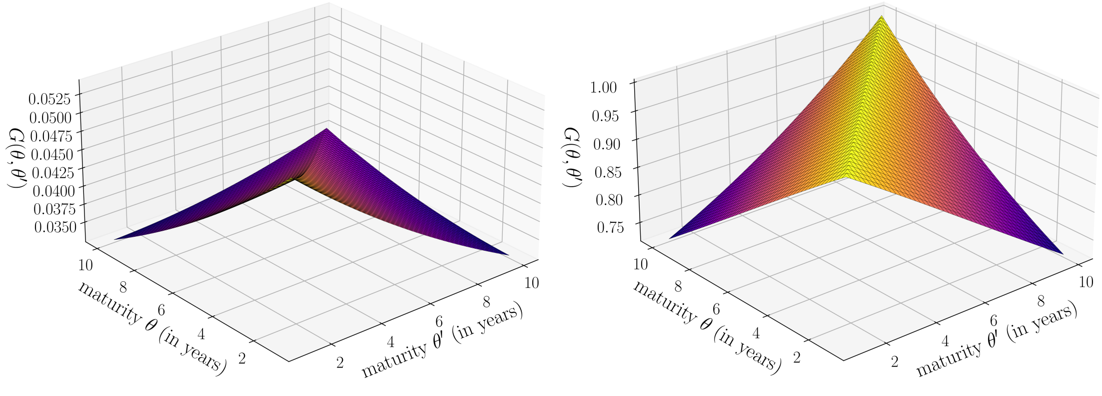


Figure 5.3: A Graphical comparison of the field's propagator $G(\theta,\theta')$ (a) and normalized propagator $\frac{G(\theta,\theta')}{\sqrt{G(\theta,\theta)G(\theta',\theta')}}$ (b), both with $\mu = 0.0566$ year^{-1}.

5.2.1 HJM Limit

To see how this field theory reduces to the original one-factor HJM model, we must look at the behavior in the limit $\mu \to 0$. Since the rigidity of the forward rate curve was given a $\frac{1}{\mu^2}$, this limit corresponds to an infinitely rigid "string". As μ approaches 0, the propagator in eq. (5.23) behaves as $\mu/\tanh\mu$, which means that

$$\lim_{\mu\to 0} G(x,x';t,t+T_{FR}) = 1.$$

Furthermore, if we define

$$\eta(t) = \int_t^{t+T_{FR}} \mathrm{d}x\,\mathcal{A}(t,x), \tag{5.25}$$

which loosely translates to $t \mapsto \mathcal{A}(t,x)$ being a Gaussian noise, we find that

$$\begin{aligned}
\lim_{\mu\to 0}\mathcal{S}[\mathcal{A}] &= \lim_{\mu\to 0} -\frac{1}{2}\int_{t_i}^{\infty}\mathrm{d}t\int_t^{t+T_{FR}}\mathrm{d}x\,\mathrm{d}x'\,\mathcal{A}(t,x)G(x,x';t,T_{FR})\mathcal{A}(t,x') && (5.26)\\
&= -\frac{1}{2}\int_{t_i}^{\infty}\mathrm{d}t\int_t^{t+T_{FR}}\mathrm{d}x\,\mathrm{d}x'\,\mathcal{A}(t,x)\mathcal{A}(t,x') && (5.27)\\
&= -\frac{1}{2}\int_{t_i}^{\infty}\mathrm{d}t\,\eta(t)^2 && (5.28)
\end{aligned}$$

[3]This specific choice for μ will become clear once we discuss the calibration results in chapter 6.

which exactly the action given in eq. (4.19) used in the path integral formulation of the HJM model (where we set $t_f = t_2 \to \infty$).

Moreover, in the case of $T_{FR} \to \infty$, we see that

$$\begin{aligned} G(x,x';t) &\equiv \lim_{T_{FR}\to\infty} G(x,x';t,t+T_{FR}) \\ &= \lim_{T_{FR}\to\infty} \mu \frac{\cosh[\mu(T_{FR}-|x-x'|)] + \cosh[\mu(T_{FR}-(x+x'-2t))]}{2\sinh \mu T_{FR}} \\ &= \frac{\mu}{2}\left[e^{-\mu|x-x'|} + e^{-\mu(x+x'-2t)}\right], \end{aligned} \tag{5.29}$$

and, for $\theta = x - t,\ \theta' = x' - t$, we have

$$G(\theta,\theta') = \frac{\mu}{2}\left[e^{-\mu|\theta-\theta'|} + e^{-\mu(\theta+\theta')}\right], \tag{5.30}$$

where we used the fact that $\cosh\phi = (e^{\phi}+e^{-\phi})/2$ and $\sinh\phi = (e^{\phi}-e^{-\phi})/2$. Equation (5.29) shows that the fluctuations of the field $\mathcal{A}$ in the x-direction are correlated with *correlation time* μ^{-1}. It is clear that, in the limit $\mu \to 0$, this correlation time goes to infinity and thus all fluctuations are perfectly correlated, which is indeed the case in the HJM model.

5.3 No Arbitrage Condition, One Last Time

We will now show that the condition on the drift term of the forward rates $\alpha(t,x)$ for the absence of arbitrage is very similar to eqs. (2.55) and (4.33) but with one small alteration: the implementation of the propagator. The approach is very similar to that of section 4.2.1. Again consider the risk-neutral price at time t_0 of a ZCB with maturity s. This is still given the same as in eq. (4.20), namely

$$P(t_0,s) = \mathbb{E}_{[t_0,t_*]}\left[e^{-\int_{t_0}^{t_*} r(t)\mathrm{d}t} P(t_*,s)\middle|\mathcal{F}_{t_0}\right] \tag{5.31}$$

for $t_i \le t_0 \le t \le t_* \le s \le t_i + T_{FR}$. Hence, by rewriting the expectation value using the probability measure for the field theory,

$$= \frac{1}{Z}\int \mathcal{D}[\mathcal{A}] e^{-\int_{t_0}^{t_*} r(t)\mathrm{d}t} P(t_*,s) e^{\mathcal{S}[\mathcal{A}]}. \tag{5.32}$$

In an analogous manner as the derivation for eq. (4.32) and by using eq. (5.20) we can write

$$\begin{aligned} e^{\int_{\mathcal{T}} \alpha(t,x)} &= \frac{1}{Z}\int \mathcal{D}[\mathcal{A}] e^{-\int_{\mathcal{T}} \sigma(t,x)\mathcal{A}(t,x)} e^{\mathcal{S}[\mathcal{A}]} \\ &= e^{\frac{1}{2}\int_{t_0}^{t_*}\mathrm{d}t \int_t^s \mathrm{d}x\mathrm{d}x' \sigma(t,x)G(x,x';t,t+T_{FR})\sigma(t,x')}, \end{aligned}$$

where $\sigma(t,x)$ was considered to be a source function. The integration region $\mathcal{T}$ is shown in fig. 5.4 and can be seen to be embedded within the larger domain $\mathcal{F}$. Dropping the integration over t on both sides yields

$$e^{\int_t^s \mathrm{d}x \alpha(t,x)} = e^{\frac{1}{2}\int_t^s \mathrm{d}x\mathrm{d}x' \sigma(t,x)G(x,x';t,t+T_{FR})\sigma(t,x')} \tag{5.33}$$

and so we find that the condition for the drift term α to be arbitrage free takes on the following form in the field theory description:

$$\boxed{\alpha(t,x) = \sigma(t,x)\int_t^x \mathrm{d}x'\, G(x,x';t,t+T_{FR})\sigma(t,x'),} \tag{5.34}$$

with $G(x, x'; t, t + T_{FR})$ still given by eq. (5.23). This condition holds true for the drift velocity in any general Gaussian field model and yields the following result for the forward interest rates

$$f(t,x) = f(t_i, x) + \int_{t_i}^{t} \mathrm{d}t' \, \sigma(t', x) \int_{t'}^{x} \mathrm{d}x' \, G(x, x'; t', t' + T_{FR}) \sigma(t', x') + \int_{t_i}^{t} \mathrm{d}t' \, \sigma(t', x) \mathcal{A}(t', x). \tag{5.35}$$

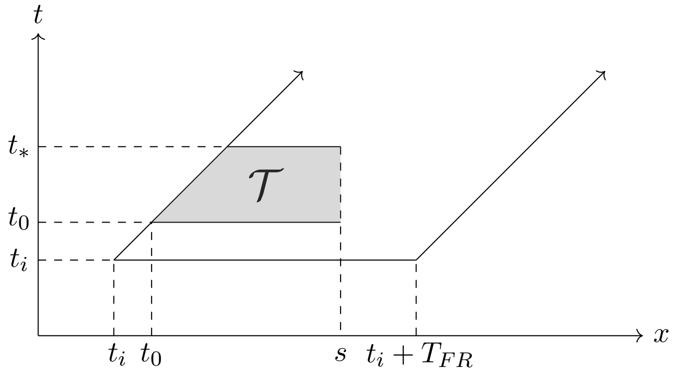


Figure 5.4: Graphical representation of the domain $\mathcal{T}$ embedded in the region $\mathcal{F}$.

5.4 Change of Numeraire in Field Theory

As already explained in section 2.2.3, one can change the quantity with which a discounted value is calculated. In this section we will explain in detail how a change of numeraire from the cash bank account to the ZCB price at present time t_0 manifests itself in the risk-neutral drift and the field theory's action. It is easy to see how the quantity $P(t, s)/B(t_0, t)$ is a martingale, i.e. obeying the martingale condition defined in section 2.1.3. Since $B(t_0, t_0) = 1$, we indeed see, in analogy to eq. (2.25)

$$\begin{aligned} P(t_0, s) &= \mathbb{E}_{[t_0, t_*]} \left[e^{-\int_{t_0}^{t_*} \mathrm{d}t r(t)} P(t_*, s) \middle| \mathcal{F}_{t_0} \right] \\ \frac{P(t_0, s)}{B(t_0, t_0)} &= \mathbb{E}_{[t_0, t_*]} \left[\frac{P(t_*, s)}{B(t_0, t_*)} \middle| \mathcal{F}_{t_0} \right], \end{aligned} \tag{5.36}$$

for some $t_* \in [t_0, s]$.

If we now consider the case in which, at time t_0, all ZCBs are discounted by the ZCB $P(t_0, t_N)$ with $t_0 \le t_N \le s$, the risk-neutral price for $P(t_0, s)$ is given by eq. (2.26)

$$\begin{aligned} P(t_0, s) &= P(t_0, t_N) \mathbb{E}_N \left[P(t_N, T) | \mathcal{F}_{t_0} \right] \\ &= \frac{P(t_0, t_N)}{Z_N} \int \mathcal{D}[\mathcal{A}] P(t_N, T) e^{\mathcal{S}_N[\mathcal{A}]} \end{aligned}$$

where $\mathbb{E}_N$ denotes the expectation value with respect to the new risk-neutral probability measure corresponding to the numeraire $P(t_0, t_N)$, determined by $\mathcal{S}_N$. Z_N is the partition function with respect to this new action. After such a change of numeraire, the quantity $P(t, s)/P(t, t_N)$ is a martingale by the exact same reasoning that explains (5.36).

The risk-neutral drift $\alpha(t, x)$ is however not invariant under a change of numeraire. If we denote the risk-neutral drift corresponding to the new numeraire, $P(t_0, t_N)$, as $\alpha_N(t, x)$, we can perform an analogous derivation as the one for eq. (5.34) that yields

$$\alpha_N(t,x) = \sigma(t,x) \int_{t_N}^{x} \mathrm{d}x'\, G(x,x';t,t+T_{FR})\sigma(t,x'). \tag{5.37}$$

Using this expression for the new risk-neutral drift α_N, we will now derive a relation between the risk-neutral weights $e^{\mathcal{S}}/Z$ and $e^{\mathcal{S}_N}/Z_N$. This can be achieved as follows. For our own convenience and without loss of generality, will work in the $T_{FR} \to \infty$ limit.

First, we adjust the region on which the field theory will be defined such that we only consider the domain relevant for this scenario, i.e. $t_F = t_N$. Denote this region[4] $\mathcal{F}_N$:

$$\mathcal{F}_N = \left\{(t,x) \in \mathbb{R}^2 \middle| t_i \leq t \leq t_N \text{ and } t \leq x \leq \infty\right\}. \tag{5.38}$$

Remember now that the action $\mathcal{S}[f]$, defined on the region $\mathcal{F}_N$, is given as

$$\mathcal{S}[f] = -\frac{1}{2}\int_{\mathcal{F}_N} \left[\frac{\frac{\partial f(t,x)}{\partial t} - \alpha(t,x)}{\sigma(t,x)}\right] G^{-1}(x,x';t) \left[\frac{\frac{\partial f(t,x')}{\partial t} - \alpha(t,x')}{\sigma(t,x')}\right], \tag{5.39}$$

where $\int_{\mathcal{F}_N}$ again shows that the integration is performed over the entire region $\mathcal{F}_N$. The action $\mathcal{S}_N[f]$, on the other hand, is obtained by substituting α_N for α in eq. (5.39).

Let us now define the quantity

$$\delta\alpha \equiv \alpha - \alpha_N = \sigma(t,x) \int_{t}^{t_N} \mathrm{d}x'\, G(x,x';t)\sigma(t,x') \tag{5.40}$$

such that

$$\left(\frac{\delta\alpha}{\sigma}\right)(t,x) = \int_{t}^{t_N} \mathrm{d}x'\, G(x,x';t)\sigma(t,x'). \tag{5.41}$$

This can be used to rewrite $\mathcal{S}_N[f]$ as

$$\begin{aligned}
\mathcal{S}_N[f] &= -\frac{1}{2}\int_{\mathcal{F}_N} \left[\frac{\frac{\partial f(t,x)}{\partial t} - \alpha(t,x) + \delta\alpha(t,x)}{\sigma(t,x)}\right] G^{-1}(x,x';t) \left[\frac{\frac{\partial f(t,x')}{\partial t} - \alpha(t,x') + \delta\alpha(t,x')}{\sigma(t,x')}\right] \\
&= \mathcal{S}[f] - \int_{\mathcal{F}_N} \left[\frac{\frac{\partial f(t,x)}{\partial t} - \alpha(t,x)}{\sigma(t,x)}\right] G^{-1}(x,x';t) \left[\frac{\delta\alpha}{\sigma}\right](t,x') \\
&\quad - \frac{1}{2}\int_{\mathcal{F}_N} \left[\frac{\delta\alpha}{\sigma}\right](t,x) G^{-1}(x,x';t) \left(\frac{\delta\alpha}{\sigma}\right)(t,x').
\end{aligned} \tag{5.42}$$

Now, per construction of the propagator, we have

$$G^{-1}(x,x';t) = \left(1 - \frac{1}{\mu^2}\frac{\partial^2}{\partial x^2}\right)\delta(x-x'), \tag{5.43}$$

which, together with eq. (5.41), yields

$$\begin{aligned}
G^{-1}(x,x';t)\left(\frac{\delta\alpha}{\sigma}\right)(t,x) &= G^{-1}(x,x';t) \int_{t}^{t_N} \mathrm{d}x'\, G(x,x';t)\sigma(t,x') \\
&= \left(1 - \frac{1}{\mu^2}\frac{\partial^2}{\partial x^2}\right)\delta(x-x') \int_{t}^{t_N} \mathrm{d}x'\, G(x,x';t)\sigma(t,x') \\
&= \delta(x-x') \int_{t}^{t_N} \mathrm{d}x\, \delta(x-x')\sigma(t,x')
\end{aligned}$$

[4]Not to be confused with the filtration up to time t_N, which would be denoted $\mathcal{F}_{t_N}$.

$$
\begin{aligned}
&= \delta(x-x')\begin{cases}\sigma(t,x) & \text{for} \quad x\in[t,t_N]\\ 0 & \text{for} \quad x\notin[t,t_N]\end{cases}\\
&= \delta(x-x')\sigma(t,x)\Theta(t_N-x), \qquad (5.44)
\end{aligned}
$$

where Θ denotes the step function. Equation (5.44) can be written in terms of the step function since $t\le x$ is always true. Now plugging eq. (5.44) into eq. (5.42) yields

$$
\begin{aligned}
\mathcal{S}_N[f] = \mathcal{S}[f] &- \int_{\mathcal{F}_N}\left[\frac{\frac{\partial f(t,x)}{\partial t}-\alpha(t,x)-\delta\alpha(t,x)}{\sigma(t,x)}\right]\delta(x-x')\sigma(t,x)\Theta(t_N-x)\\
&-\frac{1}{2}\int_{\mathcal{F}_N}\left(\frac{\delta\alpha}{\sigma}\right)(t,x)\delta(x-x')\sigma(t,x)\Theta(t_N-x).
\end{aligned}
$$

Using eq. (5.41) results in

$$
\begin{aligned}
\mathcal{S}_N[f] = \mathcal{S}[f] &- \int_{\mathcal{F}_N}\left[\frac{\partial f(t,x)}{\partial t}-\alpha(t,x)\right]\delta(x-x')\Theta(t_N-x)\\
&-\frac{1}{2}\int_{\mathcal{F}_N}\left[\int_t^{t_N}\mathrm{d}x''\,G(x,x'';t)\sigma(t,x'')\right]\delta(x-x')\sigma(t,x)\Theta(t_N-x)\\
= \mathcal{S}[f] &- \int_{\mathcal{F}_N}\frac{\partial f(t,x)}{\partial t}\delta(x-x')\Theta(t_N-x)\\
&+\int_{\mathcal{F}_N}\left[\alpha(t,x)-\frac{1}{2}\sigma(t,x)\int_t^{t_N}\mathrm{d}x''\,G(x,x'';t)\sigma(t,x'')\right]\delta(x-x')\Theta(t_N-x).
\end{aligned}
$$

The no-arbitrage condition on the drift term in eq. (5.34) and a change in notation of the integration variable x'' to x' for convenience gives

$$
\begin{aligned}
\mathcal{S}_N[f] = \mathcal{S}[f] &- \int_{T_i}^{t_N}\mathrm{d}t\int_t^{t_N}\mathrm{d}x\,\frac{\partial f(t,x)}{\partial t}+\int_{t_i}^{t_f}\mathrm{d}t\int_t^{t_N}\mathrm{d}x\,\sigma(t,x)\left[\int_t^x\mathrm{d}x'\,G(x,x';t)\sigma(t,x')\right.\\
&\left.-\frac{1}{2}\int_t^x\mathrm{d}x'\,G(x,x';t)\sigma(t,x')-\frac{1}{2}\int_x^{t_N}\mathrm{d}x'\,G(x,x';t)\sigma(t,x')\right]\\
= \mathcal{S}[f] &- \int_{t_i}^{t_N}\mathrm{d}t\int_t^{t_N}\mathrm{d}x\,\frac{\partial f(t,x)}{\partial t}\\
&+\frac{1}{2}\int_{t_i}^{t_N}\mathrm{d}t\int_t^{t_N}\mathrm{d}x\,\sigma(t,x)\left[\int_t^x\mathrm{d}x'\,G(x,x';t)\sigma(t,x')-\int_x^{t_N}\mathrm{d}x'\,G(x,x';t)\sigma(t,x')\right].
\end{aligned}
$$

Since the propagator is symmetric under the interchange $x\leftrightarrow x'$ we can write

$$
\begin{aligned}
\mathcal{S}_N[f] = \mathcal{S}[f] &- \int_{t_i}^{t_N}\mathrm{d}t\int_t^{t_N}\mathrm{d}x\,\frac{\partial f(t,x)}{\partial t}\\
&+\frac{1}{2}\int_{t_i}^{t_N}\mathrm{d}t\left[\int_t^{t_N}\mathrm{d}x\int_t^x\mathrm{d}x'\,G(x,x';t)\sigma(t,x')\sigma(t,x)-\int_t^{t_N}\mathrm{d}x\int_x^{t_N}\mathrm{d}x'\,G(x',x;t)\sigma(t,x)\sigma(t,x')\right].
\end{aligned}
$$

We can now relabel the integration variables in the last term to obtain

$$
\begin{aligned}
\mathcal{S}_N[f] = \mathcal{S}[f] &- \int_{t_i}^{t_N}\mathrm{d}t\int_t^{t_N}\mathrm{d}x\,\frac{\partial f(t,x)}{\partial t}\\
&+\frac{1}{2}\int_{t_i}^{t_N}\mathrm{d}t\left[\int_t^{t_N}\mathrm{d}x\int_t^x\mathrm{d}x'\,G(x,x';t)\sigma(t,x')\sigma(t,x)-\int_t^{t_N}\mathrm{d}x'\int_{x'}^{t_N}\mathrm{d}x\,G(x,x';t)\sigma(t,x')\sigma(t,x)\right].
\end{aligned}
$$

By carefull inspection of the integration boundaries of the two terms inside the square-brackets, we notice that we integrate the same function over the same region, hence these terms cancel. We are thus left with

$$\mathcal{S}_N[f] = \mathcal{S}[f] - \int_{T_i}^{t_N} \mathrm{d}t \int_t^{t_N} \mathrm{d}x \, \frac{\partial f(t,x)}{\partial t}. \tag{5.45}$$

The double integral can be rewritten by looking at the integration domain and applying Fubini's theorem [1]

$$\int_{t_i}^{t_N} \mathrm{d}t \int_t^{t_N} \mathrm{d}x \, \frac{\partial f(t,x)}{\partial t} = \int_{t_i}^{t_N} \mathrm{d}x \int_{t_i}^{x} \mathrm{d}t \, \frac{\partial f(t,x)}{\partial t}.$$

By the fundamental theorem of calculus

$$\int_{t_i}^{t_N} \mathrm{d}t \int_t^{t_N} \mathrm{d}x \, \frac{\partial f(t,x)}{\partial t} = \int_{t_i}^{t_N} \mathrm{d}x \left[f(x,x) - f(t_i,x) \right].$$

Using the fact that $r(t) = f(t,t)$, yields

$$\mathcal{S}_N[f] = \mathcal{S}[f] - \int_{t_i}^{t_N} \mathrm{d}t \, r(t) + \int_{t_i}^{t_N} \mathrm{d}x \, f(T_i, x). \tag{5.46}$$

By eq. (2.30), we can ultimately write

$$e^{\mathcal{S}_N[f]} = \frac{e^{-\int_{t_i}^{t_N} \mathrm{d}t r(t)}}{P(t_i,t_N)} e^{\mathcal{S}}[f]. \tag{5.47}$$

Equation (5.47) relates the weights of the paths the forward rate can take for each of the numeraires. These are different because, as stated in section 2.2.3, to each specific numeraire corresponds a certain martingale measure. This relation will prove especially useful when deriving the closed-form pricing formula for the ZCB option.

5.5 ZCB Option, One Last Time

We will now use the main result eq. (5.47) from section 5.4 to derive the pricing formula of a European call option on a ZCB for one last time. Consider again the price at time t_i of a European call option with strike K and expiry[5] t_N on a ZCB of maturity s, which we know is given by

$$\begin{aligned} \mathrm{ZCBC}(t_i; K, t_N, s) &= \mathbb{E}_{[t_i,t_N]}\left[e^{-\int_{t_i}^{t_N} \mathrm{d}t r(t)} (P(t_N,s) - K)^+ \middle| \mathcal{F}_{t_i} \right] \\ &= \int \mathcal{D}[f] e^{\mathcal{S}[f]} e^{-\int_{t_i}^{t_N} \mathrm{d}t r(t)} (P(t_N,s) - K)^+. \end{aligned} \tag{5.48}$$

Plugging in the relation found in eq. (5.47) gives

$$\begin{aligned} \mathrm{ZCBC}(t_i; K, t_N, s) &= \int \mathcal{D}[f] e^{\mathcal{S}_N[f]} P(t_i,t_N)(P(t_N,s) - K)^+ \\ &= P(t_i,t_N) \mathbb{E}_N \left[(P(t_N,s) - K)^+ \middle| \mathcal{F}_{t_i} \right]. \end{aligned} \tag{5.49}$$

[5] Note that, instead of fixing the option expiry to the maturity of the new numeraire, one could just as easily fix the maturity of the numeraire to the expiry of the option.

Because $P(t_i, t_N)$ has, at present time t_i, a deterministic value, using this as the numeraire, such as in eq. (5.49), greatly simplifies the derivation of the ZCB option price as opposed to using the old-fashioned DCF as numeraire, such as in eq. (5.48).

We will now rederive the price of the call option once using (5.49). Start by reminding ourselves that the payoff of the ZCB call option can be written as in eq. (4.36) and that the forward interest rate is given as (2.49). This means we can write

$$\int_{t_N}^{s} \mathrm{d}x\, f(t_N, x) = \int_{t_N}^{s} \mathrm{d}x\, f(t_i, x) + \int_{t_N}^{s} \mathrm{d}x \int_{t_i}^{t_N} \mathrm{d}t\, \alpha_N(t,x) + \int_{t_N}^{s} \mathrm{d}x \int_{t_i}^{t_N} \mathrm{d}t\, \sigma(t,x)\mathcal{A}(t,x). \tag{5.50}$$

Plugging eq. (5.50) into eq. (4.36) yields

$$[P(t_N, s) - K]^+ = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}G\, \mathrm{d}p\, e^{\mathrm{i}p\left[G + \int_{t_N}^{s}\mathrm{d}x f(t_i,x) + \int_{\mathcal{R}} \alpha_N(t,x) + \int_{\mathcal{R}} \sigma(t,x)\mathcal{A}(t,x)\right]}(e^G - K)^+, \tag{5.51}$$

where we denoted the region of integration by $\mathcal{R} = \{(t,x) \in \mathbb{R}^2 | t_i \leq t \leq t_N \text{ and } t_N \leq x \leq s\}$. Since $\mathcal{A}$ is the only stochastic variable, the expectation value in eq. (5.49) becomes

$$\mathbb{E}_N\left[(P(t_N,s) - K)^+ \middle| \mathcal{F}_{t_i}\right] = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}G\, \mathrm{d}p\, e^{\mathrm{i}p\left[G + \int_{t_N}^{s}\mathrm{d}x f(t_i,x) + \int_{\mathcal{R}} \alpha_N(t,x)\right]}(e^G - K)^+ \mathbb{E}_N\left[e^{\mathrm{i}p\int_{\mathcal{R}}\sigma(t,x)\mathcal{A}(t,x)} \middle| \mathcal{F}_{t_i}\right]. \tag{5.52}$$

We can evaluate the expectation value as follows. We start by writing the expectation value in the field theory formalism

$$\mathbb{E}_N\left[e^{\mathrm{i}p\int_{\mathcal{R}}\sigma(t,x)\mathcal{A}(t,x)} \middle| \mathcal{F}_{t_i}\right] = \frac{1}{Z_N}\int \mathcal{D}[\mathcal{A}] e^{\mathrm{i}p\int_{\mathcal{R}}\sigma(t,x)\mathcal{A}(t,x)} e^{\mathcal{S}_N[\mathcal{A}]}.$$

Considering $\mathrm{i}p\sigma(t,x)$ as a source function allows us to use eq. (5.20) which yields

$$= \exp\left[-\frac{p^2}{2}\int_{t_i}^{t_N}\mathrm{d}t \int_{t_N}^{s} \mathrm{d}x\, \mathrm{d}x'\, \sigma(t,x) G(x,x';t)\sigma(t,x')\right].$$

For notational purposes, we now define

$$q^2 \equiv \int_{t_i}^{t_N}\mathrm{d}t \int_{t_N}^{s} \mathrm{d}x\, \mathrm{d}x'\, \sigma(t,x) G(x,x';t)\sigma(t,x'). \tag{5.53}$$

Furthermore, we see that, by rearranging the integration boundaries

$$q^2 = \int_{t_i}^{t_N}\mathrm{d}t \int_{t_N}^{s}\mathrm{d}x \left[\int_{t_N}^{x} \mathrm{d}x'\, \sigma(t,x)G(x,x';t)\sigma(t,x') + \int_{x}^{s}\mathrm{d}x' \sigma(t,x)G(x,x';t)\sigma(t,x')\right].$$

Because of the symmetry of the propagator under the interchange of x and x', we can write

$$\begin{aligned} &= \int_{t_i}^{t_N}\mathrm{d}t \left[\int_{t_N}^{s}\mathrm{d}x \int_{t_N}^{x}\mathrm{d}x'\, \sigma(t,x)G(x,x';t)\sigma(t,x') + \int_{t_N}^{s}\mathrm{d}x' \int_{x'}^{s}\mathrm{d}x\, \sigma(t,x)G(x,x';t)\sigma(t,x')\right] \\ &= 2\int_{t_i}^{t_N}\mathrm{d}t \int_{t_N}^{s}\mathrm{d}x \int_{t_N}^{x}\mathrm{d}x'\, \sigma(t,x)G(x,x';t)\sigma(t,x'). \end{aligned}$$

Hence, using the expression for α_N in eq. (5.37)

$$\int_{\mathcal{R}} \alpha_N(t,x) = \int_{t_i}^{t_N}\mathrm{d}t \int_{t_N}^{s}\mathrm{d}x \left[\sigma(t,x)\int_{t_N}^{x}\mathrm{d}x'\, G(x,x';t)\sigma(t,x')\right]$$

$$\int_{\mathcal{R}} \alpha_N(t,x) = \frac{q^2}{2}. \tag{5.54}$$

Equation (5.54) is exactly the identity used in section 4.2.2 in eq. (4.47), which we have now successfully proven (albeit in a field theory setting). Returning to eq. (5.52) and plugging in eq. (5.54) gives

$$\mathbb{E}_N\left[(P(t_N,s)-K)^+\middle|\mathcal{F}_{t_i}\right] = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}G\,\mathrm{d}p\, e^{\mathrm{i}p\left[G+\int_{t_N}^{s}\mathrm{d}x f(t_i,x)+\frac{q^2}{2}\right]}(e^G-K)^+ e^{-\frac{q^2}{2}p^2}.$$

Note that here G is nothing more than an integration variable coming from the exponential expression of the Dirac delta and has nothing to do with the propagator $G(x,x';t)$. Performing the Gaussian integration over p yields

$$= \frac{1}{\sqrt{2\pi q^2}}\int_{-\infty}^{\infty} \mathrm{d}G\, e^{-\frac{1}{2q^2}\left[G+\frac{q^2}{2}+\int_{t_N}^{s}\mathrm{d}x f(t_i,x)\right]}(e^G-K)^+. \tag{5.55}$$

We now see that the integral in eq. (5.55) is exactly the same integral as in eq. (4.49), apart from the pre-factor. The result of evaluating eq. (5.55) is thus given by

$$\mathbb{E}_N\left[(P(t_N,s)-K)^+\middle|\mathcal{F}_{t_i}\right] = F\Phi(d_+) - K\Phi(d_-), \tag{5.56}$$

where

$$F = e^{-\int_{t_N}^{s}\mathrm{d}x f(t_i,x)} \tag{5.57}$$

and

$$d_\pm = \frac{1}{q}\left[\ln\frac{F}{K} \pm \frac{q^2}{2}\right]. \tag{5.58}$$

Because the price the call option on the ZCB is actually given as eq. (5.49), we ultimately find:

$$\boxed{\mathrm{ZCB}C(t_i;K,t_N,s) = P(t_i,t_N)\left[F\Phi(d_+) - K\Phi(d_-)\right].} \tag{5.59}$$

Note that this is again the "Black-Scholes"-like pricing formula we would expect from eqs. (2.70) and (4.53), however there is one small difference. The way we defined the quantity q^2 in eq. (5.53) now takes the propagator from eq. (5.23) into account. This enables us to allow for a non-trivial correlation between forward rates with different maturity times x. These correlations are generated through the field $\mathcal{A}(t,x)$ via its two-point correlation function in eq. (5.22).

We have now successfully derived the pricing formula of a call option on a ZCB in three different settings:

1. the "classical" setting, as was first proposed by Heath, Jarrow and Morton in [32];
2. the path-integral setting where we reformulated the forward rate dynamics as a Wiener path integral (although written in a formalism that eased the transition to the field theory description);
3. the field theory setting, where the one-factor HJM model was extended to a (statistical) field theory.

A simple comparison of the "classical" and field theory with constant volatility ZCB option price as a function of the option's time to expiry $\theta_* = t_* - t_0$ can be seen in fig. 5.5. The prices in the different models or model versions (i.e. $T_{FR} = 30Y$ or $T_{FR} \to \infty$) are very similar. Even more so, the difference in "field theory price" between using $G(x, x'; t, t + T_{FR})$ and $G(x, x'; t)$ as given by eqs. (5.23) and (5.29) respectively, is negligible. This shows that the field theory model can produce reliable ZCB option prices and is consistent with well established ZCB option pricing formulas. We can also see that the price of the option increases for longer times to expiry. This is because the holder of the contract has more time for the ZCB price to move above (or below) the strike value [59]. Also, the field theory price with $\mu = 0$ coincides (almost) exactly with the HJM price, as expected. The small discrepancy can be a consequence of the machine accuracy and numerical integration algorithms. The strength in the field theory model lies in the incorporated propagator. This accounts for non-trivial correlations of forward rates with different maturities which allows for better pricing of more exotic interest rate derivatives. Furthermore, it is important to note that, to plot the price of a ZCB option, one requires interest rate data for the entirety of the bond's lifetime, i.e. from t_i to s. Such data was kindly provided by KBC Bank and will be examined in chapter 6.

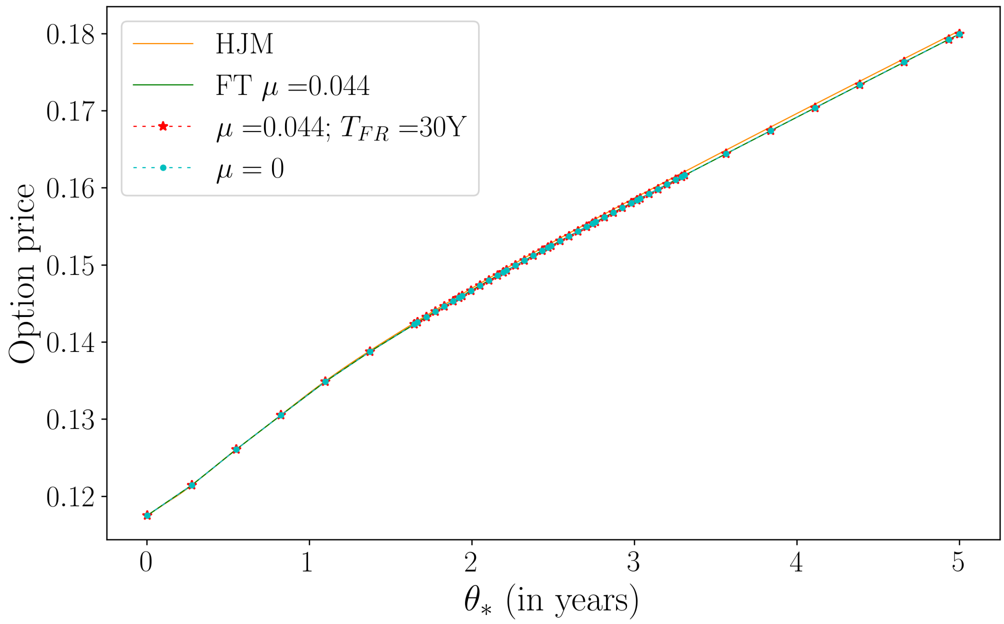


Figure 5.5: A Graphical comparison of the ZCB option price as a function of its time to expiry with constant volatility. The volatility is taken $\sigma = 0.005$ year$^{\frac{1}{2}}$ for both the one-factor HJM and field theory model. The field theory models take $\mu = 0.0566$ year^{-1}.

Chapter 6

Model Calibration

Even though the idea of a field theory of forward interest rates seems compelling and yields acceptable theoretical results, a brilliant physicist and idol of mine once said

> "It doesn't matter how beautiful your theory is, it doesn't matter how smart you are. If it doesn't agree with experiment, it's wrong.
>
> - *Richard P. Feynman*"

To honor this truthful quote, we will test the field theory against market data by calibrating the ZCB option price in eq. (5.59) and comparing the results with conventional models such as the one-factor HJM model with a constant volatility and a Vasicek volatility (see eq. (2.57)). In this chapter, we explain the final pieces of underlying theory necessary to perform the desired calibration, discuss the data that was used and how it was manipulated. Next, the calibration procedure is explained and we conclude with a discussion of the general results.

6.1 Final Pieces of Theory

As mentioned in section 2.2.2, ZCBs (and by extension ZCB options) are theoretical products and are not actually available for trading. However, there exists an interest rate derivative, closely related to the ZCB option, that *is* traded. This derivative is the *interest rate cap*. Without delving too much into the details, an interest rate cap, or cap, is a financial contract that ensure the *floating* interest rate (i.e. the interest rates as set by the market) not to exceed a certain threshold or *strike rate* [59]. A cap is constructed out of many smaller payments that are made periodically (often quarterly) called *caplets*. A cap can be considered as the sum of all these individual caplets and each of these caplets is linked to a ZCB option in the following way. Consider the price at time t_0 of a caplet with *fixing time*[1] t_*, payment date s and threshold value K for a total notional amount N, is given by [15]

$$\mathrm{Cplt}(t_0; K, t_*, s, N) = \tilde{N}\mathrm{ZCB}P(t_0; \tilde{K}, t_*, s), \tag{6.1}$$

[1] The fixing time of a caplet is the moment at which the floating rate is tested against the predetermined strike rate and is similar to the expiry of a bond option.

where

$$\tilde{N} = N(1 + K\tau) \tag{6.2}$$
$$\tilde{K} = \frac{1}{1 + K\tau}, \tag{6.3}$$

and $\tau = (s - t_*)/365$ is the *daycount fraction* between fixing time t_* and payment date s. This daycount fraction is calculated by dividing by the duration in days by the total number of days in one *trading year*. There are many different conventions for the amount of *trading days* in one year and we use the *Actual/365* conventions, meaning we take there to be 365 days in one trading year. The arbitrage-free price of a caplet can thus be written as $\tilde{N}$ times the price of a corresponding (European) ZCB put option with expiry t_*, strike $\tilde{K}$ and bond maturity s. The put option price was denoted $P(t_0; \tilde{K}, t_*, s)$ and can be obtained from eq. (5.59) via the Put-Call Parity relation:

$$\text{ZCB}P(t_0; \tilde{K}, t_*, s) = P(t_0, t_*)[-F\Phi(-d_+) + \tilde{K}\Phi(-d_-)], \tag{6.4}$$

with F and $d_\pm$ given by eqs. (5.57) and (5.58) respectively.

Caplet prices are not directly observable in the market but they can however be *stripped* from the associated cap by a method called *caplet stripping* [59]. This allows one to extract the caplet prices and their associated characteristics. These caplets can thus be used as calibration reference. If we want to calculate the caplet prices numerically, an extra piece of data is required: the instantaneous forward interest rate curve at time t_0. Unfortunately, also the instantaneous forward interest rates are, as mentioned in section 2.2.4, artificial quantities and therefore not directly observable in the market. But they can be derived from the yield to maturity $R(t, T)$ that *is* observed in the market.

The yield to maturity $R(t, T)$ was discussed in section 2.2.4 and can be seen as the average rate on a zero coupon loan starting at time t over a period of length T. From this we can derive a corresponding DCF $D(t, s)$, with $s = t + T$, using [12]

$$D(t, t + T) = [1 + R(t, T)]^{-T}. \tag{6.5}$$

Next, this is related to the the forward rates $F(t; s_1, s_2)$, with $s_{1,2} = t + T_{1,2}$, by

$$F(t; t + T_1, t + T_2) = \frac{-\ln\left(\frac{D(t,t+T_2)}{D(t,t+T_1)}\right)}{T_2 - T_1}. \tag{6.6}$$

The instantaneous forward interest rates $f(t, t + T)$ were defined in section 2.2.4 by the limit of $F(t; t + T_1, t + T_2)$ for $T_1 \to T_2$. This limit can be approximated by taking $T_2 = T_1 + \epsilon$, where ϵ equals one trading day, i.e. $\epsilon = 1/365$ in the Actual/365 convention. Hence

$$f(t, t + T) \approx F(t; t + T, t + T + \epsilon). \tag{6.7}$$

6.2 The Market Data

The data we use are yields to maturity, $R(t, T)$, given for daily values of t over a long period of time, from the 2^{nd} of January 2006 to the 23^{rd} of May in 2023, for various tenors T, ranging from 1 day to 50 years. The "raw" yields to maturity as observed in the market are displayed in fig. 6.1 as a function of time t for various tenors T (see fig. 6.1a) and as a function of the tenor T for different dates t (see fig. 6.1b).

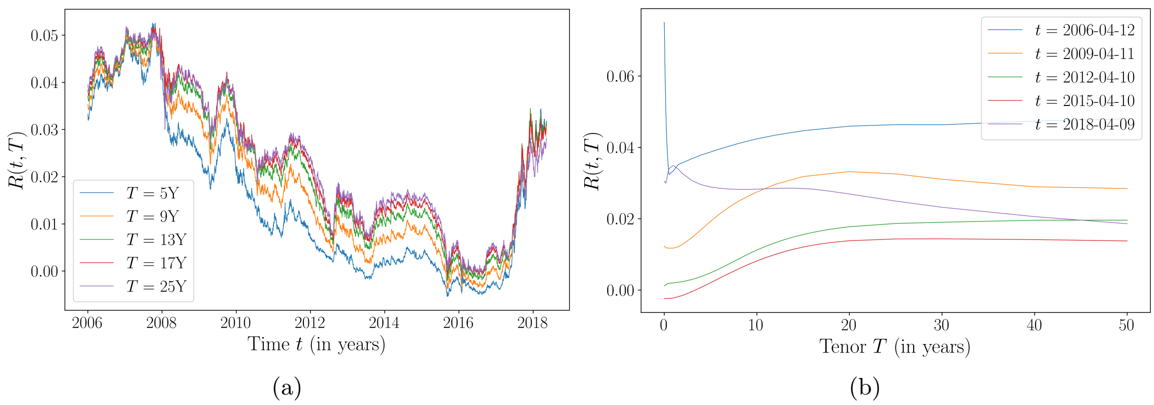


Figure 6.1: A Graphical representation of the raw historical yields to maturity as a function of time t (a) and as a function of the tenor T (b). The extreme behavior at small tenors T is (partly) due to influence from Central Banks.

Using eqs. (6.5) and (6.6), we can transform the yields to maturity to the corresponding forward rates $F(t; t + T_i, t + T_{i+1})$, where T_i runs over all available tenors. We then approximate the instantaneous forward interest rates by using eq. (6.7). The data is provided only for some specific tenors T_i for which it is generally the case that $T_{i+1} - T_i \ggg \epsilon$. Hence, they can not be used for a good (enough) approximation of the instantaneous forward interest rates. A solution is to perform a cubic spline interpolation on the data over the tenor values such that we obtain data points for daily values of T_i, i.e. $T_{i+1} - T_i = \epsilon$. Such an interpolation can be achieved using the "scipy" python library and yields smooth instantaneous forward rate curves $T \mapsto f(t, t + T)$, where both t and T take daily values. This interpolation is carried out only over a portion of the data to mitigate computational requirements. The results of an interpolation for the final 5 years of the data-set, i.e. 2018 until 2023, over daily tenor values ranging from 3 up until 13 years are visualized in fig. 6.2. In fig. 6.2a, the instantaneous forward rates $f(t, t + T)$ are shown as a function of time t for different tenor values, while in fig. 6.2b, $f(t, t + T)$ is shown as a function of the tenor T for different dates t. Note that in the latter, one can see small "bumps" in the curves as a result of the cubic spline interpolation. The corresponding interpolated instantaneous forward rate surface $(t, T) \mapsto f(t, t + T)$ is shown in fig. 6.3. Note that the same historical data was used to calculate (with eq. (6.5)) and display the "discount surface" in fig. 2.4 as for the instantaneous forward rate surface in fig. 6.3. Furthermore, the data is only provided for the tenors and not explicitly for maturity times s. It is therefore a lot more convenient to transform our equations to that same convention. This is achieved rather easily by performing the transformations $s \to \theta_s = s - t_0$ and $t_* \to \theta_* = t_* - t_0$, which makes data manipulation a lot less complicated.

The interpolated instantaneous forward interest rates $f(t_0, t_0 + T)$ can be used to calculated the price of a ZCB as a function of its time to maturity T at different dates t_0 via eq. (2.30). The integration is carried out by means of the "scipy.integrate" module of the python "scipy" library. The results can be seen in fig. 6.4. Note that here, we only had to interpolate to curve at a specific date t_0 one at a time, instead of an entire surface. Therefore, we did not encounter computational difficulties and managed to interpolate over the tenors ranging from $T = 0$ to $T = 36$.

Another essential piece of data are the caplet prices and their corresponding characteristics. We received the caplet data on May 10th, 2023 and we will set this date equal to t_0. This means that all caplets are priced on t_0. Each of the caplets has a different fixing and payment date, t_{*i} and s_i respectively, but they all have 6 months between fixing and payments, i.e. $s_i - t_{*i} = 6M$, $\forall i$.

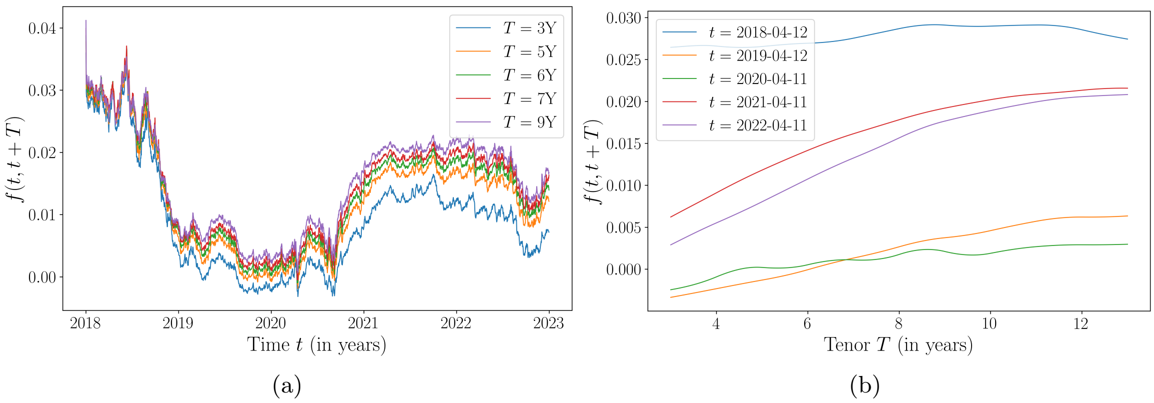


Figure 6.2: Graphical representation of the interpolated historical instantaneous forward interest rates as a function of time t (a) and as a function of the tenor T (b). The instantaneous forward interest rates are derived from the historical yields to maturity via eqs. (6.5) to (6.7).

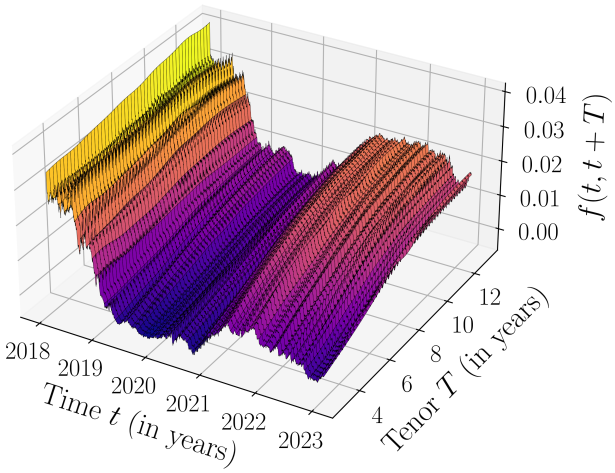


Figure 6.3: Graphical representation of the interpolated instantaneous forward rate surface. Both the time t- and tenor T axes have daily value-points.

They also share the same ATM strike K. This ATM strike for caplets is based on the threshold of the encompassing cap and is defined as [15]

$$K = \frac{\sum_i (s_i - t_{*i}) D(t_{*i}, s_i) F(t_0; t_{*i}, s_i)}{\sum_i (s_i - t_{*i}) D(t_{*i}, s_i)}, \tag{6.8}$$

where both sums run over all available caplets. Note that, for every i, we have that $s_i - t_{*i} \approx 0.5$.

Accompanied with these characteristics is each of the caplet's respective *cashflow*, i.e. the caplet's price. The notional amount is always $N = 100\,000\,000$. This numerical data is displayed in table 6.1. A plot of the caplet prices in function of their time to the payment date $\theta_s \equiv s - t_0$ can be seen in fig. 6.5. We see that the first two caplets appear to be divergent from the rest. This is due to their short time to payment date and the fact that short-term interest rates are heavily influenced by central banks [59]. We will therefore discard these two data points.

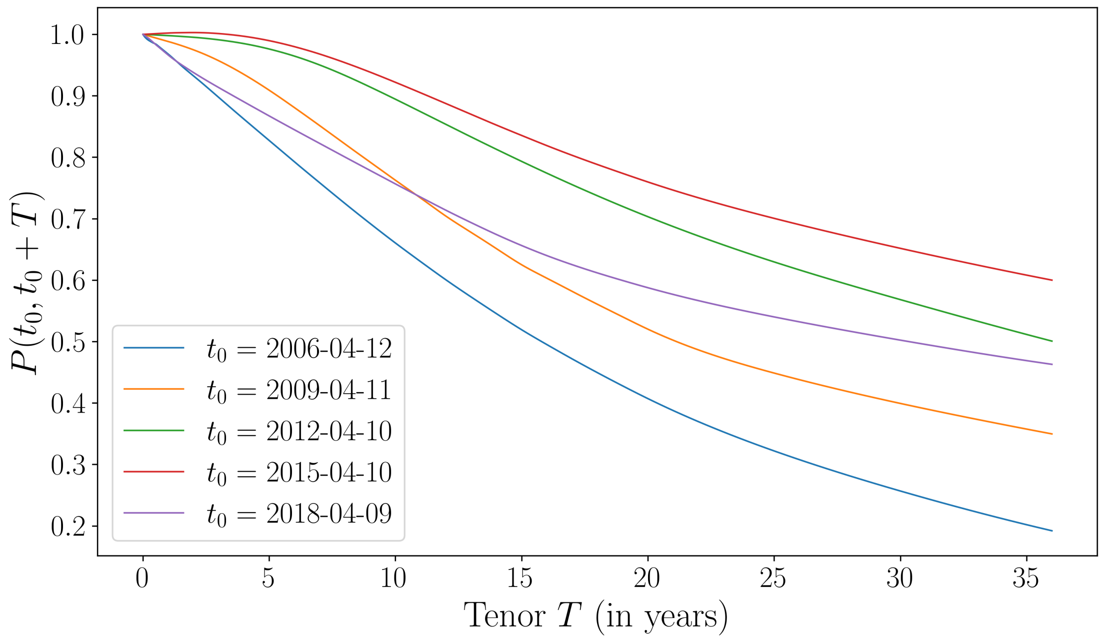


Figure 6.4: Graphical representation of the ZCB price at different dates, calculated from interpolated instantaneous forward rate data.

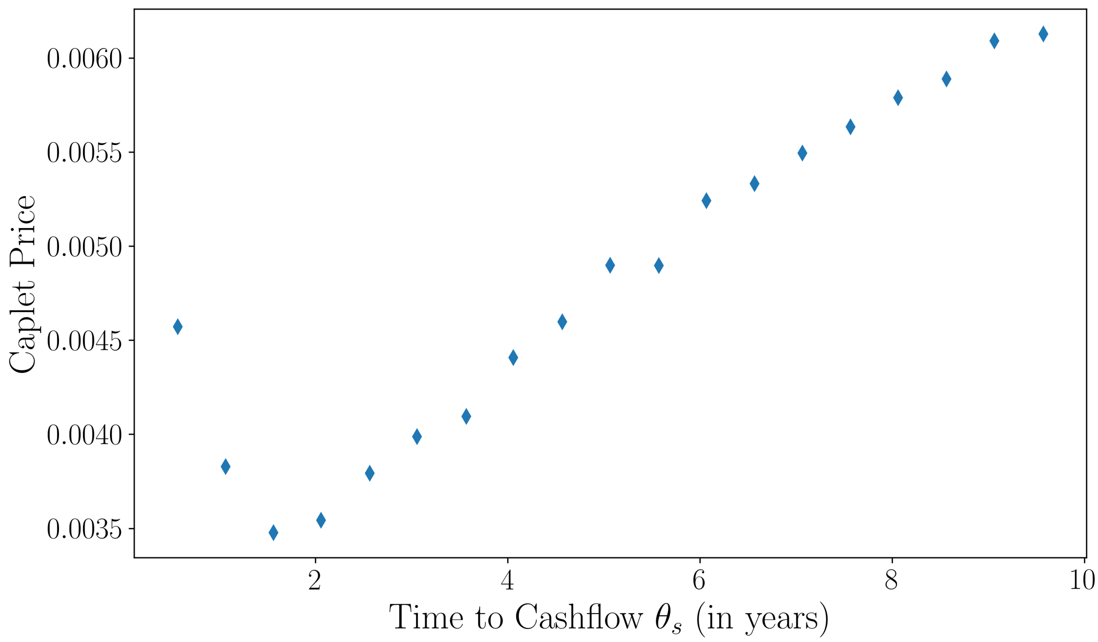


Figure 6.5: Plot of the caplet market prices in function of their time to the payment date $\theta_s = s - t_0$ on May 10, 2023.

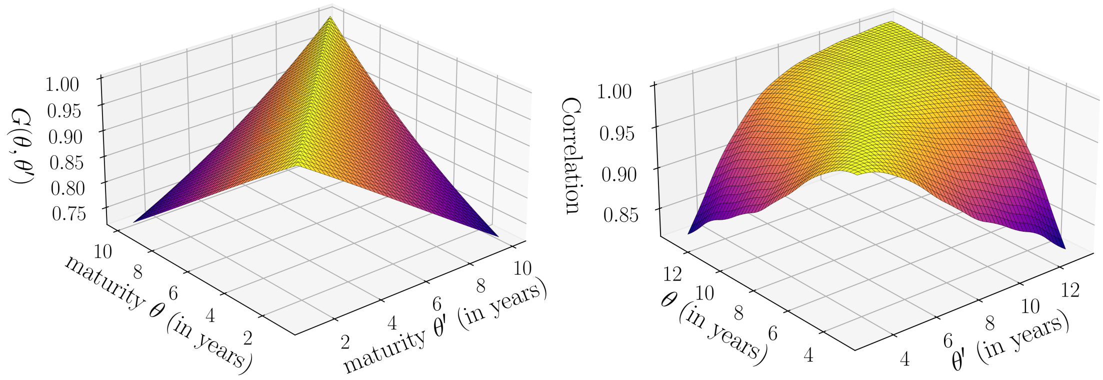
Table 6.1: Numerical ATM caplet data as observed in the market on May 10, 2023.

Fixing Date	Payment Date	Strike (%)	$F(t_0; t_{*i}, s_i)$ (%)	$D(t_{*i}, s_i)$	$N\times$Price (USD)
2/11/2023	6/05/2024	2,94	3,7497	0,96487	457.249,59
2/05/2024	4/11/2024	2,94	3,3529	0,94983	382.831,70
31/10/2024	5/05/2025	2,94	3,0062	0,93670	347.785,66
30/04/2025	4/11/2025	2,94	2,8146	0,92413	354.369,28
31/10/2025	4/05/2026	2,94	2,7738	0,91286	379.442,63
29/04/2026	4/11/2026	2,94	2,7207	0,90112	398.786,75
2/11/2026	4/05/2027	2,94	2,7103	0,89029	409.506,88
30/04/2027	4/11/2027	2,94	2,7458	0,87872	440.893,21
2/11/2027	4/05/2028	2,94	2,7766	0,86770	459.879,53
2/05/2028	6/11/2028	2,94	2,7949	0,85604	489.899,31
2/11/2028	4/05/2029	2,94	2,8121	0,84509	489.720,77
2/05/2029	5/11/2029	2,94	2,8335	0,83362	524.261,56
1/11/2029	6/05/2030	2,94	2,8609	0,82254	533.252,20
2/05/2030	4/11/2030	2,94	2,8942	0,81129	549.595,05
31/10/2030	5/05/2031	2,94	2,9252	0,80021	563.411,15
30/04/2031	4/11/2031	2,94	2,9547	0,78879	579.029,50
31/10/2031	4/05/2032	2,94	2,9846	0,77757	588.856,68
30/04/2032	4/11/2032	2,94	3,0168	0,76609	609.138,14
2/11/2032	4/05/2033	2,94	3,0491	0,75489	612.767,72

6.3 Visual & Numerical Tests

Visual Test

A first short but interesting test is to have a look a the empirical correlations between the instantaneous forward rates. According to the one-factor HJM model, these correlations will all be equal 1 and thus the correlation surface would be the flat surface $(t, s) \mapsto 1$. However, what we observe is a lot different. This empirical correlation surface can be easily calculated by means of the "numpy" python library and the resulting surface is shown in fig. 6.6b. Clearly, this is far from the flat correlation surface one would expect from the HJM model. Moreover, when comparing fig. 6.6b to the surface determined by the propagator in fig. 5.3b, it is clear that correlations in field theory carry a lot more structure. Nevertheless, there is still a lot of room for improvement and we will discuss some suggestions in chapter 7.



(a) (b)

Figure 6.6: A Graphical comparison of the theoretically predicted correlation surface as induced by the propagator $G(\theta, \theta')$ (a) and the correlation surface of the interpolated historical instantaneous forward interest rates $(\theta, \theta') \mapsto \langle f(t, t + \theta) f(t, t + \theta') \rangle$ (b).

Numerical Test

The second and main test of this thesis is the calibration of the field theory ZCB option price against the caplet market prices shown in fig. 6.5. All the ingredients one needs to calculate the theoretical caplet price in the different models, using eqs. (4.53), (5.59) and (6.1) are given in table 6.1. The theoretical caplet prices are fine-tuned in such a way as to minimize the normalized *root mean square error* (RMSE):

$$\text{RMSE} = \sqrt{\sum_i \left(\frac{\text{Cplt}_i - \tilde{N}_i \text{ZCB}P(t_0; \tilde{K}_i, t_{*i}, s_i)}{\text{Cplt}_i} \right)^2}, \tag{6.9}$$

where Cplt_i denotes the market price of the caplet corresponding to the characteristics (X_i, t_{*i}, s_i). This RMSE is actually a function of the parameters that reside within the put option price. What parameters these are depends on the model one considers. For our model comparison we will examine three different cases:

Case 1. **Constant volatility one-factor HJM model**: Constant volatility one-factor HJM model: the ZCB option price is determined[2] by eq. (2.70) and the RMSE is thus a function of the

[2] Note that by the put-call parity relation in eq. (2.3), we can convert the call price to a put price.

constant volatility σ; RMSE(σ).

Case 2. **Vasicek-like one-factor HJM model**: the volatility structure is given by eq. (2.57) and the ZCB option price can be found by plugging

$$\begin{aligned} q^2 &= \int_{t_0}^{t_0+\theta_*} \mathrm{d}t \left[\int_{t_*-t}^{s-t} \mathrm{d}\theta \, \rho e^{-k\theta} \right]^2 \\ &= \frac{\rho^2}{2k^3} \left(e^{-k\theta_*} - e^{-k\theta_s} \right)^2 \left(e^{k\theta_*} - 1 \right), \end{aligned} \tag{6.10}$$

with $\theta_* = t_* - t_0$ and $\theta_s = s - t_0$ into eq. (4.53). The RMSE is thus a function of ρ and k; RMSE(ρ, k).

Case 3. **Constant volatility field theory**: the ZCB option price is given by eq. (5.59) with

$$\begin{aligned} q^2 &= \sigma^2 \int_{t_0}^{t_0+\theta_*} \mathrm{d}t \int_{\theta_*+t_0-t}^{\theta_s+t_0-t} \mathrm{d}\theta \, \mathrm{d}\theta' \, G(\theta, \theta'), \\ &= \sigma^2 \int_{t_0}^{t_0+\theta_*} \mathrm{d}t \int_{\theta_*+t_0-t}^{\theta_s+t_0-t} \mathrm{d}\theta \, \mathrm{d}\theta' \, \frac{\mu}{2} \left[e^{-\mu|\theta-\theta'|} + e^{-\mu(\theta+\theta')} \right] \\ &= \sigma^2 \left[(\theta_s - \theta_*)\theta_* - \frac{1}{\mu} \left(1 - e^{-\mu(\theta_s-\theta_*)} \right) \theta_* + \frac{1}{4\mu^2} \left(1 - e^{-\mu(\theta_s-\theta_*)} \right)^2 \left(1 - e^{-2\mu\theta_*} \right) \right] \end{aligned} \tag{6.11}$$

where again $\theta_* = t_* - t_0$ and $\theta_s = s - t_0$ and $G(\theta, \theta')$ is given by eq. (5.30). Note that we have written the analytical result for q^2 in such a way that it can conveniently be related to the data. Hence the RMSE is a function of σ and μ; RMSE(σ, μ).

Each of these cases will have its RMSE minimized by finding the optimal values for its respective parameters. The minimization is achieved through the *Nelder-Mead* algorithm provided by the "scipy" python library. The minimized RMSE-values with their corresponding parameters are displayed in table 6.2. The associated price graphs are shown in fig. 6.7. Cases 1, 2 and 3 have a respective RMSE of 0.02981, 0.01646 and 0.01598 and can thus all three be considered a reasonable fit. Naturally, the fits of Cases 2 and 3 are better than the fit of Case 1 since they each have two parameters. However, this is not a completely fair comparison and the field theory model is actually in a disadvantage compared to the Vasicek-like HJM model (and even then the RMSE of the field theory model is slightly smaller than that of the Vasicek like one-factor HJM model). If we would like to level the playing field, we should plug the Vasicek-like volatility structure, used in Case 2, into eq. (5.59) such that the expression for q^2 from eq. (5.53) is altered. This would then increase the accuracy of the field theory model even further (but might not be worth the higher computational complexity). Nevertheless, the simplest field theory model that one can construct immediately shows its ability to accurately reproduce the caplet market prices.

Table 6.2: Numerical calibration of each of the three cases.

	RMSE	σ (year$^{\frac{1}{2}}$)	μ (year^{-1})	ρ (year$^{\frac{1}{2}}$)	k (year^{-1})
Constant volatility model	0.02981	0.01269			
Vasicek-like volatility model	0.01646			0.01907	0.01399
Field theory model	0.01598	0.05750	0.05662		

One might wonder why Case 2 yields a calibrated fit of such accuracy since, after all, Case 2 is also a one-factor HJM model with the same "problematic" exact correlation. We just imposed more structure on the volatility. While this is true and the Vasicek-like HJM model proves useful for a mere pricing calibration of simple caplets, this model will fall of rapidly in both accuracy and utility when dealing with more exotic interest rate derivatives. One example of such an exotic derivative is a *spread-option* on a *constant maturity swap* (CMS), which is a financial instrument that explicitly depends on the correlation between two forward rates of different tenor. A one-factor HJM model with a Vasicek-like volatility structure would fail to deal with such derivatives' correlation surface in an appropriate manner [30].

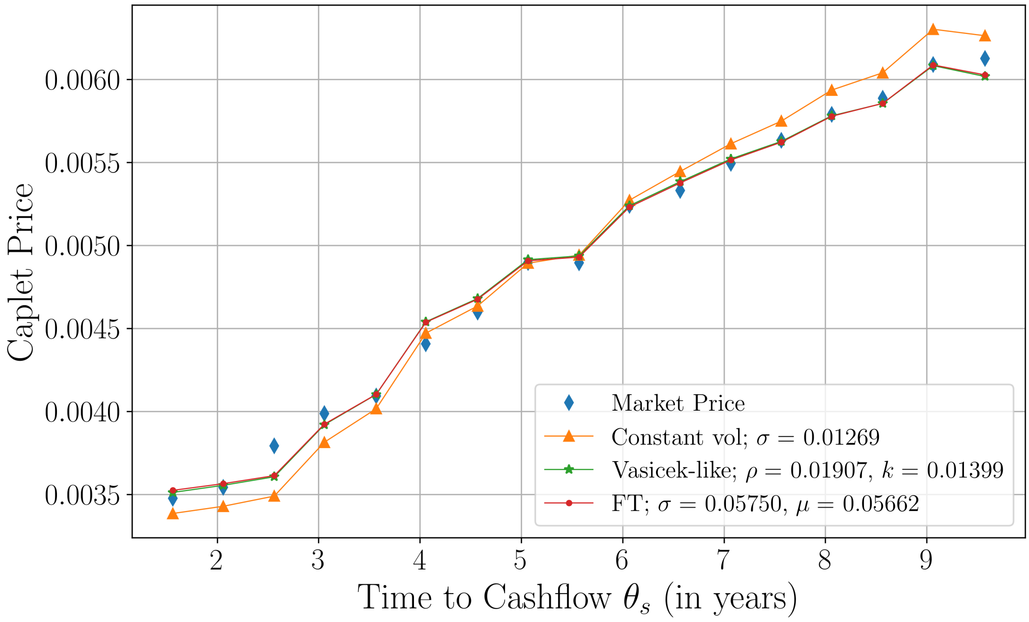


Figure 6.7: Comparison of the model fits to the market caplet prices. The figure shows the best fits of the constant volatility one-factor HJM model (1 parameter), Vasicek-like one-factor HJM model (2 parameters) and field theory model (2 parameters) to the caplet prices as determined by the market on May 10, 2023.

Chapter 7

Conclusion & Further Prospects

7.1 Conclusion

We started this thesis with a discussion about the relevance of theories in physics for finance and why it would be interesting to study the inner workings behind the "time-value of money". This naturally led us to the concept of interest rates, cash bank accounts and DCFs to which we gave a basic introduction in chapter 2. Here, we saw that, on the contrary to regular stock, interest rates (except for the spot rate) and the DCFs have an infinite degrees of freedom, i.e. each (time to) maturity or duration of a loan. This allows us to represent these values by means of a field or better, because they are considered to be stochastic in nature, a statistical field. We introduced two familiar interest rate models, Vasicek and one-factor HJM, and derived closed-form pricing formulas for the ZCB and ZCB option within these models, given in eqs. (2.43) and (2.68) respectively. For the HJM model, a simple no-arbitrage argument was given to obtain the form of the risk-neutral drift in eq. (2.55). The derivations followed a "classical" approach as proposed by the authors in their original papers (see [32,56]). These models, however showed a major flaw: they assume exact correlations between rates for different maturities, which is, as displayed in fig. 6.6b, not observed in the market. The field representation of forward interest rates and their apparent non-trivial correlation surface guided us in the direction of a field theory description.

As we aimed to extend the HJM model to a field theory of interest rates, we introduced stochastic processes and their path integral description. This formalism was later used to translate the earlier examined interest rate models to the respective path integral formulation. On top of that, we briefly introduced the formalisms used in statistical and quantum field theory. The main takeaway of this discussion was the comparison of the different path integrals and how they all took a "Feynman-Kac" form when using the appropriate Lagrangian and action. Afterwards, some neat computational tricks using partition functionals and the propagator for the harmonic oscillator were derived in the context of QFT. The harmonic oscillator, considered as a (0+1) field, and its propagator with Neumann boundary conditions in eq. (3.113) later proved essential in describing the simplest field theory for interest rates.

We applied this path integral description of SDE to both the Vasicek and HJM model, allowing us to rederive the previously obtained closed-form pricing formulas for the ZCB and ZCB option in eqs. (4.53) and (4.13) respectively. The risk-neutral drift was rederived as well in eq. (4.33). We explained that the contents of chapter 4 are formally a translation to a path integral description of two, already well-established, interest rate models. Nevertheless, this translation familiarized us with the path integrals and its formalism to ease the further transition to a field theory of interest rates.

We justified such a transition by again stating that the interest rate curve can be seen as a "system" with an infinite amount of degrees of freedom (and thus the field theory formalism may prove useful). No multi-factor HJM models exist that can capture the complete term structure for general volatility structures. Finite-factor models theoretically allow for strange hedging possibilities and their inability to parsimoniously incorporate a non-trivial correlation between forward rates of different maturities. From theoretical descriptions of stochastic strings and phenomenological studies of the interest rate curve (see [14]), we managed to construct a Lagrangian for the forward interest rate field $f(t,x)$ in eq. (5.4) and, by a change of variables, also for the velocity field $\mathcal{A}(t,x)$ in eq. (5.13). The field Lagrangians are characterized by a single rigidity parameter μ, which specifies how much the field theory deviates from a regular one-factor HJM model. We indicated that, the harmonic oscillator is equivalent to the field theory of forward rates and that its propagator can thus be recycled (up to some parameter renaming). We showed that, taking the limit $\mu \to 0$, the propagator reduces to 1 and, by extension, the field theory to the one-factor HJM model. The no-arbitrage condition on the drift term for the theory was rederived in a field theoretical setting which incorporated the propagator as shown in eq. (5.34). Before tending to the ZCB option pricing formula for one last time, we discussed the concept of changing the numeraire in field theory in which we explained how a ZCB with a useful maturity can be utilized as a discounting tool and how this change established a relation between two actions. This relationship was given in eq. (5.47). Finally, we rederived the field theory price for the ZCB option under this new choice of numeraire. The result was again a "Black-Scholes"-like pricing formula, now including the effects of the non-trivial correlations over different maturities induced by the propagator. The closed-form field theory pricing formula for a ZCB option of expiry t_*, strike K and bond maturity s is given at time t_0 as

$$\mathrm{ZCB}C(t_0; K, t_*, s) = P(t_0, t_*)\left[F\Phi(d_+) - K\Phi(d_-)\right],$$

where Φ is the standard normal cumulative distribution function and

$$\begin{aligned} P(t_0,t_*) &= e^{-\int_{t_0}^{t_*}\mathrm{d}x f(t_0,x)}, \\ F &\equiv e^{-\int_{t_*}^{s}\mathrm{d}x f(t_0,s)} \\ d_\pm &= \frac{1}{q}\left[\ln\frac{F}{K} \pm \frac{q^2}{2}\right], \\ q^2 &\equiv \int_{t_0}^{t_*}\mathrm{d}t\int_{t_*}^{s}\mathrm{d}x\,\mathrm{d}x'\,\sigma(t,x)G(x,x';t)\sigma(t,x'), \end{aligned}$$

for a propagator

$$G(x,x';t,t+T_{FR}) = \mu\frac{\cosh[\mu(T_{FR}-|x-x'|)] + \cosh[\mu(T_{FR}-(x+x'-2t))]}{2\sinh\mu T_{FR}},$$

with T_{FR} to maximum future time for which ZCBs are observed in the market.

Having established all the necessary theory for the valuation of a ZCB in the context of a forward interest rate's field, we set up two possible tests. This was done using yields to maturity and caplet data provided by KBC Bank. By constructing a cubic spline interpolation and using the relations between the different types of interest rates, we managed to approximate the instantaneous forward interest rates and the corresponding surface was displayed in fig. 6.3. Using this interpolated instantaneous forward interest rates data we suggested two tests to verify the validity and/or usefulness of the field theory for interest rates. The first one was a brief visual comparison of the non-trivial market correlation between the forward rates and the correlation induced by the propagator. The two surfaces were displayed next to each other in fig. 6.6. Clearly, the field theory has more structure and resemblance in its correlation surface than the exact correlations as suggested by the one-factor HJM models. The second test was a numerical calibration. We used the data to compute the theoretical price of put options on a ZCB (see eq. (6.4)), which are related to the caplet prices by eq. (6.1). This relation lies at the basis of the calibration and allowed us to optimize the parameter values to get the best possible fit. The numerical results were given in table 6.2 and the model-predicted prices were fitted against the market data in fig. 6.7. We saw that the field theory had the ability to valuate the caplets with a very high accuracy (RMSE=0.01598), comparable to and even slightly better than that of a Vasicek-like one-factor HJM model (RMSE=0.01646). However, the field theory model has far greater structure in its correlation surface which would make it a lot more useful in pricing more complex interest rate derivatives, such as, for example, spread options on a CMS that take correlation into account. Nevertheless, the correlation surface induced by the field theory's propagator still has some major shortcomings, such as its non-smooth nature on the diagonal $\theta = \theta'$, which is clearly visible in fig. 5.3b. Some suggestions for plausible improvements are discussed in the next section.

7.2 Further Prospects

To conclude this thesis, we discuss some suggestions on how to improve on the field theory for forward interest rates studied here and how this field theory approach can be deployed in settings other than the simple ZCBs. None of these suggestions will be worked out in detail but we rather sketch the main ideas.

Non-Constant Parameters

A first thought could be to make the rigidity μ, and/or volatility σ time- or rather "time-to-maturity-dependent", i.e. $\mu(\theta)$ and/or $\sigma(\theta)$. One could either work with an educated guess on the form of these functions or, a somewhat simpler approach is to take these functions to be piece-wise constant. This way, μ and/or σ would take a specific constant value for every available market tenor. The pricing accuracy of the model would be greatly improved and even exact for the caplet market prices used as calibration instruments. There would, however, be almost no difference with an adaptation of the Vasicek-like one-factor HJM model, where we also make ρ and k time-dependent and piece-wise constant in nature as this would then also match the prices of the calibration instruments exactly. Even the constant volatility one-factor HJM model with a piece-wise constant volatility would match the prices exactly. This is because one would always plug in the so-called *implied volatility*, which is the volatility one needs to plug into a pricing formula to obtain the observed market price [59]. Hence, there would be no discrepancy between the theoretical prices for the market instruments but there would be a difference in prices of other derivatives valuated with the three models.

To further indulge the idea of non-constant volatility, we could take it one step further and let the volatility be a stochastic quantity. In [3], the volatility $\sigma(t, x)$ is introduced as an independent stochastic field. On the contrary to interest rates nowadays, volatility is in fact always positive. For this reason, another field $h(t, x)$ is introduced such that

$$\sigma(t, x) = \sigma_0 e^{-h(t,x)} \quad \text{for } h(t, x) \in [-\infty, \infty]. \tag{7.1}$$

The system thus consists of two interacting fields $f(t, x)$ and $h(t, x)$. It is argued that the corresponding Lagrangian of such a system should maintain certain features:

- The Lagrangian should contain a parameter ξ quantifying the stochasticity of $h(t, x)$. As $\xi \to 0$, $h(t, x)$ reduces to a deterministic function.

- The Lagrangian should contain a parameter κ which is the volatility's counterpart of μ, governing the fluctuations of $\sigma(t,x)$ in the x-direction.
- The Lagrangian should contain a parameter $\rho \in [-1,1]$ quantifying the correlation between $f(t,x)$ and $h(t,x)$
- The volatility has a (risk-neutral) drift term $\beta(t,x)$, completely analogous to $\alpha(t,x)$.

From this, a *possible*[1] Lagrangian is suggested to be of the form

$$\begin{aligned}\mathcal{L}[f,h] = &-\frac{1}{2(1-\rho^2)}\left[\frac{\frac{\partial f}{\partial t}-\alpha}{\sigma} - \rho\frac{\frac{\partial h}{\partial t}-\beta}{\xi}\right]^2 - \frac{1}{2}\left(\frac{\frac{\partial h}{\partial t}-\beta}{\xi}\right)^2 \\ &- \frac{1}{2\mu^2}\left[\frac{\partial}{\partial x}\left(\frac{\frac{\partial f}{\partial t}-\alpha}{\sigma}\right)\right]^2 - \frac{1}{2\kappa^2}\left[\frac{\partial}{\partial x}\left(\frac{\frac{\partial h}{\partial t}-\beta}{\xi}\right)\right]^2,\end{aligned} \tag{7.2}$$

where the (t,x)-dependence is left implicit for notational simplicity. The corresponding action reads

$$\mathcal{S}[f,h] = \int_{\mathcal{F}} \mathcal{L}[f,h]. \tag{7.3}$$

The necessary boundary conditions for the forward rates field $f(t,x)$ remain the same as first specified in section 5.1, while those for $\sigma(t,x)$ are given by $\sigma(t_i,x)$ and $\sigma(t_f,x)$ for $t=t_i$ and $t=t_f$ respectively and, on the other hand, we have

$$\frac{\partial}{\partial x}\left[\frac{\partial h(t,x)}{\partial t} - \beta(t,x)\right], \tag{7.4}$$

for $t_i \le t \le t_f$ and $x=t$ or $x=t+T_{FR}$. The integration measure and partition function are given rather straightforward by

$$\begin{aligned}\int \mathcal{D}[f]\mathcal{D}[\sigma^{-1}] &= \prod_{(t,x)\in\mathcal{F}} \int_{-\infty}^{\infty} \mathrm{d}f(t,x)\,\mathrm{d}\sigma^{-1}(t,x) && (7.5)\\ &= \prod_{(t,x)\in\mathcal{F}} \int_{-\infty}^{\infty} \mathrm{d}f(t,x)\,\mathrm{d}h(t,x)\,e^{h(t,x)}, && (7.6)\end{aligned}$$

and

$$Z = \int \mathcal{D}f\mathcal{D}\sigma^{-1}e^{\mathcal{S}[f,h]}. \tag{7.7}$$

Such a model could be useful in valuating more exotic interest rate derivatives.

Higher Order Terms

Apart from imposing a time-dependence on the volatility or rigidity, one could also include higher order terms in the original Lagrangian in eq. (5.4). Including a term with a higher order derivative in x would govern the *stiffness* of fluctuations in $x \mapsto f(t,x)$. This "stiffness" would further quantify the correlations of the shocks that points on the curve at nearby maturities experience. Such a system was studied in [2] and the action reads

$$\mathcal{S}[\mathcal{A}] = \int_{\mathcal{F}}\left[\mathcal{A}(t,x)^2 + \frac{1}{\mu^2}\left(\frac{\partial\mathcal{A}(t,x)}{\partial x}\right)^2 + \frac{1}{\lambda^4}\left(\frac{\partial^2\mathcal{A}(t,x)}{\partial x^2}\right)^2\right] \tag{7.8}$$

[1] We say "*possible*" because the Lagrangian for such a system obeying all the stated features is not unique [3].

It is shown that the extra "stiffness term" actually eliminates the *kink* in the propagator on the diagonal but to be fully consistent with market data, μ and λ can not be considered constant in x. This thus allows for a better description of the correlation surface observed in the market. This non-constant nature is explained by the idea that market participants do not perceive the time-to-maturity in a uniform manner.

Perturbation Expansion

Another natural question to ponder is whether or not perturbation expansions can be practical in this story. And in fact, as discussed in [4, 6], one can evaluate the price of a (zero) coupon bond option by considering a perturbation around the *forward coupon bond price.* Now, the price of a coupon bond $\mathcal{B}(t, s_N)$, that pays N coupons a_i at times s_i and a principal L at time s can be given as a portfolio of ZCBs

$$\mathcal{B}(t,s) = \sum_{i=1}^{N} c_i P(t, s_i),$$

where $c_i = a_i$ for $i = 1, \ldots, N-1$, $c_N = a_N + L$ and $s_N = s$. This can be rewritten as

$$\begin{aligned}\mathcal{B}(t,s) &= \sum_{i=1}^{N} c_i F_i + \sum_{i=1}^{N} c_i [P(t,s_i) - F_i] \\ &= F + V,\end{aligned}$$

with

$$F_i = e^{-\int_t^{s_i} f(t,x)\mathrm{d}x} \quad \text{and} \quad V \equiv \sum_{i=1}^{N} c_i [P(t,s_i) - F_i]. \tag{7.9}$$

The term V acts as a potential in the effective action. The pay function of a call option with expiry $t_* \leq s$ and strike K reads

$$\begin{aligned}[\mathcal{B}(t,s) - K]^+ &= (F + V - K)^+ & (7.10)\\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\tilde{V}\, \mathrm{d}J\, e^{\mathrm{i}J(V-\tilde{V})}(F + \tilde{V} - K)^+, & (7.11)\end{aligned}$$

where the exponential expression for the Dirac delta $\delta(V - \tilde{V})$ was used. The call option is thus given as

$$\mathrm{ZCBC}(t_0; K, t_*, s) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\tilde{V}\, \mathrm{d}J\, e^{\mathrm{i}J(V-\tilde{V})}(F + \tilde{V} - K)^+ \mathcal{Z}[J], \tag{7.12}$$

with

$$\mathcal{Z}[J] = \frac{1}{Z}\int \mathcal{D}[f] e^{\mathcal{S}_{\text{eff}}} \tag{7.13}$$

and

$$\mathcal{S}_{\text{eff}} = \mathcal{S}[\mathcal{A}] - \int_{t_0}^{t_*} r(t)\, \mathrm{d}t + \mathrm{i}JV. \tag{7.14}$$

The idea is now to expand $\mathcal{Z}[J]$ in terms of J considering σ is small ($\approx 10^{-2}$ [4]). The expansion coefficients are given as

$$G_{ij} \equiv \int_{t_0}^{t_*} \mathrm{d}t \int_{t_*}^{s_i} \mathrm{d}x \int_{t_*}^{s_j} \mathrm{d}x'\, \sigma(t,x) G(x,x';t) \sigma(t,x') \tag{7.15}$$

and can be represented by specific diagrams. This procedure allows for a perturbative expanded price for the coupon bond option (and more complex interest rate derivatives) and yields easily calculated price approximations.

Other Interest Rate Derivatives

Beside the coupon bonds and coupon bond options, there are many other interest rate derivatives that are traded in high volumes such as *swaps, accordion swaps, swaptions, Bermudan swaptions, caps and floors, CMS, CMS spreads,...* [59]. It is however the case that, for example, swaps can also be constructed from a portfolio of ZCBs. This way, their pricing formulas can be derived from the work done in section 5.5. An interesting research project might be to build a valuation framework for these derivatives in the context of a field theory for forward interest rates and see how the known pricing formulas are altered by the presence of the non-trivial correlation induced by the theory's propagator. This would be especially interesting for interest rate derivatives with and explicit dependence on the correlation between forward rates of different maturities, such as a CMS spread option. For further information on exotic interest rate derivatives we refer to [15, 30, 59].

With these briefly explained outlooks and suggestions, we hope to convince the reader that there is much more to this field theory description of interest rates and their derivatives beyond the scope of this work.

Bibliography

[1] Robert A Adams and Christopher Essex. *Calculus: A Complete Course, with My Matlab*. Pearson, 2013.

[2] Belal Baaquie and Jean-Philippe Bouchaud. Stiff field theory of interest rates and psychological future time. *arXiv preprint cond-mat/0403713*, 2004.

[3] Belal E Baaquie. Quantum field theory of forward rates with stochastic volatility. *Physical Review E*, 65(5):056122, 2002.

[4] Belal E Baaquie. Feynman perturbation expansion for the price of coupon bond options and swaptions in quantum finance. i. theory. *Physical Review E*, 75(1):016703, 2007.

[5] Belal E Baaquie. *Quantum finance: Path integrals and Hamiltonians for options and interest rates*. Cambridge University Press, 2007.

[6] Belal E Baaquie and Cui Liang. Feynman perturbation expansion for the price of coupon bond options and swaptions in quantum finance. ii. empirical. *Physical Review E*, 75(1):016704, 2007.

[7] Louis Bachelier. *Theory of Speculation*. Princeton University Press, 1900.

[8] Martin Baxter, Andrew Rennie, and Andrew JO Rennie. *Financial calculus: an introduction to derivative pricing*. Cambridge university press, 1996.

[9] Fischer Black, Emanuel Derman, and William Toy. A one-factor model of interest rates and its application to treasury bond options. *Financial analysts journal*, 46(1):33–39, 1990.

[10] Fischer Black and Piotr Karasinski. Bond and option pricing when short rates are lognormal. *Financial Analysts Journal*, 47(4):52–59, 1991.

[11] Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *Journal of political economy*, 81(3):637–654, 1973.

[12] Stephen Blyth. *An introduction to quantitative finance*. Oxford University Press, 2013.

[13] Jean-Philippe Bouchaud and Marc Potters. *Theory of financial risk and derivative pricing: from statistical physics to risk management*. Cambridge university press, 2003.

[14] Jean-Philippe Bouchaud, Nicolas Sagna, Rama Cont, Nicole El-Karoui, and Marc Potters. Phenomenology of the interest rate curve. *Applied Mathematical Finance*, 6(3):209–232, 1999.

[15] Damiano Brigo, Fabio Mercurio, et al. *Interest rate models: theory and practice*, volume 2. Springer, 2001.

[16] Alberto Bueno-Guerrero, Manuel Moreno, and Javier F Navas. Stochastic string models with continuous semimartingales. *Physica A: Statistical Mechanics and its Applications*, 433:229–246, 2015.

[17] Alberto Bueno-Guerrero, Manuel Moreno, and Javier F Navas. The stochastic string model as a unifying theory of the term structure of interest rates. *Physica A: Statistical Mechanics and its Applications*, 461:217–237, 2016.

[18] Andrew JG Cairns. *Interest rate models: an introduction*, volume 10. Princeton University Press, 2004.

[19] Masud Chaichian and Andrei Demichev. *Path integrals in physics: Volume I stochastic processes and quantum mechanics*. CRC Press, 2018.

[20] Masud Chaichian and Andrei Demichev. *Path Integrals in Physics: Volume II Quantum Field Theory, Statistical Physics and other Modern Applications*. CRC Press, 2018.

[21] ByoungSeon Choi and MY Choi. General solution of the black–scholes boundary-value problem. *Physica A: Statistical Mechanics and its Applications*, 509:546–550, 2018.

[22] Rama Cont. Modeling term structure dynamics: an infinite dimensional approach. *International Journal of theoretical and applied finance*, 8(03):357–380, 2005.

[23] John C Cox, Jonathan E Ingersoll Jr, and Stephen A Ross. A theory of the term structure of interest rates. In *Theory of valuation*, pages 129–164. World Scientific, 2005.

[24] Ashok Das. *Field theory: a path integral approach*, volume 83. World Scientific, 2006.

[25] Yasar Demirel. *Nonequilibrium thermodynamics: transport and rate processes in physical, chemical and biological systems*. Elsevier, 2007.

[26] Richard Phillips Feynman and Laurie M Brown. *Feynman's thesis: a new approach to quantum theory*. World Scientific, 2005.

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

[28] The Nobel Committee for Physics. For groundbreaking contributions to our understanding of complex physical systems. 2021.

[29] Robert S Goldstein. The term structure of interest rates as a random field. *The Review of financial studies*, 13(2):365–384, 2000.

[30] Pierre Hanton and Marc Henrard. Cms spread options and similar options in multi-factor hjm framework. *Available at SSRN 1604389*, 2010.

[31] Uwe Hassler et al. Stochastic processes and calculus. *Springer Texts in Business and Economics*, 2016.

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

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

[34] John C. Hull. *Options, Futures and other Derivatives*, volume 8. Pearson, 2012.

[35] John C. Hull and Alan White. Pricing interest-rate-derivative securities. *The review of financial studies*, 3(4):573–592, 1990.

[36] Claude Itzykson and Jean-Michel Drouffe. *Statistical Field Theory Volume 1: From Brownian motion to renomalization and lattice gauge theory.* Cambridge University Press, 1989.

[37] Jessica James and Nick Webber. *Interest Rate Modelling.* Wiley, 2000.

[38] Mehran Kardar. *Statistical physics of fields.* Cambridge University Press, 2007.

[39] Douglas P Kennedy. The term structure of interest rates as a gaussian random field. *Mathematical Finance*, 4(3):247–258, 1994.

[40] Robert L Kimmel. Modeling the term structure of interest rates: A new approach. *Journal of Financial Economics*, 72(1):143–183, 2004.

[41] Aaron Klug. Rosalind franklin and the discovery of the structure of dna. *Nature*, 219(5156):808–810, 1968.

[42] Daragh McInerney and Tomasz Zastawniak. *Stochastic interest rates.* Cambridge University Press, 2015.

[43] Bank of England. What are negative interest rates, and how would they affect me? https://www.bankofengland.co.uk/explainers/what-are-negative-interest-rates, Sep 2021.

[44] Lars Onsager and Stefan Machlup. Fluctuations and irreversible processes. *Physical Review*, 91(6):1505, 1953.

[45] Hugh Osborn. Advanced quantum field theory lent term lecture notes 2013, 2013.

[46] Matthias Otto. Using path integrals to price interest rate derivatives. *arXiv preprint cond-mat/9812318*, 1998.

[47] Michael Peskin and Daniel Schroeder. *An introduction to quantum field theory.* CRC press, 2018.

[48] Riccardo Rebonato. *Interest Rate Option Models.* Wiley, 1998.

[49] Pedro Santa-Clara and Didier Sornette. The dynamics of the forward interest rate curve with stochastic string shocks. *The Review of Financial Studies*, 14(1):149–185, 2001.

[50] Wim Schoutens. *Lévy processes in finance: pricing financial derivatives.* Wiley Online Library, 2003.

[51] Wim Schoutens. *Fundamentals of Financial Mathematics.* KULeuven, 2021.

[52] Didier Sornette. String formulation of the dynamics of the forward interest rate curve. *The European Physical Journal B-Condensed Matter and Complex Systems*, 3:125–137, 1998.

[53] Didier Sornette. *Why stock markets crash: critical events in complex financial systems.* Princeton university press, 2009.

[54] H Eugene Stanley and Rosario N Mantegna. *An introduction to econophysics.* Cambridge University Press, Cambridge, 2000.

[55] Nicolaas Godfried Van Kampen. *Stochastic processes in physics and chemistry*, volume 1. Elsevier, 1992.

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

[57] Johannes Voit and Robert W Lourie. The statistical mechanics of financial markets. *Physics today*, 55(8):8–52, 2002.

[58] Paul Wilmott. *Frequently asked questions in quantitative finance*. John Wiley & Sons, 2010.

[59] Paul Wilmott. *Paul Wilmott on quantitative finance*. John Wiley & Sons, 2013.

[60] Horacio S Wio. *Path integrals for stochastic processes: An introduction*. World Scientific, 2013.

[61] Volker Ziemann. *Physics and Finance*. Springer, 2021.

[62] Jean Zinn-Justin. *Quantum field theory and critical phenomena*, volume 171. Oxford university press, 2021.

Institute for Theoretical Physics
Celestijnenlaan 200D
3001 LEUVEN, BELGIË
tel. + 32 16 32 72 32
www.kuleuven.be