D I P L O M A R B E I T

Stochastic Volatility in Fixed-Income Markets (Version 1.0.1)

ausgeführt am Institut für

Wirtschatfsmathematik

der Technischen Universität Wien

unter Anleitung von Ao. Univ. - Prof. Dr. Josef Teichmann

durch

Richard Warnung

Leebgasse, 7/18

Wien - Favoriten

Dedicated to oh so many people - I was not able to pick a small number. Certainly among them: my family that supported me so much and still does, my friends who accompanied me on my way so far, some of my lecturers who showed me the beauty of using one's reason and that it might be worth all the trouble, Ernest Hemingway and Charles Bukowski for making me read other stuff besides math books and to Daniela who helps me to relax in these weeks and months in which I write this ... and thanks to Josef for his never ending motivation.

Abstract

My diploma thesis which I wrote under the supervision of Prof. Teichmann in the winter term 2004/2005 covers several topics in stochastic analysis and interest rate theory. The first chapter is dedicated to an introduction to stochastic analysis. Brownian motion is introduced and the construction of the Itô integral is described. The Stratonovich integral is defined and the interpretation and connection of both integrals is explained. The text continues with Itô's formula. Then a deep analysis of conditions under which existence and uniqueness of the solution of a stochastic differential equation can be guaranteed follows. The text covers the case of Lipschitz continuity as well as Hölder continuity of the volatility term.

The text continues with the introduction of the generator of an Itô diffusion and one of its applications - the Dynkin formula - which is applied to the analysis of recurrence and transience of Brownian motion in dimension n. We show that an Itô diffusion has the Markov property and finish the introductory part with Girsanov's theorem and the martingale representation theorem which have many applications in interest rate theory.

The second chapter starts with an introduction to arbitrage-free pricing of zero-coupon bonds and various ideas of interest rates. For the main purpose of this text - the pricing of caps and floors - the idea of LIBOR rates is crucial. Therefore we continue with a definition of the LIBOR rate for a certain period. We construct an arbitrage-free model for a LIBOR rate market in a discrete as well as in a continuous tenor setting.

The main chapter of this text is the chapter on interest rate derivatives. It starts with the definition of swap rates, caps and floors and captions and gives formulas of their value. Those contracts can be seen as European style options. We therefore continue with Black's formula for European options on bonds. We then define and illustrate the notion of implied volatility. This notion gives us a new measure of accuracy of our models and shows us the need to develop and analyze more sophisticated models such as the stochastic volatility model. We start following the ideas of Andersen and Brotherton-Ratcliffe. They introduce a model of a LIBOR rate whose variance process is modeled by a mean reverting stochastic differential equation in the form used in the Cox-Ingersoll-Ross model. The Brownian motions driving the SDE for the LIBOR rate and the SDE for the variance process respectively are assumed to be uncorrelated. They therefore start to develop the price of a cap with constant variance first. We cite their result of a partly explicit formula up to order 3 of $\sqrt{T-t}$. Andersen and Brotherton-Ratcliffe then start to define a certain Laplace transform of the variance process and develop it in terms of the volatility of variance parameter up to order 4. We now analyze the Laplace transform in a different way. For integer dimension of the CIR-model we find the representation as element of the second Wiener - Itô chaos. We then use a formula from quantum field theory to find a closed expression for its

Laplace transform where a Carleman-Fredholm determinant appears. We finally want to motivate this approach in order to price the cap in the stochastic volatility model. The appendix provides theorems and proofs of various fields of probability theory such as the martingale problem and the chaos expansion of an L^2-random variable. An introduction of Bessel processes and the connection to the CIR-model can be found in detail. Version 1.01 is simply edited with some minor corrections in notation and expression.

Contents

1 Introduction to Stochastic Analysis **3**

1.1 The Brownian Motion . . . 3

1.1.1 The Wiener Process . . . 3

1.1.2 Brownian Motion, Definition and Properties . . . 4

1.2 Stochastic Integration . . . 7

1.3 The Itô Interpretation . . . 9

1.3.1 The Itô Integral . . . 9

1.3.2 Itô's formula . . . 13

1.4 The Stratonovich Interpretation . . . 15

1.5 Stochastic Differential Equations . . . 17

1.5.1 Existence and Uniqueness of the Solution of an SDE . . . 17

1.5.2 Differential Operators . . . 24

1.5.3 Notation . . . 25

1.5.4 Solution Methods . . . 25

1.5.5 Supplements on the Existence and Uniqueness Theorem . . . 26

1.6 Properties of the solution of an SDE . . . 28

1.6.1 What a Diffusion Really Does . . . 28

1.6.2 The Generator of an Itô Process . . . 29

1.6.3 The Dynkin Formula . . . 32

1.6.4 The Markov Property . . . 34

1.7 Girsanov's Theorem . . . 36

1.8 Martingale Representation Theorem . . . 37

2 Interest Rate Theory **40**

2.1 Zero-coupon Bonds, Forward Rates and the Short Rate . . . 40

2.1.1 Zero-Coupon Bonds . . . 40

2.1.2 Term Structure of Interest Rates . . . 41

2.1.3 Forward Interest Rates . . . 41

2.1.4 Short-term Interest Rate . . . 42

2.1.5 Arbitrage-free Pricing of Zero-coupon Bonds . . . 42

2.2 LIBOR Rates . . . 47

2.2.1 The Mathematical Setting . . . 47

2.2.2 Definition of Spot and Forward Martingale measure . . . 50

2.2.3 Arbitrage-free Properties and the Implied Savings Account . . . 50
2.2.4 Bond Price Volatility . . . 53
2.2.5 Forward LIBOR Rates . . . 54
2.2.6 The Discrete Tenor Model . . . 55
2.2.7 The Continuous Tenor Case . . . 58

3 Interest Rate Derivatives **61**
3.1 Forward Swap Rates . . . 61
3.1.1 A Swap Contract . . . 61
3.1.2 Interest Rate Swaps . . . 62
3.1.3 Model of Forward Swap Rates . . . 64
3.2 More Interest Rate Derivatives . . . 68
3.2.1 Caps and Floors . . . 68
3.2.2 Captions . . . 69
3.2.3 Swaptions . . . 69
3.3 Option Valuation . . . 69
3.3.1 The Arbitrage Price of a Contingent Claim . . . 70
3.3.2 A Version of Black's Formula for Bond Options . . . 70
3.4 Black's Formula for Caplets and the Volatility Smile . . . 73
3.4.1 Implied Volatility . . . 73
3.4.2 Volatility-Smile . . . 73
3.5 The Stochastic Volatility Model . . . 74
3.5.1 The Observed Model . . . 75
3.5.2 An Asymptotic Solution for V constant . . . 76
3.5.3 Stochastic V . . . 78

4 Appendix **82**
4.1 Proofs of Classical Theorems . . . 82
4.2 Theorems from Measure Theory . . . 85
4.3 Wiener-Itô Chaos Expansion . . . 87
4.4 Mathematical Tools . . . 89
4.4.1 Solution Methods of ODE's . . . 89
4.4.2 Lévy's Criterion for Brownian Motion . . . 90
4.4.3 The Martingale Problem . . . 90
4.4.4 Random Time Change . . . 93
4.4.5 Bessel Processes . . . 95
4.4.6 The Connection between Bessel processes and the CIR model . 97
4.4.7 The Carleman-Fredholm determinant . . . 98

Chapter 1

Introduction to Stochastic Analysis

1.1 The Brownian Motion

Most of the theorems and definitions are taken from Teichmann [16] and Øksendal [13].

1.1.1 The Wiener Process

The Wiener process is named after the mathematician Norbert Wiener and it is defined as follows:

Definition 1.1.1 (The Wiener Process). *A stochastic process $(W_t)_{t\geq 0}$ is called Wiener process if it is Gaussian and if*

$$\begin{aligned} \mathbf{E}[W_t] &= 0 \\ \mathbf{E}[W_t W_s] &= s \wedge t. \end{aligned} \tag{1.1}$$

Theorem 1.1.1 (Existence of the Wiener Process). *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a sequence of i.i.d. (independent identically distributed) random variables $(X_n)_{n\geq 0}$ such that each X_n is N(0,1)-distributed then there is a Wiener process $(W_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, \mathbb{P})$.*

Proof.

- First choose an orthonormal basis $(e_i)_{i\geq 0}$ for the Hilbert space $L^2(\mathbb{R}_{\geq 0}, \mathcal{B}(\mathbb{R}_{\geq 0}), dx)$.
- Then we define a unique isometry η such that $\eta(e_i) = X_i$ for $i \geq 0$.
- We have to remark that the span of the X_i is a space of Gaussian random variables and obviously every finite selection is multinormally distributed.
- Therefore $\eta : L^2(\mathbb{R}_{\geq 0}, \mathcal{B}(\mathbb{R}_{\geq 0}), dx) \to L^2(\Omega, \mathcal{F}, \mathbb{P})$ is a well defined linear map with the isometry property

$$\mathbf{E}[\eta(f)\eta(g)] = \int_0^\infty f(s)g(s)ds.$$

- We now define $W_t := \eta(\mathbf{1}_{[0,t]})$ for $t \geq 0$.
- The resulting process then is Gaussian by definition of η satisfying that $\mathbf{E}[W_t] = 0$ and by the isometry

$$\mathbf{E}[W_t W_s] = \int_0^\infty \mathbf{1}_{[0,t]}(u)\mathbf{1}_{[0,s]}(u)du = s \wedge t,$$

and therefore a Wiener process.

□

1.1.2 Brownian Motion, Definition and Properties

Definition 1.1.2 (A Version of a Process). *A stochastic process X_t is a version of another process Y_t if $\forall t$ $\mathbb{P}(X_t = Y_t) = 1$. So they coincide outside of a null set for each t.*

Definition 1.1.3 (Indistinguishability). *A stochastic process X_t is indistinguishable from another process Y_t if $\mathbb{P}(\forall t, X_t = Y_t) = 1$. So they coincide outside of one null set.*

Definition 1.1.4 (Usual Conditions). *A filtration of a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is said to fulfill the usual conditions if*

- $\mathcal{F}_{t+} = \bigcap_{s>t} \mathcal{F}_s = \mathcal{F}_t$ *(right continuity)*
- $\mathcal{F}_t$ *is complete meaning that it contains all $\mathbb{P}$-null sets.*

Definition 1.1.5 (Martingale). *A process M_t is a martingale if it fulfills the following properties:*

1. *M_t is adapted to $\mathcal{F}_t$.*
2. *M_t is integrable.*
3. $\mathbf{E}[M_t|F_s] = M_s, t \geq s$.

Definition 1.1.6 (Stopping Time). *Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a filtration $(\mathcal{F}_t)_{t\geq 0}$ then a random variable $\tau : \Omega \longrightarrow \mathbb{R}_{\geq 0}$ is called a stopping time with respect to $(\mathcal{F}_t)_{t\geq 0}$ if the set of events $\{\tau \leq t\} \in \mathcal{F}_t \quad \forall t$.*

Definition 1.1.7 (Local Martingale). *A process M_t is a local martingale if and only if there exists a sequence of stopping times T_n tending to infinity such that $M_{t\wedge T_n}$ - the stopped processes - are martingales for all n.*

Remark 1.1.1. *Note that a stochastic process M_t is called bounded if there exists a constant K such that for all ω and $t \geq 0$ $|M_t| \leq K$ holds.*

Theorem 1.1.2. *A bounded local martingale is a martingale.*

Proof. Let T_n be the sequence of stopping times of the definition of the local martingale M_t. Then T_n tends to infinity and $M_{t\wedge T_n}$ tends to M_t pointwise.
Moreover $\lim_{n\to\infty}\mathbf{E}[M_{t\wedge T_n}|\mathcal{F}_s] = \mathbf{E}[M_t|\mathcal{F}_s]$ by dominated convergence (the constant can be taken as dominating function). By definition $M_{t\wedge T_n}$ is a martingale and so $\mathbf{E}[M_{t\wedge T_n}|\mathcal{F}_s] = M_{s\wedge T_n}$. Putting all this together we get

$$\mathbf{E}[M_t|\mathcal{F}_s] = \lim_{n\to\infty}\mathbf{E}[M_{t\wedge T_n}|\mathcal{F}_s] = \lim_{n\to\infty} M_{s\wedge T_n} = M_s$$

which completes the proof. □

Definition 1.1.8 (Brownian Motion)**.** *An adapted process $(B_t)_{t\in\mathbb{R}_{\geq 0}}$ is called Brownian Motion if it fulfills the following properties:*

1. *$B_0 = 0$ almost surely (a.s.).*
2. *The increment $B_t - B_s$ is independent of $\mathcal{F}_s$ (the filtration at time s) for $t > s$.*
3. *$B_t - B_s$ is normally distributed with mean zero and variance $t-s$. The distribution is stationary.*
4. *There is a continuous version. Continuous in the sense that the map $t \to B_t(\omega)$ is continuous almost surely.*

With these properties one can show that the process B_t is a martingale with respect to its natural filtration (the filtration generated by the process itself).

$\mathbf{E}[B_t - B_s|\mathcal{F}_s] = \mathbf{E}[B_t - B_s] = 0 \Longrightarrow \mathbf{E}[B_t|\mathcal{F}_s] = B_s$ (by independence and adaptedness).

The question whether Brownian Motion exists is answered by the Kolmogorov-Centsov Theorem (see Øksendal [13]) (we have a Wiener process which is moreover adapted and by Kolmogorov-Centsov we have continuity).

Definition 1.1.9 (Quadratic Variation)**.** *The quadratic variation Q_t of a stochastic process $(X_t)_{t\geq 0}$ is defined by*

$$Q_t = \lim_{\Delta\to 0}\sum_{i=0}^{n-1}(X_{t_{i+1}} - X_{t_i})^2$$

where the limit is taken over a partition of $[0,t]$ with mesh $\Delta := \max_{0\leq i\leq n-1}|t_{i+1} - t_i|$ in probability.

Theorem 1.1.3. *The quadratic variation $\sum_{k=0}^{2^n-1}(B_{t\frac{k+1}{2^n}} - B_{t\frac{k}{2^n}})^2$ of a Brownian Motion B_t converges to t in L^2. Thus the total variation $\sum_{k=0}^{n-1}|B_{t_{k+1}} - B_{t_k}|$ of a Brownian Motion is a.s. infinite.*

We will moreover need to define a Brownian Motion in d Dimensions .

Definition 1.1.10. *A process $(B_t)_{t\geq 0}$ with $B_t : \Omega \mapsto \mathbb{R}^d$ is called a d-dimensional Brownian Motion with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$ which satisfies the usual conditions, if the following properties hold:*

- $B_0 = 0$ *(d-dimensional)*
- B_t *is* $\mathcal{F}_t$*-measurable*
- *the process has continuous trajectories a.s. (this means that* $t \mapsto B_t(\omega)$ *is continuous almost surely)*
- $B_t - B_s$ *is independent of* $\mathcal{F}_s, t \geq, s \geq 0$ *and multinormally distributed with expectation 0 and covariance matrix of the form:* $(t-s)\cdot E$ *where* E *denotes the d-dimensional identity matrix.*

Now the question could arise, whether such a process exists. I will show that $B_t = (B_t^1, \ldots, B_t^d)$ where each B_t^i is a 1-dimensional BM is a d-dimensional BM: We first choose the probability space:

- the space $\Omega := \Omega^1 \times \ldots \times \Omega^d$
- the sigma-algebra $\mathcal{F} := \mathcal{F}^1 \bigotimes \ldots \bigotimes \mathcal{F}^d$
- the prob. measure $\mathbb{P} := \mathbb{P}^1 \bigotimes \ldots \bigotimes \mathbb{P}^d$
- the filtration $(\mathcal{F}_t) := \mathcal{F}_t^1 \bigotimes \ldots \bigotimes \mathcal{F}_t^d$

The requirements of the definition are met:

- $B_0 = 0$
- B_t is $\mathcal{F}_t$ measurable since B_t^i is $\mathcal{F}_t^i$ measurable.
- The trajectories are a.s. continuous because they are a.s. continuous in each component.
- The independence of $B_t - B_s$ holds since the measures are product measures on $(\Omega, \mathcal{F}, \mathbb{P})$ and $B_t^i - B_s^i$ is independent of $\mathcal{F}_s^i, s \leq t$
  $B_t^i - B_s^i$ is $N(0, t-s)$ distributed and since $(B_t^i)_{t\geq 0}$ is independent of $(B_t^j)_{t\geq 0}$ we have that $\mathbf{E}[(B_t^i - B_s^i)(B_t^j - B_s^j)] = \delta_{ij}(t-s)$.

Remark 1.1.2. *It is clear that BM is a Wiener process. For a BM one needs moreover the filtration - there is no word mentioned in the definition above about a filtration in the setting of a Wiener process. Finally we have the almost surely continuous paths of BM which are so important in many notions.*

1.2 Stochastic Integration

In applications in finance and other fields of science we need to give a meaning to so called stochastic integrals with respect to Brownian motion. The dynamics of processes are often modeled in the following way:

$$X_t = X_0 + \int_0^t b(s, X_s)ds + \text{"} \int_0^t \sigma(s, X_s)dB_s\text{"}.$$

Thus we need to understand problems and methods of the construction of such a stochastic integral. We will first see where the problems arise.

Lemma 1.2.1 (Moments of Increments of the Brownian Motion)**.** *Let* $(B_t)_{t\geq 0}$ *denote a BM. Then*

$$\mathbf{E}[(B_t - B_s)^{2k}] = \frac{(2k)!}{k!2^k}(t-s)^k \quad . \tag{1.2}$$

Proof. $B_t - B_s$ is normally distributed with mean zero and variance t-s - this follows from the stationary of the distribution ($B_t - B_s$ has the same distribution as $B_{t-s} - B_0$). Therefore we have to calculate the moments of a normally distributed random variable with these parameters. From the characteristic function: $\mathbf{E}[e^{iuX}] = e^{-\frac{u^2}{2}(t-s)}$ we find the result by comparison of the coefficients. □

Theorem 1.2.1 (Theorem of Wiener)**.** *Let* $t > 0$ *be a fixed point in time and* $(B_s)_{s\geq 0}$ *a BM, then*

$$\lim_{n\to\infty} \sum_{i=0}^{2^n-1} (B_{\frac{t(i+1)}{2^n}} - B_{\frac{ti}{2^n}})^2 = t \tag{1.3}$$

almost surely.

Proof. For $n \geq 1$ we define

$$S_n = \sum_{i=0}^{2^n-1} (B_{\frac{t(i+1)}{2^n}} - B_{\frac{ti}{2^n}})^2.$$

One easily sees that (think of $\mathbf{E}[(B_t - B_s)^2] = t - s$ for $t \geq s$):

$$\begin{aligned}
\mathbf{E}[S_n] &= t \\
\mathbf{E}[S_n^2] &= \sum_{i=0}^{2^n-1} \mathbf{E}[(B_{\frac{t(i+1)}{2^n}} - B_{\frac{ti}{2^n}})^4] + \sum_{i\neq j=0}^{2^n-1} \mathbf{E}[(B_{\frac{t(i+1)}{2^n}} - B_{\frac{ti}{2^n}})^2(B_{\frac{t(j+1)}{2^n}} - B_{\frac{tj}{2^n}})^2] \\
&= t^2(3\frac{2^n}{2^{2n}} + \frac{2^n-1}{2^n}) = t^2(\frac{2}{2^n} + 1)
\end{aligned}$$

Further basic calculations lead to

$$\begin{aligned}\mathbf{E}[(S_n - t)^2] &= \mathbf{E}[S_n^2] - 2\mathbf{E}[S_n]t + t^2 = \\ t^2(\frac{2}{2^n} + 1 - 2 + 1) &= \frac{t^2}{2^{n-1}}\end{aligned}$$

for $n \geq 1$. Tchebychev's inequality gives:

$$\mathbb{P}((S_n - t)^2 \geq \frac{1}{2^{\frac{n}{2}}}) \leq 2^{\frac{n}{2}} \frac{t^2}{2^{n-1}} = \frac{1}{2^{\frac{n}{2}}} 2t^2.$$

Using the convergence part of Borel-Cantelli (First Borel Cantelli Lemma) thinking of the fact that the sum $\sum_{i=1}^{\infty} \frac{1}{2^{\frac{i}{2}}}$ converges one sees that the set of events ω for which $(S_n - t)^2 \geq \frac{1}{2^{\frac{n}{2}}}$ for infinitely many n has $\mathbb{P}$-measure 0. Therefore on a set of measure 1 (this is almost surely) we have that $\lim_{n\to\infty} S_n = t$. $\square$

Remark 1.2.1. *Using dyadic partitions of time we were able to assure that the sum converges but every refining sequence of partitions leads to this result.*

Remark 1.2.2. *It is shown directly that the quadratic variation of a BM tends to t for $n \to \infty$ in L^2 (!) by using the independence of the increments, for any sequence of partitions with mesh tending to zero.*

The important task now is to decide how to interpret the integral w.r.t. the Brownian motion. As mentioned before the Brownian motion has a.s. infinite total variation. Thus a pathwise integral in the Lebesgue-Stieltjes sense is not possible. To say it in other more illustrative words: the paths of BM are too rough to define a Lebesgue-Stieltjes integral w.r.t. BM. Consider an integral of the form $\int_S^T f_t dB_t$. If we assume that f has the following form:

$$\phi_t = \sum_{j\geq 0} e_j \cdot \mathbf{1}_{(j\cdot 2^{-n},(j+1)\cdot 2^{-n}]}(t). \tag{1.4}$$

Then there is more than only one interpretation of

$$\int_S^T \phi_t dB_t = \sum_{j\geq 0} e_j (B_{t_{j+1}} - B_{t_j}) \tag{1.5}$$

where the $t_k's$ are the points in time defined in the equation above. The procedure to deal with such integrals is to approximate the function f by $\sum_j f(t_j^*) \cdot \mathbf{1}_{(t_j,t_{j+1}]}(t)$ where the points t_j^* belong to the interval $[t_j, t_{j+1}]$. Then we define the integral as the limit of $\sum_j f(t_j^*)[B_{t_{j+1}} - B_{t_j}]$ in a certain sense that will be explained later. So far I just want to mention that different choices of the t_j^*'s lead to different results.

- $t_j^* = t_j$ (the left end point), which leads to the Itô integral denoted by $\int_S^T f(t)dB_t$.

Summary 1.1. *We can summarize that it is useful to define an integral w.r.t. BM but it can not be defined pathwise because the paths of BM are too rough. So we need a new interpretation.*

1.3 The Itô Interpretation

1.3.1 The Itô Integral

This section is based on a lecture given by Josef Teichmann: Stochastic Analysis [16].

Itô decided to use the evaluation at the left end point of the interval to define the approximating step function. By doing this he achieves that the value of a function adapted to the Brownian filtration is independent of the increment of the Brownian motion. This will become clear with more formalism.

Definition 1.3.1 (A Progressively Measurable Process). *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(\mathcal{F}_t)_{t\geq 0}$ a filtration. Then a stochastic process $(\phi_t)_{t\geq 0}$ is progressively measurable if:*

- $\phi : \Omega \times \mathbb{R}_{t\geq 0} \to \mathbb{R}^d$ *is measurable with respect to* $\mathcal{F} \bigotimes \mathcal{B}(\mathbb{R}_{t\geq 0})$ *and*
- $\phi : \Omega \times [0, t] \to \mathbb{R}^d$ *is measurable w.r.t.* $\mathcal{F}_t \bigotimes \mathcal{B}([0, t])$

Where $\mathcal{B}$ denotes the Borel-sigma-algebra. The last property can be seen as: $\phi_s \cdot \mathbf{1}_{[0,t]}(s)$ lies in $L^2(\Omega \times [0, t], \mathcal{F}_t \bigotimes \mathcal{B}([0, t], dt \bigotimes \mathbb{P})$.

Additionally we need an integrability property:

$$\mathbf{E}[\int_0^\infty \phi(s)^2 ds] = \int_\Omega \int_0^\infty \phi(s, \omega)^2 ds d\mathbb{P}(\omega) < \infty.$$

Example 1.1. *To give an illustrative example for a progressively measurable process we take the following:*

$$\begin{gathered} (X_t)_{t\geq 0} \textit{ adapted to } \mathcal{F}_t \\ (X_t)_{t\geq 0} \textit{ has continuous trajectories} \\ \mathbf{E}[\textstyle\int_0^\infty X_t^2 dt] \leq \infty \\ \implies (X_t)_{t\geq 0} \textit{ is progressively measurable.} \end{gathered}$$

Proof. $(X_t)_{t\geq 0}$ seen as a measurable function $\Omega \times \mathbb{R}_{\geq 0} \mapsto \mathbb{R}^n$ can be approximated by

$$X_s^n = \mathbf{1}_0(s) X_0 \cdot + \sum_{i=0}^{N-1} \mathbf{1}_{(t_i^n, t_{i+1}^n]}(s) \cdot X_{t_i}^n \tag{1.6}$$

With $0 \leq t_0^n, \ldots, t_{N_n}^n < \infty$ with a mesh refining for $n \to \infty$. This process is progressively measurable and from continuity of the paths we get:

$$\lim_{n\to\infty} X^n = X.$$

□

Additionally we introduce a certain class of functions, which will be used as our set from which we take integrands to define the Itô-integral $\int_0^t f(s)dB_s$ for an $f : [0,t] \times \Omega \to \mathbb{R}^d$ bounded and measurable, and B_t a d-dimensional BM.

Definition 1.3.2 (Simple Predictable Processes). *Let $\mathcal{E}$ denote the set*

$$\{f : \Omega \times \mathbb{R}_{\geq 0} \to \mathbb{R}^d | f(s) = \sum_{i=0}^{n-1} F_i \mathbf{1}_{(t_i,t_{i+1}]}(s)\} \tag{1.7}$$

with $0 = t_0 \leq t_1 \leq \ldots < t_n = t$ and F_i is $\mathcal{F}_{t_i}-$measurable and square integrable.

So the elements of $\mathcal{E}$ are step functions with $\mathcal{F}_{t_i}$-measurable jumps. $\mathcal{E} \subset L^2([0,t] \times \Omega, \mathbb{R}^d)$ but it is not dense at all!

For such an $f \in \mathcal{E}$ we can define the Itô integral in the following way:

$$I(f) = \int_0^\infty f_t \; dB_t = \sum_{j=1}^{d} \sum_{i=0}^{n-1} F_i^j (B^j_{t_{i+1}} - B^j_{t_i}) \tag{1.8}$$

Theorem 1.3.1. $\mathcal{E} \subset L^2_{prog}(\Omega \times \mathbb{R}_{\geq 0}, \mathcal{F}_{prog}, P \bigotimes dt)$

Proof: On $(\Omega \times \mathbb{R})$ we have the sigma-algebra $\mathcal{F}_{prog}$ which is generated by all progressively measurable processes. So we can write $L^2_{prog}(\Omega \times \mathbb{R}_{\geq 0}, \mathcal{F}_{prog}, P \bigotimes dt)$.

Theorem 1.3.2. *(***Itô's Lemma***)*

The Itô integral $I : \mathcal{E} \mapsto L^2(\Omega, \mathcal{F}, \mathbb{P})$ is an isometry in the following sense: Let $f, g \in \mathcal{E}$ then for d = 1

$$\mathbf{E}[I(f)I(g)] = \mathbf{E}[\int_0^\infty f(t)g(t)dt]. \tag{1.9}$$

Where the last expectation can be seen as $\mathbf{E}[\omega \mapsto \int_0^\infty f(t,\omega)g(t,\omega)dt]$. In general:

$$\mathbf{E}[I(f)I(g)] =< f, g >_{L^2_{prog}} . \tag{1.10}$$

Proof of the Lemma for the case d = 1, the general case is proved in the same way but additionally uses the independence of the components of the d-dimensional BM:
$\mathbf{E}[I(f)I(g)] =$
$\mathbf{E}[(\sum_{i=0}^{n-1} F_i(B_{t_{i+1}} - B_{t_i}))(\sum_{j=0}^{n-1} G_j(B_{t_{j+1}} - B_{t_j}))] =$
$\mathbf{E}[\sum_{i=0}^{n-1} F_iG_i(B_{t_{i+1}} - B_{t_i})^2] + 2\mathbf{E}[\sum_{0=i<j}^{n-1} F_iG_j(B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j})] =$
$|F_i, G_i \; \mathcal{F}_{t_i} - measurable,$ and the increment of the BM is independent of $\mathcal{F}_{t_i}$
and finally we use that $\mathbf{E}[B_{t_{j+1}} - B_{t_j} | \mathcal{F}_{t_j}] = 0| =$
$\mathbf{E}[\mathbf{E}[\sum_{i=0}^{n-1} F_iG_i(B_{t_{i+1}} - B_{t_i})^2 | \mathcal{F}_{t_i}]] + 2\mathbf{E}[\mathbf{E}[\sum_{0=i<j}^{n-1} F_iG_j(B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j}) | \mathcal{F}_{t_i}]] =$
$\mathbf{E}[\sum_{i=0}^{n-1} F_iG_i\mathbf{E}[(B_{t_{i+1}} - B_{t_i})^2]] + 2\mathbf{E}[\mathbf{E}[\sum_{0=i<j}^{n-1} F_iG_j(B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j}) | \mathcal{F}_{t_j}]] =$
$\mathbf{E}[\sum_{i=0}^{n-1} F_iG_i(t_{i+1} - t_i)] + 2\mathbf{E}[\sum_{0=i<j}^{n-1} F_iG_j(B_{t_{i+1}} - B_{t_i})\mathbf{E}[B_{t_{j+1}} - B_{t_j} | \mathcal{F}_{t_j}]] =$
$\mathbf{E}[\sum_{i=0}^{n-1} F_iG_i(t_{i+1} - t_i)]\square$

Theorem 1.3.3. *The closure of $\mathcal{E}$ equals $L^2_{prog}(\Omega \times \mathbb{R}_{\geq 0}, \mathcal{F}_{prog}, \mathbb{P} \bigotimes dt)$. Formally we have that*

$$\overline{\mathcal{E}} = L^2_{prog}(\Omega \times \mathbb{R}_{\geq 0}, \mathcal{F}_{prog}, \mathbb{P} \bigotimes dt). \tag{1.11}$$

Proof. For each element of $L^2_{prog}(\Omega \times \mathbb{R}_{\geq 0})$ we need a sequence in $\overline{\mathcal{E}}$ that converges to it.

- First we replace $\mathbb{R}_{\geq 0}$ by $[0, 1]$.
- Take μ a measurable process in $L^2(\Omega \times [0, 1], \mathcal{F}_{prog}, \mathbb{P} \bigotimes dt)$.
- This process can be approximated by a bounded progressively measurable process $\mu' = \mu \wedge n \to \mu$.
- This μ' can be approximated by a continuous adapted process which finally can be approximated by a simple predictable step process $\in \mathcal{E}$.
- To make the last point clearer: we define $\mu'_h(t) = \frac{1}{h} \int_{t-h}^{t} \mu'(s)ds$ for $t, h > 0$ and $\mu'_h(s) = 0$ for $s \leq 0$. This $\mu'_h(t)$ tends to μ' for $h \to 0$ by Lebesgue's differentiation theorem. The integration makes it continuous!
- Thus for a fixed $\omega \in \Omega$, $\mu'_h(t)(\omega) \to \mu(t)(\omega)$ holds almost everywhere in $[0, 1]$ and so
- $\int_0^1 |\mu'_h(s)(\omega) - \mu'(s)(\omega)|^2 ds \to 0$ by dominated convergence and
- $\mathbf{E}[\int_0^1 |\mu'_h(s)(\omega) - \mu'(s)(\omega)|^2 ds] \to 0$ again by dominated convergence and by boundedness.
- Finally Fubini's theorem tells us that if this integral exists, the iterated ones do so, too, and this is the measure in $L^2_{prog}(dt \bigotimes dP)$.
- Continuity and adaptedness of the μ'_h complete the proof.

□

Definition 1.3.3. *The closure of $\mathcal{E}$ in L^2_{prog} is denoted by $L(B)$.*

Remark 1.3.1. *The definite integral of a progressively measurable process $u(s)$ is defined by:*

$$\int_0^t u(s)dB_s = \int_0^\infty \mathbf{1}_{[0,t]}(s)u(s)dB_s$$

for $t \geq 0$ which is well defined for a progressively measurable process.

Theorem 1.3.4. *The process $M_t := \int_0^t u(s)dB_s$ is a martingale w.r.t. the filtration $(\mathcal{F}_s)_{s\geq 0}$ for $u \in L(B)$.*

Proof. Let $u \in \mathcal{E}$ be given by

$$u(s) = \sum_{i=0}^{n-1} F_i \mathbf{1}_{(t_i,t_{i+1}]}(s)$$

then the stochastic integral is calculated by

$$M_t = \sum_{i=0}^{n-1} F_i (B_{t\wedge t_{i+1}} - B_{t\wedge t_i}).$$

Given $s \geq t$ and assume that there is a $k \leq n$ with $t_k = t$ then

$$\begin{aligned}
\mathbf{E}[M_s|\mathcal{F}_t] &= \sum_{i=0,t_{i+1}\leq t}^{n-1} F_i(B_{s\wedge t_{i+1}} - B_{s\wedge t_i}) \\
&+ \sum_{i=0,t_{i+1}>t}^{n-1} \mathbf{E}[F_i \mathbf{E}[B_{s\wedge t_{i+1}} - B_{s\wedge t_i}|\mathcal{F}_{t_i}]|\mathcal{F}_t] \\
&= \sum_{i=0}^{k-1} F_i(B_{s\wedge t_{i+1}} - B_{s\wedge t_i}) = M_t.
\end{aligned}$$

The assumptions are trivial because we can always refine the partition. By L^2 convergence of the conditional expectation we obtain the result. $\square$

Theorem 1.3.5. *The process $M_t := \int_0^t u(s)dB_s$ has a version with continuous paths.*

Proof. Analogous to the proof before we write

$$M_t = \sum_{i=0}^{n-1} F_i (B_{t\wedge t_{i+1}} - B_{t\wedge t_i})$$

for $t \geq 0$ which is continuous due to the continuity of BM. We now consider a Cauchy sequence $u_n \in \varepsilon$ converging to u and denote the associated martingales by M^n, then by Doob's martingale inequality we have

$$\mathbb{P}(\sup_{t\leq T} |M_t^n - M_t^m| \geq \epsilon) \leq \frac{1}{\epsilon^2}\mathbf{E}[|M_T^n - M_T^m|^2].$$

By the Itô isometry this equals

$$\frac{1}{\epsilon^2}\mathbf{E}[\int_0^T (u_n(s) - u_m(s))^2 ds] \to 0$$

for $n, m \to \infty$. So, we can find a subsequence n_k such that

$$\mathbb{P}(\sup_{t\leq T} |M_t^{n_k} - M_t^{n_{k+1}}| \geq \frac{1}{2^k}) \leq \frac{1}{2^k}$$

for $k > 0$. By the Borel-Cantelli Lemma we obtain that the set of events, where this supremum is greater than $\frac{1}{2^k}$ for infinitely many k has measure zero (because of convergence of the sum). So we can define $I(u)$ almost surely as uniform limit of continuous processes - which yields continuity. $\square$

Remark 1.3.2. *All simple processes $u \in \mathcal{E}$ are progressively measurable by definition. Given $u \in L^2_{prog}$ with continuous paths, then we can approximate the process $(u_s \mathbf{1}_{[0,t]}(s))_{s \geq 0}$ by elements in ε of the form*

$$u_s^n := \sum_{i=0}^{2^n - 1} u_{\frac{ti}{2^n}} \mathbf{1}_{(\frac{ti}{2^n}, \frac{t(i+1)}{2^n}]}(s),$$

which converge to u_s by continuity and in L^2_{prog} by dominated convergence. Therefore we can calculate the Itô integral for processes in L^2_{prog} with continuous paths by

$$\lim_{n \to \infty} \int_0^t u_s^n dB_s = \lim_{n \to \infty} \sum_{i=0}^{2^n - 1} u_{\frac{ti}{2^n}} (B_{\frac{t(i+1)}{2^n}} - B_{\frac{ti}{2^n}}).$$

We shall always take a continuous version.

Summary 1.2. *We can state that there is a proper definition of an integral of a progressively measurable process w.r.t. BM by Itô. This Itô integral leads to continuous martingales.*

1.3.2 Itô's formula

Definition 1.3.4 (Itô process). *A process $(X_t)_{t \geq 0}$ of the form*

$$X_t = X_0 + \int_0^t v(s) dB_s + \int_0^t u(s) ds \tag{1.12}$$

with $u, v \in L^2_{prog}$is called an Itô process.

Theorem 1.3.6. *Let $f \in C_b^2(\mathbb{R})$ (bounded, with bounded first and second derivative) and X_t an Itô process. Then*

$$\begin{aligned} f(X_t) &= f(X_0) + \int_0^t f'(X_s) u(s) dB_s + \int_0^t f'(X_s) v(s) ds \\ &+ \frac{1}{2} \int_0^t f''(X_s) u^2(s) ds \end{aligned} \tag{1.13}$$

holds.

For the proof see for example Øksendal [13]. First it assumes $f \in C_b^\infty(\mathbb{R})$ to use the Taylor expansion of $f(x)$ and careful computations of the limits lead to the result. Finally f is approximated by $C_b^2(\mathbb{R})-$functions.

Remark 1.3.3. *Itô's formula is often used to describe $f(X_t)$ by means of integrals. This often gives a better picture of the process.*

Definition 1.3.5. *We introduce the following notations:*

$$\begin{aligned} dX_t &= u(t)dB_t + v(t)dt \text{ for an Itô process and} && (1.14)\\ df(X_t) &= f'(X_t)dX_t + \frac{1}{2}f''(X_t)(dX_t)^2 \text{ for Itô's formula} && (1.15)\end{aligned}$$

for convenience.

Theorem 1.3.7 (Refined Version of Itô's Formula)**.** *Let $f(t,x) \in \mathbb{R}^{1,2}$ (thus once differentiable in t and two times in x) and X_t an Itô process then*

$$f(t, X_t) = f(0, X_0) + \int_0^t \frac{\partial f}{\partial t} ds + \int_0^t \frac{\partial f}{\partial x} dX_s + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(dX_s)^2. \tag{1.16}$$

Remark 1.3.4. *Like this we have a formula for functions that depend on an Itô process and time explicitly.*

Example 1.2. *The so called stochastic exponential of a deterministic process $h \in L^2(\mathbb{R}_{\geq 0}, \mathcal{B}(\mathbb{R}_{\geq 0}), dx)$ is defined in the following way:*

$$\varepsilon(h)_t = \exp(\int_0^t h(s)dB_s - \frac{1}{2}\int_0^t h(s)^2 ds) \tag{1.17}$$

for $t \geq 0$. If we define the Itô process $X_t = h(t)dB_t - \frac{1}{2}h(t)^2dt, X_0 = 0$ and take $f(x) = \exp(x)$ then Itô's formula shows that

$$\begin{aligned} d(\exp(X_t)) &= \exp(X_t)dX_t + \frac{1}{2}\exp(X_t)(dX_t^2)\\ &= \exp(X_t)h(t)dB_t. && (1.18)\end{aligned}$$

One thing has to be observed: $\exp(x)$ is not bounded at all. But it behaves nicely - we can stop the process at a stopping time and then apply Itô's formula on the stopped process and then let the stopping time tend to infinity. So this process can be expressed in terms of an Itô integral. This is remarkable.

Stochastic Integration by Parts

The so called integration by parts is a useful method for calculating certain stochastic integrals.

Corollary 1.3.1 (Stochastic Integration by Parts)**.** *Let a deterministic $f(s)$ be continuously differentiable then*

$$\int_0^t f(s)dB_s = f(t)B_t - \int_0^t B_s f'(s)ds \tag{1.19}$$

holds.

Proof. Take the Itô process $dX_t = dB_t$ and the function $F(t,x) = f(t)\cdot x$. Then Itô's formula leads to

$$\begin{aligned} F(t,X_t) &= \underbrace{f(0)\cdot B_0}_{=0} + \int_0^t \frac{\partial F}{\partial s} ds + \int_0^t \frac{\partial F}{\partial x} dX_s + \int_0^t \frac{1}{2} \underbrace{\frac{\partial^2 F}{\partial x^2}}_{=0} (dX_s)^2 \\ f(t)B_t &= \int_0^t f'(s)B_s ds + \int_0^t f(s) dB_s. \end{aligned}$$

□

Summary 1.3. *In this section we got to know Itô processes and learned how to represent functions of Itô processes using Itô's formula. In a corollary we saw an integration by parts formula which has convenient applications.*

1.4 The Stratonovich Interpretation

We define another integral idea, written down in terms of an Itô integral and the covariance process. Read more details in [15]. First we give a condition for the existence of the quadratic variation of a martingale M (denoted by $\langle M, M\rangle$):

Theorem 1.4.1. *A continuous and bounded martingale M is of finite quadratic variation and $\langle M, M\rangle$ is the unique continuous increasing adapted process vanishing at zero such that $M^2 - \langle M, M\rangle$ is a martingale.*

We can quote a similar theorem for a local martingale:

Theorem 1.4.2. *If M is a continuous and bounded local martingale, there exists a unique continuous increasing adapted process $\langle M, M\rangle$, vanishing at zero, such that $M^2 - \langle M, M\rangle$ is a local martingale.*

Definition 1.4.1 (Semi-Martingale)**.** *A stochastic $\mathbb{R}^d$-valued process $X = (X^1, \ldots, X^d)$ is called a semi-martingale if each component has a unique decomposition of the form*

$$X^i = X_0^i + M^i + A^i$$

where M^i is a local martingale and A^i an adapted process of finite variation, with $M_0^i = A_0^i = 0$.

The following proposition can be found in [15]:

Proposition 1.4.1. *A continuous semimartingale $X = M + A$ has finite quadratic variation and $\langle X, X\rangle = \langle M, M\rangle$.*

Proof. If Δ is a subdivision of $[0,t]$ then

$$|\sum_i (M_{t_{i+1}} - M_{t_i})(A_{t_{i+1}} - A_{t_i})| \leq (\sup_i |M_{t_{i+1}} - M_{t_i}|)\mathtt{Var}_t(A)$$

where $\mathtt{Var}_t(A)$ is the variation of A on $[0,t]$, and this converges to 0 when $|\Delta|$ tends to zero because of continuity of M. Likewise $\lim_{|\Delta|\to 0} \sum_i (A_{t_{i+1}} - A_{t_i})^2 = 0$. □

Definition 1.4.2 (Covariation Process)**.** *The covariation process (or the bracket) of 2 continuous semi-martingales $X = M + A$ and $Y = N + B$ is denoted by $\langle X, Y\rangle$ and equals $\langle M, N\rangle$. $\langle X, Y\rangle$ is the limit in probability of $\sum_i (X_{t_{i+1}} - X_{t_i})(Y_{t_{i+1}} - Y_{t_i})$.*

It can be derived from the quadratic variation of the processes by the following formula:

$$\langle M, N\rangle_t = \frac{\langle M+N, M+N\rangle_t - \langle M-N, M-N\rangle_t}{4}.$$

Conclusion 1.1. *The quadratic variation process of dB_t equals dt and in general the covariation process of an Itô diffusion of the form $dM_t = a_t dt + b_t dB_t$ equals $b_t^2 dt$.*

It can be proved that the Stratonovich integral is the limit of the following form:

$$\lim_{n\to\infty} \sum_{i=0}^{n} M_{\frac{(t_{i+1}+t_i)}{2}} (B_{t_{i+1}} - B_{t_i}) \quad = \quad \int_0^t M_t \circ dB_t \tag{1.20}$$

with $t_i = \frac{t}{n} i$. So it can be interpreted as a stochastic integral w.r.t. BM where we don't evaluate M at the left end point of the interval but at the middle point.

Definition 1.4.3 (The Stratonovich Integral)**.** *The Stratonovich integral of a semi-martingale M_t w.r.t. the BM B_t is defined as*

$$\int_0^t M_s \circ dB_s := \int_0^t M_s dB_s + \frac{1}{2}\langle M, B\rangle_t. \tag{1.21}$$

Remark 1.4.1. *Note that it is crucial for the process M_t for which one wants to define the Stratonovich integral to be a semi-martingale. This is a condition that shrinks the class of integrands (from progressively measurable to semi-martingales additionally).*

Whereas the Itô formulation is convenient for the analysis of probabilistic properties the Stratonovich interpretation shows the geometric behavior. The reader will see this in the sections later in the text.

1.5 Stochastic Differential Equations

1.5.1 Existence and Uniqueness of the Solution of an SDE

First we have to do some definitions on how a process can be a solution of an SDE. Additionally we will define two notions of uniqueness of such a solution. These definitions can be found in [8]. The intent is to assign a meaning to an SDE of the form

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t \tag{1.22}$$

where $B = \{B_t; 0 \leq t < \infty\}$ is an r-dimensional Brownian motion and $X = \{X_t; 0 \leq t < \infty\}$ is a suitable process with continuous sample paths and values in $\mathbb{R}^d$ - the 'solution' of the equation.

Definition 1.5.1 (Strong Solution). *A* ***strong solution*** *of the SDE 1.22 on the given probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and with respect to the fixed Brownian motion B and initial condition ζ, is a process $X = \{X_t; 0 \leq t < \infty\}$ with continuous sample paths and the following properties:*

1. *X is adapted to the filtration $\mathcal{F}_t$ (the augmentation of the filtration generated by B and the initial condition),*
2. *$\mathbb{P}(X_0 = \zeta) = 1$,*
3. *$\mathbb{P}(\int_0^t b_i(s, X_s) + \sigma_{ij}^2(s, X_s)ds < \infty) = 1$ holds for every $1 \leq i, j \leq d, 1 \leq j \leq r$ and $0 \leq t < \infty$ and*
4. *$X_t = X_0 + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dB_s$ holds almost surely.*

On the other hand we define a weak solution:

Definition 1.5.2 (Solution in the Weak Sense). *A* ***solution in the weak sense*** *of the SDE 1.22 is a triple $(X, B), (\Omega, \mathcal{F}, \mathbb{P}), \mathcal{F}_t$, where*

1. *$(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, and $\mathcal{F}_t$ is a filtration of sub-σ-fields of $\mathcal{F}$ satisfying the usual conditions,*
2. *$X = \{X_t, \mathcal{F}_t; 0 \leq t < \infty\}$ is a continuous, adapted $\mathbb{R}^d-$valued process, $B = \{B_t, \mathcal{F}_t; 0 \leq t < \infty\}$ is an r-dimensional Brownian motion, and*
3. *$\mathbb{P}(\int_0^t b_i(s, X_s) + \sigma_{ij}^2(s, X_s)ds < \infty) = 1$ holds for every $1 \leq i, j \leq d, 1 \leq j \leq r$ and $0 \leq t < \infty$ and*
4. *$X_t = X_0 + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dB_s$ holds almost surely.*

The difference of the definitions is that the filtration in the definition of a weak solution is not necessarily the augmentation of the filtration generated by B and the initial condition. But as B is a Brownian motion w.r.t. $\mathcal{F}_t$, another connection is given. It is clear that strong solvability implies weak solvability.

We now give two definitions of uniqueness (from [15]):

Definition 1.5.3 (Pathwise Uniqueness). *There is* ***pathwise uniqueness*** *of 1.22 if whenever (X, B) and (X', B') are two solutions defined on the same filtered space with $B = B'$ and $X_0 = X'_0$ a.s. then X and X' are indistinguishable.*

Some authors refer to this kind of uniqueness as strong uniqueness.

Definition 1.5.4 (Uniqueness in Law). *There is* ***uniqueness in law*** *of 1.22 if, for every $x \in \mathbb{R}^d$, whenever (X, B) and (X', B') are two solutions such that $X_0 = x$ and $X'_0 = x$ a.s., then the laws of X and X' are equal.*

Some authors refer to this kind of uniqueness as weak uniqueness.

The Case of Lipschitz Coefficients

In this section I want to prove the existence and uniqueness of the solution of an SDE. The proof can be found in Øksendal [13] and Teichmann [16]. The classical theorems needed can be found in Williams' book [17]. Additionally we will have to recall the fixed point theorem of Banach:

Theorem 1.5.1. *(Banach's Fixed Point Theorem) Let (M, d) be a metric space moreover let $F : M \to M$ be a function satisfying $d(F(x), F(y)) \leq K \cdot d(x, y)$, $K < 1$. Then there exists a unique fixed point x which can be found by iterating the function F starting with an initial value x_0. More formally:*

$$\exists! x, F(x) = x,$$
$$x = \lim_{n\to\infty} F^n(x_0).$$

The condition on the function F is a Lipschitz condition with a Lipschitz constant smaller than one. Besides that we will need the lemma of Gronwall.

Lemma 1.5.1. *(Lemma of Gronwall) Given a function $v : [0, T] \to \mathbb{R}$ being non-negative and continuous and satisfying $v(t) \leq F + A \cdot \int_0^t v(s)ds, t \in [0, T]$ and $F, A \in \mathbb{R}_{\geq 0}$, then $v(t) \leq F \cdot exp(At)$ holds.*

In the proof of the main theorem we will moreover need some essential theorems from probability theory:

Theorem 1.5.2. *(***Doob's Martingale Inequality***) Let M_t be an L^p-martingale for $p \in [1, \infty)$ with continuous trajectories, then*

$$\mathbb{P}(sup_{0\leq t\leq T}|M_t| \geq \lambda) \leq \frac{1}{\lambda^p} E(|M_T|^p) \tag{1.23}$$

holds.

Theorem 1.5.3. *(***Tchebychev's Inequality***) Let X be an L^2- random variable then*

$$\mathbb{P}(|X| \geq \lambda) \leq \frac{1}{\lambda^2}\mathbf{E}[X^2] \tag{1.24}$$

holds.

Theorem 1.5.4. *(Borel-Cantelli Lemma) Let $(\Omega, \mathcal{F}, P)$ be a probability space and A_n be a sequence of events ($A_n \in \mathcal{F}$). Then the following two assertions hold:*

$$a) \sum_n \mathbb{P}(A_n) < \infty \qquad \Rightarrow \mathbb{P}(\limsup_n A_n) = 0 \tag{1.25}$$

$$b) \sum_n \mathbb{P}(A_n) = \infty \quad \text{and the } (A_n)_{n\in\mathbf{N}} \text{ are independent}$$

$$\Rightarrow \mathbb{P}(\limsup_n A_n) = 1. \tag{1.26}$$

Theorem 1.5.5. (Existence and uniqueness theorem for stochastic differential equations) *Let $T > 0$ and $b : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n, \sigma : [0, T] \times \mathbb{R}^n \to \mathbb{R}^{n\times m}$ be measurable functions satisfying*

$$\|b(t, x)\| + \|\sigma(t, x)\| \leq C(1 + \|x\|); \quad x \in \mathbb{R}^n, t \in [0, T] \tag{1.27}$$

for some constant C, (where $\|\sigma\|^2 = \sum |\sigma_{ij}|^2$) and such that

$$\|b(t, x) - b(t, y)\| + \|\sigma(t, x) - \sigma(t, y)\| \leq D\|y - x\|; \quad x, y \in \mathbb{R}^n, t \in [0, T] \tag{1.28}$$

for some constant D. Let Z be a random variable which is independent of $\mathcal{F}_\infty$, the $\sigma-$algebra generated by $B_s, s \geq 0$ and such that

$$\mathbf{E}[\|Z\|^2] < \infty.$$

Then the stochastic differential equation

$$X_t = b(t, X_t)dt + \sigma(t, X_t)dB_t, \quad 0 \leq t \leq T, X_0 = Z \tag{1.29}$$

has a unique t-continuous solution $X_t(\omega)$ with the property that $X_t(\omega)$ is adapted to the filtration $\mathcal{F}_t^Z$ generated by Z and $B_s; s \leq t$ and

$$\mathbf{E}[\int_0^T \|X_t\|^2 dt] < \infty. \tag{1.30}$$

Proof. **uniqueness:**
Let $(X_t^1)_{t\geq 0}$and $(X_t^2)_{t\geq 0}$ be solutions of the SDE of the form

$$X_k = b(t, X_t)dt + \sigma(t, X_t)dB_t$$

and $X_0^1 = X_0^2 = Z$ satisfying the L^2 condition of a solution. Then by using the following tools

$$(A+B)^2 \leq 3 \cdot (A^2 + B^2)$$
$$\int_0^t 1 \cdot f(s)ds \leq t^{\frac{1}{2}} \cdot (\int_0^t f(s)^2 ds)^{\frac{1}{2}}$$
$$\text{the Itô isometry}$$

in the following computations:

$$\begin{aligned}
\mathbf{E}[\|X_t^1 - X_t^2\|^2] = \mathbf{E}[\| \int_0^t (b(s, X_s^1) - b(s, X_s^2))ds &\quad + \quad \int_0^t (\sigma(s, X_s^1) - \sigma(s, X_s^2))dB_s\|^2] \\
\leq 3 \cdot (\mathbf{E}[\| \int_0^t (b(s, X_s^1) - b(s, X_s^2))ds\|^2] &\quad + \quad \mathbf{E}[\| \int_0^t (\sigma(s, X_s^1) - \sigma(s, X_s^2))dB_s\|^2]) \\
\leq 3 \cdot t \cdot \int_0^t \mathbf{E}[\|(b(s, X_s^1) - b(s, X_s^2))\|^2]ds &\quad + \quad 3 \cdot \int_0^t \mathbf{E}[\|(\sigma(s, X_s^1) - \sigma(s, X_s^2))\|^2]ds) \\
\leq 3 \cdot t \cdot D^2 \int_0^t \mathbf{E}[\|X_s^1 - X_s^2\|^2]ds &\quad + \quad 3 \cdot D^2 \int_0^t \mathbf{E}[\|X_s^1 - X_s^2\|^2]ds
\end{aligned}$$

Now we can apply Gronwall's lemma with $F = 0$
So $v(s) \leq 3 \cdot (1+t) \cdot D^2 \int_0^t v(s)ds$
$\implies v(s) = 0$ on $[0, t]$. This implies that $X_t^1 = X_t^2$ a.s. - in other words $(X_t^1)_{t\geq 0}$ is indistinguishable from $(X_t^2)_{t\geq 0}$.

existence:
The proof of the existence is done by the Picard-Lindelöff Iteration:
We start with $Y_t^{(0)} = Z$. Further iterations lead to $Y_t^{(k)}$ and $Y_t^{(k+1)}, k \geq 0$. The following relationship obviously holds:

$$Y_t^{(k+1)} = Z + \int_0^t b(s, Y_s^{(k)})ds + \int_0^t \sigma(s, Y_s^{(k)})dB_s.$$

Similar calculations as in the uniqueness proof lead to:

$$\mathbf{E}[\|Y_t^{(k+1)} - Y_t^{(k)}\|^2] \leq 3 \cdot (T+1) \cdot D^2 \int_0^t \mathbf{E}[\|Y_s^{(k)} - Y_s^{(k-1)}\|^2]ds.$$

So we have for $t \leq T$:

$$\begin{aligned}
\mathbf{E}[\|Y_t^{(1)} - Y_t^{(0)}\|^2] &\leq 2C^2t^2 \cdot (1 + \mathbf{E}[\|Z\|^2] + 2C^2t(1 + \mathbf{E}[\|Z\|^2]) \\
&\leq A_1 \cdot t
\end{aligned}$$

and induction shows that the following holds:

$$\mathbf{E}[\|Y_t^{(k+1)} - Y_t^{(k)}\|^2] \leq \frac{t^{k+1}}{(k+1)!} A_2^{k+1}; \quad k \geq 0, 0 \leq t \leq T.$$

Where A_1 depends on $\mathbf{E}[\|Z\|^2], C$ and T and A_2 only depends on $\mathbf{E}[\|Z\|^2], C, D$ and T.

Moreover

$$\begin{aligned} \sup_{0\leq t\leq T} \|Y_t^{(k+1)} - Y_t^{(k)}\| &\leq \int_0^T \|b(s, Y_s^{(k)}) - b(s, Y_s^{(k-1)})\| ds \\ &+ \sup_{0\leq t\leq T} \|\int_0^t (\sigma(s, Y_s^{(k)}) - \sigma(s, Y_s^{(k-1)})) dB_s\|. \end{aligned}$$

Where the last expression is a martingale.

$$\begin{aligned} \mathbb{P}(\sup_{0\leq t\leq T} \|Y_t^{(k+1)} - Y_t^{(k)}\| &> \frac{1}{2^k}) \\ \leq \mathbb{P}((\int_0^T \|b(s, Y_s^{(k)}) - b(s, Y_s^{(k-1)})\| ds)^2 &> \frac{1}{2^{2k+2}}) \\ +\mathbb{P}(\sup_{0\leq t\leq T} \|\int_0^t (\sigma(s, Y_s^{(k)}) - \sigma(s, Y_s^{(k-1)})) dB_s\| &> \frac{1}{2^{k+1}}) \end{aligned}$$

By Doob's maximal martingale inequality, Tchebychev's inequality and the Itô isometry we get that this is:

$$\begin{aligned} &\leq 2^{2k+2} \cdot T \cdot \int_0^T \mathbf{E}[\|b(s, Y_s^{(k)}) - b(s, Y_s^{(k-1)})\|^2] ds \\ &+ 2^{2k+2} \cdot \int_0^T \mathbf{E}[\|\sigma(s, Y_s^{(k)}) - \sigma(s, Y_s^{(k-1)})\|^2] ds \\ &\leq 2^{2k+2} D^2 (T+1) \int_0^T \frac{A_2^k t^k}{k!} dt \leq \frac{(4A_2T)^{k+1}}{(k+1)!} \text{ , if } A_2 \geq D^2(T+1). \end{aligned}$$

This sum converges and so we can apply the Borel-Cantelli Lemma. It gives that the set of ω such that $\sup_{0\leq t\leq T} \|Y_t^{(k+1)} - Y_t^{(k)}\| > \frac{1}{2^k}$ for infinitely many k has measure zero. Thus for almost all ω there exists $k_0 = k_0(\omega)$ such that $\sup_{0\leq t\leq T} \|Y_t^{(k+1)} - Y_t^{(k)}\| \leq \frac{1}{2^k} \quad k \geq k_0$. Therefore the sequence

$$Y_t^{(n)} = Y_t^{(0)} + \sum_{k=0}^{n-1} (Y_t^{k+1} - Y_t^k)$$

is uniformly convergent in $[0, T]$ for almost all ω. Let us denote this limit by X_t. X_t is continuous because $Y_t^{(n)}$ is continuous for all n and it is $\mathcal{F}_t$-measurable for the same reason. Moreover we want to show the adaptedness and the integrability:

$$\mathbf{E}[\|Y_t^{(n)} - Y_t^{(m)}\|^2]^{\frac{1}{2}} = \|Y_t^{(n)} - Y_t^{(m)}\|_{L^2(P)} = \|\sum_{k=n}^{m-1}(Y_t^{(k+1)} - Y_t^{(k)})\|_{L^2(P)}$$
$$\leq \sum_{k=n}^{m-1}\|Y_t^{(k+1)} - Y_t^{(k)}\|_{L^2(P)} \leq \sum_{k=n}^{\infty}[\frac{(A_2t)^{k+1}}{(k+1)!}]^{\frac{1}{2}} \to 0 \text{ as } n \to \infty.$$

So $Y_t^{(n)}$ converges in $L^2(P)$ to a limit say Y_t. So a subsequence converges $\omega-$pointwise to Y_t and therefore we must have $Y_t = X_t$ a.s. So X_t fulfills the adaptedness and integrability property.

The last thing to show is that this limit really fulfills the SDE: By Fatou and some results before we have:

$$\mathbf{E}[\int_0^T \|X_t - Y_t^{(n)}\|^2 dt] \leq \limsup_{m\to\infty} \mathbf{E}[\int_0^T \|Y_t^{(m)} - Y_t^{(n)}\|^2 dt] \to 0$$

as $n \to \infty$. By the Itô isometry it follows that

$$\int_0^t \sigma(s, Y_s^{(n)})dB_s \to \int_0^t \sigma(s, X_s)dB_s$$

and by the Hölder inequality that

$$\int_0^t b(s, Y_s^{(n)})ds \to \int_0^t b(s, X_s)ds$$

in $L^2(P)$. So by taking the limit in the representation

$$Y_t^{(n+1)} = X_0 + \int_0^t b(s, Y_s^{(n)})ds + \int_0^t \sigma(s, Y_s^{(n)})dB_s.$$

we get the final result. $\square$

Remark 1.5.1. *In the notions of uniqueness and existence which we got to know we can say that this theorem gives conditions for a strongly unique strong solution.*

Example 1.3. Ornstein-Uhlenbeck process *The solution of the following SDE is called Ornstein-Uhlenbeck process:*

$$dX_t = a \cdot (b - X_t) \cdot dt + \sigma \cdot dB_t,\ X_0 = x > 0\ a, b, \sigma \in \mathbb{R}. \tag{1.31}$$

The requirements for a unique solution are fulfilled:
The uniqueness requirement:

$$|a \cdot (b - X) - a \cdot (b - Y)| + |\sigma - \sigma| = a \cdot |X - Y|$$

The growth requirement:

$$\begin{aligned}|a\cdot(b-X)|+|\sigma| &= |ab-aX|+|\sigma| \\ &\leq \max(|ab|,|a|)\cdot|1+X|+|\sigma|.\end{aligned}$$

So this equation has a unique solution. It can be found by the method of variation of constants.

Example 1.4. Black-Scholes Model *This model describes the value process of one unit of a stock. The equation is:*

$$dS_t = \mu\cdot S_t\cdot dt + \sigma\cdot S_t\cdot dB_t,\ S_0\in\mathbb{R}_{\geq 0}. \tag{1.32}$$

Example 1.5. Vasicek Model *This models the overnight interest rate r_t by an Ornstein-Uhlenbeck process. The rate is modeled by the following equation:*

$$dr_t = a\cdot(b-r_t)\cdot dt + \sigma\cdot dB_t,\ r_0\in\mathbb{R}_{\geq 0}. \tag{1.33}$$

Theorems for the Case of Hölder-continuous Coefficients

Theorems on existence and uniqueness in the more general case of Hölder-continuous coefficients can be found in the book of Karatzas and Shreve [8] and in the book of Revuz and Yor [15].

Remark 1.5.2. *The existence of a solution for an SDE with a square root appearing in the volatility term can not be guaranteed by the theorems in the last section (although needed e.g. in the CIR-model). The square root is Hölder of order $\frac{1}{2}$. Let without loss of generality $x\geq y$ then $(\sqrt{x}-\sqrt{y})^2 = x-2\sqrt{xy}+y\leq x-2y+y = x-y$ and therefore $\sqrt{x}-\sqrt{y}\leq\sqrt{x-y}$.*

Proposition 1.5.1 (Yamada & Watanabe (1971))**.** *Let us suppose that the coefficients of the one-dimensional equation*

$$dX_t = b(t,X_t)dt + \sigma(t,X_t)dB_t$$

satisfy the conditions

$$\begin{aligned}|b(t,x)-b(t,y)| &\leq K|x-y| \\ |\sigma(t,x)-\sigma(t,y)| &\leq h(|x-y|)\end{aligned}$$

for every $0\leq t<\infty$ and $x\in\mathbb{R}, y\in\mathbb{R}$, where K is a positive constant and $h:[0,\infty)\rightarrow[0,\infty)$ is a strictly increasing function with $h(0)=0$ and

$$\int_0^\epsilon h^{-2}(u)du = \infty, \forall\epsilon>0$$

then strong uniqueness holds.

One easily sees that $h(x) = x^\alpha$ for $\alpha \geq \frac{1}{2}$ is a possible choice and this enables us to cope with the square root.

Remark 1.5.3. *Yamada and Watanabe actually established proposition 1.5.1 under a weaker condition on $b(t,x)$ namely*

$$|b(t,x) - b(t,y)| \leq \kappa(|x-y|), 0 \leq t < \infty, x \in \mathbb{R}, y \in \mathbb{R}$$

where $\kappa : [0,\infty) \to [0,\infty)$ is strictly increasing and concave with $\kappa(0) = 0$ and $\int_0^\epsilon \kappa^{-1}(u)du = \infty, \forall \epsilon > 0$.

Karatzas and Shreve [8](page 323) quote a theorem on the existence of a weak solution in the case of a time-homogeneous SDE with bounded continuous coefficients.

Theorem 1.5.6 (Skorohod (1965), Strook & Varadhan (1969))**.** *Consider the SDE*

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t$$

where the coefficients $b_i, \sigma_{ij} : \mathbb{R}^d \to \mathbb{R}$ are bounded and continuous functions. Corresponding to every initial distribution μ on $\mathcal{B}(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} \|x\|^{2m}\mu(dx) < \infty$ for some $m > 1$, there exists a weak solution.

1.5.2 Differential Operators

In this section I want to repeat and make clear how vector fields and differential operators are connected. This was explained in Teichmann's lecture [16].

Definition 1.5.5. *For $V : \mathbb{R}^n \mapsto \mathbb{R}^n$ (a vector field) and a smooth function $f : \mathbb{R}^n \mapsto \mathbb{R}$ we can define the value of a differential operator V on a function in the following way:*

$$\begin{aligned} Vf(x) &= V(x) \cdot \mathit{grad}\ f \\ &= \sum_{i=1}^n v_i(x) \frac{\partial}{\partial x_i} f(x). \end{aligned}$$

Remark 1.5.4. *The importance of this will become clear if one wants to apply a function to the solution of an SDE.*

Definition 1.5.6. *We define the differential operator D applied to a function*

$V : \mathbb{R}^n \mapsto \mathbb{R}^n$ *by* $DV = \begin{pmatrix} \frac{\partial V_1}{\partial x_1} & \cdots & \frac{\partial V_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial V_n}{\partial x_1} & \vdots & \frac{\partial V_n}{\partial x_n} \end{pmatrix}$.

1.5.3 Notation

We will use two ways to write down SDE's:

- $$dX_t = V(X_t)dt + \sum_{i=1}^{d} V_i(X_t)dB_t^i \quad X_0 = x$$
  With the vector fields $V : \mathbb{R}^n \mapsto \mathbb{R}^n$ (the Itô drift) and $V_i : \mathbb{R}^n \mapsto \mathbb{R}^n, i = 1, \dots d$ (the volatility vector fields) and d BM's.
  In this notation we introduce $V^0 = V + \frac{1}{2}\sum_{i=1}^{d} DV_i \cdot V_i$ the so called Stratonovich corrected Itô drift. This drift shows the geometric drift of the process which is the Itô drift plus terms coming from the covariation process.

- $$dX_t = b(X_t)dt + \sigma(X_t)dB_t, \quad 0 \le t \le T, X_0 = Z$$
  with $b : \mathbb{R}^n \to \mathbb{R}^n$ and $\sigma : \mathbb{R}^n \to \mathbb{R}^{n \times d}$.

Those definitions are equivalent and we will use the notation that is most convenient.

1.5.4 Solution Methods

As in the case of an ODE there are only a few equations that have an explicit solution here I point out 2 methods for SDE's of a special structure.

Variation of Constants

The variation of constants enables us to solve an SDE of the following form

$$dX_t = (a - bX_t)dt + cdB_t \quad a, b, c \in \mathbb{R}. \tag{1.34}$$

Such an equation is called scalar SDE. This particular equation is know as the Lagrange equation, the solution is the mean reverting Ornstein-Uhlenbeck process. The solution is given by

$$X_t = e^{-bt} \cdot (X_0 + \int_0^t ae^{bs}ds + \int_0^t ce^{bs}dB_s). \tag{1.35}$$

Proof. The homogeneous equation $dz_t = -bz_t dt$ has the solution $z_t = z_0 \cdot e^{-bt}$ So a variation of the constants shows that

$$\begin{aligned}
X_t &= z_0(t) \cdot e^{-bt} \\
dX_t &= e^{-bt}dz_0(t) - bX_t dt \text{ (Itô's formula)} \\
\text{by } e^{-bt}dz_0(t) &= adt + cdB_t \\
z_0(t) &= X_0 + \int_0^t ae^{bs}ds + \int_0^t ce^{bs}dB_s.
\end{aligned}$$

By multiplying by e^{-bt} we get the result. $\square$

A Formula for n Dimensions and d Brownian Motions

Theorem 1.5.7. *Let f and $V^1, \dots, V^d : [0,T] \to \mathbb{R}^n$ be smooth functions and A be a constant $N \times N$ matrix, then the unique solution of the SDE*

$$dX_t = (f(t) + AX_t)dt + \sum_{i=1}^{d} V^i(t)dB_t^i \tag{1.36}$$

with initial value $X_0 \in \mathbb{R}^n$ is given by

$$X_t = \exp(At)X_0 + \int_0^t \exp(A(t-s))f(s)ds + \sum_{i=1}^{d} \int_0^t \exp(A(t-s))V^i(s)dB_s^i.$$

The proof is a simple application of Itô's formula.

1.5.5 Supplements on the Existence and Uniqueness Theorem

For this section see Arnold [3]. We needed the Lipschitz condition to guarantee that $b(t,x)$ and $\sigma(t,x)$ do not change faster than x itself. This excludes discontinuous functions as coefficients and even a large class of continuous ones. To enlarge the class of functions for which we can guarantee a solution we need the following corollary.

Corollary 1.5.1. *The existence and uniqueness theorem remains valid if we replace the Lipschitz condition with the more general condition that, for every $N > 0$, there exists a constant K_N s.t., for all $t \in [0,T], \|x\| \leq N$ and $\|y\| \leq N$*

$$\|b(x,t) - b(y,t)\| + \|\sigma(x,t) - \sigma(y,t)\| \leq K_N \|x-y\|. \tag{1.37}$$

Another corollary helps to recognize when a function fulfills a Lipschitz condition.

Corollary 1.5.2. *For the Lipschitz condition in the existence and uniqueness theorem or its generalization (1.37) to be satisfied it is sufficient that the functions $b(x,t)$ and $\sigma(x,t)$ have continuous partial derivatives of first order w.r.t. the components of x for every $t \in [0,T]$ and that these be bounded on $[0,T] \times \mathbb{R}^n$ (or in the case of the generalization on $[0,T] \times \{\|x\| \leq N\}$).*

Proof. This is proved by the mean-value theorem for each component. □

We now focus on the second condition on the coefficients - the so called explosion condition. This condition allows at most linear increase of the functions b and σ. If this condition is violated one speaks of the phenomenon of explosion of the solution.

Example 1.6. *Consider the scalar ODE*

$$dX_t = X_t^2 dt, X_0 = c.$$

The solution is $X_t = 0$ for $c = 0$ and $X_t = (1/c - t)^{-1}$ for $c \neq 0$. So we see that the solution is only defined for $c > 0$ on the interval $[0, 1/c)$. At $t = 1/c$ the explosion takes place.

Definition 1.5.7 (Global Solution)**.** *If the functions b and σ are defined on $[0,\infty)\times\mathbb{R}^n$ and if the assumptions of the existence and uniqueness theorem hold on every finite subinterval $[0,T]$ of $[0,\infty)$, then the SDE*

$$dX_t = b(X_t,t)dt + \sigma(t,X_t)dB_t, \quad X_0 = Z$$

has a unique solution X_t defined on the entire half-line $[0,\infty)$. Such a solution is called a global solution.

Example 1.7. *Consider the autonomous (meaning that there is no explicit dependence of the coefficients on time) SDE*

$$dX_t = -1/2\exp(-2X_t)dt + \exp(-X_t)dB_t.$$

For $x<0$ the coefficients do not satisfy any Lipschitz or growth condition. So explosions are possible. The function $X_t = \log(B_t + e^{X_0})$ is a unique solution on the interval $[0,\eta)$ with instant of explosion

$$\eta = \inf\{t : B_t = -e^{X_0}\} > 0.$$

This can be verified by Itô's formula.

The following is shown in a theorem of McKean:

Theorem 1.5.8. *The autonomous SDE*

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t, X_0 = c$$

where b and σ are continuously differentiable functions, has a unique local solution that is defined up to a (random) explosion time η in the interval $0\le\eta\le\infty$. If $\eta<\infty$ then $X_{\eta^-} = -\infty$ or ∞.

One more theorem will help us to understand the behavior of a solution up to a stopping-time.

Theorem 1.5.9. *Let the vector fields $V, V_i : U\subset\mathbb{R}^n\to\mathbb{R}^n$ with U open, $i=1,\dots,d$ and let them be in C^∞. Then the SDE*

$$dX_t^x := V(X_t^x)dt + \sum_{i=1}^{d} V_i(X_t^x)dB_t^i \tag{1.38}$$

has a unique strong solution up to a stopping time $\tau>0$.

By a strong solution X_t^x we mean a process X_t starting at x that fulfills the SDE with a version of the BM given in advance and constructed by the Picard-Lindelöff procedure in the existence theorem. We introduce the pair $[(X_t^x),\tau^x]$, $\tau^x>0, \mathbb{P}-a.s.$ of a process X_t starting at x fulfilling $\mathbf{E}[\int_0^{\tau^x} X_t^2]<\infty$ with a stopping time until which it is defined. With these conventions we can start to prove this theorem.

Proof. We choose a $K \subset U$ compact such that the interior of K is not empty. For each x in the interior of K we can find a pair consisting of a strong solution and the corresponding stopping time until which it is defined $[(X_t^x), \tau^x]$ such that X_t^x fulfills the SDE 1.38. We now define the vector fields

$$\begin{aligned}
\tilde{V}(x) &:= V(x)\psi(x) \\
\tilde{V}_1(x) &:= V_1(x)\psi(x) \\
&\vdots \\
\tilde{V}_d(x) &:= V_d(x)\psi(x).
\end{aligned}$$

With a function $\psi : U \to \mathbb{R}$ such that $\psi|_K = 1$ and ψ has compact support in U and such that the vector fields $\tilde{V}_i(x)$ and $\tilde{V}(x)$ are still C^∞ but now bounded. Therefore by the existence theorem we have a solution $(\tilde{X}_t^x)_{t\geq 0}$ that fulfills equation 1.38 with the tilde vector fields plugged in. If we now define the stopping time $\tau^x := \inf\{t | X_t^x \notin K\}$ then we have for x in the interior of K that $\tau^x > 0$ a.s. and that the process $\tilde{X}_t^x$ solves equation 1.38 up to the stopping time τ^x. We did not specify the compact set in U so we can take any and there is a solution for every $x \in U$.

If we now have 2 pairs $[X_t^x, \tau^x]$ and $[\hat{X}_t^x, \hat{\tau}^x]$ where both processes are strong solutions up to the respective stopping time then they have to coincide up to the minimum of these times (which is again a stopping time) by uniqueness and the pair $[X_t^x, \min(\tau^x, \hat{\tau}^x)]$ is unique. $\square$

Summary 1.4. *In this section we saw some cases in which we can guarantee a solution of an SDE. 2 standard methods of solving SDE's of a certain form were presented and finally some theorems about local solutions.*

1.6 Properties of the solution of an SDE

Most of the theorems and definitions found here are taken from Øksendal [13] and Teichmann [16].

1.6.1 What a Diffusion Really Does

A rotation in $\mathbb{R}^2$ can be described by the following system of ODE's:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} \tag{1.39}$$

with $x(0) = x_0 \in \mathbb{R}$ and $y(0) = y_0 \in \mathbb{R}$ as initial conditions. This system can be solved by methods from the theory of ODE's like the Picard-Lindelöff iteration which leads to the solution

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} x_0 \cos(t) + y_0 \sin(t) \\ -x_0 \sin(t) + y_0 \cos(t) \end{pmatrix}$$

We now have a look at a noisy rotation as an example:

$$\begin{aligned} dx &= y dB_t, \quad x(0) = 1 \\ dy &= -x dB_t, \quad y(0) = 0 \end{aligned} \tag{1.40}$$

this system has the solution:

$$\begin{aligned} x(t) &= e^{\frac{t}{2}} \cos(B_t) \\ y(t) &= e^{-\frac{t}{2}} \sin(B_t) \end{aligned}$$

In this example we see that the solution, though it is a local martingale (only Itô integrals appear in the SDE system), leaves the unit disc and the distance to the origin tends to ∞.

If we look at another example :

$$\begin{aligned} dx &= -\frac{1}{2} x dt + y dB_t, \quad x(0) = x_0 \\ dy &= -\frac{1}{2} y dt - x dB_t, \quad y(0) = y_0 \end{aligned} \tag{1.41}$$

This system is solved by

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \exp(B_t \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}) \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \tag{1.42}$$

Observe that $(x(t), y(t))$ remain on the unit disc for all t.

So what is the difference between these two systems? It is the drift term. There is always another drift coming from the quadratic variation of the BM. It is canceled by the added drift in the second stochastic example and in the first process one sees what happens if one neglects this drift. We can say that the real behavior of the solution of a SDE is determined by the so called Stratonovich corrected Itô drift. Lets calculate for example the drift of (1.40):

$$\begin{aligned} &V(x, y) = 0, V_1(x, y) = (y, -x) \Longrightarrow \\ &V_0(x, y) = 0 - \frac{1}{2}(-x, -y) = \frac{1}{2}(x, y) \end{aligned}$$

This is exactly the drift that was subtracted in the system (1.41) leading to solutions that stay on the unit disc.

1.6.2 The Generator of an Itô Process

One fundamental thing to understand is the connection between an Itô process and a second order differential operator. This connection is given by the generator of the Itô process denoted by A:

Definition 1.6.1. *The infinitesimal generator A of an Itô process X_t in $\mathbb{R}^n$ is defined in the following way:*

$$Af(x) = \lim_{t \downarrow 0} \frac{\mathbf{E}[f(X_t)] - f(x)}{t}; \quad x \in \mathbb{R}^n. \tag{1.43}$$

We denote the set of functions $f : \mathbb{R}^n \to \mathbb{R}$ for which this limit exists by $\mathcal{D}_A(x)$. $\mathcal{D}_A$ denotes the set of functions for which the limit exists for all $x \in \mathbb{R}^n$. We need some more results, to express this connection in terms of the vector fields associated to the process.

Lemma 1.6.1. *Let Y_t be an Itô process in $\mathbb{R}^n$ of the form*

$$Y_t^x = x + \int_0^t u_s ds + \int_0^t v_s dB_s.$$

Let $f \in C_2(\mathbb{R})$ with compact support and moreover τ be a stopping time and assume that $\mathbf{E}[\tau] < \infty$. Assume that $u(t,\omega)$ and $v(t,\omega)$ are bounded on the set of (t,ω) such that $Y(t,\omega)$ belongs to the support of f.

Then $\mathbf{E}[f(Y_\tau)] =$

$$f(x) + \mathbf{E}[\int_0^\tau (\sum_i u_i(s,\omega) \frac{\partial f}{\partial x_i}(Y_s) + \frac{1}{2} \sum_{i,j} (vv^T)_{i,j}(s,\omega) \frac{\partial^2 f}{\partial x_i \partial x_j}(Y_s)) ds]. \tag{1.44}$$

Proof. If we put $Z = f(Y)$ and suppress some confusing notation then we get by applying Itô's formula that

$$\begin{aligned} dZ &= \sum_i \frac{\partial f}{\partial x_i}(Y) dY_i + \frac{1}{2} \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(Y) dY_i dY_j \\ &= \sum_i u_i \frac{\partial f}{\partial x_i} dt + \frac{1}{2} \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j} (vdB)_i (vdB)_j + \sum_i \frac{\partial f}{\partial x_i} (vdB)_i. \end{aligned}$$

The next line follows easily from previous observations and the following:

$$\begin{aligned} (vdB)_i \cdot (vdB)_j &= (\sum_k v_{ik} dB_k)(\sum_n v_{jn} dB_n) \\ &= (\sum_k v_{ik} v_{jk}) dt = (vv^T)_{ij} dt, \end{aligned}$$

so

$$\begin{aligned} f(Y_t) = f(Y_0) &+ \int_0^t (\sum_i u_i \frac{\partial f}{\partial x_i} + \frac{1}{2} \sum_{i,j} (vv^T)_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j}) ds \\ &+ \sum_{i,k} \int_0^t v_{ik} \frac{\partial f}{\partial x_i} dB_k. \end{aligned}$$

Hence

$$\begin{aligned}\mathbf{E}[f(Y_\tau] = f(x) \quad &+ \quad \mathbf{E}[\int_0^\tau (\sum_i u_i \frac{\partial f}{\partial x_i}(Y) + \frac{1}{2}\sum_{i,j}(vv^T)_{ij}\frac{\partial^2 f}{\partial x_i \partial x_j}(Y))ds] \\ &+ \quad \sum_{i,k}\mathbf{E}[\int_0^\tau v_{ik}\frac{\partial f}{\partial x_i}(Y)dB_k].\end{aligned}$$

If g is a bounded Borel function, $|g| \leq M$, then for all integers k we have

$$\mathbf{E}[\int_0^{\tau\wedge k} g(Y_s)dB_s] = \mathbf{E}[\int_0^k \mathbf{1}_{\{s<\tau\}} g(Y_s)dB_s] = 0$$

since g and $\mathbf{1}_{\{s<\tau\}}$ are both $\mathcal{F}_s$-measurable. Moreover it follows by monotone convergence of $\tau \wedge k \to \tau$ that

$$\begin{aligned}\mathbf{E}[(\int_0^\tau g(Y_s)dB_s - \int_0^{\tau\wedge k} g(Y_s)dB_s)^2] &= \mathbf{E}[(\int_{\tau\wedge k}^t g(Y_s)dB_s)^2] \\ \underbrace{=}_{\text{Itô isometry}} \mathbf{E}[\int_{\tau\wedge k}^\tau g^2(Y_s)ds] &\leq M^2 \cdot \mathbf{E}[\tau - \tau\wedge k] \underbrace{\to}_{k\to\infty} 0.\end{aligned}$$

Therefore

$$\mathbf{E}[\int_0^\tau g(Y_s)dB_s] = \lim_{k\to\infty}\mathbf{E}[\int_0^{\tau\wedge k} g(Y_s)dB_s] = 0.$$

$\square$

Theorem 1.6.1. *Let X_t be an Itô process of the form $X_t = b(X_t)dt + \sigma(X_t)dB_t$. If $f \in C_0^2(\mathbb{R})$ then $f \in \mathcal{D}_A$ and*

$$Af(x) = \sum_i b_i(x)\frac{\partial f}{\partial x_i} + \frac{1}{2}\sum_{i,j}(\sigma\sigma^T)_{i,j}(x)\frac{\partial^2 f}{\partial x_i \partial x_j}.$$

Proof. This follows from the last lemma if one takes $\tau = t$ and then looks at the definition of the generator:

$$\begin{gathered}Af(x) = \lim_{t\downarrow 0}\tfrac{\mathbf{E}[f(X_t)]-f(x)}{t} \\ \lim_{t\downarrow 0}\tfrac{1}{t}\cdot(f(x) + \mathbf{E}[\int_0^t(\sum_i b_i(X_s)\tfrac{\partial f}{\partial x_i} + \tfrac{1}{2}\sum_{i,j}(\sigma\sigma^T)_{i,j}(X_s)\tfrac{\partial^2 f}{\partial x_i\partial x_j})ds] - f(x)) = \\ \lim_{t\downarrow 0}\mathbf{E}[\tfrac{1}{t}\int_0^t(\sum_i b_i(X_s)\tfrac{\partial f}{\partial x_i} + \tfrac{1}{2}\sum_{i,j}(\sigma\sigma^T)_{i,j}(X_s)\tfrac{\partial^2 f}{\partial x_i\partial x_j})ds] = \\ \sum_i b_i(x)\tfrac{\partial f}{\partial x_i} + \tfrac{1}{2}\sum_{i,j}(\sigma\sigma^T)_{i,j}(x)\tfrac{\partial^2 f}{\partial x_i\partial x_j}\end{gathered}$$

The last line holds by continuity and for $t \to 0$ the process is $x \in \mathbb{R}^n$ thus the expectation loses its meaning . Together with the definition of A we get the result. $\square$

Theorem 1.6.2. *If we write the SDE in the way* $dX_t = V(X_t)dt + \sum_{i=1}^{d} V_i(X_t)dB_t^i$ *described before then we can define the generator in the following way:*

$$
\begin{aligned}
Af &= (V_0 - \sum_{i=1}^{d} V_i^2)f && (1.45)\\
&= V_0 \cdot gradf - \sum_{i=1}^{d} V_i(V_i \cdot gradf) && (1.46)
\end{aligned}
$$

where V_0 *is the Stratonovich corrected Itô drift and the* V_i *are the volatility vector fields. This leads to the same object as the definitions before.*

Example 1.8. *The n-dimensional BM is the solution of* $dX_t = dB_t$*, so we have* $b = 0$ *and* $\sigma = I_n$ *(the identity matrix in* $R^{n\times n}$*). Thus the generator is the following:*

$$Af = \frac{1}{2}\sum \frac{\partial^2 f}{\partial x_i^2} = \frac{1}{2}\sum \Delta f.$$

1.6.3 The Dynkin Formula

With the results of the previous section we can propose a very useful theorem. Before that we have to define an object.

Definition 1.6.2 (Stopping Time)**.** *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space, and* $\mathcal{F}_t$ *a filtration. Then a random variable* τ *is called a stopping time w.r.t.* $\mathcal{F}_t$ *if*

$$\{\tau \leq t\} \in \mathcal{F}_t \quad \text{for all } t \geq 0. \tag{1.47}$$

Theorem 1.6.3 (Dynkin's Formula)**.** *Let* $f \in C_0^2(\mathbb{R}^n)$*. Suppose* τ *is a stopping time with* $\mathbf{E}[\tau] < \infty$*. Then*

$$\mathbf{E}[f(X_\tau] = f(x) + \mathbf{E}[\int_0^\tau Af(X_s)ds]. \tag{1.48}$$

There is nothing left to prove after the last section on the generator. I just want to give a very illustrative example.

Example 1.9. *We look at a n-dimensional BM starting at* $a \in \mathbb{R}^n$ *furthermore we assume* $|a| < R$*. Thus the BM starts inside of a ball of radius R centered at the origin. By applying Dynkin's formula we will be able to answer the following question: What is* $\mathbf{E}[\tau_K]$ *if* τ_K *denotes the exit time of the ball* $K_R = \{x \in \mathbb{R}^n; |x| < R\}$ *?*
We introduce a stopping time $\sigma_k = min(k, \tau_K)$ *for an arbitrary* $k \in \mathbb{N}$ *furthermore we need an* $f \in C_0^2(\mathbb{R}^n)$ *- we take* $|x|^2$ *for* $|x| < R$*. We don't have to define f for other x because we will only integrate the process with values inside of the ball. Now we can apply Dynkin's formula with the stopping time* σ_k *because we know that it has finite*

expectation:

$$\begin{aligned}\mathbf{E}[f(B\sigma_k)] &= f(a) + \mathbf{E}[\int_0^{\sigma_k} \frac{1}{2}\Delta f(B_s)ds] \\ &= |a|^2 + \mathbf{E}[\int_0^{\sigma_k} n \cdot ds] = |a|^2 + n\mathbf{E}[\sigma_k].\end{aligned}$$

One sees this easily be computing the 2_{nd} *derivatives of the square of the norm in* $\mathbb{R}^n$ *and summing them up. If we now think of the fact that* $|x| \le R$ *until the stopping time and that* $E[|x|^2] \le R^2$ *we arrive at* $\mathbf{E}[\sigma_k] \le \frac{1}{n}(R^2 - |a|^2)$ *for all* k*. For* $k \to \infty$ $\sigma_k \to \tau_K$ *a.s. we see that*

$$\mathbf{E}[\tau_K] = \frac{1}{n}(R^2 - |a|^2).$$

This follows from the fact that $|x|$ *is not less or equal* R *like in the case of the* σ_k *where we took the minimum of* τ *and* k *but if we stop at* τ $|x| = R$ *holds.*

We can answer more questions by applying Dynkin's formula:
Assume $n \ge 2$ *and* $|b| > R$ *- what is the probability that a BM starting at b hits the ball* K *? Let* α_k *be the first exit time from the annulus* $A_k = \{x, R < |x| < 2^k R\}$*;* $k = 1, 2, \ldots$ *and we put* $T_K = \inf\{t > 0; X_t \in K\}$*. Next we choose a function* f *- again with compact support. We choose it for* x *in the annulus in the following way:*
f(x) $= \begin{cases} -log|x| \text{ for } & n = 2 \\ |x|^{2-n} \text{ for } & n > 2 \end{cases}$*. Dynkin's formula simplifies, because of* $\triangle f = 0$ *inside of the* A_k'*s, as basic calculations show to*

$$\mathbf{E}[f(B_{\alpha_k})] = f(b) \quad \textit{for all } k.$$

The BM can exit the annulus either to the ball inside it or to the exterior, thus we define

$$p_k = \mathbb{P}(|B\alpha_k| = R), q_k = \mathbb{P}(|B\alpha_k| = 2^k R).$$

We have to consider the case $n = 2$ *and* $n \ge 2$ *separately:*
$\mathbf{n = 2}$:

$$\begin{aligned}\mathbf{E}[f(B_{\alpha_k})] &= f(b) \\ -\log R \cdot p_k - (\log R + k \cdot \log 2)q_k &= -\log|b|\end{aligned}$$

for all k*.*

So for $k \to \infty$ q_k has to tend to 0, so that

$$\begin{aligned}p_k &\to 1 \\ \mathbb{P}(T_K < \infty) &= 1.\end{aligned}$$

This phenomenon of coming back a.s. is called recurrence. The BM is recurrent in $\mathbb{R}^2$. Now the case
$\mathbf{n > 2}$:
In the same way as before we get that

$$p_k \cdot R^{2-n} + q_k \cdot (2^k R)^{2-n} = |b|^{2-n}.$$

Since $0 \le q_k \le 1$ we get by $k \to \infty$

$$\lim_{k\to\infty} p_k = \mathbb{P}(T_K < \infty) = (\frac{|b|}{R})^{2-n}.$$

So it does not come back with probability one. This is called transience and we conclude that the BM is transient in $\mathbb{R}^n$ for $n > 2$.

1.6.4 The Markov Property

To define this property we need some notation: $X_t^{s,x}$ is the process at time t that had the value x at the point in time s.

Definition 1.6.3. *An Itô process X_t is called time-homogeneous diffusion if $X_t(\omega) = X(t,\omega) : [0,\infty) \times \Omega \to \mathbb{R}^n$ satisfies the SDE of the form*

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t, \quad t \ge s, \quad X_s = x. \tag{1.49}$$

Where B_t is a m-dimensional BM and $b : \mathbb{R}^n \to \mathbb{R}^n$ and $\sigma : \mathbb{R}^n \to \mathbb{R}^{n\times m}$. Note that the associated vector fields do not depend on t explicitly. This leads to time-homogeneity in the following sense:

$$\begin{aligned} X_{s+h}^{s,x} &= x + \int_s^{s+h} b(X_u^{s,x})du + \int_s^{s+h} \sigma(X_u^{s,x})dB_u \\ &= x + \int_0^h b(X_{s+v}^{s,x})dv + \int_0^h \sigma(X_{s+v}^{s,x})d\widetilde{B}_v, \quad (u = s+v) \end{aligned} \tag{1.50}$$

where $\widetilde{B}_v = B_{s+v} - B_s; v \ge 0$ and the filtration is $(\widetilde{\mathcal{F}}_v) = (\mathcal{F}_{s+v})$.

It is clear that $(\widetilde{B}_v)_{v\ge0}$ and $(B_v)_{v\ge0}$ have the same $\mathbb{P}$-distribution and therefore $(X_{s+h}^{s,x})_{h\ge0}$ and $(X_h^{0,x})_{h\ge0}$ have the same $\mathbb{P}$-distribution.

Theorem 1.6.4 (The Markov Property for Itô Diffusions)**.** *Let f denote a Borel function from $\mathbb{R}^n$ to $\mathbb{R}$ and X_t an Itô diffusion then for $t, h \ge 0$*

$$\mathbf{E}[f(X_{t+h})|\mathcal{F}_t] = \mathbf{E}[f(X_h)]|_{X_0=X_t(\omega)}. \tag{1.51}$$

Proof. We have for $r \geq t$

$$X_r(\omega) = X_t(\omega) + \int_t^r b(X_u)du + \int_t^r \sigma(X_u)dB_u$$

and by uniqueness

$$X_r(\omega) = X_r^{t,X_t}(\omega).$$

If we define

$$F(x,t,r,\omega) = X_r^{t,x}(\omega); \quad r \geq t,$$

then we can rewrite the theorem as

$$\mathbf{E}[f(F(X_t,t,t+h,\omega))|\mathcal{F}_t] = \mathbf{E}[f(F(x,0,h,\omega))]|_{x=X_t}.$$

Now we put

$$g(x,\omega) = f \circ F(x,t,t+h,\omega)$$

then g is measurable.

So we can approximate g by bounded functions of the form

$$\sum_{k=1}^{m} \phi_k(x)\psi_k(\omega).$$

Hence by usage of the properties of the conditional expectation we get

$$\begin{aligned}
\mathbf{E}[g(X_t,\omega)|\mathcal{F}_t] &= \mathbf{E}[\lim \sum \phi_k(X_t)\psi_k(\omega)|\mathcal{F}_t] \\
&= \lim \sum \phi_k(X_t)\mathbf{E}[\psi_k(\omega)|\mathcal{F}_t] \\
&= \lim \sum \mathbf{E}[\phi_k(y)\psi_k(\omega)|\mathcal{F}_t]|_{y=X_t} \\
&= \mathbf{E}[g(y,\omega)|\mathcal{F}_t]|_{y=X_t} \underbrace{=}_{g \text{ indep. of } \mathcal{F}_t} \mathbf{E}[g(y,\omega)]|_{y=X_t}.
\end{aligned}$$

Finally by time-homogeneity of the process we get

$$\begin{aligned}
\mathbf{E}[f(F(X_t,t,t+h,\omega))|\mathcal{F}_t] &= \mathbf{E}[f(F(y,t,t+h,\omega))]|_{y=X_t} \\
&= \mathbf{E}[f(F(y,0,h,\omega))]|_{y=X_t}.
\end{aligned}$$

□

The next theorem shows that this property can be generalized from a fixed point in time t to a stopping time τ. But before that we need the definition of the filtration of a stopping time.

Definition 1.6.4. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and $\mathcal{F}_t$ a filtration. Let τ be a stopping time w.r.t. $\mathcal{F}_t$. Then the sigma-algebra $\mathcal{F}_\tau$ consists of the sets $A \in \mathcal{F}$ satisfying*

$$A \bigcap \{\tau \leq t\} \in \mathcal{F}_t \quad \text{for all } t \geq 0. \tag{1.52}$$

Theorem 1.6.5 (The Strong Markov Property)**.** *Let f be a bounded Borel function on $\mathbb{R}^n$, τ a stopping time $< \infty$ a.s. Then*

$$\mathbf{E}[f(X_{\tau+h})|\mathcal{F}_\tau] = \mathbf{E}[f(X_h)]|_{X_0 = X_\tau}, \qquad \text{for all } h \geq 0. \tag{1.53}$$

Proof. The proof tries to redo the prove of the Markov Property but needs more technical care in order to deal with the stopping time. It can be found in Øksendal [13] for example. □

Summary 1.5. *In this section we learned about the meaning of the Stratonovich corrected Itô drift to understand the geometric behavior of a diffusion. Then we got to know Dynkin's formula which is a useful tool in combination with stopping times. We got some results on the difference of the geometric behavior of the BM in 2 dimensions and in more than 2 dimensions. Finally we saw that an Itô diffusion fulfills the Markov property.*

1.7 Girsanov's Theorem

This theorem has a lot of applications in finance and other fields where stochastic calculus is used, so it is essential to quote a version of it.

Theorem 1.7.1 (Girsanov's Theorem)**.** *Let $Y_t \in \mathbb{R}^n$ be a n-dimensional process of the form:*

$$dY_t = a_t dt + dB_t; \quad t \leq T, Y_0 = 0.$$

Where T is a given constant $\leq \infty$ and B_t is an n-dimensional Brownian motion. Put

$$M_t = \exp(-\int_0^t a_s dB_s - \frac{1}{2}\int_0^t a_s^2 ds); \quad t \leq T.$$

Moreover assume that $a(s, \omega)$ satisfies the so called Novikov condition

$$\mathbf{E}[\exp(\frac{1}{2}\int_0^T a_s^2 ds)] < \infty.$$

Where $\mathbf{E} = \mathbf{E}_\mathbb{P}$ denotes the expectation w.r.t. $\mathbb{P}$ - the original probability law - then we can define a measure Q on $(\Omega, \mathcal{F}_T)$ by

$$d\mathbb{Q} = M_T d\mathbb{P}.$$

Then Y_t is a Brownian Motion w.r.t. to the probability law Q for $t \leq T$.

Remark 1.7.1. *So what does this theorem tell us? It says: If I change the drift of BM from 0 to a drift a_t then how does the law change and can I find a law such that it remains a BM? Note that the change of measure corresponds to the stochastic exponential of the process $-a_t$! Furthermore compare this to the standardization of a random variable $X \sim N(\mu, \sigma^2)$.*

1.8 Martingale Representation Theorem

Definition 1.8.1 (Brownian martingale)**.** *A Brownian martingale is a martingale w.r.t. to the filtration induced by the Brownian Motion.*

Theorem 1.8.1 (Martingale Representation Theorem for Brownian Martingales)**.** *Let $(\mathcal{F}_t)_{t=0}^T$ be the augmented (set of $\mathbb{P}$-null sets added)Brownian filtration assume that M_t is a square-integrable Brownian martingale, then there exists an adapted process ϕ_t with*

$$\mathbf{E}[\int_0^T \phi_t^2 dt] < \infty$$

and

$$M_t = M_0 + \int_0^T \phi_s dB_s.$$

Remark 1.8.1. *There is a more general version of this theorem where martingales as such are covered, but we don't need it in this text.*

Proof. **part (1)**: First assume $M_0 = 0 \Longrightarrow \mathbf{E}[M_T|\mathcal{F}_t] = M_t, \mathbf{E}[M_T] = 0$. Next consider the space $L^2([0,T]) = \{H|H$ is adapted and fulfills $\mathbf{E}[\int_0^T H_s^2 ds] < \infty\}$
which is a Hilbert space with $\langle H, \tilde{H}\rangle = \mathbf{E}[\int_0^T H_s \tilde{H}_s ds]$.
Moreover consider $L^{2,0}(\Omega, \mathcal{F}_T, \mathbb{P}) = L^2(\Omega, \mathcal{F}_T, \mathbb{P}) \bigcap \{R|\mathbf{E}[R] = 0\}$ the space of centered L^2 random variables.
We define a map:

$$I : L^2([0,T]) \to L^{2,0}(\Omega, \mathcal{F}_T, \mathbb{P})$$
$$I(X) \mapsto \int_0^T X_s dW_s.$$

$L^2([0,T])$ is a Hilbert space and the map I is an isometry (the Itô isometry) therefore $I(L^2([0,T])) =: V$ is a complete (closed) subspace of $L^{2,0}(\Omega, \mathcal{F}_T, \mathbb{P})$.

To show the theorem it is enough to show that $V = L^{2,0}(\Omega, \mathcal{F}_T, \mathbb{P})$, because this implies that each centered square integrable r.v. is representable as a stochastic integral w.r.t. BM. Especially: $M_T = \int_0^T \phi_t dB_t \Longrightarrow M_t = \int_0^t \phi_s dB_s$.

We show this now: Let Z be an element of $L^{0,2} \bigcap V^\perp$ (the orthogonal complement of V). To show that $L^{0,2} = V$ we have to show that $Z = 0$. Consider $Z_t := \mathbf{E}[Z|\mathcal{F}_t]$ then Z_t is a right continuous martingale by assumption on the filtration (right continuity).

Let $H \in L^2([0,T])$: $N_T = \int_0^T H_t dB_t$ and $N_t = \mathbf{E}[N_T|\mathcal{F}_t] = \int_0^t H_s dB_s$ (which is a continuous martingale). Moreover τ is a bounded stopping time $\Longrightarrow N_\tau = \int_0^\tau H_t dB_t = \int_0^T \mathbf{1}_{[0,\tau]}(t) H_t dB_t$ and therefore $N_\tau \in V$ and

$$\begin{aligned} 0 \underbrace{=}_{\text{orthogonality}} & \mathbf{E}[ZN_\tau] \\ \underbrace{=}_{\text{tower property}} & \mathbf{E}[N_\tau \mathbf{E}[Z|\mathcal{F}_\tau]] \\ \underbrace{=}_{\text{per definition}} & \mathbf{E}[N_\tau Z_\tau]. \end{aligned}$$

This holds for all bounded stopping times, therefore $N_t Z_t$ is a martingale.

part (2): Consider the function $f(t,x) = \exp(i\theta x + \frac{1}{2}\theta^2 t)$ with $i = \sqrt{-1}$ and $\theta \in \mathbb{R}$. Then we define

$$\begin{aligned} M_t^\theta := f(t, B_t) &= \exp(i\theta B_t + \frac{1}{2}\theta^2 t) \\ &= \exp(\frac{1}{2}\theta^2 t) \cdot (\cos(\theta B_t) + i \sin(\theta B_t)). \end{aligned}$$

Obviously $|M_t^\theta| = \exp(\frac{1}{2}\theta^2 t)$ is bounded for $t \le T$. By Itô we get:

$$M_t^\theta = 1 + i \cdot \theta \int_0^t f(s, B_s) dB_s + \frac{1}{2}\int_0^t \theta^2 f(s, B_s) ds - \frac{1}{2}\int_0^t \theta^2 f(s, B_s) ds.$$

Remember that a continuous local martingale is a martingale if it is bounded therefore M_t^θ is a martingale for all $\theta \in \mathbb{R}$ and by part(1) $\mathbf{E}[Z_t M_t^\theta] = Z_s M_s^\theta$

$$\begin{aligned} \Longrightarrow \mathbf{E}[Z_t \exp(i \cdot \theta B_t + \frac{1}{2}\theta^2 t)|\mathcal{F}_s] &= Z_s \cdot \exp(i \cdot \theta B_s + \frac{1}{2}\theta^2 s) \\ \Longrightarrow \mathbf{E}[Z_t \exp(i \cdot \theta (B_t - B_s))|\mathcal{F}_s] &= Z_s \cdot \exp(-\frac{1}{2}\theta^2 (t-s)). \end{aligned} \tag{1.54}$$

Now we look at the partition: $0 = t_0 < t_1 < \ldots < t_n = T$ and $B_{t_k} - B_{t_{k-1}} =: \Delta_k$

$$\begin{aligned} \mathbf{E}[Z_T \cdot \exp(i \sum_{k=1}^n \theta_k \Delta_k)] &= \mathbf{E}[\underbrace{\mathbf{E}[Z_T \cdot \exp(i\theta_n \Delta_n)|\mathcal{F}_{t_{n-1}}]}_{=(1.54) Z_{t_{n-1}} \cdot \exp(-\frac{1}{2}\theta_n^2 (T - t_{n-1}))} \exp(i \sum_{k=1}^{n-1} \theta_k \Delta_k)] \\ = \ldots &= \exp(i \sum_{k=1}^n \theta_k^2 (t_k - t_{k-1}) \cdot \mathbf{E}[Z_0] \underbrace{=}_{Z_0 = 0 \text{ a.s.}} 0 \end{aligned}$$

$\Longrightarrow Z_T \perp \exp(i \sum_{k=1}^n \theta_k \Delta_k), \forall n, \forall \theta_k \in \mathbb{R}$. Which finally gives together with the uniqueness of the Fourier transform that $Z_T = 0$ a.s.

uniqueness: Assume that there exist two processes ϕ, ϕ' such that

$$\begin{aligned}
\int_0^t \phi_s dB_s &= \int_0^t \phi'_s dB_s. \text{ Then} \\
\int_0^t (\phi_s - \phi'_s) dB_s &= 0 \text{ a.s} \\
\Longrightarrow \mathbf{E}[(\int_0^t (\phi_s - \phi'_s) dB_s)^2] &\underbrace{=}_{\text{Itô}} \mathbf{E}[\int_0^t (\phi_s - \phi'_s)^2 ds] = 0 \\
\Longrightarrow \phi_s = \phi'_s & \quad ds \otimes d\mathbb{P} \text{ a.s.}
\end{aligned}$$

□

Chapter 2

Interest Rate Theory

2.1 Zero-coupon Bonds, Forward Rates and the Short Rate

The sources of the following theorems were the lecture of Prof.Grandits [4], the book of Lamberton and Lapeyre [10], the book by Kijima [9] and the book of Musiela and Rutkowski [11]. Many concepts of interest rates have been developed so far. I decided to explain the concept of a zero-coupon bond first. Then the abstract notion of a instantaneous forward rate and finally the rather classical notion of an instantaneous short rate.

2.1.1 Zero-Coupon Bonds

Let T^* be a fixed point in time as ultimate time horizon for all market activities.

Definition 2.1.1 (Zero-coupon Bond). *A zero-coupon bond of maturity T defines a security paying one unit of cash to its holder at the fixed date T in the future.*

Remark 2.1.1. *The expression* zero-coupon bond *will often be abbreviated by z.c.b. - zero-coupon bonds are also called discount bonds. No intermediate payments (coupon payments) are done.*

By convention the bond's principal (also called face value or nominal value) will be 1 unit of a currency. The price of a z.c.b at time $t \leq T$ will be denoted by $P(t,T)$. It is obvious that $P(T,T) = 1$ for any $T \leq T^*$. We will assume that for any $T \leq T^*$ the bond price $P(s,T)_{s=0}^{T}$ follows a strictly positive and adapted process on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a filtration $\mathcal{F} = (\mathcal{F}_t)_{t=0}^{T^*}$. Two kinds of bonds are distinguished - corporate bonds and default-free bonds. In corporate bonds there is the risk of default, whereas default-free bonds are issued by governments and other institutions such that one can assume that this default risk vanishes.

2.1.2 Term Structure of Interest Rates

If we consider a z.c.b. with maturity $T \leq T^*$ then the simple rate of return from holding the bond over the period $[t, T]$ is given by:

$$\frac{1 - P(t,T)}{P(t,T)} \quad = \quad \frac{1}{P(t,T)} - 1.$$

Compare this to the solution for r of

$$P(t,T)(1 + r) = 1.$$

The equivalent rate of return with continuous compounding is commonly referred to as a yield to maturity on a bond. This is defined formally in the following definition.

Definition 2.1.2. *The adapted process $Y(t,T)$ defined by*

$$Y(t,T) = -\frac{1}{T-t} \ln P(t,T), \quad \forall t \in [0,T), \tag{2.1}$$

is called the yield-to-maturity on a zero-coupon bond maturing at time T.

Remark 2.1.2. *Compare this to the solution for r of*

$$P(t,T) \exp(r(T-t)) = 1.$$

The term structure of interest rates, known as yield curve, is the function that relates the yield $Y(t,T)$ to the maturity T. It is obvious that for an arbitrary fixed maturity T and a given yield to maturity $Y(t,T)$

$$P(t,T) = \exp(-Y(t,T)(T-t)), \quad \forall t \in [0,T)$$

holds. This discount function relates a bond price to a maturity.

In practice the yield curve is derived from the prices of several actively traded interest rate instruments. It is determined for one day only by the prices quoted on that day. The shape of a historically observed yield curve varies over time. This shows the complexity of finding a model for the stochastic behavior of the term structure of interest rates.

2.1.3 Forward Interest Rates

We introduce the notion of a forward rate $f(t,T)$ at time $t \leq T$. It should be interpreted as the interest rate over the infinitesimal interval $[T, T + dT]$ as seen from time t. It is a mathematically idealized concept rather than an observable quantity. But widely accepted due to an approach to modeling a bond price by Heath, Jarrow and Morton.

Given a family $f(t,T), t \leq T \leq T^*$ of instantaneous forward rates the bond prices are defined by

$$P(t,T) = \exp(-\int_t^T f(t,u)du), \quad \forall t \in [0,T]. \tag{2.2}$$

So if we consider the family of bond prices $P(t,T)$ to be C^1 w.r.t. T then the implied instantaneous forward interest rate $f(t,T)$ can be defined.

Definition 2.1.3 (Forward Rate)**.** *The family of instantaneous forward rates $f(t,T), t \leq T \leq T^*$ is formally defined by*

$$f(t,T) = -\frac{\partial \ln P(t,T)}{\partial T}. \tag{2.3}$$

Remark 2.1.3. *Thus we assume the map $T \mapsto P(t,T)$ to be C^1.*

2.1.4 Short-term Interest Rate

This notion is the most traditional. We denote the instantaneous interest rate (short term interest rate, spot interest rate) by r_t. Meaning the rate for a risk-free borrowing or lending over the infinitesimal interval $[t, t+dt]$. In a stochastic setup it is an adapted process defined on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We moreover assume r_t to be a stochastic process such that almost all sample paths are integrable on $[0, T^*]$ w.r.t. the Lebesgue measure. With these assumptions we can introduce the adapted process P of finite variation and with continuous sample paths given by

$$P_t = \exp(\int_0^t r_u du), \quad \forall t \in [0, T^*]. \tag{2.4}$$

Equivalently for almost all $\omega \in \Omega$ P_t solves the differential equation

$$dP_t = r_t P_t dt$$

with $P_0 = 1$ by convention. Financially it can be interpreted as the price process of a risk-free security which continuously compounds at the rate r_t.

Remark 2.1.4. *The connection to a forward rate is given by $r_t = f(t,t)$.*

2.1.5 Arbitrage-free Pricing of Zero-coupon Bonds

The price processes $P(t,T)$ fulfill the properties described before and moreover we specialize the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ the be filtered by the $\mathbb{P}$-completed filtration generated by the Brownian motion. We will write

$$P^*(t,T) = \frac{P(t,T)}{P_t} \quad \text{with } P_t = \exp(\int_0^t r_u du) \tag{2.5}$$

for the discounted price of the z.c.b. with maturity T.

Definition 2.1.4. *For two probability measures on a measurable space* $(\Omega, \mathcal{F})$ $\mathbb{Q} \sim \mathbb{P}$ *means that the measures have the same null sets.*

In interest rate theory arbitrage free pricing is difficult to characterize. Therefore we define it as the existence of an equivalent martingale measure.

Definition 2.1.5 (Arbitrage Free Prices of Zero Coupon Bonds)**.** *A family of zero-coupon bond prices is arbitrage free (relative to r) if there exists a measure $\mathbb{Q} \sim \mathbb{P}$ such that the discounted bond prices $P^*(t,T)$ are martingales.*

Conclusion 2.1. *Let $P(t,T)$ a family of bond prices. Then the following statements are equivalent:*

1. *the family is arbitrage free*
2. $\exists \mathbb{Q} \sim \mathbb{P}$ *such that* $P(t,T) = \mathbf{E}_{\mathbb{Q}}[\exp(-\int_t^T r_u du)|\mathcal{F}_t]$.

Proof. (1) $\Longrightarrow$ (2) : $\exists \mathbb{Q} : P^*(t,T)$ is a $\mathbb{Q}$-martingale (definition of arbitrage freeness)

$$\begin{aligned} P(t,T) \quad & \underbrace{=}_{\text{by def.}} \quad P_t \cdot P^*(t,T) = P_t \mathbf{E}_{\mathbb{Q}}[P^*(T,T)|\mathcal{F}_t] = \\ P_t \mathbf{E}_{\mathbb{Q}}[\frac{P(T,T)}{P_T}|\mathcal{F}_t] \quad & \underbrace{=}_{P(T,T)=1} \quad \mathbf{E}_{\mathbb{Q}}[\frac{P_t}{P_T}|\mathcal{F}_t] = \mathbf{E}_{\mathbb{Q}}[\exp(-\int_t^T r_u du)|\mathcal{F}_t)] \end{aligned}$$

(2) $\Longrightarrow$ (1) :

$$\begin{aligned} \mathbf{E}_{\mathbb{Q}}[P^*(T,T)|\mathcal{F}_t] \quad & \underbrace{=}_{\text{by def.}} \quad \mathbf{E}_{\mathbb{Q}}[\frac{P(T,T)}{P_T}|\mathcal{F}_t] = \\ \mathbf{E}_{\mathbb{Q}}[\frac{P_t}{P_t P_T}|\mathcal{F}_t] \quad & = \quad \frac{1}{P_t}\mathbf{E}_{\mathbb{Q}}[\exp(-\int_t^T r_u du)|\mathcal{F}_t] \underbrace{=}_{\text{by assertion 2}} \\ \frac{1}{P_t} P(t,T) \quad & = \quad P^*(t,T) \end{aligned}$$

So the discounted bond prices are martingales w.r.t. $\mathbb{Q}$. □

We now consider interest rate markets driven by a d-dimensional Brownian motion.

Lemma 2.1.1. *Let* $(M)_{t=0}^T$ *be a continuous martingale and let*

$$\mathbb{P}(M_t > 0) = 1 \ \forall t \in [0,T] \Longrightarrow \mathbb{P}(M_t > 0, \forall t \in [0,T]) = 1.$$

Proof. Just a sketch of the proof: Take a stopping time that indicates the time, when the process equals zero. Use Doob's optional stopping and by definition of the conditional expectation and continuity you get the result. □

Lemma 2.1.2. *Let $(\mathcal{F}_t)_{t=0}^T$ be the augmented Brownian filtration moreover let $\mathbb{Q} \ll \mathbb{P}$ and consider $L_t = \mathbf{E}_{\mathbb{P}}[\frac{d\mathbb{Q}}{d\mathbb{P}}|\mathcal{F}_t]$, the density process, then the following notions are equivalent:*

1. $\mathbb{Q} \sim \mathbb{P}$
2. $\exists q_s$ *with* $\mathbb{P}(\int_0^T q_s ds < \infty) = 1$ *such that* $L_t = \exp(\int_0^t q_s dB_s - \frac{1}{2}\int_0^t q_s^2 ds)$.

Proof. $(\mathbf{2}) \Longrightarrow (\mathbf{1})$:
This is clear since $L_t > 0$ a.s. i.e. the two measures have the same null sets.
$(\mathbf{1}) \Longrightarrow (\mathbf{2})$:
Thinking of the martingale representation theorem we know that

$$\exists \gamma_s \text{ such that } L_t = L_0 + \int_0^t \gamma_s dB_s$$

with $\mathbb{P}(\int_0^T \gamma_s^2 ds < \infty) = 1$. We know moreover that $L_0 = \mathbf{E}_{\mathbb{P}}[\frac{d\mathbb{Q}}{d\mathbb{P}}|\mathcal{F}_0] = \mathbf{E}_{\mathbb{P}}[\frac{d\mathbb{Q}}{d\mathbb{P}}] = 1$. From (1) we have that $L_t > 0$ a.s. $\forall t \in [0,T]$. So by lemma (2.1.1) we have that $\mathbb{P}(L_t > 0, \forall t \in [0,T]) = 1$. Now we apply Itô's formula with $f(x) = \ln(x)$:

$$\begin{aligned}
\ln(L_t) &= \underbrace{0}_{\ln(L_0)=\ln(1)} + \int_0^t \underbrace{\frac{1}{L_s}\gamma_s}_{:=q_s} dB_s - \frac{1}{2}\int_0^t \underbrace{\frac{1}{L_s^2}\gamma_s^2}_{=q_s^2} ds \\
\ln(L_t) &= \int_0^t q_s dB_s - \frac{1}{2}\int_0^t q_s^2 ds \\
L_t &= \exp(\int_0^t q_s dB_s - \frac{1}{2}\int_0^t q_s^2 ds)
\end{aligned}$$

with $\mathbb{P}(\int_0^T q_s^2 ds < \infty) = 1$ and q_s well-defined ensured by lemma (2.1.1). □

The main goal in the following lemma and theorem is to find an SDE with respect to the physical measure $\mathbb{P}$.

Lemma 2.1.3. *Let $P(t,T)$ be a family of arbitrage-free bond prices, then we have the representation:*

$$P(t,T) = \mathbf{E}_{\mathbb{P}}[\exp(-\int_t^T r_u du + \int_t^T q_u dB_u - \frac{1}{2}\int_t^T q_u^2 du)|\mathcal{F}_t]$$

where r_u is the overnight rate and q_u is from the equivalent martingale measure of lemma (2.1.2).

Proof.

$$P(t,T) = \mathbf{E}_{\mathbb{Q}}[\exp(-\textstyle\int_t^T r_u du)|\mathcal{F}_t] \underbrace{=}_{\text{Bayes' formula (lemma 4.2.1)}} \mathbf{E}_{\mathbb{P}}[\exp(-\textstyle\int_t^T r_u du)\frac{L_T}{L_t}|\mathcal{F}_t] =$$

$$= \mathbf{E}_{\mathbb{P}}[\exp(-\textstyle\int_t^T r_u du) \cdot \exp(\int_0^T q_u dB_u - \frac{1}{2}\int_0^T q_u^2 du) \cdot \exp(-\int_0^t q_u dB_u + \frac{1}{2}\int_0^t q_u^2 du)|\mathcal{F}_t] =$$

$$= \mathbf{E}_{\mathbb{P}}[\exp(-\textstyle\int_t^T r_u du + \int_t^T q_u dB_u - \frac{1}{2}\int_t^T q_u^2 du)|\mathcal{F}_t]$$

$\square$

Remark 2.1.5. *In this proof Bayes' formula (lemma 4.2.1) is applied. See the appendix for the proof and details.*

Theorem 2.1.1. *Let $P(t,T)$ be an arbitrage-free family of zero coupon bonds, then there exists for all T an adapted process σ_t^T with the integrability assumption*

$$\mathbb{P}(\int_0^T (\sigma_t^T)^2 ds < \infty) = 1$$

such that

$$P(t,T) = P(0,T) + \int_0^t (r_u - q_u\sigma_u^T)P(u,T)du + \int_0^t \sigma_u^T P(u,T)dB_u \tag{2.6}$$

or in other terms:

$$\frac{dP(t,T)}{P(t,T)} = \underbrace{(r_t - q_t\sigma_t^T)}_{\text{so this is the yield in expectation}} dt + \sigma_t^T dB_t.$$

Proof. $P^*(t,T)$ is a $\mathbb{Q}$-martingale $\implies P^*(t,T) \cdot L_t$ is a $\mathbb{P}$-martingale. Thus by the martingale representation theorem and lemma (2.1.2) we get that

$$\begin{aligned}
\exists\theta_t^T \text{ such that } \frac{P^*(t,T)L_t}{P^*(0,T)\underbrace{L_0}_{=1}} &= \exp(\int_0^t \theta_u^T dB_u - \frac{1}{2}\int_0^t (\theta_u^T)^2 du) \\
\implies \frac{P^*(t,T)}{P^*(0,T)} &= \exp(\int_0^t (\theta_u^T - q_u)dB_u - \frac{1}{2}\int_0^t ((\theta_u^T)^2 - q_u^2)du).
\end{aligned}$$

Thinking of $P^*(t,T) = \frac{P(t,T)}{\exp(-\int_0^t r_u du)}$ we get

$$\begin{aligned}
\frac{P(t,T)}{P(0,T)} &= \exp(\int_0^t r_u - \frac{1}{2}((\theta_u^T)^2 - q_u^2)du + \int_0^t (\theta_u^T - q_u)dB_u) \\
&=: \exp(X_t).
\end{aligned}$$

Then by Itô's formula it follows that

$$\begin{aligned}
X_t &= \int_0^t K_s ds + \int_0^t H_s dB_s \\
\exp(X_t) &= 1 + \int_0^t \exp(X_s)K_s ds + \int_0^t \exp(X_s)\frac{1}{2}H_s^2 ds + \int_0^t \exp(X_s)H_s dB_s.
\end{aligned}$$

Finally we find that

$$\begin{aligned}
\frac{P(t,T)}{P(0,T)} &= 1+\int_0^t \frac{P(s,T)}{P(0,T)}K_s ds+\frac{1}{2}\int_0^t \frac{P(s,T)}{P(0,T)}H_s^2 ds+\int_0^t \frac{P(s,T)}{P(0,T)}H_s dB_s \\
P(t,T) &= P(0,T)+\int_0^t P(s,T)[r_s-\frac{1}{2}((\sigma_s^T)^2-q_s^2)+\frac{1}{2}(\theta_s^T-q_s)^2]ds+ \\
& \int_0^t P(s,T)\underbrace{(\theta_s^T-q_s)}_{=\sigma_s^T}dB_s \\
&= P(0,T)+\int_0^t P(s,T)(r_s+q_s^2-\theta_s^T q_s)ds+\int_0^t P(s,T)\sigma_s^T dB_s \\
&= P(0,T)+\int_0^t P(s,T)(r_s-q_s\sigma_s^T)ds+\int_0^t P(s,T)\sigma_s^T dB_s.
\end{aligned}$$

□

Remark 2.1.6. *The $-q_s$ in this formula can be seen as some kind of risk premium.*

Theorem 2.1.2. *Let $\tilde{B}_t := B_t-\int_0^t q_u du$ then*

1. *$\tilde{B}_t$ is a $\mathbb{Q}-$Brownian Motion.*

2. *the formula of theorem 2.1.1 becomes:*

$$\frac{dP(t,T)}{P(0,T)} = r_t dt+\sigma_t^T d\tilde{B}_t.$$

Proof. 1. $L_t = \mathbf{E}[\frac{d\mathbb{Q}}{d\mathbb{P}}|\mathcal{F}_t]$ which equals the stochastic exponential of q_s so the assertion follows by Girsanov's theorem.
2.

$$\begin{aligned}
\frac{dP(t,T)}{P(0,T)} &= (r_t-q_t\sigma_t^T)dt+\sigma_t^T(d\tilde{B}_t+q_t dt) \\
&= r_t dt+\sigma_t^T d\tilde{B}_t.
\end{aligned}$$

□

Remark 2.1.7. *There are models that focus on the short rate (e.g. Vasicek Model, Cox-Ingersoll-Ross Model) and others that focus on the forward rate (the so called Heath-Jarrow-Morton methodology).*

2.2 LIBOR Rates

The notions described in the section before assume the existence of instantaneous rates. This assumption requires a certain degree of smoothness with respect to the tenor (i.e.maturity) of bond prices and their volatilities. Thus the step that we want to do now is to construct families of arbitrage free bond prices without referring to instantaneous, continuously compounded rates, which is more suitable in some circumstances.

2.2.1 The Mathematical Setting

We assume a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a filtration $(\mathcal{F}_t)_{t\in[0,T^*]}$ which fulfills the usual conditions. Moreover the process B is a $d-$dimensional standard Brownian motion defined on this probability space and the filtration is the $\mathbb{P}$-augmented natural filtration of B. We will use the following notations:

$\mathcal{M}_{loc}(\mathbb{P})$ for the class of all real valued local martingales

$\mathcal{M}(\mathbb{P})$ for the class of all real valued martingales

$\mathcal{V}$ for the class of all real valued adapted processes of finite variation

$\mathcal{A}$ for the class of all real valued predictable processes of finite variation

$\mathcal{S}_p(\mathbb{P})$ for the class of real valued special semi-martingales, i.e. $X \in \mathcal{S}_p(\mathbb{P})$ means that X admits a decomposition $X = X_0 + M + A$ where $M \in \mathcal{M}(\mathbb{P})$ and $A \in \mathcal{A}$.

The superscript $^+$ stands for the collection of strictly positive processes of a certain class. The subscript $_c$ stands for the collection of processes of a certain class with continuous sample paths. For example $\mathcal{M}(\mathbb{P})^+_c$ will denote the class of strictly positive martingales with continuous paths. Attention: $\mathcal{S}^+_p(\mathbb{P})$ will denote the class of special martingales which are strictly positive and (!) the process of left-hand limits is also strictly positive. Note that the class $\mathcal{S}_p(\mathbb{P})$ as well as $\mathcal{S}^+_p(\mathbb{P})$ is invariant w.r.t. a change of measure to an equivalent measure. $\mathbb{Q}$ and $\mathbb{P}$ are equivalent if the Radon-Nikodym derivative

$$\Lambda_t = \mathbf{E}[\frac{d\mathbb{Q}}{d\mathbb{P}}|\mathcal{F}_t] \quad \forall t \in [0, T^*]$$

is a locally bounded process. Since the filtration is the one generated by the Brownian motion, Λ_t will follow a continuous exponential martingale - hence it will be locally bounded (see the section on the martingale representation theorem, section 1.8). Therefore we will simply write $\mathcal{S}_p$ and $\mathcal{S}^+_p$.

Assumptions on the Bond Price

The family of bond prices fulfills the assumptions that we had before. It is a family of strictly positive real-valued adapted processes $P(t,T), t \in [0,T]$ with $P(T,T) = 1$ for every $T \in [0,T^*]$. In this section the family of bond prices will be assumed to be given - meaning already constructed by a certain procedure. But we make the following assumptions:

(BP.1) For any maturity date $T \in [0,T^*]$, the bond price $P(t,T), t \in [0,T]$ belongs to the class S_p^+.

(BP.2) For any fixed $T \in [0,T^*]$, the forward process

$$F_P(t,T,T^*) \quad := \quad \frac{P(t,T)}{P(t,T^*)}, \quad \forall t \in [0,T]$$

follows a martingale under $\mathbb{P}$, or equivalently

$$P(t,T) \quad = \quad \mathbf{E}_{\mathbb{P}}[\frac{P(t,T^*)}{P(T,T^*)}|\mathcal{F}_t], \quad \forall t \in [0,T]. \tag{2.7}$$

Remark 2.2.1. *We see the equivalence by:*

$$\mathbf{E}_{\mathbb{P}}[F_P(T,T,T^*)|\mathcal{F}_t] \underbrace{=}_{martingale\ property} F_P(t,T,T^*) = \tfrac{P(t,T)}{P(t,T^*)}$$
$$\Leftrightarrow$$
$$\mathbf{E}_{\mathbb{P}}[F_P(T,T,T^*)P(t,T^*)|\mathcal{F}_t] = P(t,T) \text{ by adaptedness of } P(t,T^*).$$

So by those assumptions the introduced process $F_P(t,T,T^*), t \in [0,T]$ follows a continuous, strictly positive $\mathbb{P}$-martingale w.r.t. the filtration generated by the d-dimensional BM, so that $F_P \in \mathcal{M}_c^+(\mathbb{P})$. By the martingale representation theorem (d-dimensional version) we know that there exists an $\mathbb{R}^d$-valued predictable process $\gamma(t,T,T^*), t \in [0,T]$ such that

$$F_P(t,T,T^*) = F_P(0,T,T^*)\varepsilon_t(\gamma) =$$
$$= F_P(0,T,T^*)\exp(\textstyle\int_0^t \gamma(u,T,T^*)dB_u - \frac{1}{2}\int_0^t |\gamma(u,T,T^*)|^2 du). \tag{2.8}$$

Put in another way using Itô we see that for a fixed maturity T

$$dF_P(t,T,T^*) = F_P(t,T,T^*)\gamma(t,T,T^*)dB_t. \tag{2.9}$$

If we consider any 2 maturities $T,U \in [0,T^*]$ then we define the forward process $F_P(t,T,U)$ by

$$F_P(t,T,U) \quad := \quad \frac{F_P(t,T,T^*)}{F_P(t,U,T^*)}, \quad \forall t \in [0,T \wedge U]. \tag{2.10}$$

Suppose $U > T$ - then the amount

$$f_s(t,T,U) = (U-T)^{-1}(F_P(t,T,U)-1) \tag{2.11}$$

is the add-on (annualized) forward rate over the future time interval $[T,U]$ prevailing at time t and

$$f(t,T,U) = \frac{\ln F_P(t,T,U)}{U-T}$$

is the (continuously compounded) forward rate at time t.

Remark 2.2.2. *For a better understanding of equation 2.11 and the following, think of a solution for r of*

$$P(t,T) = P(t,U)(1 + r(U-T))$$

and

$$P(t,T) = P(t,U)\exp(r(U-T))$$

respectively.

On the other hand if $U < T$ then $F_P(t,T,U)$ represents the value at time t of the forward price of a T-maturity bond for a forward contract that settles at time U.

Lemma 2.2.1. *Given 2 maturities $T, U \in [0,T^*]$ then the SDE describing the dynamics of the forward process under $\mathbb{P}$ is given by*

$$dF_P(t,T,U) = F_P(t,T,U)\gamma(t,T,U)\cdot(dB_t - \gamma(t,U,T^*)dt), \tag{2.12}$$

with

$$\gamma(t,T,U) = \gamma(t,T,T^*) - \gamma(t,U,T^*)$$

for every $t \in [0, T\wedge U]$.

Proof. The proof follows easily by applying Itô's formula. □

Using Girsanov's theorem we can write equation (2.12) like this:

$$dF_P(t,T,U) = F_P(t,T,U)\gamma(t,T,U)\cdot dB_t^U \tag{2.13}$$

with

$$B_t^U = B_t - \int_0^t \gamma(u,U,T^*)du$$

for every $t \in [0,U]$.

This process B_t^U is a standard BM on the probability space $(\Omega, (\mathcal{F}_t)_{t\in[0,U]}, \mathbb{P}_U)$ with $\mathbb{P}_U \sim \mathbb{P}$ and the Radon-Nikodym derivative

$$\frac{d\mathbb{P}_U}{d\mathbb{P}} = \varepsilon_U(\gamma(u,U,T^*)) \quad \mathbb{P} \text{ a.s..}$$

This measure $\mathbb{P}_U$ is called a forward measure. Solving equation (2.13) we find that

$$F_P(t,T,U) = F_P(0,T,U)\varepsilon_t(\gamma(u,U,T^*))$$

For $t \in [0, T\wedge U]$. Where the stochastic exponential is meant w.r.t. the BM B_t^U.

Remark 2.2.3. *Note that* $\mathbb{P}_{T^*} = \mathbb{P}$ *and* $B_t^{T^*} = B_t$ *by the definition of* $F_P(t, T, T^*)$ *and* $(BP.2)$.

2.2.2 Definition of Spot and Forward Martingale measure

We want to have clear definitions for the measures that appeared in the theorems before. So we sum up:

Definition 2.2.1 (Forward Martingale Measure)**.** *Let* U *be a fixed maturity date.* $\mathbb{Q}_U \sim \mathbb{P}$ *a probability measure on* $(\Omega, \mathcal{F}_U)$ *is called a forward martingale measure for the date* U *if for any maturity* $T \in [0, T^*]$, *the forward process* $F_P(t, T, U), t \in [0, T \wedge U]$, *is a local martingale under* $\mathbb{Q}_U$.

In the setting above obviously $\mathbb{P}$ is a forward martingale measure for the date T^*.

Definition 2.2.2 (Spot Martingale Measure)**.** *A spot martingale measure for the set up* $BP.1 - BP.2$ *is any probability measure* $\mathbb{P}^* \sim \mathbb{P}$ *on* $(\Omega, \mathcal{F}_{T^*})$ *for which there exists a process* $P^* \in \mathcal{A}^+$, *with* $P_0^* = 1$, *and such that for any maturity* $T \in [0, T^*]$ *the bond price* $P(t, T)$ *satisfies*

$$P(t, T) = \mathbf{E}_{\mathbb{P}^*}[P_t^*/P_T^*|\mathcal{F}_t], \quad \forall t \in [0, T].$$

2.2.3 Arbitrage-free Properties and the Implied Savings Account

Arbitrage-freeness

We will get to know two kinds of arbitrage-freeness. One without the presence of cash (a pure bond market) and another one that includes a cash account.

Definition 2.2.3 (Weak No-Arbitrage Condition)**.** *A family of bond prices* $P(t, T)$ *is said to satisfy the weak no-arbitrage condition if and only if there exists an equivalent probability measure* $\mathbb{Q} \sim \mathbb{P}$ *on* $(\Omega, \mathcal{F}_{T^*})$ *s.t. for any maturity* $T \leq T^*$ *the forward process* $F_P(t, T, T^*) = P(t, T)/P(t, T^*)$ *belongs to* $\mathcal{M}_{loc}(\mathbb{Q})$. *We say that the family satisfies the no-arbitrage condition if in addition* $P(T, U) \leq 1$ *holds for any maturities* $T \leq U \in [0, T^*]$.

The inequality $P(T, U) \leq 1$ is equivalent to $F_P(T, U, T) \leq 1$ and leads to

$$F_P(t, U, T) = \mathbf{E}_{\mathbb{P}}[F_P(T, U, T)|\mathcal{F}_t] \leq 1$$

for every $t \in [0, T]$. Since almost all sample paths of the forward process are continuous we can reformulate this condition in the following way:

(BP.3) For any two maturities $T \leq U$ with probability 1 $P(t, U) \leq P(t, T), \quad \forall t \in [0, T]$ holds.

The contrary inequality would lead to arbitrage if one assumes the presence of an increasing savings account. I.e. if $P(t,U) > P(t,T)$ then I could issue a U-maturity bond and buy a T-maturity bond at time t. I would have $P(t,U) - P(t,T)$ units of currency and to meet the liability at time U it is enough to carry one unit of currency, received at time T over the period $[T,U]$. I try to make this clear in the following table:

points in time	t	T	U
	$+P(t,U)$ $-P(t,T)$	$-P(T,U)$ $+1$	-1 1
sum	$P(t,U) - P(t,T) > 0$	$-P(T,U) + 1$	$-1 + 1 = 0$

One sees that in case of presence of a risk-free savings account I could have earned interest from my surplus at time t and could easily have put away the 1 unit at time U and I would have had a profit equal to this interest payment.

So we see that no-arbitrage in a bond market with presence of cash is strongly related to the existence of a savings account implied by the family $P(t,T)$.

Definition 2.2.4 (Implied savings Account)**.** *A savings account implied by the family $P(t,T)$ of bond prices is an arbitrary process P^* which belongs to $\mathcal{A}^+$, with $P^*_0 = 1$, and s.t. there exists a probability measure $\mathbb{P}^* \sim \mathbb{P}$ on $(\Omega, \mathcal{F}_{T^*})$ under which the relative bond price*

$$P^*(t,T) = P(t,T)/P^*_t, \quad \forall t \in [0,T]$$

is a martingale for any maturity $T \in [0,T^*]$.

Compare this definition with the definition of a spot martingale measure.

The Implied Savings Account

In this section we will investigate about the existence of an implied savings account under the assumptions $(BP.1) - (BP.3)$.

We first need some preliminary results for the terminal discounting factor $D_t = P^{-1}(t,T^*), t \in [0,T^*]$. Note that D belongs to $\mathcal{S}_p^+$.

Lemma 2.2.2. *Under the assumptions $(BP.1) - (BP.3)$, the terminal discounting factor D is a strictly positive supermartingale under the forward martingale measure $\mathbb{P}$.*

Proof. With $(BP.2)$ and $(BP.3)$ we obtain

$$P(t,U) = \mathbf{E}_{\mathbb{P}}[\frac{P(t,T^*)}{P(U,T^*)}|\mathcal{F}_t] \leq \mathbf{E}_{\mathbb{P}}[\frac{P(t,T^*)}{P(T,T^*)}|\mathcal{F}_t] = P(t,T),$$

so that

$$\mathbf{E}_{\mathbb{P}}[D_U|\mathcal{F}_t] \leq \mathbf{E}_{\mathbb{P}}[D_T|\mathcal{F}_t] \quad \forall t \leq T \leq U \leq T^*.$$

Setting $t = T$ in the last inequality, we find that

$$\mathbf{E}_{\mathbb{P}}[D_U|\mathcal{F}_T] \leq \mathbf{E}_{\mathbb{P}}[D_T|\mathcal{F}_T] = D_T$$

for every $T \leq U \leq T^*$ so that D is a $\mathbb{P}$- supermartingale. $\square$

To show the existence of an implied savings account we will use the following standard result from Itô stochastic calculus.

Proposition 2.2.1. *Suppose X belongs to the class $\mathcal{S}_p^+$, with $X_0 = 1$. There exists a unique pair (M, A) of stochastic processes such that $X = MA$, the process M belongs to $\mathcal{M}(\mathbb{P})^+_{loc}$, with $M_0 = 1$, and A belongs to $\mathcal{A}^+$, with $A_0 = 1$. If in addition X is a supermartingale, then A is a decreasing process.*

Corollary 2.2.1. *Under $(BP.1)-(BP.2)$, there exists a predictable process ξ integrable w.r.t. the BM B, and such that the terminal discount factor D admits the unique decomposition*

$$D_t = D_0 \tilde{A}_t \tilde{M}_t = D_0 \tilde{A}_t \varepsilon_t(\xi), \quad \forall t \in [0, T^*],$$

where $\tilde{M}_t \in \mathcal{M}(\mathbb{P})^+_{c,loc}$ and $\tilde{A}_t \in \mathcal{A}^+$, with the initial value 1 for both processes. If in addition condition $(BP.3)$ is met, then $\tilde{A}$ is a decreasing process.

This corollary is a consequence of the last lemma and the last proposition and the representation theorem of a strictly positive martingale w.r.t. the Brownian filtration.

Now we come to the main proposition of this section:

Proposition 2.2.2. *Let the family of bond prices $P(t,T)$ satisfy $(BP.1) - (BP.3)$. Assume that the process $\tilde{M}$ defined by the multiplicative decomposition (corollary 2.2.1) of the terminal discount factor D, is a martingale (not only a local martingale) under $\mathbb{P}$. Let $P^* = 1/\tilde{A}$ be an increasing predictable process uniquely determined by corollary 2.2.1. Then the following holds:*

- *P^* represents a savings account implied by the family $P(t,T)$.*
- *P^* is associated with the spot martingale measure $\mathbb{P}^*$, given by*

$$\frac{d\mathbb{P}^*}{d\mathbb{P}} := \tilde{M}_{T^*} = P^*_{T^*} P(0, T^*), \quad \mathbb{P} - a.s. \tag{2.14}$$

- *The relative price process $P(t,T^*)/P^*_t$ follows a martingale under the forward martingale measure $\mathbb{P}$ for the date T^*.*

Lemma 2.2.3. *Let P^* and $\hat{P}$ be two processes of $\mathcal{A}^+$ such that for every $T \in [0, T^*]$*

$$\mathbf{E}_{\mathbb{P}^*}[P^*_t/P^*_T|\mathcal{F}_t] = \mathbf{E}_{\hat{\mathbb{P}}}[\hat{P}_t/\hat{P}_T|\mathcal{F}_t]$$

for every $t \in [0, T]$, where $\mathbb{P}^* \sim \hat{\mathbb{P}}$. If $P^*_0 = \hat{P}_0$ then $P^* = \hat{P}$.

Corollary 2.2.2. *Under $(BP.1)-(BP.2)$, the uniqueness of an implied savings account holds.*

We sum it all up by the last corollary of this section:

Corollary 2.2.3. *Under $(BP.1) - (BP.2)$, the following are equivalent:*

- *The bond price $P(t,T)$ is a non-increasing function of the maturity T.*
- *The forward process $F_P(t,T,U), t \leq T \leq U$ is never strictly less than 1.*
- *The bond price $P(t,T)$ is never strictly greater than 1.*
- *The implied savings account follows an increasing process.*

Summary 2.1. *The last corollary sums it up perfectly. Having the two fundamental assumptions on the bond prices we get a lot more by asking the price not to be strictly greater than one. The main observation is that this assumption that might appear to be trivial nevertheless implies a riskless savings account.*

2.2.4 Bond Price Volatility

We assume $(BP.1) - (BP.3)$ to hold. Then we define a bond price volatility.

Definition 2.2.5 (Bond Price Volatility)**.** *An R^d-valued process $b(t,T)$ is called a bond price volatility for maturity T if the bond price admits the representation*

$$dP(t,T) = P(t,T)b(t,T) \cdot dB_t + dC_t^T \tag{2.15}$$

where C_t^T is a predictable process of finite variation.

Under $(BP.1) - (BP.2)$ the existence and uniqueness of the bond price volatility follows from the canonical decomposition of the special semi-martingale. It is also invariant under the change to an equivalent probability measure. More precisely we have that $b(t,T)$ stays the same while the BM changes to a BM in the new measure and the process C_t^T is changed too. But we assumed $(BP.3)$ to hold as well and so we have that there exists a process P_t^* and a spot martingale measure $\mathbb{P}^*$ such that the relative bond prices $(Z^*(t,T))$ follow a local martingale. Thus the relative bond prices can be expressed as stochastic integrals with respect to the BM B_t^*

$$Z^*(t,T) = P(0,T)\varepsilon_t(b(t,T)).$$

By setting $t = T$ we can easily find an expression for P_t^* in terms of the bond price volatility. Finally we note that for any maturities $T, U \in [0, T^*]$ the forward volatility (volatility of the forward process) is given by

$$\gamma(t,T,U) = b(t,T) - b(t,U), \forall t \in [0, T \wedge U].$$

2.2.5 Forward LIBOR Rates

The general assumptions made in the section before ((BP.1) and (BP.2)) still hold in this notion of forward rates. We have a family of bond prices, denoted by $P(t,T)$ and thereby a collection $F_P(t,T,U)$ of forward processes. In this setting we define the notion of a LIBOR rate.

Definition 2.2.6 (LIBOR Rate)**.** *For a strictly positive real number δ the so called δ-LIBOR rate for the date $T \leq T^* - \delta$ prevailing at time t is denoted by $L(t,T)$ and defined by*

$$1 + \delta L(t,T) = F_P(t,T,T+\delta), \quad \forall t \in [0,T]. \tag{2.16}$$

Remark 2.2.4. *The abbreviation LIBOR stands for London Interbank Offered Rate. These rates are most commonly used in international financial markets.*

Remark 2.2.5. *Typical choices for δ are 0.25 or 0.5 which leads to so called 3-month LIBOR rates and 6-month LIBOR rates respectively.*

If we compare this definition with equation (2.11) in the last section then we see that

$$L(t,T) = f_s(t,T,T+\delta)$$

meaning that the LIBOR rate $L(t,T)$ represents the add-on forward rate prevailing at time t for the interval $[T,T+\delta]$. For any maturity $T \in [T^* - \delta, T^*]$ we can express the LIBOR rate in terms of bond prices:

$$1 + \delta L(t,T) = \frac{P(t,T)}{P(t,T+\delta)} \quad \forall t \in [0,T].$$

The initial structure if a LIBOR rate is given by

$$L(0,T) = f_s(0,T,T+\delta) = \delta^{-1}(\frac{P(0,T)}{P(0,T+\delta)} - 1).$$

In the same manner as in the section before we get the following SDE under the forward probability measure $\mathbb{P}_{T+\delta}$:

$$dL(t,T) = \delta^{-1} F_P(t,T,T+\delta)\gamma(t,T,T+\delta) \cdot dB_t^{T+\delta},$$

with $B_t^{T+\delta}$ defined analogous as before.

Plugging in the definition of $L(t,T)$ leads to the SDE

$$dL(t,T) = \delta^{-1}(1 + \delta L(t,T))\gamma(t,T,T+\delta) \cdot dB_t^{T+\delta} \tag{2.17}$$

for the process L. Suppose that LIBOR rates are strictly positive then we can rewrite equation (2.17) as

$$dL(t,T) = L(t,T)\lambda(t,T) \cdot dB_t^{T+\delta} \tag{2.18}$$

where for any $t \in [0, T]$

$$\lambda(t, T) = \frac{1 + \delta L(t, T)}{\delta L(t, T)} \gamma(t, T, T + \delta). \tag{2.19}$$

This shows that a model for LIBOR rates depends on the collection of forward rates. So we can find two characterizing properties:

(LR.1) For any maturity $T \leq T^* - \delta$, we are given an $\mathbb{R}^d$-valued, bounded, deterministic function $\lambda(t, T)$ which represents the volatility of the forward LIBOR rate process $L(t, T)$ for $t \in [0, T]$.

(LR.2) We assume a strictly decreasing and strictly positive initial term structure $\tilde{P}(0, T), T \in [0, T^*]$, and by that an initial term structure $L(0, T)$ of forward LIBOR rates

$$L(0, T) = \delta^{-1}\left(\frac{\tilde{P}(0, T)}{\tilde{P}(0, T + \delta)} - 1\right), \quad \forall T \in [0, T^* - \delta].$$

Remark 2.2.6. *For the function λ a stochastic function could be chosen as well, but in this introduction we focus on a model where λ is deterministic. This leads to the so called lognormal model. Later on we have a look at what we can do if the volatility is assumed to follow a stochastic process.*

Advantages of Using LIBOR rates

Using LIBOR rates has several advantages and therefore they have become one of the most important notions in the financial industry.

1. Market models for LIBOR rates are based on observable market rates whereas continuously compounded instantaneous rates are not observable.

2. Most of the derivative contracts are based on a certain LIBOR rate (floors/caps, swaps).

3. LIBOR market models are consistent with the market convention of quoting caps using Black's formula.

2.2.6 The Discrete Tenor Model

We assume that the horizon date T^* is a multiple of δ, say $T^* = M\delta$ with $M \in \mathbb{N}$. We will consider a finite number of dates, $T^*_{m\delta} = T^* - m\delta$ for $m = 1, \ldots, M - 1$. This is what is meant by discrete tenor. The LIBOR rates still follow a continuous-time process. We start by defining the LIBOR rate for the longest maturity, $L(t, T^*_\delta)$. We postulate that $L(t, T^*_\delta)$ follows the following SDE under $\mathbb{P}$:

$$dL(t, T^*_\delta) = L(t, T^*_\delta)\lambda(t, T^*_\delta) \cdot dB_t, \tag{2.20}$$

with

$$L(0,T_\delta^*) = \delta^{-1}(\frac{\tilde{P}(0,T_\delta^*)}{\tilde{P}(0,T^*)} - 1).$$

By solving equation (2.20), we postulate that for every $t \in [0, T_\delta^*]$

$$L(t,T_\delta^*) = \delta^{-1}(\frac{\tilde{P}(0,T_\delta^*)}{\tilde{P}(0,T^*)} - 1)\varepsilon_t(\lambda(\cdot,T_\delta^*)). \tag{2.21}$$

Remark 2.2.7. *Here one sees that $L(t,T_\delta^*)$ is lognormally distributed by thinking of the definition of a lognormally distributed random variable and the definition of the stochastic exponential $\varepsilon_t()$.*

Since $\tilde{P}(0,T_\delta^*) > \tilde{P}(0,T^*)$ it is clear that $L(t,T_\delta^*) \in M_c^+(\mathbb{P})$. Also by assumption for fixed $t \leq T^* - \delta$, $L(t,T_\delta^*)$ has lognormal probability law under $\mathbb{P}$.

The next step is to define the forward LIBOR rate for the date $T_{2\delta}^*$ using the relation (2.19) with $T = T_\delta^*$, i.e.

$$\gamma(t,T_\delta^*,T^*) = \frac{\delta L(t,T_\delta^*)}{1 + L(t,T_\delta^*)}\lambda(t,T_\delta^*), \quad \forall t \in [0, T^* - \delta]. \tag{2.22}$$

We know from the general properties of the forward process (equation 2.9) that the forward process $F_P(t,T_\delta^*,T^*)$ solves the following SDE under $\mathbb{P}$:

$$dF_P(t,T_\delta^*,T^*) = F_P(t,T_\delta^*,T^*)\gamma(t,T_\delta^*,T^*) \cdot dB_t \tag{2.23}$$

with the initial condition $F_P(0,T_\delta^*,T^*) = \tilde{P}(0,T_\delta^*)/\tilde{P}(0,T^*)$. The forward process belongs to $\mathcal{M}(\mathbb{P})_c^+$ since the volatility γ is a bounded process.

We introduce a d-dimensional process $B^{T_\delta^*}$ by

$$B^{T_\delta^*} = B_t - \int_0^t \gamma(u,T_\delta^*,T^*)du, \quad \forall t \in [0, T_\delta^*].$$

By assumption on the volatility function $\gamma(t,T_\delta^*,T^*)$ is bounded and by Girsanov's theorem we get the existence of the process $B^{T_\delta^*}$ and the associated probability measure $\mathbb{P}_{T_\delta^*}$ under which $B^{T_\delta^*}$ is a BM. We have that the change from $\mathbb{P}$ to the equivalent measure $\mathbb{P}_{T_\delta^*}$ is $\mathbb{P}$-a.s. given by

$$\frac{d\mathbb{P}_{T_\delta^*}}{d\mathbb{P}} = \varepsilon_{T_\delta^*}(\gamma(\cdot,T_\delta^*,T^*))$$

Now we are able to specify the dynamics of the forward LIBOR rate for the date $T_{2\delta}^*$ under the forward probability measure $\mathbb{P}_{T_\delta^*}$. Analogously to (2.20), we set

$$dL(t,T_{2\delta}^*) = L(t,T_{2\delta}^*)\lambda(t,T_{2\delta}^*) \cdot dB_t^{T_\delta^*}$$

with the initial condition

$$L(0,T_{2\delta}^*) = \delta^{-1}(\frac{\tilde{P}(0,T_{2\delta}^*)}{\tilde{P}(0,T_{\delta}^*)} - 1).$$

Solving this equation and comparing with (2.19) for $T = T_{2\delta}^*$, we get

$$\gamma(t,T_{2\delta}^*,T_{\delta}^*) = \frac{\delta L(t,T_{2\delta}^*)}{1+\delta L(t,T_{2\delta}^*)}\lambda(t,T_{2\delta}^*), \quad \forall t \in [0,T_{2\delta}^*].$$

To find $\gamma(t,T_{2\delta}^*,T^*)$, we compare with (2.19) and using the relation

$$\gamma(t,T_{2\delta}^*,T_{\delta}^*) = \gamma(t,T_{2\delta}^*,T^*) - \gamma(t,T_{\delta}^*,T^*), \quad \forall t \in [0,T_{2\delta}^*]$$

we get the process.

Given the process $\gamma(t,T_{2\delta}^*,T_{\delta}^*)$, we can define $B^{T_{2\delta}^*}$ and $\mathbb{P}_{T_{2\delta}^*}$ corresponding to the date $T_{2\delta}^*$ and so forth. Working backwards to the first date $T_{(M-1)\delta}^* = \delta$ we construct a family of forward LIBOR rates $L(t,T_{m\delta}^*), \quad m = 1,\dots,M-1$. Notice that the lognormal probability law of every process $L(t,T_{m\delta}^*)$ under the corresponding forward measure $\mathbb{P}_{T_{(m-1)\delta}^*}$ is ensured. We have for any $m = 1,\dots,M-1$

$$dL(t,T_{m\delta}^*) = L(t,T_{m\delta}^*)\lambda(t,T_{m\delta}^*)\cdot dB_t^{T_{(m-1)\delta}^*}$$

where $B^{T_{(m-1)\delta}^*}$ is a standard BM under $\mathbb{P}_{T_{(m-1)\delta}^*}$. This completes the lognormal model of forward LIBOR rates.

The Implied Savings Account in the Discrete Model

Simultaneously we constructed a family of forward LIBOR rates and the associated forward processes. Therefore it is interesting to examine the existence and uniqueness of an implied savings account. In a discrete time setting it is a process $P_t^*, t = 0,\delta,\dots,T^* = M\delta$. From the definition of the forward rate

$$F_P(t,T_j,T_{j+1}) = \frac{F_P(t,T_j,T^*)}{F_P(t,T_{j+1},T^*)} = \frac{P(t,T_j)}{P(t,T_{j+1})}$$

with $T_j = j\delta$. This yields by setting $t = T_j$

$$F_P(T_j,T_j,T_{j+1}) = \frac{1}{P(T_j,T_{j+1})}.$$

Thus the price of a bond $P(T_j,T_{j+1})$ is uniquely defined in the model. Although a bond that matures at time T_j does not physically exist, it seems justifiable to see $F_P(T_j,T_j,T_{j+1})$ as its forward value at time T_j for the date T_{j+1}. In other words the spot value of one unit of cash at time T_{j+1} received at time T_j equals $P^{-1}(T_j, T_{j+1})$. The discrete savings account thus equals

$$P^*_{T_k} = \prod_{j=1}^{k} F_P(T_{j-1}, T_{j-1}, T_j) = \prod_{j=1}^{k} \frac{1}{P(T_{j-1}, T_j)}$$

for $k = 0, \ldots, M-1$. (remember $P^*_0 = 1$).

It is easily seen that by $F_P(T_j, T_j, T_{j+1}) = 1 + \delta L(T_j, T_{j+1}) > 1$ for $j = 1, \ldots, M-1$ and by $P^*_{T_{j+1}} = F_P(T_j, T_j, T_{j+1}) P^*_{T_j}$ the discrete savings account process is a strictly increasing process. Moreover we define a measure $\mathbb{P}^* \sim \mathbb{P}$ on $(\Omega, \mathcal{F}_{T^*})$ by the formula from the section on the implied savings account by

$$\frac{d\mathbb{P}^*}{d\mathbb{P}} = P^*_{T^*} P(0, T^*), \mathbb{P} - a.s. \tag{2.24}$$

We see this, if we consider the condition

$$P(T_l, T_k) = \mathbf{E}_{\mathbb{P}^*}[P^*_{T_l}/P^*_{T_k} | \mathcal{F}_{T_l}]$$

for every $l \leq k \leq M$, then in the case $l = k-1$ this condition coincides with the condition on the implied savings account.

2.2.7 The Continuous Tenor Case

In a continuous model all forward LIBOR rates $L(t, T)$ with $T \in [0, T^*]$ are specified. This is done using the procedure described before and filling the gaps between the discrete maturities. The construction of a model in which each forward LIBOR rate $L(t, T)$ follows as lognormal process under the forward measure for the date $T + \delta$ is done by a backward induction.

First Step

We construct a discrete tenor model using the method described in the section before.

Second Step

We first fill the gap for maturities $T \in (T^*_\delta, T^*)$. We do not have to take the forward LIBOR rate $L(t, T)$ into account, because we don't have any for T in this interval. But we are given values of the implied savings account for the dates T^*_δ and T^* from the previous construction. Note that those values $P^*_{T^*_\delta}$ and $P^*_{T^*}$ are $\mathcal{F}_{T^*_\delta}$ measurable. We define a spot martingale measure $\mathbb{P}^*$ using formula 2.24. We search for an increasing function $\alpha : [T^*_\delta, T^*]$ such that $\alpha(T^*_\delta) = 0$ and $\alpha(T^*) = 1$ since the model has to match a given initial term structure. The process

$$\ln P^*_t = (1 - \alpha(t)) \ln P^*_{T^*_\delta} + \alpha(t) \ln P^*_{T^*}, \quad \forall t \in [T^*_\delta, T^*],$$

satisfies $\tilde{P}(0,t) = \mathbf{E}_{\mathbb{P}^*}[1/P_t^*]$ for every $t \in [T_\delta^*, T^*]$. From $0 < P^*_{T^*_\delta} < P^*_{T^*}$ and $\tilde{P}(0,t), t \in [T_\delta^*, T^*]$, is assumed to be strictly decreasing, we see that such a function α exists uniquely.

Third Step

In the second step we have constructed the implied savings account for every $T \in [T_\delta^*, T^*]$. Hence a forward martingale measure for any date $T \in (T_\delta^*, T^*)$ can be defined by

$$\frac{d\mathbb{P}_T}{d\mathbb{P}^*} = \frac{1}{P_T^*\tilde{P}(0,T)}, \mathbb{P}^* - a.s. \tag{2.25}$$

Combining this with formula 2.24 we get

$$\frac{d\mathbb{P}_T}{d\mathbb{P}} = \frac{d\mathbb{P}_T}{d\mathbb{P}^*}\frac{d\mathbb{P}^*}{d\mathbb{P}} = \frac{P^*_{T^*}\tilde{P}(0,T^*)}{P^*_T\tilde{P}(0,T)}, \mathbb{P} - a.s.$$

for every $T \in [T_\delta^*, T^*]$.
Moreover we have

$$\frac{d\mathbb{P}_T}{d\mathbb{P}}_{|\mathcal{F}_t} = \mathbf{E}_{\mathbb{P}}[\frac{P^*_{T^*}\tilde{P}(0,T^*)}{P^*_T\tilde{P}(0,T)}|\mathcal{F}_t], \quad \forall t \in [0,T].$$

By exponential representation of the martingale (the change of measure process) we get

$$\frac{d\mathbb{P}_T}{d\mathbb{P}}_{|\mathcal{F}_t} = \frac{\tilde{P}(0,T^*)}{\tilde{P}(0,T)}\varepsilon_t(\gamma(\cdot,T,T^*)), \quad \forall t \in [0,T].$$

This process γ defines the forward volatility for any $T \in (T_\delta^*, T^*)$. We are even able to define the $\mathbb{P}_T$ -BM B^T. Given those objects we define the forward LIBOR rate process $L(t,T_\delta)$ by

$$dL(t,T_\delta) = L(t,T_\delta)\lambda(t,T_\delta)\cdot dB_t^T$$

with $T_\delta = T - \delta$ and initial condition

$$L(0,T_\delta) = \delta^{-1}(\frac{\tilde{P}(0,T_\delta)}{\tilde{P}(0,T)} - 1).$$

Finally we can set

$$\gamma(t,T_\delta^*,T^*) = \frac{\delta L(t,T_\delta^*)}{1+\delta L(t,T_\delta^*)}\lambda(t,T_\delta^*), \quad \forall t \in [0,T_\delta^*].$$

To define the forward measure $\mathbb{P}_U$ and the corresponding BM B^U for any maturity $U \in (T_{2\delta}^*, T_\delta^*)$ we put

$$\gamma(t,U,T) = \gamma(t,T_\delta,T) = \frac{\delta L(t,T_\delta)}{1+\delta L(t,T_\delta)}\lambda(t,T_\delta), \quad \forall t \in [0,T_\delta],$$

where $U = T_\delta$ so that $T = U + \delta$ belongs to (T^*_δ, T^*). Note that we get $\gamma(t, U, T^*)$ in the same manner as before by the relation

$$\gamma(t, U, T^*) = \gamma(t, U, T) - \gamma(t, T, T^*), \quad \forall t \in [0, U].$$

By proceeding this backward induction we specify a fully continuous family of forward LIBOR rates. Simultaneously we define the forward volatilities $\gamma(t, T, T^*)$ and by that the forward processes $F(t, T, T^*)$ which fulfill the SDE

$$dF(t, T, T^*) = F(t, T, T^*)\gamma(t, T, T^*) \cdot B_t.$$

By formally setting $P(t, T) = F(t, T, t)$ the collection of forward process admits an associated family of bond prices. The bond prices obtained in this way fulfill the weak no-arbitrage condition but the no-arbitrage with cash ($F_P(T, T, U) \geq 1$ for $U \geq T$) may fail to hold. See papers of Musiela and Rutkowski for detailed information.

Summary 2.2. *In this section we got to know LIBOR rates as interest rates for periods in a discrete setting as well as in a continuous setting. They are useful in some applications and assuming deterministic volatility their models lead to lognormal distributions.*

Chapter 3

Interest Rate Derivatives

3.1 Forward Swap Rates

3.1.1 A Swap Contract

A swap contract (or swap) is an agreement between two parties to exchange cash flows at future points in time according to a prearranged formula. The two most popular kinds of swaps are standard **interest swaps** and **cross-currency swaps** (differential swaps). In a so called plain vanilla interest swap party A agrees to pay to party B amounts of money determined by a fixed interest rate on some principal at each of the payment dates whereas party B pays interest at a floating reference rate on the same principal for the same period of time. Such an interest swap can be used to transform a fixed interest rate loan into a floating-rate loan or vice versa. Swaps are also distinguished in who pays the fixed and who pays the floating rate:

Payer Swap: The fixed rate is payed at the end (or beginning) of each period and the floating rate is received. Also called fixed-for-floating swap.

Receiver Swap: The investor pays a floating rate and receives a fixed one.

It is also distinguished whether those agreements are settled in arrears (at the end of each period) or in advance (in the beginning).

Given a set of points in time $T_0, T_1, \dots T_n$ a **forward start swap** (or simply forward swap) is a swap contract entered at some date $t < T_0$ with payment dates corresponding to the set of dates (corresponding to settlement in arrear or in advance - n dates always.). The **forward swap rate** is the value of the fixed rate which makes the value of the forward swap zero. The market makes quotes of these rates for several maturities. The most typical option contract associated with swaps is a swaption - an option on the value of the underlying swap or equivalently on the (forward) swap rate.

3.1.2 Interest Rate Swaps

We consider a forward start payer swap settled in arrear with notional principal N. Moreover we consider a finite collection of dates $T_j, j = 1, \ldots, n$. The distance of those dates has the constant value δ. The floating rate $L(T_j)$ is received at time T_{j+1} and set at time T_j by looking at the price of a zero-coupon bond over that period. Thus $L(T_j)$ satisfies

$$P(T_j, T_{j+1})^{-1} = 1 + (T_{j+1} - T_j)L(T_j) = 1 + \delta L(T_j). \tag{3.1}$$

We see that this formulation agrees with the market quotations of LIBOR and $L(T_j)$ can be seen as the spot LIBOR rate prevailing at time T_j for the period δ. We remember the more general definition of a LIBOR rate

$$1 + \delta L(t, T_j) = \frac{P(t, T_j)}{P(t, T_{j+1})}$$

and see that $L(T_j)$ coincides with $L(T_j, T_j)$. The cash flows of the payer are at any of the dates $T_j, j = 1, \ldots, n$ and the amounts are $L(T_{j-1})\delta N$ and $-\kappa\delta N$, where κ stands for the preassigned fixed rate of interest. We make some notation clear:

n **:** The number of payments is referred to as the length of the swap.

$T_0, \ldots, T_{n-1}$ **:** Those points in time are called the reset dates.

$T_1, \ldots, T_n$ **:** Those points in time are called the settlement dates.

T_0 **:** The first reset date is referred to as the start date of a swap.

$[T_{j-1}, T_j]$ **:** Such an interval is referred to as the j^{th} accrual period.

N **:** The notional principal is assumed to be 1 without loss of generality.

The value of a forward start payer swap at time t - denoted by $\mathbf{FS}_t$ or $\mathbf{FS}_t(\kappa)$ - is given by

$$\begin{aligned} \mathbf{FS}_t(\kappa) &= \mathbf{E}_{\mathbb{P}^*}[\sum_{j=1}^{n} \frac{P_t}{P_{T_j}}(L(T_{j-1}) - \kappa)\delta|\mathcal{F}_t] \\ &= \sum_{j=1}^{n} \mathbf{E}_{\mathbb{P}^*}[\frac{P_t}{P_{T_j}}(P(T_{j-1}, T_j)^{-1} - \tilde{\delta})|\mathcal{F}_t], \end{aligned} \tag{3.2}$$

with $\tilde{\delta} = 1 + \kappa\delta$. Using the adaptedness of the bond price we get:

$$\begin{aligned} \mathbf{FS}_t(\kappa) &= \sum_{j=1}^{n} \mathbf{E}_{\mathbb{P}^*}[P(T_{j-1}, T_j)^{-1} \frac{P_t}{P_{T_{j-1}}} \mathbf{E}_{\mathbb{P}^*}[\frac{P_{T_{j-1}}}{P_{T_j}}|\mathcal{F}_{T_{j-1}}]|\mathcal{F}_t] \\ &- \sum_{j=1}^{n} \tilde{\delta}\mathbf{E}_{\mathbb{P}^*}[\frac{P_t}{P_{T_j}}|\mathcal{F}_t] = \sum_{j=1}^{n} P(t, T_{j-1}) - \tilde{\delta}P(t, T_j). \end{aligned} \tag{3.3}$$

Rearranging leads to

$$\mathbf{FS}_t(\kappa) = P(t,T_0) - \sum_{j=1}^{n} c_j P(t,T_j) \tag{3.4}$$

for every $t \in [0,T]$, with $c_j = \kappa\delta$ for $j = 1,\ldots,n-1$ and $c_n = \tilde{\delta}$. Like this one sees that a swap contract settles in arrears can be seen as receiving a z.c.b. and to deliver a specified coupon bearing bond.

The situation is a bit more complicated in a forward start payer swap that is settled in advance. This means that the reset dates and the settlement dates coincide. There are various convention in how to discount the payments. In the U.S. and in many European markets the cash flows at the settlement dates $T_j, j = 0,\ldots,n-1$ are $L(T_j)\delta(1+L(T_j)\delta)^{-1}$ and $-\kappa\delta(1+L(T_j)\delta)^{-1}$. The value of this contract is denoted by $\mathbf{FS}_t^{**}(\kappa)$ and at time t it equals:

$$\begin{aligned}
\mathbf{FS}_t^{**}(\kappa) &= \mathbf{E}_{\mathbb{P}^*}[\sum_{j=0}^{n-1} \frac{P_t}{P_{T_j}} \frac{(L(T_j)-\kappa)\delta}{1+L(T_j)\delta} | \mathcal{F}_t] \\
&= \mathbf{E}_{\mathbb{P}^*}[\sum_{j=0}^{n-1} \frac{P_t}{P_{T_j}} (L(T_j)-\kappa)\delta P(T_j,T_{j+1}) | \mathcal{F}_t] \\
&= \mathbf{E}_{\mathbb{P}^*}[\sum_{j=1}^{n} \frac{P_t}{P_{T_j}} (L(T_{j-1})-\kappa)\delta | \mathcal{F}_t] \qquad (3.5)
\end{aligned}$$

which equals the value of a swap settled in arrears.
Finally I just mention the convention in Australia where the cash flows are $L(T_j)\delta(1+L(T_j)\delta)^{-1}$ and $-\kappa\delta(1+\kappa\delta)^{-1}$. This leads to the value of a swap settled in arrears but discounted with the fixed rate κ.

Definition 3.1.1 (Forward Swap Rate)**.** *The forward swap rate $\kappa(t,T_0,n)$ at time t for the date T_0 is the value for the fixed rate κ which makes the value of the forward swap zero. This is the value of κ such that $\mathbf{FS}_t(\kappa) = 0$.*

Using equation (3.4) we easily get

$$\kappa(t,T_0,n) = (P(t,T_0) - P(t,T_n))(\delta \sum_{j=1}^{n} P(t,T_j))^{-1}. \tag{3.6}$$

Definition 3.1.2 (Swap Rate)**.** *A swap (swap rate) is the forward swap rate with $t = T_0$.*

It equals:

$$\kappa(T_0,T_0,n) = (1 - P(T_0,T_n))(\delta \sum_{j=1}^{n} P(T_0,T_j))^{-1}. \tag{3.7}$$

To finish this general section we have a look at a 1-period swap. With $n = 1$ and $T = T_j$ the formula for the forward swap rate gives:

$$\kappa(t, T_j, 1) = \frac{P(t, T_j) - P(t, T_{j+1})}{\delta P(t, T_{j+1})}, \quad \forall t \in [0, T_j].$$

We see the identity $\kappa(t, T_j, 1) = L(t, T_j)$ so it coincides with the LIBOR rate over the period $[T_j, T_{j+1}]$.

3.1.3 Model of Forward Swap Rates

The model described in this section was developed by Jamishidian. For reference look at [6]. We consider a forward start fixed-for-floating interest rate swap which starts at time T_j of a given collection of points in time $T_j = j\delta, j = 1, \dots, M$. This contract has $M - j$ accrual periods. From equation (3.6) we know that the forward swap rate $\kappa(t, T_j, M)$ - the value for the fixed rate κ such that the value of the contract is zero - is given by

$$\kappa(t, T_j, M) = (P(t, T_j) - P(t, T_M))(\delta \sum_{l=j+1}^{M} P(t, T_l))^{-1}$$

for every $t \in [0, T_j]$ and every $j = 1, \dots, M - 1$.

We will consider a family of swap rates $\tilde{\kappa}(t, T_j) = \kappa(t, T_j, M-j)$ with $j = 1, \dots, M-1$. By this definition the swaps differ in length but mature at the same time $T^* = T_M$. We define $T^*_{k\delta} := T^* - k\delta$ with $T^*_0 = T^*$. The forward swap rate for a date $T^*_{m\delta}$ is then given by:

$$\tilde{\kappa}(t, T^*_{m\delta}) = \frac{P(t, T^*_{m\delta}) - P(t, T^*)}{\delta(P(t, T^*_{(m-1)\delta}) + \dots + P(t, T^*))}, \quad \forall t \in [0, T^*_{m\delta}]. \tag{3.8}$$

We assume that the bond prices $P(t, T^*_{m\delta}), m = 1, \dots, M-1$ are given on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a BM B. We assume $\mathbb{P} = \mathbb{P}_{T^*}$ to be the forward measure for the date T^* and let B^{T^*} be its forward BM. We define for any $m = 1, \dots, M-1$ the coupon process

$$G_t(m) := \sum_{k=0}^{m-1} P(t, T^*_{k\delta}), \quad \forall t \in [0, T^*_{(m-1)\delta}].$$

The forward swap measure $\tilde{\mathbb{P}}_{T^*_{(m-1)\delta}}$ for the date $T^*_{(m-1)\delta}$ is the equivalent measure to $\mathbb{P}$, which corresponds to the choice of $G(m)$ as numeraire. In other words under this measure the relative bond prices

$$Z_m(t, T^*_{k\delta}) := \frac{P(t, T^*_{k\delta})}{G_t(m)} \quad t \in [T^*_{k\delta} \wedge T^*_{m\delta}]$$

with m fixed and $k = 0, \ldots, M-1$ follow a local martingale. Obviously $G_t(1) = P(t, T^*)$ holds and by $Z_1(t, T^*_{k\delta}) = F_P(t, T^*_{k\delta}, T^*)$ we see that the measure $\tilde{\mathbb{P}}_{T^*}$ can be chosen to be the forward martingale measure $\mathbb{P}_{T^*}$. Finally from the definition of $\tilde{\kappa}(t, T^*_{m\delta})$ and $G_t(m)$ we see that $\tilde{\kappa}(t, T^*_{m\delta})$ is a local martingale under $\tilde{\mathbb{P}}_{T^*_{(m-1)\delta}}$:

$$\tilde{\kappa}(t, T^*_{m\delta}) = \delta^{-1}(Z_m(t, T^*_{m\delta}) - Z_m(t, T^*)), \quad \forall t \in [0, T^*_{m\delta}].$$

The aim is to construct a direct model of forward swap rates of the form

$$d\tilde{\kappa}(t, T^*_{(m+1)\delta}) = \tilde{\kappa}(t, T^*_{(m+1)\delta})\nu(t, T^*_{(m+1)\delta}) \cdot d\tilde{B}_t^{T^*_{m\delta}}$$

for every $m = 0, \ldots, M-2$, where $\tilde{B}_t^{T^*_{m\delta}}$ is a BM under the forward swap measure $\tilde{\mathbb{P}}_{T^*_{m\delta}}$. The bounded deterministic functions $\nu(\cdot, T^*_{m\delta}) : [0, T^*_{m\delta}] \to \mathbb{R}, \quad m = 1, \ldots, M-1$ form the family of volatility functions of the swap rates. The initial condition is given in terms of the initial term structure $\tilde{P}(0, T^*_{m\delta}), \quad m = 1, \ldots, M-1$ and in order to be consistent the initial condition is

$$\tilde{\kappa}(0, T^*_{(m+1)\delta}) = \frac{\tilde{P}(0, T^*_{(m+1)\delta}) - \tilde{P}(0, T^*)}{\delta(\tilde{P}(0, T^*_{m\delta}) + \cdots + \tilde{P}(0, T^*))}. \tag{3.9}$$

The first step is to find an SDE for the forward swap rate $\tilde{\kappa}(t, T^*_\delta)$ by

$$d\tilde{\kappa}(t, T^*_\delta) = \tilde{\kappa}(t, T^*_\delta)\nu(t, T^*_\delta) \cdot d\tilde{B}_t^{T^*}$$

with the initial condition

$$\tilde{\kappa}(0, T^*_\delta) = \frac{\tilde{P}(0, T^*_\delta) - \tilde{P}(0, T^*)}{\delta \tilde{P}(0, T^*)}.$$

Remark 3.1.1. *Note that* $\tilde{B}_t^{T^*} = B_t^{T^*} = B_t$

The next step is to define an SDE for $\tilde{\kappa}(t, T^*_{2\delta})$. Therefore we have to define the forward measure and the BM for the time $T^*_{2\delta}$. Additionally we will need the following lemma.

Lemma 3.1.1. *Let G and H be real-valued adapted processes, such that $dG_t = G_t g_t \cdot dB_t$ and $dH_t = H_t h_t \cdot dB_t$. Assume that $H > -1$. Then the process $Y_t = G_t/(1 + H_t)$ satisfies*

$$dY_t = Y_t(g_t - \frac{H_t h_t}{1 + H_t}) \cdot (dB_t - \frac{H_t h_t}{1 + H_t} dt). \tag{3.10}$$

Proof. The proof follows easily by Itô's formula. □

To define the process $\tilde{\kappa}(t, T^*_{2\delta})$ we note that $Z_1(t, T^*_{k\delta}) = F_P(t, T^*_{k\delta}, T^*)$ for each possible k follows a strictly positive local martingale under the measure $\tilde{\mathbb{P}}_{T^*} = \mathbb{P}_{T^*}$. Like this we have

$$dZ_1(t, T^*_{k\delta}) = Z_1(t, T^*_{k\delta})\gamma_1(t, T^*_{k\delta}) \cdot B_t^{T^*}$$

for some process $\gamma_1(t, T^*_{k\delta})$. If we now define the process Z_2 by

$$Z_2(t, T^*_{k\delta}) = \frac{P(t, T^*_{k\delta})}{P(t, T^*_{k\delta}) + P(t, T^*)} = \frac{Z_1(t, T^*_{k\delta})}{1 + Z_1(t, T^*_{\delta})}.$$

We postulate that the process $Z_2(t, T^*_{k\delta})$ follows a local martingale under $\tilde{P}_{T^*_\delta}$. Applying lemma 3.1.1 and Girsanov's theorem we see that this holds true and that the associated BM looks like this:

$$\tilde{B}_t^{T^*_\delta} = \tilde{B}_t^{T^*} - \int_0^t \frac{Z_1(u, T^*_\delta)\gamma(u, T^*_\delta)}{1 + Z_1(u, T^*_\delta)} du \quad \forall t \in [0, T^*_\delta].$$

The measure $\tilde{\mathbb{P}}_{T^*_\delta}$ is also found by Girsanov's theorem. For an explicit expression for $\tilde{B}_t^{T^*_\delta}$ note that

$$Z_1(t, T^*_\delta) = \frac{P(t, T^*_\delta)}{P(t, T^*)} = \delta\tilde{\kappa}(t, T^*_\delta) + Z_1(t, T^*) = \delta\tilde{\kappa}(t, T^*_\delta) + 1.$$

From the SDE's for the respective expressions we get

$$Z_1(t, T^*_\delta)\gamma(t, T^*_\delta) = \delta\tilde{\kappa}(t, T^*_\delta)\nu(t, T^*_\delta).$$

So we get the following explicit expression:

$$\tilde{B}_t^{T^*_\delta} = \tilde{B}_t^{T^*} - \int_0^t \frac{\delta\tilde{\kappa}(u, T^*_\delta)}{\delta\tilde{\kappa}(u, T^*_\delta) + 2}\nu(u, T^*_\delta)du, \forall t \in [0, T^*_\delta].$$

We are now able to define the forward swap measure $\tilde{\mathbb{P}}_{T^*_\delta}$ and the process $\tilde{\kappa}(t, T^*_{2\delta})$ by

$$d\tilde{\kappa}(t, T^*_{2\delta}) = \tilde{\kappa}(t, T^*_{2\delta})\nu(t, T^*_{2\delta}) \cdot d\tilde{B}_t^{T^*_\delta}$$

with the initial condition

$$\tilde{\kappa}(0, T^*_{2\delta}) = \frac{\tilde{P}(0, T^*_{2\delta}) - \tilde{P}(0, T^*)}{\delta(\tilde{P}(0, T^*_\delta) + \tilde{P}(0, T^*))}.$$

Before we do the general step we consider the third step explicitly. We want to find a swap forward measure and an associated BM for the process $\tilde{\kappa}(t, T^*_{3\delta})$. We define

$$Z_3(t, T^*_{k\delta}) = \frac{P(t, T^*_{k\delta})}{P(t, T^*_{2\delta}) + P(t, T^*_\delta) + P(t, T^*)} = \frac{Z_2(t, T^*_{k\delta})}{1 + Z_2(t, T^*_{2\delta})}.$$

The associated BM is then given by

$$\tilde{B}_t^{T^*_{2\delta}} = \tilde{B}_t^{T^*_\delta} - \int_0^t \frac{Z_2(u, T^*_{2\delta})}{1 + Z_2(u, T^*_{2\delta})} \gamma_2(u, T^*_{2\delta}) du$$

for $t \in [0, T^*_{2\delta}]$. Note that

$$Z_2(t, T^*_{2\delta}) = \frac{P(t, T^*_{2\delta})}{P(t, T^*_\delta) + P(t, T^*)} = \delta\tilde{\kappa}(t, T^*_{2\delta}) + Z_2(t, T^*)$$

with

$$Z_2(t, T^*) = \frac{Z_1(t, T^*)}{\delta\tilde{\kappa}(t, T^*_\delta) + Z_1(t, T^*) + 1},$$

where $Z_1(t, T^*)$ is already defined in the previous step.

In general, if we assume that we have defined forward swap rates $\tilde{\kappa}(t, T^*_\delta), \ldots, \tilde{\kappa}(t, T^*_{m\delta})$, the forward swap measure $\tilde{\mathbb{P}}_{T^*_{(m-1)\delta}}$ and the associated BM $B_t^{T^*_{(m-1)\delta}}$. The aim is to determine the forward swap measure $\tilde{\mathbb{P}}_{T^*_{m\delta}}$, $B_t^{T^*_{m\delta}}$ and of course the forward swap rate $\tilde{\kappa}(t, T^*_{(m+1)\delta})$. We postulate that

$$Z_{m+1}(t, T^*_{k\delta}) = \frac{P(t, T_{k\delta})}{P(t, T_{m\delta}) + \cdots + P(t, T^*)} = \frac{Z_m(t, T^*_{k\delta})}{1 + Z_m(t, T^*_{m\delta})}$$

follow local martingales under $\tilde{\mathbb{P}}_{T^*_{m\delta}}$. With help of lemma 3.1.1 we find the BM:

$$B_t^{T^*_{m\delta}} = B_t^{T^*} - \int_0^t \frac{Z_m(u, T^*_{m\delta})}{1 + Z_m(u, T^*_{m\delta})} \gamma_m(u, T^*_{m\delta}) du$$

for $t \in [0, T^*_{m\delta}]$. As in the step before we see that

$$Z_m(t, T^*_{m\delta}) = \frac{P(t, T^*_{m\delta})}{P(t, T^*_{(m-1)\delta}) + \cdots + P(t, T^*)} = \delta\tilde{\kappa}(t, T^*_{m\delta}) + Z_m(t, T^*)$$

with

$$Z_m(t, T^*) = \frac{Z_{m-1}(t, T^*)}{\delta\tilde{\kappa}(t, T^*_{(m-1)\delta}) + Z_{m-1}(t, T^*) + 1}.$$

$Z_{m-1}(\cdot, T^*)$ is a rational function of the swap rates that have already been found and therefore the BM is defined. Now we easily find the forward martingale measure and thereby an SDE to define $\tilde{\kappa}(t, T^*_{(m+1)\delta})$.

Summary 3.1. *In this section we got to know swap contracts and (forward) swap rates. Finally we saw that it is possible to define a direct model (not explicitly via the LIBOR model) for forward swap rates by martingale methods.*

3.2 More Interest Rate Derivatives

In the following section on interest derivatives the pricing will be done by discounting with the money market account P_t and the measure $\mathbb{P}^*$ under which the assets discounted using the savings account are martingales.

3.2.1 Caps and Floors

A interest rate cap is an agreement in which the grantor (seller) has the obligation to pay cash to the holder (buyer) if a particular interest rate exceeds a certain level at some future date.

A interest rate floor is an agreement in which the grantor (seller) has the obligation to pay cash to the holder (buyer) if a particular interest rate is below a certain level at some future date.

To make some notions clear: a caplet (resp. a floorlet) is then one so called leg of the cap (resp. floor). This means that if we have a cap over N periods then it consists of N caplets. Caps and floors can either be settled in arrears or in advance. We place ourselves in a similar set-up as in a swap agreement. We have points in time $T_j, j = 1, \ldots, n$ where $T_j - T_{j-1} = \delta$ with a fixed $\delta > 0$. The cash flow at time T_j is $N(L(T_{j-1}) - \kappa)^+\delta$. Where κ is the so called cap strike rate, meaning the crucial level of an interest rate and $L(T_{j-1})$ is the interest rate for the period $[T_{j-1}, T_j]$ determined at the reset date T_{j-1}. It satisfies

$$\frac{1}{P(T_{j-1}, T_j)} = 1 + L(T_{j-1})\delta.$$

The arbitrage price of at time $t \leq T_0$ of a forward cap is then ($N = 1$ by assumption)

$$\mathbf{FC}_t = \sum_{j=1}^{n} \mathbf{E}_{\mathbb{P}^*}[\frac{P_t}{P_{T_j}}(L(T_{j-1}) - \kappa)^+\delta|\mathcal{F}_t].$$

Let's consider on caplet with reset date T and settlement date $T_1 = T + \delta$. The value at time t for this caplet then equals:

$$\begin{aligned}
\mathbf{Cpl}_t &= \mathbf{E}_{\mathbb{P}^*}[\frac{P_t}{P_{T_1}}((P(T,T_1)^{-1} - 1)\delta^{-1} - \kappa)^+\delta|\mathcal{F}_t] \\
&= \mathbf{E}_{\mathbb{P}^*}[\frac{P_t}{P_{T_1}}(\frac{1}{P(T,T_1)} - \tilde{\delta})^+|\mathcal{F}_t] \\
&= \mathbf{E}_{\mathbb{P}^*}[\frac{P_t}{P_T}(\frac{1}{P(T,T_1)} - \tilde{\delta})^+\mathbf{E}_{\mathbb{P}^*}[\frac{P_T}{P_{T_1}}|\mathcal{F}_T]|\mathcal{F}_t] \\
&= \mathbf{E}_{\mathbb{P}^*}[\frac{P_t}{P_T}(1 - \tilde{\delta}P(T,T_1))^+|\mathcal{F}_t] \\
&= P(t,T)\mathbf{E}_{\mathbb{P}_T}[(1 - \tilde{\delta}P(T,T_1))^+|\mathcal{F}_t]
\end{aligned}$$

Where the last equality comes from the representation in the forward measure $\mathbb{P}_T$. Like this one sees that a caplet can either be seen as put option on a zero-coupon bond or as an option on a one-period swap.

3.2.2 Captions

In the last section we saw that a caplet can be seen as a put option on a zero-coupon bond. There exist call options on a cap. Such a **caption** is thus a call option on a portfolio of put options. Its pay-off can be described by

$$\mathbf{CC}_T = (\sum_{j=1}^{n} \mathbf{Cpl}_T^j - K)^+$$

with $\mathbf{Cpl}_T^j$ standing for the price at time T of the j^{th} caplet of the cap. T is of course the expiry date of the call option and K is the strike price.

3.2.3 Swaptions

A **swaption** is an option on a swap agreement. The owner of a payer (receiver respectively) swaption with strike rate κ maturing at time $T = T_0$ has the right to enter at time T the underlying forward payer (receiver respectively) swap settled in arrears. From the section before we know that $\mathbf{FS}_T(\kappa)$ is the value at time T of a payer swap with fixed rate κ. It is now clear that the price of a swaption at time t equals

$$\mathbf{PS}_t = \mathbf{E}_{\mathbb{P}^*}[\frac{P_t}{P_T}(\mathbf{FS}_T(\kappa))^+|\mathcal{F}_t].$$

This can be written more explicitly in the following terms:

$$\mathbf{PS}_t = \mathbf{E}_{\mathbb{P}^*}[\frac{P_t}{P_T}(\mathbf{E}_{\mathbb{P}^*}[\sum_{j=1}^{n}\frac{P_T}{P_{T_j}}(L(T_{j-1}) - \kappa)\delta)|\mathcal{F}_T])^+|\mathcal{F}_t].$$

For the receiver swaption we have the price

$$\mathbf{RS}_t = \mathbf{E}_{\mathbb{P}^*}[\frac{P_t}{P_T}(-\mathbf{FS}_T(\kappa))^+|\mathcal{F}_t].$$

Summary 3.2. *In this section we got to know a certain number of interest rate derivatives. We showed some principles how (forward) prices of these derivatives can be expressed and found a structure that resembles the form of an option.*

3.3 Option Valuation

In this section we will assume the bond price volatilities to be bounded deterministic functions. This will be weakened as the final result of this text.

3.3.1 The Arbitrage Price of a Contingent Claim

Definition 3.3.1 (Forward Contract)**.** *A forward contract is an agreement, established at time $t \leq T$, to pay or receive a preassigned payoff - say X - at time T, at an agreed forward price.*

Remark 3.3.1. *The times of course fulfill $t \leq T \leq T^*$. The forward price at time t of the payoff X at time T will be denoted by $F_X(t,T)$.*

Lemma 3.3.1. *The forward price at time t for the date T of an attainable contingent claim X which settles at time T equals*

$$F_X(t,T) = \mathbf{E}_{\mathbb{P}_T}[X|\mathcal{F}_t], \forall t \in [0,T], \tag{3.11}$$

provided that X is $\mathbb{P}_T$-integrable.

Definition 3.3.2 (Arbitrage Price)**.** *The arbitrage-free price of an contingent claim X is called its arbitrage price.*

Lemma 3.3.2. *The arbitrage price of a contingent claim X is given by the formula*

$$\pi_t(X) = P(t,T)\mathbf{E}_{\mathbb{P}_T}[X|\mathcal{F}_t], \forall t \in [0,T]. \tag{3.12}$$

Just some thoughts about the proof. We have that under the forward measure any asset with maturity T discounted with the bond price is a martingale. So for time t we have the discount factor $\frac{P(t,T)}{P(T,T)}$ and the rest follows by measurability of $P(t,T)$. Another way to look at this formula is to think of the spot martingale measure and the savings account and then do a change of measure from the spot to the forward measure as done in sections before.

Corollary 3.3.1. *Let X be an arbitrary attainable contingent claim which settles at time U. If $U \leq T$, then the price of X at time $t \leq U$ equals*

$$\pi_t(X) = P(t,T)\mathbf{E}_{\mathbb{P}_T}[XP(U,T)^{-1}|\mathcal{F}_t].$$

3.3.2 A Version of Black's Formula for Bond Options

The pay-off of a European call option written on a zero-coupon bond which matures at time $U \geq T$ at expiry time T equals

$$C_T = (P(T,U) - K)^+.$$

Because of $P(T,U) = F_P(T,U,T)$ this can be reexpressed by

$$C_T = (F_P(T,U,T) - K)^+ = F_P(T,U,T)\mathbf{1}_D - K\mathbf{1}_D$$

where

$$D = \{P(T,U) > K\}.$$

Theorem 3.3.1. *Assume that the bond price volatilities $b(t,T)$ and $b(t,U)$ are bounded deterministic functions. The arbitrage price at time $t \in [0,T]$ of a European call option with expiry date T and strike price K, written on a zero-coupon bond which matures at time $U \geq T$, equals*

$$C_t = P(t,U)N(h_1(P(t,U),t,T)) - KP(t,T)N(h_2(P(t,U),t,T)), \tag{3.13}$$

where

$$h_{1,2}(b,t,T) = \frac{\ln(b/K) - \ln(P(t,T)) \pm \frac{1}{2}v_U^2(t,T)}{v_U(t,T)} \tag{3.14}$$

for $(b,t) \in \mathbb{R}_+ \times [0,T]$, and

$$v_U^2(t,T) = \int_t^T |b(u,U) - b(u,T)|^2 du, \forall t \in [0,T]. \tag{3.15}$$

The arbitrage price of the corresponding European put option written on a zero-coupon bond equals

$$P_t = KP(t,T)N(-h_2(P(t,U),t,T)) - P(t,U)N(-h_1(P(t,U),t,T)).$$

Proof. From lemma 3.3.2 we know that we have to evaluate the conditional expectation:

$$C_t = P(t,T)\mathbf{E}_{\mathbb{P}_T}[F_P(T,U,T)\mathbf{1}_D|\mathcal{F}_t] - KP(t,T)\mathbb{P}_T(D|\mathcal{F}_t) := I_1 - I_2.$$

Furthermore we know about the dynamics of $F_P(t,U,T)$ under $\mathbb{P}_T$ and thus find

$$F_P(T,U,T) = F_P(t,U,T)\exp(\int_t^T \gamma(u,U,T)\cdot dB_u^T - \frac{1}{2}\int_t^T |\gamma(u,U,T)|^2 du),$$

where $\gamma(u,U,T) = b(u,U) - b(u,T)$. This can be rewritten as

$$F_P(T,U,T) = F_P(t,U,T)\exp(\zeta(t,T) - \frac{1}{2}v_U^2(t,T)).$$

In this formula $F_P(t,U,T)$ is of course $\mathcal{F}_t$- measurable, and
$\zeta(t,T) = \int_t^T \gamma(u,U,T)) \cdot dB_u^T$ is a Gaussian random variable under $\mathbb{P}_T$, independent of $\mathcal{F}_t$ with expectation zero and variance $v_U^2(t,T)$. Using those properties and the properties of conditional expectation we find that

$$\mathbb{P}_T(D|\mathcal{F}_t) = \mathbb{P}_T(\zeta(t,T) < \ln(\frac{F_P(t,U,T)}{K}) - \frac{1}{2}v_U^2(t,T)).$$

Finally we see that

$$I_2 = KP(t,T)N(\frac{\ln(\frac{F_P(t,U,T)}{K}) - \frac{1}{2}v_U^2(t,T)}{v_U(t,T)}).$$

To finish the proof we have to show the expression for I_1. To do this we introduce the equivalent measure $\tilde{\mathbb{P}}_T \sim \mathbb{P}_T$ on $(\Omega, \mathcal{F}_t)$ defined by

$$\frac{d\tilde{\mathbb{P}}_T}{d\mathbb{P}_T} = \exp(\int_0^T \gamma(u,U,T) \cdot dB_u^T - \frac{1}{2}\int_0^T |\gamma(u,U,T)|^2 du) := \tilde{\eta}_T.$$

By Girsanov it is clear that

$$\tilde{B}_t^T = B_t^T - \int_0^t \gamma(u,U,T)du, \forall t \in [0,T]$$

is a standard BM under $\tilde{\mathbb{P}}_T$. For the forward price we get the expression

$$F_P(T,U,T) = F_P(t,U,T)\exp(\int_t^T \gamma(u,U,T) \cdot d\tilde{B}_u^T + \frac{1}{2}\int_t^T |\gamma(u,U,T)|^2 du)$$

so that

$$F_P(T,U,T) = F_P(t,U,T)\exp(\tilde{\zeta}(t,T) - \frac{1}{2}v_U^2(t,T)).$$

$\tilde{\zeta}(t,T) = \int_t^T \gamma(u,U,T) \cdot d\tilde{B}_u^T$ is a Gaussian random variable under $\tilde{\mathbb{P}}_T$, independent of $\mathcal{F}_t$ with expectation zero and variance $v_U^2(t,T)$. Again looking at lemma 3.3.2 we get (note that $P(t,T)$ cancels in the multiplication with $F_P(t,U,T)$)

$$I_1 = P(t,U)\mathbf{E}_{\mathbb{P}_T}[\mathbf{1}_D \exp(\int_t^T \gamma(u,U,T) \cdot dB_u^T + \frac{1}{2}\int_t^T |\gamma(u,U,T)|^2 du)|\mathcal{F}_t].$$

Rewritten this is

$$I_1 = P(t,U)\mathbf{E}_{\mathbb{P}_T}[\tilde{\eta}_T \tilde{\eta}_t^{-1}\mathbf{1}_D|\mathcal{F}_t]$$

which is by using the abstract Bayes rule

$$I_1 = P(t,U)\tilde{\mathbb{P}}_T(D|\mathcal{F}_t).$$

So we can finish the proof with the last two statements:

$$\tilde{\mathbb{P}}_T(D|\mathcal{F}_t) = \tilde{\mathbb{P}}_T(\tilde{\zeta}(t,T) \leq \ln(\frac{F_P(t,U,T)}{K}) + \frac{1}{2}v_U^2(t,T)),$$

and thus

$$I_1 = P(t,U)N(\frac{\ln(F_P(t,U,T)/K) + \frac{1}{2}v_U^2(t,T)}{\frac{1}{2}v_U(t,T)}).$$

The price of the put option can be calculated analogously. □

3.4 Black's Formula for Caplets and the Volatility Smile

We already mentioned cap contracts and caplets. We consider a cap that pays if the LIBOR rate exceeds a strike rate K. If we assume the lognormal model for the LIBOR rate then we can find a price for a caplet using Black's formula.

Remark 3.4.1. *If we speak of the lognormal model then the volatility function is assumed to be deterministic.*

Proposition 3.4.1. *If we price a caplet at time t for the period $[T, T+\delta]$ with strike rate K using Black's formula then we have the following:*

$$P_{C,t} = P(t, T+\delta)\delta[L(t,T)\Phi(d_+) - K\Phi(d_-)],$$

where Φ is the cumulative normal distribution function and

$$d_\pm = \frac{\log(L(t,T)/K) \pm \frac{1}{2}\sigma_T^2(T-t)}{\sigma_T\sqrt{T-t}}.$$

3.4.1 Implied Volatility

Financial contracts such as derivatives are freely traded on various markets and in various currencies. Often there exist more or less useful models to price those derivatives. Prices can be observed in the market and reflect the situation best. Working with stochastic differential equations one has some volatility term. Then one gets an equation of the following type:

$$P_{\text{Market},t}(T, K) = P_{\text{Model},t}(T, K, \sigma), \tag{3.16}$$

where T stands for the maturity, σ for the volatility and K for the strike price (in case of a derivative of the form of an option). If we have a closed expression for the price in the model then we can solve it (usually some numerical method is needed) for the volatility σ. This solution for the unknown volatility is then called the **implied volatility**.

Remark 3.4.2. *The implied volatility can be specified more precisely. If e.g. Black's formula was used to find the price of the option then we call it the Black-implied volatility.*

3.4.2 Volatility-Smile

Comparing model prices to market prices implied volatility reveals some systematics. One could think that this indicator of mispricing does not follow any rules. But it is well known and documented in various papers (such as the one by Jarrow, Li, and Zhao [7]) that the implied volatility of equity options as well as the implied volatility of caplets show significant patterns. The authors of the article [7] give empirical evidence that the implied volatility of LIBOR rates if plotted in dependence on the strike rate form some kind of smile. Meaning that the implied volatility is small if the cap is at the money (so the strike rate equals the LIBOR rate) and significantly higher in case of caps out of the money (LIBOR is smaller than the strike rate) and in the money (LIBOR is higher than the strike rate). So the graph shows the so called **volatility smile**. There is also empirical evidence that this smile is more recognizable after the terrorist attacks of the 11^{th} of September 2001. Moreover there is evidence that this phenomenon is stronger in case of short times to maturity. Finally they observed that this smile is not symmetric. ITM (in the money) caps seem to have a much stronger skew than OTM (out of the money) caps. By observing this phenomenon we found a new task for LIBOR market models. It is desirable to be able to capture the volatility smile. The standard lognormal model is not able to do this. One model that has much more flexibility is described in the following section.

Remark 3.4.3. *The article shows that even the most sophisticated LIBOR models can not fully capture the volatility smile of caps.*

Remark 3.4.4. *Of course the evidence is not only given in form of some plots. There has been made multivariate regression analysis to estimate the dependence on maturity and moneyness of the implied volatility and other statistical methods have been used. My aim here is only to make clear how the need for other models than the lognormal model is motivated.*

3.5 The Stochastic Volatility Model

The stochastic volatility model in this section was defined by L.Andersen and R.Brotherton-Ratcliffe [2]. They assume the volatility to solve a scalar mean-reverting SDE. Namely

$$V(t) \quad = \quad \kappa(\theta - V(t))dt + \eta\psi(V(t))dZ_{n+1}(t), \tag{3.17}$$

where θ, η and κ are positive constants. $Z_{n+1}(t)$ is a Brownian motion under the forward martingale measure $\mathbb{P}_{n+1}$. $\psi : \mathbb{R}^+ \to \mathbb{R}^+$ is a well-behaved function (a concrete choice for ψ will follow). To make sure that V is nonnegative we must assume that $\psi(0) = 0$. In applications it is often natural to scale the process such that $\theta = 1$. Doing that one gives the quantity $1 - V(t)$ the meaning of the percentage at which the volatility differs form the long-term mean.

Definition 3.5.1. *The stochastic volatility model for LIBOR rates of Andersen and Brotherton-Ratcliffe (2001) assumes that the dynamics of the LIBOR rate are given by the following SDE:*

$$dL_n(t) = \varphi(L_n(t))\sqrt{V(t)}\lambda_n(t)^T dB_{n+1}(t) \tag{3.18}$$

where B_{n+1} is a d-dimensional Brownian motion under the forward measure $\mathbb{P}_{n+1}$. $L_n(t) = L(t, T_n)$ is the forward LIBOR rate for the period $[T_n, T_{n+1}]$ of some given partition of the time interval of the model. $\varphi : \mathbb{R}^+ \to \mathbb{R}^+$ is a function satisfying certain regularity conditions. $V(t)$ is the volatility process described in equation 3.17, and $\lambda_n(t)$ is a deterministic d-dimensional function.

Andersen and Brotherton-Ratcliffe moreover assume that $Z_{n+1}(t)$ and $B_{n+1}(t)$ are uncorrelated. This is consistent with the evidence that in all major fixed income markets the correlation between short-dated forward rates and their volatilities is indistinguishable form zero. They propose $\varphi(x) = x^\alpha$ with $0 < \alpha < 1$ as a choice for φ and thereby they were able to find closed-form expressions for the prices of options.

Jarrow, Li and Zhao [7] specialized the model defining the following equations:

$$dL_n(t) = \varphi(L_n(t))\sqrt{V(t)}dB_{n+1}(t), \tag{3.19}$$

$$dV(t) = \kappa(\theta - V(t))dt + \eta\sqrt{V(t)}dZ_{n+1}(t). \tag{3.20}$$

The choice of φ is the same: $\varphi(x) = x^\gamma$ with $\gamma \in (0,1)$. This generates a downward sloping volatility skew. The uncorrelated volatility process $V(t)$ produces a symmetric volatility smile. The combination of both produces an asymmetric smile as desired.

3.5.1 The Observed Model

We want to specialize the general model of Andersen and Brotherton-Ratcliffe. We assume the volatility to follow equation (3.20) - thus we specialize the function $\psi(V(t))$ to $\sqrt{V(t)}$ - and we assume the LIBOR rate to follow

$$dL_n(t) = \varphi(L_n(t))\sqrt{V(t)}\lambda_n(t)^T dB_{n+1}(t). \tag{3.21}$$

We want to price a caplet with strike rate X in this model. The arbitrage-free price is given by

$$C(t) = P(t, T_{n+1})\delta\mathbf{E}^{\mathbb{P}_{n+1}}[(L_n(T_n) - X)^+ | \mathcal{F}_t]. \tag{3.22}$$

We define the function $G(t, L_n(t), V(t)) := \mathbf{E}^{\mathbb{P}_{n+1}}[(L_n(T_n) - X)^+ | \mathcal{F}_t]$. In order to price a European style option we can replace the term $\lambda_n(t)^T dB_{n+1}(t)$ by $\|\lambda_n(t)\| dY(t)$ - where $Y(t)$ is a one dimensional Brownian motion uncorrelated to $Z_{n+1}(t)$. From the Feynman-Kac theorem we find the following PDE for $G(t, L, V)$:

$$\frac{\partial G}{\partial t} + \kappa(\theta - V)\frac{\partial G}{\partial V} + \frac{1}{2}\epsilon^2 V\frac{\partial^2 G}{\partial V^2} + \frac{1}{2}\varphi(L_n)^2 V\|\lambda_n(t)\|^2\frac{\partial^2 G}{\partial L_n^2} = 0,$$

with the boundary condition $G(T_n, L_n, V) = (L_n - X)^+$. There are 2 things to remark:

1. The absence of a mixed derivative of the form $\frac{\partial^2 G}{\partial L_n \partial V}$ comes from the assumption of the vanishing correlation between the two Brownian motions.

2. We have $\psi(V)^2 = V$.

In the following asymptotic calculations I follow the article [2] and use some of their notation dropping the subscript n and writing λ (a scalar) instead of $\|\lambda_n(t)\|$. The PDE then becomes:

$$\frac{\partial G}{\partial t} + \kappa(\theta - V)\frac{\partial G}{\partial V} + \frac{1}{2}\epsilon^2 V \frac{\partial^2 G}{\partial V^2} + \frac{1}{2}\varphi(L)^2 V \lambda^2 \frac{\partial^2 G}{\partial L^2} = 0, \tag{3.23}$$

with the boundary condition $G(T, L, V) = (L - X)^+$.

3.5.2 An Asymptotic Solution for V constant

In this section we assume V to be constant and set it without loss of generality equal to one. Moreover we set $\lambda(t)^2 = c$ for a constant c. The PDE (3.23) then becomes:

$$\frac{\partial G}{\partial t} + \frac{1}{2}\varphi(L)^2 c \frac{\partial^2 G}{\partial L^2} = 0 \tag{3.24}$$

with the boundary condition $G(T, L) = (L - X)^+$.

In this setting the authors of [2] state a proposition for constant volatility:

Proposition 3.5.1. *Let $\tau = T - t$. An asymptotic expansion for the solution to (3.24) is $G(t, L) = g(t, L, c)$ where*

$$g(t, L, c) = L\Phi(d_+) - X\Phi(d_-), d_\pm = \frac{\ln(L/X) \pm \frac{1}{2}\Omega(t, L, c)^2}{\Omega(t, L, c)} \tag{3.25}$$

where Φ is the commulative Gaussian distribution function, and

$$\begin{aligned}
\Omega(t, L, c) &= \Omega_0(L) c^{\frac{1}{2}} \tau^{\frac{1}{2}} + \Omega_1(L) c^{\frac{3}{2}} \tau^{\frac{3}{2}} + \mathcal{O}(\tau^{\frac{5}{2}}), \\
\Omega_0(L) &= \frac{\ln(L/X)}{\int_X^L \varphi(u)^{-1} du}, \\
\Omega_1(L) &= -\frac{\Omega_0(L)}{(\int_X^L \varphi(u)^{-1} du)^2} \ln(\Omega_0(L)(\frac{LX}{\varphi(L)\varphi(X)})^{\frac{1}{2}}).
\end{aligned} \tag{3.26}$$

Proof. The result of this proposition represents an expansion around the volatility term in the Black formula which we got to know in sections before (for the choice $\varphi(L) = L$). $\frac{\Omega(t,F,c)}{\sqrt{\tau}}$ thus represents the implied Black volatility. For $\tau \to 0$ we must require $\Omega(t, F, c) \sim \tau^{\frac{1}{2}}$ so we seek for a solution for small τ of the form

$$\Omega(t, L, c) = \sum_{i \geq 0} c^{i+\frac{1}{2}} \tau^{i+\frac{1}{2}} \Omega_i(L). \tag{3.27}$$

The weights of the integer powers in this series vanish because of the structure of the resulting homogeneous differential equations. Substituting 3.27 into 3.24 and matching terms of order $\tau^{\frac{1}{2}}$ gives

$$L^2\Omega_0^2 = \varphi^2(L)(1 - \frac{\Omega_0'}{\Omega_0} L \ln(L/X))^2 \tag{3.28}$$

where Ω_0' denotes the derivative w.r.t. L. Squaring and rearranging of 3.28 leads to an ODE of the Bernoulli type. The boundary condition (the limit $F \to X$ for Ω_0 must be finite) and discarding the negative solution leads to the expression for Ω_0 in the proposition. For the terms of order $\tau^{\frac{3}{2}}$ in 3.27 we get

$$2L\Omega_1 = \frac{1}{2}\varphi^2(L)(L\Omega_0'' + \Omega_0') - \varphi(L) L \ln(L/X)(\frac{\Omega_1'}{\Omega_1} - \frac{\Omega_1\Omega_0'}{\Omega_0^2}), \tag{3.29}$$

leading again to a Bernoulli ODE. The solution of this ODE considering all constraints leads to the expression in the proposition. □

The authors furthermore introduce the notion of the **implied skew volatility** . As mentioned in the proof $\Omega/\sqrt{T-t}$ represents the implied Black volatility σ_{imp} which is often quoted in interest rate caps markets. Thus we can find a new definition:

Definition 3.5.2 (Implied Skew Volatility). *The solution of $g(t, \lambda_{imp}^2) = G(t)$ where $G(t)$ is the market-quoted cap price is called the implied skew volatility (a scalar in this context).*

The authors give a connection between σ_{imp} and λ_{imp} which can be used for quick calibration of the model if the implied volatility is quoted. The relation is as follows:

$$\sigma_{\text{imp}} \quad \approx \quad \Omega_0(L)\lambda_{\text{imp}} + \Omega_1(L)\lambda_{\text{imp}}^3(T-t). \tag{3.30}$$

One sees this easily be plugging into the proposition: We had that $\Omega/\sqrt{\tau}$ represents the implied volatility. Then we can rearrange to:

$$\sigma_{\text{imp}}\tau^{\frac{1}{2}} \quad = \quad \Omega_0(L)\lambda_{\text{imp}}\tau^{\frac{1}{2}} + \Omega_1(L)\lambda_{\text{imp}}^3\tau^{\frac{3}{2}} + \mathcal{O}(\tau^{\frac{5}{2}})$$

and get the assertion.

λ not Constant in Time

Andersen and Brotherton-Ratcliffe continue their analysis assuming $\lambda(t)^2$ not to be constant anymore in time. V still equals 1. The PDE then becomes:

$$\frac{\partial G}{\partial t} + \frac{1}{2}\varphi(L)^2\lambda(t)^2\frac{\partial^2 G}{\partial L^2} = 0, \tag{3.31}$$

with the boundary condition $G(T, L) = (L - X)^+$. Then the following proposition can be quoted:

Proposition 3.5.2. *An asymptotic solution to (3.31) is given by*

$$G(T,L) = g(t,L,(T-t)^{-1}\int_t^T \lambda(u)^2 du),$$

where g is defined in proposition 3.5.1.

Proof. Performing a deterministic time-change (compare [1]) one finds that the proposition holds unchanged if one sets $c\tau = \int_t^T \lambda(u)^2 du$. This is also seen from the structure of the pricing PDE. □

3.5.3 Stochastic V

We now let V follow the SDE 3.20. The assumption that the Brownian motion of the LIBOR rate and the Brownian motion of the variance process are uncorrelated allows us to write an asymptotic solution as

$$G(t,L,V) = \mathbf{E}[g(t,L,(T-t)^{-1}U(T)|\mathcal{F}_t] \tag{3.32}$$

where $U(T) = \int_t^T \lambda(u)^2 V(u)du$. The expectation is take w.r.t. the forward martingale measure $\mathbb{P}_{n+1}$. The density of this $U(T)$ is not that easy to generate but with some known results on the variance process we can find the first moment:

$$\begin{aligned} \mu_U(t,V(t)) &:= \mathbf{E}[U(T)|\mathcal{F}_t] = \int_t^T \lambda(u)^2 \mathbf{E}[V(u)|\mathcal{F}_t]du && (3.33) \\ &= \int_t^T \lambda(u)^2(\theta + (V(t)-\theta)\exp(-\kappa(u-t)))du && (3.34) \end{aligned}$$

To establish higher moments we now introduce the Laplace transform of $U(T)$ which is given by $\mathbf{E}[\exp(-sU(T))] := L(t,V,s)$, with complex valued s. Andersen and Brotherton-Ratcliffe then introduce the centered Laplace transform $l(t,V,s) = \mathbf{E}[a^{-s[U(T)-\mu_U(t,V)]}]$ and expand it in orders of ϵ^2 - the squared volatility of variance parameter. They use this expansion of the centered Laplace transform to get expressions for the centered moments of higher order of $U(T)$ and finally give a Taylor style expansions using these centered moments. The approach that we want to analyze now is to interpret the variance process as element of the second Wiener - Itô chaos and get expressions for the Laplace transform via that way.

The Variance Process as Element of the Second Chaos

If we consider the Markov property, then it can be proved that for a bounded and continuous function r

$$\mathbf{E}[\exp(\int_t^T r(u,V(u)^x)du)|\mathcal{F}_t]$$

equals

$$\mathbf{E}[\exp(\int_0^{T-t} r(u+t, V(u)^x)du)]|_{x=V(t)}.$$

We can therefore consider the non-conditional expectation and insert $V(t)$ afterwards (for details see [10]).

We have that $V(t)$ can be transformed to a squared Bessel process BESQ $^\delta$. $V(t)$ can be expressed as $e^{-at}Y_{\alpha(t)}$ with Y a BESQ $^\delta$ with $\delta = 4ab/c^2$ and $\alpha(t) = \frac{c^2}{4a}(e^{at}-1)$. We assume $\delta \in \mathbb{N}$ and $\delta \geq 2$. We will write $\alpha(t) = \tau_t$ for short. Then we have a BESQ $^\delta$ Y of the form

$$Y_{\tau_t} = x + \sum_{i=1}^{\delta}(B^i_{\tau_t})^2.$$

The variance process

$$V(t) = x + \int_0^t a(b - V(s))ds + \int_0^t c\sqrt{V(t)}dB_t$$

then equals

$$e^{-at}Y_{\tau_t} = x + e^{-at}\sum_{i=1}^{\delta}(B^i_{\tau_t})^2$$

in law.

Lemma 3.5.1. *Let B be a Brownian motion, then*

$$\int_0^t\int_0^s dB_u dB_s = \int_0^t B_s dB_s = (B_t^2 - t)/2.$$

Proof. From 4.6 we know that $2\int_0^t\int_0^s dB_u dB_s = tH_2(B_t/t^{\frac{1}{2}})$. The result follows easily. ∎

We rewrite $e^{-at}Y_{\tau_t}$ to see the projections onto the chaos 0 and 2 by

$$e^{-at}Y_{\tau_t} = x + e^{-at}\sum_{i=1}^{\delta}(B^i_{\tau_t})^2 = x + 2e^{-at}\sum_{i=1}^{\delta}\int_0^{\tau_t} B^i_s dB^i_s + e^{-at}\delta\tau_t.$$

Now we bring $s\int_0^{T-t}\lambda(u+t)^2e^{-au}Y_{\tau_u}du$ into a suitable form: $s\int_0^{T-t}\lambda(u+t)^2e^{-au}Y_{\tau_u}du =$

$$2s\sum_{i=1}^{\delta}(\int_0^{T-t}\lambda(u+t)^2e^{-au}\int_0^{\tau_u}\int_0^{u_1} dB^i_{u_2}dB^i_{u_1}du) \quad + \quad s\int_0^{T-t}\lambda(u+t)^2(\delta e^{-au}\tau_u + x)du =$$

$$2s\sum_{i=1}^{\delta} I_1 \quad + \quad sI_2.$$

For I_1 we get the following expression by interchanging the order of integration:

$$\int_0^{T-t}\int_0^{\tau_u}\int_0^{u_1}\lambda(u+t)^2 e^{-au}dB^i_{u_2}dB^i_{u_1}du =$$
$$\int_0^{T-t}\int_0^{T-t}\int_0^{T-t}\lambda(u+t)^2 e^{-au}\mathbf{1}_{[0,\tau_u]}(u_1)\mathbf{1}_{[0,u_1]}(u_2)dB^i_{u_2}dB^i_{u_1}du =$$
$$\int_0^{T-t}\int_0^{T-t}\int_0^{T-t}\lambda(u+t)^2 e^{-au}\mathbf{1}_{(u_1,T-t]}(\tau_u)\mathbf{1}_{(u_2,T-t]}(u_1)dB^i_{u_2}dB^i_{u_1}du =$$
$$\int_0^{T-t}\int_0^{T-t}(\int_0^{T-t}\lambda(u+t)^2 e^{-au}\mathbf{1}_{(u_1,T-t]}(\tau_u)du)\mathbf{1}_{(u_2,T-t]}(u_1)dB^i_{u_2}dB^i_{u_1} =$$
$$\int_0^{T-t}\int_0^{u_1}(\int_0^{T-t}\lambda(u+t)^2 e^{-au}\mathbf{1}_{(u_1,T-t]}(\tau_u)du)dB^i_{u_2}dB^i_{u_1} =$$
$$\int_0^{T-t}\int_0^{u_1}(\int_0^{T-t}\lambda(u+t)^2 e^{-au}\mathbf{1}_{(u_1,T-t]}(\tau_u)du)dB^i_{u_2}dB^i_{u_1}.$$

We now symmetrize the kernel of this expression and denote it by $C(u_1,u_2)$ and multiply it by $2s$:

$$C(u_1,u_2) = \int_0^{T-t}\lambda(u+t)^2 e^{-au}2s\frac{1}{2}(\mathbf{1}_{(u_1,T-t]}(\tau_u)+\mathbf{1}_{(u_2,T-t]}(\tau_u))du$$
$$C(u_1,u_2) = \int_0^{T-t}\lambda(u+t)^2 e^{-au}s(\mathbf{1}_{(u_1,T-t]}(\tau_u)+\mathbf{1}_{(u_2,T-t]}(\tau_u))du. \tag{3.35}$$

For I_2 we get a useful expression by plugging in $\delta = \frac{4ab}{c^2}$ and $\tau_u = \alpha(u) = \frac{c^2}{4a}(e^{au}-1)$:

$$I_2 = \int_0^{T-t}\lambda(u+t)^2(b(1-e^{-au})+x)du. \tag{3.36}$$

In order to calculate the Laplace transform we want to calculate $\mathbf{E}[\exp(-Y)]$. If Y is out of the second chaos then we have a useful theorem.

Exponentiated Second Chaos

The following theorem on exponentiated second chaos can be found in a paper of Graselli and Hurd [5]. Please have a look into the appendix for the notation. We want to find the expectation of an L^2-random variable of the form $\exp(-Y)$ with

$$Y = A + \int_\Delta B(\tau_1)dB_{\tau_1} + \int_{\Delta_2} C(\tau_1,\tau_2)dB_{\tau_1}dB_{\tau_2}. \tag{3.37}$$

We introduce the space $\mathcal{H}^{(+)}_{\leq 2} \subset \mathcal{H}_{\leq 2} := \mathcal{H}_0 \oplus \mathcal{H}_1 \oplus \mathcal{H}_2$. If we moreover define C in 3.37 to fulfill $C(\tau_1,\tau_2) = C(\tau_2,\tau_1)$ whenever $\tau_1 > \tau_2$ then C is the kernel of a symmetric integral operator on $L^2(\Delta)$, i.e.

$$[Cf](\tau) = \int_0^\infty C(\tau,\tau_1)f(\tau_1)d\tau_1.$$

Then we say that $Y \in \mathcal{H}_{\leq 2}$ is in $\mathcal{H}_{\leq 2}^{(+)}$ if C is the kernel of a symmetric Hilbert-Schmidt operator on $L^2(\Delta)$ such that $1 + C$ has non-negative spectrum.

Remark 3.5.1. *Recall: Hilbert-Schmidt operators are finite norm operators under the norm*

$$\|C\|_{HS}^2 = \int_{\Delta^2} |C(\tau_1, \tau_2)|^2 d\tau_1 d\tau_2.$$

We assume this Y to be element of a subspace $\mathcal{H}_{\leq 2}^{(+)}$.

Proposition 3.5.3. *Let $Y \in \mathcal{H}_{\leq 2}^{(+)}$. Then*

$$\begin{aligned}\mathbf{E}[\exp(-Y)] &= [det_2(1+C)]^{-\frac{1}{2}} \\ & \exp(-A + \frac{1}{2}\int_{\Delta^2} B(\tau_1)^T (1+C)^{-1}(\tau_1, \tau_2) B(\tau_2) d\tau_1 d\tau_2).\end{aligned}$$

Remark 3.5.2. *The Carleman-Fredholm determinant is defined as the extension of the formula*

$$det_2(1+C) = det(1+C)\exp(-Tr(C))$$

from finite rank operators to bounded Hilbert-Schmidt operators.

The Laplace Transform

Applying the last theorem with

- $A = s\int_0^{T-t} \lambda(u+t)^2 (b(1-e^{-au}) + x) du$ (compare equation 3.36)
- $B = 0$, and
- $C(u_1, u_2) = s\int_0^{T-t} \lambda(u+t)^2 e^{-au} (\mathbf{1}_{(u_1, T-t]}(\tau_u) + \mathbf{1}_{(u_2, T-t]}(\tau_u)) du$ (compare equation 3.35)

leads to an expression for the Laplace transform of a compact form.

Proposition 3.5.4. $\mathbf{E}[\exp(-s\int_t^T \lambda(u)^2 V(u)^x du)|\mathcal{F}_t]$ *can be expressed in the following form:*

$$[det_2(1+C)]^{-\frac{1}{2}} \exp(-A).$$

Proof. This proposition is direct consequence of proposition 3.5.3 and the preceding calculations. ☐

The last proposition applied to the Laplace transform leads to:

$$L(t, V, s) = [det_2(1+C)]^{-\frac{1}{2}} \exp(-s\int_0^{T-t} \lambda(u+t)^2 (b(1-e^{-au}) + V(t)^x) du)$$

where $C(t_1, t_2)$ is defined as above and $V(t)^x$ is the variance process at time t started at x.

Chapter 4

Appendix

4.1 Proofs of Classical Theorems

Proof: 4.1.1. Proof of the Gronwall Lemma *Given a function $v : [0,T] \to \mathbb{R}$ being non-negative and continuous and satisfying $v(t) \leq F + A \cdot \int_0^t v(s)ds, t \in [0,T]$, then $v(t) \leq F \cdot exp(At)$ holds.*

Proof of the lemma: Let's assume $A \neq 0$ (otherwise the lemma is trivial). We define a function $w(t) := \int_0^t v(s)ds$. Thus $w(t) \in C^1([0,T], \mathbb{R})$. Then $w'(t) = v(t), t \in [0,T]$ and thus $w'(t) \leq F + A \cdot \int_0^t v(s)ds$.

From partial integration ($\int_0^t (f(s) \cdot g(s))'ds = \int_0^t h(s)g'(s)ds + \int_0^t g(s)f'(s)ds$)we know that

$$exp(-At)w(t) = \int_0^t exp(-As)w'(s)ds + \int_0^t w(s)(-A)exp(-As)ds$$

- thinking of the fact that $w(0) = 0$.

This is $\leq \int_0^t exp(-As)(F + Aw(s))ds + \int_0^t w(s)(-A)exp(-As)ds$ by the assumption on $w'(t)$. This equals $F \cdot \int_0^t exp(-As)ds$ and finally by integration it follows that $w(t) \leq \frac{F}{A}(exp(At) - 1)$. Which yields the proposition for $w'(t)$ which is identical with $v(t)$. $\square$

Proof: 4.1.2. Proof of Tchebychev's Inequality *Let X be a L^2- random variable then*

$$\mathbb{P}(|X| \geq \lambda) \leq \frac{1}{\lambda^2}\mathbf{E}[X^2]. \tag{4.1}$$

*Proof: We have that $\mathbf{E}[X^2] = \int_{\mathbb{R}} x^2 dF(x) \geq \int_{|x| \geq \lambda} x^2 dF(x) \geq$
$\lambda^2 \cdot \int_{|x|\geq\lambda} dF(x) = \lambda^2 \cdot \mathbb{P}(|X| \geq \lambda)$.*$\square$

Proof: 4.1.3. Proof of the First Borel-Cantelli Lemma

*Let $(E_n : n \in \mathbb{N})$ be a sequence of events such that $\sum \mathbb{P}(E_n) < \infty$.
Then $\mathbb{P}(\limsup E_n) = 0$.*

Proof: We denote $G_m = \bigcup_{n \geq m} E_n$ *and* $G = \limsup(E_n)$*. Then for each* m $\mathbb{P}(G) \leq \mathbb{P}(G_m)$ *by monotone convergence and* $\mathbb{P}(G_m) \leq \sum_{n \geq m} \mathbb{P}(E_n)$ *by basic properties of the probability measure. If we let* $m \to \infty$ *then we have the result.* $\square$

Proof: 4.1.4. Proof of the Second Borel-Cantelli Lemma
If $(E_n : n \in \mathbb{N})$ *is a sequence of independent events, then*

$$\sum \mathbb{P}(E_n) = \infty \Rightarrow \mathbb{P}(\limsup E_n) = 1$$

Proof: First

$$(\limsup E_n)^c = \liminf E_n^c = \bigcap_m \bigcup_{n \geq m} E_n^c.$$

With p_n *denoting* $\mathbb{P}(E_n^c)$*, we have*

$$\mathbb{P}(\bigcup_{n \geq m} E_n^c) = \prod_{n \geq m} (1 - p_n).$$

For $x \geq 0$*,* $1 - x \leq exp(-x)$*, so that, since* $\sum p_n = \infty$

$$\prod_{n \geq m} (1 - p_n) \leq \exp(-\sum_{n \geq m} p_n) = 0$$

So, $\mathbb{P}[(\limsup E_n)^c] = 0$. $\square$

Proof: 4.1.5. Proof of Doob's Martingale Inequality
Let M_t *be an* L^p-*martingale for* $p \in [1, \infty)$ *with continuous trajectories, then*

$$\mathbb{P}(sup_{0 \leq t \leq T} \|M_t\| \geq \lambda) \leq \frac{1}{\lambda^p} E(\|M_T\|^p) \tag{4.2}$$

holds.

To proof this, we need another lemma, the so called

Lemma 4.1.1. Hunt Property *Let a stochastic process* $(S_s)_{s \geq 0}$ *be nonnegative with continuous paths and adapted to a filtration* $(\mathcal{F}_s)_{s \geq 0}$ *and moreover let the submartingale property be fulfilled:* $\mathbf{E}(S_t | \mathcal{F}_s) \geq S_s$ *then for 2 stopping times* σ, τ *with* $\sigma \leq \tau \leq T$ *a.s.* $\mathbf{E}[S_\sigma] \leq \mathbf{E}[S_\tau]$ *holds.*

Proof of the lemma:

Assume that $\sigma \leq \tau \leq T$ *take finitely many values* $0 = t_0 < t_1 < \ldots < t_n = T$*. Then* $\mathbf{E}[S_\tau - S_\sigma] = \mathbf{E}[\sum_{i=1}^n (S_{t_i} - S_{t_{i-1}}) \cdot \underbrace{\mathbf{1}_{\{\sigma \leq t_{i-1} < \tau\}}}_{\mathcal{F}_{t_{i-1}}-measurable}] =$

$\sum_{i=1}^{n} \mathbf{E}[\underbrace{\mathbf{E}[S_{t_i} - S_{t_{i-1}}|\mathcal{F}_{t_{i-1}}]}_{\geq 0} \cdot \mathbf{1}_{\{\sigma \leq t_{i-1} < \tau\}}] \geq 0$. *Now consider the stopping times*

$$\tau^n := \{\frac{\tau_i}{2^n} \text{ if } \frac{\tau_{(i-1)}}{2^n} < \tau \leq \frac{\tau_i}{2^n}\} \ i = 1, \dots, 2^n.$$

Then $\mathbf{E}[S_{\tau^n} - S_{\sigma^n}] \to \mathbf{E}[S_\tau - S_\sigma]$ *for* $n \to \infty$ *pointwise by the continuity of the paths. Moreover* $\{S_\tau | \tau \leq T, \ \tau$ *a stopping time taking finitely many values*$\}$ *is a uniformly integrable family:*

$\mathbf{E}[S_\tau \mathbf{1}_{\{S_\tau \geq \lambda\}}] = \sum_{i=1}^{n} \mathbf{E}[S_{t_i} \mathbf{1}_{\{S_{t_i} \geq \lambda\}} \mathbf{1}_{\{\tau = t_i\}}] \underbrace{\leq}_{submartingale} \sum_{i=1}^{n} \mathbf{E}[S_T \mathbf{1}_{\{S_{t_i} \geq \lambda\}} \mathbf{1}_{\{\tau = t_i\}}] = \mathbf{E}[S_T \mathbf{1}_{\{S_T \geq \lambda\}}] \leq \mathbf{E}[S_T \mathbf{1}_{\{\max_{0 \leq s \leq T} S_s \geq \lambda\}}] \to 0$ *as* $\lambda \to \infty$.

Now we can prove Doob's inequality: If $(M_t)_{t \geq 0}$ *is a martingale with continuous paths, then* $S_t = |M_t|$ *is a submartingale with cont. paths and* $S_t \geq 0$. *One sees this easily using Jensen:*

$$\mathbf{E}[|M_t||\mathcal{F}_s] \underbrace{\geq}_{Jensen} |\mathbf{E}[M_t|\mathcal{F}_s]| = |M_s|, \quad t \geq s$$

We introduce a stopping time:

$$\sigma = \inf\{s \geq 0 | |M_s| \geq \lambda\} \wedge T$$

Then we see that:

$$\mathbb{P}(max_{0 \leq t \leq T} |M_t| \geq \lambda) \lambda^p \leq \mathbf{E}[|M_\sigma|^p \frac{1}{\lambda^p}] \lambda^p \leq \mathbf{E}[|M_T|^p] \text{ (by Hunt)}$$

Remark 4.1.1. $(|M_t|^p)_{t \geq 0}$ *is a submartingale.*

$\square$

4.2 Theorems from Measure Theory

The knowledge of fundamental notions such as the notion of a σ-algebra and of a measurable function is presupposed.

Theorem 4.2.1 (Monotone Convergence)**.** *Let f_n be a monotone, nondecreasing sequence of nonnegative, measurable functions and let f denote its limit. Moreover let $\mathbb{P}$ be a probability measure. Then*

$$\int f d\mathbb{P} = \lim_n \int f_n d\mathbb{P}.$$

In terms of an expectation:

$$\mathbf{E}[lim_n f_n] = \mathbf{E}[f] = lim_n \mathbf{E}[f_n]$$

Theorem 4.2.2 (Monotone Convergence - extended)**.** *Let f_n be a monotone , nondecreasing sequence of measurable functions moreover let g be a measurable function with $\int g d\mathbb{P} > -\infty$ and $f_n \geq g \quad \forall n$ and let f denote its limit, then*

$$\int f d\mathbb{P} = \lim_n \int f_n d\mathbb{P}.$$

Theorem 4.2.3 (Dominated Convergence)**.** *Let f_n be a sequence of measurable functions, moreover let f and g be measurable and let $|f_n| \leq g$ a.s. and $\int |g| d\mathbb{P} < \infty$ then $lim_n f_n = f$ a.s. implies*

$$lim_n \int f_n d\mathbb{P} = \int f d\mathbb{P}$$

or in terms of expectation:

$$\mathbf{E}[f] = lim_n \mathbf{E}[f_n].$$

Remark 4.2.1. *Those theorems tell us under which condition we are allowed to 'pull out' a limit. By that we get properties of a sequence for its limit.*

Theorem 4.2.4 (Theorem of Radon-Nikodym)**.** *Let $\mathbb{P}$ and $\mathbb{Q}$ be two probability measures on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathbb{Q}$ be absolutely continuous w.r.t. $\mathbb{P}$ then there is an $\mathbb{P}$- a.s. unique random variable X such that*

$$\mathbb{Q}(A) = \int_A X d\mathbb{P}$$

for all sets of events $A \in F$.

Remark 4.2.2. *This theorem tells us how a change of measures is done and under which circumstances it can be done. This is often used to change from a natural probability measure $\mathbb{P}$ to a martingale measure $\mathbb{Q}$.*

Now assume two equivalent measures $\mathbb{P}$ and $\mathbb{Q}$ defined on a common measure space $(\Omega, \mathcal{F})$. Let the Radon-Nikodym derivative of $\mathbb{Q}$ w.r.t. $\mathbb{P}$ equal

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \eta$$

$\mathbb{P}$-a.s. This r.v. η is strictly positive $\mathbb{P}$-a.s. (by equivalence) and $\mathbb{P}$-integrable with $\mathbf{E}_{\mathbb{P}}[\eta] = 1$. Finally it is clear that $\mathbf{E}_{\mathbb{Q}}[\psi] = \mathbf{E}_{\mathbb{P}}[\eta\psi]$ holds for any $\mathbb{Q}$-integrable r.v. ψ.

Lemma 4.2.1. *Let $\mathcal{G}$ be a sub-σ-algebra of the σ-algebra $\mathcal{F}$, and let ψ be a r.v. integrable w.r.t. $\mathbb{Q}$. Then the following abstract version of the Bayes formula holds:*

$$\mathbf{E}_{\mathbb{Q}}[\psi|\mathcal{G}] = \frac{\mathbf{E}_{\mathbb{P}}[\eta\psi|\mathcal{G}]}{\mathbf{E}_{\mathbb{P}}[\eta|\mathcal{G}]} \tag{4.3}$$

Proof. It is easily checked that $\mathbf{E}_{\mathbb{P}}[\eta|\mathcal{G}]$ is strictly positive $\mathbb{P}$-a.s. so that the right hand side is well defined. By assumption $\eta\psi$ is $\mathbb{P}$-integrable and we have to show that

$$\mathbf{E}_{\mathbb{P}}[\eta\psi|\mathcal{G}] = \mathbf{E}_{\mathbb{Q}}[\psi|\mathcal{G}]\mathbf{E}_{\mathbb{P}}[\eta|\mathcal{G}].$$

The right hand side of this formula defines a $\mathcal{G}$-measurable r.v. so we have to show that for any $A \in \mathcal{G}$, we have

$$\int_A \eta\psi d\mathbb{P} = \int_A \mathbf{E}_{\mathbb{Q}}[\psi|\mathcal{G}]\mathbf{E}_{\mathbb{P}}[\eta|\mathcal{G}]d\mathbb{P}.$$

But for every $A \in \mathcal{G}$, we get:

$$\begin{aligned}\int_A \eta\psi d\mathbb{P} = \int_A \psi d\mathbb{Q} &= \int_A \mathbf{E}_{\mathbb{Q}}[\psi|\mathcal{G}]d\mathbb{Q} = \int_A \mathbf{E}_{\mathbb{Q}}[\psi|\mathcal{G}]\eta d\mathbb{P} \\ &= \int_A \mathbf{E}_{\mathbb{P}}[\mathbf{E}_{\mathbb{Q}}[\psi|\mathcal{G}]\eta|\mathcal{G}]d\mathbb{P} = \int_A \mathbf{E}_{\mathbb{Q}}[\psi|\mathcal{G}]\mathbf{E}_{\mathbb{P}}[\eta|\mathcal{G}]d\mathbb{P}.\end{aligned} \tag{4.4}$$

□

4.3 Wiener-Itô Chaos Expansion

The Wiener-Itô chaos expansion leads to an orthogonal decomposition of an L^2-random variable. A very clear introduction to this topic is given in [14]. Let B_t be a $1-$dimensional Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with the filtration $(\mathcal{F}_t)_{t\geq 0}$ generated by $B_s, s \leq t$. Fix $T > 0$. We consider functions g of the form $g : [0,T]^n \to \mathbb{R}$. Such a g is called symmetric if $g(x_1, \ldots, x_n) = g(x_{\pi(1)}, \ldots, x_{\pi(n)})$ for all permutations π on $\{1, \ldots, n\}$. If in addition

$$\|g\|^2_{L^2([0,T]^n)} := \int_{[0,T]^n} g^2(x_1, \ldots, x_n) dx_1 \ldots dx_n < \infty$$

then we say that $g \in \hat{L}^2([0,T]^n)$, the space of symmetric square integrable functions on $[0,T]^n$. We define the set $S_n = \{(x_1, x_2, \ldots, x_n) \in [0,T]^n; 0 \leq x_1 \leq x_2 \leq \ldots \leq x_n \leq T\}$. This set occupies the fraction $\frac{1}{n!}$ of the whole n-dimensional box $[0,T]^n$ and therefore for $g \in \hat{L}^2([0,T]^n)$ we have

$$\|g\|^2_{L^2([0,T]^n)} = n! \int_{S_n} g^2(x_1, \ldots, x_n) dx_1 \ldots dx_n = n! \|g\|^2_{S_n}. \tag{4.5}$$

For any real function f defined on $[0,T]^n$ we define the symmetrization of $\tilde{f}$ of f by

$$\tilde{f} = \frac{1}{n!} \sum_{\sigma} f(x_{\sigma_1}, \ldots, x_{\sigma_n}),$$

where the sum is taken over all permutations of $\{1, \ldots, n\}$.

Example 4.1. *The following examples explain how to proceed in order to symmetrize a function:*

- *Consider the function* $f(x_1, x_2) = x_1^2 + x_2 \sin(x_1)$. *Its symmetrization is*

$$\tilde{f} = \frac{1}{2}[x_1^2 + x_2^2 + x_2 \sin(x_1) + x_1 \sin(x_2)].$$

- *Now, consider the function* $f(t_1, t_2) = \mathbf{1}_{\{t_1 < t < t_2\}}$ *then its symmetrization is*

$$\tilde{f} = \frac{1}{2}(\mathbf{1}_{\{t_1 < t < t_2\}} + \mathbf{1}_{\{t_2 < t < t_1\}}).$$

For a deterministic f defined on $S_n, n \geq 1$ such that

$$\|f\|^2_{L^2(S_n)} := \int_{S_n} f^2(t_1, \ldots, t_n) dt_1 \ldots dt_n < \infty$$

we can define the n-fold Itô integral by

$$J_n(f) = \int_0^T \int_0^{t_n} \cdots \int_0^{t_2} f(t_1, \ldots, t_n) dB_{t_1} \cdots dB_{t_{n-1}} dB_{t_n}.$$

This object is well-defined because at each Itô integration w.r.t. B_{t_i} the integrand is $\mathcal{F}_{t_i}-$adapted and square-integrable w.r.t. $dP \times dt_i, 1 \leq i \leq n$. Iterated application of the Itô isometry leads to

$$\mathbf{E}[J_n^2(h)] = \|h\|^2_{L^2(S_n)}$$

for a deterministic $h \in L^2(S_n)$. For $g \in L^2(S_m)$ and $h \in L^2(S_n)$ with $m < n$ we get

$$\mathbf{E}[J_m(g)J_n(h)] = 0.$$

For a $g \in \hat{L}^2([0,T]^n)$ we define

$$I_n(g) = \int_{[0,T]^n} g(t_1, \ldots, t_n) dB_t^{\bigotimes n} := n! J_n(g).$$

Obviously we have from previous results that

$$\mathbf{E}[I_n^2(g)] = \mathbf{E}[(n!)^2 J_n^2(g)] = (n!)^2 \|g\|^2_{L^2(S_n)} = n! \|g\|^2_{L^2([0,T]^n)}$$

for all $g \in \hat{L}^2([0,T]^n)$.

Moreover for each $n \geq 0$ we have the n_{th} **Hermite polynomial** defined by

$$H_n(x) = (-1)^n e^{x^2/2} \frac{d^n}{dx^n} e^{-x^2/2}.$$

Then $\mathcal{H}^{(n)}$ (the n_{th} **Wiener chaos**) denotes the space $\text{span}\{H_n(\int_0^T h_t dB_t) | h \in L^2([0,T])\}$, $n > 0$ and $\mathcal{H}^{(0)}$ denotes the space of constant random variables. One can show that these spaces form an orthogonal decomposition of the space $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$ i.d.

$$L^2(\Omega, \mathcal{F}_T, \mathbb{P}) = \oplus_{n=0}^{\infty} \mathcal{H}^{(n)}.$$

There is a useful formula for the special case that the integrand is the tensor power of a function $g \in L^2([0,T])$:

$$n! \int_0^{\infty} \int_0^{t_n} \cdots \int_0^{t_2} g(t_1) g(t_2) \cdots g(t_n) dB_{t_1} dB_{t_2} \cdots dB_{t_n} = \|g\|^n H_n\left(\frac{\int_0^T g(t) dB_t}{\|g\|}\right), \quad (4.6)$$

where the norm is the norm in $L^2([0,T])$. The formula simplifies if $\|g\|_{L^2([0,T])} = 1$ is assumed.

Example 4.2. *We can write an element of* $\mathcal{H}^{(3)}$ *if we assume that* f *is one-dimensional as*

$$\int_0^T \int_0^{s_3} \int_0^{s_2} f_3(s_1, s_2, s_3) dB_{s_1} dB_{s_2} dB_{s_3}.$$

Theorem 4.3.1 (The Wiener-Itô chaos expansion)**.** *Let φ be an $\mathcal{F}_T-$measurable random variable such that $\|\varphi\|^2_{L^2(\Omega)} := \mathbf{E}_{\mathbb{P}}[\varphi^2] < \infty$. Then there exists a unique sequence $(f_n)_{n=0}^{\infty}$ of deterministic functions $f_n \in \hat{L}^2([0,T]^n)$ such that*

$$\varphi = \sum_{n=0}^{\infty} I_n(f_n) \quad \textit{(convergence in } L^2\textit{)}.$$

Moreover, we have the isometry

$$\mathbf{E}_{\mathbb{P}}[\varphi^2] = \sum_{n=0}^{\infty} n! \|f_n\|^2_{L^2([0,T]^n)}.$$

Graselli and Hurd ([5]) use a different notation to express stochastic integrals w.r.t. Brownian motion of dimension n. They define $\Delta := \mathbb{R}_+ \times \{1, \ldots, N\}$ and for $h \in L^2(\Delta)$ the integral $\int h(\tau)dB_\tau := \sum_\mu \int_0^\infty f(s,\mu)dB_s^\mu$ where τ denotes the pair $(s,\mu) \in \Delta$.

An element of the 2_{nd} chaos is then written as $\int_{\Delta_2} C(\tau_1, \tau_2) dB_{\tau_1} dB_{\tau_2}$, where we have the definition $\Delta_n := \{(\tau_1, \ldots, \tau_n) | \tau_i = (s_i, \mu_i) \in \Delta, 0 \leq s_1 \leq \ldots \leq s_n < \infty\}$.

4.4 Mathematical Tools

4.4.1 Solution Methods of ODE's

Bernoulli ODE's

A Bernoulli ordinary differential equation is an equation of the form (we have y is $y(x)$):

$$y' + yf(x) = y^a g(x) \tag{4.7}$$

with $a \neq 0, 1$. The substitution $z = y^{1-a}, z' = (1-a)y^{-a}y'$ leads to a linear ODE of the type: $z' + (1-a)zf(x) = (1-a)g(x)$.

Linear ODE's

A linear ODE is an ODE of the form

$$y' + a(x)y = f(x).$$

The solution of such an ODE is:

$$y(x) = \exp(-A(x))(\int \exp(A(x))f(x)dx + C)$$

with $A(x) = \int a(x)dx$.

4.4.2 Lévy's Criterion for Brownian Motion

Let B_t be a process adapted to the filtration $\mathcal{F}_t$. Let B_t have the following properties:

1. B_t has continuous sample paths.
2. B_t is a martingale.
3. B_t has quadratic variation t.

Then B_t is a Brownian motion.

Proof. It is enough to show that for every $\alpha \in \mathbb{R}$ $\exp(\alpha B_t - \frac{1}{2}\alpha^2 t)$ is a martingale. We have that for $t > s \geq 0$

$$E[\exp(\alpha(B_t - B_s))|\mathcal{F}_s] = \exp(\frac{1}{2}\alpha^2(t-s)).$$

We see this using the martingale property and rearranging the expressions. This shows that

- $B_t - B_s$ is independent of $\mathcal{F}_s$ because the right hand side is constant.
- $B_t - B_s$ is N(0,t-s) - thinking of the characteristic function.

To see that the expression is a martingale we define $f(x,t) = \exp(\alpha x - \frac{1}{2}\alpha^2 t)$. Then we see by Itô's formula that $f(B_t, t)$ is a martingale because $\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2} = 0$. □

4.4.3 The Martingale Problem

For a short introduction to this problem read [13].

If X_t is an Itô diffusion in $\mathbb{R}^n$ of the form $dX_t = b(X_t)dt + \sigma(X_t)dB_t$ with generator A and if $f \in C_0^2(\mathbb{R}^n)$ then

$$f(X_t) = f(x) + \int_0^t Af(X_s)ds + \int_0^t \nabla f^T(X_s)\sigma(X_s)dB_s.$$

Then we can define $M_t = f(X_t) - \int_0^t Af(X_r)dr$ and this process is a martingale w.r.t. the filtration $\mathcal{M}_t$ which is generated by the process X_t. We can identify each $\omega \in \Omega$ with $X_t^x(\omega)$ and thereby identify the space $(\Omega, \mathcal{M}, \mathbb{Q}^x)$ with $(C[0,\infty)^n, \mathcal{B}(C[0,\infty)^n), \tilde{\mathbb{Q}}^x)$. Regarding the law of X_t^x as a probability measure $\tilde{\mathbb{Q}}^x$ on $\mathcal{B}$ we can formulate a theorem:

Theorem 4.4.1. *If $\tilde{\mathbb{Q}}^x$ is the probability measure induced by the law $\mathbb{Q}^x$ of an Itô diffusion X_t, then for all $f \in C_0^2(\mathbb{R}^n)$ the process M_t (as defined above) is a $\tilde{\mathbb{Q}}^x-$martingale. w.r.t. the Borel $\sigma-$algebra of $C[0,\infty)^n$.*

We can put this theorem into a bit different words and make a definition:

Definition 4.4.1 (Martingale Problem)**.** *Let L be a semi-elliptic differential operator of the form*

$$L = \sum b_i \frac{\partial}{\partial x_i} + \sum_{i,j} a_{ij} \frac{\partial^2}{\partial x_i \partial x_j}$$

where the coefficients b_i, a_{ij} are locally bounded Borel measurable functions on $\mathbb{R}^n$. Then we say that a probability measure $\tilde{P}^x$ on $(C[0,\infty)^n, \mathcal{B}(C[0,\infty)^n))$ solves the martingale problem for L (starting at x) if the process

$$M_t = f(\omega_t) - \int_0^t Lf(\omega_r)dr, M_0 = f(x) \quad \tilde{P}^x\text{-a.s.}$$

is a $\tilde{P}^x$ martingale w.r.t. $\mathcal{B}_t$, for all $f \in C_0^2(\mathbb{R}^n)$. The martingale problem is called well-posed if there is a unique measure $\tilde{P}^x$ solving the martingale problem.

Like this we see that $\tilde{\mathbb{Q}}^x$ solves the martingale problem for A whenever X_t is a weak solution of the SDE

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t. \tag{4.8}$$

Conversely it can be proved that if $\tilde{\mathbb{P}}^x$ solves the martingale problem for an operator L starting at x, for all $x \in \mathbb{R}^n$, then there exists a weak solution of the SDE 4.8. The following result of Strook and Varadhan gives conditions if a martingale problem has a solution.

Proposition 4.4.1. *$L = \sum b_i \frac{\partial}{\partial x_i} + \sum_{i,j} a_{ij} \frac{\partial^2}{\partial x_i \partial x_j}$ has a unique solution of the martingale problem if $[a_{ij}]$ is everywhere positive definite, a_{ij} is continuous, $b(x)$ is measurable and there exists a constant D such that*

$$|b(x)| + |a(x)|^{\frac{1}{2}} \leq D(1 + |x|) \quad \textit{for all } x \in \mathbb{R}^n.$$

Conditions on the Existence of a Weak Solution

Karatzas and Shreve [8] give various conditions when the solvability of the martingale problem and the existence of a weak solution to an SDE of the form

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t \tag{4.9}$$

are implied by each other. We have the following assertions:

1. There exists a weak solution to 4.9 with initial distribution μ.

2. There exists a solution $\mathbb{P}$ to the local martingale problem associated with the differential operator A with $\mathbb{P}(y(0) \in \Gamma) = \mu(\Gamma); \Gamma \in \mathcal{B}(\mathbb{R}^n)$.

3. There exists a solution $\mathbb{P}$ to the martingale problem associated with A with $\mathbb{P}(y(0) \in \Gamma) = \mu(\Gamma); \Gamma \in \mathcal{B}(\mathbb{R}^n)$.

Proposition 4.4.2. *Condition (1) and (2) are equivalent and are implied by (3). Furthermore (1) implies (3) under either of the additional assumptions:*

- *For each* $T < \infty$ $\|\sigma(t, y)\| \leq K_T; 0 \leq t \leq T, y \in C[0, \infty)^n$, *where* K_T *is a constant depending on* T.
- *Each* σ_{ij} *is of the form* $\sigma_{ij}(t, y) = \tilde{\sigma}_{ij}(t, y(t))$ *where the Borel-measurable functions* $\tilde{\sigma}_{ij} : [0, \infty) \times \mathbb{R}^n \to \mathbb{R}$ *are bounded on compact sets.*

4.4.4 Random Time Change

Øksendal [13] describes a random time change in the following way. We have an $\mathcal{F}_t$-adapted r.v. $c_t \geq 0$. We define $\beta_t = \int_0^t c_s ds$. Then we call this β_t a **random time change** with **time change rate** c_t. β_t is also $\mathcal{F}_t$-adapted and the mapping $t \to \beta_t$ is non-decreasing. If we now define a r.v. α_t by $\alpha_t = \inf\{s; \beta_s > t\}$. Then α_t is the right inverse of β_t i.e. $\beta(\alpha_t) = t, \quad \forall t \geq 0$ and the mapping $t \to \alpha_t$ is right continuous. If we assume that $c_s > 0$ for almost all s then $t \to \beta_t$ is strictly increasing, $t \to \alpha_t$ is continuous and α_t is also the left inverse of β_t i.e. $\alpha(\beta_t) = t, \quad \forall t \geq 0$. We now ask how and under which condition a process can be transformed into a BM by a random time change. We have the following theorem:

Theorem 4.4.2. *Let $dY_t = v_t dB_t, v \in \mathbb{R}^{n\times m}, B_t \in R^m$ be an Itô integral in $\mathbb{R}^n$,$Y_0 = 0$ and assume that*

$$vv^T(t) = c_t I_n$$

for some process $c_t \geq 0$. Let α_t, β_t be as in the definition of a random time change. Then Y_{α_t} is an n-dimensional BM.

Corollary 4.4.1. *Let $c_t \geq 0$ be given and define $Y_t = \int_0^t \sqrt{c_s} dB_s$, where B_s is an n-dimensional BM. Then Y_{α_t} is also an n-dimensional BM.*

We now know how we get a new BM but for applications it is of great interest how Itô integrals change. So we change time in an Itô integral which will lead to an Itô integral again but driven by a different BM $\tilde{B}$.

Lemma 4.4.1. *Suppose $s \to \alpha_s$ is continuous, $\alpha_0 = 0$ for almost all ω. Fix $t > 0$ such that $\beta_t < \infty$ and assume that $\mathbf{E}[\alpha_t] < \infty$. For $k = 1, 2, \dots$ put*

$$t_j = \begin{cases} j2^{-k} & \textit{if } j2^{-k} \leq \alpha_t, \\ t & \textit{if } j2^{-k} > \alpha_t. \end{cases}$$

and choose r_j such that $\alpha_{r_j} = t_j$. Suppose $f \geq 0$ is $\mathcal{F}_s$-adapted, bounded and continuous in s for almost all ω. Then

$$\lim_{k\to\infty} \sum_j f(\alpha_j)\Delta B_{\alpha_j} = \int_0^{\alpha_t} f_s dB_s \quad a.s., \tag{4.10}$$

where $\alpha_j = \alpha_{r_j}, \Delta B_{\alpha_j} = B_{\alpha_{j+1}} - B_{\alpha_j}$ and the limit is in $L^2(\Omega, \mathbb{P})$.

Proof. For all k we have

$$\begin{aligned} \mathbf{E}[(\sum_j f(\alpha_j)\Delta B_{\alpha_j} - \int_0^{\alpha_t} f_s dB_s)^2] &= \mathbf{E}[(\sum_j \int_{\alpha_j}^{\alpha_{j+1}} (f(\alpha_j) - f_s) dB_s)^2] \\ \sum_j \mathbf{E}[(\int_{\alpha_j}^{\alpha_{j+1}} (f(\alpha_j) - f_s) dB_s)^2] &= \sum_j \mathbf{E}[\int_{\alpha_j}^{\alpha_{j+1}} (f(\alpha_j) - f_s)^2 ds] \\ = \mathbf{E}[\int_0^{\alpha_t} (f - f_k)^2 ds], \end{aligned}$$

where $f_k(s) = \sum_j f(t_j)\mathbf{1}_{(t_j,t_{j+1}]}(s)$ is the elementary (predictable) approximation to f. □

As main result of this section we state a **time change formula for Itô integrals**:

Theorem 4.4.3. *Suppose c_s and α_s are continuous in s and $\alpha_0 = 0$ for almost all ω and that $\mathbf{E}[\alpha_t] < \infty$. Let B_s be an n-dimensional BM and let $v_s \in \mathcal{V}_{\mathcal{H}}^{n\times m}$ be bounded and continuous in s. Define*

$$\tilde{B}_t = \lim_{k\to\infty} \sum_j \sqrt{c_{\alpha_j}}\Delta B_{\alpha_j} = \int_0^{\alpha_t} \sqrt{c_s} dB_s. \tag{4.11}$$

Then $\tilde{B}_t$ is an $\mathcal{F}^{(m)}_{\alpha_t}$ BM and

$$\int_0^{\alpha_t} v_s dB_s = \int_0^t v_{\alpha_r}\sqrt{\alpha_r'} d\tilde{B}_r, \mathbb{P} - a.s. \tag{4.12}$$

where α_r' is the derivative of α_r w.r.t. r, so that $\alpha_r' = \frac{1}{c_{\alpha_r}}$ for almost all $r \geq 0$ and almost all $\omega \in \Omega$.

Remark 4.4.1. *This theorem is taken from [13] and there is no use to explain the notation in detail. The author uses the space $\mathcal{V}_{\mathcal{H}}^{n\times m}$ to denote the space of properly adapted integrands for an Itô integral. The filtration $\mathcal{F}^{(m)}$ is the one generated by the m-dimensional BM.*

Proof. The existence of the limit in equation 4.11 and the identity follow by applying the last lemma to the function $f_s = \sqrt{c_s}$. Then by the last corollary we have that $\tilde{B}$ is an $\mathcal{F}^{(m)}_{\alpha_t}$ BM. It remains to prove equation 4.12:

$$\begin{aligned}
\int_0^{\alpha_t} v_s dB_s &= lim_{k\to\infty} \sum_j v(\alpha_j)\Delta B_{\alpha_j} \\
&= lim_{k\to\infty} \sum_j v(\alpha_j)\sqrt{\frac{1}{c(\alpha_j)}}\sqrt{c(\alpha_j)}\Delta B_{\alpha_j} \\
&= lim_{k\to\infty} \sum_j v(\alpha_j)\sqrt{\frac{1}{c(\alpha_j)}}\Delta \tilde{B}_j \\
&= \int_0^t v(\alpha_r)\sqrt{\frac{1}{c(\alpha_r)}} d\tilde{B}_r.
\end{aligned}$$

□

4.4.5 Bessel Processes

This introduction to the theory of Bessel process is taken from [15]. To define and analyze Bessel processes we need the following notation: BM^{δ} will denote a $\delta-$dimensional Brownian motion. The process ρ_t denotes the modulus of BM^{δ} and $\mathbb{P}^x$ denotes the measure of the Brownian motion started at x. $\mathcal{F}_t$ denotes the complete Brownian filtration. Now we can state a proposition on the density of ρ_t:

Proposition 4.4.3. *For every $\delta \geq 1$ the process ρ_t, $t \geq 0$, is a homogeneous $\mathcal{F}_t-$Markov process with respect to each $\mathbb{P}^x, x \in \mathbb{R}^{\delta}$. For $\delta \geq 2$, its semi-group P_t^{δ} is given on $[0, \infty)$ by the densities*

$$p_t^{\delta}(a,b) = (a/t)(b/a)^{\delta/2} I_{\delta/2-1}(ab/t) \exp(-(a^2+b^2)/2t)$$

for $a, b > 0$, where I_ν is the modified Bessel function of index ν, and

$$p_t^{\delta}(0,b) = \Gamma(\delta/2) 2^{\delta/2-1} t^{-\delta/2} b^{\delta-1} \exp(-b^2/2t).$$

Definition 4.4.2 (Bessel Process)**.** *A Markov process with semi-group P_t^{δ} is called a $\delta-$ dimensional Bessel process.*

We will denote such a process by BES^{δ} and $\mathrm{BES}^{\delta}(x)$ if it starts at $x \geq 0$. Itô's formula shows that ρ solves

$$\rho_t^2 = \rho_0^2 + 2\sum_{i=1}^{\delta} \int_0^t B_s^i dB_s^i + \delta t.$$

For $\delta > 1, \rho_t$ is a.s. greater than zero and for $\delta = 1$ the set $\{s : \rho_s = 0\}$ has a.s. zero Lebesgue measure. So we can consider the process

$$\beta_t = \sum_{i=1}^{\delta} \int_0^t (B_s^i/\rho_s) dB_s^i$$

which is a Brownian motion since $\langle \beta, \beta \rangle = t$. Therefore ρ_t^2 satisfies the SDE

$$\rho_t^2 = \rho_0^2 + 2\int_0^t B_s^i d\beta_s + \delta t.$$

If we set, for any real $\delta \geq 0$ and $x \geq 0$, the process

$$Z_t = x + 2\int_0^t \sqrt{|Z_s|} d\beta_s + \delta t$$

then it can be shown that this SDE has a unique strong solution which is ≥ 0 thus the absolute value can be discarded.

Definition 4.4.3. *For every $\delta \geq 0$ and $x \geq 0$ the unique strong solution of the equation*

$$Z_t = x + 2\int_0^t \sqrt{Z_s}d\beta_s + \delta t$$

is called the square of the $\delta-$dimensional Bessel process started at x and is denoted by $BESQ^\delta(x)$.

Remark 4.4.2. *The law of $BESQ^\delta(x)$ on $C(\mathbb{R}_{\geq 0}, \mathbb{R})$is denoted by $\mathbb{Q}_x^\delta$. The number $\nu = (\delta/2) - 1$ is called the index of the corresponding process. We write $BESQ^{(\nu)}$ instead of $BESQ^\delta$ sometimes for convenience.*

We have the following theorem on the convolution of measures of squared Bessel processes:

Theorem 4.4.4. *For every $\delta, \delta' \geq 0$ and $x, x' \geq 0$,*

$$\mathbb{Q}_x^\delta * \mathbb{Q}_{x'}^{\delta'} = \mathbb{Q}_{x+x'}^{\delta+\delta'}.$$

Corollary 4.4.2. *If μ is a measure on $\mathbb{R}_{\geq 0}$ such that $\int_0^\infty (1+t)d\mu(t) < \infty$, there exist two numbers A_μ and $B_\mu > 0$ such that*

$$\mathbb{Q}_x^\delta[\exp(-\int_0^\infty X_t d\mu(t))] = A_\mu^x B_\mu^\delta$$

where X is the coordinate process.

Proof. We call the left-hand side $\phi(x, \delta)$. The condition on μ entails that

$$\phi(x, \delta) \geq \exp(-\mathbb{Q}_x^\delta(\int_0^\infty X_t d\mu(t))) = \exp(-(\int_0^\infty (x + \delta t d\mu(t))) > 0.$$

Moreover from the theorem we have $\phi(x+x', \delta+\delta') = \phi(x, \delta)\phi(x', \delta')$, so that $\phi(x, \delta) = \phi(x, 0)\phi(0, \delta)$. These functions are multiplicative and equal to 1 at 0. Furthermore they are monotone and thus measurable. $\square$

We now want to compute the constants A_μ and B_μ explicitly.

Computation of the Constants

If μ is a Radon measure on $[0, \infty)$ then the differential equation (in the distribution sense) $\phi'' = \phi\mu$ has a unique solution ϕ_μ which is positive, non increasing on $[0, \infty)$ and such that $\phi_\mu(0) = 1$. The function ϕ_μ is convex so the right-hand side derivative exists and is ≤ 0. Since ϕ_μ is non increasing the limit $\phi_\mu(\infty) := \lim_{x\to\infty} \phi_\mu(x)$ exists and belongs to $[0, 1]$. We can assume $\phi_\mu(\infty) < 1$ because otherwise ϕ_μ is identical 1 and $\mu = 0$. We still assume μ to satisfy $\int_0^\infty (1 + x)d\mu(x) < \infty$ which leads to $\phi_\mu(\infty) > 0$ (in the proof found in [15]). We define $X_\mu = \int_0^\infty X_t d\mu(t)$.

Theorem 4.4.5. *Under the preceding assumptions we have that*

$$\mathbb{Q}_x^\delta[\exp(-\frac{1}{2}X_\mu)] = \phi_\mu(\infty)^{\delta/2}\exp(\frac{x}{2}\phi'_\mu(0)).$$

To illustrate this result we will prove the following corollary:

Corollary 4.4.3.

$$\mathbb{Q}_x^\delta[\exp(-\frac{b^2}{2}\int_0^1 X_s ds)] = (\cosh b)^{-\delta/2}\exp(-\frac{1}{2}xb\tanh b).$$

Proof. We must compute ϕ_μ for $\mu(ds) = b^2 ds$ on $[0,1]$. Thus on $[0,1]$ we must have $\phi_\mu(t) = \alpha\cosh bt + \beta\sinh bt$ and $\phi_\mu(0) = 1$ forces $\alpha = 1$. Since ϕ_μ is constant on $[1,\infty)$ and ϕ'_μ is continuous we need $\phi'_\mu(1) = 0$ which is $b\sinh b + \beta b\cosh b = 0$ and thus $\beta = -\tanh b$. Finally we have $\phi_\mu(t) = \cosh bt - \tanh b \sinh bt$ on $[0,1]$. Then $\phi_\mu(\infty) = \phi_\mu(1) = (\cosh b)^{-1}$ and $\phi'_\mu(0) = -b\tanh b$. □

4.4.6 The Connection between Bessel processes and the CIR model

If we have a square root process r_t of the form

$$dr_t = a(b - r_t)dt + c\sqrt{r_t}dB_t, \quad r_0 \in \mathbb{R}$$

and a BESQ$^\delta$ X of the form

$$X_t = x + \delta t + 2\int_0^t \sqrt{X_s}dB_s.$$

We now consider the process at the point in time $\alpha(t) = \frac{c^2}{4a}(e^{at} - 1)$:

$$X_{\alpha(t)} = x + \delta\alpha(t) + 2\int_0^{\alpha(t)} \sqrt{X_s}dB_s.$$

By theorem 4.4.3 we see the following ($\sqrt{\alpha(r)'} = \frac{c}{2}e^{ar/2}$):

$$X_{\alpha(t)} = x + \delta\alpha(t) + \int_0^t \sqrt{X_{\alpha_r}}2\frac{c}{2}e^{ar/2}d\tilde{B}_r$$
$$X_{\alpha(t)} = x + \delta\alpha(t) + \int_0^t c\sqrt{X_{\alpha_r}e^{ar}}d\tilde{B}_r.$$

The Brownian motion $\tilde{B}$ is defined by $\tilde{B}_t = \int_0^{\alpha(t)} \frac{2}{c}e^{-as/2}dB_s$. If we now set $r_t = e^{-at}X_{\alpha(t)}$ then we get by Itô's formula:

$$dr_t = d(e^{-at}X_{\alpha(t)}) = -ae^{-at}X_{\alpha(t)}dt + e^{-at}dX_{\alpha(t)}$$
$$dr_t = (-ar_t + e^{-at}\delta\alpha(t)')dt + e^{-at}c\sqrt{X_{\alpha(t)}}e^{at/2}d\tilde{B}_t$$
$$dr_t = (-ar_t + e^{-at}\frac{4ab}{c^2}\frac{c^2}{4}e^{at})dt + c\sqrt{X_{\alpha(t)}}e^{-at/2}d\tilde{B}_t$$
$$dr_t = a(b - r_t)dt + c\sqrt{r_t}d\tilde{B}_t.$$

By the existence of a weak solution of the SDE for a BESQ$^\delta$ X we get the weak existence of the SDE for $e^{-at}X_{\alpha(t)} =: r_t$. Thus we get the existence of the solution of the square root process by the Bessel process. Uniqueness is shown by the theorem of Yamada and Watanabe 1.5.1.

4.4.7 The Carleman-Fredholm determinant

David Nualart [12] gives a short description of the Carleman-Fredholm determinant and how it can be calculated. We have a measure space $(T, \mathcal{B}, \mu)$ and $K \in L^2(T \times T)$. Moreover we assume that the Hilbert space $H = L^2(T)$ is separable and $\{e_i, i \geq 1\}$ to be a complete orthonormal system in H. In the case $K = \sum_{i,j=1}^n a_{ij} e_i \otimes e_j$ the Carelman-Fedholm determinant of $I + K$ is defined as

$$det_2(I + K) = det(I + A) \exp(-Tr(A)),$$

where $A = (a_{ij})_{1 \leq i,j \leq n}$. It can be proved that the mapping $K \rightarrow det_2(I + K)$ is continuous in $L^2(T \times T)$. Consequently it can be extended to the whole space $L^2(T \times T)$. A useful formula to compute $det_2(I + K)$, where $K \in L^2(T \times T)$, is the following:

$$det_2(I + K) = 1 + \sum_2^\infty \frac{\gamma_n}{n!},$$

where $\gamma_n = \int_{T^n} det(\hat{K}(t_i, t_j))\mu(dt_1) \dots \mu(dt_n)$. Here $\hat{K}(t_i, t_j) = K(t_i, t_j)$, if $i \neq j$ and $K(t_i, t_i) = 0$.

Bibliography

[1] Leif Andersen and Jesper Andreasen, *Volatility Skews and Extensions of the Libor Market Model*, Applied Mathematical Finance, 7, 1-32 (2000)

[2] Leif Andersen and Rupert Brotherton-Ratcliffe, *Extended LIBOR Market Models with Stochastic Volatility*, Paper (December 2001)

[3] Ludwig Arnold, *Stochastic Differential Equations: Theory and Applications*, John Wiley & Sons (1974)

[4] Peter Grandits, *Advanced Mathematics of Finance 2*, lecture at the Technical University of Vienna, check for notes at http://www.fam.tuwien.ac.at/

[5] Matheus R. Grasselli and Tom R. Hurd, *The Wiener Chaos Expansion for the Cox-Ingersoll-Ross Model*, Paper (2003)

[6] Farshid Jamishidian, *LIBOR and swap market models and measures*, Applied Mathematical Finance, 3,93-115 (1997)

[7] Robert Jarrow, Haitao Li, and Feng Zhao, *Interest Rate Caps "Smile Too"! But Can the LIBOR Market Models Capture It?*, Paper (January 2003)

[8] Ioannis Karatzas, Steven E.Shreve, *Brownian Motion and Stochastic Calculus*, Springer (1988, 1991)

[9] Masaaki Kijima, *Stochastic Processes with Applications to Finance*, Chapman & Hall/CRC (2003)

[10] Damien Lamberton, Bernard Lapeyre, *Introduction to Stochastic Calculus Applied to Finance*, Chapman and Hall (1996) - originally in French *Introdution au calcul stochastique appliqué à la finance*, Édition Marketing S.A.(1991), published by Ellipses

[11] Marek Musiela, Marek Rutkowski, *Martingale Methods in Financial Modelling*, Springer (corrected second printing 1998)

[12] David Nualart, *The Malliavin Calculus and Related Topics*, Springer (1995)

[13] Bernt Øksendal, *Stochastic Differential Equations*, Springer (1995)

[14] Bernt Øksendal, *An Introduction to Malliavin Calculus with Applications to Economics*, paper May 1997

[15] Daniel Revuz, Marc Yor, *Continuous Martingales and Brownian Motion*, Springer (1991)

[16] Josef Teichmann, *Stochastic Analysis*, lecture at the Technical University of Vienna, check for notes at http://www.fam.tuwien.ac.at/~jteichma/

[17] David Williams, *Probability with Martingales*, Cambridge Mathematical Textbooks (1991)

Index

Arbitrage free, 42

Banach's Fixed Point Theorem, 18
Bessel process, 95
Bond option
- explicit formula, 70
Borel-Cantelli Lemma, 19
- proof of the second, 83
- proof of the first, 82
Brownian Motion, 5
- quadratic variation, 5
- in d Dimensions, 5
- recurrence, 34
- transience, 34

Cap, 68
- caption, 69
Covariation process, 16

Doob's Martingale Inequality, 18
- proof, 83
Dynkin's Formula, 32

Floor, 68
Forward
- swap rate
  - definition, 63
- martingale measure, 50
- measure, 49
- process, 48
- rate, 41
- swap rate
  - market model, 64

Girsanov's Theorem, 36
Gronwall Lemma, 18
- proof, 82

Implied savings account, 51
Infinitesimal generator, 29

LIBOR rate, 47
- continuous tenor model, 58
- definition, 54
- discrete tenor model, 55

Markov Property, 34
- strong Markov Property, 36
Martingale
- martingale problem, 91
- Brownian martingale, 37
- definition, 4
- local martingale, 4
- martingale representation theorem, 37
- semi-martingale, 15

Novikov's condition, 36

Picard-Lindelöff, 20, 28

Quadratic Variation, 5

Random time change, 93

SDE
- existence and uniqueness, 19
- global solution, 27
- local solution, 27
- solution in the weak sense, 17
- strong solution, 17
Short Rate, 42
Spot martingale measure, 50
Swap, 61
- swaption, 69
Swap Rate

- definition, 63

Tchebychev's Inequality, 19
Time homogeneity, 34

Uniqueness
- pathwise uniqueness, 18
- uniqueness in law, 18

Volatility
- bond price volatility, 53
- implied, 73
- implied skew volatility, 77
- smile, 73
- stochastic, 74

Weak no-arbitrage condition, 50
Wiener Process, 3

Yield-to-maturity, 41

Zero-coupon bond, 40