Positive numerical integration of Stochastic Differential Equations

Diploma Thesis

Christian Kahl

Supervisor ABN AMRO London
Dr. Thilo Roßberg

Supervisor University of Wuppertal
Prof. Dr. Michael Günther

University of Wuppertal
Faculty of Mathematics and Natural Science
Research Group Numerical Analysis
September 9, 2004

Acknowlegdement

The present diploma thesis arises from a cooperation between the University of Wuppertal, Faculty of Mathematics and Natural Sciences and ABN AMRO, especially the Quantitative Analytics Group under the direction of Dr. Marcel van Regenmortel.
Special thanks go to Dr. Thilo Roßberg. He spared no effort to support me in solving this interesting problem. Within an uncountable number of telephone calls he facilitates my access to financial mathematics.
The contact to ABN AMRO has been arranged by Prof. Dr. Michael Günther. He was not only the source for this cooperation but also the supervisor of this project. Whenever I needed some advice Mr. Günther spent his time. Furthermore I would like to thank the members of the research group Numerical Analysis of the University of Wuppertal. My cordial thanks go to Prof. Dr. Silke Schlosser-Haupt who introduced me the multifarious world of numerics and made the contact to Prof. Dr. Michael Günther.
Additionally I would like to thank Cathrin van Emmerich for the helpful discussions and her support in completing this thesis. Last not least thanks are due to my family for the unfatiguing assistance and encouragement during my studies.

Wuppertal, May 2004 Christian Kahl

Table of Contents

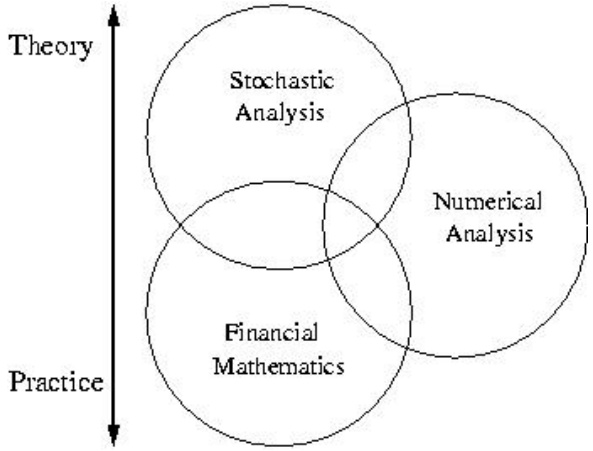
List of Figures

List of Tables

Introduction

This thesis deals with positivity preserving numerical integration schemes for stochastic differential equations (SDEs), particularly in the context of pricing interest rate derivatives. What is at issue here is a part of computational finance as a subarea of applied mathematics.

Applied mathematics can be understood as a link between theory and praxis. Stochastic analysis is the theoretical background in this thesis, especially stochastic differential equations as indicated by the title. Stochastic differential equations allow to model observables that are driven by non-deterministic phenomena.
On the other hand this thesis deals with financial mathematics. In fact the problem of deriving numerical integration schemes that preserves positivity arises from pricing interest rate derivatives. This requires the knowledge of another area of applied mathematics: numerical analysis.



Numerical integration of stochastic differential equations is one particular part of numerical analysis. As for deterministic systems, geometric integration schemes are mandatory if essential structural properties of the underlying system have to be preserved. In this thesis we focus on positive

integration schemes as the analytical solution of the SDE is positive as well. This problem is rather independent of financial mathematics as it appears in various other applications. Thus also from a general point of view one may aim at answering the question whether it is possible to develop a numerical integration scheme which guarantees positivity, if the considered stochastic process is positive, as well.

In this thesis the mathematical modelling of interest rates is based on the Libor market model. In recent years the Libor market model has been continuously refined. There are quite a few extensions but I want to concentrate on Constant Elasticity of Variance (CEV) and Displaced Diffusion (DD). Introducing stochastic volatility improves the dynamics of the model. Stochastic volatility then turns out to be a main part of the study.
All extensions have in common that numerical positivity of the integration scheme decisively influences the quality of results.

The thesis is structured as follows:
The first chapter recalls some basic facts about Ito Diffusions, among other things the theorems of Feynman-Kac and Girsanov. Moreover some criteria for analytical positivity will be developed.
The second part deals with financial mathematics and motivates the need for numerical schemes when pricing interest rate derivatives.
The third chapter gives a detailed description of numerical schemes for stochastic differential equations. In order to develop some stochastic integration schemes stochastic Taylor expansions are studied. Another important topic is the Milstein theorem.
The next chapter is the very heart of this thesis. The integration schemes developed above are examined concerning positivity.
The last chapter verifies the theoretical results by numerical tests. The thesis is concluded by pointing out the main results in a summary.

CHAPTER I

Diffusion theory

Diffusion theory describes one special class of stochastic processes. It is of much interest in this thesis as it provides many tools to model financial mathematical relations. This first chapter will only repeat some results of diffusion theory which are essential for further modelling in finance. The reader who is not familiar with stochastic processes will find the necessary facts about stochastic analysis in appendix A.

There are two different topics discussed in this introductory chapter. We start with a presentation of some essential results from diffusion theory, especially the theorems of Feynman-Kac and Girsanov. Secondly we consider the aspects of analytical positivity of a diffusion process to enable an adequate modelling.

1.1 Basic results

A stochastic process is called a diffusion process if it satisfies the (strong) Markov property and if its paths X_t are continuous functions[KT81].
We do not need this abstract definition of a diffusion process but only a special case.

Definition 1.1 (Ito diffusion) An Ito diffusion is a time homogeneous stochastic process $X_t : [t_0, \infty) \times \Omega \to \mathbb{R}^n$ which is the solution of the following stochastic differential equation

$$dX_t = a(X_t)dt + b(X_t)dB_t \text{ with } X_{t_0} = x, t > t_0. \tag{1.1}$$

Where B_t is an m-dimensional Brownian motion. The coefficient $a : \mathbb{R}^n \to \mathbb{R}^n$

is called the drift and $b : \mathbb{R}^n \to \mathbb{R}^{n\times m}$ the diffusion. In this chapter we need only the one-dimensional case with $n = m = 1$.

It is essential in the theory of Ito diffusions to study the analytical behaviour only based on the knowledge of drift a and diffusion b. Therefore it is not necessary to know the exact solution of the stochastic differential equation. In this context the fundamental principle is the so-called generator of a stochastic process X_t.

Definition 1.2 Let X_t be an Ito diffusion. Then the generator A is defined by

$$Af(x) = \lim_{t\to 0} \frac{E^x\left[f(X_t)\right] - f(x)}{t}. \tag{1.2}$$

We denote the set of all functions such that the limit exists by $\mathcal{D}_A$. Later on we see that $C_0^2(\mathbb{R}) \subset \mathcal{D}_A$.

This definition is generally quite complicated. To find a relation between the generator A and the coefficients a and b we need some basic calculation. First we define a function $f \in C_0^2(\mathbb{R})$ and apply Ito's theorem A.25

$$\begin{aligned} f(X_t) = f(X_0) \quad &+ \int_0^t \left(a(X_s)\frac{\partial f}{\partial x}(X_s) + \frac{1}{2}b^2(X_s)\frac{\partial^2 f}{\partial x^2}(X_s) \right) ds \\ &+ \int_0^t b(X_s)\frac{\partial f}{\partial x}(X_s) dB_s. \end{aligned}$$

Setting up the expectation one gets the following equation by using the martingale property of the Ito integral (see (A.14))

$$E\left[f(X_t)\right] = f(X_0) + E\left[\int_0^t \left(a(X_s)\frac{\partial f}{\partial x}(X_s) + \frac{1}{2}b^2(X_s)\frac{\partial^2 f}{\partial x^2}(X_s) \right) ds \right].$$

This schematic approach is not mathematically rigid. A complete proof can be found in chapter 7 of [Øks00]. This equation leads to a simple representation of the generator of the Ito diffusion.

Theorem 1.3 Let X_t be an Ito diffusion given by

$$dX_t = a(X_t)dt + b(X_t)dB_t.$$

Then the equation

$$Af(x) = a(x)\frac{\partial f}{\partial x}(x) + \frac{1}{2}b^2(x)\frac{\partial^2 f}{\partial x^2} \tag{1.3}$$

holds for the generator A and $f \in C_0^2(\mathbb{R})$. Applying the calculation above one gets in an analogous manner

Theorem 1.4 (Dynkin's formula) Let $f \in C_0^2(\mathbb{R})$ and X_t be an Ito diffusion then

$$E^x\left[f(X_t)\right] = f(x) + E^x\left[\int_0^t Af(X_s)ds\right].$$

Dynkin's formula is the starting-point to deduce one of the main results in the theory of stochastic analysis. This is the theorem of Feynman-Kac. For that reason we define a function u as follows

$$u(t,x) = E^x\left[f(X_t)\right].$$

Looking at Dynkin's formula and computing the partial derivative with respect to t one see that

$$\frac{\partial u}{\partial t} = E^x\left[Af(X_t)\right].$$

The generator A of the Ito diffusion commutates with the expectation. Hence one gets the so-called Kolmogorov's backward equation.

Theorem 1.5 (Kolmogorov's backward equation) Let $f \in C_0^2(\mathbb{R})$ and X_t be an Ito diffusion. Then the following statements hold:

(i) Define

$$u(t,x) = E^x\left[f(X_t)\right]. \tag{1.4}$$

Then $u(t,\cdot) \in \mathcal{D}_A$ for each t and

$$\frac{\partial u}{\partial t} = Au,\ t > t_0, \tag{1.5}$$

$$u(0,x) = f(x). \tag{1.6}$$

(ii) Otherwise if $w(t,x) \in C^{1,2}(\mathbb{R}\times\mathbb{R})$ is a bounded function solving the equations (1.5) and (1.6) then $u(t,x) = w(t,x)$ in (1.4).

This connection is quite amazing because it shows that it is possible to get the expectation of a stochastic process by solving a partial differential equation.
If one is interested in a numerical solution of this problem the Kolmogorov backward equation allows to choose between two ways to reach it. Firstly one can do a Monte Carlo simulation to calculate the expectation of the stochastic process. Secondly one can use a discretisation scheme to solve the partial differential equation. Numerical integration schemes as the main interest of this thesis are described detailed in chapter 3.

With a little more effort we obtain the following useful generalisation of Kolmogorov's backward equation:

Theorem 1.6 (Feynman-Kac) Let $f \in C_0^2(\mathbb{R})$ and assume that $q \in C(\mathbb{R})$ is a lower bounded function. Then the following statements hold for an Ito diffusion X_t:

(i) Define

$$v(t,x) = E^x \left[\exp\left(-\int_0^t q(X_s)ds \right) f(X_t) \right], \tag{1.7}$$

then for all $t > t_0$:

$$\frac{\partial v}{\partial t} = Av - qv, \tag{1.8}$$

$$v(0,x) = f(x). \tag{1.9}$$

(ii) If there is a bounded function $w(t,x) \in C^{1,2}(\mathbb{R} \times \mathbb{R})$, satisfying (1.8) and (1.9), then $u(t,x) = w(t,x)$ in equation (1.7).

Remark 1.7 Demanding the boundary condition (1.9) for the end of the interval $[0,T]$ instead of the initial point

$$v(T,x) = f(x)$$

then the Feynman-Kac theorem holds if one adapts the equation (1.8) as follows:

$$-\frac{\partial v}{\partial t} = Av - qv.$$

Example 1.8 (Heat conduction) The next equation describes a one-dimensional heat conduction

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} - qu \text{ with a source term } qu.$$

If one likes to apply the Feynman-Kac theorem it is necessary to find the stochastic process with the appropriate generator

$$Au = \frac{\partial^2 u}{\partial x^2}.$$

The Brownian motion has got the right property as one can easily verify. Hence the solution is given by

$$u(t,x) = E^x \left[\exp\left(-\int_0^t q(B_s)ds \right) f(B_t) \right] = \exp(-qt) E^x \left[f(B_t) \right].$$

As the Brownian motion is $\mathcal{N}(0,t)$-distributed for a fixed t, we get

$$u(t,x) = (2\pi t)^{-1/2} \exp(-qt) \int_{\mathbb{R}} f(t) \exp\left(-\frac{(z-x)^2}{2t} \right) dt.$$

This way we are able to reach a closed solution with respect to an appropriate boundary condition f.

Another important tool used in stochastic analysis is still missing to solve the problems in financial mathematics. This is Girsanov´s theorem. It describes what happens to the coefficients of a stochastic process if the probability measure is changed. In this case also the Radon-Nikodym (see equation (A.1)) derivative plays an essential role.

Lemma 1.9 Let $\varphi \in \mathcal{V}$ satisfy the so-called Novikov condition

$$E\left[\exp\left(\frac{1}{2} \int_0^T \varphi^2(s,\omega)ds \right) \right] < \infty \tag{1.10}$$

where $\mathcal{V}$ is the set of all Ito integrable functions (see def. A.21). Then the stochastic process

$$Z_t = \exp\left\{ \int_0^t \varphi(s,\omega)dB_s - \frac{1}{2} \int_0^t \varphi^2(s,\omega)ds \right\} \text{ with } 0 \le t \le T$$

is a martingale if $\varphi(s,\omega) \cdot Z_t \in \mathcal{V}$.

Proof: Applying Ito's theorem leads to:

$$\begin{aligned} dZ_t &= -\frac{1}{2}\varphi^2(s,\omega)Z_t dt + \varphi(s,\omega)Z_t dB_t + \frac{1}{2}\varphi^2(s,\omega)Z_t(dB_t)^2 \\ &= \varphi(s,\omega)Z_t dB_t. \end{aligned}$$

Now one can use Lemma A.24 which proves the statement. □

Theorem 1.10 (Girsanov) Let Y_t be an Ito process represented by the following differential equation

$$dY_t = a(s,\omega)dt + b(s,\omega)dB_t.$$

Assume there exist processes σ and γ such that

$$b(s,\omega)\sigma(s,\omega) = a(s,\omega) - \gamma(s,\omega),$$

with σ satisfying the Novikov condition. Setting

$$Z_t = \exp\left\{ -\int_0^t \sigma(s,\omega)dB_s - \frac{1}{2}\int_0^t \sigma^2(s,\omega)ds \right\},$$

this defines a new measure

$$dQ(\omega) = Z_T(\omega)dP(\omega), \tag{1.11}$$

and

$$\tilde{B}_t := \int_0^t u(s,\omega)ds + B(t) \tag{1.12}$$

is a Brownian motion w.r.t. Q. The stochastic process Y_t now has a new representation in terms of $\tilde{B}_t$:

$$dY_t = \gamma(s,\omega)dt + b(s,\omega)d\tilde{B}_t. \tag{1.13}$$

The transformation $P \to Q$ is called the Girsanov transformation of measures.

Proof: One can find the proof in chapter 8 of [Øks00]. A more general version can be found in [HT94, KS88] □

The Girsanov theorem allows to change the measure of the underlying probability space. This important result has many applications in economics especially in the context of arbitrage in the financial markets. There is often the situation that one knows both measures P, Q. Then the martingale property of the Radon-Nikodym derivative has to be checked to change the measures in the same manner. The Girsanov theorem completes the most fundamental results in the theory of Ito diffusions.

1.2 Analytical positivity

If we model scientific processes with stochastic differential equations we usually want to get a reproduction of reality which is as detailed as possible. Thereby the differential equation should correctly describe the interactions. Furthermore the stochastic process as the solution of the equation must have the analytical properties given by the model.
A simple example is a stochastic differential equation describing population. It is quite clear that the stochastic process must be nonnegative. One should bear in mind this basic fact.
Even in financial mathematics there is a bunch of processes which have to fulfil certain regularity requirements, e. g., the chart of a stock or an interest rate must take positive values. The volatility of a stock can be modelled by a mean-reverting process. In this case as well it does not make sense to allow negative values.
Our aim is now to develop criteria for the drift and diffusion term of a given stochastic process. We obtain some results about the boundary behaviour without necessarily knowing the exact solution. A rather complete description and corresponding proofs can be found in [KT81].

Again an Ito diffusion is the starting-point. Drift and diffusion term are assumed to be known and continuous.

Definition 1.11 (Hitting time) Let X_t be a stochastic process with initial value x_0. Then the hitting time is given by

$$T_x = \begin{cases} \infty & : \text{if } X(t) \neq x, \text{ for all } t \\ \inf\{t \geq 0 \ , \ X_t = x\} & : \text{else} \end{cases}$$

Furthermore we define

$$T^* = \min\{T_a, T_b\}$$

if $a < x_0 < b$.

It is necessary to describe the generator (1.3) of an Ito diffusion in a slightly different way to study the problem of analytical positivity. Therefore we define the following functions

$$s(x) = \exp\left(-\int_{x_0}^{x} \frac{2a(t)}{b^2(t)} dt\right), \tag{1.14}$$

$$m(x) = \frac{1}{s(x)b^2(x)}. \tag{1.15}$$

We call s the scale function and m the speed density. Obviously the following equation holds for s:

$$\frac{s'(x)}{s(x)} = -\frac{2a(x)}{b^2(x)}.$$

Then, following a classical approach, we achieve a new representation for the generator

$$Lf(x) = \frac{1}{2}\left(\frac{1}{1/(b^2(x)s(x))}\right)\frac{d}{dx}\left[\frac{1}{s(x)}\frac{df}{dx}(x)\right].$$

To obtain a more succinct and meaningful expression for L, write $dM = m(x)dx$ and $dS = s(x)dx$. In terms of these differentials the operator L is simply

$$Lf(x) = \frac{1}{2}\frac{d}{dM}\left[\frac{df}{dS}(x)\right]. \tag{1.16}$$

To receive the relation between the generator and the boundary behaviour of the diffusion process we have to consider the next two problems.

Proposition 1.12 (Problem A) Let

$$u(x) = P\{T_b < T_a | X(0) = x\} = E^{0,x}\left[1_{\{T_b<T_a\}}(x)\right].$$

This is the probability of reaching the point b before a. The theorem of Feynman-Kac proofs that this is equivalent to the boundary-value problem

$$0 = a(x)\frac{du}{dx} + \frac{1}{2}b^2(x)\frac{d^2u}{dx^2} = Lu, \tag{1.17}$$

with $u(a) = 0$ and $u(b) = 1$. Hereby the solution is given by

$$u(x) = \frac{S(x) - S(a)}{S(b) - S(a)}.$$

Proof: Having the representation of the generator (1.16) the partial differential equation (1.17) is equivalent to

$$\frac{1}{2}\frac{d}{dM}\left[\frac{du}{dS}(x)\right] = 0.$$

Integration leads to

$$\begin{aligned}
& \tfrac{du}{dS} &=& K \\
\iff \quad & du &=& K\,dS \\
\iff \quad & \int_a^x du &=& \int_a^x K\,dS \\
\iff \quad & u(x) - u(a) &=& K(S(x) - S(a)).
\end{aligned}$$

The boundary condition $u(b) = 1$ completes the proof. □

Proposition 1.13 (Problem B) Let

$$v(x) = E^{0,x}\left[T^*\right].$$

Thus v is the expected time to reach a or b. This can be written as a boundary-value problem, too:

$$-1 = a(x)\frac{dv}{dx} + \frac{1}{2}b^2(x)\frac{d^2v}{dx^2} = Lv \tag{1.18}$$

with the boundary conditions $v(a) = v(b) = 0$. The solution is given by

$$v(x) = 2\left(u(x)\int_x^b S[z,b]dM(z) + (1-u(x))\int_a^x S[a,z]dM(z)\right).$$

Proof: One obtains this result in an analogous manner as in problem A. For a detailed proof see [KT81]. □

The results so far enable us to classify the boundary behaviour of a stochastic process with respect to the functions u and v. This way we can distinguish between what is attractive and what is attainable.

Definition 1.14 Let

$$S(0,x] := \lim_{a\to 0} S[a,x], \tag{1.19}$$

$$\Sigma(0) := \int_0^x S(0,z]dM(z) = \int_0^x M[z,x]dS(z). \tag{1.20}$$

Then the boundary 0 is called attractive if

$$S(0,b] < \infty.$$

We classify 0 as attainable if

$$\Sigma(0) < \infty.$$

The difference between attractivity and attainability is that the stochastic process does not necessarily reach the value of an attractive boundary in finite time. Otherwise if a boundary is attainable it is attractive as well.

$$\begin{aligned} \Sigma(0) < \infty &\implies S(0,x] < \infty \\ S(0,x] = \infty &\implies \Sigma(0) = \infty \end{aligned}$$

Remark 1.15 The assertions of the boundary 0 are independent from the value x. Further it is possible to generalise the boundary behaviour for arbitrary lower bounds $\ell \in \mathbb{R}$. One gets the equation (1.20) for $\Sigma(0)$ by partial integration as a short calculation shows

$$\begin{aligned}
\int_0^x S(0,z]dM(z) &= \int_0^x (S(z) - S(0))\, m(z)dz \\
&= S(x)M(x) - S(0)M(0) \\
&\quad - \int_0^x s(z)M(z)dz - \int_0^x S(0)m(z)dz \\
&= S(x)M(x) - S(0)M(x) - \int_0^x s(z)M(z)dz \\
&= \int_0^x s(z)M(x) - s(z)M(z)dz \\
&= \int_0^x M[z,x]dS(z).
\end{aligned}$$

Now we can analyse the positivity of a diffusion process in an elementary way by just using drift and diffusion term.
The following examples are presented particularly with regard to modelling in financial mathematics. The technique used above can be applied to other problems as well.

Application in mathematical finance

The easiest example is a geometric Brownian motion. In this case it is not necessary to prove the positivity because it can be directly deduced from the exact solution.

Lemma 1.16 Let X_t be a geometric Brownian motion

$$dX_t = \mu X_t dt + \sigma X_t dW \tag{1.21}$$

with $\mu, \sigma \in \mathbb{R}^+$. Then the boundary 0 is not attainable.

Proof: With $a(x) = \mu x$ and $b(x) = \sigma x$ we get

$$\begin{aligned} s(x) &= \exp\left(-\int_{x_0}^{x} \frac{2\mu z}{\sigma^2 z^2} dz\right) \\ &= \exp\left(-\frac{2\mu}{\sigma^2}\log(x)\right) \\ &= x^{-\frac{2\mu}{\sigma^2}}. \end{aligned}$$

Therefore

$$S(0,x] = \infty \text{ if } \mu > \sigma^2/2 \text{ and } S(0,x] < \infty \text{ else.}$$

So far we have not got a clear indication of the boundary 0. Thus we have to study the behaviour of $\Sigma(0)$. Let $\lambda = -\frac{2\mu}{\sigma^2}$ then

$$\begin{aligned} \Sigma(0) &= \int_0^x \left(\int_0^y s(z)dz\right) m(y)dy \\ &= \int_0^x \frac{1}{1-\lambda} y^{1-\lambda} \frac{y^\lambda}{\sigma^2 y^2} dy \\ &= \int_0^x C\frac{1}{y} dy \\ &= \infty. \end{aligned}$$

As demonstrated above the stochastic process never reaches 0. Hence it is strictly positive. □

In the next chapter we will use two further stochastic differential equations for modelling in financial mathematics. Already at this point we present their analytical properties.

Proposition 1.17 The stochastic process X_t given by

$$dX_t = (a + bX_t)dW \tag{1.22}$$

where $a \in \mathbb{R}$ and $b \in \mathbb{R}^+$ has the following characteristics:

- $X_t \in [-a/b, \infty]$
- $-a/b$ is an attractive boundary but not attainable.

Proof: The calculation can be done in an analogous way as for the geometric Brownian motion. □

Proposition 1.18 Now let X_t be given by

$$dX_t = X_t^{\alpha} dW \ , \ \alpha \geq 0. \tag{1.23}$$

Then X_t possesses the following properties

- $0 < \alpha < 1 \Rightarrow 0$ is an attainable boundary,
- $1 \geq \alpha \Rightarrow 0$ is an unattainable boundary,
- $\forall \alpha \in \mathbb{R}_0^+$ 0 is an attractive boundary.

Proof: One can verify these assertions by basic calculation. □

Mean-reverting

A mean-reverting process is one of the most often used stochastic differential equations in scientific applications. The conception is that the mean-reverting process varies about a mean often called equilibrium. For example the volatility in financial applications can be described by such a process. In this case the positivity of the stochastic process is of much interest, too.

Definition 1.19 The class of mean-reverting processes is the solution of the following differential equation:

$$dX_t = (\alpha(t) - \beta(t)X_t)dt + \sigma(t)X_t^p dW.$$

where $\alpha, \beta, \sigma : \mathbb{R} \longrightarrow \mathbb{R}^+$ are continuous functions and $p \in \mathbb{R}^+$.

The mean-reverting processes can also be examined by the methods we developed above.

Lemma 1.20 Let the stochastic process X_t be given by the next stochastic differential equation with $\alpha, \beta, \sigma \in \mathbb{R}^+$:

$$dX_t = (\alpha - \beta X_t)dt + \sigma X_t^{\frac{1}{2}} dW. \tag{1.24}$$

Then the boundary 0 is not attractive if $\alpha > \sigma^2/2$.

Proof: We start with computing the function s

$$\begin{aligned} s(x) &= \exp\left(-\int_{x_0}^{x} \frac{2(\alpha-\beta z)}{\sigma^2 z} dz\right) \\ &= \exp\left(-\log(x)\frac{2\alpha}{\sigma^2} + x\frac{2\beta}{\sigma^2}\right) \\ &= x^{-\frac{2\alpha}{\sigma^2}} \exp\left(x\frac{2\beta}{\sigma^2}\right). \end{aligned}$$

This calculation leads to the following result for measure S

$$S(0,x] = \infty \text{ if } \alpha > \sigma^2/2 \text{ and } S(0,x] < \infty \text{ else.}$$

Accordingly we must only check the case where $\alpha \leq \sigma^2/2$ (to simplify notation let $\lambda = \frac{2\alpha}{\sigma^2}$):

$$\begin{aligned} \Sigma(0) &= \int_0^x \left(\int_0^y s(z)dz\right) m(y)dy \\ &= \int_0^x C y^{1-\lambda} \frac{y^\lambda}{\sigma^2 y} dy \\ &= \int_0^x C/\sigma^2 dy \\ &< \infty. \end{aligned}$$

Thus the lower bound 0 is always attainable if $\alpha \leq \sigma^2/2$. □

Proposition 1.21 The stochastic process given by

$$dX_t = (\alpha - \beta X_t)dt + \sigma X_t^p dW \tag{1.25}$$

with $\alpha, \beta, \sigma, p \in \mathbb{R}^+$, has an unattractive bound 0 if

1. $p = \frac{1}{2}$ and $\alpha > \sigma^2/2$,
2. $p > \frac{1}{2}$.

Proof: The first assertion follows from Lemma 1.20. To prove the second one, we must calculate the function s:

$$s(y) = \exp\left(-\frac{2\alpha}{\sigma^2(1-2p)}y^{1-2p}\right)\exp\left(\frac{2\beta}{\sigma^2(2-2p)}y^{2-2p}\right).$$

The second term is bounded below on $[0,x]$ by a constant C. However a $x_0 > 0$ exists so that

$$\exp\left(-\frac{2\alpha}{\sigma^2(1-2p)}x_0^{1-2p}\right) = x_0^{-1}.$$

Hence we get for an arbitrary $x > x_0$

$$\begin{aligned} S(0,x] &> S(0,x_0] \\ &= \int_0^{x_0} C\exp\left(-\frac{2\alpha}{\sigma^2(1-2p)}y^{1-2p}\right)dy \\ &> \int_0^{x_0} y^{-1}dy \\ &= \infty. \end{aligned}$$

We notice that in the case $p > \frac{1}{2}$ the boundary 0 is not even attractive. □

To sum up we could clear the problem of analytical positivity for a class of stochastic processes which are of great importance in mathematical finance. That leads to the question if there are numerical schemes which preserve this regularity property.
We will deal with this problem in chapter 4. Doing so we will mainly concentrate on the processes (1.21),(1.22), (1.23),(1.24) and (1.25).

CHAPTER II

Mathematical modelling of interest rates derivatives

In finance there is a wide range of different derivatives to hedge the risk of losing money. A put on a stock, e. g., can protect portfolios against falling stock prices. In consideration of the fact that a stock option is the easiest example, appendix B explains this concept more detailed. Not only stocks are exposed to random fluctuations but interest rates as well. The interest rate underlying the European market is the Euribor (see fig. 2.1).

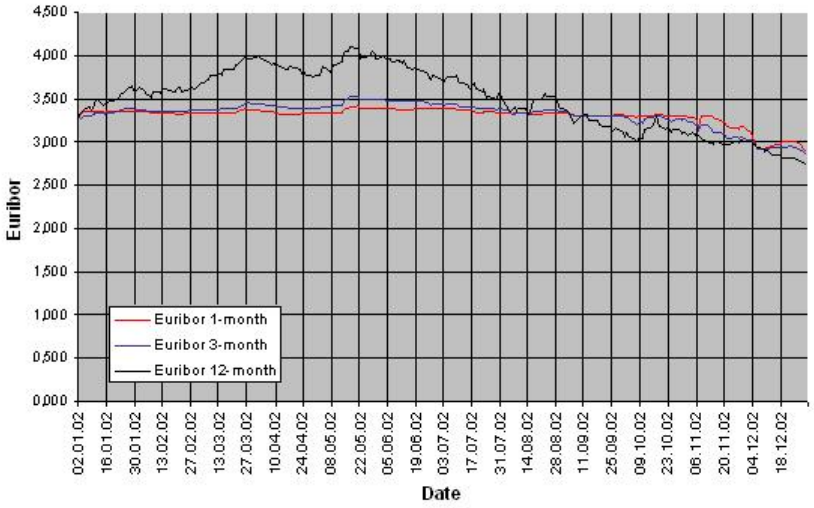


Figure 2.1: *Euribor* chart from the year 2002; data provided by www.euribor.org

The first question we have to deal with is how to trade interest rates. As interest rates are not directly tradable, the market uses zero-coupon bonds $B(t,T)$ as a substitute:

$$B(t,T) = \text{ price at time } t \text{ paying 1 at time } T. \tag{2.1}$$

The different dates of expiration T define the so-called maturity structure

$$0 = T_0 < T_1 < ... < T_N < T_{N+1} = T^*. \tag{2.2}$$

Usually there are 3-months periods between the different dates of expiration. Knowing the price of a zero-coupon bond a further description can be given.

Definition 2.1 (Libor forward rates) The forward rate $F_k(t)$ is the arbitrage-free interest rate in the time period from T_k to T_{k+1}

$$1 + \delta_k F_k(t) = \frac{B(t,T_k)}{B(t,T_{k+1})} \text{ with } \delta_k = T_{k+1} - T_k.$$

Obviously this leads to the following expression for $F_k(t)$

$$F_k(t) = \frac{1}{\delta_k}\left(\frac{B(t,T_k)}{B(t,T_{k+1})} - 1\right).$$

Therefore the bond prices determine the interest curve and vice versa, the forward rates describe the forward curve as shown in figure 2.2.

The prices of zero-coupon bonds contain certain expectations about the future development of interest rates. As in the equity market, see appendix B, we want to find a stochastic model covering these expectations.

2.1 Libor market model

There are options on forward rates to hedge against the fluctuations of interest rates. Considering the stock options call and put as the basic derivatives (see again app. B), the equivalents in the interest rate market are caplets (call) and floorlets (put).

Definition 2.2 (Caplet) A caplet guarantees a payoff at T_k to the holder if the forward rate F is higher than the strike K

$$C(T_k) = \delta_k B(0,T_{k+1})(F(T_k) - K)^+.$$

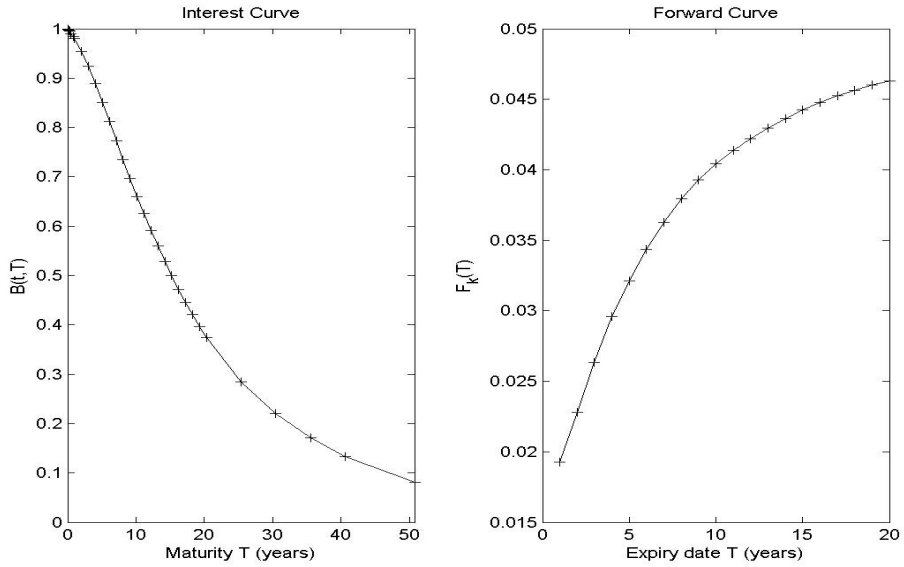


Figure 2.2: Interest curve and forward curve, prices of zero-coupon bonds denoted in Euro at 2004/03/30, data provided by Dr. Thilo Roßberg ABN AMRO London.

Analogously the payoff function f of a floorlet is defined by

$$F(T_k) = \delta_k B(0, T_{k+1})(K - F(T_k))^+.$$

So we can use a caplet as a hedging tool against interest rates fluctuations at a fixed date. If we are interested in hedging at every moment T_i we need a combination of caplets. Such a combination is called cap.
Considering the modelling of forward rates we have to deal with two problems. On the one hand we have to describe an arbitrage-free market. On the other hand the model has to fit the market prices.
Similar to stock options we observe that the relative caplet prices are different depending on the strike K. Likewise we find that there is a variation which depends on maturity. Exactly this problem will be of further interest later on.
The first aim is setting up an arbitrage-free model. To do so we need to define an appropriate measure.

Definition 2.3 (Forward measure) The so-called forward measure $\mathbb{Q}_{T_k}$ is the measure associated to the zero-coupon bond $B(t, T_k)$.

One gets an arbitrage-free market if the forward rate F_k fulfils the following stochastic differential equation.

$$dF_k(t) = \sigma_k(t)F_k(t)dB^{T_{k+1}} \text{ with } t < T_k \tag{2.3}$$

where $B^{T_{k+1}}$ is a Brownian motion with regard to the measure $\mathbb{Q}_{T_{k+1}}$.

Using this definition one gets N stochastic differential equations modelling N different forward rates from $t = 0$ to T_N. If one only has to price caplets or caps it is not necessary to construct a common measure for all forward rates. But determining the value of a more complex option requires the same measure for all forward rates.

First we analyse the stochastic differential equation for F_k changing the measure from $\mathbb{Q}_{T_{k+1}}$ to $\mathbb{Q}_{T_k}$. We consider the Radon-Nikodym derivative Z (see equation (A.1))

$$\begin{aligned} Z(t) = \frac{d\mathbb{Q}_{T_k}}{d\mathbb{Q}_{T_{k+1}}} &= \frac{B(t,T_k)/B(0,T_k)}{B(t,T_{k+1})/B(0,T_{k+1})} \\ &= \frac{B(t,T_k)/B(t,T_{k+1})}{B(0,T_k)/B(0,T_{k+1})} \\ &= \frac{1+\delta_k F_k(t)}{1+\delta_k F_k(0)}. \end{aligned}$$

Differentiating Z leads to

$$dZ(t) = \frac{\delta_k dF_k(t)}{1+\delta_k F_k(0)} = \frac{\delta_k F_k(t)}{1+\delta_k F_k(0)}\sigma_k(t)dB^{T_{k+1}}.$$

It follows that Z is a martingale and that

$$\frac{dZ(t)}{Z(t)} = \underbrace{\frac{\delta_k \sigma_k(t)F_k(t)}{1+\delta_k F_k(t)}}_{\alpha(t)} dB^{T_{k+1}}.$$

Consequently a geometric Brownian motion describes Z. Obviously the solution is

$$Z_t = \exp\left\{\int_0^t \alpha(s)dB_s^{T_{k+1}} - \frac{1}{2}\int_0^t \alpha^2(s)ds\right\}.$$

Remembering the Girsanov theorem and particularly the equations (1.11) and (1.12)

$$dB^{T_k} = -\frac{\delta_k F_k(t)}{1+\delta_k F_k(0)}\sigma_k(t)dt + dB^{T_{k+1}} \tag{2.4}$$

holds under the new measure $\mathbb{Q}_{T_k}$.
A suitable common measure is the so-called terminal measure $\mathbb{Q}_{T_{N+1}}$. By using equation (2.4) for the different forward rates we get

$$\begin{aligned}
\frac{dF_k(t)}{F_k(t)} &= \sigma_k(t)\left(dB^{T_{k+1}} - dB^{T_{N+1}}\right) + \sigma_k(t)dB^{T_{N+1}} \\
&= \sigma_k(t)\sum_{i=k+1}^{N}\left(dB^{T_i} - dB^{T_{i+1}}\right) + \sigma_k(t)dB^{T_{N+1}} \\
&= -\sigma_k(t)\sum_{i=k+1}^{N}\frac{\delta_i F_i(t)}{1+\delta_i F_i(0)}\sigma_i(t)dt + \sigma_k(t)dB^{T_{N+1}}.
\end{aligned}$$

Analogous calculation permits to change into any other forward measure. The Libor market model allows a closed solution for a single caplet. For that purpose we describe the price as a conditional expectation. Therefore let $F_k(t)$ be the current forward rate

$$C(t) = \delta_k B(t,T_{k+1}) E^{F_k(t),\mathbb{Q}_{T_{k+1}}}\left[(F_k(T_k)-K)^+\right]. \tag{2.5}$$

As F_k is lognormal-distributed, we can apply the Black-Scholes formula (see app. B and particularly example B.4). This leads to the caplet price C

$$C(t) = \delta_k B(t,T_{k+1})\left[F_k(t)\Phi\big(d(t)\big) - K\Phi\big(d(t)-V_k(t)\big)\right] \tag{2.6}$$

with the auxiliary functions V_k and d

$$\begin{aligned}
V_k(t) &= \int_t^{T_k}\sigma_k(s)ds, \\
d(t) &= \frac{\log(\frac{F_k(t)}{K}) + \frac{V_k^2(t)}{2}}{V_k(t)}.
\end{aligned}$$

In this case Φ is the cumulative standard normal distribution (B.7).

Additionally the Black-Scholes formula allows to calculate the volatility σ if the market price of the caplet is known. Volatility which is calculated like this is named implicit volatility (see also appendix B).

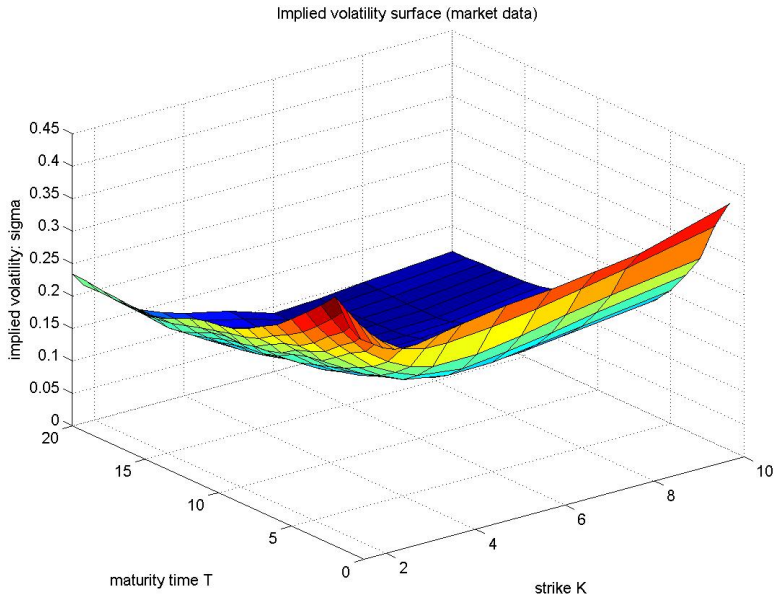


Figure 2.3: Volatility surface, prices of caplets on Euro-forwards, 2004/03/30, provided by Dr.Thilo Roßberg, ABN AMRO London.

Thus it is necessary to extend the caplet model to fit with the observed market prices.

2.2 Extended Libor market model

In this section different approaches to include the market observation will be discussed. For the first time we cannot achieve closed formulas. Thus numerics comes into the game. The model extensions can be classified as follows

1. Displaced Diffusion (DD),
2. Constant Elasticity of Variance (CEV),
3. Stochastic Volatility.

The displaced diffusion and the constant elasticity of variance are two possibilities to face the problem of non-lognormal distributed forward rates. The stochastic volatility can be considered as an additional idea. It does not take too much effort to combine the different extensions. Considering the method

of stochastic volatility, a stochastic process V_t is added to the diffusion term. By this principle it becomes possible to adapt the implied volatility surface to the market data.

2.2.1 Displaced Diffusion

Definition 2.4 (Displaced diffusion (DD)) The forward rates under the forward measure $\mathbb{Q}_{T_{k+1}}$ are given by

$$dF_k(t) = \sigma_{k,DD}(t)\,(F_k(t) + m_k)\,dB^{T_{k+1}} \text{ with } t < T_k. \tag{2.7}$$

Using Girsanov's theorem one can change the measure $\mathbb{Q}_{T_{k+1}}$ to $\mathbb{Q}_{T_k}$ by

$$\begin{aligned} Z(t) &= \frac{B(t,T_k)/B(0,T_k)}{B(t,T_{k+1})/B(0,T_{k+1})} = \frac{1+\delta_k F_k(t)}{1+\delta_k F_k(0)} \\ dZ(t) &= \frac{\delta_k(F_k(t)+m_k)}{1+\delta_k F_k(0)}\sigma_{k,DD}(t)dB^{T_{k+1}}. \end{aligned}$$

This leads to the next equation under the new measure $\mathbb{Q}_{T_k}$:

$$dB^{T_k} = -\frac{\delta_k(F_k(t)+m_k)}{1+\delta_k F_k(0)}\sigma_{k,DD}(t)dt + dB^{T_{k+1}}.$$

It is reasonable to choose the terminal measure as a common measure for all forward rates:

$$\begin{aligned} \frac{dF_k(t)}{F_k(t)+m_k} &= \sigma_{k,DD}(t)\left(dB^{T_{k+1}} - dB^{T_{N+1}}\right) + \sigma_{k,DD}(t)dB^{T_{N+1}} \\ &= \sigma_{k,DD}(t)\sum_{i=k+1}^{N}\left(dB^{T_i} - dB^{T_{i+1}}\right) + \sigma_{k,DD}(t)dB^{T_{N+1}} \\ &= -\sigma_{k,DD}(t)\sum_{i=k+1}^{N}\frac{\delta_i(F_i(t)+m_i)}{1+\delta_i F_i(0)}\sigma_{i,DD}(t)dt + \sigma_{k,DD}(t)dB^{T_{N+1}}. \end{aligned}$$

To calculate the price of a caplet we ascertain that $\hat{F}_k(t) = F_k(t) + m_k$ is lognormal distributed. With the notation $\hat{K} = K + m_k$ and the two auxiliary functions V_k and d

$$\begin{aligned} V_k(t) &= \int_t^{T_k}\sigma_k(s)ds, \\ d(t) &= \frac{\log(\frac{\hat{F}_k(t)}{\hat{K}}) + \frac{V_k^2(t)}{2}}{V_k(t)} \end{aligned}$$

we can apply the Black-Scholes formula:

$$C(t) = \delta_k B(t, T_{k+1}) \left[\hat{F}_k(t) \Phi\big(d(t)\big) - \hat{K} \Phi\big(d(t) - V_k(t)\big) \right]. \tag{2.8}$$

Depending on the parameter m_k the model of displaced diffusion provides a skew in the implied volatility surface. This behaviour is exemplified in the figures 2.4 and 2.5.

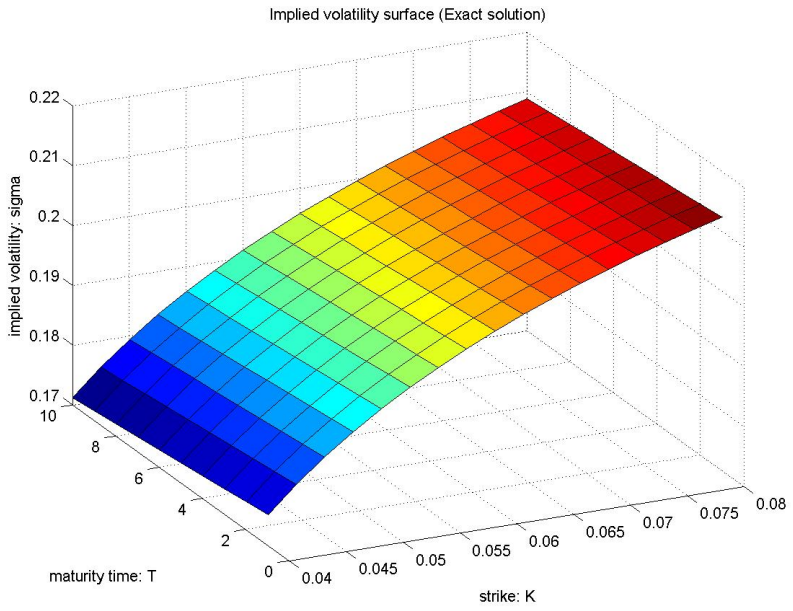


Figure 2.4: Implied volatility surface: $F_k(0) = 0.06$, $\sigma_k = 0.2$, $T_k = 1, 2, ..., 10$, $K = 0.04, ..., 0.08$ and $m_k = -0.02$.

Evidently the parameter $m_k < 0$ causes the caplet prices to increase with the value of the strike K. On the other hand choosing $m_k > 0$ the skew in the surface is in the opposite direction as we can see in figure 2.5.

Investigation into the analytical behaviour of displaced diffusion shows that the forward rates live on the interval $[-m, \infty) \subset \mathbb{R}$. This assertion has already been made in the first chapter (see equation (1.22)). This means that the stochastic process can reach negative values if m is greater zero. Looking ahead to the problem of numerical simulation we have to find an integration scheme which keeps the numerical values in the interval $[-m, \infty)$. But the other extensions of the Libor market model are confronted with similar problem, too.

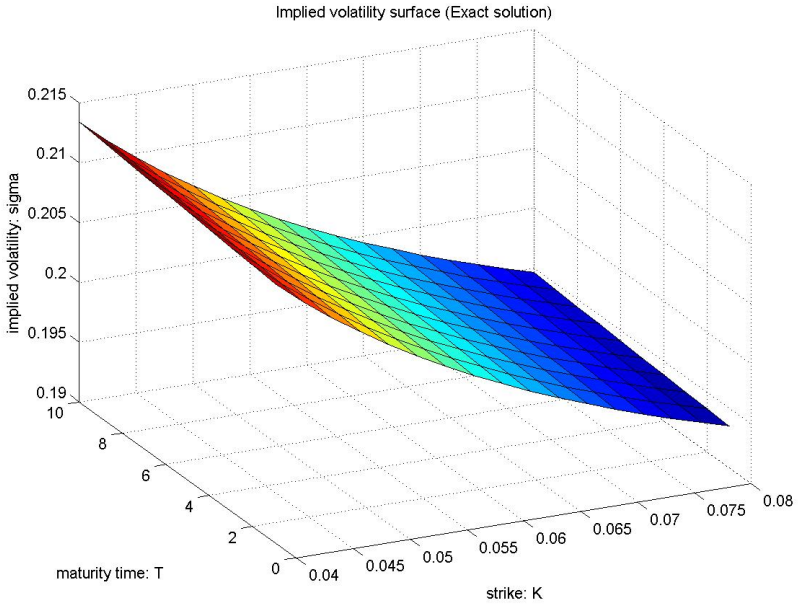


Figure 2.5: Implied volatility surface: $F_k(0) = 0.06$, $\sigma_k = 0.15$, $T_k = 1, 2, ..., 10$, $K = 0.04, ..., 0.08$ and $m_k = 0.02$.

2.2.2 Constant Elasticity of Variance

Definition 2.5 (Constant Elasticity of Variance (CEV)) In this model the forward rates are given under the forward measure $\mathbb{Q}_{T_{k+1}}$ by

$$dF_k(t) = \sigma_{k,CEV}(t)F_k(t)^{\alpha_k}dB^{T_{k+1}} \text{ with } t < T_k. \tag{2.9}$$

Girsanov's theorem defines also in this case the change from the measure $\mathbb{Q}_{T_{k+1}}$ to $\mathbb{Q}_{T_k}$:

$$\begin{aligned} Z(t) &= \frac{B(t,T_k)/B(0,T_k)}{B(t,T_{k+1})/B(0,T_{k+1})} = \frac{1+\delta_k F_k(t)}{1+\delta_k F_k(0)}, \\ dZ(t) &= \frac{\delta_k F_k(t)^{\alpha_k}}{1+\delta_k F_k(0)}\sigma_{k,CEV}(t)dB^{T_{k+1}}. \end{aligned}$$

Under the new measure $\mathbb{Q}_{T_k}$ we obtain:

$$dB^{T_k} = -\frac{\delta_k F_k(t)^{\alpha_k}}{1+\delta_k F_k(0)}\sigma_{k,CEV}(t)dt + dB^{T_{k+1}}.$$

The same argumentation leads to the forward rates under the terminal measure:

$$\begin{aligned}\frac{dF_k(t)}{F_k(t)^{\alpha_k}} &= \sigma_{k,CEV}(t)\left(dB^{T_{k+1}} - dB^{T_{N+1}}\right) + \sigma_{k,CEV}(t)dB^{T_{N+1}} \\ &= \sigma_{k,CEV}(t)\sum_{i=k+1}^{N}\left(dB^{T_i} - dB^{T_{i+1}}\right) + \sigma_{k,CEV}(t)dB^{T_{N+1}} \\ &= -\sigma_{k,CEV}(t)\sum_{i=k+1}^{N}\frac{\delta_i F_i(t)^{\alpha_i}}{1+\delta_i F_i(0)}\sigma_{i,CEV}(t)dt + \sigma_{k,CEV}(t)dB^{T_{N+1}}.\end{aligned}$$

The CEV provides closed formulas for the caplet prices as well even though it is more complicated than in the case of displaced diffusion. We define the parameter d and V_k as usual

$$\begin{aligned}V_k &= \int_t^{T_k}\sigma_k(s)ds, \\ d(t) &= \frac{\log(\frac{F_k(t)}{K}) + \frac{V_k^2(t)}{2}}{V_k(t)}.\end{aligned}$$

Additionally we need three auxiliary variables a, b, c:

$$a = \frac{K^{2(1-\alpha)}}{(1-\alpha)^2 V_k(t)}\ ,\ b = \frac{1}{1-\alpha}\ ,\ c = \frac{F_k(t)^{2(1-\alpha)}}{(1-\alpha)^2 V_k(t)}.$$

This culminates in the following representation of the caplet prices in dependence of α:

$$C_k(t) = \delta_k B(t, T_{k+1})\begin{cases}[F_k(t)(1-\chi^2(a, b+2, c)) - K\chi^2(c, b, a)] & :0<\alpha<1 \\ \left[F_k(t)\Phi\big(d(t)\big) - K\Phi\big(d(t) - V_k(t)\big)\right] & :\alpha = 1 \\ [F_k(t)(1-\chi^2(c, -b, a)) - K\chi^2(a, 2-b, c)] & :\alpha > 1.\end{cases}$$

Remark 2.6 Here $\chi^2(x, y, z)$ is the non-central chi-square distribution with three grades of freedom. A complete proof can be found in [JK81, Sch89].

The parameter α enables us to model a skew into the implied volatility surface and the quantity of α determines the direction as the next two examples illustrate.

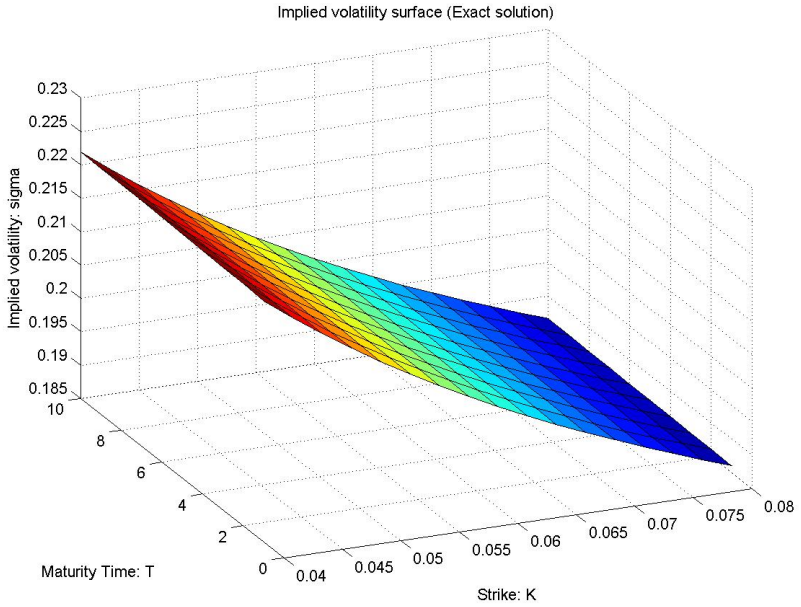


Figure 2.6: Implied volatility surface: $F_k(0) = 0.06$, $\sigma_k = 0.04899$, $T_k = 1, 2, ..., 10$, $K = 0.04, ..., 0.08$ and $\alpha_k = 0.5$.

In this case we see a downward trend with increasing strike. In contrast $\alpha > 1$ causes another direction in the volatility surface, shown in figure 2.7.

From the analytical point of view the CEV model is superior to the DD model because it averts negative forward rates (see equation (1.23)). This advantage is reduced by the fact that 0 is an attainable bound if $\alpha < 1$. Again the numerical simulation has to take this fact into account. In the end of this chapter a short introduction to stochastic volatility will be given. The results presented below mainly refer to the article "Extended Libor market model with Stochastic volatility" by Andersen and Brotherton-Ratcliffe [ABR01].

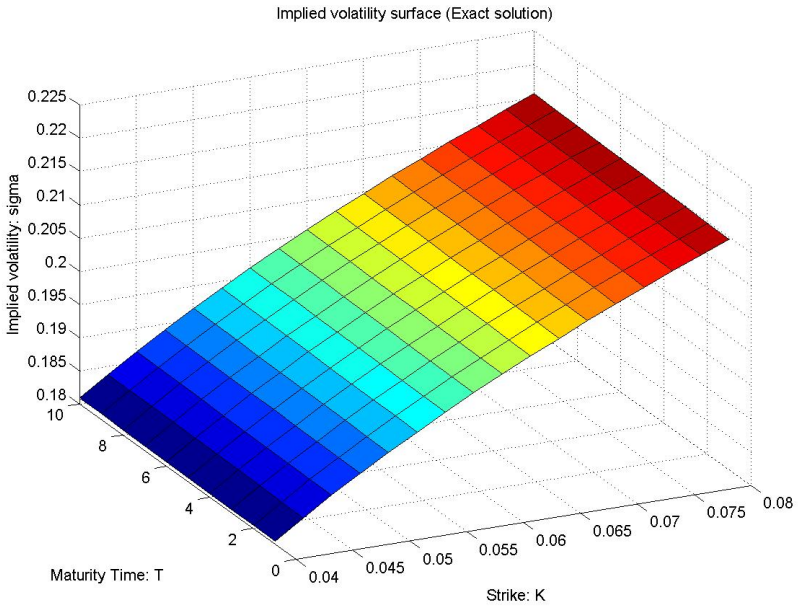


Figure 2.7: Implied volatility surface: $F_k(0) = 0.06$, $\sigma_k = 0.8161$, $T_k = 1, 2, ..., 10$, $K = 0.04, ..., 0.08$ and $\alpha_k = 1.5$

.

2.2.3 Stochastic Volatility

Definition 2.7 (Stochastic volatility) The Libor market model with stochastic volatility defines the forward rates as follows

$$dF_k(t) = \sigma_k(t)\sqrt{V(t)}\varphi(F_k(t))dB^{T_{k+1}} \text{ with } t < T_k. \tag{2.10}$$

The function φ can be given either by the displaced diffusion or as constant elasticity of variance. The volatility V is described by the stochastic differential equation

$$dV(t) = \kappa(\theta - V(t))dt + \epsilon V(t)^{\beta}dZ. \tag{2.11}$$

Here Z is a Brownian motion and $\beta \in (0.5, 1]$.

Using stochastic volatility makes closed pricing formulas impossible. The best we can get is an appropriate approximation. A detailed description will not be given in this thesis but an intensive study of this topic can be found in [ABR01]. Indeed the equation for stochastic volatility is interesting for another reason too. Yet in section 1.2 it became clear that a mean-reverting process only takes positive values under certain conditions. Because

of the absence of closed formulas numerical simulation becomes essential. The maintaining of numerical positivity is an important requirement to get exact approximations. In chapter 4 we will see that it is possible to find a numerical integration scheme providing this property. Numerical simulation of caplet prices is a focal point in chapter 5. The next picture gives a first impression of the impact of stochastic volatility on the implied volatility surface.

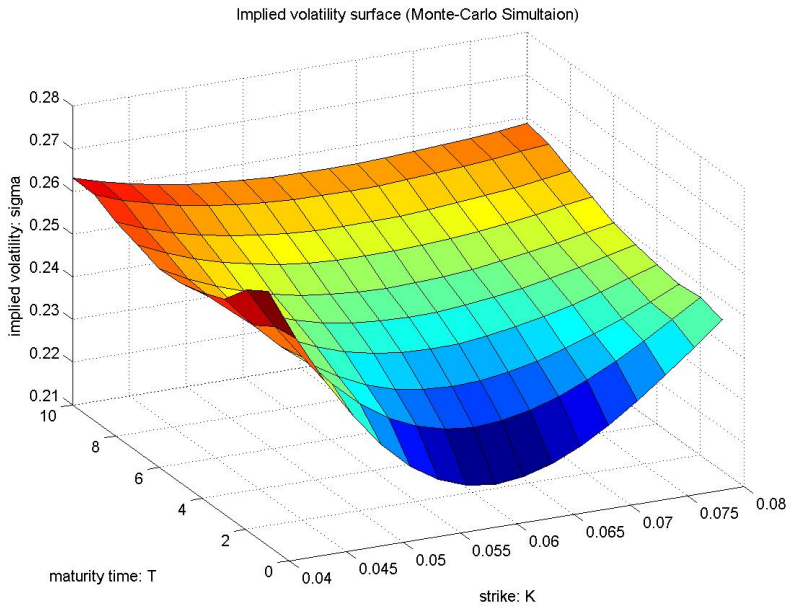


Figure 2.8: Implied volatility surface: $F_k(0) = 0.06$, $\sigma_k = 0.2$, $T_k = 1, 2, ..., 10$, $K = 0.04, ..., 0.08$, $V(0) = \kappa = \theta = 1$, $\epsilon = 1.4$, $\beta = 0.75$, $\Delta_t = 0.25$, Monte Carlo simulation with 5000000 paths.

CHAPTER III

Numerical integration of stochastic differential equations

In general there is no closed solution of a stochastic differential equation. Consequently the Black-Scholes formula B.4 is rather an exception as the underlying process is described by a geometric brownian motion.
Thus, having a more complex stochastic differential equation, it is necessary to use numerical integration schemes to approximate the solution. We know that a stochastic process possesses a density p at each time t. To get a numerical approximation for this density a Monte-Carlo simulation can be done. Schematically one must consider two aspects

1. Discretisation (of the SDE),
2. Simulation (of N paths).

First it is necessary to discretise the stochastic differential equation. This point is strongly connected to stochastic Taylor expansion. Secondly a finite number of paths must be simulated to get an approximation of the density.

But both points contain sources of error so that we have to be careful to make sure that the approximated solution agrees with the analytical one. It is quite clear that the number of paths directly corresponds with the accuracy and the variance of the approach: *"Increase the number of paths to decrease the oscillation of the error."* Hence it is necessary to simulate a suitable number of paths to reach a sensible numerical solution. In the following we assume this point as settled and we will pay attention to the other source of error: the discretisation of the differential equation.

To deal with this problem it is necessary to develop stochastic Taylor expansion as the main tool to study discretisation error. For that reason the usual concept of multiindices is introduced.

3.1 Stochastic Taylor expansion

Starting-point is an arbitrary Ito process X_t

$$dX = a(t,X)dt + b(t,X)dW \text{ with } X_{t_0} = x_0. \tag{3.1}$$

To motivate the stochastic Taylor expansion we start with the easiest example. For this purpose let X_t be an Ito diffusion with its integral representation

$$X_t = X_0 + \int_0^t a(X_s)ds + \int_0^t b(X_s)dW_s.$$

Applying Ito's formula to a function $f \in C_0^2(\mathbb{R})$ leads to

$$\begin{aligned} f(X_t) = f(X_0) \quad &+ \quad \int_0^t \left(a(X_s)\frac{\partial}{\partial x}f(X_s) + \frac{1}{2}b(X_s)^2\frac{\partial^2}{\partial x^2}f(X_s) \right) ds \\ &+ \quad \int_0^t b(X_s)\frac{\partial}{\partial x}f(X_s)dW_s. \end{aligned}$$

To simplify the notation we introduce the following operators

$$L^0 = a\frac{\partial}{\partial x} + \frac{1}{2}b^2\frac{\partial^2}{\partial x^2} \text{ and } L^1 = b\frac{\partial}{\partial x}.$$

Choosing $a = f$ and $b = f$

$$\begin{aligned} X_t \quad &= X_0 \quad + \int_0^t \left(a(X_0) + \int_0^s L^0 a(X_u)du + \int_0^s L^1 a(X_u)dW_u \right) ds \\ &\qquad + \int_0^t \left(b(X_0) + \int_0^s L^0 b(X_u)du + \int_0^s L^1 b(X_u)dW_u \right) dW_s \\ &= X_0 \quad + a(X_0)\int_0^t ds + b(X_0)\int_0^t dW_s + R. \end{aligned}$$

The remainder R possesses only terms of higher order

$$\begin{aligned} R &= \int_0^t \int_0^s L^0 a(X_u) du ds + \int_0^t \int_0^s L^1 a(X_u) dW_u ds \\ &+ \int_0^t \int_0^s L^0 b(X_u) du dW_s + \int_0^t \int_0^s L^1 b(X_u) dW_u dW_s. \end{aligned}$$

To point out the connection of the remainder and the stochastic Taylor expansion it is necessary to declare some further notations.

Definition 3.1 (Multiindices I) A multiindex is a vector α with entries $j_i \in \{0, 1, ..., d\}$

$$\alpha = (j_1, ..., j_m). \tag{3.2}$$

The length of a multiindex is defined by

$$\ell(\alpha) = \ell((j_1, ..., j_m)) = m. \tag{3.3}$$

Additionally we write $\alpha = \{\nu\}$ for the vector of length 0. Finally we notate the set of all multiindices with

$$\mathcal{M} := \{(j_1, ..., j_m) \text{ for all } m = 0, 1, ...\}. \tag{3.4}$$

Example 3.2 Let $\alpha = (0, 0, 1, 2, 4)$ then $\ell(\alpha) = 5$. Further let $-\alpha$ be the multiindex obtained by deleting the first component of α. Thus $-(0, 0, 1, 2, 4) = (0, 1, 2, 4)$ and by the same token $(0, 0, 1, 2, 4)- = (0, 0, 1, 2)$. The concatenation of multiindices is notated by $*$. Therefore $(0, 1) * (2, 3) = (0, 1, 2, 3)$.

With this notation it is possible to describe the integration in the Taylor expansion in a particularly simple way.

Definition 3.3 The multiple Ito integral is recursively defined by

$$I_\alpha[f(\cdot)]_{\rho,\tau} = \begin{cases} f(\tau) & : \ell = 0 \\ \int_\rho^\tau I_{\alpha-} f[(\cdot)]_{\rho,\tau} ds & : \ell \geq 1 \text{ and } j_\ell = 0 \\ \int_\rho^\tau I_{\alpha-} f[(\cdot)]_{\rho,\tau} dW_s^{j_\ell} & : \ell \geq 1 \text{ and } j_\ell \geq 1 \end{cases} \tag{3.5}$$

with $\alpha = (j_1, ..., j_\ell)$.

Example 3.4 Let W^1, W^2 and W^3 be independent Brownian motions

$$\begin{aligned} I_{(0,0)}[g(\cdot)] &= \iint g(s_1)ds_1ds_2, \\ I_{(1,1,0)}[g(\cdot)] &= \iiint g(s_1)dW_{s_1}dW_{s_2}ds_3, \\ I_{(0,1,2,3)}[g(\cdot)] &= \iiiint g(s_1)ds_1dW^1_{s_2}dW^2_{s_3}dW^3_{s_4}. \end{aligned}$$

Summing up these facts we can write the first step in the stochastic Taylor expansion as follows

$$X_t = X_0 + a(X_0)I_{(0)}[1]_{0,t} + b(X_0)I_{(1)}[1]_{0,t} + R. \tag{3.6}$$

The generalisation of this concept needs the introduction of more notations.

Definition 3.5 A nonempty set $\mathcal{A} \subset \mathcal{M}$ is called hierarchical if:

$$\sum_{\alpha\in\mathcal{A}} \ell(\alpha) < \infty, \tag{3.7}$$

$$-\alpha \in \mathcal{A} \text{ for all } \alpha \in \mathcal{A} \setminus \{\nu\}. \tag{3.8}$$

The remainder set of $\mathcal{A}$ is given by

$$B(\mathcal{A}) = \{\alpha \in \mathcal{M} \setminus \mathcal{A} : -\alpha \in \mathcal{A}\}. \tag{3.9}$$

Example 3.6 The next two sets are hierarchical as one can easily verify

1. $\mathcal{A} = \{\{\nu\}, (0), (1)\}$,
2. $A = \{\{\nu\}, (0), (1), (11)\}$

with the corresponding remainder sets

1. $B(\mathcal{A}) = \{(00), (10), (01), (11)\}$,
2. $B(\mathcal{A}) = \{(00), (10), (01), (11), (011), (111)\}$.

Looking at the first example we ascertain that the indices agree with the first step of the stochastic Taylor expansion. This principle can be generalised. To do so it is necessary to analyse the relation between different Ito integrals. For that purpose let

$$\begin{aligned} I_{\alpha,t} &= I_\alpha[1]_{0,t}, \\ I_{\alpha,\Delta} &= I_\alpha[1]_{\rho,\tau} \text{ with } \Delta = \tau - \rho. \end{aligned}$$

The next proposition can be found in [KP92].

Proposition 3.7 Let $j_1, ..., j_m \in \{1, ..., d\}$ and $\alpha = (j_1, ..., j_m) \in \mathcal{M}$ where $m \geq 1$. Then the next equation holds:

$$I_{(j),t} I_{\alpha,t} = \sum_{i=0}^{m} I_{(j_1,...,j_i,j,j_{i+1},...,j_m),t} + \sum_{\substack{i=0 \\ j_i = j \neq 0}}^{m} I_{(j_1,...,j_{i-1},0,j_{i+1},...,j_m),t}.$$

Remark 3.8 This result will be needed to prove the convergence of balanced stochastic integration schemes. Before presenting some more examples we will analyse a special case of this proposition.

Corollary 3.9 Let $\alpha = (j_1, ..., j_m)$ with $j_1 = j_2 = ... = j_m = j$ where $m > 2$. Then

$$I_{\alpha,t} = \begin{cases} \frac{1}{m!} t^m & : j = 0, \\ \frac{1}{m} \left(I_{(j),t} I_{\alpha-,t} - I_{(0),t} I_{(\alpha-)-,t} \right) & : j \neq 0. \end{cases} \tag{3.10}$$

Proof: The case $j = 0$ is obvious and the Proposition 3.7 shows that

$$\begin{aligned} I_{(0),t} I_{(\alpha-)-,t} &= \sum_{i=0}^{m-2} I_{(j_1,...,j_i,0,j_{i+1},...,j_{m-2}),t}, \\ I_{(j),t} I_{\alpha-,t} &= m I_{\alpha,t} + \sum_{i=1}^{m-1} I_{(j_1,...j_{i-1},0,j_{i+1},...,j_{m-1}),t}. \end{aligned}$$

Renumbering and differentiating proves the statement. □

Example 3.10 The following examples are selected especially with respect to further requirements:

$$\begin{aligned} I_{(11),t} &= \frac{1}{2} \left(I_{(1),t}^2 - I_{(0),t} \right), \\ I_{(1),t} I_{(01),t} &= I_{(101),t} + 2 I_{(011),t} + I_{(00),t}, \\ I_{(1),t} I_{(11),t} &= 3 I_{(111),t} + I_{(10),t} + I_{(01),t}, \\ I_{(11),t} I_{(11),t} &= \frac{1}{2} \left(I_{(1),t}^2 - I_{(0),t} \right) I_{(11),t}, \\ &= 6 I_{(1111),t} + 2 \left(I_{(011),t} + I_{(101),t} + I_{(110),t} \right) + I_{(00),t}. \end{aligned}$$

There is only one last tool missing to represent the stochastic Taylor expansion. Already at the beginning of this chapter we introduced the operators L^0 and L^1.

This principle can be transferred to the multidimensional case

$$L^0 = \frac{\partial}{\partial t} + \sum_{k=1}^{d} a_k \frac{\partial}{\partial x_k} + \frac{1}{2} \sum_{k,\ell=1}^{d} \sum_{j=1}^{m} b_{k,j} b_{\ell,j} \frac{\partial^2}{\partial x_k \partial x_\ell}.$$

For $j \neq 0$ we get the following operator

$$L^j = \sum_{k=1}^{d} b_{k,j} \frac{\partial}{\partial x_k}.$$

Definition 3.11 The so-called Ito coefficient functions are recursively defined by with $\alpha = (j_1, ..., j_m)$

$$f_\alpha = \begin{cases} f & \ell = 0, \\ L^{j_1} f_{-\alpha} & \ell \geq 1. \end{cases} \tag{3.11}$$

Example 3.12 Later on we will examine numerical integration schemes and therefore we will always choose $f(t, x) = x$ in the stochastic Taylor expansion:

$$f_{(0)} = a \ , \ f_{(1)} = b \text{ and } f_{(11)} = bb'.$$

Applying all these notations we are able to set up the general stochastic Taylor expansion.

Theorem 3.13 (Stochastic Taylor expansion) Let X_t be a stochastic process given by the equation (3.1) and let $\mathcal{A}$ be a hierarchical set. Then the following statement holds

$$f(\tau, X_\tau) = \sum_{\alpha \in \mathcal{A}} I_\alpha[f_\alpha(\rho, X_\rho)]_{\rho,\tau} + \sum_{\alpha \in B(\mathcal{A})} I_\alpha[f_\alpha(\cdot, X_\cdot)]_{\rho,\tau}. \tag{3.12}$$

Proof: See [KP92]. □

In the next section this theorem is particularly interesting for the construction of suitable integration schemes with an appropriate error of the remainder. Before continuing some more definitions are necessary.

Definition 3.14 (Multiindices II) Let α be a given multiindex. Then α^+ is defined by leaving out all zero components. Further let $k_i(\alpha)$ be the distance between the i-th and $(i+1)$-th nonzero entries.

Example 3.15 The next two examples illustrate this

1. $\alpha = (0,1,2,0) \Longrightarrow \alpha^+ = (1,2)$ and $k_0 = 1, k_1 = 0, k_2 = 1$
2. $\alpha = (2,3) \Longrightarrow \alpha^+ = (2,3)$ and $k_0 = 0, k_1 = 0, k_2 = 0$.

Especially the moments of Ito integrals will play essential roles later on. Therefore we define

$$w(\alpha,\beta) = \ell(\alpha^+) + \sum_{i=0}^{\ell(\alpha^+)} \left(k_i(\alpha) + k_i(\beta)\right). \tag{3.13}$$

This section will be completed by presenting two main results of the approximation behaviour of Ito integrals.

Theorem 3.16 Let C_1 be a suitable constant. The expectation of an Ito integral is

$$E\left[I_{\alpha,\Delta}|\mathcal{F}_\rho\right] \leq \begin{cases} 0 & : \text{if } \ell(\alpha) \neq \ell(\alpha^+) \\ C_1\Delta^{\ell(\alpha)} & : \text{else.} \end{cases} \tag{3.14}$$

Theorem 3.17 In the latter case $(\cdot,\cdot)$ is the usual Euclidean scalar product. Then we get with an appropriate constant C_2:

$$E\left[(I_{\alpha,\Delta}, I_{\beta,\Delta})|\mathcal{F}_\rho\right] \leq \begin{cases} 0 & : \alpha^+ \neq \beta^+ \\ C_2\Delta^{w(\alpha,\beta)} & : \text{else.} \end{cases} \tag{3.15}$$

Proof: See again [KP92] especially for a detailed description of C_1 and C_2. □

According to this theoretical introduction we now have the necessary tools to develop stochastic integration schemes. Particularly in cooperation with the theorem of Milstein the characterisation of stochastic Taylor expansion by multiindices is a useful tool. Now it is possible to prove the convergence of an integration scheme with only little effort.

3.2 Basic numerical integration schemes

The aim of this section is to develop numerical integration schemes. First of all we will present the explicit Euler and Milstein methods which can be directly deduced from the stochastic Taylor expansion. But we will also present implicit methods especially with regard to the problem of numerical

positivity. The implicit schemes can be generalised further: These are the balanced methods.

The most familiar stochastic integration scheme is the Euler scheme.

Definition 3.18 (Euler method) The stochastic Euler method for solving equation (3.1) is defined by:

$$X_{n+1} = X_n + a(t_n, X_n)\Delta_{t_n} + b(t_n, X_n)\Delta W. \tag{3.16}$$

Thereby $\Delta_{t_n} := t_{n+1} - t_n$ is called the stepsize and ΔW is a $\mathcal{N}(0, \Delta_{t_n})$ distributed random variable.

Hence the Euler scheme corresponds with the hierarchical set

$$\mathcal{A} = \{(0), (1)\}. \tag{3.17}$$

Considering the remainder set $B(\mathcal{A})$ one can observe that for all $\alpha, \beta \in B(\mathcal{A})$ and for suitable constants C_1, C_2 the next inequalities hold:

$$\begin{aligned} E\left[I_{\alpha,\Delta_{t_n}}|\mathcal{F}_{t_n}\right] &\leq C_1\Delta_{t_n}^{p_1} \text{ with } p_1 \geq 2, \\ \left(E\left[(I_{\alpha,\Delta_{t_n}}, I_{\beta,\Delta_{t_n}})|\mathcal{F}_{t_n}\right]\right)^{1/2} &\leq C_2\Delta_{t_n}^{p_2} \text{ with } p_2 \geq 1. \end{aligned}$$

This leads to the question if there is in general a connection between the values of p_1 and p_2. To give an answer it is necessary to define what the convergence of a stochastic integration scheme is.

Definition 3.19 (Strong convergence) A numerical integration scheme X_n converges strongly with order γ_1 towards the exact solution Y_{T_N} if

$$\lim_{\Delta_t \to 0} E\left[\left|Y_{T_N} - X_N\right||\mathcal{F}_0\right] \leq C\Delta_t^{\gamma_1}. \tag{3.18}$$

In this inequality $\Delta_t := \max_{i=1,\ldots N-1}\{\Delta_{t_i}\}$ and T_N is an arbitrary value in $[0, T]$ and independent of the constant C.

Definition 3.20 (Weak convergence) Again X_n is a numerical integration scheme and Y_{T_N} the exact solution. Then the scheme has the weak order of convergence γ_2 if

$$\lim_{\Delta_t \to 0}\left|E\left[Y_{T_N}|\mathcal{F}_0\right] - E\left[X_N|\mathcal{F}_0\right]\right| \leq C\Delta_t^{\gamma_2} \tag{3.19}$$

with Δ_t, T_N and C as above.

But how can we prove the convergence order of a stochastic integration scheme? As in the deterministic case the Taylor expansion is the essential key. A rather simple method is using hierarchical sets $\mathcal{A}$. These hierarchical sets enable us to make some error estimations. Doing this the Theorems 3.16 and 3.17 turn out to be crucial to calculate the local convergence order.

In the deterministic case we call this behaviour the consistence of the numerical scheme. But in the case of stochastic processes consistence is not sufficient to guarantee convergence, either. Here the first and second moments of the local error must converge. The Milstein theorem illustrates this interplay between local and global properties. Before we set out to explain this properly we have to say a few words about local convergence.

Definition 3.21 The local mean error is the error of the one-step discretisation

$$\left| E\left[Y_{\tau_{n+1}}^{\tau_n, Y_n} - X_{n+1} | \mathcal{F}_{\tau_n} \right] \right| \leq C \Delta_{t_n}^p \tag{3.20}$$

where $Y_{\tau_{n+1}}^{\tau_n, X_n}$ is the value of the exact solution at time τ_{n+1} starting at τ_n with the value X_n. On the other hand X_{n+1} is the $(n+1)$-th step of the integration scheme. It is necessary to take the conditional expectation to be independent of the behaviour of the path before.

Theorem 3.22 (Milstein) Let X_n be a numerical integration scheme for the solution of the stochastic differential equation (3.1). The value of the exact solution at τ_n is Y_{τ_n}. Assume that the first two moments of the one-step discretisation for all $N = 1, 2..$ and all $n = 0, ..., N-1$ fulfil the inequalities

$$\left| E\left[Y_{\tau_{n+1}}^{\tau_n, X_n} - X_{n+1} | \mathcal{F}_{\tau_n} \right] \right| \leq \left(1 + |Y_n|^2\right)^{1/2} \Delta_{t_n}^{p_1} \tag{3.21}$$

and

$$\left(E\left[|Y_{\tau_{n+1}}^{\tau_n, X_n} - X_{n+1}|^2 | \mathcal{F}_{\tau_n} \right] \right)^{1/2} \leq \left(1 + |Y_n|^2\right)^{1/2} \Delta_{t_n}^{p_2} \tag{3.22}$$

with $p_2 \geq \frac{1}{2}$ and $p_1 \geq p_2 + \frac{1}{2}$. Then the global error satisfies for all $k = 0, ..., N$

$$\left(E\left[|Y_{\tau_k}^{0, X_0} - X_k|^2 | \mathcal{F}_0 \right] \right)^{1/2} \leq \left(1 + |X_0|^2\right)^{1/2} \Delta_t^{p_2 - \frac{1}{2}}. \tag{3.23}$$

So the integration scheme has a strong convergence order of $\gamma_1 = p_2 - \frac{1}{2}$.

Proof: One can find the proof in [Mil95]. □

In the last section we already discussed how it is possible to describe the stochastic Taylor expansion with hierarchical sets. In particular we get estimations for the remainder R. Therewith it is possible to construct a stochastic integration scheme with an arbitrary strong order of convergence γ. Hence it is only necessary to find the appropriate hierarchical sets which provide a proper local convergence error.

Definition 3.23 (Strong Ito-Taylor multiindices) Let the following hierarchical sets be defined for $\gamma = 0.5, 1, 1.5, ...$

$$\mathcal{A}_\gamma = \{\alpha \in \mathcal{M} : \ell(\alpha) + n(\alpha) \leq 2\gamma \text{ or } \ell(\alpha) = n(\alpha) = \gamma + 0.5\}. \tag{3.24}$$

Example 3.24 The first hierarchical sets are presented below to give a short illustration

$$\begin{aligned}
\mathcal{A}_{0.5} &= \{\nu, (0), (1)\}, \\
\mathcal{A}_{1.0} &= \{\nu, (0), (1), (1,1)\}, \\
\mathcal{A}_{1.5} &= \{\nu, (0), (1), (1,1), (0,1), (1,0), (0,0), (1,1,1)\}.
\end{aligned}$$

Lemma 3.25 The remainder set for $\mathcal{A}_\gamma$ is:

$$B(\mathcal{A}_\gamma) = \{\alpha \in \mathcal{M} : n(\alpha) + \ell(\alpha) \in [2\nu + 1, 2\nu + 2] \text{ or } n(\alpha) = \ell(\alpha) = \nu + 1.5\}.$$

Proof: Let $\beta = -\alpha \in \mathcal{A}$. Then we have to check that $\alpha = (j) * \beta \in B(\mathcal{A})$. In the most simple case $n(\beta) \neq \ell(\beta)$, we can directly conclude that

$$\begin{aligned}
n(\alpha) + \ell(\alpha) &= n(\beta) + \ell(\beta) + 2 \leq 2\nu + 2 \text{ for } j = (0), \\
n(\alpha) + \ell(\alpha) &= n(\beta) + \ell(\beta) + 1 \leq 2\nu + 1 \text{ for } j \neq (0).
\end{aligned}$$

In the case that $n(\beta) = \ell(\beta)$ we have to differentiate further. If $n(\beta) = \ell(\beta) = \nu + 0.5$ we get:

$$\begin{aligned}
n(\alpha) = \ell(\alpha) &= n(\beta) + 1 = \nu + 1.5 \text{ for } j = (0), \\
n(\alpha) + \ell(\alpha) &= n(\beta) + \ell(\beta) + 1 \leq 2\nu + 2 \text{ for } j \neq (0),
\end{aligned}$$

otherwise $n(\beta) = \ell(\beta) \leq \nu$. Therefore the result is:

$$\begin{aligned}
n(\alpha) = \ell(\alpha) &= n(\beta) + 1 = \nu + 1 \leq 2\nu + 2 \text{ for } j = (0), \\
n(\alpha) + \ell(\alpha) &= n(\beta) + \ell(\beta) + 1 \leq 2\nu + 1 \text{ for } j \neq (0).
\end{aligned}$$

We notice that the proof of this lemma needs only basic calculation but it allows to verify the next theorem in a simple way. □

Theorem 3.26 The numerical integration scheme

$$X_{n+1} = \sum_{\alpha \in \mathcal{A}_\gamma} f_\alpha(\tau_n, X_n) I_\alpha \tag{3.25}$$

possesses a strong order of convergence γ.

Proof: Using the Milstein theorem we need only to calculate the two constants p_1 and p_2. Therefore it is indispensable that $p_2 \geq \gamma + \frac{1}{2}$ and $p_1 \geq \gamma + 1$. Let us begin with a general observation:

$$\begin{aligned} X_{\tau_{n+1}}^{\tau_n, X_n} - X_{n+1} &= X_{\tau_{n+1}}^{\tau_n, X_n} - \sum_{\alpha \in \mathcal{A}_\gamma} f_\alpha(\tau_n, X_n) I_\alpha \\ &= \sum_{\alpha \in B(\mathcal{A}_\gamma)} f_\alpha(\tau_n, X_n) I_\alpha \end{aligned}$$

where $X_{\tau_{n+1}}^{\tau_n, X_n}$ denotes the exact solution. With this calculation we can use the lemma above to verify the inequalities for p_1 and p_2. Firstly take a closer look at p_1. Let $\alpha \in B(\mathcal{A})$ be arbitrary. If furthermore $n(\alpha) \neq \ell(\alpha)$ it follows from Theorem (3.14) that

$$E\left[I_{\alpha, \Delta_{t_n}} | \mathcal{F}_\rho\right] = 0. \tag{3.26}$$

Otherwise if $n(\alpha) = \ell(\alpha)$ then

$$E\left[I_{\alpha, \Delta_{t_n}} | \mathcal{F}_\rho\right] \leq C \Delta_{t_n}^{\ell(\alpha)}. \tag{3.27}$$

From this it follows that $p_1 \geq \gamma + 1$. The derivation of p_2 is not complicated, either. Considering that

$$E\left[(I_{\alpha, \Delta_{t_n}}, I_{\beta, \Delta_{t_n}}) | \mathcal{F}_\rho\right] \leq C_2 \Delta_{t_n}^{w(\alpha, \beta)},$$

we can assume that $\alpha = \beta$ and therefore we obtain

$$w(\alpha, \alpha) = \ell(\alpha) - n(\alpha) + 2n(\alpha) = \ell(\alpha) + n(\alpha) \geq 2\gamma + 1.$$

Thus $p_2 \geq \gamma + 0.5$. □

Basic schemes

If we remember the Euler method we recognize that the integration scheme agree with the hierarchical set $\mathcal{A}_{0.5}$. The next higher stochastic integration scheme is called Milstein method.

Definition 3.27 (Milstein method) The stochastic integration scheme which belongs to the hierarchical set $\mathcal{A}_{1.0}$ is the Milstein method

$$\begin{aligned} X_{n+1} = X_n &+ a(t_n, X_n)\Delta_{t_n} + b(t_n, X_n)\Delta W_n \\ &+ \frac{1}{2} b(t_n, X_n)\frac{\partial b}{\partial x}(t_n, X_n)\left((\Delta W_n)^2 - \Delta_{t_n}\right). \end{aligned}$$

The Milstein method will be of further interest later particularly with regard to preserving positivity.
Up to now we have only developed explicit integration schemes. As in the deterministic case there is often the problem that these schemes cannot avert instabilities caused by stiffness (see [HW87]).
So it is clear that the same kind of problems also appears in the stochastic case. Additionally, we have to handle the problem of numerical positivity. Therefore we need to analyse and develop some more integration schemes.

3.3 Advanced schemes

3.3.1 Drift implicit methods

Drift implicit methods can be directly deduced from their explicit counterpart up to a strong convergence order of $\gamma_1 = 2$. But it will be only necessary to use schemes with an order less than $\gamma_1 = 1$.

Definition 3.28 (Implicit Euler) The drift implicit Euler method is defined by the following integration scheme

$$X_{n+1} = X_n + a(t_{n+1}, X_{n+1})\Delta_{t_n} + b(t_n, X_n)\Delta W_n.$$

The only difference to the explicit one is the exchange of the drift. Taking the next integration step into the drift confronts with the problem of solving nonlinear equation systems. This computation is usually quite expensive. Later on we will see that in the case of a mean-reverting process the implicit Milstein integration scheme seems to be the method of choice.

Definition 3.29 (Implicit Milstein) The drift implicit Milstein method is defined by

$$\begin{aligned} X_{n+1} = X_n &+ a(t_{n+1}, X_{n+1})\Delta_{t_n} + b(t_n, X_n)\Delta W_n \\ &+ \frac{1}{2} b(t_n, X_n)\frac{\partial}{\partial x} b(t_n, X_n)\left((\Delta W_n)^2 - \Delta_{t_n}\right). \end{aligned}$$

It is possible to generalise implicit methods furthermore. This leads to the so-called balanced methods.

3.3.2 Balanced methods

The difference is that balanced methods allow to choose functions to control the behaviour of the integration. This way we can provide positivity. Yet, in most cases we have to accept a decreased convergence speed and higher computational cost.

To simplify the notation we consider only Ito diffusions

$$dX_t = a(X_t)dt + b(X_t)dW$$

in the following but the balanced methods can be obviously applied to Ito processes as well.

Definition 3.30 (Balanced implicit method (BIM)) The integration scheme for the BIM is given as follows

$$\begin{aligned} X_{n+1} &= X_n + a(X_n)\Delta_{t_n} + b(X_n)\Delta W + (X_n - X_{n+1})C_n(X_n) \\ C_n(X_n) &= c_0(X_n)\Delta + c_1(X_n)|\Delta W|. \end{aligned}$$

In this method the functions c_0 and c_1 are called control functions. The control functions must be bounded and have to satisfy the inequality

$$1 + c_0(X_n)\Delta_{t_n} + c_1(X_n)|\Delta W| > 0. \tag{3.28}$$

The BIM is based on the Euler method and has been intensively analysed in [MPS98, Sch96, Sch97]. We can make use of the idea of the BIM and combine it with the Milstein method.

Definition 3.31 (Balanced Milstein method (BMM)) The BMM is given by the following integration scheme

$$\begin{aligned} X_{n+1} = X_n &+ a(X_n)\Delta_{t_n} + b(X_n)\Delta W + \frac{1}{2}b(X_n)b'(X_n)\left((\Delta W_n)^2 - \Delta_{t_n}\right) \\ &+ (X_n - X_{n+1})C_n(X_n). \end{aligned}$$

C_n consists again of control functions but here we need different terms to guarantee convergence:

$$C_n(X_n) = c_0(X_n)\Delta_{t_n} + c_2(X_n)\left((\Delta W_n)^2 - \Delta_{t_n}\right).$$

We need further the restriction

$$1 + c_0(X_n)\Delta_{t_n} + c_2(X_n)\left((\Delta W_n)^2 - \Delta_{t_n}\right) > 0 \tag{3.29}$$

on the control functions c_0 and c_2.

To substantiate the next propositions we have to remember the Milstein theorem and especially the local convergence orders p_1 and p_2.

Theorem 3.32 The BIM method has the following properties

1. local convergence order $p_1 = 1.5$,
2. local square convergence order $p_2 = 1$,
3. global strong convergence order $\gamma_1 = 0.5$.

Proof: See [MPS98]. The proof is in a similar manner to the one for the BMM method. □

Theorem 3.33 The following statements hold for the BMM method

1. local convergence order $p_1 = 2$,
2. local square convergence order $p_2 = 1.5$,
3. global strong convergence order $\gamma_1 = 1.0$.

Proof: Statement 3 is a direct conclusion from the Milstein theorem and the assertions 1 and 2. Therefore we only need to verify p_1 and p_2. In the following let X^M be the local Milstein approximation

$$X^M_{n+1} = X_n + a(X_n)\Delta_{t_n} + b(X_n)\Delta W + \frac{1}{2}b(X_n)b'(X_n)\left((\Delta W_n)^2 - \Delta_{t_n}\right).$$

This integration step is called local because it is based on an arbitrary integration step X_n which will be the BMM method in the following. In addition to this let X denote the exact solution. Then

$$|X_{\tau_{n+1}}^{\tau_n,Y_n} - X_{n+1}| \leq \underbrace{|X_{\tau_{n+1}}^{\tau_n,Y_n} - X_{n+1}^M|}_{H_1} + \underbrace{|X_{n+1}^M - X_{n+1}|}_{H_2}. \tag{3.30}$$

Hence the problem is splitting up into two minor ones. The assertions for p_1 and p_2 are obvious for the term H_1, because this is only the Milstein scheme. The term H_2 demands further calculation. To simplify the notation we will use multiindices so that the schemes can now be reread as

$$\begin{aligned}
X_{n+1}^M &= X_n + aI_{(0)} + bI_{(1)} + bb'I_{(11)} \\
X_{n+1} &= X_n + aI_{(0)} + bI_{(1)} + bb'I_{(11)} + (X_n - X_{n+1})(c_0I_{(0)} + c_2I_{(11)}) \\
&= \left(1 + c_0I_{(0)} + c_2I_{(11)}\right)^{-1} \\
&\quad \Big(X_n + aI_{(0)} + bI_{(1)} + bb'I_{(11)} + X_n(c_0I_{(0)} + c_2I_{(11)})\Big).
\end{aligned}$$

This leads to the following expressions for the difference:

$$\begin{aligned}
X_{n+1}^M - X_{n+1} &= \left(1 + c_0I_{(0)} + c_2I_{(11)}\right)^{-1} \\
&\quad \Big[\left(1 + c_0I_{(0)} + c_2I_{(11)}\right) \left(X_n + aI_{(0)} + bI_{(1)} + bb'I_{(11)}\right) \\
&\quad -\Big(X_n + aI_{(0)} + bI_{(1)} + bb'I_{(11)} + X_n \left(c_0I_{(0)} + c_2I_{(11)}\right) \Big)\Big] \\
&= \left(1 + c_0I_{(0)} + c_2I_{(11)}\right)^{-1} \\
&\quad \Big[c_0I_{(0)} \left(X_n + aI_{(0)} + bI_{(1)} + bb'I_{(11)}\right) \\
&\quad +c_2I_{(11)} \left(X_n + aI_{(0)} + bI_{(1)} + bb'I_{(11)}\right) \\
&\quad -X_n \left(c_0I_{(0)} + c_2I_{(11)}\right) \Big] \\
&= \left(1 + c_0I_{(0)} + c_2I_{(11)}\right)^{-1} \\
&\quad \Big[c_0I_{(0)} \left(aI_{(0)} + bI_{(1)} + bb'I_{(11)}\right) \\
&\quad +c_2I_{(11)} \left(aI_{(0)} + bI_{(1)} + bb'I_{(11)}\right)] \\
&= \left(1 + c_0I_{(0)} + c_2I_{(11)}\right)^{-1} \\
&\quad \Big[c_0aI_{(00)} + c_0b \left(I_{(10)} + I_{(01)}\right) + c_0bb' \left(I_{(011)} + I_{(101)} + I_{(110)}\right) \\
&\quad +c_2a \left(I_{(011)} + I_{(101)} + I_{(110)}\right) + c_2b \left(3I_{(111)} + I_{(01)} + I_{(10)}\right) \\
&\quad +c_2bb' \left(6I_{(1111)} + 2 \left(I_{(011)} + I_{(101)} + I_{(110)}\right) + I_{(00)}\right) \Big].
\end{aligned}$$

Thus we have all the means we need to calculate p_1 and p_2

$$\begin{aligned}\left|E\left[X_{\tau_{n+1}}^{\tau_n,Y_n} - X_{n+1}|\mathcal{F}_{\tau_n}\right]\right| &\leq \left|E\left[X_{\tau_{n+1}}^{\tau_n,Y_n} - X_{n+1}^{M}|\mathcal{F}_{\tau_n}\right]\right| \\ &\quad + \left|E\left[X_{n+1}^{M} - X_{n+1}|\mathcal{F}_{\tau_n}\right]\right| \\ &\leq \left(1+|X_n|^2\right)^{1/2}\Delta_{t_n}^2.\end{aligned}$$

In the same way we obtain

$$\begin{aligned}\left(E\left[|X_{\tau_{n+1}}^{\tau_n,Y_n} - X_{n+1}|^2|\mathcal{F}_{\tau_n}\right]\right)^{1/2} &\leq \left(E\left[|X_{\tau_{n+1}}^{\tau_n,Y_n} - X_{n+1}^{M}|^2|\mathcal{F}_{\tau_n}\right]\right)^{1/2} \\ &\quad + \left(E\left[|X_{n+1}^{M} - X_{n+1}|^2|\mathcal{F}_{\tau_n}\right]\right)^{1/2} \\ &\leq \left(1+|X_n|^2\right)^{1/2}\Delta_{t_n}^{1.5}.\end{aligned}$$

To get both inequalities it is indispensable that

$$1 + c_0 I_{(0)} + c_2 I_{(11)} > 0.$$

Hence it is important for the convergence of balanced methods to restrict the control functions (see equation (3.28) and (3.29)). □

Summary

So far we have introduced a lot of different integration schemes. Obviously there have been many more endeavors in literature to integrate stochastic differential equations. Numerical positivity (or life time) of these schemes is of special interest in the next chapter.
We will recognize that in most cases the Milstein schemes preserve the analytical behaviour of a stochastic process. Consequently, we developed the BMM method to get an integration scheme combining both advantages. On the one hand we have the possibility to control the integration steps like in the BIM method. On the other hand the BMM method is based on the Milstein method. This is why it shows the same regularity behaviour.

CHAPTER IV

Positive numerical integration schemes

The numeric of stochastic differential equations analyses an integration scheme with regard to three basic principles:

1. Rate of convergence,
2. Stability,
3. Positivity (preserving the analytical behaviour).

Both, rate of convergence and stability, have been widely and intensively discussed in literature (see [GL97, HMGR99, KP92, Mil95, Sch97]). But in this thesis numerical positivity is the most important aspect. Already in chapter 1 we developed criteria to understand the analytical behaviour of a stochastic process. Therefore we only need to know drift and diffusion but we do not need to know the exact solution.

This chapter goes one step further. Actually, in the majority of cases we do not know a closed solution for a stochastic differential equation. Thus we need numerical integration schemes to do Monte-Carlo simulations. This leads to the problem of numerical positivity.
A simple example is the chart of a stock usually described by a geometric Brownian motion. If we integrate the corresponding differential equation numerically we have to maintain the analytical positivity because a negative stock price does not make sense. So one part of simulation is to take into account the characteristics of the model.

We keep up the notation of chapter 3.

4.1 Finite and eternal life time

Firstly we have to define numerical positivity. Schurz [Sch97] introduced the concept of numerical positivity in 1995.

Definition 4.1 Let X_t be a stochastic process with

$$P\left(\{X_t > 0 \text{ for all } t\}\right) = 1. \tag{4.1}$$

Then the stochastic integration scheme possesses an eternal life time if

$$P(\{X_{n+1} > 0 | X_n > 0\}) = 1. \tag{4.2}$$

Otherwise it has a finite life time.

So the life time depends only on one step of the integration scheme. Evidently we can define the life time for an arbitrary manifold $G \in \mathbb{R}^n$, too. Therefore the analytical behaviour of the stochastic process is as follows:

$$P(\{X_t \in G \text{ for all } t\}) = 1. \tag{4.3}$$

Consequently, a numerical integration scheme has an eternal life time if

$$P(\{X_{n+1} \in G | X_n \in G\}) = 1. \tag{4.4}$$

To get a better understanding of the life time of an integration scheme we should start with the easiest case, the Euler method. Unfortunately, this scheme possesses a finite life time for all stochastic differential equations.

Proposition 4.2 The Euler method has a finite life time for all stochastic differential equations.

Proof: Consider one integration step of the Euler scheme:

$$X_{n+1} = a(t_n, X_n)\Delta_{t_n} + b(t_n, X_n)\Delta W.$$

It suffices to prove the proposition for $a(t_n, X_n)b(t_n, X_n) > 0$. Then

$$X_{n+1} < 0 \Longleftrightarrow \Delta W < -\frac{a(t_n, X_n)}{b(t_n, X_n)}\Delta_{t_n}.$$

occurs with a positive probability. □

What is the key problem in this case? It is not too difficult to point out that the main problem is that ΔW takes all values $c \in \mathbb{R}$ with a positive probability.

There are mainly three different concepts to avert numerical negativity:

1. Balancing (BIM),
2. Dominating (Milstein),
3. Combining (BMM).

On the one hand the BIM method allows to choose appropriate control functions c_0 and c_1. Especially the function c_1 and the appendant random term $|\Delta W|$ allow to balance the term ΔW.

The other idea is the Milstein method because the term

$$b(X_n)b'(X_n)\left((\Delta W)^2 - \Delta_{t_n}\right)$$

can dominate the random term ΔW. Obviously we can combine both methods to be able to balance and to dominate.

4.2 Balanced implicit method

Firstly we sum up some facts about the BIM. A detailed description can be found in [MPS98, Sch96].

Lemma 4.3 The BIM has an eternal life time for the geometric Brownian motion (1.21)

$$dX_t = \lambda X_t dt + \sigma X_t dW$$

if the control functions c_0 and c_1 are such that

$$1 + (c_0 + \lambda)\Delta_{t_n} > 0 \text{ and } c_1 > \sigma. \tag{4.5}$$

Proof: Remembering the integration scheme of the BIM and deleting the implicitness leads to

$$\begin{aligned} X_{n+1} &= X_n + \lambda X_n \Delta_{t_n} + \sigma X_n \Delta W + (X_n - X_{n+1})(c_0\Delta_{t_n} + c_1|\Delta W|) \\ &= X_n \left(\frac{1 + (\lambda + c_0)\Delta_{t_n} + \sigma\Delta W + c_1|\Delta W|}{1 + c_0\Delta_{t_n} + c_1|\Delta W|} \right). \end{aligned}$$

Then the positivity for X_{n+1} directly follows from (4.5) □

In the case of the geometric Brownian motion it is simple to choose the functions c_0 and c_1 to provide positivity. But it is not possible to generalise

so that it is valid for any arbitrary differential equation as we can see in the next example:

$$dX_t = (\alpha - \beta X_t)dt + \sigma X_t^{\frac{1}{2}} dW. \tag{4.6}$$

In this case it is necessary to choose $c_1(x) = \sigma x^{-\frac{1}{2}}$. But if $x \to 0$, c_1 is an unbounded function and therefore we cannot guarantee convergence. So the BIM cannot achieve a finite life time for this process but it satisfies a weaker positivity concept: the ϵ-positivity.

Definition 4.4 Let $\epsilon > 0$ be an arbitrary constant. We call an integration scheme ϵ-positive if

$$P(\{X_{n+1} > 0 | X_n > \epsilon\}) = 1.$$

Remark 4.5 Evidently a scheme with eternal life time is ϵ-positive.

Lemma 4.6 The BIM method is ϵ-positive for the stochastic process (4.6) if we choose the control functions as follows:

$$\begin{aligned} c_0 &= \beta, \\ c_1(x) &= \begin{cases} \sigma x^{-\frac{1}{2}} & : \text{if } x > \epsilon, \\ \sigma \epsilon^{-\frac{1}{2}} & : \text{else.} \end{cases} \end{aligned}$$

Proof: Also in this case deleting the implicitness leads to

$$\begin{aligned} X_{n+1} &= \frac{X_n + (\alpha - \beta X_n)\Delta_{t_n} + \sigma X_n^{\frac{1}{2}} \Delta W + X_n(\beta \Delta_{t_n} + \sigma X_n^{-\frac{1}{2}} |\Delta W|)}{1 + \beta \Delta_{t_n} + \sigma X_n^{-\frac{1}{2}} |\Delta W|} \\ &= \frac{X_n + \alpha \Delta_{t_n} + \sigma X_n^{\frac{1}{2}} (\Delta W + |\Delta W|)}{1 + \beta \Delta_{t_n} + \sigma X_n^{-\frac{1}{2}} |\Delta W|}. \end{aligned}$$

Hence the BIM is ϵ-positive. □

4.3 Milstein method

The Milstein method has two advantages in comparison with the BIM. On the one hand the convergence speed is twice as high as in the BIM and on the other hand positivity can be achieved without using control functions. This point is of special importance because numerical tests show that using

c_0 and c_1 leads to a decreased convergence speed. When applying the Milstein method we also have to check the numerical positivity for the stochastic processes needed to model financial applications.

We start with a general result:

Theorem 4.7 The Milstein method

$$X_{n+1} = X_n + a(X_n)\Delta_{t_n} + b(X_n)\Delta W + \frac{1}{2}b(X_n)b'(X_n)\left((\Delta W)^2 - \Delta_{t_n}\right) \tag{4.7}$$

has an eternal life time if the following properties hold:

$$b(x)b'(x) > 0, \tag{4.8}$$

$$x > \frac{b(x)}{2b'(x),} \tag{4.9}$$

$$\Delta_{t_n} < \frac{2xb'(x) - b(x)}{\big(b(x)b'(x) - 2a(x)\big)b'(x)}. \tag{4.10}$$

The last condition is only necessary if the denominator is positive.

Proof: Let $x = X_n > 0$ and define $g(z) := b(x)z + \frac{1}{2}b(x)b'(x)z^2$. Then (4.7) can be written as

$$X_{n+1} = x + \left(a(x) - \frac{1}{2}b(x)b'(x)\right)\Delta_{t_n} + g(\Delta W).$$

According to (4.8) g possesses a global minimum. For that purpose an obvious calculation shows that

$$g'(z) = b(x) + b(x)b'(x)z.$$

Hence we get

$$\tilde{z} = -\frac{1}{b'(x)} \text{ with } g(\tilde{z}) = -\frac{b(x)}{2b'(x)}.$$

For this reason we can calculate the lower bound for all random terms ΔW. This enables us to exchange the value of $g(\Delta W)$ by its minimum

$$X_{n+1} \geq x + \left(a(x) - \frac{1}{2}b(x)b'(x)\right)\Delta_{t_n} - \frac{b(x)}{2b'(x)}.$$

Considering the requirements (4.9) and (4.10) we get that $X_{n+1} > 0$. □

Corollary 4.8 The Milstein method has an eternal life time for the numerical integration of the geometric Brownian motion (1.21) if $\Delta_{t_n} < \frac{1}{\sigma^2 - 2\mu}$.

Proof: We only need to verify the requirements of the theorem above, thus $a(x) = \mu x$ as well as $b(x) = \sigma x$ and $b'(x) = \sigma$. So obviously the requirements (4.8) and (4.9) are satisfied and we get (4.10) with a short calculation:

$$\begin{aligned}\frac{2xb'(x)-b(x)}{\big(b(x)b'(x)-2a(x)\big)b'(x)} &= \frac{2x\sigma - x\sigma}{(\sigma^2 x - 2\mu x)\sigma} \\ &= \frac{1}{\sigma^2 - 2\mu} > \Delta_{t_n}.\end{aligned}$$

In most financial applications we have $\sigma < 1$. Hence the step size restriction allows a $\Delta_{t_n} > 1$ which is of course satisfied as we need a finer discretisation to get an adequate approximation. □

In the case of the Brownian motion the question of numerical positivity does not arise because the exact solution is already known. That is one big difference to the class of mean-reverting processes.

4.4 Mean-reverting processes

We have already noticed that the BIM only guarantees ϵ-positivity for a mean-reverting process. Thus the applicability is restricted.
To check the suitability of the Milstein method we consider an arbitrary mean-reverting process

$$dX_t = (\alpha - \beta X_t^q)dt + \sigma X_t^p dW.$$

With this arbitrary parameter constellation it is not possible to get a closed solution. The question of analytical positivity has already been studied in chapter 1. First we will take a closer look at the most common mean-reverting process. In scientific applications this process presents the usual model if X_t oscillates around a mean

$$dX_t = (\alpha - \beta X_t)dt + \sigma X_t^{\frac{1}{2}} dW. \tag{4.11}$$

In chapter 1 we have analysed this process and proved its positivity in the case that $\alpha > \frac{\sigma^2}{2}$. The explicit Milstein scheme is not the right one for the numerical integration because of the negative sign in front of X_t in the drift. Therefore it is necessary to apply the implicit Milstein method.

$$X_{n+1} = X_n + a(X_{n+1})\Delta_{t_n} + b(X_n)\Delta W + \frac{1}{2}b(X_n)b'(X_n)\left((\Delta W)^2 - \Delta_{t_n}\right).$$

Theorem 4.9 The implicit Milstein method has an eternal life time for the integration of the stochastic process (4.11) independent of stepsize Δ_{t_n}.

Proof: One integration step is as follows

$$X_{n+1} = X_n + (\alpha - \beta X_{n+1})\Delta_{t_n} + \sigma X_n^{\frac{1}{2}} \Delta W + \frac{1}{4}\sigma^2 \left((\Delta W)^2 - \Delta_{t_n} \right).$$

In an elementary way we can eliminate the implicitness

$$X_{n+1} = \frac{X_n + \alpha\Delta_{t_n} + \sigma X_n^{\frac{1}{2}} \Delta W + \frac{1}{4}\sigma^2 \left((\Delta W)^2 - \Delta_{t_n} \right)}{1 + \beta\Delta_{t_n}} = \frac{N(X_n)}{D(X_n)}.$$

Now we only have to verify that the numerator $N(X_n)$ is positive. Employing the idea of Theorem 4.7 with $a(x) = \alpha$ and $b(x) = \sigma x^{\frac{1}{2}}$ we obtain

$$\begin{aligned}
N(X_n) &= X_n + \left(\alpha - \frac{1}{4}\sigma^2\right)\Delta_{t_n} + g(\Delta W) \\
&\geq X_n + \left(\alpha - \frac{1}{4}\sigma^2\right)\Delta_{t_n} + \min_{\Delta W \in \mathbb{R}} g(\Delta W) \\
&= X_n + \left(\alpha - \frac{1}{4}\sigma^2\right)\Delta_{t_n} + g(\Delta \tilde{W}) \\
&= X_n + \left(\alpha - \frac{1}{4}\sigma^2\right)\Delta_{t_n} - X_n \\
&= \left(\alpha - \frac{1}{4}\sigma^2\right)\Delta_{t_n} > 0,
\end{aligned}$$

since the analytical positivity causes $\alpha > \frac{1}{2}\sigma^2$. □

This result is rather surprising because the Milstein method provides numerical positivity without any restrictions. Now we can deal with a more general case

$$dX_t = (\alpha - \beta X_t)dt + \sigma X_t^p dW, \tag{4.12}$$

where $p > \frac{1}{2}$. If this is so the Milstein method can also preserve positivity.

Lemma 4.10 The step size adapted implicit Milstein method has an eternal life time for the integration of (4.12).

Proof: Again the first step of the proof is eliminating implicitness

$$X_{n+1} = \frac{X_n + \alpha\Delta_{t_n} + \sigma X_n^p \Delta W + \frac{1}{2}p\sigma^2 X_n^{2p-1} \left((\Delta W)^2 - \Delta_{t_n} \right)}{1 + \beta\Delta_{t_n}} = \frac{N(X_n)}{D(X_n)}.$$

Also in this case we have to check the positivity only for the numerator. Remembering Theorem 4.7 we ascertain that the drift $a(x) = \alpha$ and the diffusion $b(x) = \sigma x^p$ fulfils

$$b(x)b'(x) = \sigma^2 p x^{2p-1} > 0,$$

just as

$$\frac{b(x)}{2b'(x)} = \frac{x}{2p} < x.$$

So the restriction (4.10) on Δ_{t_n} gives us the step size specification

$$\begin{aligned} \Delta_{t_n} &< \frac{2xb'(x) - b(x)}{\big(b(x)b'(x) - 2a(x)\big)b'(x)} \\ &= \frac{2\sigma p x^p - \sigma x^p}{(\sigma^2 p x^{2p-1} - 2\alpha)\sigma x^{p-1}} \\ &= \frac{2p-1}{p\sigma^2 x^{2p-2} - 2\alpha x^{-1}}. \end{aligned}$$

This requirement strictly depends on the parameter p:

Case 1: $(p = 1)$
Considering the assumption $\sigma^2 > 2\alpha x^{-1}$ we get

$$\frac{1}{\sigma^2 - 2\alpha x^{-1}} \geq \frac{1}{\sigma^2} > \Delta_{t_n}.$$

This enables us to fix the stepsize at the beginning.

Case 2: $(\frac{1}{2} < p < 1)$
Δ_{t_n} has only to be restricted if the denominator is positive in restriction (4.10). So this leads to

$$\begin{aligned} & p\sigma^2 x^{2p-2} > 2\alpha x^{-1} \\ \iff \quad & x^{2p-1} > \tfrac{2\alpha}{p\sigma^2} \\ \iff \quad & x > \left(\tfrac{2\alpha}{p\sigma^2}\right)^{\frac{1}{2p-1}} =: \ell. \end{aligned}$$

Therefore x is lower bounded. Hence it is only necessary to analyse the behaviour of the denominator in the interval $I = [\ell, \infty)$:

$$N(\ell) = 0 \text{ and } \lim_{x \longrightarrow \infty} N(x) = 0.$$

The denominator is a continuous function on a compact interval. Therefore it has got a maximum. Due to the fact that the numerator is constant this maximum gives us the restriction for the stepsize Δ_{t_n}. Basic calculation verifies that

$$\tilde{x} = \left(\frac{2\alpha}{(2-2p)p\sigma^2} \right)^{\frac{1}{2p-1}}$$

is the maximiser of the denominator. This term seems to be quite difficult but we only need an appropriate approach. For that reason we see that the restriction for the step size in the case $p = 1$ is a sufficient condition for $\frac{1}{2} < p < 1$ as well.

Case 3: $(p > 1)$
The calculation is similar to case 2. But now the problem is that the denominator increases with an increasing x. As x does not have to be bounded above we cannot achieve a lower bound for Δ_{t_n} at the beginning of the procedure. In this case it is only possible to use a step size adapted integration scheme recalculating Δ_{t_n} in each step.

This proof shows that the parameter p of a mean-reverting process has a great impact on the numerical behaviour of the Milstein scheme. □

Remark 4.11 Alternatively one can use the BIM method if $p > 1$. Numerical tests show that in this case also the Milstein method is the better choice because of its higher convergence order and the use of control functions in the BIM.

4.5 Integration in the extended Libor market model

Concluding this chapter we will study the numerical behaviour of different stochastic differential equations which model the forward rates in the extended Libor market model.

Integration of Displaced Diffusion

The displaced diffusion models the forward rates as follows (see also section 2.2.1 and especially the equation (2.7)):

$$dX_t = \sigma_{DD}(t)(X_t + m)dB.$$

Remembering the analytical behaviour we know that this process takes values in $[-m, \infty)$.

Proposition 4.12 The explicit Milstein method has an eternal life time for the integration of equation (4.5).

Remark 4.13 Eternal life time has to be understood in a broader sense because the analytical structure is the interval $[-m, \infty)$.

Proof: The Milstein method can be written as (see equation (4.3) in the proof of Theorem 4.7):

$$\begin{aligned} X_{n+1} &= x + \left(a(x) - \frac{1}{2}b(x)b'(x)\right)\Delta_{t_n} + g(\Delta W) \\ &\geq x + \left(a(x) - \frac{1}{2}b(x)b'(x)\right)\Delta_{t_n} + \min_{\Delta W \in \mathbb{R}} g(\Delta W) \\ &= x + \left(a(x) - \frac{1}{2}b(x)b'(x)\right)\Delta_{t_n} - \frac{b(x)}{2b'(x)}. \end{aligned}$$

The next step is to insert the parameters $a = 0$ and $b = \sigma_{DD}(x + m)$

$$\begin{aligned} X_{n+1} &= x - \frac{1}{2}\sigma_{DD}^2(x+m)\Delta_{t_n} - \frac{m+x}{2} \\ &= \frac{1}{2}x(1 - \sigma_{DD}^2\Delta_{t_n}) - \frac{1}{2}m(1 + \sigma_{DD}^2\Delta_{t_n}) \\ &\geq \frac{1}{2}m(1 - \sigma_{DD}^2\Delta_{t_n}) - \frac{1}{2}m(1 + \sigma_{DD}^2\Delta_{t_n}) \\ &= m. \end{aligned}$$

Therefore the Milstein method has an eternal life time for the displaced diffusion model. □

Integration of Constant Elasticity of Variance

The model of constant elasticity of variance describes the forward rates as follows (see also section 2.2.2 and equation (2.9)):

$$dX_t = \sigma_{CEV}(t)X_t^{\alpha}dB \tag{4.13}$$

In this case the stochastic process takes only positive values. But the boundary 0 has a different behaviour with respect to α (see equation 1.23)).

Proposition 4.14 The Milstein method has the following properties for the integration of equation (4.13):

- $\alpha \leq \frac{1}{2} \Rightarrow$ finite life time,
- $\alpha > \frac{1}{2} \Rightarrow$ eternal life time if step size is adapted.

Proof: The same calculation as in the case of the DD model leads to

$$\begin{aligned} X_{n+1} &= x - \frac{1}{2}\alpha^2\sigma_{CEV}^2 x^{2\alpha-1}\Delta_{t_n} - \frac{x}{2\alpha} \\ &= x\left(1 - \frac{1}{2\alpha}\right) - \frac{1}{2}\alpha^2\sigma_{CEV}^2 x^{2\alpha-1} \end{aligned}$$

with $a = 0$ and $b = \sigma_{CEV}x^{\alpha}$. If $\alpha \leq \frac{1}{2}$ both terms are negative so that we get a finite life time. In case that $\alpha > \frac{1}{2}$ we can calculate further:

$$X_{n+1} = \frac{x}{2\alpha}\left(2\alpha - 1 - \alpha^2\sigma_{CEV}^2 x^{2\alpha-2}\Delta_{t_n}\right).$$

Then the step size adaption is obviously

$$\Delta_{t_n} < \frac{2\alpha - 1}{\alpha^2\sigma_{CEV}^2 x^{2\alpha-2}}.$$

In this inequality we have to distinguish between two cases. Firstly if $\frac{1}{2} < \alpha < 1$ then the discretisation size Δ_{t_n} must be small if X_n is small. On the other hand if $1 < \alpha$ we get a decreasing step size with an increasing X_n. □

Proposition 4.15 The BIM has the following properties for the integration of equation (4.13):

- $\alpha < 1 \Rightarrow \epsilon$-positive,
- $\alpha \geq 1 \Rightarrow$ eternal life time.

First let $\alpha < 1$. Then the parameters c_0 and c_1 can be chosen as

$$\begin{aligned} c_0 &= 0, \\ c_1(x) &= \begin{cases} \sigma_{CEV}x^{\alpha-1} & \text{if } x > \epsilon, \\ \sigma_{CEV}\epsilon^{\alpha-1} & \text{else.} \end{cases} \end{aligned}$$

For $\alpha \geq 1$ the choice is much easier

$$\begin{aligned} c_0 &= 0, \\ c_1(x) &= \sigma_{CEV}x^{\alpha-1}. \end{aligned}$$

Proof: Straight forward calculation verifies the statements. □

Integration of stochastic volatility

The dynamic of the Libor market model can be further improved by introducing stochastic volatility. Usually the stochastic process describing the volatility is a mean-reverting process. Since this extension is used as an additive in the Libor market model it is sufficient to study the numerical characteristics of mean-reverting processes. This has already been done in this chapter (see section 4.4).

Summary

The Milstein scheme is an integration method that preserves the analytical behaviour of a stochastic process in a natural way. It is not necessary to use control functions to prevent the numerical values from becoming negative. Numerical tests show that this property strongly influences the convergence speed of integration schemes as the use of control functions leads to higher computational cost and to a reduced approximation accuracy.

CHAPTER V

Numerical tests

The last chapter of this thesis deals with numerical tests. Thereby the theoretical results should be verified in practice. Special attention is paid to the aspect of numerical positivity.
The tests are mainly done with regard to two different aspects. On the one hand we study how exact the Monte-Carlo approximation can approach the caplet prices. This comparison can be done quite easily because we already have a closed solution for caplets.
On the other hand we analyse the convergence speed and the preserving of positivity without any direct financial applications. This will be done by integrating different mean-reverting processes.

Remark 5.1 The Balanced Milstein method is introduced in this thesis but it is not incorporated in the numerical tests. The main reason is that in the majority of cases the use of control function decreases the convergence speed and increases the computational cost. As the Milstein method already preserves positivity it is not necessary to use the BMM.

5.1 Application in financial mathematics

The numerical simulation of caplet prices is an important indicator to study the applicability of the different integration schemes in the Libor market model. Thus the approximation behaviour is of interest as well as the computational time. The caplet price approximation is connected with the weak convergence speed because we compare expectations. To make this point clear let $F_k(t)$ be the analytical forward rate and $F_k^N(t)$ the numerical one.

Then the difference of the caplet prices is given by

$$\left| \underbrace{\delta_k B(t,T_{k+1}) E\left[(F_k(T_k)-K)^+\right]}_{C(t)} - \delta_k B(t,T_{k+1}) E\left[(F_k^N(T_k)-K)^+\right] \right|$$

with the exact caplet price $C(t)$. Thus the obvious thing to do would be comparing caplet prices. In consideration of the fact that we are interested in pricing path-dependent interest rates derivatives the strong convergence order is more important. In the most simple case of the Libor market model the forward rates are driven by a geometric Brownian motion. Hence we have a closed formula. This enables us to compare the strong convergence order:

$$E\left[|F_k(T_k) - F_k^N(T_k)|\right].$$

So the applicability of the different integration schemes will be tested for the caplet price approximation as well as for the strong convergence towards the forward rates.

Integration schemes

In this section we compare the following four integration schemes:

1. Euler scheme,
2. Milstein scheme,
3. BIM,
4. log-Euler scheme.

The log-Euler schemes is constructed by transforming the original stochastic differential equation using the Ito theorem and the logarithm function. This leads to the following method:

Definition 5.2 The log-Euler method for the integration of the stochastic differential equation

$$dF_k = \sigma_k \varphi(F_k) dB^{T_{k+1}},$$

is given by:

$$F_{n+1} = F_n \exp\left(-\frac{1}{2}\left(\sigma_k \frac{\varphi(F_n)}{F_n}\right)^2 \Delta_{t_n} + \frac{\varphi(F_n)}{F_n}\Delta W \right).$$

Obviously we get the CEV model for $\varphi(x) = x^\alpha$ and the DD model for $\varphi(x) = x + m$.

Proof: For detailed proof see [BR97]. We will only look at a short sketch of the proof. Define $g(t,x) = \log(x)$ and apply Ito's theorem on the forward rate $F_k(t)$, then

$$\begin{aligned} dg(F_k(t)) &= \frac{1}{F_k(t)} dF_k(t) - \frac{1}{2(F_k(t))^2} (dF_k(t))^2 \\ &= \sigma \frac{\varphi(F_k(t))}{F_k(t)} dW - \frac{1}{2} \left(\sigma \frac{\varphi(F_k(t))}{F_k(t)} \right)^2 dt. \end{aligned}$$

Thus we can directly deduce an integration scheme for $g(F_k(t))$:

$$\begin{aligned} g(F_{n+1}) &= g(F_n) + \left(-\frac{1}{2} \left(\sigma_k \frac{\varphi(F_n)}{F_n} \right)^2 \Delta_{t_n} + \frac{\varphi(F_n)}{F_n} \Delta W \right) \\ \Rightarrow \quad F_{n+1} &= F_n \exp \left(-\frac{1}{2} \left(\sigma_k \frac{\varphi(F_n)}{F_n} \right)^2 \Delta_{t_n} + \frac{\varphi(F_n)}{F_n} \Delta W \right). \end{aligned}$$

This motivates the log-Euler method but there is still no answer to the question of numerical convergence. □

Remark 5.3 All numerical tests use an equidistant discretisation size Δ_{t_n} which we denote with Δ_t.

5.1.1 Forward rate approximation

First of all we will analyse the strong convergence speed of the different integration schemes. Our test equation is:

$$dF_k(t) = 0.2 F_k(t) dB^{T_{k+1}} \text{ with } F_k(0) = 0.06.$$

The exact solution is

$$F_k(t) = F_k(0) \exp \left(-\frac{1}{2} \sigma^2 t + \sigma W_t \right).$$

Remark 5.4 Integrating this stochastic differential equation the log-Euler scheme provides accurate numerical results because the transformation with the exponential function leads to the analytical solution. Therefore we will only compare the other three methods.

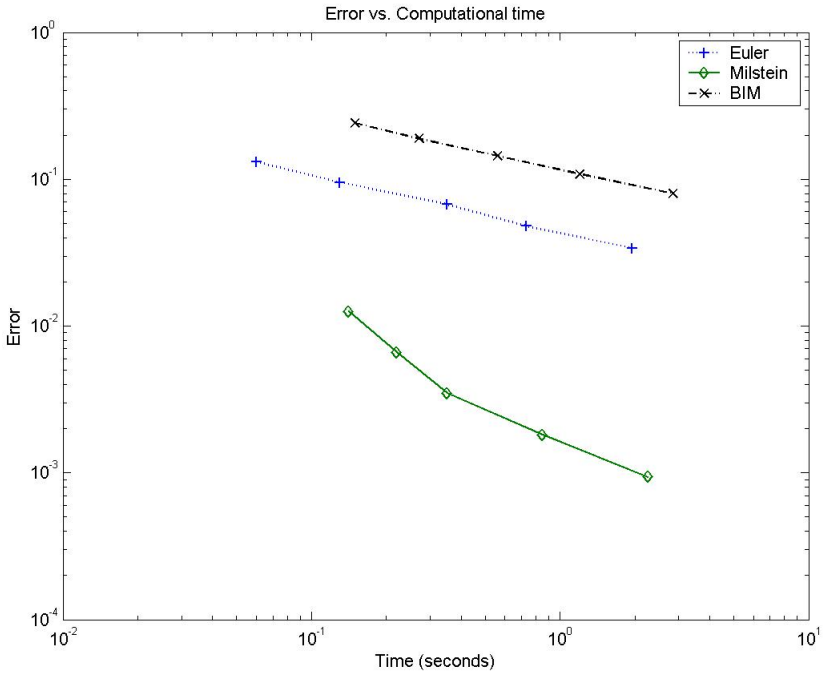


Figure 5.1: $F_k(0) = 0.06$, $\sigma_k = 0.2$, number of paths 2000000, Error $= E\left[|F_k(T_k) - F_k^N(T_k)|\right]$, discretisation stepsizes: $\Delta_t = 1, 0.5, 0.25, 0.125, 0.0625$.

Picture 5.1 confirms our expectation. The Milstein method is superior in this comparison. With fixed computational time it has the smallest error and vice versa with fixed error it needs least computational time. By the way the Milstein method preserves the analytical positivity see corollary 4.8.

Unfortunately we have a closed solution for the forward rates only in the most simple case. Thus we can only compare the caplet price approximation in the extended version of the Libor market model.

5.1.2 Constant Elasticity of Variance

Firstly we analyse the numerical behavior in the CEV extension. This model is characterised by the fact that the forward rates take only nonnegative values. Indeed the boundary 0 is attainable if $0 < \alpha < 1$.

CEV model with $\alpha = 0.5$

I	Maturity	T = 1	T = 2	T = 4	T = 6	T = 10	Comp.
	Exact	21.878	29.124	36.506	39.623	40.171	time
$\Delta_t = \frac{1}{2}$	Euler	21.891	29.163	36.581	39.686	40.242	1.8
	Milstein	21.900	29.161	36.570	39.695	40.239	2.0
	BIM	18.230	24.185	30.218	32.802	33.270	3.3
	log-Euler	21.903	29.163	36.559	39.703	40.282	3.2
$\Delta_t = \frac{1}{4}$	Euler	21.885	29.141	36.531	39.637	40.197	5.0
	Milstein	21.894	29.147	36.547	39.663	40.211	5.5
	BIM	19.061	25.313	31.701	34.417	34.913	8.3
	log-Euler	21.888	29.121	36.545	39.654	40.224	8.0
$\Delta_t = \frac{1}{8}$	Euler	21.880	29.149	36.513	39.611	40.193	16.5
	Milstein	21.878	29.126	36.516	39.635	40.177	17.7
	BIM	19.754	26.295	32.930	35.727	36.266	23.3
	log-Euler	21.903	29.151	36.525	39.601	40.158	22.6

Table 5.1: $dF_k(t) = \sigma_k F_k(t)^\alpha dB^{T_{k+1}}$ with: $F_k(0) = 0.06$, $\sigma_k = 0.04899$, $\alpha = 0.5$, tenor spacing $T = 0.5$, maturity times $T_k = 1, 2, 4, 6, 10$, strike $K = 0.06$, Monte Carlo simulation with 5000000 paths.

In table 5.1 we see that the choice of the integration scheme does not make any decisive difference as long as we do not use the BIM. At first glance the other three schemes seems to be indistinguishable. But taking a more precise look we notice that with a decreased stepsize the Milstein scheme is the best choice. Especially with the smallest discretisation we get a rather close approach.
Two reason for this behaviour can be identified. The first is obviously the higher convergence order of the Milstein scheme. The second reason is on the one hand the numerical instability of the Euler scheme and on the other hand the point that the log-Euler scheme enforces the numerical positivity. This enforcing averts better results. We will see this handicap in the other numerical simulations, too.
Another important aspect is the computational time. We notice that the difference between the Euler and the Milstein scheme is marginal. The BIM is no appropriate choice with regard to the computational cost either.

The relation between approximation error and computational cost becomes clear by a a comparison of these two aspects. This is done in the next figure 5.2.

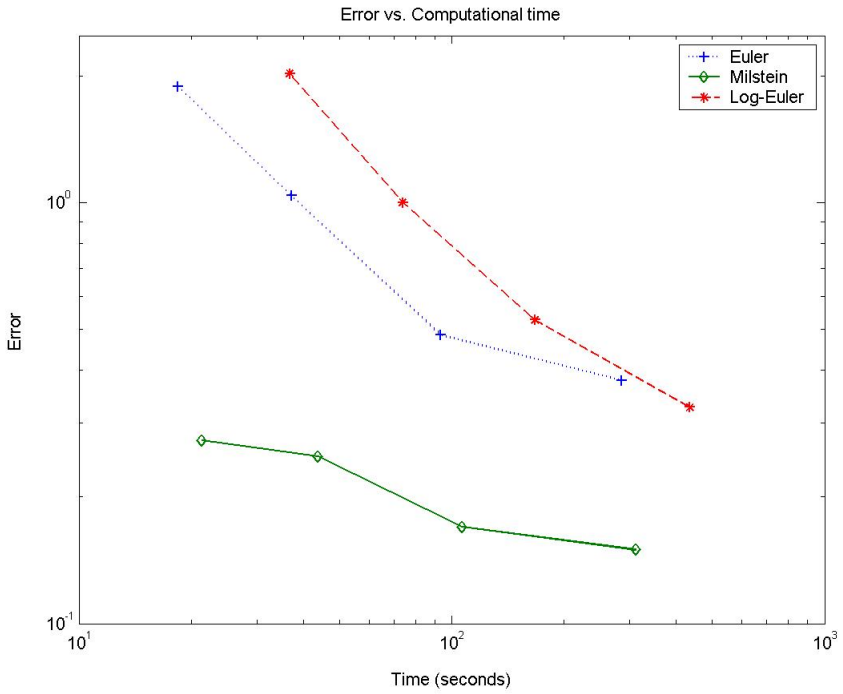


Figure 5.2: $dF_k(t) = \sigma_k F_k(t)^\alpha dB^{T_{k+1}}$ with: $F_k(0) = 0.06$, $\sigma_k = 0.04899$, $\alpha = 0.5$, tenor spacing $T = 0.5$, maturity times $T_k = 1, 2, 4, 6$, strikes $K = 0.04, 0.05, 0.06, 0.07, 0.08$, number of paths 2000000, Error $= ||C_{\text{exact}} - C_{MC}||_{2, K \times T_k}$ (grid points: strike $\times$ maturity times), discretisation stepsizes: $\Delta_t = 1, 0.5, 0.25, 0.125$.

Comparing these three methods shows that using the Milstein scheme leads to an exact approximation in shortest time. Particularly for strikes which are not at-the-money the Milstein scheme provides the best results. This fact is just mentioned as it makes no sense to present more tables with pure data.

Another important aspect is the error of the integration scheme compared with the number of simulated paths. Here we have to answer the question how many paths we need to get a sufficient approach.

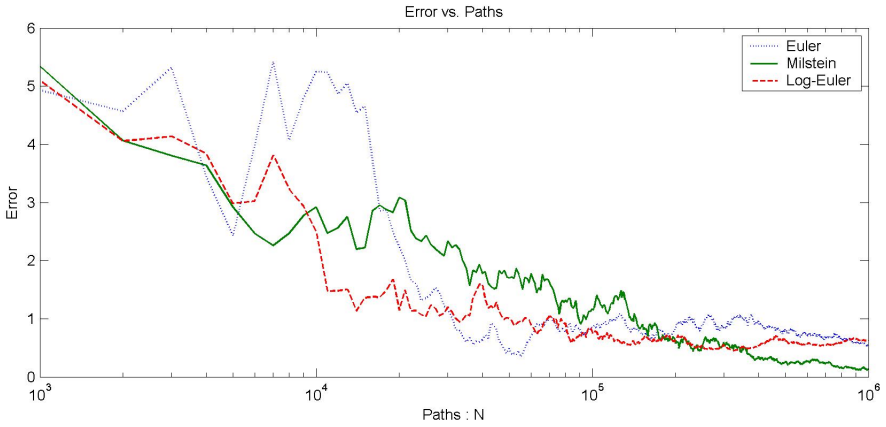


Figure 5.3: $F_k(0) = 0.06$, $\sigma_k = 0.04899$, $\alpha = 0.5$, tenor spacing $T = 0.5$, maturity times $T_k = 1, 2, 4, 6, 10$, strikes $K = 0.04, 0.05, 0.06, 0.07, 0.08$, integration steps $\Delta_t = 0.25$, Error $= ||C_{\text{exact}} - C_{MC}||_{2,K\times T_k}$ (grid points: strike $\times$ maturity times)

The figure 5.3 shows the error of the numerical caplet calculation against the number of simulated paths. We see that a minimum of 100000 paths is necessary to be able to interpret the convergence behaviour of the different schemes. Simulating less paths the oscillation makes an analysis of the approach impossible.

We notice that 1000000 paths are sufficient to guarantee a stable approximation. Thus the simulation of 5000000 paths in the first example table 5.1 and 2000000 in figure 5.2 is surely an appropriate choice to interpret the different integration schemes.

A careful consideration of this test shows that the systematic error of the Milstein scheme is the lowest and the simulation of an exceeding number of paths stabilises this result.

CEV model with $\alpha = 1.5$

The situation in the CEV model for $\alpha = 1.5$ is quite different as we can see in table 5.2 on the next page.

II	Maturity	T = 1	T = 2	T = 4	T = 6	T = 10	Comp.
	Exact	21.867	29.113	36.488	39.604	40.151	time
$\Delta_t = \frac{1}{2}$	Euler	22.115	29.638	37.245	40.498	43.545	2.3
	Milstein	21.885	29.060	36.387	39.496	40.092	2.5
	BIM	18.392	24.347	30.456	33.018	33.522	3.9
	log-Euler	21.904	29.174	36.567	39.606	39.350	2.2
$\Delta_t = \frac{1}{4}$	Euler	22.067	29.444	36.918	40.094	43.063	6.1
	Milstein	21.830	29.049	36.405	39.475	40.138	6.5
	BIM	19.140	25.432	31.904	34.611	35.100	10.3
	log-Euler	21.860	29.087	36.498	39.550	39.509	5.9
$\Delta_t = \frac{1}{8}$	Euler	21.948	29.266	36.634	39.758	41.470	18.7
	Milstein	21.878	29.123	36.495	39.589	40.152	19.6
	BIM	19.839	26.359	33.013	35.864	36.336	35.3
	log-Euler	21.873	29.116	36.488	38.629	39.322	18.4

Table 5.2: $F_k(0) = 0.06$, $\sigma_k = 0.8161$, $\alpha = 1.5$, tenor spacing $T = 0.5$, maturity times $T_k = 1, 2, 4, 6, 10$, $K = 0.06$, Monte Carlo simulation with 5000000 paths

First we notice that in this case the computational times are nearly equal with a little advantage for the log-Euler scheme. Again the comparison of the computational effort with the approximation error is more significant.

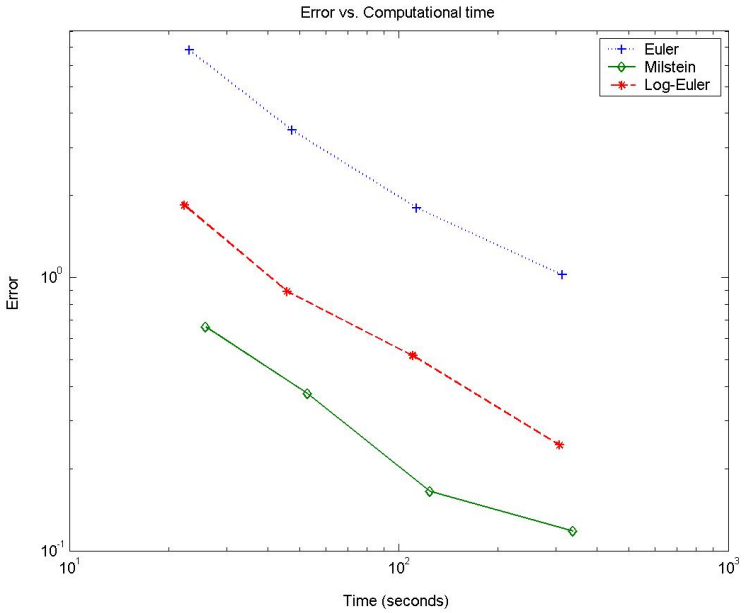


Figure 5.4: $\sigma_k = 0.8161$, $\alpha = 1.5$, all other parameters as in figure 5.2.

The Milstein scheme provides in this case also the best results in the lowest time.

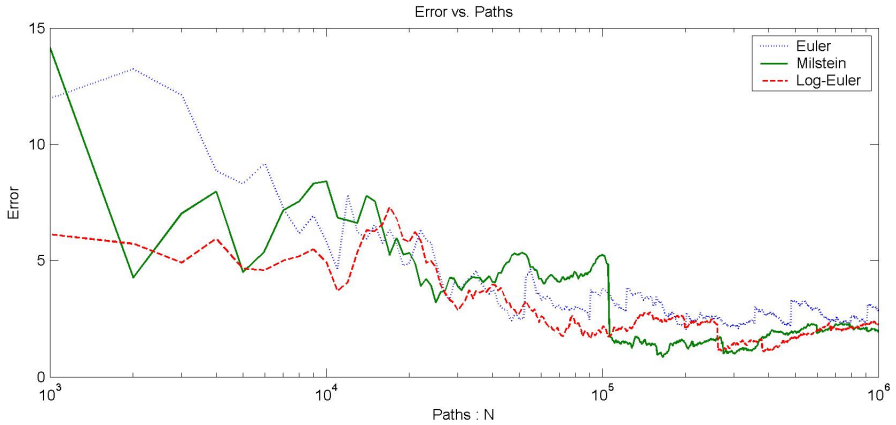


Figure 5.5: $F_k(0) = 0.06$, $\sigma_k = 0.8161$, $\alpha = 1.5$, tenor spacing $T = 0.5$, maturity times $T_k = 1, 2, 4, 6, 10$, strikes $K = 0.04, 0.05, 0.06, 0.07, 0.08$, integration steps $\Delta_t = 0.25$, Error $= ||C_{\text{exact}} - C_{MC}||_{2,K\times T_k}$ (grid points: strike $\times$ maturity times)

Comparing the error behaviour in the last CEV test, figure 5.5 above, we notice that all in all the error is higher than in the case of $\alpha = 0.5$. Here as well we see that the error behaviour stabilises when we reach a number of 100000 simulated paths.

Conclusion CEV

The CEV model shows that different choices of the parameter α have a great impact on the numerical simulation. In the case that $\alpha = 0.5$ all integration schemes provide sufficient caplet price approximations. Also for great discretisation stepsizes Δ_t the approaches are sufficient.
In contrast the quality of approximation significantly depends on integration scheme and discretisation stepsize when $\alpha = 1.5$.

Summing up we ascertain that the Milstein scheme turned out to be the best method to simulate the forward rates in the CEV model.

5.1.3 Displaced Diffusion

The model of displaced diffusion is a second possibility to get a smile in the implicit volatility surface. Indeed the problem is that the forward rates can take negative values as the stochastic process lives on the interval $[-m, \infty)$. Therefore it makes no sense to try to preserve positive values in the numerical integration. The analysis of the Milstein method in the chapter before shows that it provides an eternal life time with respect to the interval $[-m, \infty)$. This raises up the question if the numerical approximation of caplet prices profits from this fact.

Remark 5.5 The numerical results of the BIM are also inadequate in the displaced diffusion model. Accordingly the outcomes are not presented.

DD model with $m = 0.02$

III	Maturity	T = 1	T = 2	T = 4	T = 6	T = 10	Comp.
	Exact	29.159	38.805	48.600	52.707	53.362	time
$\Delta_t = \frac{1}{2}$	Euler	29.318	39.062	49.027	53.214	53.899	1.2
	Milstein	29.195	38.795	48.563	52.693	53.358	1.2
$\Delta_t = \frac{1}{4}$	Euler	29.268	38.988	48.887	53.036	53.690	4.0
	Milstein	29.170	38.816	48.623	52.768	53.385	4.0
$\Delta_t = \frac{1}{8}$	Euler	29.215	38.894	48.780	52.846	53.407	14.6
	Milstein	29.188	38.805	48.629	52.698	53.354	14.6

Table 5.3: $F_k(0) = 0.06$, $\sigma_k = 0.2$, $m = 0.02$, tenor spacing $T = 0.5$, maturity times $T_k = 1, 2, 4, 6, 10$, strike $K = 0.06$, Monte Carlo simulation with 5000000 paths.

In the first example, table 5.3, we only present the results for the Euler and the Milstein scheme as the log-Euler has great numerical instabilities. The reason for these instabilities is that the log-Euler enforces numerical positivity but in this case the stochastic process takes values in $[-0.02, \infty)$. Comparing the Euler and the Milstein method we obtain a familiar result. The Milstein scheme is superior for the approximation and on the other hand the computational times are nearly equal. This result becomes more meaningful comparing error and computational time.

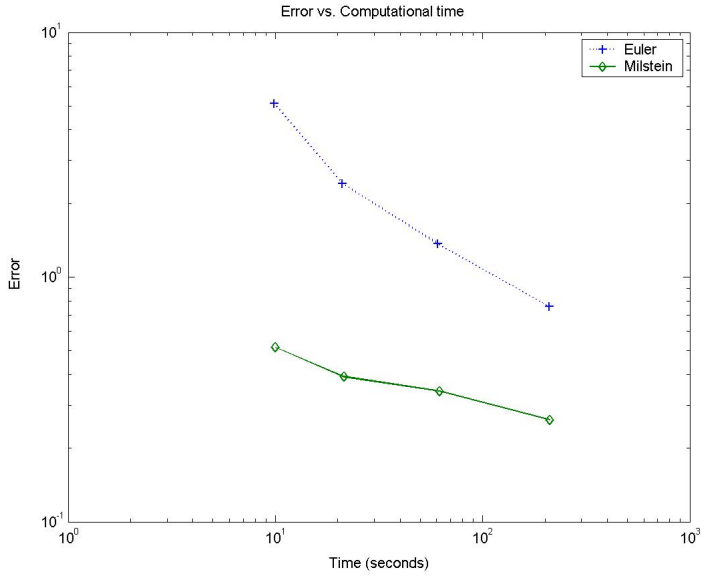


Figure 5.6: $F_k(0) = 0.06$, $\sigma_k = 0.2$, $m = 0.02$, tenor spacing $T = 0.5$, maturity times $T_k = 1, 2, 4, 6$, strikes $K = 0.04, 0.05, 0.06, 0.07, 0.08$, number of paths 2000000, Error $= ||C_{\text{exact}} - C_{MC}||_{2,K\times T_k}$ (grid points: strike $\times$ maturity times)

Also for the DD model the number of necessary paths is of great interest for the numerical simulation. Picture 5.7 below shows that again 100000 paths are the lower bound to reach a satisfactory approach.

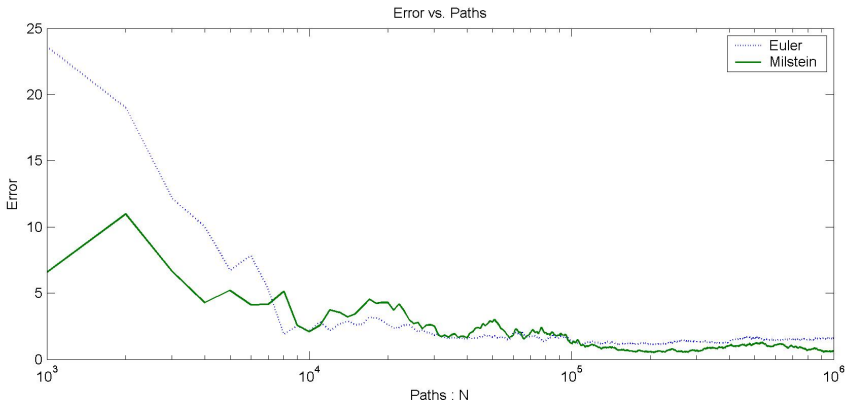


Figure 5.7: $F_k(0) = 0.06$, $\sigma_k = 0.2$, $m = 0.02$, tenor spacing $T = 0.5$, maturity times $T_k = 1, 2, 4, 6, 10$, strikes $K = 0.04, 0.05, 0.06, 0.07, 0.08$, integration steps $\Delta_t = 0.25$, Error $= ||C_{\text{exact}} - C_{MC}||_{2,K\times T_k}$ (grid points: strike $\times$ maturity times)

DD model with $m = -0.02$

A negative choice of the parameter m is a new challenge for the different integration schemes as the forward rates live in the interval $[-m, \infty)$.

IV	Time	T = 1	T = 2	T = 4	T = 6	T = 10	Comp.
	Exact	14.580	19.402	24.298	26.353	26.681	Cost.
$\Delta_t = \frac{1}{2}$	Euler	14.658	19.571	24.513	26.580	26.890	1.2
	Milstein	14.601	19.412	24.313	26.391	26.698	1.2
	log-Euler	14.588	19.432	24.317	26.383	26.700	1.3
$\Delta_t = \frac{1}{4}$	Euler	14.644	19.432	24.317	26.383	26.700	4.0
	Milstein	14.586	19.415	24.304	26.367	26.674	4.0
	log-Euler	14.575	19.392	24.296	26.322	26.691	4.1
$\Delta_t = \frac{1}{8}$	Euler	14.616	19.459	24.367	26.414	26.742	14.6
	Milstein	14.561	19.383	24.312	26.368	26.707	14.6
	log-Euler	14.577	19.410	24.308	26.338	26.628	14.8

Table 5.4: $F_k(0) = 0.06$, $\sigma_k = 0.2$, $m = -0.02$, tenor spacing $T = 0.5$, maturity times $T_k = 1, 2, 4, 6, 10$, strike $K = 0.06$, Monte Carlo simulation with 5000000 paths.

Calculating the caplet prices shown in table 5.4 is particularly characterised by rather good approximations no matter which integration scheme we use. But giving more attention we notice that actually the Milstein method provides the best results. One reason is that only the Milstein scheme makes the numerical approximation to stay in the interval $[0.02, \infty)$.

Figure 5.8 on the next page compares the computational effort with the quality of the approximation and again the Milstein method turns out to be the best.

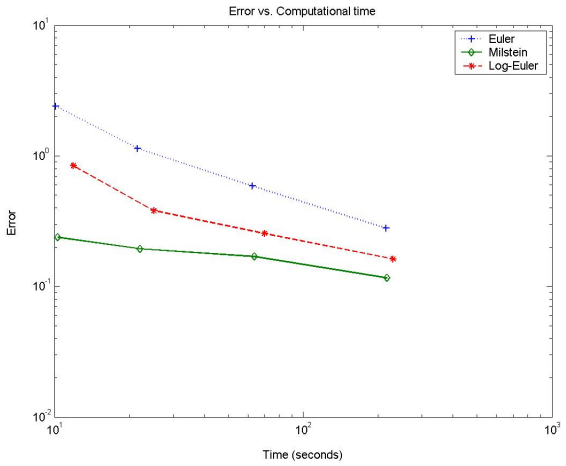


Figure 5.8: $F_k(0) = 0.06$, $\sigma_k = 0.2$, $m = -0.02$, tenor spacing $T = 0.5$, maturity times $T_k = 1, 2, 4, 6$, strikes $K = 0.04, 0.05, 0.06, 0.07, 0.08$, number of paths 2000000, Error $= ||C_{\text{exact}} - C_{MC}||_{2,K\times T_k}$ (grid points: strike $\times$ maturity times)

Comparing error and number of paths in the displaced diffusion model in the figure 5.9 we get accustomed results. As the number of paths increases the error stabilises. With 1000000 paths or more we obtain a rather good approximation.

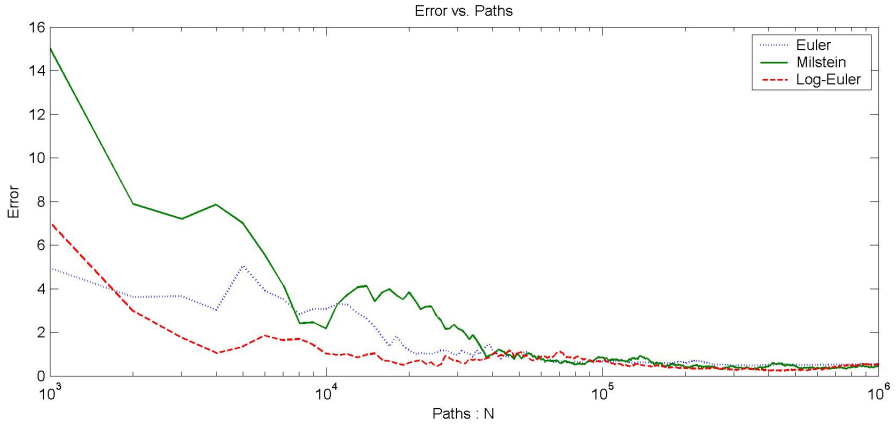


Figure 5.9: $F_k(0) = 0.06$, $\sigma_k = 0.2$, $m = -0.02$, tenor spacing $T = 0.5$, maturity times $T_k = 1, 2, 4, 6, 10$, strikes $K = 0.04, 0.05, 0.06, 0.07, 0.08$, integration steps $\Delta_t = 0.25$, Error $= ||C_{\text{exact}} - C_{MC}||_{2,K\times T_k}$ (grid points: strike $\times$ maturity times)

Conclusion DD

In the case of numerical integration in the displaced diffusion model the Milstein method is always superior to the Euler method. Firstly the computational costs are significantly lower. Secondly the Milstein method provides the better approximation to the exact caplet prices and additionally it preserves the analytical structure of the stochastic process. The log-Euler is problematic because we can use it only if $m < 0$.

5.2 Mean-reverting processes

The following section is quite independent from financial mathematics. The integration schemes are tested particularly with regard to the class of mean-reverting processes. Most important criteria is guaranteeing positivity. Let us start with a first example:

$$dX_t = (1 - X_t)dt + 1.4X_t^{\frac{1}{2}}dW \text{ with } X_0 = 1.$$

This stochastic differential equation is a stochastic process typically used to simulate stochastic volatility in the extended Libor market model (see [ABR01]).

		Euler		BIM		Milstein		imp. Milstein	
Time	Stepsize	Error	Negative	Error	Negative	Error	Negative	Error	Negative
$T = 1$	$\Delta_t = \frac{1}{2}$	0.2754	27.29 %	0.5187	0 %	0.2464	23.14 %	0.1503	**0 %**
	$\Delta_t = \frac{1}{4}$	0.1926	25.82 %	0.4339	0 %	0.1166	7.45 %	0.0677	**0 %**
	$\Delta_t = \frac{1}{8}$	0.1370	**21.59 %**	**0.3426**	0 %	0.0558	0.65 %	**0.0333**	**0 %**
$T = 2$	$\Delta_t = \frac{1}{2}$	0.3290	45.54 %	0.7118	0 %	0.2639	34.72 %	0.2080	**0 %**
	$\Delta_t = \frac{1}{4}$	0.2241	43.39 %	0.5832	0 %	0.1269	12.12 %	0.0849	**0 %**
	$\Delta_t = \frac{1}{8}$	0.1563	**38.88 %**	**0.4499**	0 %	0.0607	1.18 %	**0.0397**	**0 %**
$T = 4$	$\Delta_t = \frac{1}{2}$	0.3435	69.18 %	1.2734	0 %	0.2767	53.12 %	0.3174	**0 %**
	$\Delta_t = \frac{1}{4}$	0.2333	67.21 %	1.0188	0 %	0.1305	20.63 %	0.1019	**0 %**
	$\Delta_t = \frac{1}{8}$	0.1610	**62.44 %**	**0.7415**	0 %	0.0633	2.22 %	**0.0435**	**0 %**

Table 5.5: Time: [0,T], stepsize: integration stepsize Δ_t, Error: integration error compared with implicit Milstein ($\Delta_t = \frac{1}{4096}$), Negative: percentage of negative paths, BIM: $c_0(x) = 1$ and $c_1(x) = 1.4x^{-\frac{1}{2}}$.

The results of the test in table 5.5 clearly show that the implicit Milstein method is superior to the other methods concerning convergence speed as

well as remaining positivity.
As expected the Euler scheme cannot carry forward the analytical positivity to numerics. Because of the convergence order of $\gamma_1 = 0.5$ the convergence speed is not as high as the one of the Milstein methods.
The BIM can indeed remain positivity. Admittedly the error is even bigger than in the Euler method because an extensive use of the control functions c_0 and c_1 is necessary.

The parameter constellation of the next example

$$dX_t = (1 - X_t)dt + 1.4X_t dW \text{ with } X_0 = 1$$

is slightly different as in the example above.

		Euler		BIM		Milstein		imp. Milstein	
Time	Stepsize	Error	Negative	Error	Negative	Error	Negative	Error	Negative
$T = 1$	$\Delta_t = \frac{1}{2}$	0.5250	28.77 %	0.6130	0 %	0.4484	10.35 %	0.2427	**0 %**
	$\Delta_t = \frac{1}{4}$	0.3997	25.17 %	0.5046	0 %	0.2291	0 %	0.1266	**0 %**
	$\Delta_t = \frac{1}{8}$	0.2767	**13.76 %**	**0.3585**	0 %	0.1031	0 %	**0.0716**	**0 %**
$T = 2$	$\Delta_t = \frac{1}{2}$	0.8021	48.89 %	0.8930	0 %	0.5533	24.71 %	0.3505	**0 %**
	$\Delta_t = \frac{1}{4}$	0.5211	42.95 %	0.6912	0 %	0.2732	0 %	0.1500	**0 %**
	$\Delta_t = \frac{1}{8}$	0.3301	**24.30 %**	**0.4415**	0 %	0.1120	0 %	**0.0819**	**0 %**
$T = 4$	$\Delta_t = \frac{1}{2}$	1.1377	74.81 %	1.9005	0 %	0.7252	47.74 %	0.5270	**0 %**
	$\Delta_t = \frac{1}{4}$	0.6278	66.63 %	1.2736	0 %	0.2743	0 %	0.1537	**0 %**
	$\Delta_t = \frac{1}{8}$	0.3602	**41.55 %**	**0.5903**	0 %	0.1163	0 %	**0.0852**	**0 %**

Table 5.6: Time: [0,T], stepsize: integration stepsize Δ_t, error: integration error compared with implicit Milstein ($\Delta_t = \frac{1}{4096}$), Negative: percentage of negative paths, BIM: $c_0(x) = 1$ and $c_1(x) = 1.4$.

The results in this test are quite similar to the results of the first example. The question arises if the better integration behaviour of the implicit Milstein method negatively influences the computational time.
Measuring the computational effort we notice that Euler and Milstein method take roughly the same time to simulate one path. In contrast the computational time the BIM takes is circa three times as high because more calculation is needed.
Thus the computational effort is no argument against the implicit Milstein method. To make that perfectly clear figure 5.10 shows the relation between error and computational time.

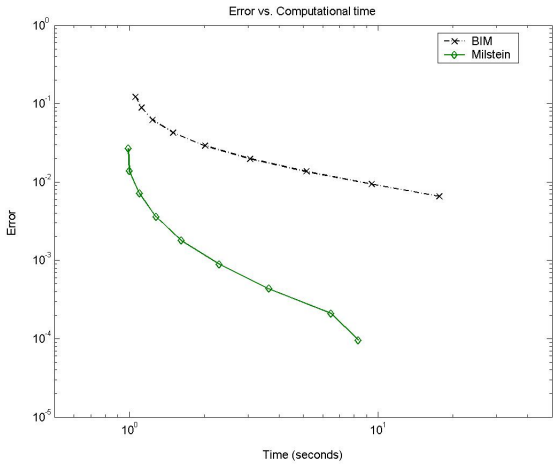


Figure 5.10: $dX_t = (1 - X_t)dt + 1.4\sqrt{X_t}dW$, path wise error,100000 paths, BIM: $c_0 = 1$ and $c_1(x) = 1.4x^{-\frac{1}{2}}$

To conclude there is a last example, table 5.7 below. The chosen parameter constellation shows even clearer the disadvantages of Euler method and BIM.

$$dX_t = (0.125 - 4X_t)dt + 0.5\sqrt{X_t}dW \text{ with } X_0 = 0.1$$

Δ_t	E-$\mathcal{S}$	E-$\mathcal{N}$	BIM-$\mathcal{S}$	BIM-$\mathcal{N}$	M-$\mathcal{S}$	M-$\mathcal{N}$	iM-$\mathcal{S}$	iM-$\mathcal{N}$
2^{-3}	0.0103349	67.62 %	0.0136404	0 %	0.0089764	53.56 %	0.0110619	0 %
2^{-4}	0.0071670	59.51 %	0.0110236	0 %	0.0043066	18.62 %	0.0033908	0 %
2^{-5}	0.0 050365	52.93 %	0.0088861	0 %	0.0020941	0.03 %	0.0014815	0 %
2^{-6}	0.0035999	46.80 %	0.0072702	0 %	0.0010266	0.01 %	0.0007614	0 %
2^{-7}	0.0025479	41.77 %	0.0057562	0 %	0.0005148	0 %	0.0003927	0 %
2^{-8}	0.0018158	37.28 %	0.0045040	0 %	0.0002616	0 %	0.0002021	0 %
2^{-9}	0.0012823	33.76 %	0.0033995	0 %	0.0001339	0 %	0.0001034	0 %
2^{-10}	0.0009086	30.89 %	0.0026244	0 %	6.8521e-05	0 %	5.2497e-05	0 %
2^{-11}	0.0006399	28.30 %	0.0019556	0 %	3.5276e-05	0 %	2.6341e-05	0 %
2^{-12}	0.0004512	26.10 %	0.0014425	0 %	1.8552e-05	0 %	1.2706e-05	0 %

Table 5.7: Δ_t : integration stepsize, E-$\mathcal{S}$: strong approximation error Euler scheme, E-$\mathcal{N}$: percentage of negative paths Euler scheme, BIM: Balanced implicit method ($c_0(x) = 4$ and $c_1(x) = 0.5x^{-\frac{1}{2}}$), M: Milstein scheme, iM: implicit Milstein

Also with a very small discretisation stepsize the Euler method cannot guarantee positivity. Additionally with decreasing stepsize the higher convergence order of the Milstein scheme becomes more and more profitable compared to Euler and BIM.

Conclusion Mean-reverting

The examples make clear that the appropriate choice of an integration scheme is essential for the right numerical calculation of a mean-reverting process. Hereby the convergence order is as important as the non-negativity. The implicit Milstein method can meet both requirements. Furthermore it is nearly at optimal computational cost. For this reason the implicit Milstein scheme is the method of choice to integrate a mean-reverting process.

CHAPTER VI

Summary

This diploma thesis analysed the topic of positive numerical integration of stochastic differential equations, a topic which is not only of theoretical interest as there are a lot of applications. In this particular case the motivation comes from pricing interest rates derivatives as a task in financial mathematics.

We considered the Libor market model and different extensions as a possibility to model interest rates by forwards rates. Naturally the forward rates are modelled by stochastic differential equations. Hence we had to answer two question:

- When is a stochastic process analytically positive?
  (Analytical positivity)
- How can analytical positivity preserved numerically?
  (Numerical positivity)

We answered both questions in the framework of the Libor market model in this thesis.

Stochastic processes describing the forward rates can be classified with respect to analytical positivity by using diffusion theory. Following a classical approach enables us to characterise the analytical behaviour by basic calculations.

But this thesis goes one step further. As the extended Libor market model does not provide closed solutions to price interest rate derivatives we need numerical approximations. To be able to tackle the problem of numerical

positivity we studied different numerical integration schemes. Combining the BIM with the Milstein method we developed the balanced Milstein method (BMM).

The analysis of positive numerical integration is the very heart of this thesis. After defining numerical positivity we studied different integration methods especially with regard to this characteristic. As already known, the BIM preserves the positivity of some stochastic processes. But within this thesis we went another way. Analysing the Milstein method we obtained the surprising result that this method maintains positivity in a natural way.

Numerical tests concluded our analysis of numerical positivity. On the one hand we tested different integration methods on their applicability in financial applications. It turned out that preserving positivity is essential for an accurate approximation. In the case of the forward rates simulation the Milstein method is superior. On the other hand we considered mean-reverting processes. In this case the implicit Milstein scheme is the method of choice. Indeed both results are not surprising as preserving of geometric properties leads to an improved convergence speed.

This thesis exhibits a first step to numerical analysis of Stochastic Differential Algebraic Equations (SDAEs). But analytical positivity may only be one geometric aspect of a stochastic process which should be preserved by numerical approximation. The application in financial mathematics makes perfectly clear that we have to pay attention to the analytical behaviour of a stochastic process to get a proper numerical approximation. Thus the choice of the right integration scheme is one essential part of the simulation.

After all it is a moot question whether it is possible to guarantee numerical positivity for every positive stochastic process. Maybe this problem could be solved by using integration schemes of higher order, stochastic Runge-Kutta methods or multi-step schemes.

APPENDIX A

Stochastic analysis

The stochastic analysis is the mathematical base for a sensible modelling of financial derivatives. The majority of results will be given without a proof. Corresponding elucidations are in [Bau92, HT94, KS88, KT74, KT81, Øks00].

Probability theory and stochastic processes

Definition A.1 Let Ω be given and $\mathcal{F} \subset \Omega$, then we call $\mathcal{F}$ a σ-algebra if the following properties hold

(i) $\emptyset \in \mathcal{F}$,

(ii) $F \in \mathcal{F} \Rightarrow F^c \in \mathcal{F}$ with $F^c = \Omega \setminus F$,

(iii) $F_1, F_2, ... \in \mathcal{F} \Rightarrow \bigcup_{i \in \mathbb{N}} F_i \in \mathcal{F}$.

For an arbitrary family of subsets $\mathcal{U} \subset \Omega$ a smallest σ-algebra $\mathcal{H}_\mathcal{U}$ exists which includes all $\mathcal{U}$

$$\mathcal{H}_\mathcal{U} = \bigcap \{\mathcal{H} : \mathcal{H} \ \sigma\text{-algebra with } \mathcal{U} \subset \mathcal{H}\}.$$

This is the σ-algebra generated by $\mathcal{U}$.

Definition A.2 The open subsets $U \subset \Omega$ generate a well known σ-algebra, the Borel-algebra.

We call $(\Omega, \mathcal{F})$ a measurable space. A probability measure on a measurable space is a function

$$P : \mathcal{F} \longrightarrow [0, 1],$$

so that the following holds:

1. $P(\Omega) = 1$,
2. let $F_1, F_2, ... \in \mathcal{F}$ be a countable number of disjoint sets, then

$$P\left(\bigcup_i F_i\right) = \sum_i P(F_i).$$

The triple $(\Omega, \mathcal{F}, P)$ is called a probability space.

We can interpret each element $F \in \mathcal{F}$ as one particular event. Hence the probability measure P associates each event with a value in $[0, 1]$. This value can be understood as the probability of that event.

Random variable is the next important concept.

Definition A.3 Let $(\Omega, \mathcal{F}, P)$ be a probability space.

$$X : \Omega \longrightarrow \mathbb{R}^n$$

is measurable with respect to $\mathcal{F}$, if and only if

$$X^{-1}(U) = \{\omega \in \Omega : X(\omega) \in U\} \in \mathcal{F}$$

holds for all Borel-sets U. X is called random variable.

In a natural way every random variable induces a probability measure μ_X on $\mathbb{R}^n$ by

$$\mu_X(B) = P(X^{-1}(B)) \text{ , } B \text{ Borel-set.}$$

Thus μ_X is called the distribution of X.

Having two different measures on one single measurable space we could ask if it is possible to transform one measures into the other. The Radon-Nikodym theorem answers this question.

Theorem A.4 (Radon-Nikodym) Let P and Q be probability measures on the σ-algebra $(\Omega, \mathcal{F})$ and let

$$P(B) = 0 \Longleftrightarrow Q(B) = 0,$$

for every $B \in \mathcal{F}$. Then a non-negative random variable Z exists, such that

$$Q(A) = \int_A Z(\omega) dP(\omega) \text{ for all } A \in \mathcal{F}. \tag{A.1}$$

Z is called the Radon-Nikodym derivative from Q with respect to P.

The density p of a random variable is the unique Radon-Nikodym derivative of the induced distribution μ_X of X w.r.t. the Lebesgue measure

$$\mu_X(B) = \int_B p(x) dx.$$

Definition A.5 Let X be a random variable on a probability space $(\Omega, \mathcal{F}, P)$. Then the expectation of X is defined by

$$E[X] = \int_\Omega X(\omega) dP(\omega) = \int_{\mathbb{R}^n} x d\mu(x) = \int_{\mathbb{R}^n} x p(x) dx. \tag{A.2}$$

The expectation plays an important role in pricing financial derivatives. It further holds for an arbitrary function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$

$$E[f(X)] = \int_\Omega f(X(\omega)) dP(\omega) = \int_{\mathbb{R}^n} f(x) d\mu(x) = \int_{\mathbb{R}^n} f(x) p(x) dx.$$

The variance V of a random variable is as important as the expectation:

$$V(X) = E\left[(X - E[X])^2\right] = E\left[X^2\right] - E[X]^2.$$

Therewith we have accumulated all required basics of probability theory. Before analysing stochastic processes we study one example to illustrate the concept of density.

Definition A.6 A normal distributed random variable $X \sim \mathcal{N}(\mu, \sigma^2)$ on $\mathbb{R}$ possesses the density

$$p(x) = (2\pi\sigma^2)^{-1/2} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

Hereby μ is the expectation and σ^2 the variance (see fig. A.1).

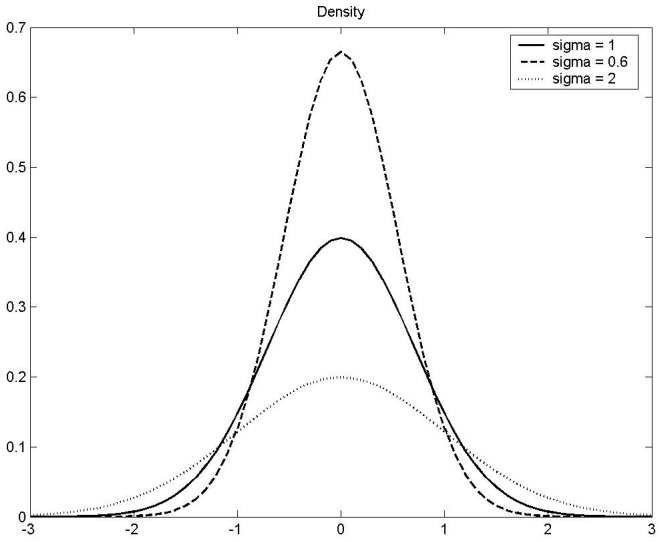


Figure A.1: Normal distribution: $X \sim \mathcal{N}(0, \sigma^2)$

Definition A.7 A stochastic process is a quadruple $(\Omega, \mathcal{F}, P, (X_t)_{t\in I})$ where $(\Omega, \mathcal{F}, P)$ is a probability space and $(X_t)_{t\in I}$ is a family of random variables

$$X_t : \Omega \to \mathbb{R}^n \tag{A.3}$$

with I being $[0, \infty)$ in this thesis.

There are two ways of interpreting a stochastic process. On the one hand the mapping

$$\omega \to X_t(\omega)$$

is a random variable for fixed t . On the other hand for fixed ω the stochastic process

$$t \to X_t(\omega)$$

defines a mapping from $\mathbb{R}$ to $\mathbb{R}^n$. This map is called the path of a stochastic process.

To adapt the concept of measurability to stochastic processes we have to add a filtration to $\mathcal{F}$.

Definition A.8 A filtration $\{\mathcal{F}_t\}_{t\geq 0}$ is a non-descending sequence of sub σ-algebras such that

$$\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F} \text{ for } s \leq t.$$

Definition A.9 Let $\mathcal{F}_t$ be a filtration. Then a stochastic process X_t is called adapted by $\mathcal{F}_t$, if X_t is measurable w.r.t. $\mathcal{F}_t$ for all $t \in I$.

The measurability of stochastic processes becomes clear in context with conditional expectation.

But we can also see it from a different point of view. Then a stochastic process naturally induces a σ-algebra.

Definition A.10 Let X_t be a stochastic process. Then the σ-algebra $\mathcal{F}_T$ generated by $(X_t)_{t\in I}$ is the smallest σ-algebra containing all sets of the form

$$\{\omega : X_{t_1}(\omega) \in F_1, ..., X_{t_n}(\omega) \in F_n\}$$

where $0 \leq t_j \leq T$, $j \leq n = 1, 2, ...$ and $F_j \subset \mathbb{R}^n$ Borel-sets.

Actually this requires the definition of a product σ-algebra on $\tilde{\Omega} = (\mathbb{R}^n)^{[0,T]}$. The construction is rather complicated but the reader can find a detailed description in [Par72].

Accordingly a stochastic process defines a filtration as

$$\mathcal{F}_S \subseteq \mathcal{F}_T \text{ with } 0 \leq S \leq T.$$

It is also clear that every stochastic process X_t is adapted to its own filtration $\mathcal{F}_t$.

Another essential concept is the conditional expectation.

Definition A.11 (Conditional expectation) Let $H \in F$ be a σ-algebra. Then $E[X|H]$ is the unique H measurable function with

$$\int_G E[X|H]\, dx = \int_G X dx \text{ for all } G \in H. \tag{A.4}$$

The proof of uniqueness directly follows from the Radon-Nikodym theorem (see equation A.1).

There is an easy interpretation of this abstract definition. For that reason the σ-algebra H is a kind of information about the stochastic process which will be explained by the following statements.

Theorem A.12 Let X and Y be random variables and H a σ-algebra. Then

1. $E[X+Y|H] = E[X|H] + E[Y|H]$,
2. $E[X|H] = X$ if X is measurable w.r.t. H,
3. $E[X|H] = E[X]$ if X is independent of H.

Proof: The proof can be found in [Øks00]. For detailed description about conditional expectation see [Bau92]. □

Definition A.13 (Martingale) Let X_t be an F_t adapted stochastic process. Then X_t is called a martingale if

$$E[X_t|F_s] = X_s \text{ for all } s \leq t. \tag{A.5}$$

The martingale property of a stochastic process means that the process possesses a neutral behaviour concerning future events. Thus we can only state as much about the process as there is information in the σ-algebra F_s.

Brownian motion and Ito-Calculus

One important stochastic process is the Brownian motion.

Definition A.14 Let B_t be an $\mathcal{F}_t$ adapted process. Then we call B_t a Brownian motion if the following holds for all $0 \leq s < t$:

(i) $B_0 = 0$,

(ii) $B_t - B_s$ is independent of F_s,

(iii) $B_t - B_s$ is $\mathcal{N}(0, t-s)$ distributed.

To get a clear picture of a Brownian motion figure A.2 illustrates the density in the course of time t.

Corollary A.15 The Brownian motion B_t is a martingale with respect to its own filtration $\mathcal{F}_t$.

Proof: A simple calculation leads to

$$\begin{aligned} E[B_t|\mathcal{F}_s] &= E[B_s + B_t - B_s|\mathcal{F}_s] \\ &= E[B_s|\mathcal{F}_s] + E[B_t - B_s|\mathcal{F}_s] = B_s. \end{aligned}$$

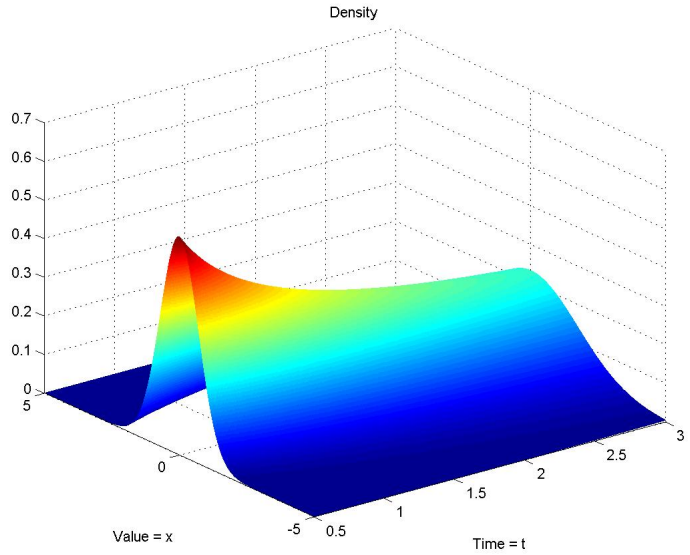


Figure A.2: Brownian motion

The Brownian motion is a basic stochastic process. This raises the question if it is possible to define more complex stochastic processes based on a Brownian motion. Therefore we must regard stochastic processes from another point of view, namely as the solution of stochastic differential equations.

Definition A.16 (Stochastic differential equation (SDE)) Let $T > 0$, $a : [0, T] \times \mathbb{R} \to \mathbb{R}$ and $b : [0, T] \times \mathbb{R} \to \mathbb{R}$. Then we call the following equation a stochastic differential equation

$$dX_t = a(t, X_t)dt + b(t, X_t)dB_t \text{ with } X_{t_0} = x_0. \tag{A.6}$$

The function a is the drift and b is the diffusion.

Remark A.17 If we define a stochastic process by an SDE then the process starts at time t_0 in x_0. At $t > t_0$ we can determine the density of the process in x. This is called the transition density from (t_0, x_0) to (t, x)

$$p(t_0, t, x_0, x)dx = P(X_t \in dx) = P(x < X_t < x + dx). \tag{A.7}$$

For calculating the expectation we also have to consider where the process starts:

$$E^{t_0,x_0}[X_t] = E[X|X_{t_0} = x_0]. \tag{A.8}$$

Two special cases of stochastic differential equations are already known. Firstly with $b = 0$ it is an ordinary differential equation. Secondly with $a = 0$ and $b = 1$ we get the SDE $dX_t = dB_t$ with the obvious solution $X_t = B_t$.

Here a questions comes up: How can we deal with the term dB_t? To handle this problem it is necessary to write the stochastic differential equation as an integral equation and to define the Gaussian measure dB_s

$$X_t = X_0 + \int_0^t a(s, X_s)ds + \int_0^t b(s, X_s)dB_s.$$

The integral

$$\int_0^t b(s, X_s)dB_s$$

is a so-called stochastic integral. The idea of constructing such an integral is similar to the construction of the Riemann-Stieltjes integral. It is not possible to explain the whole theory of Ito integration here. Nevertheless we will recapitulate some main results in short. For further reading see [Arn73, Bog98, KS88, Øks00].

First of all we have to mention some facts about the Brownian motion. To simplify the notation let $J = [0, T]$ be an interval. Further let $P(J)_N$ be a partition of this interval

$$P(J)_N = 0 = t_0 < t_1 < ... < t_N = T$$

and $\mathcal{P}_N$ is the set of all partitions of length N.

Proposition A.18 The following statements hold for the Brownian motion:

(A) the Brownian motion is not mean-square integrable,

(B) the paths are not differentiable w.r.t. P a. s.,

(C) the process possesses an unlimited variance on each bounded interval $J = [0, T]$

$$\sup_N \sup_{P(J)_N \in \mathcal{P}_N} \sum_{j=0}^{N-1} |B_{t_{j+1}} - B_{t_j}| = \infty.$$

Proof: Verifying (A) needs only basic calculation

$$\lim_{h\to 0} E\left[\left(\frac{B_{t+h}-B_t}{h}\right)^2\right] = \lim_{h\to 0}\frac{1}{h} = \infty.$$

The proof of (B) and (C) can be found in literature. □

Thus it is not possible to define the stochastic integral as an ordinary Riemann-Stieltjes integral

$$\int_0^T b(s,X_s)dB_s = \lim_{N\to\infty} \sum_{\substack{P(J)_N\in\mathcal{P}_N \\ \xi\in[t_i,t_{i+1}]}} b(\xi,X_\xi)(B_{t_{i+1}}-B_{t_i}).$$

To simplify the notation we define further:

$$S_N := \sum_{\substack{P(J)_N\in\mathcal{P}_N \\ \xi\in[t_i,t_{i+1}]}} b(\xi,X_\xi)(B_{t_{i+1}}-B_{t_i}).$$

The mean-square convergence is denoted by

$$L^2 - \lim_{N\to\infty} X_N := \lim_{N\to\infty} E\left[(X_N)^2\right].$$

The mean-square integrability is exactly an appropriate definition for the stochastic integral. But there is still the problem that the integral strongly depends on the choice of $\xi \in [t_i, t_{i+1}]$. In the case of Riemann-Stieltjes we can take any value in the interval, here we have to fix the evaluation point ξ to get a well defined integral. The following two choices have turned out to be the most appropriate ones:

Definition A.19 (Ito integral) With $\xi = t_j$ (left point) we get the Ito integral.

Definition A.20 (Stratonovich integral) Choosing $\xi = \frac{t_j+t_{j+1}}{2}$ (center) leads to the Stratonovich integral.

This enables us to define a stochastic integral as follows:

$$\int_0^T b(s,X_s)dB_s = L^2 - \lim_{N\to\infty} \sum_{\substack{P(I)_N\in\mathcal{P}_N \\ \xi\in[t_i,t_{i+1}]}} b(\xi,X_\xi)(B_{t_{i+1}}-B_{t_i}). \tag{A.9}$$

It is not difficult to transform the Ito integral into the Stratonovich and vice-versa. One important characteristic of the Ito integral is the martingale property (see A.13 below). The big advantage of the Stratonovich integral is that differentiating can be done in the usual way. Obviously not all functions are integrable w.r.t dB_s.

Definition A.21 The set of all Ito integrable functions is denoted by $\mathcal{V} = \mathcal{V}(0,T)$. One element

$$f(t,\omega) : [0,\infty) \times \Omega \to \mathbb{R} \tag{A.10}$$

is characterised by the following properties:

- $(t,\omega) \to f(t,\omega)$ is measurable w.r.t. $B \times \mathcal{F}$ with B the Borel sets on $[0,\infty)$.
- $f(t,\omega)$ is adapted by $\mathcal{F}_t$.
- $E\left[\int_0^T f(t,\omega)^2 dt\right] < \infty$.

Definition A.22 (Ito process) An Ito process is the stochastic process which solves the stochastic differential equation

$$dX_t = a(t,X_t)dt + b(t,X_t)dB_t \text{ with } X_{t_0} = x_0. \tag{A.11}$$

The existence of such a process requires $b(t,\omega) \in \mathcal{V}$ and

$$P\left(\int_0^t |a(s,\omega)|ds < \infty \text{ for all } t \geq 0\right) = 1.$$

First we have a closer look at the properties of an Ito process.

Lemma A.23 (Ito isometry) Let $f : [0,\infty) \times \Omega :\to \mathbb{R} \in \mathcal{V}$. Then

$$E\left[\left(\int_0^T f(t,\omega)dB_t\right)^2\right] = E\left[\int_0^T f^2(t,\omega)dt\right]. \tag{A.12}$$

In consideration of the Ito theorem (see A.25 below) the symbolic calculation rule $(dB_t)^2 = dt$ seems sensible. In financial mathematics it is of great importance that the Ito integral is a martingale.

Lemma A.24 Let $f(t,\omega) \in \mathcal{V}$. Then the stochastic process M_t defined by

$$M_t = \int_0^t f(s,\omega) dB_s \tag{A.13}$$

is a martingale with respect to the filtration induced by M_t. So we get

$$E\left[\int_0^t f(s,\omega) dB_s\right] = M_0. \tag{A.14}$$

This behaviour can be condensed in the sentence:

An Ito integral does not anticipate future.

Although from a theoretical point of view these results are quite decisive they cannot help how to construct a new stochastic process from a given one. The fundamental theorem in Ito calculus helps.

Theorem A.25 (Ito's theorem) Let X_t be an Ito process and $g(t,x) \in C^{1\times 2}([0,\infty)\times\mathbb{R})$. Then

$$Y_t = g(t, X_t)$$

is again Ito process, and

$$dg(t,X_t) = \frac{\partial g}{\partial t}(t,X_t)dt + \frac{\partial g}{\partial x}(t,X_t)dX_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t,X_t)(dX_t)^2 \tag{A.15}$$

holds with $(dt)^2 = dtdB = 0$ and $(dB)^2 = dt$.

Example A.26 Using Ito's theorem enables us to calculate the integral

$$\int_0^t B_s dB_s.$$

Define $g(t,x) = x^2$ and apply Ito's theorem on $X_t = B_t$ we easily verify

$$\begin{aligned} & d(B_t)^2 && = 2B_t dB_t + (dB_t)^2 \\ \iff & \int_0^t B_s dB_s && = \frac{1}{2}\left((B_t)^2 - t\right). \end{aligned}$$

There are only a few examples where the Ito integral has a closed solution. One of these processes which are as elementary as important in financial mathematics is the geometric Brownian motion as the solution of the following stochastic differential equation

$$dX_t = \alpha X_t dt + \beta X_t dB_t \text{ with } X_0 = x. \tag{A.16}$$

Applying Ito's theorem verifies the solution as

$$X_t = x \exp\left(\left(\alpha - \frac{\beta^2}{2}\right) t + \beta B_t\right). \tag{A.17}$$

Starting with the density of the Brownian motion B_t we are able to calculate the density of the geometric Brownian motion:

$$P(B_t - B_{t_0} \in db) = \frac{1}{\sqrt{2\pi(t-t_0)}} \exp\left(-\frac{b^2}{2(t-t_0)}\right) db,\ t > t_0.$$

Now we choose a proper transformation according to the closed solution with $X_{t_0} = x$:

$$\begin{aligned} & y &&= x \exp\left(\left(\alpha - \frac{\beta^2}{2}\right)(t-t_0) + \beta b\right) \\ \iff\ & b &&= \frac{1}{\beta}\left[\log\left(\frac{y}{x}\right) - \left(\alpha - \frac{\beta^2}{2}\right)(t-t_0)\right]. \end{aligned}$$

Obviously we get

$$\frac{dy}{db} = \beta y \iff db = \frac{dy}{\beta y}. \tag{A.18}$$

After these preliminaries we only need to put the different pieces together:

$$\begin{aligned} P(X_t \in dy) \;=\; & \frac{1}{\beta y \sqrt{2\pi(t-t_0)}} \cdot \\ & \exp\left(-\frac{1}{2\beta^2(t-t_0)}\left[\log\left(\frac{y}{x}\right) - \left(\alpha - \frac{\beta^2}{2}\right)(t-t_0)\right]^2\right) dy. \end{aligned}$$

APPENDIX B

Equity market

The modelling of stock prices is certainly the most natural connection between stochastic analysis and financial mathematics. Looking at a stock chart (see fig. B.1) we recognise that it is unpredictable to a certain extent.

Figure B.1: *DAX* chart from 2001/05/06-2004/05/06.

That is exactly the phenomenon we want to describe with stochastic analysis. Actually it is rather the target to obtain the fair value of a stock option than predicting future. We start with the most simple kind of stock option: the call.

Definition B.1 (Call) A European call is a contract that allows the buyer to acquire a particular share at a previously fixed price K (strike) on a fixed date T in future.

Putting this definition into mathematical formulas leads to the following payoff function at $t = T$ (see fig. B.2)

$$C(T) = \max(S(T) - K, 0) = (S(T) - K)^+.$$

Here S_T is the share price at time $t = T$.

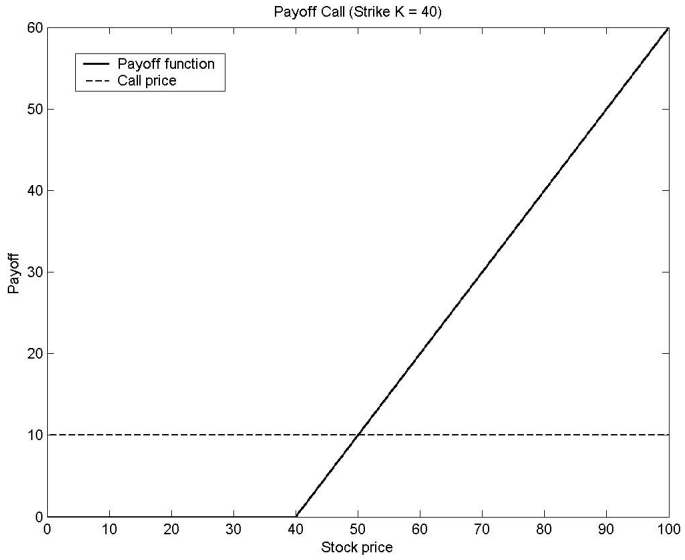


Figure B.2: Payoff function of a European call.

This provides the option price $C(t)$ at T. But what is the fair price at $t \in [0, T)$? To answer this question first of all we have to define what "fair" means in this context. The term "arbitrage" answers this question.

Definition B.2 (Arbitrage) Arbitrage is the opportunity of drawing immediately a riskless profit. When no instant riskless profit is possible we call the market arbitrage-free.

The assumption of arbitrage-free markets is one of the most basic assumptions we have in the pricing of derivatives.

A financial investment at a fixed interest rate without risk of default is the easiest way of investing risklessly. Certainly this is no arbitrage as the profit is not immediately. Therefore the discount rate is given by

$$D(t,T) = \exp\left(-\int_t^T r(\tau)d\tau\right). \tag{B.1}$$

Assuming we describe a stock with a stochastic process S_t with a starting price $S_{t_0} = s_0$, this stochastic process will be specified later on, the discounted conditional expectation gives us the fair price of a European call

$$\begin{aligned} C(t,S) &= E^{t_0,s_0}\left[D(t,T)(S_T-K)^+|\mathcal{F}_t\right] \\ &= E^{t,S(t)}\left[D(t,T)(S_T-K)^+\right]. \end{aligned}$$

So the option price agrees with the discounted profit expectation of the share. This discounting makes our model arbitrage free.

Remark B.3 In general we can define any payoff function and get

$$C(t,S) = E^{t,S(t)}\big[D(t,T)f(S_T)\big]. \tag{B.2}$$

Therewith we can formulate the following options:

$$f(x) = \begin{cases} (K-S_T)^+ & : \text{Put}, \\ (K-S_T)^+ + (S_T-K)^+ & : \text{Straddle}, \\ (K_1-S_T)^+ + (S_T-K_2)^+ & : \text{Strangle}. \end{cases}$$

These are only a few possibilities. Actually there are multitudes of options. In fact we can give our fancy full scope in constructing new options. See also [GJ03, Nel96, Wil00a, Wil00b]. All options presented above have in common that their payoff functions f depend on $t = T$. Needless to say that options exist whose values depend on the stock prices $S(t)$ with $0 \le t \le T$, e. g. an Asian option

$$C(T,S) = \left(\frac{1}{T}\int_0^T S_\tau d\tau - K\right)^+.$$

We still have to face the question which stochastic process is capable of modelling the stock chart. Here in particular it makes sense that the stochastic process takes only positive values. Hence a geometric Brownian motion, see also fig. B.3, seems to be appropriate. We have already proved the non-negativity in Lemma 1.16.

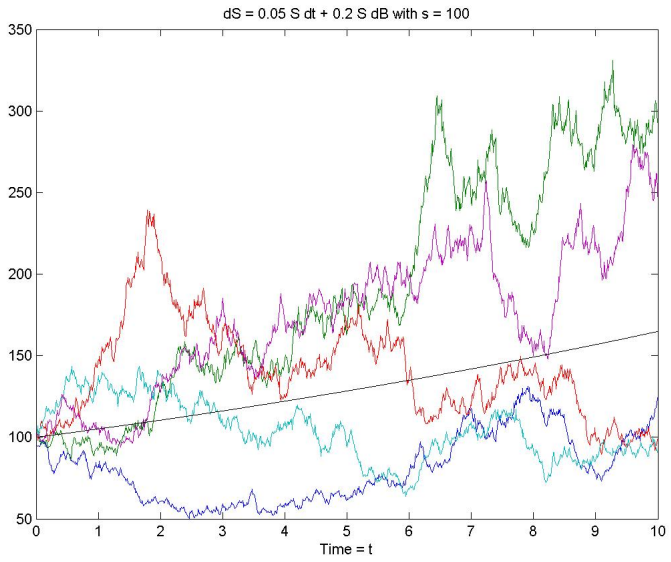


Figure B.3: Paths of the geometric Brownian motion.

Example B.4 (Black-Scholes) Let the stock prices be given by the following stochastic differential equation

$$dS_t = \mu S_t dt + \sigma S_t dB_t \text{ with } S_{t_0} = s_0. \tag{B.3}$$

One can construct the following self-financing portfolio Π consisting of shares and options.

$$\Pi = \Delta S - C(t, S).$$

An elaborate explanation about self-financing portfolios can be found in [KK01]. Initially Δ is a non-specified factor. Later on it becomes clear why it is often called the hedge factor.
The next equation is not obvious. An exact derivation can be found in the relevant literature

$$d\Pi = \Delta dS - dC(t, S).$$

Assuming an arbitrage-free market, the constructed portfolio is supposed to gain as much profit as a riskless investment at rate r

$$d\Pi = r\Pi dt.$$

This allows to equate the first differential equation to the second. We get

$$r\Pi dt = \Delta dS - dC(t, S).$$

As the option price depends on time t as well as on the stock price S, we can calculate the differential dC using Ito's formula. This leads to:

$$r\Pi dt = \Delta dS - \left(\frac{\partial C}{\partial t} dt + \frac{\partial C}{\partial S} dS + \frac{1}{2} \frac{\partial^2 C}{\partial S^2} (dS)^2 \right).$$

Employing dS, Π and $(dS)^2 = \sigma^2 S^2 dt$ we get

$$r(\Delta S - C)dt = \left(\Delta - \frac{\partial C}{\partial S} \right) dS - \left(\frac{\partial C}{dt} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} \right) dt.$$

There is only one non-deterministic term left. But we can choose Δ so that we eliminate the risk completely:

$$\Delta = \frac{\partial V}{\partial S}. \tag{B.4}$$

With this particular Δ we achieve the Black-Scholes differential equation:

$$-\frac{\partial C}{\partial t} = rS \frac{\partial C}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} - rC. \tag{B.5}$$

The bottom line is that we can reach a closed solution of this equation if we choose the geometric Brownian motion to model the stock chart. Transforming the Black-Scholes differential equation into heat conduction equation is one possible way of proceeding. In this case we use the corresponding argumentation from the theory of partial differential equations.
But remembering the Feynman-Kac theorem provides a more elegant alternative. Using this theorem we see that the equation above is equivalent to

$$C(t, S) = E^{t, S_t} \left[\exp \left(- \int_t^T r d\tau \right) (S_T - K)^+ \right] \tag{B.6}$$

because the geometric Brownian motion has the appropriate generator. Indeed a more intuitive point of view leads to the same result, see equation (B.2). As we want to get an explicit formula we call back to mind the density of the geometric Brownian motion:

$$p(t, T, x, y) = \frac{1}{\sigma y \sqrt{2\pi}} \exp \left(- \frac{1}{2\sigma^2 (T - t)} \left[\log \left(\frac{y}{x} \right) - \left(r - \frac{\sigma^2}{2} \right) (T - t) \right]^2 \right).$$

Additionally, we need the following notation

$$d(t)_{1/2} = \frac{\log(\frac{S_t}{K}) + (r \pm \sigma^2/2)(T-t)}{\sigma\sqrt{(T-t)}}$$

and

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp(-\frac{y^2}{2}) dy. \tag{B.7}$$

Having in mind the above we can examine the option price with only little effort as a short calculation shows

$$\begin{aligned} C(t,S) &= \exp\big(-r(T-t)\big) E^{t_0,s_0}\left[(S_T-K)^+ | \mathcal{F}_t\right] \\ &= D(t,T) E^{t,S_t}\left[(S_T-K)^+\right] \\ &= D(t,T) \int_{\mathbb{R}} (y-K)^+ p(t,T,S_t,y) dy \\ &= S_t \Phi(d(t)_1) - D(t,T) K \Phi(d(t)_2). \end{aligned}$$

Recapitulating the Black-Scholes formula leads to a closed solution for the fair price of a call if we assume that the stock prices behave like a geometric Brownian motion. But does this calculated price fit with the market prices? There are still two questions: how can we compare prices and (if they do not fit) how can we extend the model so that it replicates the market prices?

Implied volatility and model extension

In the Black-Scholes model there are some parameters we can directly get from the market. That is current price, duration, strike and to a certain extent interest rate as well. But how can we determine the volatility of a share? One approach is to refer to historical volatility.

Definition B.5 The historical volatility is defined by

$$\sigma_{\text{hist}} = \sqrt{N} \left(\frac{1}{N-1} \sum_{i=1}^{N} (y_i - \bar{y})^2 \right)^{1/2}. \tag{B.8}$$

In this case N is the number of days, $y_i = \ln(S_{i+1}) - \ln(S_i)$ where S_i describes the stock price at time i and $\bar{y}$ denominates the average over y_i.

On the other hand the market provides call prices. Thus we can use the adapted Black-Scholes formula to calculate the volatility the market expects. That is why this kind of volatility is called implied. It is an enormous advantage that prices become comparable by implied volatilities.
Comparing implied volatilities in respect of different maturities and strikes we obtain the volatility surface.

There is still the second question we must find an answer to: How can the model be extended to achieve a better adjustment to market data? A whole string of modifications has been developed:

1. Stochastic Volatilities,
2. Non-lognormal Distribution,
3. Jump Diffusion.

These extensions make it possible to reproduce volatility surfaces. The parameters can be validated with market data. So we use the basic derivatives (e.g. put and call) to identify the parameters of the accordant model. Therefore we should choose a model that allows a fast calculation of plain vanilla prices. Once we know the parameters the model can be used to price other options as well.
This is usually the point when numerics comes into play. A schema of this procedure could be

1. Model set up (stochastic volatility, jump diffusion, etc.),
2. Parameter Estimation (fitting the model to market data),
3. Numerical Simulation.

Bibliography

[AA98] L. Andersen and J. Andreasen. Volatility skews and extensions of the Libor market model. *Working paper, General Re Financial Products*, 1998.

[ABR01] L. Andersen and R. Brotherton-Ratcliffe. Extended Libor market models with stochastic volatility. *Working paper, General Re Financial Products*, 2001.

[Arn73] L. Arnold. *Stochastische Differentialgleichungen*. R. Oldenbourg Verlag, 1973.

[Bau92] H. Bauer. *Wahrscheinlichkeitstheorie*. de Gruyter, 1992.

[Bog98] V. Bogachev. *Gaussian Measures*. American Mathematical Society, 1998.

[BR97] R. Brotherton-Ratcliffe. The bgm model for path-dependent swaps. *Working Paper, General Re Financial Products*, 1997.

[GJ03] M. Günther and A. Jüngel. *Finanzderivate mit MATLAB*. Vieweg, 2003.

[GL97] J.G. Gaines and T.J. Lyons. Variable step size control in the numerical solution of stochastic differential equations. *SIAM*, 57(5):1455–1484, 1997.

[HMGR99] N. Hofmann, T. Müller-Gronbach, and K. Ritter. Optimal approximation of stochastic differential equations by adaptive stepsize control. *Mathematics of Computation*, 69(231):1017–1034, 1999.

[HT94] W. Hakenbroch and A. Thalmaier. *Stochastische Analysis*. B.G. Teubner Stuttgart, 1994.

[HW87] E. Hairer and G. Warner. *Solving Ordinary Differential Equations II*. Springer, 1987.

[Jam97] F. Jamshidian. Libor and swap market models and measures. *Finance and Stochastics*, 1:293–330, 1997.

[JK81] N.L. Johnson and S. Kotz. *Continuous Univariate Distributions*. Wiley and Sons., 1981.

[KA00] P. Knabner and L. Angermann. *Numerik partieller Differentialgleichungen*. Springer-Verlag, 2000.

[KK01] R. Korn and E. Korn. *Optionsbewertung und Portfoliooptimierung*. Vieweg, 2001.

[KP92] P.E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer-Verlag, 1992.

[KS88] I. Karatzas and S. Shreve. *Brownian Motion and Stochastic Calculus*. Springer-Verlag, 1988.

[KT74] S. Karlin and H.M. Taylor. *A first course in stochastic processes*. Academic Press, 1974.

[KT81] S. Karlin and H.M. Taylor. *A second course in stochastic processes*. Academic Press, 1981.

[Mil95] G.N. Milstein. *Numerical solution of Stochastic Differential Equations*. Kluwer, 1995.

[MPS98] G.N. Milstein, E. Platen, and H. Schurz. Balanced implicit methods for stiff stochastic systems. *SIAM*, 38(3):1010–1019, 1998.

[MR97] M. Musiela and M. Rutkowski. Continuous-time term structure models:forward measure approach. *Finance and Stochastics*, 1(4):261–291, 1997.

[MSS97] K. Miltersen, K. Sandmann, and D. Sondermann. Closed form solutions for term structure derivatives with lognormal interes rates. *Journal of Finance*, 52(1):409–430, 1997.

[Nel96] I. Nelken. *The Handbook of Exotic Options*. IRWIN Professional Publishing, 1996.

[Øks00] B. Øksendal. *Stochastic Differential Equations*. Springer-Verlag, 2000.

[Par72] K. Parthasarathy. *Probability measures on metric spaces*. Academic Press, 1972.

[Sch89] M. Schroder. Computing the constant elasticity of variance option pricing formula. *Journal of Finance*, 44(1):211–219, 1989.

[Sch96] H. Schurz. Numerical regularization for SDE's: Construction of nonnegative solutions. *Dyn. Syst. Appl.*, 5:323–352, 1996.

[Sch97] H. Schurz. *Stability,stationarity, and boundedness of some implicit numerical methods for stochastic differential equations and applications*. Logos, 1997.

[Sey00] R. Seydel. *Einführung in die numerische Berechnung von Finanz-Derivaten*. Springer-Verlag, 2000.

[Ste01] M. Steele. *Stochastic Calculus and Financial Applications*. Springer-Verlag, 2001.

[vR02] M. van Regenmortel. BGM with a smile: Theory and practice. *Internal paper*, 2002.

[Wil00a] P. Wilmott. *Quantative Finance*, volume 1. Wiley, 2000.

[Wil00b] P. Wilmott. *Quantative Finance*, volume 2. Wiley, 2000.