UNIVERSITÄT DER BUNDESWEHR MÜNCHEN
Fakultät für Elektrotechnik und Informationstechnik

Multi-dimensional Markov-functional models in option pricing

Thomas Seifert

Vorsitzender des Promotionsausschusses:	Prof. Dr.-Ing. D. Gerling
1. Berichterstatter:	Prof. Dr.-Ing. Dr.rer.nat. S. Schäffler
2. Berichterstatter:	Prof. Dr.rer.nat. A. Sachs

Tag der Prüfung: 6.5.2004

Mit der Promotion erlangter akademischer Grad:
Doktor-Ingenieur
(Dr.-Ing.)

Neubiberg, den 17. Mai 2004

UNIVERSITÄT DER BUNDESWEHR MÜNCHEN
Fakultät für Elektrotechnik und Informationstechnik

Multi-dimensional Markov-functional models in option pricing

Thomas Seifert

Vorsitzender des Promotionsausschusses:	Prof. Dr.-Ing. D. Gerling
1. Berichterstatter:	Prof. Dr.-Ing. Dr.rer.nat. S. Schäffler
2. Berichterstatter:	Prof. Dr.rer.nat. A. Sachs

Tag der Prüfung: 6.5.2004

Mit der Promotion erlangter akademischer Grad:
Doktor-Ingenieur
(Dr.-Ing.)

Neubiberg, den 17. Mai 2004

The theory of martingales began life with the aim of providing insight into the apparent impossibility of making money by placing bets on a fair game.

J. Michael Steele

Introduction

30 years ago, a revolution took place at Wall Street. Until then European options, financial derivatives that allow one party to buy an asset from another party at a future date and at a prespecified price, have been known and traded, but the pricing depended heavily on the market opinion of each trader. In 1973, however, Fisher Black and Myron Scholes published their famous formula yielding an exact price for each European option. This formula has been extended by Robert Merton in a number of significant ways. The ideas behind their formula are fundamental until today and represent the first model of the evolution of interest rates.
A lot of progress has been made since then. The next step in the evolution of interest rate models where the so called short-rate models. They were based on the assumption that the prices of assets traded in the market depend solely on a single stochastic process, the so called short-rate $\{r_t; t \in [0,T]\}$ for some $T > 0$. Precisely, if D_{tT} denotes time-t value of the zero coupon bond maturing at time T then

$$D_{tT} = \exp\Big(-\int_t^T r_u du\Big).$$

The short-rate itself is usually modeled by some stochastic differential equation. Short-rate models are easy to understand and easy to implement. but calibration of these models requires some precaution. Further, evaluation is cumbersome, usually requiring a Monte Carlo simulation. Another disadvantage from a modeling point of view is that the short-rate actually cannot be observed in the market. Things changed radically in 1997 when Brace, Gatarek, Musiela and later Jamshidian introduced Market models. Instead of modeling the whole term structure by a single short-rate their model considered only a discrete subset of rates of the term structure with each rate given by a single stochastic differential equation:

$$dL_t^{(i)} = \sigma_t^{(i)} dW_t^{(i)}, \mathcal{S}^{(i)}, t \in [0, T_i],$$

where $\{L_t^{(i)}; t \in [0,T_i]\}$ denotes the i-th interest rate maturing on T_i under some probability measure $\mathcal{S}^{(i)}$. The corresponding zero coupon bond values may be easily derived thereof. Each stochastic differential equation is driven by a one-dimensional Wiener process $\{W_t^{(i)}; t \in [0,T_i]\}$. These models are easy to calibrate since the currently prevailing interest rate is part of the model input, contrary to short-rate models. However, Market models may also only be evaluated by means of Monte Carlo simulation, but the approach of modeling single interest rates that can be observed in the market is a clear advantage over short-rate models.

Interest rate models usually make the assumption of a log-normal distribution of rates as Black and Scholes did. Unfortunately at the latest since the Asia crisis in 1997 this is not the case anymore as statistical tests show. The distribution of interest rates observed in the market have fatter tails compared with the log-normal distribution. Consequently extreme events are more likely to happen as predicted by our models. Closing this gap between modeling and reality is one of the greatest challenges in financial mathematics nowadays. One successful attempt has been brought about lately by Hunt and Kennedy. They introduced a general class of interest rate models known as Markov-functional models that capture the implied distribution of interest rates. Further, they allow for a fast numerical implementation using numerical integration. The class of models proposed by Hunt and Kennedy allows only for one source of randomness in the market. This is sufficient for simple products but for more complex derivatives it is desirable to more sources of randomness. An extreme example would be to model n different rates with n different, not necessarily independent, sources of randomness. This work proposes an extension of the Markov-functional models allowing exactly for this.

Contents

Chapter 1

Preliminaries

Before introducing the multi-dimensional Markov-functional model we start with some notation and a short introduction to probability theory. Our main focus is on martingales, Markov processes and stochastic differential equations. These are the ingredients to construct the Markov-functional model.

1.1 Notation

We start by introducing some notation that will be relevant throughout this work. In this section let $m, n, k \in \mathbb{N}$.

Notation 1.1 ($\mathbb{R}^+, \overline{\mathbb{R}}, \mathbb{N}$). Let $\mathbb{R}$ denote the set of real numbers, let $\mathbb{R}^+$ denote the set of positive real numbers,

$$\mathbb{R}^+ := \{x \in \mathbb{R}; x > 0\}$$

and $\mathbb{R}_0^+ := \mathbb{R}^+ \cup \{0\}$. Let $\overline{\mathbb{R}}$ denote the set $\mathbb{R} \cup \{\pm\infty\}$, similarly $\overline{\mathbb{R}_0^+} := \mathbb{R}_0^+ \cup \{+\infty\}$.

$$\mathbb{N} := \{1, 2, 3, \ldots\}$$

denotes the set of natural numbers without zero and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.

Notation 1.2 (Matrices, vectors). Let $A \in \mathbb{R}^{n \times m}$ and $i, j \in \mathbb{N}$, where $1 \leq i \leq n, 1 \leq j \leq m$. Then $A_{i\cdot}$ denotes the i-th row of A, $A_{\cdot i}$ the i-th column and A_{ij} the entry on position (i, j).
$\mathbb{E}_n$ denotes the n-dimensional identity matrix.
Let $v \in \mathbb{R}^n$. Then $v^{(i)}$ denotes the i-th entry of v. For $v \in \mathbb{R}^n$, $\operatorname{diag}(v^{(1)}, \ldots, v^{(n)})$ denotes the matrix having the entries $v^{(1)}, \ldots, v^{(n)}$ on the diagonal and 0 elsewhere, i.e.,

$$\operatorname{diag}(v^{(1)}, \ldots, v^{(n)}) = \begin{pmatrix} v^{(1)} & & 0 \\ & \ddots & \\ 0 & & v^{(n)} \end{pmatrix}.$$

Remark 1.3. We will identify $\mathbb{R}^n$ with $\mathbb{R}^{n\times 1}$, thus the elements of $\mathbb{R}^n$ are column vectors.

Notation 1.4 (Power set). Let A be an arbitrary set. $\mathfrak{P}(A)$ denotes the power set of A, i.e., the set of all subsets of A.

Notation 1.5 (Indicator function). We denote the indicator function of any set A by $\mathbf{1}_{\{A\}}$, i.e.,

$$\mathbf{1}_{\{A\}}(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A. \end{cases}$$

Notation 1.6 (Norms).

1. $\|\cdot\|_2$ denotes the Euclidean norm on $\mathbb{R}^n$,

$$\|x\|_2 := \sqrt{\sum_{i=1}^{n} (x^{(i)})^2}.$$

2. $\|\cdot\|_F$ denotes the Frobenius norm on $\mathbb{R}^{n\times m}$,

$$\|A\|_F := \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{m} A_{ij}^2}.$$

Remark 1.7. The topology on $\mathbb{R}^n$ is the topology induced by the Euclidean norm.

Notation 1.8 (Neighborhood of $x_0 \in \mathbb{R}^n$, open set). For $\epsilon > 0$ and $x_0 \in \mathbb{R}^n$, we define the *ϵ-neighborhood of x_0* by $U_\epsilon(x_0) := \{x \in \mathbb{R}^n; \|x - x_0\|_2 < \epsilon\}$. A set $U \subset \mathbb{R}^n$ is called *open*, if for all $x \in U$ there exists an $\epsilon > 0$ such that $U_\epsilon(x) \subset U$.

After defining a topology on $\mathbb{R}^n$ we may also define spaces of continuous functions on $\mathbb{R}^n$.

Notation 1.9 (Spaces of continuous functions). Let U be an subset of $\mathbb{R}^k$ and V be an subset of $\mathbb{R}^n$. We define

1. $C(U, V) := \{f : U \to V; \text{f is continuous}\}$,
2. $C(U) := C(U, \mathbb{R})$.

If U and V are open, then we set for $q \in \mathbb{N}$

3. $C^q(U, V) := \{f : U \to V; \text{f is } q\text{-times continuous differentiable}\}$,
4. $C^q(U) := C^q(U, \mathbb{R})$.

Notation 1.10 (Derivative). Let $U \subset \mathbb{R}^n$ be open and $f \in C^1(U, \mathbb{R}^m)$. We denote the derivative of f in x by

$$Df(x) := \begin{pmatrix} \frac{\partial f^{(1)}}{\partial x^{(1)}}(x) & \cdots & \frac{\partial f^{(m)}}{\partial x^{(1)}}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial f^{(1)}}{\partial x^{(n)}}(x) & \cdots & \frac{\partial f^{(m)}}{\partial x^{(n)}}(x) \end{pmatrix}.$$

If a function and its inverse are differentiable, we call it a diffeomorphism:

Definition 1.11 (Diffeomorphism). *Let $U, V \subset \mathbb{R}$ be open subsets and $f \in C^n(U, V), n \in \mathbb{N}$. If the inverse f^{-1} exists, f is called* diffeomorphism.

It is known from real analysis, that under the assumptions of the last definition, f^{-1} is also differentiable and $f^{-1} \in C^n(V, U)$.

The notion of monotonic increasing functions will be very important throughout this work.

Definition 1.12 (Strictly monotonic increasing function). *Let $x, y \in \mathbb{R}^n$. $x < y$ is defined component wise,*

$$x < y :\Leftrightarrow \forall\, 1 \leq i \leq n : x^{(i)} \leq y^{(i)} \wedge \exists\, i \in \{1, \ldots, n\} : x^{(i)} < y^{(i)}.$$

A function $f : \mathbb{R} \to \mathbb{R}$ is strictly monotonic increasing, *if*

$$x < y \Rightarrow f(x) < f(y) \text{ for all } x, y \in U.$$

A function $f : \mathbb{R}^n \to \mathbb{R}$ is strictly monotonic increasing, *if for all $x < y, x, y \in \mathbb{R}^n$,*

$$f(x^{(1)}, \ldots, x^{(i)}, \ldots, x^{(n)}) < f(x^{(1)}, \ldots, y^{(i)}, \ldots, x^{(n)})$$

holds for all $1 \leq i \leq n$ such that $x^{(i)} < y^{(i)}$

Due to the importance of the definition above we give a slightly different interpretation. Choose some $y \in \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$. f is strictly monotonic increasing in the sense of this definition if the projections $f_i : \mathbb{R} \to \mathbb{R}$, $x \mapsto f(y^{(1)}, \ldots, y^{(i-1)}, x, y^{(i+1)}, y^{(n)})$ are strictly monotonic increasing scalar functions in x for every $y \in \mathbb{R}^n$. We thus may reduce the question, whether a function defined on $\mathbb{R}^n$ is strictly monotonic increasing or not, to the usual one-dimensional case by examining the function componentwise for every such $y \in \mathbb{R}^n$. Other generalizations to the multi-dimensional case are possible, but the one above will be the most convenient for our purposes.

1.2 Probability theory

Most important in the theory of derivative pricing is the modeling of the uncertainty in an economy. This is accomplished by the means of probability theory which is therefore a cornerstone of this work. In this chapter a short introduction to probability theory and especially stochastic integration and stochastic differential equations with respect to the Itô-Integral is given. An introduction to interest rate derivatives based on the definitions made here will be given in the next chapter. A more detailed and thorough course in probability theory and stochastic integration is provided by [Schäffler, Sturm 1994], [Schäffler, Sturm 1995], [Schäffler 1996], [Bauer 1991], [Karatzas, Shreeve 2001] and [Protter 1990]. All proofs for the theorems in this section may be found in [Bauer 1991] and [Bauer 1992] unless stated otherwise.

1.2.1 Preliminaries

Notation 1.13 (Borel-σ-field). Let $\mathcal{B}^n$ denote the Borel-σ-field in $\mathbb{R}^n$, $n \in \mathbb{N}$, i.e., the σ-field generated by the open subsets in $\mathbb{R}^n$. $\overline{\mathcal{B}} := \{A \in \mathfrak{P}(\overline{\mathbb{R}}); A \cap \mathbb{R} \in \mathcal{B}\}$ denotes the extended σ-field on $\overline{\mathbb{R}}$. Further, let λ^n denote the Lebesgue-measure of dimension n.

In the remainder of this section we will work with a special probability space, the *Wiener space*. This is the space Ω^n of all continuous functions $F : \mathbb{R}_0^+ \to \mathbb{R}^n$ equipped with the σ-field $\mathcal{B}(\Omega^n)$. $\mathcal{B}(\Omega^n)$ is the smallest σ-field $\mathcal{S}$ over Ω^n such that the maps $p_t : \Omega^n \to \mathbb{R}^n, F \mapsto F(t), t \in \mathbb{R}_0^+$, are $\mathcal{S}$-$\mathcal{B}^n$-measurable. To define a measure on this measurable space, the *Wiener measure*, we require some more definitions.

Definition 1.14 ($\mathcal{P}$-almost surely). *Let $(\Omega, \mathcal{S}, \mathcal{P})$ be a probability space. A property is said to hold* $\mathcal{P}$-almost surely, *if a set $A \in \mathcal{S}$ of measure 0 exists such that the property holds on the complement of A.*

Definition 1.15 (Extended function). *A function $f : A \to \overline{\mathbb{R}}$ defined on a non-empty set $A \subseteq \Omega^n$ is called* extended function.

Definition 1.16 (Random variable). *A function $X : \Omega^n \to \mathbb{R}^n$ is called* (n-dimensional) random variable, *if it is $\mathcal{B}(\Omega^n)$-$\mathcal{B}^n$-measurable. Given an $\mathbb{R}^n$-valued random variable X, we denote by $\sigma(X)$ the smallest σ-field $\mathcal{A}$ such that X is $\mathcal{A}$-$\mathcal{B}^n$-measurable.*

Central to this work is the concept of stochastic processes.

Definition 1.17 (Stochastic process). *Let $X_t : \Omega^n \to \mathbb{R}^n$ be an n-dimensional random variable for all $t \geq 0$. The set $\{X_t;\ t \geq 0\}$ is called* stochastic process *or simply* process, *if the mapping*

$$(t, \omega) \mapsto X_t(\omega) : \mathbb{R}_0^+ \times \Omega^n \to \mathbb{R}^n$$

is $\mathcal{B}(\mathbb{R}_0^+) \otimes \mathcal{B}(\Omega^n)$-$\mathcal{B}^n$*-measurable.*

We will abbreviate a stochastic process $\{X_t;\, t \geq 0\}$ by simply writing X if the choice of t is clear from the context.

When freezing an $\omega \in \Omega^n$, one obtains the path of a stochastic process.

Definition 1.18 ((Sample) Path of a stochastic process). *Let* $\{X_t;\, t \geq 0\}$ *be a stochastic process. The map* $X^\omega : \mathbb{R}_0^+ \to \mathbb{R}^n, t \mapsto X_t(\omega), \omega \in \Omega^n$, *is called* (sample) path of the stochastic process $\{X_t; t \geq 0\}$.

Definition 1.19 (Continuous stochastic process). *If every path* X^ω *is a continuous function a stochastic process* $\{X_t;\, t \geq 0\}$ *is* continuous.

If all information about the past is contained in the current σ-field, we get an increasing sequence of σ-fields, a filtration.

Definition 1.20 (Filtration). *Let* $(\Omega,\, \mathcal{S},\, \mathcal{P})$ *be a probability space,* $U \subseteq \mathbb{R}$ *and* $\{\mathcal{F}_s; s \in U\}$ *a set of* σ*-fields in* Ω*, where* $\mathcal{F}_s \subseteq \mathcal{S}$ *for all* $s \in U$*. Further, let* $\mathcal{F}_s \subseteq \mathcal{F}_r$ *for all* $s, r \in U$ *satisfying* $s \leq r$*. Then* $\{\mathcal{F}_s; s \in U\}$ *is a* filtration in $\mathcal{S}$.

Definition 1.21 (Adapted stochastic process). *Let* $(\Omega,\, \mathcal{S},\, \mathcal{P})$ *be a probability space,* $U \subseteq \mathbb{R}$*, and* $\{X_s; s \in U\}$ *a stochastic process and* $\{\mathcal{F}_s; s \in U\}$ *a filtration in* $\mathcal{S}$. $\{X_s; s \in U\}$ *is* adapted to the filtration $\{\mathcal{F}_s; s \in U\}$ *if* X_s *is* $\mathcal{F}_s$-$\mathcal{B}^n$*-measurable for each* $s \in U$.

Given a stochastic process $\{X_s; s \in U\}$, $U \subseteq \mathbb{R}$, the simplest filtration is that generated by the process itself, i.e.,

$$\mathcal{F}_t^X := \sigma(X_s; s \in U, s \leq t).$$

This is the smallest σ-field with respect to which X_s is measurable for every $s \in U$.

Definition 1.22 (The usual conditions). *Let* $(\Omega,\, \mathcal{S},\, \mathcal{P})$ *be a probability space and* $\{\mathcal{F}_t; t \geq 0\}$ *a filtration in* Ω. $\{\mathcal{F}_t; t \geq 0\}$ *is* right-continuous *if* $\mathcal{F}_t = \bigcap_{\epsilon > 0} \mathcal{F}_{t+\epsilon}$ *for every* $t \geq 0$*. If* $\{\mathcal{F}_t; t \geq 0\}$ *is right-continuous and* $\mathcal{F}_0$ *contains all* $\mathcal{P}$*-negligible events* $\{\mathcal{F}_t; t \geq 0\}$ *is said to satisfy* the usual conditions.

It is easy to see that any given filtration $\{\mathcal{F}_t; t \geq 0\}$ can be augmented such that the augmented filtration $\{\tilde{\mathcal{F}}_t; t \geq 0\}$ satisfies the usual conditions and $\mathcal{F}_t \subseteq \tilde{\mathcal{F}}_t$ for all $t \geq 0$. Throughout this work we assume that we always work with a filtration satisfying the usual conditions. If a filtration does not satisfy the usual conditions, we work with the corresponding augmented filtration.

Another important point is the independency of random variables. We will introduce this notion first for events and after that for random variables.

Definition 1.23 (Independent events). *Let $(\Omega, \mathcal{S}, \mathcal{P})$ be a probability space and $A_1, \ldots, A_n \in \mathcal{S}, n \in \mathbb{N}$. The events $A_1, \ldots, A_n$ are* independent*, if for each $k \in \mathbb{N}, k \leq n$, and for each $i_j \in \mathbb{N}, 1 \leq j \leq k$, where $1 \leq i_1 < \ldots < i_k \leq n$,*

$$\mathcal{P}\left(\bigcap_{j=1}^{k} A_{i_j}\right) = \prod_{j=1}^{k} \mathcal{P}(A_{i_j})$$

holds.

We derive now the definition of independency of a set of events.

Definition 1.24 (Independency of a set of events). *Let $(\Omega, \mathcal{S}, \mathcal{P})$ be a probability space and $\{A_i \in \mathcal{S}; i \in I\}, I \neq \emptyset$, a set of events. These events are* independent*, if $A_{i_1}, \ldots, A_{i_n}$ are independent for each $n \in \mathbb{N}$, with $n \leq |I|$ and for each set $\{i_1, \ldots, i_n\} \subseteq I$.*

Definition 1.25 (Independency of a set of subsets). *Let $(\Omega, \mathcal{S}, \mathcal{P})$ be a probability space and $\{\mathcal{F}_i; i \in I\}, I \neq \emptyset, \mathcal{F}_i \subseteq \mathcal{S}$, a set of subsets of Ω. These sets are* independent*, if for arbitrary $A_{i_k} \in \mathcal{F}_{i_k}, k = 1, \ldots, n$, the events $A_{i_1}, \ldots, A_{i_n}$ are independent for each $n \in \mathbb{N}$, with $n \leq |I|$ and for each set $\{i_1, \ldots, i_n\} \subseteq I$.*

Definition 1.26 (Independency of random variables). *Let $(\Omega, \mathcal{S}, \mathcal{P})$ be a probability space and $(\Omega', \mathcal{S}')$ a measurable space. The random variables $X_i : \Omega \to \Omega', i \in I$, are* independent*, if the sets $\{\sigma(X_i); i \in I\}$ are independent.*

We define now the distribution and expected value of random variables on a probability space.

Definition 1.27 (Distribution). *Let $(\Omega, \mathcal{S}, \mathcal{P})$ be a probability space and $X : \Omega^n \to \mathbb{R}^n$ a random variable thereon. The measure μ on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \lambda^n)$ defined by*

$$\mu(A) := \mathcal{P}(\{\omega \in \Omega^n; X(\omega) \in A\})$$

for all $A \in \mathcal{B}(\mathbb{R}^n)$ is called the distribution of X.

Definition 1.28 (Expected value). *Let $X : \Omega^n \to \mathbb{R}^n$ be a random variable. Then the* expected value *of X is defined via the integral*

$$E_{\mathcal{P}}[X] := \int_{\Omega^n} X d\mathcal{P},$$

if the above integral exists.

An analogous definition as above holds for stochastic processes.

Definition 1.29 (Mean value, covariance function). *Let $\{X_t; t \geq 0\}$ be a stochastic process on some probability space $(\Omega,\ \mathcal{S},\ \mathcal{P})$ and suppose $E_{\mathcal{P}}[X_0^2] < \infty$. We introduce the* mean vector *and* covariance matrix *functions*

$$\begin{aligned} m(t) &:= E_{\mathcal{P}}[X_t], \\ cov(s,t) &:= E_{\mathcal{P}}[(X_s - m(s))(X_t - m(t))^T], \\ V(t) &:= cov(t,t), \end{aligned}$$

for some $s, t \geq 0$, if the above integrals exist.

The notion of probability density functions is a very useful one.

Definition 1.30 (Density function). *Let $(\Omega, \mathcal{S})$ be a measurable space and $\mathcal{P}$ and μ measures on $\mathcal{S}$. If a non-negative , $\mathcal{S}$-$\overline{\mathcal{B}}$-measurable function $f : \Omega \to \overline{\mathbb{R}}$ exists, satisfying*

$$\mathcal{P}(A) = \int_A f d\mu \text{ for all } A \in \mathcal{S},$$

it is called the density function *of the probability measure $\mathcal{P}$ with respect to the measure μ.*

An important example is the density of the normal distribution.

Definition 1.31 (Normal distribution). *Let $(\Omega,\ \mathcal{S},\ \mathcal{P})$ be a probability space and $X : \Omega \to \mathbb{R}^n$, $n \in \mathbb{N}$. X is* normally distributed *with mean 0 and covariance matrix $t\mathbb{E}_n$ if X has the density function*

$$\phi(x) = (2\pi t)^{-\frac{n}{2}} \exp\Big(-\frac{\|x\|_2^2}{2t}\Big)$$

with respect to the n-dimensional Lebesgue measure. We write $X \sim \mathbf{N}(0, t\mathbb{E}_n)$. X is said to be log-normally distributed, *if $\exp(X)$ is normal distributed. Further, we denote the cumulative distribution function of the normal distribution by $\Phi(x)$.*

Lemma 1.32. *Let $X : \Omega \to \mathbb{R}$ be a $\mathbf{N}(m, \sigma^2)$-distributed random variable on our usual probability space $(\Omega,\ \mathcal{S},\ \mathcal{P})$. The expected value of $\exp(X)$ is*

$$E_{\mathcal{P}}[\exp(X)] = \exp\Big(m + \frac{\sigma^2}{2}\Big).$$

Proof. See [Hoffmann-Jørgensen 1994], Section 4.25. □

To verify whether a density function exists for a given measure, we need the following two definitions.

Definition 1.33 (Absolutely continuous measure). *Let $(\Omega, \mathcal{S})$ be a measurable space and $\mathcal{Q}$ a measure on $\mathcal{S}$. $\mathcal{P}$ is* absolutely continuous with respect to $\mathcal{Q}$, *if for each $A \in \mathcal{S}$*

$$\mathcal{Q}(A) = 0 \Rightarrow \mathcal{P}(A) = 0$$

holds. The two measures are equivalent, *if $\mathcal{P}$ is absolutely continuous with respect to $\mathcal{Q}$ and vice versa. The equivalence of two measures $\mathcal{Q}$ and $\mathcal{P}$ will be denoted by $\mathcal{Q} \sim \mathcal{P}$.*

Definition 1.34 (σ-finite measure). *Let $(\Omega, \mathcal{S})$ be a measurable space. A measure $\mathcal{Q}$ on $\mathcal{S}$ is* σ-finite *if a sequence $\{A_i\}_{i\in\mathbb{N}}, A_i \in \mathcal{S}$, of subsets of Ω exists such that*

$$\bigcup_{i=1}^{\infty} A_i = \Omega, \; \mathcal{Q}(A_i) < \infty \textit{ for all } i \in \mathbb{N}.$$

The following theorem gives us the conditions under which a density function for a measure $\mathcal{P}$ with respect to a measure $\mathcal{Q}$ exists.

Theorem 1.35 (Radon-Nikodym). *Let $(\Omega, \mathcal{S}, \mathcal{P})$ be a probability space and $\mathcal{Q}$ a σ-finite measure on $\mathcal{S}$. $\mathcal{P}$ has a probability density f with respect to $\mathcal{Q}$ iff $\mathcal{P}$ is absolutely continuous with respect to $\mathcal{Q}$. In this case we use the notation $\frac{d\mathcal{P}}{d\mathcal{Q}} := f$. $\frac{d\mathcal{P}}{d\mathcal{Q}}$ is called* Radon-Nikodymderivative *of $\mathcal{P}$ with respect to $\mathcal{Q}$.*

Throughout this work we need only probability densities with respect to the Lebesgue measure λ^n. If the probability distribution of a random variable X has a probability density f with respect to the Lebesgue measure, we say that X has the density function f.

Based on the above theorem we define now the conditional expection of a random variable.

Definition 1.36 (Conditional expectation). *Let $(\Omega, \mathcal{S}, \mathcal{P})$ be a probability space and $X : \Omega \to \mathbb{R}^n$ a random variable such that $E_{\mathcal{P}}[\|X\|_2] < \infty$. If $\mathcal{H} \subset \mathcal{S}$ is a sub-σ-field of $\mathcal{S}$, then the ($\mathcal{P}$-a.s. unique)* conditional expectation of X given $\mathcal{H}$, *denoted by $E_{\mathcal{P}}[X|\mathcal{H}]$, is defined by the following properties:*

1. *$E_{\mathcal{P}}[X|\mathcal{H}]$ is $\mathcal{H}$-measurable.*
2. *$\int_H E_{\mathcal{P}}[X|\mathcal{H}] d\mathcal{P} = \int_H X d\mathcal{P}$ for all $H \in \mathcal{H}$.*

The existence and uniqueness of the conditional expectation can be proved using the Radon-Nikodym theorem. A proof can be found in [Øksendal 1998].

Random times and stopping times have been brought up to 'tame' continuous time processes and especially martingales.

Definition 1.37 (Random time, stopping time). *Let* $(\Omega, \mathcal{S}, \mathcal{P})$ *be a probability space equipped with the filtration* $\{\mathcal{F}_t; t \geq 0\}$. *An extended function* $\tau : \Omega \to \overline{\mathbb{R}}$ *is a* random time, *if* τ *is* $\mathcal{S}$-$\overline{\mathcal{B}}$*-measurable. A* stopping time *is a random time* τ*, such that* $\{\tau \leq t\} \in \mathcal{F}_t$ *for all* $t \geq 0$.

We now define an important stochastic process, the Wiener process or Brownian motion. First we define the Wiener process on a general probability space, then we show the existence of a measure, the Wiener measure, on $(\Omega^n, \mathcal{B}(\Omega^n))$ such that a Wiener process exists.

Definition 1.38 (Wiener process on $(\Omega, \mathcal{S}, \mathcal{P})$**).** *Let* $(\Omega, \mathcal{S}, \mathcal{P})$ *be a probability space and let* $X_t : \Omega \to \mathbb{R}^n$ *be* $\mathcal{S}$-$\mathcal{B}^n$*-measurable for each* $t \in \mathbb{R}_0^+$. *Then a stochastic process* $\{X_t; t \geq 0\}$ *with the following properties is called* (standard) n-dimensional Wiener process *on* $(\Omega, \mathcal{S}, \mathcal{P})$:

1. $X_0 = 0$ $\mathcal{P}$*-almost surely,*

2. *for each* $k \in \mathbb{N}$ *and each vector* $(t_1, t_2, \ldots, t_k)^T \in \mathbb{R}^k$*, where* $0 \leq t_1 < t_2 < \ldots < t_k$*, the increments* $X_{t_1}, X_{t_2} - X_{t_1}, \ldots, X_{t_k} - X_{t_{k-1}}$ *are independent,*

3. *for each* $s, t \in \mathbb{R}_0^+, s < t$*, the random variable* $X_t - X_s$ *is* $\mathbf{N}(0, (t-s)\mathbb{E}_n)$ *distributed,*

4. *all paths of* $\{X_t; t \geq 0\}$ *are continuous.*

The following theorem shows the existence and uniqueness of a measure $\mathcal{W}^n$ on the Wiener space, such that a Wiener process exists on this certain space.

Theorem 1.39 (Wiener measure). *Let* $\{W_t; t \geq 0\}$ *be a stochastic process, such that* $W_t(\omega) = \omega(t)$ *for each* $t \geq 0$ *and* $\omega \in \Omega^n$*. Then, for each* $n \in \mathbb{N}$*, there exists exactly one measure* $\mathcal{W}^n$ *on* $(\Omega^n, \mathcal{B}(\Omega^n))$*, called* Wiener measure, *with the properties 1 to 4 as stated in Definition 1.38.*

We can now proceed with the following definition.

Definition 1.40 (Wiener process on $(\Omega^n, \mathcal{B}(\Omega^n), \mathcal{W}^n)$**).** *The stochastic process* $\{W_t; t \geq 0\}$*, where* $W_t(\omega) = \omega(t)$ *for each* $t \geq 0$ *and* $\omega \in \Omega^n$*,* $n \in \mathbb{N}$*, is called* n-dimensional Wiener process *on* $(\Omega^n, \mathcal{B}(\Omega^n), \mathcal{W}^n)$.

From now on we assume that the probability space we are working with supplies an n-dimensional Wiener process.

When 'pinning down' a Wiener process at some future time $T > 0$ the result is a Wiener process connecting two points, the Brownian bridge:

Theorem and Definition 1.41 (Brownian bridge). *Let $(\Omega, \mathcal{S}, \mathcal{P})$ be a probability space and $\{W_t; t \geq 0\}$ a one-dimensional Wiener process. For any $T \in \mathbb{R}^+$ and $a, b \in \mathbb{R}$ the process*

$$B_t^{a\to b} := a\Big(1 - \frac{t}{T}\Big) + b\frac{t}{T} + \Big(W_t - \frac{t}{T}W_T\Big), \quad t \in [0,T],$$

defines the Brownian bridge from a to b on $[0,T]$.
The process $B_t^{a\to b}$ is normal distributed with $\mathcal{P}$-almost surely continuous paths, expectation function

$$m(t) := E_{\mathcal{P}}[B_t^{a\to b}] = a\Big(1 - \frac{t}{T}\Big) + b\frac{t}{T}, \quad t \in [0,T],$$

and covariance function

$$cov(s,t) = \min(s,t) - \frac{st}{T}, \quad s,t \in [0,T].$$

1.2.2 Martingales

Martingales are a class of processes defined by a conditional expectation that we will use frequently.

Definition 1.42 ((Sub-)Martingale). *Let $(\Omega, \mathcal{S}, \mathcal{P})$ be a probability space equipped with the filtration $\{\mathcal{F}_t; t \geq 0\}$. A real valued stochastic process $\{X_t; t \geq 0\}$ is a* martingale, *if it is adapted to the filtration $\{\mathcal{F}_t; t \geq 0\}$, $E_{\mathcal{P}}[|X_t|] < \infty$ for all $t \geq 0$ and the martingale property*

$$X_s = E_{\mathcal{P}}[X_t|\mathcal{F}_s] \tag{1.1}$$

holds for all $s, t \in \mathbb{R}_0^+$, $s \leq t$. If $\{X_t; t \geq 0\}$ is adapted to $\{\mathcal{F}_t; t \geq 0\}$, $E_{\mathcal{P}}[|X_t|] < \infty$ for all $t \geq 0$ and $X_s \leq E_{\mathcal{P}}[X_t|\mathcal{F}_s]$ holds instead of (1.1), $\{X_t; t \geq 0\}$ is called submartingale.

An important subset is the set of square-integrable martingales.

Definition 1.43 (Square-integrable martingale). *Let $(\Omega, \mathcal{S}, \mathcal{P})$ be a probability space, $\{\mathcal{F}_t; t \geq 0\}$ a filtration in $\mathcal{S}$ and $\{M_t; t \geq 0\}$ a continuous, real valued martingale adapted to $\{\mathcal{F}_t; t \geq 0\}$. M is* square-integrable, *if $E_{\mathcal{P}}[|M_t|^2] < \infty$ for every $t \geq 0$.*

Examples of square-integrable martingales are the Wiener process and the Brownian bridge for any choice of T, a and b.

For square-integrable martingales we define the quadratic variation and cross-variation processes.

Theorem and Definition 1.44 (Quadratic variation). *Let* $(\Omega, \mathcal{S}, \mathcal{P})$ *be a probability space,* $\{\mathcal{F}_t; t \geq 0\}$ *a filtration in* $\mathcal{S}$ *and* $\{M_t; t \geq 0\}$ *a continuous, square-integrable martingale adapted to* $\{\mathcal{F}_t; t \geq 0\}$ *satisfying* $M_0 = 0$ $\mathcal{P}$*-a.s. Then a stochastic process* $\{Q_t; t \geq 0\}$ *with the following properties exists:*

1. $\{Q_t; t \geq 0\}$ *is adapted to the filtration* $\{\mathcal{F}_t; t \geq 0\}$*,*

2. $Q_0 = 0$ $\mathcal{P}$*-a.s.,*

3. *the mapping* $t \mapsto Q_t(\omega)$ *is a non-decreasing, right-continuous function for* $\mathcal{P}$*-almost all* $\omega \in \Omega$*,*

4. $E_{\mathcal{P}}[Q_t] < \infty$ *for every* $t \geq 0$*,*

5. $\{M_t^2 - Q_t; t \geq 0\}$ *is a martingale.*

We set $\langle M \rangle_t := Q_t$ *for every* t *and call the stochastic process* $\{\langle M \rangle_t; t \geq 0\}$ *the* quadratic variation *of* M*. The existence and uniqueness of* $\langle M \rangle$ *follows from the fact that* $\{M_t^2; t \geq 0\}$*,* $(t, \omega) \mapsto M_t^2(\omega)$ *for every* $t \geq 0$ *and* $\omega \in \Omega$*, is a submartingale and the Doob-Meyer-decomposition of submartingales.*

Proof. See [Karatzas, Shreeve 2001], Definition 1.5.3, Definition 1.4.4 and Theorem 1.4.10. □

The use of the term *quadratic variation* may appear to be unfounded. Indeed, a more conventional use of this term is the following. Let $(\Omega, \mathcal{S}, \mathcal{P})$ be a probability space, let $\{X_t; t \geq 0\}$ be a real valued stochastic process, fix some $t > 0$, and let $\Pi = \{t_0, t_1, \ldots, t_n\}$, with $0 = t_0 \leq t_1 \leq \ldots \leq t_n = t$, be a partition of $[0, t]$. Define the *quadratic variation of* X *over* Π to be

$$Q_t(\Pi) = \sum_{k=1}^{n} |X_{t_k} - X_{t_{k-1}}|^2.$$

Now define the mesh of the partition Π as $\|\Pi\| = \max_{1 \leq k \leq n} |t_k - t_{k-1}|$. Our justification of theorem and definition 1.44 is the following result.

Theorem 1.45. *Let* $(\Omega, \mathcal{S}, \mathcal{P})$ *be a probability space, let* $\{\mathcal{F}_t; t \geq 0\}$ *be a filtration and* $\{M_t; t \geq 0\}$ *a continuous, square-integrable martingale adapted to* $\{\mathcal{F}_t; t \geq 0\}$ *with* $M_0 = 0$ $\mathcal{P}$*-a.s. For partitions* Π *of* $[0, t]$ *we have* $\lim_{\|\Pi\| \to 0} Q_t(\Pi) = \langle M \rangle_t$ *in probability, i.e., for every* $\epsilon > 0, \eta > 0$ *there exists* $\delta > 0$ *such that*

$$\|\Pi\| < \delta \Rightarrow P(|Q_t(\Pi) - \langle M \rangle_t| > \epsilon) < \eta.$$

Proof. See [Karatzas, Shreeve 2001], Theorem 1.5.8. □

Having introduced the quadratic variation of continuous martingales, we define now the cross-variation of two continuous martingales.

Definition 1.46 (Cross-variation). *Let* $(\Omega, \mathcal{S}, \mathcal{P})$ *be a probability space, let* $\{\mathcal{F}_t; t \geq 0\}$ *be a filtration and* $\{M_t; t \geq 0\}$ *and* $\{N_t; t \geq 0\}$ *two continuous, square-integrable martingales adapted to* $\{\mathcal{F}_t; t \geq 0\}$*. Again we assume* $M_0 = N_0 = 0$ $\mathcal{P}$*-a.s. We define the* cross-variation process $\{\langle M, N\rangle_t; t \geq 0\}$ *of* M *and* N *by*

$$\langle M, N\rangle_t := \frac{1}{4}[\langle M+N\rangle_t - \langle M-N\rangle_t]$$

for all $t \geq 0$*.*

Lemma 1.47 (Properties of the Cross-variation). *The cross-variation is a symmetric bilinear form, i.e., let* $\{X_t; t \geq 0\}$*,* $\{Y_t; t \geq 0\}$ *and* $\{Z_t; t \geq 0\}$ *be continous, square-integrable martingales as in Definition 1.46 and* $\alpha, \beta \in \mathbb{R}$*. Then*

$$\begin{aligned} \langle \alpha X + \beta Y, Z\rangle_t &= \alpha\langle X, Z\rangle_t + \beta\langle Y, Z\rangle_t, \\ \langle X, Y\rangle_t &= \langle Y, X\rangle_t \end{aligned}$$

for every $t \geq 0$*.*

Proof. See [Karatzas, Shreeve 2001], Definition 1.5.5 and Problem 1.5.7. □

Remark 1.48 (Quadratic variation of the Wiener process). Let $\{W_t; t \geq 0\}$ be an one-dimensional Wiener process on some probability space $(\Omega, \mathcal{S}, \mathcal{P})$. The following properties are known from probability theory:

1. $\{W_t; t \geq 0\}$ is a square-integrable, continuous martingale.
2. $\langle W\rangle_t = t$ for all $t \geq 0$.

If $W^{(1)}$ and $W^{(2)}$ are two independent Wiener processes, then the cross-variation process is $\mathcal{P}$-a.s. zero,

$$\langle W^{(1)}, W^{(2)}\rangle_t = 0 \ \mathcal{P}\text{-a.s. for all } t \geq 0.$$

Important about martingales is the possibility to change the time-scale such that the resulting process is a Wiener process under certain assumptions. We give a multi-dimensional version of this theorem that also covers the one-dimensional case.

Theorem 1.49 (Knight's theorem). *Let* $(\Omega, \mathcal{S}, \mathcal{P})$ *be a probability space,* $\{\mathcal{F}_t; t \geq 0\}$ *a filtration in* $\mathcal{S}$ *and* $\{(M_t^{(1)}, \ldots, M_t^{(n)}); t \geq 0\}$ *a continuous martingale adapted to* $\{\mathcal{F}_t; t \geq 0\}$*. Further, let* $\lim_{t\to\infty}\langle M^{(i)}\rangle_t = \infty$ $\mathcal{P}$*-a.s. and*

$$\langle M^{(i)}, M^{(j)}\rangle_t = 0 \text{ for all } 1 \leq i \neq j \leq n, t \geq 0. \tag{1.2}$$

Define

$$T_i(s) = \inf\{t \geq 0; \langle M^{(i)}\rangle_t > s\} \text{ for all } s \geq 0, 1 \leq i \leq n,$$

so that for each i and s, the random time $T_i(s)$ is a stopping time for the (right-continuous) filtration $\{\mathcal{F}_t; t \geq 0\}$. Then the processes

$$W_s^{(i)} := M_{T_i(s)}^{(i)}, \quad \mathcal{G}_s^{(i)} := \mathcal{F}_{T_i(s)}, \quad s \geq 0, \quad 1 \leq i \leq n$$

are independent, standard, one-dimensional Wiener processes.

Proof. See [Karatzas, Shreeve 2001], Theorem 3.4.13. □

In the special case $n = 1$ the additional condition (1.2) is dropped and Knight's theorem reduces to the usual time-change theorem for continous martingales.

1.2.3 Markov processes

A Markov process is a process such that the behavior in the near future depends not on the past of the process, given by the filtration generated by the process up to the present, but only on the present state.

Definition 1.50 (Markov process). *Let $(\Omega, \mathcal{S}, \mathcal{P})$ be a probability space, let $\{X_t; t \geq 0\}$ be a stochastic process and $\{\mathcal{F}_t^X; t \geq 0\}$ the filtration generated by $\{X_t; t \geq 0\}$. $\{X_t; t \geq 0\}$ is called* Markovian *or a* Markov process *if the Markov property*

$$\mathcal{P}(X_t \in S|\mathcal{F}_s) = \mathcal{P}(X_t \in S|\sigma(X_s)) \tag{1.3}$$

holds $\mathcal{P}$-a.s. for all $s, t \in \mathbb{R}_0^+$, $s < t$, and each $S \in \mathcal{S}$. The conditional probability $P(A|\mathcal{H})$ is defined by $E_{\mathcal{P}}[\mathbf{1}_{\{A\}}|\mathcal{H}]$ for any subset $A \in \mathcal{S}$ and any σ-field $\mathcal{H} \subseteq \mathcal{S}$.

We will derive an alternative formulation of the Markov property (1.3). This needs some preparation. Especially transition kernels will play an important role.

Definition 1.51 ((Transition) Kernel). *Let $(\Omega, \mathcal{S})$ and $(\Omega', \mathcal{S}')$ be two measurable spaces. The mapping*

$$K : \Omega \times \mathcal{S}' \to \overline{\mathbb{R}_0^+}$$

is called (transition) kernel from $(\Omega, \mathcal{S})$ to $(\Omega', \mathcal{S}')$ *if the following holds:*

1. *$\omega \mapsto K(\omega, S')$ is $\mathcal{S}$-$\mathcal{B}(\overline{\mathbb{R}_0^+})$-measurable for every $S' \in \mathcal{S}'$,*
2. *$S' \mapsto K(\omega, S')$ is a measure on $\mathcal{S}'$ for every $\omega \in \Omega$.*

K is called Markovian, *if*

$$K(\omega, \Omega') = 1$$

holds for every $\omega \in \Omega$.

Naturally, a Markovian kernel K implies a probability measure on $\mathcal{S}'$ for every $\omega \in \Omega$.
If μ is an arbitrary measure on $\mathcal{S}$, we may define a measure μ' on $\mathcal{S}'$ as follows. Choose some $S' \in \mathcal{S}'$. Define μ' by

$$\mu'(S') := \int K(\omega, S')d\mu(\omega)$$

for some kernel K. It follows from the properties of the integral that μ' is a measure on $\mathcal{S}'$. If K is Markovian and μ is a probability measure, μ' is also a probability measure. We usually denote μ' by μK and write the definition in the form

$$(\mu K)(S') := \int \mu(d\omega)K(\omega, S') := \int K(\omega, S')d\mu(\omega).$$

The kernel K also induces a mapping from the set of extended $\mathcal{S}$-$\overline{\mathcal{B}}$-measurable functions to the set of extended $\mathcal{S}'$-$\overline{\mathcal{B}}$-measurable functions. This can be verified as follows. Let $f' : \Omega' \to \overline{\mathbb{R}_0^+}$ be an extended $\mathcal{S}'$-$\overline{\mathcal{B}}$-measurable function. Then $\omega \mapsto \int f'(\omega')K(\omega, d\omega')$ is an extended $\mathcal{S}$-$\overline{\mathcal{B}}$-measurable function. We denote this mapping also by K:

$$(Kf')(\omega) := \int f'(\omega')K(\omega, d\omega').$$

for every extended $\mathcal{S}'$-$\overline{\mathcal{B}}$-measurable function f'.
As the next building block of our alternative formulation of the Markov property we introduce semi-groups of kernels.

Definition 1.52 (Semi-groups of kernels). *Let $(K_t)_{t\geq 0}$ be a set of kernels on a measurable space $(\Omega, \mathcal{S})$. If the Chapman-Kolmogorov-equations*

$$K_{s+t}(x, S) = \int K_s(x, dy)K_t(y, S),$$

or shorter

$$K_{s+t} = K_s K_t,$$

hold for every $s, t \in \mathbb{R}_0^+$ and $(x, S) \in \Omega \times \mathcal{S}$, $(K_t)_{t\geq 0}$ is a semi-group of kernels *or* transition semi-group. *If all kernels are Markovian, the semi-group $(K_t)_{t\geq 0}$ is also called* Markovian.

The finite-dimensional distributions of a stochastic process can be deduced from the transition semi-group as the following theorem shows.

Theorem 1.53. *Let $(\Omega, \mathcal{S})$ be a measurable space, μ a probability measure thereon and $(K_t)_{t\geq 0}$ a Markovian semi-group of kernels. For every finite subset of $\mathbb{R}_0^+$,*

$I := \{t_1, \dots, t_n\} \subset \mathbb{R}_0^+$, $n \in \mathbb{N}$, with $t_1 < t_2 < \dots < t_n$, and $S_{t_1}, \dots, S_{t_n} \in \mathcal{S}$ set

$$P_I(X_{t_1} \in S_{t_1}, \dots, X_{t_n} \in S_{t_n}) := \iint \dots \int \mathbf{1}_{\{S_{t_1} \times \dots \times S_{t_n}\}}(x_1, \dots, x_n) K_{t_n - t_{n-1}}(x_{n-1}, dx_n) \dots K_{t_1}(x_0, dx_1)\mu(dx_0). \tag{1.4}$$

Then $(P_I)_I$ is a family of finite-dimensional distributions on $(\Omega, \mathcal{S})$, where the index I runs over all finite subsets of $\mathbb{R}_0^+$.

We are now in the position to formulate the Markov property (1.3) in an alternative way.

Theorem 1.54 (Equivalent formulation of the Markov property). *Let $(\Omega, \mathcal{S}, \mathcal{P})$ be a probability space, $\{X_t; t \geq 0\}$ a stochastic process, $(K_t)_{t\geq 0}$ a Markovian semi-group of kernels and μ a probability measure on $\mathcal{S}$. Assume that the finite-dimensional distributions of $\{X_t; t \geq 0\}$ are determined by $(K_t)_{t\geq 0}$ and μ as in theorem 1.53. Then $\{X_t; t \geq 0\}$ is Markovian with respect to the filtration $\{\mathcal{F}_t^X; t \geq 0\}$, $\mathcal{F}_t^X := \sigma(X_s; 0 \leq s \leq t)$. Further, we have*

$$P(X_t \in S | \mathcal{F}_s) = K_{t-s}(X_s, S) \tag{1.5}$$

$\mathcal{P}$-a.s. for every $S \in \mathcal{S}$ and $s, t \in \mathbb{R}_0^+$, $s < t$.
$K_{t-s}(X_s, S)$ denotes the random variable $\omega \mapsto K_{t-s}(X_s(\omega), S)$.

Remark 1.55. Let $(\Omega, \mathcal{S}, \mathcal{P})$ be a probability space. Since

$$K(\omega, S) = \int_S K(\omega, d\omega') = \int_{\Omega'} \mathbf{1}_{\{S\}}(\omega') K(\omega, d\omega') = (K\mathbf{1}_{\{S\}})(\omega)$$

we may write (1.5) as

$$E_{\mathcal{P}}[\mathbf{1}_{\{S\}}(X_t) | \mathcal{F}_s] = (K_{t-s}\mathbf{1}_{\{S\}})(X_s).$$

for $s, t \in \mathbb{R}_0^+, s < t$. Choose some non-negative, Ω-$\mathcal{B}(\overline{\mathbb{R}_0^+})$-measurable extended function $f : \omega \to \overline{\mathbb{R}_0^+}$. Approximating f in the usual way by indicator functions and taking limits gives us

$$E_{\mathcal{P}}[f \circ X_t | \mathcal{F}_s] = (K_{t-s}f)(X_s).$$

Finally, this is the equivalent formulation of the Markov property we will use later.

An important example of a Markov process is the Wiener process defined above. Additionally, the Wiener process is also a martingale. Note, however, that there are martingales that are not Markovian and vice versa.

As shown before, a Markov process can be characterized by its transition semi-group. The most interesting for our purposes is the transition semi group of the n-dimensional Wiener process.

Theorem and Definition 1.56 (Transition semi group of Wiener process). *Let* $\{W_t; t \geq 0\}$ *be the standard* n*-dimensional Wiener process on some probability space* $(\Omega, \mathcal{S}, \mathcal{P})$. *The* transition semi-group $(K_t^W)_{t \geq 0}$ of the Wiener process *is given by*

$$(K_t^W f)(x) = \int \phi_t(x-y) f(y) d\lambda^n(y),$$

where

$$\phi_t(x) = (2\pi t)^{-\frac{n}{2}} \exp\Big(-\frac{\|x\|_2^2}{2t}\Big) \text{ for all } t > 0$$

denotes the density function of the standard n*-dimensional Wiener process and* $f : \mathbb{R}^n \to \mathbb{R}$ *is an arbitrary Borel-measurable, non-negative function, i.e.,* $f(x) \geq 0$ *for all* $x \in \mathbb{R}^n$. *We set* $\phi_0(x) = \epsilon(x)$*, the Dirac distribution, having mass 1 in* x *and mass 0 anywhere else.*

1.2.4 Stochastic integration

In this section we give a short overview over the stochastic Itô-Integral with respect to continuous, square-integrable martingales. We will not go into the details here, this section gives just an idea how the stochastic integral is constructed. A more in-depth description can be found in [Karatzas, Shreeve 2001] and [Protter 1990]. Throughout this section we assume that we are working in a probability space $(\Omega, \mathcal{S}, \mathcal{P})$ supporting a filtration $\{\mathcal{F}_t; t \geq 0\}$ that satisfies the usual conditions and a standard, one-dimensional Wiener process.
To state the existence theorem for stochastic differential equations, we need a rather technical definition.

Definition 1.57 (Progressively measurable). *If the map* $f : \mathbb{R}_0^+ \times \Omega \to \mathbb{R}^n$ *is* $\mathcal{B}(\mathbb{R}_0^+) \otimes \mathcal{S}$-$\mathcal{B}^n$*-measurable and*

$$f_s : \Omega^n \to \mathbb{R}^n, \omega \mapsto f(s, \omega) \text{ is } \mathcal{F}_s\text{-}\mathcal{B}^n\text{-measurable for each } s \geq 0,$$

then the process $\{f_t; t \geq 0\}$ *defined via the map* f_s *is called* progressively measurable *with respect to* $\{\mathcal{F}_t; t \geq 0\}$.

Since we are only working with continuous processes, there is no gap between measurable and progressively measurable processes.

Lemma 1.58. *If the stochastic process* $\{X_t; t \geq 0\}$ *is adapted to the filtration* $\{\mathcal{F}_t; t \geq 0\}$ *and every sample path is right-continuous or else every sample path is left-continuous, then* $\{X_t; t \geq 0\}$ *is also progressively measurable with respect to* $\{\mathcal{F}_t; t \geq 0\}$.

Proof. See [Karatzas, Shreeve 2001], Proposition 1.1.13 □

Modifying the stochastic process $\{X_t; t \geq 0\}$ with $\mathcal{P}$-a.s. continuous sample paths on null sets, we may assume that all sample paths are continuous. Therefore continuous, $\{\mathcal{F}_t; t \geq 0\}$-adapted processes $\{X_t; t \geq 0\}$ are also progressively measurable.

The construction of the Itô-Integral is based on the approximation of stochastic processes by simple processes.

Definition 1.59 (Simple process). *A stochastic process $\{X_t; t \geq 0\}$ is called* simple *if there exists a strictly increasing sequence of real numbers t_n, $n \in \mathbb{N}_0$, with $t_0 = 0$ and $\lim_{n\to\infty} t_n = \infty$, as well as a sequence of random variables $\zeta_n : \Omega \to \mathbb{R}$, $n \in \mathbb{N}_0$, and a non random constant $C < \infty$ with $\sup_{n\in\mathbb{N}_0} |\zeta_n(\omega)| \leq C$, for every $\omega \in \Omega$, such that*

- *ζ_n is $\mathcal{F}_{t_n}$-measurable for every $n \in \mathbb{N}_0$ and*
- $X_t(\omega) = \zeta_0(\omega)\mathbf{1}_{\{\{0\}\}}(t) + \sum_{i=0}^{\infty} \zeta_i(\omega)\mathbf{1}_{\{(t_i,t_{i+1}]\}}(t), t \in \mathbb{R}_0^+, \omega \in \Omega.$

Based on this definition we define the stochastic integral with respect to simple processes.

Definition 1.60 (Itô-Integral for simple processes). *Let $\{M_t; t \geq 0\}$ be a continuous, square-integrable martingale on $(\Omega, \mathcal{S}, \mathcal{P})$ adapted to the filtration $\{\mathcal{F}_t; t \geq 0\}$, and $\{X_t; t \geq 0\}$ a simple stochastic process with respect to $\{\mathcal{F}_t; t \geq 0\}$. Further, let t_n be s strictly increasing sequence of real numbers and ζ_n a sequence of random variables as in Definition 1.59, such that*

$$X_t(\omega) = \zeta_0(\omega)\mathbf{1}_{\{\{0\}\}}(t) + \sum_{i=0}^{\infty} \zeta_i(\omega)\mathbf{1}_{\{(t_i,t_{i+1}]\}}(t), t \in \mathbb{R}_0^+, \omega \in \Omega.$$

The Itô-integral $I_t^M(X)$ of X with respect to M at time $t \in \mathbb{R}_0^+$ *is defined as*

$$\begin{aligned} I_t^M(X) : \Omega \to \mathbb{R}, \omega \mapsto & \sum_{i=0}^{n-1} \zeta_i(M_{t_{i+1}} - M_{t_i}) + \zeta_n(M_t - M_n) \\ = & \sum_{i=0}^{\infty} \zeta_i(M_{\min(t_i,t_{i+1})} - M_{\min(t,t_i)}), \end{aligned}$$

where $n \geq 0$ is the unique integer for which $t_n \leq t \leq t_{n+1}$. We denote $I_t^M(X)$ by

$$\int_0^t X_s dM_s.$$

The stochastic integral can now be extended to a wider class of processes that shall be defined now.

Definition 1.61 (Set of admissible integrands). *Let* $\{M_t; t \geq 0\}$ *be a continuous, square-integrable martingale adapted to the filtration* $\{\mathcal{F}_t; t \geq 0\}$. $L^*(M)$ *denotes the set of equivalence classes of all* progressively measurable, $\{\mathcal{F}_t; t \geq 0\}$-*adapted stochastic processes, satisfying*

$$E_{\mathcal{P}}\Big[\int_0^T X_t^2 d\langle M\rangle_t\Big] < \infty \tag{1.6}$$

for all $T > 0$, *where the integral* $\int_0^T X_t^2 d\langle M\rangle_t$ *is defined path wise in the Lebesgue-Stieltjes sense.*

We shall follow the usual custom of not being very careful about the distinction of equivalence classes and the stochastic processes which are members of those classes. For example we have no qualms about saying, that $L^*(M)$ contains all bounded, measurable, $\{\mathcal{F}_t; t \geq 0\}$-adapted processes. Of course, $L^*(M)$ depends on the choice of M.
The integral for simple processes can be extended in a highly non-trivial, yet straightforward manner. Details may be found in [Karatzas, Shreeve 2001]. From now on we take the integral $I_t^M(X)$ for $X \in L^*(M)$ with respect to some martingale M as granted and denote it also by

$$\int_0^t X_s dM_s.$$

Having introduced the Itô-Integral we give a short characterization and show some interesting and useful properties.

Theorem 1.62 (Properties of the Itô-Integral). *Suppose* $\{M_t; t \geq 0\}$ *and* $\{N_t; t \geq 0\}$ *are two continuous, square-integrable,* $\{\mathcal{F}_t; t \geq 0\}$-*adapted martingales and take* $X \in L^*(M)$, $Y \in L^*(N)$. *Set* $I_t^M(X) := \int_0^t X_s dM_s$, $I_t^N(Y) := \int_0^t y_s dN_s$. *Then the following properties hold:*

- $\{I_t^M(X); t \geq 0\}$ *is a continuous, square-integrable,* $\{\mathcal{F}_t; t \geq 0\}$-*adapted martingale.*
- $\langle I^M(X)\rangle_t = \int_0^t X_s^2 d\langle M\rangle_s$.
- $\langle I^M(X), I^N(Y)\rangle_t = \int_0^t X_s Y_s d\langle M, N\rangle_s$.

Set now $N_t := I_t^M(X)$ *for every* $t \geq 0$ *and suppose that* $Y \in L^*(N)$. *Then* $XY \in L^*(M)$ *and*

$$I^N(Y) = I^M(XY)$$

or shorter in differential notation, if $dN = XdM$ *then*

$$YdN = XYdM.$$

Finally, suppose $X, Y \in L^*(M)$ *and* $\alpha, \beta \in \mathbb{R}$. *Then*

$$I^M(\alpha X + \beta Y) = \alpha I^M(X) + \beta I^M(Y).$$

Proof. See [Karatzas, Shreeve 2001], Proposition 3.2.10, Proposition 3.2.17 and Corollary 3.2.20. □

Stochastic calculus knows an equivalent to the chain rule from real analysis, Itô's formula.

Theorem 1.63 (Itô's formula). *Let $(\Omega, \mathcal{S}, \mathcal{P})$ be our usual probability space, $\{(M_t^{(1)}, \dots, M_t^{(n)}; t \geq 0\}$ a vector of continuous martingales, $M_t^{(i)} : \Omega \to \mathbb{R}$, adapted to some filtration $\{\mathcal{F}_t; t \geq 0\}$ and set $X_t := X_0 + M_t$, $t \geq 0$, where $X_0 : \Omega \to \mathbb{R}^n$, $n \in \mathbb{N}$, is an $\mathcal{F}_0$-$\mathcal{B}^n$-measurable random variable. Let $f(t,x) : \mathbb{R}_0^+ \times \mathbb{R}^n \to \mathbb{R}$ be of class $C^{1,2}$. Then, $\mathcal{P}$-a.s.,*

$$\begin{aligned} f(t, X_t) = & f(0, X_0) + \int_0^t \frac{\partial}{\partial t} f(s, X_s) ds + \sum_{i=1}^n \int_0^t \frac{\partial}{\partial x^{(i)}} f(s, X_s) dM_s^{(i)} + \\ & \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \int_0^t \frac{\partial^2}{\partial x^{(i)} x^{(j)}} f(s, X_s) d\langle M^{(i)}, M^{(j)} \rangle_s \end{aligned} \tag{1.7}$$

for every $t \geq 0$.

Proof. See [Karatzas, Shreeve 2001], Theorem 3.3.3. □

Remark 1.64. Equation (1.7) is often written in differential notation,

$$\begin{aligned} df(t, X_t) = & \frac{\partial}{\partial t} f(t, X_t) dt + \sum_{i=1}^n \frac{\partial}{\partial x^{(i)}} f(t, X_t) dM_t^{(i)} + \\ & \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2}{\partial x^{(i)} x^{(j)}} f(t, X_t) d\langle M^{(i)}, M^{(j)} \rangle_t, \mathcal{P}, t \in \mathbb{R}_0^+. \end{aligned}$$

To evaluate the term $d\langle M^{(i)}, M^{(j)} \rangle_t$, which is also written as

$$d\langle M^{(i)}, M^{(j)} \rangle_t = dM_t^{(i)} dM_t^{(j)},$$

we apply Theorem 1.62 and the conventional 'multiplication table'

	dt	dW_t	$d\tilde{W}_t$
dt	0	0	0
dW_t	0	dt	0
$d\tilde{W}_t$	0	0	dt

where W,$\tilde{W}$ are independent Wiener processes.

1.2.5 Stochastic differential equations

We give an existence and uniqueness result for Itô-stochastic differential equations. Again let $n, m \in \mathbb{N}$.

Definition 1.65 (Stochastic differential equation). *Let $\{W_t; t \geq 0\}$ be an m-dimensional Wiener process on our probability space $(\Omega, \mathcal{S}, \mathcal{P})$ and ξ an n-dimensional random variable independent of $\{W_t; t \geq 0\}$. Assume $E_{\mathcal{P}}[\|\xi\|_2^2] < \infty$. Further, let the mapping $b : \mathbb{R}_0^+ \times \mathbb{R}^n \to \mathbb{R}^n$ be $\mathcal{B}(\mathbb{R}_0^+) \otimes \mathcal{B}^n$-$\mathcal{B}^n$-measurable and the mapping $\sigma : \mathbb{R}_0^+ \times \mathbb{R}^n \to \mathbb{R}^{n\times m}$ $\mathcal{B}(\mathbb{R}_0^+) \otimes \mathcal{B}^n$-$\mathcal{B}^{n\times m}$-measurable. The equation*

$$X_t = \xi + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s, \mathcal{P}, t \in \mathbb{R}_0^+, X_0 = \xi, \tag{1.8}$$

is called Itô stochastic differential equation, *where $\{X_t; t \geq 0\}$ is a suitable process with continuous sample paths, the 'solution' of (1.8), which we will define immediately.*

Instead of (1.8) we usually write

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, \mathcal{P}, t \in \mathbb{R}_0^+, X_0 = \xi. \tag{1.9}$$

$b(t, X_t)$ is called *drift term*, $\sigma(t, X_t)$ is called *diffusion term.*

Usually, we denote under which measure the stochastic differential equation is valid since we will later regularly switch to equivalent measures, thus changing the stochastic differential equation in some way.
We define now the solution of (1.8) respectively (1.9).

Definition 1.66 (Solution of an SDE). *Under the assumptions of Definition 1.65 we define the* augmented filtration *$\{\mathcal{F}_t; t \geq 0\}$ as follows. Consider the filtration*

$$\mathcal{G}_t := \sigma(\xi, W_s; 0 \leq s \leq t), \quad t \geq 0,$$

as well as the collection of null sets

$$\mathcal{N} := \{N \subseteq \Omega; \exists G \in \sigma\Big(\bigcup_{t\geq 0} \mathcal{G}_t\Big) \text{ with } N \subseteq G \text{ and } \mathcal{P}(G) = 0\}$$

and create the augmented filtration

$$\mathcal{F}_t := \sigma(\mathcal{G}_t \cup \mathcal{N}), \quad t \geq 0; \quad \mathcal{F}_\infty := \sigma\Big(\bigcup_{t\geq 0} \mathcal{F}_t\Big). \tag{1.10}$$

A strong solution *of the stochastic differential equation (1.8) respectively (1.9), on the probability space $(\Omega, \mathcal{S}, \mathcal{P})$ and with respect to the fixed Wiener process W and initial condition ξ, is a stochastic process $\{X_t; t \geq 0\}$ with continuous sample paths and with the following properties:*

1. *$\{X_t; t \geq 0\}$ is adapted to the filtration $\{\mathcal{F}_t; t \geq 0\}$,*

2. *$\mathcal{P}(\{\omega \in \Omega; X_0(\omega) = \xi\}) = 1$,*

3. *$\mathcal{P}(\{\omega \in \Omega; \int_0^t b^{(i)}(s, X_s(\omega)) + \sigma_{ij}^2(s, X_s(\omega))ds < \infty\}) = 1$ holds for every $i = 1, \ldots, n$, $j = 1, \ldots, m$ and $t \geq 0$,*

4. *the integral equation (1.8) holds $\mathcal{P}$-almost surely.*

Remark 1.67. The filtration $\{\mathcal{F}_t; t \geq 0\}$ defined by (1.10) satisfies the usual conditions (see [Karatzas, Shreeve 2001]).

We can formulate conditions under which a solution of a Itô-SDE exists.

Theorem 1.68 (Existence and uniqueness). *Suppose the assumptions of Definition 1.65 and 1.66 hold and that the coefficients $b(t, x)$ and $\sigma(t, x)$ satisfy the global Lipschitz and linear growth conditions*

$$\|b(t, x) - b(t, y)\|_2 + \|\sigma(t, x) - \sigma(t, y)\|_F \leq L\|x - y\|_2 \tag{1.11}$$
$$\|b(t, x)\|_2^2 + \|\sigma(t, x)\|_F^2 \leq L^2(1 + \|x\|_2^2) \tag{1.12}$$

foe every $t \geq 0$, $x, y \in \mathbb{R}^n$, where L is a positive constant. On some probability space $(\Omega, \mathcal{S}, \mathcal{P})$, let ξ be an $\mathbb{R}^n$-valued random variable, independent of the m-dimensional Wiener process $\{W_t; t \geq 0\}$, and with finite second moment

$$E_{\mathcal{P}}[\|\xi\|_2^2] < \infty.$$

Let $\{\mathcal{F}_t; t \geq 0\}$ be as in (1.10). Then there exists a continuous, $\{\mathcal{F}_t; t \geq 0\}$-adapted process $\{X_t; t \geq 0\}$ which is a strong solution of equation (1.8) relative to W with initial condition ξ. Moreover, this process is square-integrable: for every $T > 0$, there exists a constant $C(L, T)$, depending only on L and T, such that

$$E_{\mathcal{P}}[\|X_t\|_2^2] \leq C(L, T)(1 + E_{\mathcal{P}}[\|\xi\|_2^2]) \exp(C(L, T)t), \quad t \in [0, T].$$

Let $\{Y_t; t \geq 0\}$ be another strong solution of (1.8). Then

$$\mathcal{P}(\{\omega \in \Omega; X_t(\omega) = Y_t(\omega) \text{ for } t \geq 0\}) = 1$$

holds ($\mathcal{P}$-a.s. uniqueness).

Remark 1.69. It is also possible to define a *weak solution* of (1.8). Since we are only interested in strong solutions, we omit this concept and speak simply of a *solution of* (1.8) when talking about a strong solution.

Remark 1.70 (Alternative definition of the Brownian bridge). There are other ways to define a Brownian bridge with parameters $a, b \in \mathbb{R}$ and $T \in \mathbb{R}^+$. First, define the stochastic differential equation

$$dB_t^{a\to b} = \frac{b - B_t^{a\to b}}{T_t} dt + dW_t, \mathcal{P}, t \in [0,T], B_0^{a\to b} = a. \tag{1.13}$$

It can be shown that the solution of (1.13) coincides with the definition 1.41 of the Brownian bridge given above.
Another interpretation of a Brownian bridge would be a standard Wiener process with given start *and end condition*. In differential notation,

$$dB_t^{a\to b} = dW_t, \mathcal{P}, t \in [0,T], B_0^{a\to b} = a, B_T^{a\to b} = b.$$

A Brownian bridge with additional parameters, for example

$$dB_t^{a\to b} = \sigma_t dW_t, \mathcal{P}, t \in [0,T], B_0^{a\to b} = a, B_T^{a\to b} = b,$$

or

$$dB_t^{a\to b} = B_t^{a\to b} \hat{\sigma}_t dW_t, \mathcal{P}, t \in [0,T], B_0^{a\to b} = a, B_T^{a\to b} = b,$$

is called *generalized Brownian bridge*. Of course, σ and $\hat{\sigma}$ are such that the above SDEs admit a solution.

To end this section, we present an important theorems from the theory of stochastic differential equations, Girsanov's theorem.

Theorem 1.71 (Girsanov's theorem). *Let $(\Omega, \mathcal{S}, \mathcal{P})$ be a probability space, $\{W_t; t \in [0,T]\}, T \in \mathbb{R}^+$, an n-dimensional Wiener process, $n \in \mathbb{N}$, and let the mappings h and G be defined as in Definition 1.65 such that the SDE*

$$dX_t = h(t, X_t)dt + G(t, X_t)dW_t, \mathcal{P}, t \in [0,b], X_0 = x_0,$$

admits a solution $\{X_t; t \in [0,T]\}$. Suppose there exist an $\mathbb{R}^m$-valued stochastic process $\{U_t; t \in [0,T]\}$, an $\mathbb{R}^n$-valued stochastic process $\{V_t; t \in [0,T]\}$ and a filtration $\{\mathcal{H}_t; t \in [0,t]\}$ in $\mathcal{S}$ such that $\{U_t; t \in [0,T]\}$ and $\{V_t; t \in [0,T]\}$ are adapted to $\{\mathcal{H}_t; t \in [0,T]\}$ and martingales with respect to this filtration. Assume further that

$$G(t,\omega)U(t,\omega) = h(t,\omega) - V(t,\omega) \textit{ for all } t \in [0,T] \ \mathcal{P}\text{-a.s.}$$

holds and $U(t,\omega)$ satisfies Novikov's condition

$$E_{\mathcal{P}}\Big[\exp\Big(\frac{1}{2}\int_0^T U^2(s,\omega)ds\Big)\Big] < \infty.$$

Put

$$M_t := \exp\Big(-\int_0^t U(s,\omega)dW_s - \frac{1}{2}\int_0^t U^2(s,\omega)ds\Big) \textit{ for all } t \in [0,T]$$

and define the measure $\mathcal{Q}$ *by*

$$d\mathcal{Q}(\omega) = M_T(\omega)d\mathcal{P}(\omega) \textit{ on } \mathcal{F}_T. \tag{1.14}$$

Then

$$\widehat{W}_t := \int_0^t U(s,\omega)ds + W_t, \mathcal{Q}, t \in [0,T], \widehat{W}_0 = 0,$$

is a Wiener process with respect to $\mathcal{Q}$ *and in terms of* $\widehat{W}_t$ *the process* $\{X_t; t \in [0,T]\}$ *has the stochastic integral representation*

$$dX_t = V(t,\omega)dt + G(t,\omega)d\widehat{W}_t, \mathcal{Q}, t \in [0,T].$$

Remark 1.72. In the theorem above Novikov's condition is sufficient to guarantee that $\{M_t; t \in [0,T]\}$ is a martingale. This in turn yields the existence and uniqueness of the measure $\mathcal{Q}$ as defined by (1.14). An alternative way of expressing (1.14) is $\frac{d\mathcal{Q}}{d\mathcal{P}} = M_T$ on $\mathcal{F}_T$. Since $M_T > 0$ $\mathcal{P}$-a.s. $\mathcal{P}$ is absolutely continuous with respect to $\mathcal{Q}$ and thus the two measures are equivalent, $\mathcal{P} \sim \mathcal{Q}$

Itô's formula can be viewed as a change-of-variables formula for stochastic differential equations and Girsanov's theorem tells us, how a given SDE changes under a change of measure. It is remarkable that only the drift term changes, but not the diffusion term. We will exploit this later when pricing derivatives.

1.3 Option pricing theory

Before we start pricing derivatives we must first define the underlying economy in which we are working. An *economy* with a finite time horizon $T \in \mathbb{R}^+$, which we denote by $\mathcal{E}$, consists of an underlying probability space $(\Omega, \mathcal{S}, \mathcal{P})$ supporting a Wiener process and the corresponding filtration $\{\mathcal{F}_t^W; t \in [0,T]\}$ and a set of non-dividend paying assets. We model the prices of the assets by n stochastic processes $\{A_t^{(i)}; t \in [0,T]\}$, where $A_t^{(i)} : [0,T] \times \Omega \to \mathbb{R}_0^+$ are $\mathcal{B}([0,T]) \otimes \mathcal{S}$-$\mathcal{B}(\overline{\mathbb{R}_0^+})$-measurable random variables for $i = 1, \ldots, n$. Again we use the shorter notation $A^{(i)}$ instead of $\{A_t^{(i)}; t \in [0,T]\}$ when appropriate. We also assume that there are no transaction costs, so the stochastic process $\{A_t^{(i)}; t \in [0,T]\}$ gives us the (random) price at which we may buy or sell an arbitrary amount of the i-th asset at any time $t \in [0,T]$. Further, the processes $A^{(i)}$ are assumed to be continuous and almost surely finite.
Trading in the economy is modeled by stochastic processes $\{\Psi_t^{(i)}; t \in [0,T]\}$,

where $\Psi_t^{(i)} : [0,T] \times \Omega \to \mathbb{R}$ are $\mathcal{B}([0,T]) \otimes \mathcal{S}$-$\mathcal{B}$-measurable random variables for every $i = 1, \dots, n$. We interpret $\{\Psi_t^{(i)}; t \in [0,T]\}$ as the (random) amount we hold of the i-th asset at time t. If $\Psi_t^{(i)}(\omega)$ is positive for some $t \in [0,T]$ and $\omega \in \Omega$, we bought that amount of the i-th asset, if the sign is negative we sold the adequate amount. The set of stochastic processes $\Psi := \{(\Psi_t^{(1)}, \dots, \Psi_t^{(n)}); t \in [0,T]\}$ is called *trading strategy*. We will allow trading in this economy throughout time, but we will prelude the injection of external funds in the economy - all trading strategies must be self-financing. Further, we assume that all trading strategies are $\mathcal{P}$-almost surely bounded form below by a constant not depending on t. These strategies are called *tame*. When holding a portfolio of assets according to a trading strategy Ψ, the value V_t^Ψ of the portfolio as time t is given by $V_t^\Psi := \sum_{i=1}^n \Psi_t^{(i)} A_t^{(i)}$. We call $\{V_t^\Psi; t \in [0,T]\}$ a *price process* and any price process that is almost surely positive is a *numeraire*.
We are not only interested in the assets themselves, but also in derivatives of those assets. A *derivative* is a financial instrument whose value depends only on the value of one or more underlying assets.
We will further assume that the economy is arbitrage-free. The concept of arbitrage and its absence have important consequences as stated by the following theorems.

Theorem and Definition 1.73 (Arbitrage). *An* arbitrage *is a self-financing tame trading strategy Ψ for the time interval $[0,T]$ such that $V_0^\Psi = 0$ and $V_T^\Psi > 0$ with positive probability. Further $V_t^\Psi \geq 0$ $\mathcal{P}$-almost surely for every $t \in (0,T]$.*
If a measure $\mathcal{N} \sim \mathcal{P}$ and a numeraire $\{N_t; t \in [0,T]\}$ exist such that all relative asset prices $\{\frac{A_t^{(i)}}{N_t}; t \in [0,T]\}$ are martingales with respect to $\{\mathcal{F}_t^W; t \in [0,T]\}$,

$$\frac{A_t^{(i)}}{N_t} = E_{\mathcal{N}}\Big[\frac{A_T^{(i)}}{N_T} | \mathcal{F}_t^W\Big], \quad 1 \leq i \leq n, \tag{1.15}$$

the measure $\mathcal{N}$ is called an equivalent martingale measure.
An economy $\mathcal{E}$ is arbitrage-free, *if a* numeraire pair $(N, \mathcal{N})$, *consisting of a numeraire $\{N_t; t \in [0,T]\}$ and an equivalent martingale measure $\mathcal{N}$, exists and* (1.15) *holds.*

The existence of an arbitrage implies that is possible to make a riskless profit, something we want to exclude from our model.
For our purposes it is sufficient to know that an economy is arbitrage-free if a numeraire pair exists and (1.15) holds.
An important consequence of the absence of arbitrage is the following theorem.

Theorem 1.74 (Fundamental theorem of asset pricing). *Suppose $\mathcal{E}$ is an arbitrage-free economy with a numeraire pair $(N, \mathcal{N})$ given. Let $\{V_t; t \in [0,T]\}$ denote the value of a derivative and assume that at some time $T > 0$ the price*

V_T of the derivative has been determined by the evolution of asset prices. Then the value V_t of the derivative at any time t prior to T is given by

$$V_t = N_t E_{\mathcal{N}}[V_T N_T^{-1} | \mathcal{F}_t^W]. \tag{1.16}$$

Proof. See [Hunt, Kennedy 2000]. □

It will be important to be able to change the numeraire we are working with. The following lemma gives more details.

Lemma 1.75 (Change of Numeraire). *Assume we are working in an economy $\mathcal{E}$ supporting a numeraire pair consisting of a numeraire $\{N_t; t \in [0,T]\}$ and a probability measure $\mathcal{N} \sim \mathcal{P}$ such that the price of any traded asset $A^{(i)}$ relative to N is a martingale under $\mathcal{N}$, i,e.*

$$\frac{A_t^{(i)}}{N_t} = E_{\mathcal{N}}\Big[\frac{A_T^{(i)}}{N_T} | \mathcal{F}_t^W\Big], \quad 0 \le t \le T.$$

Let $\{M_t; t \in [0,T]\}$ be an arbitrary numeraire. Then there exists a probability measure $\mathcal{M} \sim \mathcal{P}$ such that the price of any derivative $\{C_t; t \in [0,T]\}$, whose value at time T can be replicated by some trading strategy Ψ, $V_T^\Psi = C_T$, normalized by M is a martingale under $\mathcal{M}$, i.e.,

$$\frac{C_t^{(i)}}{M_t} = E_{\mathcal{M}}\Big[\frac{C_T^{(i)}}{M_T} | \mathcal{F}_t^W\Big], \quad 0 \le t \le T. \tag{1.17}$$

Equation (1.17) holds in particular if C is one of the assets in the market. Moreover, the density function $\frac{d\mathcal{M}}{d\mathcal{N}}$ defining the measure $\mathcal{M}$ with respect to the measure $\mathcal{N}$ is given by

$$\frac{d\mathcal{M}}{d\mathcal{N}} = \frac{M_T N_0}{M_0 N_T}.$$

Proof. See [Geman et al. 1995]. □

Our main interest will lie on a special kind of derivatives, namely options.

Definition 1.76 (Option). *An* option *is a contract between two counterparties that gives one party the right, but not the obligation, to buy (call-option) or sell (put-option) an asset for a pre-specified price on a pre-specified date.*

Of course the theory developed so far for pricing derivatives is also valid for pricing options.
Like other assets, options are traded continuously in the market and may be bought and sold at any time. Knowing the option value at some time T we may determine the option price at any time $t < T$ using (1.16). Usually one takes T to be the maturity of the option. In this case, V_T is called the *payoff* of the option.

Chapter 2

Interest Rate Derivatives

In this chapter we will introduce some of the basic products that define the interest rate market. The products that we consider are all traded in the market and as already mentioned in the last chapter we will also consider options based on these products. We will also introduce standard market terminology which is important for the discussions that follow. Most of this chapter is due to [Hunt, Kennedy 2000].

2.1 Deposits

We begin with the most fundamental instrument in the market, the deposit.

Definition 2.1 (Deposit). *A* deposit *is an agreement between two parties in which one pays the other a cash amount and in return receives the money back at some* pre-agreed *future date, with a* pre-agreed *additional payment of interest.*

Observe that in the above definition deposits are always made until a fixed date. Deposits are available for a range of maturities, the most common being overnight, 1 week, 1, 2, 3, 6 and 12 months. The amount of interest being paid at the end of the interest period is quoted via an *accrual factor* and an *interest rate*. These two concepts will play an important role in most of the derivatives we shall encounter.

2.1.1 Accrual factor and LIBOR

Definition 2.2 ((Spot) LIBOR). LIBOR*, also called* spot LIBOR*, is the London Interbank Offer Rate, the rate of interest that one London bank will offer to pay on a deposit made by another London bank.*

There will, in general, be a different LIBOR for each of the standard deposit maturities. LIBOR is always quoted on a per annum basis, i.e., it is the interest rate paid for a deposit maturing in exactly one year. For other maturities, especially those below one year, we have to calculate the appropriate fraction thereof. To be able to do this we introduce accrual factors, also called daycount fractions:

Definition 2.3 (Accrual factor, daycount fraction). *The total amount of interest that the depositing bank will receive is calculated by multiplying the LIBOR by the amount of time, as a proportion of a year, for which the money has been on deposit. This amount of time is called the* accrual factor *or* daycount fraction.

Accrual factors are calculated by dividing the number of days in the period by the number of days in a year. Naturally, different markets use different conventions to calculate these figures. A common example is *actual/360*. Here the number of days in the interest period is the exact number of calendar days. The whole year is taken to be 360 days long, therefore we divide by 360. Due to this convention the accrual factor may be greater than 1, since the number of calendar days in a year is 365 respectively 366 in a leap year. We will not need detailed knowledge about how the accrual factor is calculated and simply assume that an appropriate accrual factor has already been calculated. Say we made a deposit of notional amount A for a certain period of time, denoting the corresponding LIBOR by L and the accrual factor by α. At maturity along with the notional amount A we will receive an interest payment of amount $A\alpha L$. Usually we suppress the notional amount and assume it to be unity. These cashflows are illustrated in Figure 2.1.

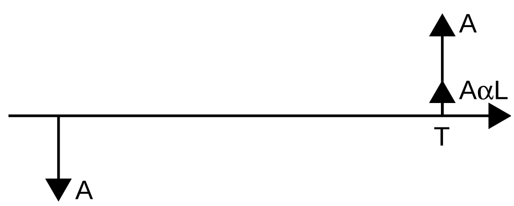

Figure 2.1: Deposit cashflows

We will often represent products in this way. The horizontal axis represents the time of cashflows; those above the axis are ones we receive, while those below the line are ones we pay.

2.1.2 Forward LIBOR

Until now we restricted ourselves to deposits starting today and maturing on some future date. But what if we want to make a deposit *starting* on some future date? The LIBOR rates introduced in Section 2.1.1 are clearly inappropriate interest rates since they implicitly assume that the deposit starts today. We have to introduce another class of interest rates, the forward LIBORs.

Definition 2.4 (Forward LIBOR). *The* forward LIBOR $L_t[T, U]$ *is the rate of interest one London bank will offer to pay at time t on a deposit made by another London bank, starting on date T and maturing on date U. Of course,* $t \leq T < U$.

Naturally the interest rate $L_t[T, U]$ is fixed once the deposit is made but its value will change through time. We will soon see how to calculate this value. Note that for $t = T$ the forward and spot LIBOR coincide.

2.2 Interest Rate Swaps

Definition 2.5 (Interest Rate Swap). *An* interest rate swap*, which we will abbreviate to* swap*, is an agreement between two counterparties to exchange a series of cashflows on pre-agreed dates in the future.*

We refer to all payments we receive as one *leg* of the swap, and those we pay constitute the other leg. To specify the swap we must know its start date (the start of the first accrual period), maturity date (the date of the last cashflow), and the payment frequency, for each of the legs. Typical maturities are 1 to 10, 12, 15, 20 and 30 years. Each leg can in general have a different payment frequency, but here we describe the case when it is the same for both legs.
Suppose there are a total of n cashflows in each leg, cashflow i occurring at time U_i. One of the legs will be a fixed leg, for which all the cashflows are known at the start of the swap. Assuming a notional unit amount the amount of cashflow i is given by $\alpha_i K$, where α_i is the accrual factor for the period $[U_{i-1}, U_i]$ and K is the *fixed rate* for the swap.
The other leg is the *floating* leg, so called because the payment amounts will be set during the life of the swap. The amount of payment i is set at time U_{i-1} and is the accrual factor α_i multiplied by $L_{U_{i-1}}[U_{i-1}, U_i]$, the LIBOR then quoted for the period $[U_{i-1}, U_i]$. For notational convenience we shall often refer to the start of the i-th accrual period as T_i (rather than U_{i-1}). Then the LIBOR appropriate for the i-th period is $L_{T_i}[T_i, U_i]$.
From now on we will use the convention that 'wavy' cashflows represent cashflows not known today in our cashflow diagrams.
Swaps are usually entered at zero initial cost for both counterparties. A swap with this property is called a *par* swap, and the value of the fixed rate K for which the swap has zero value is called the *par swap rate*. In this case when the swap start is spot (i.e., the swap starts immediately), this is often abbreviated to just the swap rate, and it is these par swap rates that are quoted on trading screens in the financial markets. A swap for which the start date is not spot is, naturally enough, referred to as a *forward start swap* on the corresponding par

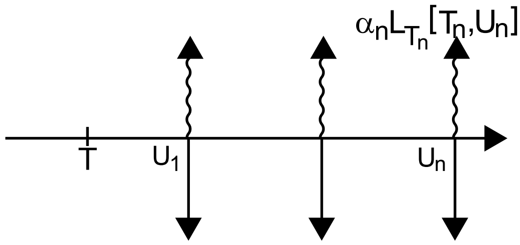

Figure 2.2: Payers interest rate swap

swap rate is the *forward swap rate.*
We will denote the (forward) par swap rate, at time t for a swap starting at date T and making *fixed* payments on dates given by the vector $U = (U_1, \ldots, U_n)$, by $S_t[T, U]$ (often just S_t when the context is clear). Note that this definition is independent of the payment frequency of the floating leg. This is indeed the case as we see when we learn how to value swaps in Section 2.4.3 .
Swaps where one pays a fixed rate and receives a floating rate are usually referred to as *payers swaps* whereas swaps where one receives a fixed rate and pays a floating rate are referred to as *receiver swaps.*

2.3 Zero coupon bonds

Definition 2.6 (Zero coupon bonds). Zero coupon bonds (ZCBs)*, also known as pure discount bonds, are assets which entitle the holder to receive a cashflow at some future date T. The amount of this cashflow is part of the contract specification, although we will often assume it to be a unit amount.*

The main value from introducing ZCBs comes from the fact that other products can be built up from them. In this sense they are fundamental. As an example, the fixed leg of a swap consists of n known cashflows and so can be thought of as n ZCBs. Thus the value of the fixed leg will be the sum of the values of each of these component ZCBs. Note that although ZCBs mature on some future date T they may be bought and sold at any time $t \leq T$. Of course, one will not receive the full (unit) notional amount when selling a ZCB at time $t < T$ but only a fraction of it.

2.4 Discount factors and valuation

In this section we introduce discount factors, the most fundamental tool for summarizing the value of interest rate products. We will see how discount factors can be used to express the value of the basic instruments we have met and also how they are related to forward LIBOR and swap rates.

Figure 2.3: Zero coupon bond

2.4.1 Discount factors

Suppose we are at time t. For any time $t \leq T$ there is a corresponding *discount factor*, denoted as D_{tT}, defined to be the value, at time t, of a ZCB paying a unit amount at time T. The value D_{tT} is thus exactly the amount we have to invest at time t to receive a unit payment at time T. Note that, for all t, $D_{tt} = 1$ since a unit cashflow now is worth a unit.
We will usually assume that the initial *discount curve*, $\{D_{0T} : T \geq 0\}$ if today time is zero, is known. This is not strictly the case since the standard liquid market instruments only give us a limited amount of information. In practice banks take these liquid market instruments and construct a full discount curve consistent with their prices. There is no unique way to do this, and what is done often depends on the use to which the discount curve is to be put. We shall not discuss the techniques used to construct discount curves in detail and shall assume that this has been done as the starting point for the rest of our analyses. When setting up a model for the discount curve we will only model the pure discount bond prices. An especially convenient way to accomplish is creating a numeraire model:

Definition 2.7 (Numeraire model). *Let $\mathcal{E}$ be an arbitrage-free economy with a numeraire pair $(N, \mathcal{N})$ given. Given a set of maturity dates $\mathcal{T} = \{T_i; 1 \leq i \leq n\}$ the term structure using a* numeraire model *is defined by*

$$D_{tT} = N_t E_{\mathcal{N}}[\frac{D_{TT}}{N_T}|\mathcal{F}_t^W] \quad (2.1)$$

$$= N_t E_{\mathcal{N}}[N_T^{-1}|\mathcal{F}_t^W], \quad T \in \mathcal{T}, \quad 0 \leq t \leq T. \quad (2.2)$$

2.4.2 Deposit valuation

A standard deposit of unit amount at time T pays at maturity U an amount $1 + \alpha L_T[T, U]$. Since there are no further payments than the initial deposit and the final redemption, these two cashflows must have the same value at time T.

Thus it follows that

$$\begin{aligned}
D_{TT} &= (1+\alpha L_T[T,U])D_{TU} = 1, \\
\Leftrightarrow D_{TU} &= (1+\alpha L_T[T,U])^{-1}, \\
\Leftrightarrow L_T[T,U] &= \frac{D_{TT}-D_{TU}}{\alpha D_{TU}} \\
&= \frac{1-D_{TU}}{\alpha D_{TU}}.
\end{aligned} \tag{2.3}$$

Analogous to the above calculations we may find the value of some forward LIBOR rate $L_t[T,U]$,

$$L_t[T,U] := \frac{D_{tT}-D_{tU}}{\alpha D_{tU}}.$$

2.4.3 Swap valuation

Again we take the notional amount to be one for simplicity.
The fixed leg of a swap consists of a series of payments on dates $U_1,\dots,U_n$, the payment at time U_i being $\alpha_i K$. If the swap starts on time T the value of the fixed leg at time $t \le T$ is given by

$$\begin{aligned}
V_t^{Fixed} &= K\sum_{j=1}^{n}\alpha_j D_{tU_j} \\
&= KP_t[T,U],
\end{aligned}$$

where $U=(U_1,\dots,U_n)$ and

$$P_t[T,U] := \sum_{j=1}^{n}\alpha_j D_{tU_j}, \quad t\in[0,T]. \tag{2.4}$$

We will refer to the expression $\{P_t[T,U]; t\in[0,T]\}$ as the *present value per basis point*, or shortly PVBP, of the swap. It represents the value of the fixed leg of the swap if the fixed rate where unity. The PVBP will play an important role in some of the products we examine later. When there can be no possible confusion we will usually abbreviate $P_t[T,U]$ to P_t.
The floating leg of the swap is more difficult to value, comprising as it does a series of floating payments. However, we can find a simple trading strategy which replicates the swaps floating payments. Suppose at time t we have an amount of cash

$$V_t^{Float} = D_{tT} - D_{tU_n}. \tag{2.5}$$

Take this cash and buy one ZCB of maturity T and sell one of maturity U_n. At time T take the unit paid by the ZCB and deposit it at LIBOR until time U_1. At time U_1 $(=T_2)$ we will receive $1+\alpha_1 L_T[T,U_1]$. The term $\alpha_1 L_T[T,U_1]$ is the

floating payment we need to replicate the swap, the extra unit of principal we deposit at LIBOR until time U_2. We repeat this at each floating payment date until time U_n we receive $1 + \alpha_n L_{T_n}[T_n, U_n]$. The LIBOR part of this payment is what we need for the last floating payment of the swap, the notional pays the amount owed on the ZCB we sold at time t.
It follows that the net value of a payers swap is exactly

$$\begin{aligned} V_t^{Payers} &= V_t^{Float} - V_t^{Fixed} \\ &= D_{tT} - D_{tU_n} - KP_t[T, U]. \end{aligned} \tag{2.6}$$

The forward swap rate $S_t[T, U]$ is the value K which sets this value to zero. Substituting $V_t^{Payers} = 0$ into (2.6) yields

$$S_t[T, U] = \frac{D_{tT} - D_{tU_n}}{P_t[T, U]}. \tag{2.7}$$

Substituting this back into (2.6), we obtain the more usual expression for the value of a payers swap,

$$V_t^{Payers} = P_t[T, U](S_t[T, U] - K). \tag{2.8}$$

The value of a receivers swap is, of course, $-V_t^{Payers}$.
As a final point on swaps, recall that we allow the floating and the fixed legs to have different payment dates. It is easy to see from the arguments above that the value of the floating leg, as given by (2.5), is independent of the floating payment frequency. It is therefore only the fixed payment frequency that matters when defining and valuing a forward swap.

2.5 Caps, floors and swaptions

Having introduced several interest rate products, we will now turn to derivatives of these products. The simplest and most liquid of these derivatives are caps, floors and swaptions.

2.5.1 Caps and floors

It is often the case that a customer is either making or receiving a series of floating payments and does not wish to convert them into a series of fixed payments. This may be because he believes future rate moves will be in his favor. However, he is then exposed if rates move against him and would like to buy some protection against this without removing the benefits of the move in rates he expects. The solution in this case is for the customer to buy a *cap* or a *floor*.
Caps and floors are similar to swaps in that they are made up of an series of

payments on regularly spaced dates $U_j, j = 1, \ldots, n$. On date U_j the holder of a cap receives a payment of amount

$$\alpha_j (L_{T_j}[T_j, U_j] - K)^+,$$

where $T_j := U_{j-1}$ is the setting time for the LIBOR which pays at time U_j and $(K)^+ := \max(K, 0)$. A floor is similar except the payment amount is given by

$$\alpha_j (K - L_{T_j}[T_j, U_j])^+.$$

The constant K in these expressions is part of the contract specification and is known as the *strike* of the option.
A counterparty who is paying LIBOR on some unit notional amount and who buys a cap (usually from a third party) has ensured that he will never pay more than $\alpha_j K$ at time U_j; one who is receiving LIBOR on some unit notional amount and buys a floor (again, usually from a third party) will never receive less than $\alpha_j K$. Each individual payment is usually referred to as a *caplet* or *floorlet*. At any time $t \leq T$, the amount $\alpha_j (L_{T_j}[T_j, U_j] - K)^+$ is referred to as the *intrinsic value* of the caplet and, similarly, $\alpha_j (K - L_{T_j}[T_j, U_j])^+$ is referred to as the intrinsic value of the floorlet. In either case, if the intrinsic value is positive then the option is said to be *in the money*, whereas if it is zero the option is said to be *out of the money*. When $L_{T_j}[T_j, U_j] = K$ both options are said to be *at the money*. Payoff profiles of caplets are floorlets are shown in Figures 2.4 and 2.5. Since caps and floors are linear combinations of caplets and floorlets, it suffices to price these single payments.

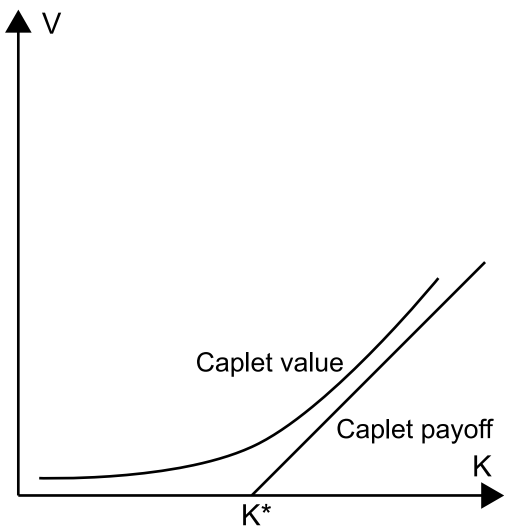

Figure 2.4: Caplet payoff profile and value

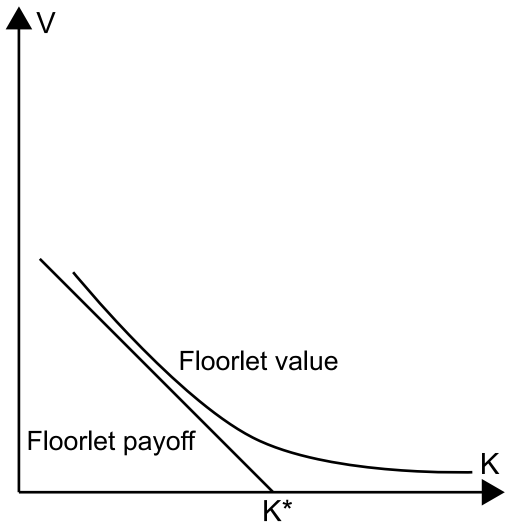

Figure 2.5: Floorlet payoff profile and value

2.5.2 Valuation of caps and floors

The payoff at some time S from a single caplet setting at time T is just

$$V_S^{Caplet} = \alpha(L_T[T,U] - K)^+$$

and so its time-t value is given, dropping the $[T,U]$ from the notation, by

$$V_t^{Caplet} = N_t E_{\mathcal{N}}[\alpha(L_T - K)^+ N_S^{-1} | \mathcal{F}_t^W], \tag{2.9}$$

for some numeraire pair $(N, \mathcal{N})$. To calculate this price we must choose a numeraire N and a suitable model for $L[T,U]$ in the measure $\mathcal{N}$. Suppose we choose $N_t = D_{tS}$, the discount bond maturing on the payment date S. The corresponding measure $\mathcal{N}$ is usually referred to as the *forward measure* and we denote it by $\mathcal{F}$. Using this measure has two important consequences. First, $N_S = D_{SS} = 1$ and so this disappears from the equation (2.9). Secondly, the forward LIBOR L, which is of the form

$$\begin{aligned} L_t &= \frac{D_{tT} - D_{tS}}{D_{tS}} \\ &= \frac{D_{tT} - D_{tS}}{N_t}, \end{aligned}$$

is a ratio of asset prices over the numeraire and so must be a martingale. As long as we model L as a martingale (under $\mathcal{F}$) we will have a model that is arbitrage-free and the caplet value will be given by (2.9). Because interest rates are always assumed to be positive, we will model L as a log-normal martingale,

$$dL_t = \sigma_t L_t dW_t, \mathcal{F}, t \in [0,T], L_0 = \xi,$$

for some deterministic σ, a Wiener process $\{W_t;\, t \in [0,T]\}$ and $\xi > 0$. This is standard market practice and yields a solution

$$L_t = L_0 \exp\Big(\int_0^t \sigma_u dW_u - \frac{1}{2}\int_0^t \sigma_u^2 du\Big).$$

Substituting into (2.9) now yields after some lengthy computations

$$\begin{aligned} V_t^{Caplet} &= D_{tS} E_{\mathcal{F}}[\alpha(L_T - K)^+ | \mathcal{F}_t^W] \\ &= D_{tS} E_{\mathcal{F}}\Big[\alpha\Big(L_t \exp\Big(\int_t^T \sigma_u dW_u - \frac{1}{2}\int_t^T \sigma_u^2 du\Big) - K\Big)^+ | \mathcal{F}_t^W\Big] \\ &= \alpha D_{tS}\big(L_t \Phi(d_1) - K\Phi(d_2)\big), \end{aligned} \tag{2.10}$$

where

$$\begin{aligned} d_1 &:= \frac{\log(L_t/K)}{\Sigma_{t;T}\sqrt{T-t}} + \frac{1}{2}\Sigma_{t;T}\sqrt{T-t}, \\ d_2 &:= \frac{\log(L_t/K)}{\Sigma_{t;T}\sqrt{T-t}} - \frac{1}{2}\Sigma_{t;T}\sqrt{T-t}, \\ \Sigma_{t;T} &:= \frac{1}{T-t}\int_t^T \sigma_u^2 du. \end{aligned} \tag{2.11}$$

This is Black's formula for caplets and was first published in [Black 1976]. Of course a cap, consisting of a series of caplets, may be valued accordingly. Pricing of floorlets is identical and yields the usual put option formula,

$$\begin{aligned} V_t^{Floorlet} &= \alpha D_{tS} E_{\mathcal{F}}[(K - L_T)^+ | \mathcal{F}_t^W] \\ &= \alpha D_{tS}\big(K\Phi(-d_2) - L_t\Phi(-d_1)\big). \end{aligned} \tag{2.12}$$

Knowing the strike K and the volatility σ we can calculate the price V_t of a caplet or floorlet using Blacks's formula (2.10) or (2.12). It is also possible to invert Black's formula and, given a market price V_t and a strike K, to retrieve the *implied volatility*. This may be done for a series of caplets of different strikes. It then turns out that the volatility σ depends not only on the time t, but also on the strike K. The effect of the volatility depending on the strike is called a *volatility smile* or *volatility skew*, depending on whether the volatility is increasing when moving away from the at-the-money-strike (i.e., the strike equal to the corresponding forward LIBOR currently quoted in the market) or decreasing for out-of-the-money strikes and increasing for in-the-money-strikes (or vice versa).

Since smile and skew effects can be observed in the market the distribution of caplet and floorlet prices is not log-normal as initially assumed but has in fact fatter tails (leptokurtosis). Modeling these effects is one of the most difficult parts when setting up an interest rate model.

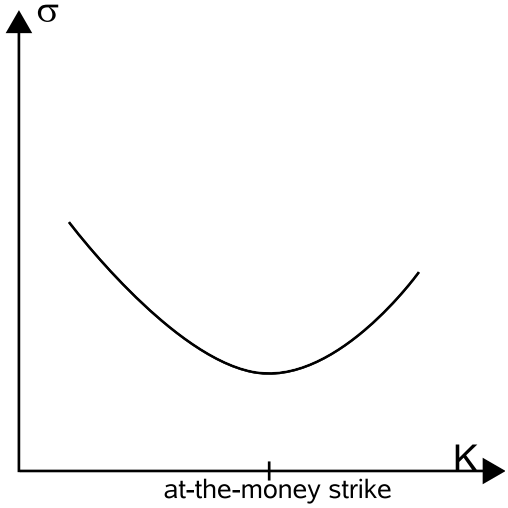

Figure 2.6: Volatility smile

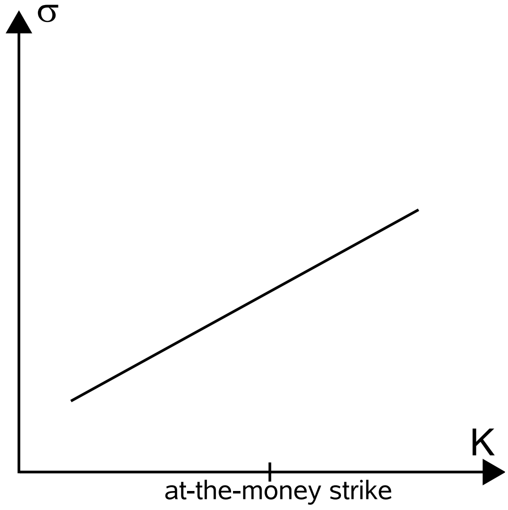

Figure 2.7: Volatility skew

2.5.3 Vanilla swaptions

Just as a caplet is an option on a ZCB, a *swaption* is an option on a swap. Swaptions are commonly traded in the market and are priced using Black's formula. There are two types of swaptions: receivers swaptions and payers swaptions . In the first case, upon exercise the option holder enters a swap in which he receives a fixed rate K, the swaption strike, and pays the floating rate; a payers swap is the reverse. We refer to this as a vanilla swaption if the underlying swap is a plain vanilla swap.
We now price a payers swaption. Let $U_t[T, U]$ be the value at time t of a vanilla payers swap which starts on date T and makes fixed payments on the dates $S = (U_1, \ldots, U_n)$. The effective payoff to the option holder, who has the right but not the obligation to enter the swap, is given at the option expiry T by

$$V_T^{PSwaption} = \max(U_T, 0)$$

since he will only enter the swap if it is to his advantage to do so. The value at time t of the swaption is, by the usual valuation formula, given by

$$\begin{aligned} V_t^{PSwaption} &= N_t E_{\mathcal{N}}[\max(U_T, 0) N_T^{-1} | \mathcal{F}_t^W] \\ &= N_t E_{\mathcal{N}}[P_T (S_T - K)^+ N_T^{-1} | \mathcal{F}_t^W], \end{aligned} \tag{2.13}$$

for some suitable numeraire pair $(N, \mathcal{N})$. The finally equality above follows by substituting (2.8) for the value of a swap, where P_T, S_T and K are, as usual, the PVBP and par swap rate at time T and the fixed rate for the underlying swap. We usually refer to K as the *swaption strike*. At any time $t \leq T$ we use the terminology *in the money*, *out of the money* and *at the money* just as we did for caplets and floorlets according to the sign of $(S_t - K)$.
To evaluate (2.13) we follow a procedure similar to that for the FRA. On this occasion we choose P as the numeraire, is which case the corresponding martingale measure is called the *swaption measure* and will be denoted by $\mathcal{S}$. This reduces (2.13) to the form

$$V_t^{PSwaption} = P_t E_{\mathcal{S}}[(S_T - K)^+ | \mathcal{F}_t^W],$$

which should by now be quite familiar. We recall from (2.7) that the forward swap rate is of the form

$$S_t = \frac{D_{tT} - D_{tU_n}}{P_t},$$

which is the ratio of asset prices over the numeraire, so must be a martingale under $\mathcal{S}$. Since we want interest rates to remain positive, we will model S to be a log-normal martingale,

$$dS_t = \sigma_t S_t dW_t, \mathcal{S}, t \in [0, T], S_0 = \xi,$$

for some deterministic function σ, a Wiener process $\{W_t;\ t \in [0,T]\}$ and $\xi > 0$. This puts us in precisely the same framework as when we priced caplets and floorlets and again yields Black's formula for the price ,

$$V_t^{PSwaption} = P_t\big(S_t\Phi(d_1) - K\Phi(d_2)\big), \tag{2.14}$$

where

$$\begin{aligned} d_1 &= \frac{\log(L_t/K)}{\Sigma_{t;T}\sqrt{T-t}} + \frac{1}{2}\Sigma_{t;T}\sqrt{T-t}, \\ d_2 &= \frac{\log(L_t/K)}{\Sigma_{t;T}\sqrt{T-t}} - \frac{1}{2}\Sigma_{t;T}\sqrt{T-t}, \\ \Sigma_{t;T} &= \frac{1}{T-t}\int_t^T \sigma_u^2 du. \end{aligned}$$

The corresponding receivers valuation formula is just

$$V_t^{RSwaption} = P_t\big(K\Phi(-d_2) - S_t\Phi(-d_1)\big).$$

2.5.4 Flexible caps and Bermudan swaptions

Until now we discussed only so called *European options*, i.e., options that may only be exercised on the maturity date S. It is possible to remove this restriction and this results in *Bermudan options*.

Definition 2.8 (Bermudan options). *An option, bought on a date T with maturity S, $S > T$, that may be exercised only on some pre-agreed dates $t \in \{T_1, \ldots, T_n\}, n \in \mathbb{N}$, is called a* Bermudan option. *The set of exercise dates is as always part of the contract specification.*

Interesting about Bermudan swaptions is that they have a special feature compared with Bermudan options on stocks. Assume we bought a Bermudan swaption maturing on date S with a set of exercise dates $\{T_1, \ldots, T_n\}$. Assume further that we may only exercise on dates on that the underlying swap is resetting. Then when exercising the option on date T_i we enter a swap resetting on date T_i with maturity date S. Contrary to options on stocks the underlying swap is a different one with different maturity after each exercise date. This additional feature makes Bermudan swaptions more difficult to value although the value may be expressed by a variant of our usual pricing formula.
Let $\mathcal{T}$ be the set of all stopping times taking values in $\{T_1, \ldots, T_n\}$. Then the value of a Bermudan swaption with strike K as described above at time t is given by

$$V_t^{Bermudan} = \sup_{\tau \in \mathcal{T}} N_t E_{\mathcal{N}}[(S_\tau - K)^+ | \mathcal{F}_t^W], \tag{2.15}$$

under a numeraire pair $(N, \mathcal{N})$. S_τ is for each $T_i \in \{T_1, \ldots, T_n\}$ the rate of a swap resetting on T_i and maturing on S. Equation (2.15) describes a problem of stochastic control theory. One will not try to find an analytical solution. Rather the equation (2.15) is solved numerically in two steps. First the *exercise boundary* is approximated. This is some function $b : [0, S] \to \mathbb{R}$ depending on time t giving the value of the swap rate for which it is optimal to exercise the option. When running a Monte Carlo simulation it is easy to check whether at time t $S_t \geq b(t)$ holds. If this is the case the option is considered exercised and the simulation for the current path is stopped. Approximating the exercise boundary, however, is cumbersome.

Similar to Bermudan swaptions, *Flexible caps* also have a kind of early exercise feature. A Flexible cap consists of n, $n \in \mathbb{N}$, caplets just like a plain vanilla cap, but only $m < n$, $m \in \mathbb{N}$, of those caplets may be exercised. Once all caplets have been exercised the Flexible cap is worthless. It is up to the owner of the Felxible cap to decide wether he wants to exercise a caplet or not, given the caplet is in the money, of course. Bermudan swaptions and Flexible caps are prototypic interest products that may be valued using Makov-functional models.

Before introducing the Markov-functional model we introduce one more type of exotic options we will need later on.

2.6 Digital options

A digital option is one which pays either one or zero at some future date, depending on the level of some index rate. The two most common examples are digital caps and floors, and digital swaptions.

2.6.1 Digital caps and floors

Just as caps and floors are made up of a series of caplets and floorlets, so *digital caps* and *floors* are made up of digital caplets and floorlets. A digital caplet is an option which pays a unit amount at time S if at T the LIBOR for the period $[T, U]$ is above some strike level K. A floorlet pays a unit amount if the LIBOR is below the strike.

The value at S of a digital caplet given by $V_S^{DigCaplet} = \mathbf{1}_{\{L_T > K\}}$, and so it follows that

$$V_t^{DigCaplet} = N_t E_{\mathcal{N}}[\mathbf{1}_{\{L_T > K\}} N_S^{-1} | \mathcal{F}_t^W].$$

Taking D_{tS} as numeraire, as we did for caps and floors, and using the same log-normal model, yields

$$\begin{aligned} V_t^{DigCaplet} &= D_{tS} E_{\mathcal{F}}[\mathbf{1}_{\{L_T > K\}} | \mathcal{F}_t^W] \\ &= D_{tS} \mathcal{F}(L_T > K | \mathcal{F}_t^W) \\ &= D_{tS} \Phi(d_2), \end{aligned} \tag{2.16}$$

where d_2 is defined as in (2.11). The analogous digital floorlet pricing formula is then

$$\begin{aligned} V_t^{DigFloorlet} &= D_{tS} E_{\mathcal{F}}[\mathbf{1}_{\{L_T<K\}} | \mathcal{F}_t^W] \\ &= D_{tS} \Phi(-d_2). \end{aligned} \tag{2.17}$$

Remark 2.9. What we have just done in (2.16) and (2.17) is to derive the price of a digital caplet and floorlet using the same log-normal model as we did to price a standard caplet and floorlet. Even when alternative models are used for the forward LIBOR there is still a fundamental relationship between the prices of caps and floors on the one hand and digital caps and floors on the other. The relationship essentially follows from the identity

$$\frac{\partial}{\partial K}(x-K)^+ = -\mathbf{1}_{\{x>K\}},$$

which can be justified as follows. Set

$$(x-K)^+ = \begin{cases} x-K & \text{if } x > K \\ 0 & \text{otherwise.} \end{cases}$$

Taking derivatives (only from the left for $x \leq K$) yields

$$\frac{\partial}{\partial K}(x-K)^+ = \left\{ \begin{matrix} -1 & \text{if } x > K \\ 0 & \text{otherwise} \end{matrix} \right\} = -\mathbf{1}_{\{x>K\}}.$$

The more general relationship will be exploited in Chapter 4.

2.6.2 Digital swaptions

Digital swaptions are less common than digital caps and floors, and they are also more difficult to price. There are, once again, two types of digital swaptions, payers and receivers. In the case of a digital payers swaption the options holder receives a unit amount at some date M if the index swap rate, S_T, is above the strike K on the setting date T. For a digital receivers the option holder receives the unit payment if the swap rate is below the strike K.
In the case of swaptions there is no obvious choice for the option payment date M relative to the first setting date T. Common choices are $M = T$ or $M = U_1$, the first payment date of the underlying swap. Considering the payers digital, the value at time M is $\mathbf{1}_{\{S_T>K\}}$ and so the time-t value of the option is

$$\begin{aligned} V_t^{DigPayers} &= N_t E_{\mathcal{N}}[\mathbf{1}_{\{S_T>K\}} N_M^{-1} | \mathcal{F}_t^W] \\ &= N_t E_{\mathcal{N}}[\mathbf{1}_{\{S_T>K\}} D_{TM} N_T^{-1} | \mathcal{F}_t^W], \end{aligned}$$

the last equality following by conditioning on $\mathcal{F}_t^W$. To value the digital consistently with standard swaptions, we would like to take the same model as in Section 2.5.3. Working in the swaption measure, we have

$$V_t^{DigPayers} = P_t E_{\mathcal{S}}[\mathbf{1}_{\{S_T > K\}} D_{TM} P_T^{-1} | \mathcal{F}_t^W]. \tag{2.18}$$

Evaluation of the expectation on the right hand side of (2.18) presents us with a new problem. The expectation involves three distinct random terms, namely S_T, D_{TM} and P_T. We know that within our model S_T is a log-normal martingale, but we have not yet specified a suitable model for the other two terms. This will be done in the next chapter.

Chapter 3

A close relative: Market models

Before introducing (multi-dimensional) Markov-functional models we will first have a look at another class of interest rate models that is closely related to the Markov-functional model. Recently several authors including Brace, Gatarek and Musiela [Brace, Gatarek, Musiela 1997] have studied a class of interest rate models parametrized by forward LIBORs. Jamshidian [Jamshidian 1997] has extended this to models parametrized by general swap rates and they are collectively known as 'Market models'. A good introduction is given in [Brigo, Mercurio 2001]. We have a closer look at the LIBOR market model, also known as BGM model.

3.1 One-dimensional LIBOR Market model

In the one-dimensional LIBOR market model LIBOR rates are taken to be log-normal distributed in their respective measures. This results in a system of stochastic differential equations having the form

$$dL_t^{(i)} = \sigma_t^{(i)} L_t^{(i)} dW_t, \mathcal{S}^{(i)}, t \in [0, T_i], L_0^{(i)} = L_i, \tag{3.1}$$

where $T_1 < T_2 < \ldots < T_n < T_{n+1}, T_i \in \mathbb{R}, i = 1, \ldots, n+1$. $\{W_t; t \in [0, T_n]\}$ is a standard, one-dimensional Wiener process and $\sigma_t^{(i)} : [0, T_i] \to \mathbb{R}^+$ is a deterministic, strictly positive function of time. It is not unusual to specify a set of n Wiener processes $\{W_t^{(i)}; t \in [0, T_i]\}$, $i = 1, \ldots, n$, the i-th Wiener process driving the the i-th SDE of the system (3.1), and a set of correlations ρ_{ij} between those processes, such that $d\langle W^{(i)}, W^{(j)} \rangle_t = \rho_{ij}\, dt$ for $i, j = 1, \ldots, n$. This approach seems more general but it can be reduced to the system of SDEs 3.1 by changing the volatility functions $\sigma_t^{(i)}$ appropriately (see [Kerkhof, Pelsser 2002]). It is thus sufficient to work with the LIBOR market model specified by (3.1).
Usually one takes the pure discount bond $D_{tT_{n+1}}$ as a numeraire and the associated measure $\mathcal{S}^{(n)}$, the so-called terminal measure. Applying Girsanov's theorem 1.71 to calculate the drift terms under the terminal measure the system of SDEs (3.1)

becomes

$$dL_t^{(i)} = \mu^i(t, L_t^{(i+1)}, \dots, L_t^{(n)})L_t^{(i)} + \sigma_t^{(i)} L_t^{(i)} dW_t, \mathcal{S}^{(n)}, t \in [0, T_i], L_0^{(i)} = L_i \quad (3.2)$$

for $1 \leq i \leq n$, where

$$\mu^i(t, L_t^{(i+1)}, \dots, L_t^{(n)}) = - \sum_{j=i+1}^{n} \frac{\alpha_j L_t^{(j)} \sigma_t^{(j)} \sigma_t^{(i)}}{1 + \alpha_j L_t^{(j)}} \text{ for } i < n,$$

$$\mu^n \equiv 0$$

(cf. [Brigo, Mercurio 2001], Proposition 6.3.1). The by now familiar application of Itô's formula (1.7) yields the solution of (3.2),

$$L_t^{(i)} = L_0^{(i)} \exp\Big(\int_0^t \mu^i(s, L_s^{(i+1)}, \dots, L_s^{(n)}) - \frac{1}{2}\big(\sigma_s^{(i)}\big)^2 ds + \int_0^t \sigma_s^{(i)} dW_s \Big).$$

Obviously the LIBOR market model calibrates automatically to the current LIBOR rates in the market, but at the price of state-dependent dynamics. Therefore evaluation of these models requires a Monte Carlo simulation. Another problem is the incorporation of non-log-normal dynamics to generate smiles or skews.
The swap market model is exactly the same but for one exception. Instead of $D_{tT_{n+1}}$ one takes P_t, the PVBP as defined in (2.4), as a numeraire and $\mathcal{S}^{(n)}$ is the associated swap measure. Under these assumptions the swap market model can be described by the same set of SDEs (3.1) as the LIBOR market model. Thus all the statements made in the following section about the LIBOR market model hold as well for the swap market model.

3.2 Two-dimensional LIBOR Market model

Of course, there is also a two-dimensional extension of the LIBOR market model (equivalently of the swap market model). A two-dimensional model is useful, even vital, when pricing more complex interest rate products. A more in-depth explanation is given in the next section where the multi-dimensional Markov-Functional model is introduced.
As before the LIBOR rates are taken to be log-normal distributed in their respective measures,

$$dL_t^{(i)} = \sum_{k=1}^{2} L_t^{(i)} \sigma_t^{(i,k)} dW_t^{(k)}, \mathcal{S}^{(i)}, t \in [0, T_i], L_0^{(i)} = L_i, \quad 1 \leq i \leq n. \quad (3.3)$$

$T_1, \dots, T_n, T_{n+1}$ are chosen as before. We extend our notation slightly, taking a two-dimensional Wiener process $\{W_t; t \in [0, T_n]\}$ under the measure $\mathcal{S}^{(i)}$ for every i. $W^{(1)}$ and $W^{(2)}$ are correlated,

$$d\langle W^{(1)}, W^{(2)} \rangle_t = \rho\, dt.$$

Again $\mathcal{S}^{(i)}$ is the measure associated with the numeraire $D_{tT_{i+1}}$. Similarly, $\sigma^{(i,k)} : [0, T_i] \to \mathbb{R}^+$, $k = 1, 2$, are the volatility functions for the i-th forward LIBOR. Under the numeraire $D_{tT_{n+1}}$ and the associated measure $\mathcal{S}^{(n)}$ the drift of the SDEs (3.3) changes as follows. The proof of the following lemma is a generalization of the one-dimensional case (see [Brigo, Mercurio 2001], Proposition 6.3.1).

Lemma 3.1 (Drift under change of numeraire). *Let the two-dimensional LIBOR Market model be defined by the system of SDEs* (3.3) *as above. Under the measure $\mathcal{S}^{(k)}$, defined by the numeraire $D_{tT_{k+1}}$, the i-th forward LIBOR has the form*

$$
\begin{aligned}
dL_t^{(i)} =& \mu^i(t, L_t^{(i+1)}, \ldots, L_t^{(n)}) L_t^{(i)} dt + \sum_{k=1}^{2} L_t^{(i)} \sigma_t^{(i,k)} dW_t^{(k)}, \mathcal{S}^{(i)}, \\
& t \in [0, \min(T_i, T_k)], L_0^{(i)} = L_i,
\end{aligned}
$$

with drift term

$$
\begin{aligned}
k < i \quad : \quad & \mu^i(t, L_t^{(i+1)}, \ldots, L_t^{(n)}) = \\
& \sum_{j=i+1}^{k} \frac{\alpha_j L_t^{(j)}}{1 + \alpha_j L_t^{(j)}} \big(\sigma_t^{(i,1)} \sigma_t^{(j,1)} + \sigma_t^{(i,2)} \sigma_t^{(j,2)} + \rho\big(\sigma_t^{(i,1)} \sigma_t^{(j,2)} + \sigma_t^{(i,2)} \sigma_t^{(j,1)}\big)\big), \\
k = i \quad : \quad & \mu^i \equiv 0, \\
k > i \quad : \quad & \mu^i(t, L_t^{(i+1)}, \ldots, L_t^{(n)}) = \\
& - \sum_{j=i+1}^{k} \frac{\alpha_j L_t^{(j)}}{1 + \alpha_j L_t^{(j)}} \big(\sigma_t^{(i,1)} \sigma_t^{(j,1)} + \sigma_t^{(i,2)} \sigma_t^{(j,2)} + \rho\big(\sigma_t^{(i,1)} \sigma_t^{(j,2)} + \sigma_t^{(i,2)} \sigma_t^{(j,1)}\big)\big).
\end{aligned}
$$

Proof. Fix some $i \in \{1, \ldots, n\}$ and take $k < i$. The i-th forward rate is driftless under the original numeraire $D_{tT_{i+1}}$. Following [Brigo, Mercurio 2001], Proposition 2.3.2, the drift term μ^i under the numeraire $D_{tT_{k+1}}$ is

$$
\mu^i dt = d\Big\langle \ln L_\cdot^{(i)}, \ln\Big(\frac{D_{\cdot T_{i+1}}}{D_{\cdot T_{k+1}}}\Big)\Big\rangle_t \tag{3.4}
$$

in differential notation, dropping the arguments from the drift term μ^i. Since $k < i$ we have

$$
\begin{aligned}
\ln\Big(\frac{D_{tT_i}}{D_{tT_k}}\Big) &= \ln\Big(\prod_{j=i+1}^{k} \big(1 + \alpha_j L_t^{(j)}\big)\Big) \\
&= \sum_{j=i+1}^{k} \ln\big(1 + \alpha_j L_t^{(j)}\big).
\end{aligned}
$$

Thus

$$
\begin{aligned}
d\Big\langle \ln L_\cdot^{(i)}, \ln\Big(\frac{D_{\cdot T_i}}{D_{\cdot T_k}}\Big)\Big\rangle_t &= d\Big\langle \ln L_\cdot^{(i)}, \sum_{j=i+1}^{k} \ln\big(1+\alpha_j L_\cdot^{(j)}\big)\Big\rangle_t \\
&= \sum_{j=i+1}^{k} d\big\langle \ln L_\cdot^{(i)}, \ln\big(1+\alpha_j L_\cdot^{(j)}\big)\big\rangle_t \\
&= \sum_{j=i+1}^{k} d\big(\ln L_t^{(i)}\big) d\big(\ln\big(1+\alpha_j L_t^{(j)}\big)\big).
\end{aligned}
$$

Before proceeding we evaluate the two differentials using Itô's formula (1.7),

$$
\begin{aligned}
d\big(\ln L_t^{(i)}\big) &= \frac{1}{L_t^{(i)}} dL_t^{(i)} - \frac{1}{2\big(L_t^{(i)}\big)^2} d\big\langle L_\cdot^{(i)}\big\rangle_t \\
&= \frac{1}{L_t^{(i)}} L_t^{(i)} \sum_{m=1}^{2} \sigma_t^{(i,m)} dW_t^{(m)} - (\ldots)dt \\
&= \sum_{m=1}^{2} \sigma_t^{(i,m)} dW_t^{(m)} - (\ldots)dt, \\
d\big(\ln\big(1+\alpha_j L_t^{(j)}\big)\big) &= \frac{\alpha_j}{1+\alpha_j L_t^{(j)}} dL_t^{(j)} - (\ldots)dt.
\end{aligned}
$$

In the equations above $(\ldots)\,dt$ denotes drift terms that have been ignored to simplify calculations. Drift terms do not contribute to the calculation of (3.4), therefore we omit them. Continue with

$$
\sum_{j=i+1}^{k} d\big(\ln L_t^{(i)}\big) d\big(\ln\big(1+\alpha_j L_t^{(j)}\big)\big) = \sum_{j=i+1}^{k} \frac{\alpha_j}{1+\alpha_j L_t^{(j)}} d\big(\ln L_t^{(i)}\big) dL_t^{(j)}.
$$

Evaluating the last differential gives (ignoring drift terms)

$$
\begin{aligned}
d\big(\ln L_t^{(i)}\big) dL_t^{(j)} &= \Big(\sum_{m=1}^{2} \sigma_t^{(i,m)} dW_t^{(m)}\Big)\Big(L_t^{(j)} \sum_{m=1}^{2} \sigma_t^{(j,m)} dW_t^{(m)}\Big) \\
&= L_t^{(j)} \Big(\sum_{m=1}^{2} \sigma_t^{(i,m)} dW_t^{(m)}\Big)\Big(\sum_{m=1}^{2} \sigma_t^{(j,m)} dW_t^{(m)}\Big) \\
&= L_t^{(j)} \Big(\big(\sigma_t^{(i,1)} dW_t^{(1)}\big)\big(\sigma_t^{(j,1)} dW_t^{(1)}\big) + \big(\sigma_t^{(i,2)} dW_t^{(2)}\big)\big(\sigma_t^{(j,1)} dW_t^{(1)}\big) \\
&\quad + \big(\sigma_t^{(i,1)} dW_t^{(1)}\big)\big(\sigma_t^{(j,2)} dW_t^{(2)}\big) + \big(\sigma_t^{(i,2)} dW_t^{(2)}\big)\big(\sigma_t^{(j,2)} dW_t^{(2)}\big)\Big) \\
&= L_t^{(j)} \big(\sigma_t^{(i,1)}\sigma_t^{(j,1)} dt + \sigma_t^{(i,2)}\sigma_t^{(j,2)} dt + \rho\sigma_t^{(i,1)}\sigma_t^{(j,2)} dt + \rho\sigma_t^{(i,2)}\sigma_t^{(j,1)} dt\big) \\
&= L_t^{(j)} \big(\sigma_t^{(i,1)}\sigma_t^{(j,1)} + \sigma_t^{(i,2)}\sigma_t^{(j,2)} + \rho\big(\sigma_t^{(i,1)}\sigma_t^{(j,2)} + \sigma_t^{(i,2)}\sigma_t^{(j,1)}\big)\big) dt.
\end{aligned}
$$

Putting this together yields

$$\begin{aligned}\mu^i dt &= d\Big\langle \ln L_\cdot^{(i)}, \ln\Big(\frac{D_{\cdot T_i}}{D_{\cdot T_k}}\Big)\Big\rangle_t \\ &= \sum_{j=i+1}^{k} \frac{\alpha_j L_t^{(j)}}{1+\alpha_j L_t^{(j)}}\big(\sigma_t^{(i,1)}\sigma_t^{(j,1)} + \sigma_t^{(i,2)}\sigma_t^{(j,2)} + \rho\big(\sigma_t^{(i,1)}\sigma_t^{(j,2)} + \sigma_t^{(i,2)}\sigma_t^{(j,1)}\big)\big)dt.\end{aligned}$$

Now take $k > i$ and note that

$$\begin{aligned}\ln\Big(\frac{D_{tT_{i+1}}}{D_{tT_{k+1}}}\Big) &= \ln\left(\frac{1}{\prod_{j=i+1}^{k}\big(1+\alpha_j L_t^{(j)}\big)}\right) \\ &= -\sum_{j=i+1}^{k} \ln\big(1+\alpha_j L_t^{(j)}\big).\end{aligned}$$

Performing the same calculations as above yields the asserted drift term. This completes the proof. □

Applying Itô's formula (1.7) again finally yields the solution of (3.3)

$$\begin{aligned}L_t^{(i)} =& L_0^{(i)} \exp\Big(\int_0^t \mu^i(s, L_s^{(i+1)}, \ldots, L_s^{(n)}) - \Big(\frac{1}{2}\big(\sigma_s^{(i,1)}\big)^2 + \rho\sigma_s^{(i,1)}\sigma_s^{(i,2)} + \frac{1}{2}\big(\sigma_s^{(i,2)}\big)^2\Big)ds \\ &+ \sum_{k=1}^{2}\int_0^t \sigma_s^{(i,k)} dW_s^{(k)}\Big).\end{aligned}$$

Chapter 4

Markov-functional Models

All interest rate models introduced here share some common assumptions. We always suppose that the uncertainty in the economy can be modeled by a (usually time-changed) Wiener process. This is the cornerstone of all models in this work. The only further information we need are market prices of some products, for example caps or swaptions, for a range of maturities and strikes. From these market prices we derive the covariance function of the underlying Wiener process by statistic methods and we use them to calibrate the model. This means that we set some model parameters such that the market prices of the products in view are replicated exactly. Therefore we may interpret an interest rate model not only as a pricing tool, but rather as an extrapolation tool since we want to determine the price of a given product relative to the prices of some other products traded in the market.
Given market prices as a model input we receive the forward LIBOR or swap rates as a model output. In case of the Markov-functional models we will not model the LIBOR or swap rates directly, but rather a set of zero coupon bonds. The corresponding LIBOR or swap rates may then be computed from these values as shown in the last chapter.

4.1 One-dimensional Markov-Functional Models

Throughout this section we will be working in a single-currency economy $\mathcal{E}$ comprising of a set of pure discount bonds D_{tT}. It will be enough to consider a finite number of bonds D_{tT}, $T \in \mathcal{T}$, where $\mathcal{T} = \{T_i; 1 \leq i \leq n\}$. We assume further that the economy admits a numeraire pair $(N, \mathcal{N})$, consisting of a numeraire N and an equivalent martingale measure $\mathcal{N} \sim \mathcal{P}$. Thus the economy is arbitrage free. Our products in view will share a common terminal time which we denote by T_{n+1}.
The models we are working with throughout this chapter are numeraire models.

This means that the term structure is defined via

$$D_{tT} = N_t E_{\mathcal{N}}[N_T^{-1} | \mathcal{F}_t^W], T \in \mathcal{T},$$

where N is the chosen numeraire and $\mathcal{N}$ the associated measure as mentioned above. See also [Hunt, Kennedy 2000].
It then follows as already laid out in section 1.3 that since the economy $\mathcal{E}$ is arbitrage free, the value V_t of any derivative at time $t < T$ is given by

$$V_t = N_t E_{\mathcal{N}}[V_T N_T^{-1} | \mathcal{F}_t^W].$$

Central to the approach of Markov-Functional models is the assumption, that the uncertainty in the economy can be captured by a low-dimensional Markov process $\{X_t;\ t \in [0,T]\}$. This is reasonable if we assume that *all* information available about an asset at some time t is contained in the asset price at time t (as it is usually done in economics). The dimension of the Markov process determines the dimension of the model. In this section we will work with a one-dimensional Markov process $\{X_t;\ t \in [0,T]\}$.
In the following definition we introduce the rather technical boundary curve. The idea behind this is the following. We will not model the whole term structure D_{tT}, $t \leq T$, $T \in \mathcal{T}$, form beginning. Instead we will model just a small number of bonds $D_{t'T'}$ for some $0 \leq t' \leq T'$ and recover the remaining ones by discounting. The time T' for which the bonds will be modeled is determined by the boundary curve. The following explanations will shed some light on this remark.

Definition 4.1. *Let $\{X_t;\ t \in [0,T]\}$ be a (time-inhomogeneous) Markov process under the measure $\mathcal{N}$. Then the pure discount bond prices D_{tS} are of the form*

$$D_{tS} = D_{tS}(X_t), \quad 0 \leq t \leq \partial(S) \leq S$$

for some boundary curve $\partial : [0, \partial^*] \to [0, \partial^*], \partial^* \in \mathbb{R}^+$, and the numeraire, itself a price process, is of the form

$$N_t = N_t(X_t), \quad 0 \leq t \leq \partial^*.$$

Therefore the processes $\{D_{tS};\ t \in [0,T]\}$ and $\{N_t;\ t \in [0,T]\}$ are defined to be functionals of $\{X_t;\ t \in [0,T]\}$.

The boundary curve $\partial_S : [0, \partial^*] \to [0, \partial^*]$ will be chosen to be appropriate for the particular pricing problem under consideration. In the examples we will discuss later the products of interest share some common terminal time T and the curve ∂_S is taken to be

$$\partial(S) := \begin{cases} S & \text{if} \quad S \leq T \\ T & \text{if} \quad S > T. \end{cases} \tag{4.1}$$

For all practical applications our models will have a boundary curve given by (4.1). From this it is clear that for a complete specification of the model it is sufficient to know:

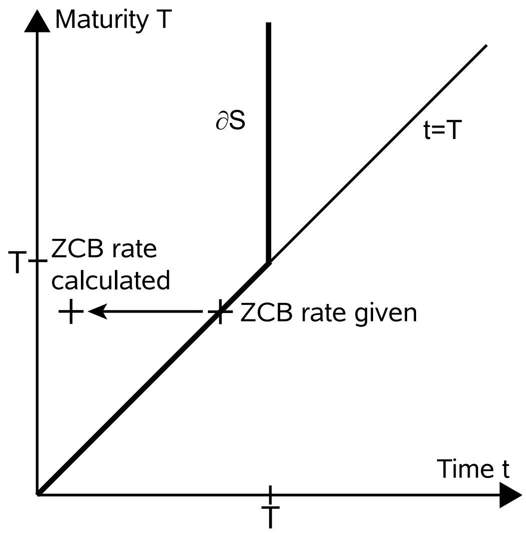

Figure 4.1: The boundary curve

(P.1) the distribution of the process $\{X_t; t \in [0, \partial^*]\}$ under $\mathcal{N}$,

(P.2) the functional form of the discount factors $D_{\partial(S)S}(X_{\partial(S)})$ on the boundary $\partial(S)$, $S \in [0, \partial^*]$,

(P.3) the functional form of the numeraire $N_t(X_t)$ for $t \in [0, \partial^*]$.

From this we can recover the remaining discount bond prices via the martingale property for numeraire rebased assets under $\mathcal{N}$,

$$D_{tS}(X_t) = N_t(X_t) E_{\mathcal{N}}\Big[\frac{D_{\partial(S)S}(X_{\partial(S)})}{N_{\partial(S)}(X_{\partial(S)})} | \mathcal{F}_t^W\Big] \tag{4.2}$$

Now assume we want to price a multi-temporal exotic product, for instance a Bermudan swaption, which depends on a set of forward swap rates or forward LIBORs, each observed at a distinct time. From Chapter 2 it is known that both forward rates can be recovered from pure discount bond prices. Therefore it is sufficient to model the pure discount bonds and we may assume that a set of swap rates $S^{(i)}$ is given. One could also set up a model working with a set of forward LIBORs $L^{(i)}$ instead. This is also possible and works along the same lines as for a set of swap rates $S^{(i)}$.
We will make two further assumptions before starting:

(P.4) the functional forms $D_{T_n S}$ are known for $T_n \leq S$ and the choice of the numeraire is such that $N_{T_n}(X_{T_n})$ can be inferred form the discount factors on the boundary, (P.5) the i-th forward swap rate at time T_i, $S_{T_i}^{(i)}$, is a strictly monotonic increasing function of X_{T_i} in the sense that $X_{T_i}(\omega) > X_{T_i}(\omega') \Rightarrow S_{T_i}^{(i)}(X_{T_i}(\omega)) > S_{T_i}^{(i)}(X_{T_i}(\omega'))$ for some $\omega, \omega' \in \Omega$ and $i = 1, \ldots, n$.

We shall see soon that assumption (P.4) is not as restrictive as one would think . The last assumption is a natural one since one wants the overall level of interest rates respectively swap rates to depend on the level of $\{X_t;\ t \in [0,T]\}$. This is crucial for the model to work and must be preserved when turning to the multi-dimensional case.

In the remainder of this section we show how market prices of the calibrating vanilla swaptions can be used to imply, numerically at least, the functional form N_{T_i} for $i = 1, \ldots, n-1$.

Equivalent to calibrating the model to vanilla swaptions is to calibrate it to the inferred market prices of digital swaptions (cf. [Dupire 1994]). This has the following interesting consequence:

Choose the PV per basis point $\{P_t; t \in [0,T]\}$ as a numeraire as before and denote the corresponding measure $\mathcal{S}$. The price of a European payers swaption on $\{S_t^{(i)}[T_i,T_n]; t \in [0,T_i]\}$ maturing on date T_i with strike K is given by

$$V_0^{Payers}(K) = P_0 E_{\mathcal{S}}[(S_{T_i}^{(i)} - K)^+].$$

Differentiating both sides with respect to K, taking only left-hand limits for $S_T \leq K$, yields

$$\frac{\partial V_0(K)}{\partial K} = P_0(-E_{\mathcal{S}}[\mathbf{1}_{\{S_T > K\}}]) = -P_0 \mathcal{S}(S_T > K),$$

which is the price of the corresponding digital swaption. Thus the swaption prices have allowed us to recover the *implied* distribution for y_T under the measure $\mathcal{S}$. This observation remains valid if the swaption prices are not log-normally distributed and incorporate smile effects.

The digital swaption based on $\{S_t^{(i)}[T_i,T_n]; t \in [0,T_i]\}$ with strike K has payoff at time T_i of

$$V_{T_i}^{(i)}(K) = P_{T_i}^{(i)} \mathbf{1}_{\{S_{T_i}^{(i)} > K\}},$$

where $P_{T_i}^{(i)}$ denotes the accrual factor of the payoff. The price of the option at time 0 is then given by

$$V_0^{(i)}(K) = N_0(X_0) E_{\mathcal{N}}\Big[\frac{P_{T_i}^{(i)}(X_{T_i})}{N_{T_i}(X_{T_i})} \mathbf{1}_{\{S_{T_i}^{(i)} > K\}}\Big] \tag{4.3}$$

To determine the functional form of $N_{T_i}(X_{T_i})$ we work back iteratively from the terminal time T_n. Consider the i-th step of this procedure. Assume that $N_{T_k}(X_{T_k}), k = i+1, \ldots, n$, have already been determined. We can also assume

$D_{T_iT_n}$ are known having been determined using (4.2) and the known conditional distributions of $X_{T_k}, k = i, \ldots, n$.
Now consider $S_{T_i}^{(i)}[T_i, T_n]$ which can be written as

$$\begin{aligned} S_{T_i}^{(i)}[T_i, T_n] &= \frac{1 - D_{T_iT_n}}{P_{T_i}^{(i)}[T_i, T_n]} \\ &= \frac{N_{T_i}^{-1} - D_{T_iT_n} N_{T_i}^{-1}}{P_{T_i}^{(i)}[T_i, T_n] N_{T_i}^{-1}}, \end{aligned} \tag{4.4}$$

since $D_{T_iT_i} = 1$. Rearranging equation (4.4) we get

$$N_{T_i}(X_{T_i}) = \frac{1}{P_{T_i}^{(i)} N_{T_i}^{-1} S_{T_i}^{(i)}(X_{T_i}) + D_{T_iT_n} N_{T_i}^{-1}},$$

dropping $[T_i, T_n]$ from the notation. Thus to determine $N_{T_i}(X_{T_i})$ it is sufficient to determine the functional form $S_{T_i}^{(i)}(X_{T_i})$.
By assumption (P.5) there exists a unique value of K, say $K^{(i)}(x^*)$, such that the set identity

$$\{X_{T_i} > x^*\} = \{S_{T_i}^{(i)}(X_{T_i}) > K^{(i)}(x^*)\} \tag{4.5}$$

holds $\mathcal{N}$-almost surely for every $x^* \in \mathbb{R}$. The set identity (4.5) is the key to implying the functional form $S_{T_i}^{(i)}(X_{T_i})$. Define

$$\begin{aligned} J_0^{(i)}(x^*) &:= N_0(X_0) E_{\mathcal{N}} \Big[\frac{P_{T_i}^{(i)}(X_{T_i})}{N_{T_i}(X_{T_i})} \mathbf{1}_{\{S_{T_i}^{(i)}(X_{T_i}) > K^{(i)}(x^*)\}} \Big] \\ &= N_0(X_0) E_{\mathcal{N}} \Big[\frac{P_{T_i}^{(i)}(X_{T_i})}{N_{T_i}(X_{T_i})} \mathbf{1}_{\{X_{T_i} > x^*\}} \Big] \end{aligned} \tag{4.6}$$

For any given $x^* \in \mathbb{R}$ we can calculate $J_0^{(i)}(x^*)$ using the known distribution of X_{T_i} under $\mathcal{N}$. Further, using market prices of digital swaptions we can find the value of K such that

$$J_0^{(i)}(x^*) = V_0^{(i)}(K). \tag{4.7}$$

Comparing (4.3) and (4.6) we can see that the value of K satisfying (4.7) is exactly $K^{(i)}(x^*)$. From (4.5) knowing $K^{(i)}(x^*)$ for any x^* is equivalent to knowing the functional form of $S_{T_i}^{(i)}(X_{T_i})$ and we are done. We could have decided to calibrate the model to vanilla swaptions instead of digital swaptions, but then it would not be clear whether a simple set identity similar to (4.5) holds. Note that we did not specify the functional form for times $t \notin \mathcal{T}$. This leaves enough freedom to capture the distribution of the market prices.
It is common market practice to use Black's formula to determine the swaption prices $V_0^{(i)}(K)$ but these techniques apply more generally. In particular smile and skew effects which stem from a not-log-normal distribution in the market are calibrated automatically. This is a major advantage compared to other interest rate models.

4.1.1 Market models and Markov-functional models

Market models present a general framework for modeling interest rate derivatives and as such Markov-functional models *are a subset thereof.* At this level Market models are just a different parametrization. However, Market models make the additional assumption that the forward LIBOR respectively swap rates are martingales in their respective measures. This assumption is stronger than ours since we only assume the martingale property on the respective fixing dates. This additional restriction in the Market models is what makes them difficult because it yields the state-dependent dynamics and makes it impossible to characterize them by a low-dimensional Markov process. Hunt and Kennedy showed in [Hunt, Kennedy 2000] that it is not possible to characterize a Market model consisting of n forward LIBORs by a one-dimensional Markov process. It is assumed, yet not proved, that the same is true for any Markov process of dimension less than n. However, when removing the state-dependency by the technique of drift approximation the resulting model is no longer arbitrage-free (because the resulting processes are in general no longer martingales in their respective measures), *but can be characterized by a low-dimensional Markov process.* These drift approximations will be used when introducing multi-dimensional Markov functional models as they approximate the LIBOR process, thus being a valuable guideline in constructing these models.

4.1.2 One-dimensional LIBOR model

In this subsection and the next we introduce two example models which can be used to value LIBOR and swap rate based interest rate derivatives.
Let $\{L_t^{(i)}; t \in [0, T_i]\}, i = 1, \ldots, n$, be a set of forward LIBORs. We model the pure discount bonds $D_{tT_i}, i = 1, \ldots, n+1$, the assets in our economy and derive the LIBORs using the known formula

$$L_t^{(i)} = \frac{1 - D_{tT_i}}{\alpha D_{tT_i}}$$

with α denoting the appropriate daycount fraction. Taking $D_{tT_{n+1}}$ as a numeraire we denote the corresponding equivalent measure by $\mathcal{S}^{(n)}, \mathcal{S}^{(n)} \sim \mathcal{P}$. Under this measure the $D_{tT_{n+1}}$-rebased assets $\{\frac{D_{tT_i}}{D_{tT_{n+1}}}; t \in [0, T_i]\}$ are martingales and the model is arbitrage free.
As laid out before we have to specify the properties (P.1)-(P.3). To be consistent with Black's formula for caplets (2.10) on $L^{(n)}$ we assume that $L^{(n)}$ is a log-normal martingale under $\mathcal{S}^{(n)}$, i.e.,

$$dL_t^{(n)} = L_t^{(n)} \sigma_t^{(n)} dW_t, \mathcal{S}^{(n)}, t \in [0, T_n], L_0^{(n)} = L_n, \tag{4.8}$$

where W_t is a standard Wiener process under $\mathcal{S}^{(n)}$ and $L_n \in \mathbb{R}^+$. $\sigma_t^{(n)} : [0, T_n] \to \mathbb{R}^+$ is the volatility function and may be specified arbitrarily at the moment as long as equation (4.8) has a solution. It follows from (4.8) and an application of Itô's formula (1.7) that

$$L_t^{(n)} = L_0^{(n)} \exp\Big(-\frac{1}{2}\int_0^t (\sigma_u^{(n)})^2 du + X_t\Big),$$

where X_t satisfies

$$dX_t = \sigma_t^{(n)} dW_t, \mathcal{S}^{(n)}, t \in [0,T], X_0 = 0. \tag{4.9}$$

$\{X_t; t \in [0,T]\}$ is therefore a deterministic time-change of the Wiener process. Further, $\{X_t; t \in [0,T]\}$ is Markovian as stated by the following theorem. This theorem and its proof may be seen as a supplement to [Hunt, Kennedy 2000], where the Markov property of $\{X_t; t \in [0,T]\}$ and the form of the transition semigroup have been stated, yet not proved.

Theorem 4.2 (Markov property of X). *Let $\sigma_t : [0,T] \to \mathbb{R}^+$ be a deterministic function of time, $T \in \mathbb{R}^+$. Then for a one-dimensional Wiener process $\{W_t; t \in [0,T]\}$ under some measure $\mathcal{P}$ the stochastic process $\{X_t; t \in [0,T]\}$ given by*

$$dX_t = \sigma_t dW_t, \mathcal{P}, t \in [0,T], X_0 = 0,$$

is a Markov process with transition semigroup

$$(K_t^X f)(x) = \int \phi_t(x-y) f(y) d\lambda(y),$$

where

$$\phi_t(x) := (2\pi\Sigma_{0;t})^{-\frac{1}{2}} \exp\Big(-\frac{x^2}{2\Sigma_{0;t}}\Big)$$

and

$$\Sigma_{s;t} := \int_s^t (\sigma_u)^2 du$$

for any non-negative Borel-measurable function $f : \mathbb{R} \to \mathbb{R}, f \geq 0$.

Proof. It is known that $\langle X \rangle_t = \langle \int_0^\cdot \sigma_u^2 dW_s \rangle_t = \int_0^t \sigma_u^2 d\langle W \rangle_u = \int_0^t \sigma_u^2 du$, since $\langle W \rangle_t = t$ for all $t \geq 0$. Hence $\lim_{t\to\infty} \langle X \rangle_t = \infty$. Further, $\{X_t; t \in [0,T]\}$ is a continuous adapted martingale and we may apply Knight's theorem 1.49. This gives $X_t = W_{<X>_t}$ for every $t \geq 0$. Set $\tau(t) := \langle X \rangle_t$. Then

$$\begin{aligned} E_{\mathcal{P}}[f \circ X_t | \mathcal{F}_s^X] &= E_{\mathcal{P}}[f \circ W_{\tau(t)} | \mathcal{F}_{\tau(s)}^W] \\ &= (P_{\tau(t)-\tau(s)} f)(W_{\tau(s)}) \end{aligned}$$

holds for any $s, t \in [0,T], s < t$, and Borel-measurable $f : \mathbb{R} \to \mathbb{R},\ f \geq 0$. Since the transition semi-group of the one-dimensional Wiener process is given by

$$\begin{aligned} (K_t^W f)(x) &= \int \phi_t(x-y) f(y) d\lambda^1(y), \\ \phi_t(x) &= (2\pi t)^{-\frac{1}{2}} \exp\Big(-\frac{x^2}{2t}\Big) \text{ for } t > 0, \end{aligned}$$

it follows that

$$(K^W_{\tau(t)-\tau(s)}f)(x) = \int \phi_{\tau(t)-\tau(s)}(x-y)f(y)d\lambda^1(y)$$

and

$$\phi_{\tau(t)-\tau(s)}(x) = (2\pi(\tau(t)-\tau(s)))^{-\frac{1}{2}} \exp\Big(-\frac{x^2}{2(\tau(t)-\tau(s))}\Big).$$

From the definition of $\tau(t)$ we have

$$\tau(t)-\tau(s) = \int_s^t \sigma_u^2 du =: \Sigma_{s;t}.$$

Back in the original time-coordinates this gives

$$(K^W_{\tau(t)-\tau(s)}f)(W_{\tau(s)}) =: (K^X_{t-s}f)(X_s),$$

where the transition semi-group K^X_t is given by

$$\begin{aligned}(K^X_t f)(x) &= \int \phi_t(x-y)f(y)d\lambda^1(y),\\ \phi_t(x) &= (2\pi\Sigma_{0;t})^{-\frac{1}{2}}\exp\Big(-\frac{x^2}{2\Sigma_{0;t}}\Big).\end{aligned}$$

This completes the proof. □

We take $\{X_t; t \in [0,T]\}$ as the driving Markov process for our model and this completes the specification of (P.1).
For this application we use the boundary curve (4.1). Thus the only functional forms needed for pure discount bonds are $D_{T_iT_i}(X_{T_i}), i = 1,\dots,n$, trivially the unit map, and $D_{T_nT_{n+1}}(X_{T_n})$, the numeraire. The latter follows from the relationship

$$D_{T_nT_{n+1}}(X_{T_n}) = \frac{1}{1+\alpha_n L^{(n)}_{T_n}},$$

which in turn yields

$$D_{T_nT_{n+1}}(X_{T_n}) = \frac{1}{1+\alpha_n L^{(n)}_0 \exp\Big(-\frac{1}{2}\int_0^{T_n}(\sigma^{(n)}_u)^2du + X_{T_n}\Big)}.$$

This completes (P.2).
It remains to find the remaining functional form of $N_{T_i}(X_{T_i}) = D_{T_iT_{n+1}}(X_{T_i})$ for $i = 1,\dots,n-1$. For this we work inductively backward from N_{T_n} to N_{T_1} and apply the techniques of Section 4.1. Since we are constructing a model of the forward LIBORs to value LIBOR based products we must choose some appropriate products traded in the market to calibrate our model (setting the

model parameters). Take the caplet on the forward LIBOR $L^{(i)}$ with strike K as calibrating instruments. The payoff of the corresponding i-th digital caplet at time T_{i+1} (remember that the LIBOR $L^{(i)}$ resets on date T_i and the caplet pays on date T_{i+1}) is $D_{T_iT_{i+1}}\mathbf{1}_{\{L^{(i)}_{T_i}>K\}}$. The digital caplet pays $\mathbf{1}_{\{L^{(i)}_{T_i}>K\}}$ and this value is discounted to time T_i since we receive the payment on date T_{i+1}, but we need the present value on date T_i. Using the Fundamental Theorem of Asset Pricing 1.74 the value of the i-th corresponding digital caplet with strike K at time 0 is given by

$$\begin{aligned} V_0^{(i)}(K) &= N_0(X_0)E_{\mathcal{S}^{(n)}}\Big[\frac{D_{T_iT_{i+1}}(X_{T_i})}{N_{T_i}(X_{T_i})}\mathbf{1}_{\{L^{(i)}_{T_i}(X_{T_i})>K\}}\Big] \\ &= D_{0T_{n+1}}(X_0)E_{\mathcal{S}^{(n)}}\Big[\frac{D_{T_iT_{i+1}}(X_{T_i})}{N_{T_i}(X_{T_i})}\mathbf{1}_{\{L^{(i)}_{T_i}(X_{T_i})>K\}}\Big]. \end{aligned}$$

If we assume that the market price is given by Black's formula (2.10) with volatility function $\sigma_t^{(i)}$, the price at time zero for this digital is

$$V_0^{(i)}(K) = D_{0T_{n+1}}(X_0)\Phi(d_2), \tag{4.10}$$

where

$$d_2 = \frac{\log(\frac{L_0^{(i)}}{K})}{\Sigma^{(i)}_{0;T_i}\sqrt{T_i}} - \frac{1}{2}\Sigma^{(i)}_{0;T_i}\sqrt{T_i}$$

and

$$\Sigma^{(i)}_{0;T} = \frac{1}{T}\int_0^T(\sigma_u^{(i)})^2du.$$

To determine the functional forms $D_{T_iT_{n+1}}(X_{T_i})$ for $i<n$ we proceed as in Section 4.1. Choose some $x^*\in\mathbb{R}$. Evaluate by numerical integration

$$\begin{aligned} J_0^{(i)}(x^*) &= D_{0T_{n+1}}(X_0)E_{\mathcal{S}^{(n)}}\Big[\frac{D_{T_iT_{i+1}}(X_{T_i})}{D_{T_iT_{n+1}}(X_{T_i})}\mathbf{1}_{\{X_{T_i}>x^*\}}\Big] \\ &= D_{0T_{n+1}}(X_0)E_{\mathcal{S}^{(n)}}\Big[E_{\mathcal{S}^{(n)}}\Big[\frac{D_{T_{i+1}T_{i+1}}(X_{T_{i+1}})}{D_{T_{i+1}T_{n+1}}(X_{T_{i+1}})}|\mathcal{F}^W_{T_i}\Big]\mathbf{1}_{\{X_{T_i}>x^*\}}\Big] \\ &= D_{0T_{n+1}}(X_0)\int_{-\infty}^{\infty}\Big[\int_{-\infty}^{\infty}\frac{1}{D_{T_{i+1}T_{n+1}}(u)}K_{T_i;T_{i+1}}(v-u)du\Big]\mathbf{1}_{\{v>x^*\}}\phi_{T_i}(v)dv \\ &= D_{0T_{n+1}}(X_0)\int_{x^*}^{\infty}\Big[\int_{-\infty}^{\infty}\frac{1}{D_{T_{i+1}T_{n+1}}(u)}K_{T_i;T_{i+1}}(v-u)du\Big]\phi_{T_i}(v)dv, \end{aligned}$$

where ϕ_{T_i} denotes the density function of X_{T_i} given by

$$\phi_t(x) = (2\pi\Sigma^{(i)}_{0;t})^{-\frac{1}{2}}\exp\Big(-\frac{x^2}{2\Sigma^{(i)}_{0;t}}\Big),$$

and $K_{T_i;T_{i+1}}(x)$, the transition density function of X_{T_i} when moving to $X_{T_{i+1}}$, is given by

$$K_{T_i;T_{i+1}}(x) = (2\pi\Sigma^{(i)}_{T_i;T_{i+1}})^{-\frac{1}{2}} \exp\Big(-\frac{x^2}{2\Sigma^{(i)}_{T_i;T_{i+1}}}\Big),$$

where

$$\Sigma^{(i)}_{s;t} = \int_s^t (\sigma^{(i)}_u)^2 du.$$

Note that in the above calculations in the next to last step the Markov property of $\{X_t;\ t \in [0,T]\}$ has been applied. Note further that the integrand depends only on $D_{T_{i+1}T_{n+1}}(X_{T_i})$, which has been determined in the previous iteration. The value of $D_{T_{i+1}T_{n+1}}(x^*)$ can now be determined as follows. Recall from equation (4.4) that to determine $N_{T_i}(x^*) = D_{T_{i+1}T_{n+1}}(x^*)$ it is sufficient to find the functional form $L^{(i)}_{T_i}(x^*)$. From equations (4.5) and (4.7) it follows that

$$L^{(i)}_{T_i}(x^*) = K^{(i)}(x^*),$$

where $K^{(i)}(x^*)$ solves

$$J^{(i)}_0(x^*) = V^{(i)}_0(K^{(i)}(x^*)). \tag{4.11}$$

We have just evaluated the left hand side of (4.11) and $K^{(i)}(x^*)$ can be obtained from (4.11) by solving the resulting equation. Formally

$$L^{(i)}_{T_i}(x^*) = L^{(i)}_0 \exp\Big(-\frac{1}{2}\Sigma^{(i)}_{0;T_i}T_i - \Sigma^{(i)}_{0;T_i}\sqrt{T_i}\Phi^{-1}\Big(\frac{J^{(i)}_0(x^*)}{D_{0T_{i+1}}(x_0)}\Big)\Big).$$

Finally, use equation (4.4) to obtain the value of $D_{T_iT_{n+1}}(x^*)$.

4.1.3 One-dimensional Swap model

For the construction of a swap model we consider the special case of a swap for which the i-th forward swap rate $S^{(i)}$, which sets on date T_i, has coupons precisely on dates $T_{i+1}, \ldots, T_n$. In this case the last swap rate $S^{(n)}$ is just the LIBOR $L^{(n)}$ for the period $[T_n, T_{n+1}]$. As in the LIBOR model we take $D_{tT_{n+1}}$ as our numeraire and assume that the forward measure $\mathcal{S}^{(n)}$ exists. The properties (P.1) and (P.2) are specified exactly as in the LIBOR model. However, the functional form for the numeraire $D_{tT_{n+1}}$ at times $T_i, i = 1, \ldots, n-1$, will need to be determined. Analogous to the LIBOR model we construct the swap model for valuing swap rate based products and must therefore choose some swap rate based, calibrating instruments from the market. Here we take the digital swaption on the swap rate $S^{(i)}$ having strike K. This digital swaption has the payoff $P_{T_i}\mathbf{1}_{\{S^{(i)}_{T_i}>K\}}$ and the value at time 0 is thus given by (applying the Fundamental Theorem of Asset

Pricing 1.74)

$$V_0^{(i)}(K) = D_{0T_{n+1}}(X_0) E_{\mathcal{S}^{(n)}}\Big[\frac{P_{T_i}^{(i)}(X_{T_i})}{D_{T_i T_{n+1}}(X_{T_i})} \mathbf{1}_{\{S_{T_i}^{(i)}(X_{T_i})>K\}}\Big],$$

where

$$P_t^{(i)}(X_t) = \sum_{j=i}^{n} \alpha_j D_{tT_{j+1}}(X_t)$$

denotes the PVBP. Note that $P_t^{(i)}$ is a linear combination of assets and therefore an asset itself. If we assume the market value is given by Black's formula (2.14), the price at time zero of the digital swaption is

$$V_0^{(i)}(K) = P_0^{(i)}(X_0)\Phi(d_2), \tag{4.12}$$

where

$$d_2 = \frac{\log(\frac{L_0^{(i)}}{K})}{\Sigma_{0;T_i}^{(i)}\sqrt{T_i}} - \frac{1}{2}\Sigma_{0;T_i}^{(i)}\sqrt{T_i}$$

and

$$\Sigma_{0;T}^{(i)} = \frac{1}{T}\int_0^T (\sigma_u^{(i)})^2 du.$$

Next choose some $x^* \in \mathbb{R}$ and evaluate by numerical integration

$$\begin{aligned}
J_0^{(i)}(x^*) &= D_{0T_{n+1}}(X_0) E_{\mathcal{S}^{(n)}}\Big[\frac{P_{T_i}^{(i)}(X_{T_i})}{D_{T_i T_{n+1}}(X_{T_i})} \mathbf{1}_{\{X_{T_i}>x^*\}}\Big] \\
&= D_{0T_{n+1}}(X_0) E_{\mathcal{S}^{(n)}}\Big[E_{\mathcal{S}^{(n)}}\Big[\frac{P_{T_{i+1}}^{(i)}(X_{T_{i+1}})}{D_{T_{i+1}T_{n+1}}(X_{T_{i+1}})}|\mathcal{F}_{T_i}^W\Big] \mathbf{1}_{\{X_{T_i}>x^*\}}\Big] \\
&= D_{0T_{n+1}}(X_0) \int_{-\infty}^{\infty}\Big[\int_{-\infty}^{\infty} \frac{P_{T_{i+1}}^{(i)}(u)}{D_{T_{i+1}T_{n+1}}(u)} K_{T_i;T_{i+1}}(v-u)du\Big] \\
&\quad \mathbf{1}_{\{v>x^*\}}\phi_{T_i}(v)dv \\
&= D_{0T_{n+1}}(X_0) \int_{x^*}^{\infty}\Big[\int_{-\infty}^{\infty} \frac{P_{T_{i+1}}^{(i)}(u)}{D_{T_{i+1}T_{n+1}}(u)} K_{T_i;T_{i+1}}(v-u)du\Big]\phi_{T_i}(v)dv.
\end{aligned}$$

$\frac{P_{T_{i+1}}^{(i)}(u)}{D_{T_{i+1}T_{n+1}}(u)}$ is strictly positive and the Markov property has been applied as in the LIBOR model. $\phi_t(x)$ and $K_{T_i;T_{i+1}}(y)$ are also defined as before.
To calculate a value for $J_0^{(i)}(x^*)$ we need to evaluate $D_{T_{i+1}T_j}(x^*)$ for $j > i$. These will have already been determined in the previous iteration. Now proceed as in Section 4.1.2 with

$$S_{T_i}^{(i)}(x^*) = K^{(i)}(x^*),$$

where $K^{(i)}(x^*)$ solves

$$J_0^{(i)}(x^*) = V_0^{(i)}(K^{(i)}(x^*)). \tag{4.13}$$

Having evaluated the left hand side of (4.13) numerically, $K^{(i)}(x^*)$ can be recovered from (4.12). Formally we have

$$S_{T_i}^{(i)}(x^*) = S_0^{(i)} \exp\Big(-\frac{1}{2}\Sigma_{0;T_i}^{(i)} T_i - \Sigma_{0;T_i}^{(i)} \sqrt{T_i}\Phi^{-1}\Big(\frac{J_0^{(i)}(x^*)}{P_0^{(i)}(x_0)}\Big)\Big).$$

The value of $D_{T_i T_{n+1}}(x^*)$ can now be calculated using (4.4).

4.2 Multi-dimensional Markov-functional models

Having introduced one-dimensional Markov-functional models we come to the core of this work, the Multi-dimensional Markov-functional models. The first point to clarify if whether one would like to increase the dimension of the driving Markov process and thus the dimension of the model.

When observing the daily changes of interest or swap rates three different main effects contributing to the changes of the interest rate or swap rate curve can be identified: parallel shift, rotation of the curve and twisting. A parallel shift raises or lowers the overall level of rates, a rotation of the curve lowers the rates for the shorter maturities and increases the rates for the longer maturities, twisting the curve makes it steeper or flatter, possibly leading to an inverted term structure for the longer maturities. These effects occur and must be modeled separately.

Simple products like caps and floors depend only on the overall level of interest rates, therefore it is sufficient to price caps and floors with a one-factor model, allowing only for parallel shifts of the curve. There are, however, more complex products depends not only on the overall level of rates, but for example on the steepness of the curve. To price these products correctly we must set up a model comprising two or three factors to catch more aspects of the changes of interest or swap rates. So far there is nothing that could not be done using some market model, taking $\{f^{(i)}(W_t); t \in [0,T]\}$ as the driving stochastic process for the i-th LIBOR rate respectively swap rate, where $\{W_t; t \in [0,T]\}$ is a two-dimensional Wiener process under some probability measure $\mathcal{P}$ and $f^{(i)} : \mathbb{R}^2 \to \mathbb{R}$ is a functional still to be specified. However, this yields some difficulties.

First, as already lined out, it is in principle possible to value products with early-exercise features like Bermudan swaptions but it is time-consuming and very cumbersome. Before valuing the product the exercise boundary has to be approximated using non-recombining trees, which is computationally expensive. It is possible to circumvent this problem and to use recombining trees, which are computationally far less expensive, but the approximation will be less accurate.

Second, there is no simple way to implement a non-lognormal distribution and thus to allow for smile and skew effects. These effects, however, have a huge impact on the option price, so one is very interested to consider these effects when designing a model. The usual way to accomplish this is to set up an SDE having the desired dynamics and to use a numerical scheme like the Euler- or Milstein-Scheme to get a numerical approximation. The question is whether there is a simpler yet computationally efficient way to construct a model with the desired properties.
Having this in mind it seems reasonable to extend the one-dimensional Markov-functional model in an appropriate way to allow for a higher-dimensional driving Markov process. Of course, all other features of the model should be preserved as there are:

- Automatic calibration to market data under consideration of non-lognormal distributions of the market prices.
- Relatively simple valuation of early-exercise products
- Fast numerical implementation.

The last point is also of great importance since it is unacceptable for a trader to wait several hours or even minutes when pricing an option. The numerical evaluation must be done within a few seconds, the less, the better. The one-dimensional Markov-functional model as proposed here allows for a fast numerical implementation (use, for example, Gauss quadrature formulae).
Our goal is therefore to construct a multi-dimensional Markov-functional model in such a way to preserve the three features mentioned above. This is done in the following section. Especially the set identity (4.5), which is the key to the further development, shall remain unchanged. Therefore we cannot simply choose an n-dimensional driving Markov process $\{X_t;\, t \in [0,T]\}$ as the set identity (4.5) would no longer hold. Instead, we choose a functional $f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ and take $\{f(X_t,t);\, t \in [0,T]\}$ as our driving process (which may not be Markovian any more). The choice of the functional f will be closely related to the drift approximations to Market models mentioned in the previous section.
The ideas presented in this section are based on a Markov process of dimension two, which is the case with the highest practical importance. They may easily be generalized to the case of higher dimensions.
Let $\{W_t; t \in [0,T]\}$ be a standard two-dimensional Wiener process,

$$W_t = \begin{pmatrix} W_t^{(1)} \\ W_t^{(2)} \end{pmatrix}, t \in [0,T],$$

where $\{W_t^{(1)}; t \in [0,T]\}$ and $\{W_t^{(2)}; t \in [0,T]\}$ are two independent, standard one-dimensional Wiener processes. Contrary to the one-dimensional case we are more limited in our choice of the volatility functions $\sigma_t^{(i)}$. One more restriction has to be imposed, separability.

Definition 4.3 (Separability). *A set of n volatility functions* $\sigma^{(i)} : [0,T_i] \to \mathbb{R}^d, i = 1,\ldots,n, n,d \in \mathbb{N}$, $T_1 < T_2 < \ldots < T_n \le T, T_i, T \in \mathbb{R}^+$, *is called* separable *if there exists a function* $\sigma : [0,T] \to \mathbb{R}$ *and vectors* $v_i \in \mathbb{R}^d, i = 1,\ldots,n$, *such that*

$$\sigma^{(i)}(t) = v_i \sigma(t)$$

for $0 \le t \le T_i, i = 1,\ldots,n$.

Assume we have chosen a set of separable $\mathbb{R}^2$-valued volatility functions $\sigma^{(i)}(t)$ as in Def. 4.3. Equivalently, we may choose a function $\sigma : [0,T] \to \mathbb{R}$ and vectors $v_i \in \mathbb{R}^2, i = 1,\ldots,n$. Then the stochastic process

$$dZ_t = \text{diag}(\sigma(t),\sigma(t))dW_t, \mathcal{P}, t \in [0,T], Z_0 = 0, \tag{4.14}$$

is Markovian and $\mathbb{R}^2$-valued.

Theorem 4.4 (Markov property of Z_t). *Let* $\sigma : [0,T] \to \mathbb{R}$, $T \in \mathbb{R}^+$ *and* $v_i \in \mathbb{R}^2, i = 1,\ldots,n$, *be a deterministic function of time and* v_i *vectors, such that* $\sigma^{(i)}(t) := v_i\sigma(t)$ *is a separable set of volatility functions. Then for a 2-dimensional Wiener process* $\{W_t; t \in [0,T]\}$ *under some measure* $\mathcal{P}$

$$dZ_t = \text{diag}(\sigma(t),\sigma(t))dW_t, \mathcal{P}, t \in [0,T], Z_0 = 0, \tag{4.15}$$

is a Markov process with transition semigroup $(K_t^Z)_{t\ge 0}$,

$$(K_t^Z f)(x) = \int \phi_t(x-y)f(y)d\lambda^2(y)$$

for any non-negative function $f : \mathbb{R}^2 \to \mathbb{R}^+$.
$\phi_t(x)$ *is defined as*

$$\phi_t(x) := (2\pi\Sigma_{0;t})^{-1}\exp\Big(-\frac{\|x\|_2^2}{2\Sigma_{0;t}}\Big)$$

and

$$\Sigma_{s;t}^{(i)} := \int_s^t \sigma(u)^2 du.$$

Proof. Let $\{\mathcal{F}_t^Z; t \in [0,T]\}$ denote the σ-field generated by $\{Z_t; t \in [0,T]\}$ and $\{\mathcal{F}_t^W; t \in [0,T]\}$ the σ-field generated by $\{W_t; t \in [0,T]\}$ as usual. Since $dZ_t = \text{diag}(\sigma(t),\sigma(t))dW_t$ it follows that $Z_t^{(i)} = \int_0^t \sigma(u)dW_u^{(i)}$ for $i = 1,2$. As in the proof of theorem 4.2 we have $\langle Z^{(i)}\rangle_t = \int_0^t \sigma(u)^2 d\langle W^{(i)}\rangle_u = \int_0^t \sigma(u)^2 du$ since $\langle W^{(i)}\rangle_t = t$ for all $t \ge 0$ and $i \in \{1,2\}$. $\{W_t^{(1)}; t \in [0,T]\}$ and $\{W_t^{(2)}; t \in [0,T]\}$ are independent and thus $\langle W^{(1)}, W^{(2)}\rangle_t = 0$ for all $t \ge 0$. Therefore

$$\langle Z^{(1)}, Z^{(2)}\rangle_t = \int_0^t \sigma(u)^2 d\langle W^{(1)}, W^{(2)}\rangle_u = 0.$$

Finally, $\{Z_t; t \in [0,T]\}$ is a continuous adapted martingale and $\lim_{t\to\infty}\langle Z^{(i)}\rangle_t = \infty$ for $i = 1,2$. We apply Knight's theorem 1.49 and have $Z_t^{(i)} = W^{(i)}_{<Z^{(i)}>_t}$ for $t \geq 0$ and $i \in \{1,2\}$. Since $\langle Z^{(1)}\rangle_t = \langle Z^{(2)}\rangle_t$ for every $t \in [0,T]$ we set $\tau(t) :=< Z^{(1)} >_t$. As before

$$\begin{aligned} E_{\mathcal{P}}[f \circ Z_t|\mathcal{F}_s^Z] &= E_{\mathcal{P}}[f \circ W_{\tau(t)}|\mathcal{F}^W_{\tau(s)}] \\ &= (P_{\tau(t)-\tau(s)}f)(W_{\tau(s)}) \end{aligned}$$

holds for any $s,t \in [0,T]$, $s < t$, and Borel-measurable $f : \mathbb{R}^2 \to \mathbb{R}$, $f \geq 0$. The transition semi-group $(K_t^W)_{t\geq 0}$ of the two-dimensional Wiener process is given by

$$\begin{aligned} (K_t^W f)(x) &= \int \phi_t(x-y)f(y)d\lambda^2(y), \\ \phi_t(x) &= (2\pi t)^{-1}\exp\Big(-\frac{\|x\|_2^2}{2t}\Big) \text{ for } t > 0, \end{aligned}$$

Applying the time-change we get

$$(K^W_{\tau(t)-\tau(s)}f)(x) = \int \phi_{\tau(t)-\tau(s)}(x-y)f(y)d\lambda^2(y),$$

where

$$\phi_{\tau(t)-\tau(s)}(x) = (2\pi(\tau(t)-\tau(s)))^{-1}\exp\Big(-\frac{\|x\|_2^2}{2(\tau(t)-\tau(s))}\Big).$$

Again

$$\tau(t) - \tau(s) = \int_s^t \sigma(u)^2 du =: \Sigma_{s;t}.$$

and changing back in the original time-coordinates we have

$$(K^W_{\tau(t)-\tau(s)}f)(W_{\tau(s)}) =: (K^Z_{t-s}f)(Z_s),$$

where the transition density function $\phi_t(x)$ of the transition semi-group $(K_t^Z)_{t\geq 0}$ is given by

$$\phi_t(x) = (2\pi\Sigma_{0;t})^{-1}\exp\Big(-\frac{\|x\|_2^2}{2\Sigma_{0;t}}\Big).$$

This completes the proof. □

We give two examples for possible choices of the function σ here.

Example 1. 1. In the one-dimensional case $d = 1$ the function σ may be specified via

$$\sigma_t^{(i)} = a_i \exp(-b(T_i - t)),\ a_i, b \in \mathbb{R}.$$

The constant b is usually referred to as the *mean reversion parameter*. It may be readily seen that the above specified volatility structure is separable.

A two-dimensional function σ may be constructed in a similar way. Choose for example

$$\sigma_t^{(i)} = \begin{pmatrix} a_1 \exp(-b(T_i - t)) \\ a_2 \exp(-b(T_i - t)) \end{pmatrix} = \begin{pmatrix} a_1 \exp(-bT_i)) \\ a_2 \exp(-bT_i)) \end{pmatrix} \exp(bt), \quad a_1, a_2, b \in \mathbb{R}.$$

2. The following choice of σ due to Rebonato is also popular amongst practitioners. It combines the mean reversion form example 1 with a linear growing term,

$$\sigma_t^{(i)} = (a(T_i - t) + b) \exp(-c(T_i - t)) + d, \; a, b, c, d \in \mathbb{R}.$$

   However, σ_t is *not* separable and the stochastic process $\{Z_t;\, t \in [0,T]\}$ will not necessarily be Markovian.

Both choices for σ_t produce volatility structures similar to those observed in the market including humps. Since the resulting process $\{Z_t;\, t \in [0,T]\}$ derived from example 2 may not be Markovian, we prefer working with functions σ_t similar to example 1. Note that increasing the mean reversion parameter b reduces the correlation of the LIBORs (equivalently swap rates) of different times (cf. [Hunt et al. 2000]).

When specifying the functional f giving us the driving process $\{f(Z_t);\, t \in [0,T]\}$, Z_t will usually appear in the form $Z_t^{(1)} + Z_t^{(2)}$ (modulo some time-dependent coefficients). This has the following reason. First, derive the functional f using a drift approximation to Market models described in Section 3. When working with a 2-factor Market model one usually models the i-th forward LIBOR rate $L^{(i)}$ (equivalently the i-th forward swap rate $S^{(i)}$) as

$$dL_t^{(i)} = \mu^{(i)}(t, L_t^{(i+1)}, \ldots, L_t^{(n)}) L_t^{(i)} dt + \sum_{j=1}^{2} L_t^{(i)} \sigma_t^{(j)} dW_t^{(j)}, \mathcal{S}^{(n)}, t \in [0, T_i], L_0^{(i)} = L_i, \tag{4.16}$$

where $\{W_t; t \in [0,T]\}$ is a two-dimensional Wiener process under the measure $\mathcal{S}^{(n)}$, $d\langle W_t^{(1)}, W_t^{(2)} \rangle_t = \rho\, dt$ as before and $\sigma^{(i)} : [0, T_i] \to \mathbb{R}^+$, $i = 1, \ldots, n$, are such that a solution of each SDE of the system (4.16) exists. Applying Itô's formula (1.7) yields the solution of (4.16)

$$\begin{aligned} L_t^{(i)} =& L_0^{(i)} \exp\Big(\int_0^t \mu^{(i)}(s, L_s^{(i+1)}, \ldots, L_s^{(n)}) - \Big(\big(\frac{1}{2}\sigma_s^{(1)}\big)^2 + \sigma_s^{(1)}\sigma_s^{(2)}\rho + \frac{1}{2}\big(\sigma_s^{(2)}\big)^2 \Big) ds \\ &+ \int_0^t \sigma_s^{(1)} dW_s^{(1)} + \int_0^t \sigma_s^{(2)} dW_s^{(2)} \Big). \end{aligned} \tag{4.17}$$

Having a look at the above equation it is tempting to reduce the driving stochastic process $\int_0^t \sigma_s^{(1)} dW_s^{(1)} + \int_0^t \sigma_s^{(2)} dW_s^{(2)}$ to $\int_0^t \sqrt{\left(\sigma_s^{(1)}\right)^2 + \left(\sigma_s^{(2)}\right)^2} d\tilde{W}_s$ for some one-dimensional Wiener process $\{\tilde{W}_t; t \in [0,T]\}$, hence reducing the model's dimension to one although having two different parameters, $\sigma_t^{(1)}$ and $\sigma_t^{(2)}$. However, this is only possible if $\rho = 0$. $\{\tilde{W}_t^{(1)}; t \in [0,T]\}$ and $\{\tilde{W}_t^{(2)}; t \in [0,T]\}$ are then uncorrelated, thus independent and evaluation of the model would be less cumbersome.
Unfortunately, the case $\rho = 0$ is not of great practical importance. The various (random) effects influencing the shape of the interest or swap rate curve like parallel shifts, rotations or twistings are not occurring independently of each other but are correlated in some way.
To summarize, we have different means of controlling the correlation of the different rates. First of all, we may change the parameter ρ to control the instantaneous correlation of the rates directly. Further, when choosing volatility functions σ_t as in example 1 above the choice of the mean reversion parameter affects the correlation of the rates, too. A higher mean reversion parameter will lead to a lower correlation and vice versa.
Note further that since (4.14) is a deterministic time-change of Brownian motion the choice of $\sigma^{(i)}$ also affects the distribution of Z_t for $t \notin \mathcal{T}$. Evaluating our model, we will not be interested in the process $\{Z_t;\ t \in [0,T]\}$ for times $t \notin \mathcal{T}$ but this gives us control of what is happening at these times.

We now define the class of functionals f that will be suitable for our purposes.

Definition 4.5. *A functional $f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}, (x,t) \mapsto f(x,t)$ is called* admissible*, if f is strictly monotonic increasing in x in the sense of Definition 1.12 and Borel-measurable.*

Note that also functionals of the form $f(x,t) = x^{(1)}$ are included in our definition. We will need them later.
Again we have to establish the properties (P.1)-(P.3) for a complete specification of the model. Property (P.1), however, is no longer valid in this form and will be substituted by

(P.1') The set of admissible functionals $f^{(i)}, i = 1, \dots, n$,

$$f^{(i)} : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}, (x,t) \mapsto f^{(i)}(x,t)$$

and the distribution of the Markov process $\{Z_t;\ t \in [0,T]\}$ under $\mathcal{N}$ are known.

Assume now we already found a set of admissible functionals $f^{(i)}$. Continue then by choosing a one-dimensional process $\{X_t;\ t \in [0,T]\} = \{X_t(Z_t);\ t \in [0,T]\}$ such that

$$X_t(Z_t) := \begin{cases} f^{(i)}(Z_t, t), & t \in \mathcal{T} \\ \text{arbitrary}, & t \notin \mathcal{T}. \end{cases}$$

Then it follows under the same assumptions as before that

$$\{X_{T_i} > x^*\} = \{S^{(i)}_{T_i}(X_{T_i}) > K^{(i)}(x^*)\}.$$

By virtue of this set identity each $f^{(i)}$ is assigned to the corresponding swap rate $S^{(i)}$. Again define the numeraire via $N_{T_i}(X_{T_i}) := D_{T_iT_{n+1}}(X_{T_i})$ and the pure discount bond prices via

$$D_{tT} = N_t E_{\mathcal{N}}[N_T^{-1}|\mathcal{F}_t^W], T \in \mathcal{T}, \tag{4.18}$$

under the equivalent martingale measure $\mathcal{N}$ associated with the numeraire N. $\{\mathcal{F}_t^W; t \in [0,T]\}$ is the filtration generated by $\{W_t;\ t \in [0,T]\}$. It can easily be seen that $D_{TT} = 1$ and $D_{tT} > 0$ holds for $T \in \mathcal{T}$. By construction the relative asset price $\{\frac{D_{tT}}{N_t}; t \in [0,T]\}$ is a martingale for every T and the model is arbitrage free. Define

$$\begin{aligned} J_0^{(i)}(x^*) &:= N_0(X_0)E_{\mathcal{N}}\Big[\frac{P^{(i)}_{T_i}(X_{T_i})}{N_{T_i}(X_{T_i})}\mathbf{1}_{\{S^{(i)}_{T_i}(X_{T_i})>K^{(i)}(x^*)\}}\Big] \\ &= N_0(X_0)E_{\mathcal{N}}\Big[\frac{P^{(i)}_{T_i}(X_{T_i})}{N_{T_i}(X_{T_i})}\mathbf{1}_{\{X_{T_i}>x^*\}}\Big]. \end{aligned}$$

We can calculate $J_0^{(i)}(x^*)$ for any given x^* using the fact that $X_{T_i} = f^{(i)}(Z_{T_i}, T_i)$ and the known distribution of Z_{T_i} under $\mathcal{N}$. Using market prices of the corresponding digital swaptions we can find the value of K such that

$$J_0^{(i)}(x^*) = V_0^{(i)}(K).$$

As before $K = K^{(i)}(x^*)$ and knowing $K^{(i)}(x^*)$ for all x^* is equivalent to knowing the functional form of $S^{(i)}_{T_i}(X_{T_i})$. Having this we are done.
Note that the two-dimensional Markov-functional model defined here is truly a generalization of the one-dimensional Markov-functional model presented in Section 4.1. To check this simply set $f^{(i)}(x,t) = x^{(1)}$. When evaluating the two-dimensional Markov-functional model using these functionals one will end up with the one-dimensional Markov-functional model.
The following example shows a two-dimensional LIBOR Markov-functional model. A swap model could be constructed along the same lines except for the different accrual factor, which would be P_{T_i} instead of $D_{T_iT_{n+1}}$. For the sake of brevity we will restrict ourselves to the two-dimensional LIBOR Markov-functional model.

4.2.1 Two-dimensional LIBOR model

We start again specifying the properties (P.1')-(P.3).
Assume we already found a set of n admissible functionals $f^{(i)} : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}, i =$

$1, \ldots, n$, as described above. Define the driving process $\{Z_t; t \in [0,T]\}$ as in Section 4.2. This completes the specification of (P.1). As before we assume than $L^{(n)}$ is a log-normal martingale under $\mathcal{S}^{(n)}$, the equivalent measure associated with the numeraire $D_{tT_{n+1}}$, i.e.,

$$dL_t^{(n)} = \sum_{k=1}^{2} L_t^{(n)} \sigma_s^{(n,k)} dW_s^{(k)}, \mathcal{S}^{(n)}, t \in [0, T_n], L_0^{(n)} = L_n.$$

$\sigma^{(i,k)} : [0, T_i] \to \mathbb{R}^+$ is a separable set of volatility functions, $\sigma_t^{(i,k)} = v_i^{(k)} \sigma(t)$ for $i = 1, \ldots, n, k = 1, 2$, such that the above SDE admits a solution. Further, $\{W_t; t \in [0, T_n]\}$ is a two-dimensional Wiener process with $d\langle W^{(1)}, W^{(2)}\rangle_t = \rho\, dt$. By virtue of Itô's formula (1.7),

$$\begin{aligned} L_t^{(n)} &= L_0^{(n)} \exp\Big(- \int_0^t \sum_{k=1}^{2} \frac{1}{2} \big(\sigma_s^{(n,k)}\big)^2 + \rho \sigma_s^{(n,1)} \sigma_s^{(n,2)} ds + \sum_{k=1}^{2} \int_0^t \sigma_s^{(n,k)} dW_s^{(k)} \Big) \\ &= L_0^{(n)} \exp\Big(- \int_0^t \sum_{k=1}^{2} \frac{1}{2} \big(v_n^{(k)} \sigma(s)\big)^2 + \rho v_n^{(1)} v_n^{(2)} \sigma(s)^2 ds + \sum_{k=1}^{2} v_n^{(k)} \int_0^t \sigma(s) dW_s^{(k)} \Big) \end{aligned} \tag{4.19}$$

Since $W^{(1)}$ and $W^{(2)}$ are correlated we cannot proceed immediately. First we rewrite the above stochastic integrals as follows.
Suppose $\{\tilde{W}_t; t \in [0, T_n]\}$ is a standard, two-dimensional Wiener process. Define a new stochastic process $\{W_t; t \in [0, T_n]\}$ by

$$W_t^{(1)} := \tilde{W}_t^{(1)}, \tag{4.20}$$

$$W_t^{(2)} := \rho \tilde{W}_t^{(1)} + \sqrt{1 - \rho^2} \tilde{W}_t^{(2)} \tag{4.21}$$

for some $\rho \in [-1, 1]$ and all $t \in [0, T_n]$. It is a simple calculation to check that

$$\begin{aligned} d\langle W^{(1)}, W^{(1)}\rangle_t &= d\langle W^{(2)}, W^{(2)}\rangle_t = dt, \\ d\langle W^{(1)}, W^{(2)}\rangle_t &= \rho\, dt. \end{aligned}$$

Hence $\{W_t; t \in [0, T_n]\}$ is a two-dimensional Wiener process with the desired correlation structure. Substituting (4.20) into equation (4.19) yields

$$\begin{aligned} L_t^{(n)} = L_0^{(n)} \exp\Big(&- \int_0^t \sum_{k=1}^{2} \frac{1}{2} \big(v_n^{(k)} \sigma(s)\big)^2 + \rho v_n^{(1)} v_n^{(2)} \sigma(s)^2 ds + \\ &\int_0^t \big(v_n^{(1)} + \rho v_i^{(2)}\big) \sigma(s) d\tilde{W}_s^{(1)} + \int_0^t \sqrt{1 - \rho^2} v_n^{(2)} \sigma(s) d\tilde{W}_s^{(2)} \Big) \\ = L_0^{(n)} \exp\Big(&- \int_0^t \sum_{k=1}^{2} \frac{1}{2} \big(v_n^{(k)} \sigma(s)\big)^2 + \rho v_n^{(1)} v_n^{(2)} \sigma(s)^2 ds + \\ &\big(v_n^{(1)} + \rho v_i^{(2)}\big) Z_t^{(1)} + \sqrt{1 - \rho^2} v_n^{(2)} Z_t^{(2)} \Big). \end{aligned}$$

We only need to define the functional form of the numeraire $D_{T_nT_{n+1}}$ to specify (P.1') as the pure discount bonds $D_{T_iT_i}(X_{T_i})$ are trivially the unit map. $D_{T_nT_{n+1}}$ is given via

$$D_{T_nT_{n+1}} = \frac{1}{1+\alpha_n L^{(n)}_{T_n}}$$

or equivalently

$$D_{T_nT_{n+1}} = \Bigg(1+\alpha_n L^{(n)}_0 \exp\Big(-\int_0^t \sum_{k=1}^{2} \frac{1}{2}\big(v^{(k)}_n \sigma(s)\big)^2 + \rho v^{(1)}_n v^{(2)}_n \sigma(s)^2 ds + \big(v^{(1)}_n + \rho v^{(2)}_i\big) Z^{(1)}_t + \sqrt{1-\rho^2} v^{(2)}_n Z^{(2)}_t\Bigg)^{-1}.$$

As before it remains to find the functional form of $N_{T_i}(X_{T_i}) = D_{T_nT_{n+1}}(X_{T_i})$ for $i = 1,\dots,n-1$. This will be accomplished using backward induction.
Taking the caplets on the forward LIBORs $L^{(i)}$ as calibrating instruments we may calculate the price of the corresponding digital caplet with strike K by

$$\begin{aligned} V^{(i)}_0(K) &= N_0(X_0) E_{\mathcal{S}^{(n)}}\Big[\frac{D_{T_iT_{i+1}}(X_{T_i})}{N_{T_i}(X_{T_i})} \mathbf{1}_{\{L^{(i)}_{T_i}(X_{T_i})>K\}}\Big] \\ &= D_{0T_{n+1}}(X_0) E_{\mathcal{S}^{(n)}}\Big[\frac{D_{T_iT_{i+1}}(X_{T_i})}{N_{T_i}(X_{T_i})} \mathbf{1}_{\{L^{(i)}_{T_i}(X_{T_i})>K\}}\Big]. \end{aligned}$$

Assuming that the market value is given by Black's formula for digital caplets (2.16) with volatility $\sigma^{(i)}_t$, the price at time zero for this digital is

$$V^{(i)}_0(K) = D_{0T_{n+1}}(X_0)\Phi(d^{(i)}_2),$$

where

$$d_2 = \frac{\log(\frac{L^{(i)}_0}{K})}{\Sigma^{(i)}_{0;T_i}\sqrt{T_i}} - \frac{1}{2}\Sigma^{(i)}_{0;T_i}\sqrt{T_i}$$

and

$$\Sigma^{(i)}_{0;T} = \frac{1}{T}\int_0^T (\sigma^{(i)}_u)^2 du.$$

To determine the functional forms $D_{T_iT_{n+1}}(X_{T_i})$ for $i < n$ we proceed as in Section 4.1.2. Choose some $x^* \in \mathbb{R}$. Evaluate by numerical integration

$$\begin{aligned}
J_0^{(i)}(x^*) &= D_{0T_{n+1}}(X_0)E_{\mathcal{S}^{(n)}}\Big[\frac{D_{T_iT_{i+1}}(X_{T_i})}{D_{T_iT_{n+1}}(X_{T_i})}\mathbf{1}_{\{X_{T_i}>x^*\}}\Big] \\
&= D_{0T_{n+1}}(X_0)E_{\mathcal{S}^{(n)}}\Big[E_{\mathcal{S}^{(n)}}\Big[\frac{D_{T_{i+1}T_{i+1}}(X_{T_{i+1}})}{D_{T_{i+1}T_{n+1}}(X_{T_{i+1}})}|\mathcal{F}_{T_i}^W\Big]\mathbf{1}_{\{X_{T_i}>x^*\}}\Big] \\
&= D_{0T_{n+1}}(X_0)E_{\mathcal{S}^{(n)}}\Big[E_{\mathcal{S}^{(n)}}\Big[\frac{1}{D_{T_{i+1}T_{n+1}}(X_{T_i})}|\mathcal{F}_{T_i}^W\Big]\mathbf{1}_{\{X_{T_i}>x^*\}}\Big] \\
&= D_{0T_{n+1}}(X_0)\int_{\mathbb{R}^2}\Big[\int_{\mathbb{R}^2}\frac{K^Z_{T_i;T_{i+1}}(u)}{D_{T_{i+1}T_{n+1}}(f^{(i+1)}(v-u,T_{i+1}))}d\lambda^2(u)\Big] \\
&\qquad \mathbf{1}_{\{f^{(i)}(v,T_i)>x^*\}}\phi_{T_i}(v)d\lambda^2(v) \\
&= D_{0T_{n+1}}(X_0)\int_{\{f^{(i)}(v,T_i)>x^*\}} \\
&\qquad \Big[\int_{\mathbb{R}^2}\frac{K^Z_{T_i;T_{i+1}}(u)}{D_{T_{i+1}T_{n+1}}(f^{(i+1)}(v-u,T_{i+1}))}d\lambda^2(y)\Big]\phi_{T_i}(v)d\lambda^2(v),
\end{aligned}$$

where $\phi_t(x)$ is defined by

$$\phi_t(x) := (2\pi\Sigma_{0;t})^{-1}\exp\Big(-\frac{\|x\|_2^2}{2\Sigma_{0;t}}\Big)$$

and $K^Z_{T_i;T_{i+1}}$ by

$$K^Z_{T_i;T_{i+1}} := (2\pi\Sigma^{(i)}_{T_i;T_{i+1}})^{-1}\exp\Big(-\frac{\|x\|_2^2}{2\Sigma^{(i)}_{T_i;T_{i+1}}}\Big),$$

where

$$\Sigma^{(i)}_{s;t} := \int_s^t (\sigma_u^{(i)})^2 du.$$

Note that $D_{T_{i+1}T_{n+1}}(f^{(i+1)}(X_{T_{i+1}},T_{i+1}))$ is a strictly positive functional depending on the Markov process $\{Z_t;\, t \in [0,T]\}$. Note further that due to the Borel-measureability of $f^{(i)}$ the set $\{x \in \mathbb{R}^2; f^{(i)}(x,T_i) > x^*\}$ is also Borel-measurable. It now remains to specify the functionals $f^{(i)}$ such that $\{x \in \mathbb{R}^2; f^{(i)}(x,T_i) > x^*\}$, the set we are integrating, has a 'simple' structure to ease integration. Some proposals for the choice of $f^{(i)}$ will be made in the next section.

4.3 On the choice of the functionals $f^{(i)}$

All functionals $f^{(i)}$ discussed in this section are based on drift approximations to the LIBOR market model. Drift approximation means simply that the path-dependent drift terms $\mu^i, i = 1,\ldots,n-1$, in the LIBOR Market model (under the measure $\mathcal{S}^{(n)}$) are substituted by 'simpler' drift terms $\hat{\mu}^i$ that are not path-dependent anymore and easier to evaluate. The reason for the use of drift approximations has been outlined in Section 3. The three drift approximations presented below are ordered by an increasing level of sophistication. We start with the simplest approximation possible.

4.3.1 Freezing the drift

When examining the stochastic differential equations of the two-dimensional LIBOR market model (and equivalently the Swap market model), assuming a separable set of volatility functions $\sigma_t^{(i,j)} := v_i^{(j)}\sigma(t)$, $i = 1, \ldots, n, j = 1, 2$, we have

$$dL_t^{(i)} = \mu^i(t, L_t^{(i+1)}, \ldots, L_t^{(n)})L_t^{(i)}dt + \sum_{j=1}^{2} L_t^{(i)}\sigma_t^{(i,j)}dW_t^{(j)}, \mathcal{S}^{(n)}, t \in [0, T_i], L_0^{(i)} = L_i,$$

where

$$\begin{aligned}
\mu^i(t, L_t^{(i+1)}, \ldots, L_t^{(n)}) &= -\sum_{j=i+1}^{n} \frac{\alpha_j L_t^{(j)}}{1 + \alpha_j L_t^{(j)}} \\
&\quad \big(\sigma_t^{(i,1)}\sigma_t^{(j,1)} + \sigma_t^{(i,2)}\sigma_t^{(j,2)} + \rho\big(\sigma_t^{(i,1)}\sigma_t^{(j,2)} + \sigma_t^{(i,2)}\sigma_t^{(j,1)}\big)\big) \\
&= -\sum_{j=i+1}^{n} \frac{\alpha_j L_t^{(j)}}{1 + \alpha_j L_t^{(j)}} \\
&\quad \big(v_i^{(1)}v_j^{(1)} + v_i^{(2)}v_j^{(2)} + \rho\big(v_i^{(1)}v_j^{(2)} + v_i^{(2)}v_j^{(1)}\big)\big)\sigma(t)^2 \text{ for } i < n, \\
\mu^n &\equiv 0,
\end{aligned}$$

and $d\langle W^{(1)}, W^{(2)}\rangle_t = \rho\, dt$. Obviously the simplest drift approximation is to 'freeze the drift'. This means that $\mu^i(t, L_t^{(i+1)}, \ldots, L_t^{(n)})$ is substituted by

$$\begin{aligned}
\hat{\mu}^i(t, L_0^{(i+1)}, \ldots, L_0^{(n)}) &= -\sum_{j=i+1}^{n} \frac{\alpha_j L_0^{(j)}}{1 + \alpha_j L_0^{(j)}} \\
&\quad \big(v_i^{(1)}v_j^{(1)} + v_i^{(2)}v_j^{(2)} + \rho\big(v_i^{(1)}v_j^{(2)} + v_i^{(2)}v_j^{(1)}\big)\big)\sigma(t)^2 \text{ for } i < n.
\end{aligned}$$

This new system of stochastic differential equations admits the solution

$$\begin{aligned}
\widehat{L}_t^{(i)} = L_0^{(i)} \exp\Big(\int_0^t & \hat{\mu}^i(t, L_0^{(i+1)}, \ldots, L_0^{(n)}) - \\
& \Big(\frac{1}{2}\big(v_i^{(1)}\sigma(s)\big)^2 + \frac{1}{2}\big(v_i^{(2)}\sigma(s)\big)^2 + \rho v_i^{(1)}v_i^{(2)}\sigma(s)^2\Big)ds \\
& + \int_0^t v_i^{(1)}\sigma(s)dW_s^{(1)} + \int_0^t v_i^{(2)}\sigma(s)dW_s^{(2)}\Big).
\end{aligned} \tag{4.22}$$

We substitute the correlated Wiener process by an uncorrelated one as shown in the last section.

Suppose $\{\tilde{W}_t; t \in [0, T_n]\}$ is a standard, two-dimensional Wiener process. Define $\{W_t; t \in [0, T_n]\}$ by

$$\begin{aligned} W_t^{(1)} &:= \tilde{W}_t^{(1)}, \\ W_t^{(2)} &:= \rho \tilde{W}_t^{(1)} + \sqrt{1-\rho^2}\tilde{W}_t^{(2)} \end{aligned}$$

for some $\rho \in [-1, 1]$ and all $t \in [0, T_n]$. Substituting those equations into the stochastic integral terms above yields

$$\begin{aligned} &\int_0^t v_i^{(1)}\sigma(s)dW_s^{(1)} + \int_0^t v_i^{(2)}\sigma(s)dW_s^{(2)} \\ &= \int_0^t v_i^{(1)}\sigma(s)d\tilde{W}_s^{(1)} + \int_0^t v_i^{(2)}\sigma(s)d\big(\rho\tilde{W}_s^{(1)} + \sqrt{1-\rho^2}\tilde{W}_s^{(2)}\big) \\ &= \int_0^t \big(v_i^{(1)} + \rho v_i^{(2)}\big)\sigma(s)d\tilde{W}_s^{(1)} + \int_0^t \sqrt{1-\rho^2}v_i^{(2)}\sigma(s)d\tilde{W}_s^{(2)}. \end{aligned}$$

We rewrite (4.22) as

$$\begin{aligned} \widehat{L}_t^{(i)} &= L_0^{(i)} \exp\Big(\int_0^t \hat{\mu}^i(t, L_0^{(i+1)}, \ldots, L_0^{(n)}) - \Big(\frac{1}{2}\big(v_i^{(1)}\big)^2 + \frac{1}{2}\big(v_i^{(2)}\big)^2 + \rho v_i^{(1)}v_i^{(2)}\Big)\sigma(s)^2 ds \\ &\quad + \int_0^t \big(v_i^{(1)} + \rho v_i^{(2)}\big)\sigma(s)d\tilde{W}_s^{(1)} + \int_0^t \sqrt{1-\rho^2}v_i^{(2)}\sigma(s)d\tilde{W}_s^{(2)}\Big) \\ &= L_0^{(i)} \exp\Big(\int_0^t \hat{\mu}^i(t, L_0^{(i+1)}, \ldots, L_0^{(n)}) - \Big(\frac{1}{2}\big(v_i^{(1)}\big)^2 + \frac{1}{2}\big(v_i^{(2)}\big)^2 + \rho v_i^{(1)}v_i^{(2)}\Big)\sigma(s)^2 ds \\ &\quad + \big(v_i^{(1)} + \rho v_i^{(2)}\big)dZ_t^{(1)} + \sqrt{1-\rho^2}v_i^{(2)}dZ_s^{(2)}\Big), \end{aligned}$$

where

$$dZ_t = \operatorname{diag}(\sigma(t), \sigma(t))d\tilde{W}_t$$

is our driving Markov process for some standard, two-dimensional Wiener process $\{\tilde{W}_t; t \in [0, T_n]\}$. Therefore we take

$$\boxed{\begin{aligned} f^{(i)}(x,t) \;=\; & L_0^{(i)} \exp\Big(\textstyle\int_0^t \hat{\mu}^i(t, L_0^{(i+1)}, \ldots, L_0^{(n)}) - \\ & \big(\tfrac{1}{2}\big(v_i^{(1)}\big)^2 + \tfrac{1}{2}\big(v_i^{(2)}\big)^2 + \rho v_i^{(1)}v_i^{(2)}\big)\sigma(s)^2 ds + \\ & \big(v_i^{(1)} + \rho v_i^{(2)}\big)x^{(1)} + \sqrt{1-\rho^2}v_i^{(2)}x^{(2)}\Big) \end{aligned}} \tag{4.23}$$

as a choice of our functional $f^{(i)}$. For simplification we write

$$f^{(i)}(x,t) = L_0^{(i)} \exp\Big(g(t) + \big(v_i^{(1)} + \rho v_i^{(2)}\big)x^{(1)} + \sqrt{1-\rho^2}v_i^{(2)}x^{(2)}\Big),$$

where $g(t)$ is chosen obviously.
We still don't know whether the functionals $f^{(i)}$ are admissible. The answer is given by the following lemma.

Lemma 4.6. *The functionals $f^{(i)}$ as defined by equation (4.23) are admissible, if v_i is component wise strictly positive for all $i = 1, \dots, n$ and $\rho \in [0,1]$.*

Proof. Clearly $f^{(i)}(x,t)$ is continuous in x and t, therefore Borel-measurable. Fix some $x^{(1)} \in \mathbb{R}$. Then the map $l^{(i)} : \mathbb{R} \to \mathbb{R}, x \mapsto \big(v_i^{(1)} + \rho v_i^{(2)}\big)x^{(1)} + \sqrt{1-\rho^2}v_i^{(2)}x$ is strictly monotonic increasing if $v_i^{(2)} > 0$ and $\rho \in [0,1]$ and so is $x \mapsto L_0^{(i)} \exp\Big(g(t) + \big(v_i^{(1)} + \rho v_i^{(2)}\big)x^{(1)} + \sqrt{1-\rho^2}v_i^{(2)}x\Big)$ for some fixed $t \in \mathbb{R}$. Remember that $L_0^{(i)} > 0$. The same holds when fixing some $x^{(2)} \in \mathbb{R}$ and verifying the map $x \mapsto \big(v_i^{(1)} + \rho v_i^{(2)}\big)x + \sqrt{1-\rho^2}v_i^{(2)}x^{(2)}$. Therefore $f^{(i)}(x,t)$ is strictly monotonic increasing in x, hence admissible, if v_i is component wise strictly positive and $\rho \in [0,1]$. □

The extra requirement that v_i should be component wise strictly positive is no severe limitation for practical purposes. For example, when choosing a mean reverting volatility structure as proposed in example 1, the constants a_1 and a_2 will be positive real numbers, otherwise the volatility $\sigma_t^{(i)}$ would be negative for all t and this is not desirable. Therefore the vectors v_i will be strictly positive, too.
When calculating $J_0^{(i)}(x^*)$ during the construction of the two-dimensional LIBOR model in Section 4.2.1 we ended up with the term

$$D_{0T_{n+1}}(X_0) \int_{\{f^{(i)}(v,T_i)>x^*\}} \Big[\int_{\mathbb{R}^2} \frac{1}{D_{T_{i+1}T_{n+1}}(f^{(i+1)}(v-u,T_{i+1}))} K^Z_{T_i;T_{i+1}}(u)d\lambda^2(u)\Big] \phi_{T_i}(v)d\lambda^2(v). \quad (4.24)$$

Having now specified $f^{(i)}(x,t)$ we may continue evaluating the integral

$$\int_{\{f^{(i)}(v,T_i)>x^*\}} h(v)\phi_{T_i}(v)d\lambda^2(v),$$

where h is chosen obviously. Since

$$\{x \in \mathbb{R}^2; f^{(i)}(x,T_i) > x^*\} = \{x \in \mathbb{R}^2; L_0^{(i)} \exp(g(T_i) + v_i^{(1)}x^{(1)} + v_i^{(2)}x^{(2)}) > x^*\}$$

we may restrict ourselves to values $x^* \in \mathbb{R}^+$. Choose some $x^* \in \mathbb{R}^+$. Then

$$\begin{aligned} &\{x \in \mathbb{R}^2; L_0^{(i)} \exp(g(T_i) + v_i^{(1)}x^{(1)} + v_i^{(2)}x^{(2)}) > x^*\} \\ =&\{x \in \mathbb{R}^2; v_i^{(1)}x^{(1)} + v_i^{(2)}x^{(2)} > \log\Big(\frac{x^*}{L_0^{(i)}}\Big) - g(T_i)\}. \end{aligned} \quad (4.25)$$

Set $\tilde{c} := \log\Big(\frac{x^*}{L_0^{(i)}}\Big) - g(T_i)$. The set (4.25) is a plane in $\mathbb{R}^2$, thus it suffices to determine

$$\{x \in \mathbb{R}^2; v_i^{(1)}x^{(1)} + v_i^{(2)}x^{(2)} = \tilde{c}\}. \quad (4.26)$$

Integrating over the sets

$$\{x \in \mathbb{R}^2; v_i^{(1)}x^{(1)} + v_i^{(2)}x^{(2)} > \tilde{c}\}$$

and

$$\{x \in \mathbb{R}^2; v_i^{(1)}x^{(1)} + v_i^{(2)}x^{(2)} \geq \tilde{c}\}$$

is equivalent since the difference set

$$\{x \in \mathbb{R}^2; v_i^{(1)}x^{(1)} + v_i^{(2)}x^{(2)} = \tilde{c}\}$$

has Lebesgue-measure zero.
The equation specifying (4.26) can be written in the form

$$\frac{v_i^{(1)}}{v_i^{(2)}}x^{(1)} + x^{(2)} = \frac{\tilde{c}}{v_i^{(2)}} := c. \tag{4.27}$$

The graph of (4.27) is a strictly monotonic decreasing line with gradient angle ψ given by $\tan\psi = \frac{v_i^{(1)}}{v_i^{(2)}}$. Using a rotation, the set $\{x \in \mathbb{R}^2; \frac{v_i^{(1)}}{v_i^{(2)}}x^{(1)} + x^{(2)} = c\}$ can be transformed into $\mathbb{R} \times \{x^{(2)} > c\}$.

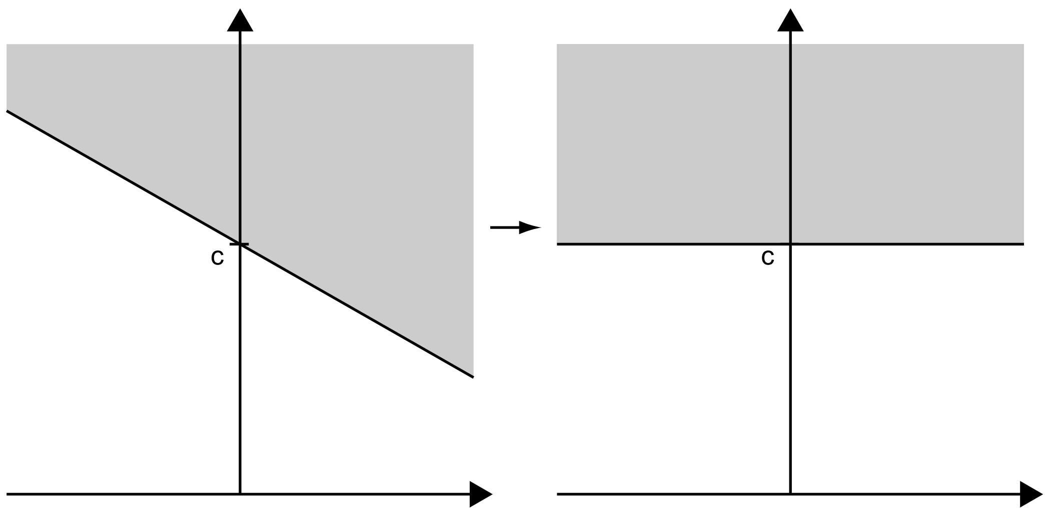

Figure 4.2: The plane integrated before and after rotation

Set

$$R := \begin{pmatrix} \cos(-\psi) & \sin(-\psi) \\ -\sin(-\psi) & \cos(-\psi) \end{pmatrix} = \begin{pmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{pmatrix}.$$

It is straightforward to check that R is a C^1-Diffeomorphism and that the mapping $R : \mathbb{R}^2 \to \mathbb{R}^2, x \mapsto Rx$ provides the desired rotation. Further the Lebesgue-measure λ is invariant under C^1-Diffeomorphisms as stated by the following theorem.

Theorem 4.7. *Let G, G' be two open subsets of $R^n, n \in \mathbb{N}$ and $\phi : G \to G'$ a C^1-Diffeomorphism. An extended function f' defined on G' is λ^n-integrable over G' iff the function $f' \circ \phi \, |detD\phi|$ is λ^n-integrable over G. Then*

$$\int_{G'} f' \lambda^n = \int_G f' \circ \phi \, | \det D\phi | d\lambda^n.$$

Proof. See [Bauer 1991], Theorem 19.4. □

Therefore the integral (4.24) becomes

$$D_{0T_{n+1}}(X_0) \int_{\mathbb{R}\times\{x^{(2)}>c\}} \Big[\int_{\mathbb{R}^2} \frac{1}{D_{T_{i+1}T_{n+1}}(f^{(i+1)}(u, T_{i+1}))} K^Z_{T_i;T_{i+1}}(Rv - u) d\lambda^2(u) \Big] \phi_{T_i}(Rv) d\lambda^2(v), \quad (4.28)$$

since $\det R = 1$. Numerical integration of (4.28) now gives $J_0^{(i)}$ and one may proceed as before by calculating $L^{(i)}_{T_i}(x^*)$ via

$$L^{(i)}_{T_i}(x^*) = L^{(i)}_0 \exp\Big(-\frac{1}{2} \Sigma^{(i)}_{0;T_i} T_i - \Sigma^{(i)}_{0;T_i} \sqrt{T_i} \Phi^{-1} \Big(\frac{J_0^{(i)}(x^*)}{D_{0T_{i+1}}(x_0)} \Big) \Big). \quad (4.29)$$

Naturally freezing the drift is the simplest, but also the crudest approximation possible since the only piece of information retained about the processes $\{L^{(i)}_t; t \in [0, T_i]\}$ is the start value $L^{(i)}_0$, which is taken as a mean value over the whole time interval. One possible way to retain some more information about the processes $\{L^{(i)}_t; t \in [0, T_i]\}$ is the following mean value drift approximation.

4.3.2 Mean value drift approximation

The idea is the same as for freezing the drift but instead of discarding everything but the value $L^{(i)}_0$ and taking it as a mean value we approximate the drift $\mu^i(t, L^{(i+1)}_t, \dots, L^{(n)}_t)$ by using the expected value of each $L^{(j)}_t$,

$$\hat{\mu}^i(t, L^{(i+1)}_t, \dots, L^{(n)}_t) = - \sum_{j=i+1}^{n} \frac{\alpha_j E_{\mathcal{S}^{(n)}}[L^{(j)}_t]}{1 + \alpha_j E_{\mathcal{S}^{(n)}}[L^{(j)}_t]} \big(v^{(1)}_i v^{(1)}_j + v^{(2)}_i v^{(2)}_j + \rho \big(v^{(1)}_i v^{(2)}_j + v^{(2)}_i v^{(1)}_j \big) \big) \sigma(t)^2$$

for $i < n$. This yields another problem, namely the calculation of the functions $m^{(j)}_t := E_{\mathcal{S}^{(n)}}[L^{(j)}_t]$. Using backward induction we start with $j = n$:

$$E_{\mathcal{S}^{(n)}}[L^{(n)}_t] = L^{(n)}_0 \exp\Big(- \int_0^t \Big(\frac{1}{2} \big(v^{(1)}_n\big)^2 + \frac{1}{2} \big(v^{(2)}_n\big)^2 + \rho v^{(1)}_n v^{(2)}_n \Big) \sigma(s)^2 ds \Big) E_{\mathcal{S}^{(n)}} \Big[\exp\Big(\int_0^t v^{(1)}_n \sigma(s) dW^{(1)}_s + \int_0^t v^{(2)}_n \sigma(s) dW^{(2)}_s \Big) \Big].$$

Applying the same technique as in the last section we write

$$\begin{aligned} X_t :=& \int_0^t v_n^{(1)} \sigma(s) dW_s^{(1)} + \int_0^t v_n^{(2)} \sigma(s) dW_s^{(2)} \\ =& \int_0^t (v_n^{(1)} + \rho v_n^{(2)}) \sigma(s) d\tilde{W}_s^{(1)} + \int_0^t \sqrt{1-\rho^2} v_n^{(2)} \sigma(s) d\tilde{W}_s^{(2)}, \end{aligned}$$

where $\{\tilde{W}_t; t \in [0, T_n]\}$ is defined as before. This is the sum of two independent Gaussian random variables, being Gaussian itself with probability distribution

$$X_t \sim \mathbf{N}\Big(0, \Big(\big(v_n^{(1)} + \rho v_n^{(2)}\big)^2 + \big(\sqrt{1-\rho^2} v_n^{(2)}\big)^2\Big) \Sigma_{0;t}\Big),$$

where $\Sigma_{s,t} := \int_0^t \sigma(u)^2 du$. This is a well known result for independent Gaussian random variables. Further,

$$Y \sim \mathbf{N}(\beta, \eta^2) \Rightarrow E_{\mathcal{P}}[\exp(Y)] = \exp\Big(\beta + \frac{\eta^2}{2}\Big)$$

under some probability measure $\mathcal{P}$ as stated in Lemma 1.32. Therefore

$$\begin{aligned} & E_{\mathcal{S}^{(n)}} \Big[\exp\Big(\int_0^t v_n^{(1)} \sigma(s) dW_s^{(1)} + \int_0^t v_n^{(2)} \sigma(s) dW_s^{(2)}\Big)\Big] \\ & \quad = \exp\Big(\frac{1}{2}\Big(\Big(\big(v_n^{(1)} + \rho v_n^{(2)}\big)^2 + \big(\sqrt{1-\rho^2} v_n^{(2)}\big)^2\Big) \Sigma_{0;t}\Big)^2\Big) \end{aligned}$$

and

$$\begin{aligned} m_t^{(n)} =& L_0^{(n)} \exp\Big(-\int_0^t \Big(\frac{1}{2}\big(v_n^{(1)}\big)^2 + \frac{1}{2}\big(v_n^{(2)}\big)^2 + \rho v_n^{(1)} v_n^{(2)}\Big) \sigma(s)^2 ds\Big) \\ & \exp\Big(\frac{1}{2}\Big(\Big(\big(v_n^{(1)} + \rho v_n^{(2)}\big)^2 + \big(\sqrt{1-\rho^2} v_n^{(2)}\big)^2\Big) \Sigma_{0;t}\Big)^2\Big). \end{aligned}$$

Our new SDE for the $(n-1)$-th forward rate

$$\begin{aligned} d\widehat{L}_t^{(n-1)} = \hat{\mu}^{(n-1)}(t, m_t^{(n)}) \widehat{L}_t^{(n-1)} + \sum_{k=1}^{2} \widehat{L}_t^{(n-1)} \sigma_t^{(n-1,k)} dW_t^{(k)}, \\ \mathcal{S}^{(n)}, t \in [0, T_{n-1}], \widehat{L}_0^{(n-1)} = L_0^{(n-1)}, \end{aligned}$$

now admits the solution

$$\begin{aligned} \widehat{L}_t^{(n-1)} =& L_0^{(n-1)} \exp\Big(-\int_0^t \frac{\alpha_n m_s^{(n)}}{1+\alpha_n m_s^{(n)}} \Big(\sigma_s^{(n,1)} \sigma_s^{(n-1,1)} + \sigma_s^{(n,2)} \sigma_s^{(n-1,2)} + \\ & \rho\big(\sigma_s^{(n,1)} \sigma_s^{(n-1,2)} + \sigma_s^{(n,2)} \sigma_s^{(n-1,1)}\big)\Big) \\ & - \Big(\frac{1}{2}\big(\sigma_s^{(n-1,1)}\big)^2 + \frac{1}{2}\big(\sigma_s^{(n-1,2)}\big)^2 + \rho \sigma_s^{(n-1,1)} \sigma_s^{(n-1,2)}\Big) ds \\ & + \int_0^t v_n^{(1)} \sigma(s) dW_s^{(1)} + \int_0^t v_n^{(2)} \sigma(s) dW_s^{(2)}\Big) \end{aligned}$$

which may serve as a functional for $j = n-1$. Formally

$$\begin{aligned}
f^{(n-1)}(x,t) =& L_0^{(n-1)} \exp\Big(- \int_0^t \frac{\alpha_n m_s^{(n)}}{1+\alpha_n m_s^{(n)}} \Big(v_n^{(1)} v_{n-1}^{(1)} + v_n^{(2)} v_{n-1}^{(2)} + \\
& \rho\big(v_n^{(1)} v_{n-1}^{(2)} + v_n^{(2)} v_{n-1}^{(1)} \big) \Big) \sigma(s)^2 \\
& - \Big(\frac{1}{2} \big(v_{n-1}^{(1)} \big)^2 + \frac{1}{2} \big(v_{n-1}^{(1)} \big)^2 + \rho v_{n-1}^{(1)} v_{n-1}^{(2)} \Big) \sigma(s)^2 ds \\
& + \big(v_{n-1}^{(1)} + \rho v_{n-1}^{(2)} \big) x^{(1)} + \sqrt{1-\rho^2} v_{n-1}^{(2)} x^{(2)} \Big)
\end{aligned}$$

Assume the approximations $m_t^{(k)}, k = i+1, \ldots, n$, have already been calculated. The k-th approximated drift term has the form

$$\begin{aligned}
& \hat{\mu}^i(t, L_t^{(i)}, m_t^{(i+1)}, \ldots, m_t^{(n)}) = \\
& \quad - \frac{\alpha_i L_t^{(i)}}{1+\alpha_i L_t^{(i)}} \big(v_i^{(1)} v_k^{(1)} + v_i^{(2)} v_k^{(2)} + \rho\big(v_i^{(1)} v_k^{(2)} + v_i^{(2)} v_k^{(1)} \big) \big) \sigma(s)^2 \\
& \quad - \sum_{j=i+1}^{n} \frac{\alpha_j m_t^{(j)}}{1+\alpha_j m_t^{(j)}} \big(v_j^{(1)} v_k^{(1)} + v_j^{(2)} v_k^{(2)} + \rho\big(v_j^{(1)} v_k^{(2)} + v_j^{(2)} v_k^{(1)} \big) \big) \sigma(s)^2 .
\end{aligned}$$

The missing term

$$\frac{\alpha_i L_t^{(i)}}{1+\alpha_i L_t^{(i)}}$$

can be approximated as above. This yields the functional

$$\boxed{\begin{aligned}
f^{(k)}(x,t) = \quad & L_0^{(k)} \exp\Big(- \textstyle\int_0^t \sum_{j=k+1}^{n} \frac{\alpha_j m_s^{(j)}}{1+\alpha_j m_s^{(j)}} \Big(v_j^{(1)} v_k^{(1)} + v_j^{(2)} v_k^{(2)} + \\
& \rho\big(v_j^{(1)} v_k^{(2)} + v_j^{(2)} v_k^{(2)} \big) \Big) \sigma(s)^2 \\
& - \Big(\tfrac{1}{2} \big(v_k^{(1)} \big)^2 + \tfrac{1}{2} \big(v_k^{(1)} \big)^2 + \rho v_k^{(1)} v_k^{(2)} \Big) \sigma(s)^2 ds \\
& + \big(v_k^{(1)} + \rho v_k^{(2)} \big) x^{(1)} + \sqrt{1-\rho^2} v_k^{(2)} x^{(2)} \Big)
\end{aligned}}$$

Now we can proceed as in the previous section.
The mean value drift approximation presented here is clearly an improvement compared with the simple freezing-the-drift approximation, since not only the values $L_0^{(i)}$ have been used but also the expected value of the LIBORs in the drift term. Now observe the following. When working backward inductively and computing an approximation $\widehat{L}_t^{(i)}$ for $L_t^{(i)}$, the approximate LIBORs $\widehat{L}_t^{(j)}$ have already been evaluated for $j > i$ using equation (4.29). When making use of these values, too, one ends up with the Brownian bridge drift approximation.

4.3.3 Brownian bridge drift approximation

Drift approximation using the Brownian bridge is in some sense the end of the road we traveled so far. The Brownian bridge exploits the fact that when working backward inductively one needs to know $L_t^{(i+1)}, \ldots, L_t^{(n)}$ when evaluating the i-th drift term. Remember the SDEs (3.2) describing the LIBOR market model under the terminal measure $\mathcal{S}^{(n)}$. The last forward rate $L^{(n)}$ may be determined exactly since

$$dL_t^{(n)} = L_t^{(n)} \sum_{k=1}^{2} \sigma_t^{(n,k)} dW_t^{(k)}, \mathcal{S}^{(n)}, t \in [0, T_n], L_0^{(n)} = L_0,$$

is fully determined by the value of $\sum_{k=1}^{2} \int_0^t \sigma_s^{(n,k)} dW_s^{(k)}$. Assume that the forward rates $L_t^{(j)}, j = i+1, \ldots, n$, have been estimated by $\widehat{L}_t^{(j)}, j = i+1, \ldots, n$. As before an estimation of the forward rate $L_t^{(i)}$ requires us to approximate the drift terms

$$\int_0^t \frac{\alpha_j L_s^{(j)}}{1+\alpha_j L_s^{(j)}} \big(\sigma_s^{(j,1)}\sigma_s^{(i,1)} + \sigma_s^{(j,2)}\sigma_s^{(i,2)} + \rho\big(\sigma_s^{(j,1)}\sigma_s^{(i,2)} + \sigma_s^{(j,2)}\sigma_s^{(i,1)}\big)\big) - \Big(\frac{1}{2}\big(\sigma_s^{(j,1)}\big)^2 + \frac{1}{2}\big(\sigma_s^{(j,2)}\big)^2 + \rho\sigma_s^{(j,1)}\sigma_s^{(j,2)}\Big) ds \tag{4.30}$$

for $j = i+1, \ldots, n$. Here the values of $L_0^{(j)}$ are known and the values of $L_t^{(j)}$ have been estimated by $\widehat{L}_t^{(j)}$. Ignoring the drift term in $L^{(j)}$ one obtains the following approximating process,

$$d\overline{L}_t^{(j)} = \overline{L}_t^{(j)} \sum_{k=1}^{2} \sigma_t^{(j,k)} dW_t^{(k)}, \mathcal{S}^{(n)}, t \in [0, T_j], \overline{L}_0^{(j)} = L_0^{(j)}, \overline{L}_{T_j}^{(j)} = \widehat{L}_{T_j}^{(j)}. \tag{4.31}$$

Proceeding as in the last sections we set

$$\begin{aligned} X_t :=& \sum_{k=1}^{2} \int_0^t \sigma_s^{(j,k)} dW_s^{(k)} \\ =& \sum_{k=1}^{2} \int_0^t v_j^{(k)} \sigma(s) dW_s^{(k)} \\ =& \int_0^t \big(v_j^{(1)} + \rho v_j^{(2)}\big)\sigma(s) d\tilde{W}_s^{(1)} + \int_0^t \sqrt{1-\rho^2} v_j^{(2)} \sigma(s) d\tilde{W}_s^{(2)} \end{aligned}$$

to see that

$$X_t \sim \mathbf{N}\Big(0, \Big(\big(v_j^{(1)} + \rho v_j^{(2)}\big)^2 + \big(\sqrt{1-\rho^2} v_j^{(2)}\big)^2\Big)\Sigma_{s;t}\Big).$$

We may therefore take $\{W_t; t \in [0, T_j]\}$ to be some standard, one-dimensional Wiener process and write (4.31) as

$$d\overline{L}_t^{(j)} = \overline{L}_t^{(j)} \sqrt{\Big(\big(v_j^{(1)} + \rho v_j^{(2)}\big)^2 + \big(\sqrt{1-\rho^2} v_j^{(2)}\big)^2\Big)} \sigma(s) dW_t^{(k)}, \\ \mathcal{S}^{(n)}, t \in [0, T_j], \overline{L}_0^{(j)} = L_0^{(j)}, \overline{L}_{T_j}^{(j)} = \widehat{L}_{T_j}^{(j)}.$$

$\overline{L}_t^{(j)}$ follows a generalized Brownian bridge. We will replace any term $L_t^{(j)}$ occurring in the integral (4.30) with the mean $m_t^{(j)}$ of $\overline{L}_t^{(j)}$. The procedure is therefore similar to the mean value drift approximation presented in the section before. Let us calculate the mean $m_t^{(j)}$ of $\overline{L}_t^{(j)}$ at time t,

$$m_s^{(j)} = E_{\mathcal{S}^{(n)}}[\overline{L}_t^{(j)} | \overline{L}_{T_j}^{(j)} = \widehat{L}_{T_j}^{(j)}].$$

To ease notation let us first determine the mean at time t of the process

$$dY_t = \sigma_t Y_t dW_t, \mathcal{S}^{(n)}, t \in [0, T_j], Y_0 = y_0, Y_{T_j} = y_j, \tag{4.32}$$

where $\{W_t; t \in [0, T_j]\}$ is a standard, one-dimensional Wiener process under $\mathcal{S}^{(n)}$ and $\sigma : [0, T_j] \to \mathbb{R}^+$ is some function, such that the SDE (4.32) admits a solution. Also $y_0, y_j \in \mathbb{R}^+$. The solution of (4.32) when dropping the additional condition $Y_{T_j} = y_j$ is given by

$$Y_t = y_0 \exp\Big(-\frac{1}{2}\int_0^t \sigma_s^2 ds + \int_0^t \sigma_s dW_s\Big).$$

As usual set $\Sigma_{s;t} := \int_s^t \sigma_u^2 du$ and $X_t := \int_0^t \sigma_s dW_s$ for $t \in [0, T_j]$. Further , note that

$$\{\omega \in \Omega; Y_{T_j} = y_j\} = \{\omega \in \Omega; X_{T_j} = \log\Big(\frac{y_j}{y_0}\Big) + \frac{1}{2}\Sigma_{0;T_j}\}.$$

and set $x_j := X_{T_j}$. $\{X_t; t \in [0, T_j]\}$ is a Wiener process according to the Knight's theorem 1.49. The time change τ is defined by

$$\begin{aligned}\tau(s) &:= \inf\{t \geq 0; \Big\langle \int_0^{\cdot} \sigma_s dW_s \Big\rangle_t > s\} \\ &= \inf\{t \geq 0; \Sigma_{0;t} > s\}.\end{aligned}$$

Working in the time changed coordinates

$$dX_{\tau(t)} = X_{\tau(t)} dW_{\tau(t)}, \mathcal{S}^{(n)}, t \in [0, \tau(T_j)], X_0 = 0, X_{\tau(T_j)} = x_j,$$

is a standard Brownian bridge for any $x_j \in \mathbb{R}$ and so, by virtue of Theorem and Definition 1.41

$$X_{\tau(t)}|_{X_{\tau(T_j)} = x_j} \sim \mathbf{N}\Big(x_j \frac{\tau(t)}{\tau(T_j)}, \tau(t) - \frac{\tau(t)^2}{\tau(T_j)}\Big).$$

Reverting to the original time coordinates this gives

$$X_t|_{X_{T_j} = x_j} \sim \mathbf{N}\Big(x_j \frac{\Sigma_{0;t}}{\Sigma_{0;T_j}}, \Sigma_{0;t} - \frac{\Sigma_{0;t}^2}{\Sigma_{0;T_j}}\Big).$$

Using this we may evaluate the mean $m_t^{(j)}$ of Y_t under the additional condition $Y_{T_j} = y_j$,

$$
\begin{aligned}
m_t^{(j)} &= E_{\mathcal{S}^{(n)}}[Y_t | Y_{T_j} = y_j] \\
&= E_{\mathcal{S}^{(n)}}\Big[y_0 \exp\Big(-\frac{1}{2}\Sigma_{0;t} + X_t\Big) | X_{T_j} = x_j\Big] \\
&= y_0 \exp\Big(-\frac{1}{2}\Sigma_{0;t}\Big) E_{\mathcal{S}^{(n)}}\Big[\exp\big(X_t\big) | X_{T_j} = x_j\Big] \\
&= y_0 \exp\Big(-\frac{1}{2}\Sigma_{0;t}\Big) \exp\Big(x_j \frac{\Sigma_{0;t}}{\Sigma_{0;T_j}} + \frac{1}{2}\Big(\Sigma_{0;t} - \frac{\Sigma_{0;t}^2}{\Sigma_{0;T_j}}\Big)\Big) \\
&= y_0 \exp\Big(-\frac{1}{2}\Sigma_{0;t}\Big) \exp(x_j)^{\frac{\Sigma_{0;t}}{\Sigma_{0;T_j}}} \exp\Big(\frac{\Sigma_{0;t}}{2\Sigma_{0;T_j}}\big(\Sigma_{0;T_j} - \Sigma_{0;t}\big)\Big) \\
&= y_0 \Big(\frac{y_j}{y_0}\Big)^{\frac{\Sigma_{0;t}}{\Sigma_{0;T_j}}} \exp\Big(\frac{\Sigma_{0;t}}{2\Sigma_{0;T_j}}\big(\Sigma_{0;T_j} - \Sigma_{0;t}\big)\Big),
\end{aligned}
$$

applying as before Lemma 1.32

$$
X \sim \mathbf{N}(\beta, \eta^2) \Rightarrow E_{\mathcal{S}^{(n)}}[\exp(X)] = \exp\Big(\beta + \frac{\eta^2}{2}\Big)
$$

and the definition of x_j,

$$
x_j = \log\Big(\frac{y_j}{y_0}\Big) + \frac{1}{2}\Sigma_{0;t}.
$$

Substituting $\overline{L}_t^{(j)}$ yields

$$
\begin{aligned}
m_t^{(j)} &= \overline{L}_0^{(j)} \Big(\frac{\overline{L}_{T_j}^{(j)}}{\overline{L}_0^{(j)}}\Big)^{\frac{\Sigma_{0;t}^{(j)}}{\Sigma_{0;T_j}^{(j)}}} \exp\Big(\frac{\Sigma_{0;t}^{(j)}}{2\Sigma_{0;T_j}^{(j)}}\big(\Sigma_{0;T_j}^{(j)} - \Sigma_{0;t}^{(j)}\big)\Big) \\
&= L_0^{(j)} \Big(\frac{\widehat{L}_{T_j}^{(j)}}{L_0^{(j)}}\Big)^{\frac{\Sigma_{0;t}^{(j)}}{\Sigma_{0;T_j}^{(j)}}} \exp\Big(\frac{\Sigma_{0;t}^{(j)}}{2\Sigma_{0;T_j}^{(j)}}\big(\Sigma_{0;T_j}^{(j)} - \Sigma_{0;t}^{(j)}\big)\Big),
\end{aligned}
$$

where

$$
\Sigma_{s;t}^{(j)} := \int_s^t \Big(\big(v_j^{(1)} + \rho v_j^{(2)}\big)^2 + \big(\sqrt{1-\rho^2} v_j^{(2)}\big)^2\Big)\sigma(s)^2 ds.
$$

Again this yields an approximation for the drift term μ^i that is time-dependent but not state-dependent. We may therefore proceed as in the previous sections to derive the functional forms.

Zusammenfassung

In der vorliegenden Arbeit wird ein Mehrfaktormodell zur Bewertung von Zinsderivaten entwickelt. Ausgangspunkt ist dabei das eindimensionale Markov-Funktional Modell von Hunt und Kennedy.
Grundlegend in diesen Modellen ist die Annahme, dass das zufällige Verhalten eines Marktes durch einen $\mathbb{R}$-wertigen Markovprozess $\{X_t; t \in [0,T]\}$ auf einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{S}, \mathcal{P})$ für festes $T > 0$ modelliert werden kann. Weiterhin wird angenommen, dass die Preise der modellierten Wirtschaftsgüter Funktionale von $\{X_t; t \in [0,T]\}$ und strikt monoton wachsend als Funktion von $\{X_t; t \in [0,T]\}$ sind. Bezeichne $L_t^{(i)}(x)$ den Preis des i-ten Wirtschaftsgutes als Funktional von x, so gilt

$$L_t^{(i)}(x) < L_t^{(i)}(x')$$

für $x < x'$. Aufgrund dieser Annahmen existiert für alle $x^* \in \mathbb{R}$ ein eindeutiger Wert $K^{(i)}(x^*) \in \mathbb{R}$, so dass die Mengengleichung

$$\{X_t > x^*\} = \{L_t^{(i)}(X_t) > K^{(i)}(x^*)\} \tag{4.33}$$

$\mathcal{P}$-fast sicher gilt. Aufgrund von (4.33) ist die Kenntnis von $K^{(i)}(x^*)$ für alle x^* equivalent zur Kenntnis des Funktionals $L_t^{(i)}$, wodurch das Modell spezifiert wäre. Der genaue Zusammenhang zwischen $K^{(i)}(x^*)$ und x^* kann durch Marktpreise für entsprechende digitale Optionen hergeleitet werden.
Dieses Modell wird nun erweitert, indem der $\mathbb{R}$-wertige Markovprozess durch einen $\mathbb{R}^n$-wertigen Markovprozess $\{Z_t; t \in [0,T]\}$ ersetzt wird, $n \in \mathbb{N}$. Um die Gültigkeit der Gleichung (4.33) zu erhalten, wird zusätzlich ein $\mathcal{B}(\mathbb{R}^n) \otimes \mathcal{B}(\mathbb{R})$-$\mathcal{B}(\mathbb{R})$-messbares Funktional $f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ gewählt und die modifizierte Gleichung

$$\{f(Z_t, t) > x^*\} = \{L_t^{(i)}(f(Z_t, t)) > K^{(i)}(x^*)\} \tag{4.34}$$

betrachtet. Zusätzlich zur Messbarkeit muss f im ersten Argument strikt monoton wachsend sei, wobei der Begriff der stirkten Monotonie in $\mathbb{R}^n$ entsprechend definiert wird. Aufgrund dieser Voraussetzungen befinden wir uns in der gleichen Situation wie im eindimensionalen Fall und die Form des Funktionals $L_t^{(i)}$ kann aus der Kenntnis von $K^{(i)}(x^*)$ für alle $x^* \in \mathbb{R}$ hergeleitet werden. Somit wäre das Modell vollständig spezifiziert.
Zur konkreten Wahl des Funktionals f werden mehrere Vorschläge gemacht.

Grundlage sind dabei immer Driftapproximationen an Marktmodelle wie das Libor-Marktmodell oder das Swap-Marktmodell. Da Marktmodelle eine Obermenge der Markov-Funktional Modelle darstellen ist diese Vorgehensweise sinnvoll und liefert entsprechende Funktionale. Nachdem das Funktional f gewählt ist, kann das resultierende Modell durch numerische Integration ausgewertet werden.

Danksagung

Diese Arbeit wäre niemals ohne die Unterstützung einer Reihe von Leuten entstanden. Ihnen allen möchte ich an dieser Stelle meinen Dank aussprechen.
An erster Stelle wäre hier zweifelsohne Prof. Dr. Dr. Stefan Schäffler zu nennen. Ohne seine Bereitschaft, mich während dieser Arbeit zu betreuen und sich in das bis dato für ihn unbekannte Gebiet der Finanzmathematik einzuarbeiten wäre diese Arbeit niemals möglich gewesen. Ihm gebührt mein Dank.
Dr. Georg Meyer hat mich von Seiten der HypoVereinsbank betreut. Aufgrund seines umfangreichen Fachwissens konnte er mir oftmals wertvolles Feedback geben. Diese Arbeit wäre niemals erfolgreich abgeschlossen worden, hätte er mir nicht einen nicht unerheblichen Teil seiner Zeit zur Verfügung gestellt. Dafür möchte ich ihm an dieser Stelle herzlich danken.
Nicht nur die kompetente und umfassende Betreuung sowohl seitens der Univerität der Bundeswehr durch Prof. Schäffler als auch seitens der HypoVereinsbank durch Dr. Meyer war entscheidend, sondern auch der Rückhalt und die Unterstützung, die mir Frau Susanne Weiner gegeben hat. Durch ihr Verständnis und ihre Unterstützung hat sie mir sehr geholfen, wofür ich ihr ebenfalls herzlich danken möchte.
Zwei anderen Doktoranden der Siemens AG, Frau Petra Ehlers und Herrn Klaus Franz möchte ich ebenso meinen Dank aussprechen. Ihr Zuspruch hat mir durch manche schwierige Zeit geholfen.
Nicht unerwähnt lassen möchte ich an dieser Stelle auch meine Kollegen aus der HypoVereinsbank. Durch ihre wohlwollende Zustimmung und ihre Unterstützung haben sie ebenfalls zum Gelingen dieser Arbeit beigetragen.
Viele andere wären an dieser Stelle noch zu nennen, Freunde, Verwandte und Bekannte, die ich nicht einzeln aufzählen möchte. Ihnen allen möchte ich ebenfalls für ihre Unterstützung danken.

Index

Bibliography

[Bauer 1991] Heinz Bauer. *Wahrscheinlichkeitstheorie*. de Gruyter, 4th edition, 1991.

[Bauer 1992] Heinz Bauer. *Maß- und Integrationstheorie*. de Gruyter, 2nd edition, 1992.

[Black 1976] F. Black. The pricing of commodity contracts. *Journal of Financial Economics*, 3:167-179, 1976.

[Brace, Gatarek, Musiela 1997] A. Brace, D. Gatarek and M. Musiela. The market model of interest rate dynamics. *Math. Finance*, 7:124-154, 1997.

[Brigo, Mercurio 2001] Damiano Brigo, Fabio Mercurio. *Interest Rate Models - Theory and Practice*. Springer, 2001.

[Dupire 1994] Bruno Dupire. Pricing with a smile. *Risk*, 9(3):18-20, 1994.

[Geman et al. 1995] H. Geman, N. El Karoui, J.C. Rochet. Changes of Numeraire, Changes of Probability Measure and Pricing of Options. *Journal of Applied Probability*, 32:443-458, 1995.

[Hoffmann-Jørgensen 1994] J. Hoffmann-Jørgensen. *Probability with a view toward statistics*. Chapman & Hall, 1994.

[Hunt et al. 2000] Phil Hunt, Joanne Kennedy, Antoon Pelsser. Markov-functional interest rate models. *Finance and Stochastics*, 4:391-408, 2000.

[Hunt, Kennedy 2000] P.J. Hunt and J.E.Kennedy. *Financial Derivatives in Theory and Practice*. Wiley, 2000.

[Jamshidian 1997] F. Jamshidian. Libor and Swap market models and measures. *Finance and Stochastics*, 1(4):293-330, 1997.

[Karatzas, Shreeve 2001] I. Karatzas and S. Shreeve, *Stochastic calculus and Brownian motion*. Springer, 2001.

[Kerkhof, Pelsser 2002] Jeroen Kerkhof and Antoon Pelsser. Observational equivalence of Discrete String Models and Market Models. *The Journal of Derivatives*, 3:55-61, 2002.

[Musiela, Rutkowski 1998] Marek Musiela, Marek Rutkowski. *Martingale Methods in Financial Modeling*. Springer, 1998.

[Øksendal 1998] Bernt Øksendal. *Stochastic differential equations - An Introduction with Applications*. Springer, 5th edition, 1998.

[Protter 1990] P. Protter. *Stochastic Integration and Differential equations*. Springer, 1990.

[Schäffler 1996] Stefan Schäffler und Thomas F. Sturm. *Stochastische Differentialgleichungen - Theorie, Numerik, Anwendungen* Schriftenreihe des Instituts für Angewandte Mathematik und Statistik der Technischen Universitaet Muenchen, volume 7, 1996.

[Schäffler, Sturm 1994] Stefan Schäffler, Thomas F. Sturm. *Wahrscheinlichkeitstheorie und Statistik I*. Schriftenreihe des Instituts für Angewandte Mathematik und Statistik der Technischen Universitaet Muenchen, volume 5, 1994.

[Schäffler, Sturm 1995] Stefan Schäffler, Thomas F. Sturm. *Wahrscheinlichkeitstheorie und Statistik II*. Schriftenreihe des Instituts für Angewandte Mathematik und Statistik der Technischen Universitaet Muenchen, volume 6, 1995.

[Shiryaev 1995] A.N. Shiryaev. *Probability*. Springer, 2nd edition, 1995.