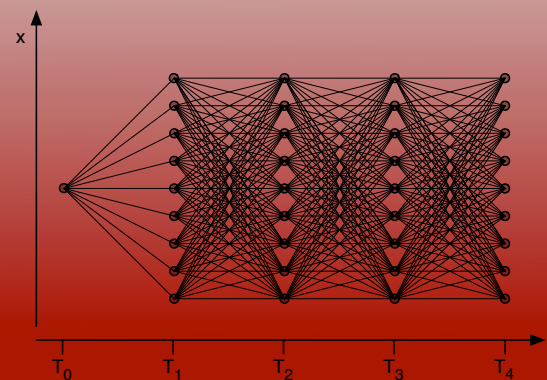
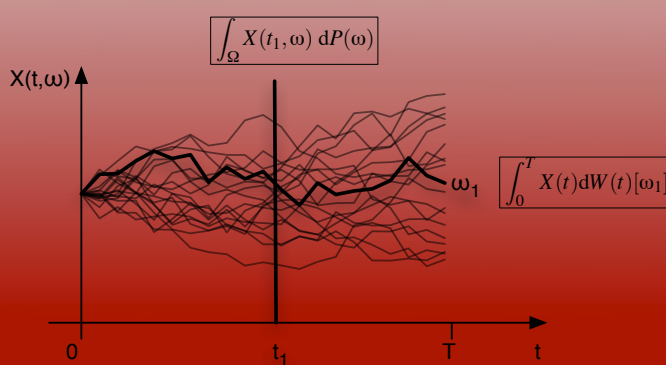


Mathematical Finance

Theory, Modeling and Implementation

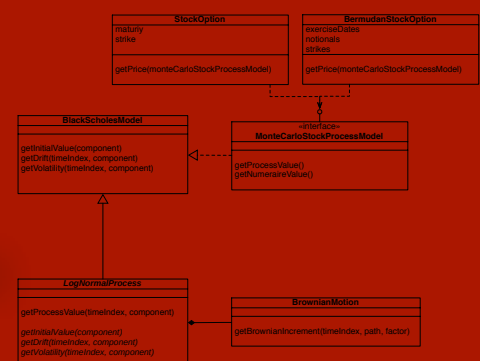
Christian P. Fries

Stochastic Processes • Interest Rate Structures • Hybrid Models
Numerical Methods • Object Oriented Implementation



$$\begin{aligned} dL_j(t) &= L_j(t) \cdot \mu_j(t) dt + L_j(t) \cdot \sigma_j(t) dW_j \quad j = 1, \dots, N \\ dFX(t) &= FX(t) \cdot \mu^{FX}(t) dt + FX(t) \cdot \sigma^{FX}(t) dW^{FX} \\ d\tilde{L}_j(t) &= \tilde{L}_j(t) \cdot \tilde{\mu}_j(t) dt + \tilde{L}_j(t) \cdot \tilde{\sigma}_j(t) d\tilde{W}_j \quad j = 1, \dots, N \end{aligned}$$

$$\begin{aligned} \mu_j(t) &= \sum_{\beta(t) \leq l \leq j} \frac{L_l \delta_l}{(1 + L_l \delta_l)} \cdot \sigma_j(t) \sigma_l(t) \rho_{j,l} \quad \int_{T_i}^{T_{i+1}} \mu^{FX}(t) dt = \log \left(\frac{1 + L_i(T_i) \delta_i}{1 + L_i(T_i) \delta_i} \right) \\ \tilde{\mu}_j(t) &= \sum_{\beta(t) \leq l \leq j} \frac{\tilde{L}_l \delta_l}{(1 + \tilde{L}_l \delta_l)} \cdot \tilde{\sigma}_j(t) \tilde{\sigma}_l(t) \rho_{j,l} - \tilde{\sigma}_j(t) \sigma^{FX}(t) \rho_{j,FX}^-(t) \end{aligned}$$



Mathematical Finance

Theory, Modeling and Implementation

**Stochastic Processes, Interest Rate Models, Hybrid Models,
Numerical Methods, Object Oriented Implementation**

Christian P. Fries

13th December 2006

preliminary, uncorrected and incomplete version
(public beta)

Note

Some parts of this manuscript are under construction.

Nowadays people know the price of everything
and the value of nothing.

Oscar Wilde
The Picture of Dorian Gray, [36].

Typesetted by the author using TeXShop for Mac OS X™. Drawings by the author
using OmniGraffle for Mac OS X™. Charts created using Java™ Code by the author.

English Version

The english translation is still in preparation. If a chapter is not available in english, its german version will be given. The time frame of the completion of the translation depends on reader feedback. Please check <http://www.christian-fries.de/finmath/book> for updates. If you like to be informed about updates to the english version send an email to [Christian Fries <email@christian-fries.de>](mailto:email@christian-fries.de)
Subject: Book Mathematical Finance.

License



This work is licensed under [Creative Commons License](#).
Lizensiert fuer <joris@math.leidenuniv.nl> (Joris Bierkens)

Suggested citation

FRIES, CHRISTIAN P.: Mathematical Finance: Theory, Modeling, Implementation.
Frankfurt am Main, 2006, <http://www.christian-fries.de/finmath/book>.

Feedback

Please help in improving this work!

Please send found errors and suggestions to

Christian Fries <email@christian-fries.de>

Thank you.



Among all replies we raffle an **iPod nano**. Your winning chance will be approximately proportional to the number of errors or suggestions sent in.
Due date is 30 September 2007.

Resources

In connection with this book the following resources are available:

- **Interactive experiments and exercises:**
<http://www.christian-fries.de/finmath/applets>

These components are under construction.



This document requires Adobe Acrobat version 6.0 or higher.

Acknowledgement

I am grateful to Andreas Bahmer, Hans-Josef Beauvisage, Michael Belledin, Dr. Urs Braun, Oliver Dauben, Peter Dellbrügger, Dr. Jörg Dickersbach, Sinan Dikmen, Dr. Dirk Ebmeyer, Dr. Lydia Fechner, Christian Ferber, Frank Genheimer, Dr. Gido Herbers, Prof. Dr. Ansgar Jüngel, JProf. Dr. Christoph Kühn, Dr. Jürgen Linde, Markus Meister, Dr. Sean Matthews, Michael Paulsen, Matthias Peter, Dr. Erwin Pier-Ribbert, Rosemarie Philippi, Dr. Radu Tunaru, Frank Ritz, Marius Rott, Thomas Schwiertz, Benedikt Wilbertz, Polina Zeydis, and Jörg Zinnegger.

Their support and their feedback as well as the stimulating discussions we had contributed significantly to this work.

Contents

1. Introduction	21
1.1. How to Read this Book	22
1.1.1. Abridged Versions	22
1.1.2. Special Sections	23
 I. Foundations	 25
2. Foundations	27
2.1. Probability Theory	27
2.2. Stochastic Processes	35
2.3. Filtration	36
2.4. Brownian Motion	37
2.5. Wiener Measure, Canonical Setup	38
2.6. Itô Calculus	40
2.6.1. Itô Integral	43
2.6.2. Itô Process	45
2.6.3. Itô Lemma and Product Rule	47
2.7. Brownian Motion with Instantaneous Correlation	50
2.8. Martingales	52
2.8.1. Martingale Representation Theorem	52
2.9. Change of Measure (Girsanov, Cameron, Martin)	53
2.10. Stochastic Integration	57
2.11. List of Symbols	59
 3. Replication	 61
3.1. Replication Strategies	61
3.1.1. Introduction	61
3.1.2. Replication in a discrete Model	65

3.2. Foundations: <i>Equivalent Martingale Measure</i>	69
3.2.1. Challenge and Solution Outline	69
English version of this part not available	
3.2.2. Steps towards the <i>Universal Pricing Theorem</i>	72
3.3. Basic Assumptions	79
3.4. Excursion: Relative Prices and Risk Neutral Measures	80
3.4.1. <i>Why relative Prices?</i>	80
3.4.2. <i>Risk Neutral Measure</i>	82
 II. First Applications	 83
 4. Pricing of an European Stock Option under the Black-Scholes Model	 85
 5. Excursion: The Density of the Underlying of an European Call Option	 89
 6. Excursion: Interpolation of European Option Prices	 91
6.1. No-Arbitrage Conditions for Interpolated Prices	91
6.2. Arbitrage Violations through Interpolation	92
6.2.1. Example (1): Interpolation of four Prices	92
6.2.2. Example (2): Interpolation of two Prices	95
6.3. Arbitrage Free Interpolation of European Option Prices	97
 7. Hedging in Continuous and Discrete Time and the Greeks	 99
English version of this part not available	
7.1. Introduction	99
7.2. Deriving the Replications Strategy from Pricing Theory	100
7.2.1. Deriving the Replication Strategy under the Assumption of a Locally Riskless Product	102
7.2.2. The Black-Scholes Differential Equation	103
7.2.3. The Derivative $V(t)$ as a Function of its Underlyings $S_i(t)$	103
7.2.4. Example: Replication Portfolio and PDE under a Black-Scholes Model	104
7.3. Greeks	107
7.3.1. Greeks einer europäischen Call-Option unter dem Black-Scholes Modell	108
7.4. Hedging in Discrete Time: Delta- and Delta-Gamma-Hedging	109
7.4.1. Delta Hedging	110
7.4.2. Error Propagation	111
7.4.3. Delta-Gamma Hedging	112

7.4.4. Vega Hedging	117
7.5. Hedging in discrete Time: Minimizing the Residual Error (Bouchaud-Sornette Method)	119
7.5.1. Minimizing the Residual Error at Maturity T	120
7.5.2. Minimizing the Residual Error in each Time Step	121
III. Interest Rate Structures, Interest Rate Products and Analytic Pricing Formulas	123
Motivation and Overview	125
8. Interest Rate Structures	127
8.1. Introduction	127
8.1.1. Fixing Times and Tenor Times	128
8.2. Definitions	128
8.3. Interest Rate Curve Bootstrapping	133
8.4. Interpolation of Interest Rate Curves	134
8.5. Implementation	135
9. Simple Interest Rate Products	137
9.1. Interest Rate Products Part 1: Products without Optionality	137
9.1.1. Fix, Floating and Swap	137
9.1.2. Money-Market Account	144
9.2. Interest Rate Products Part 2: Simple Options	145
9.2.1. Cap, Floor, Swaption	145
9.2.2. Foreign Caplet, Quanto	147
10. The Black Model for a Caplet	149
11. Pricing of a Quanto Caplet (Modeling the FFX)	151
11.1. Choice of Numéraire	151
12. Exotic Derivatives	155
12.1. Prototypical Product Properties	155
12.2. Interest Rate Products Part 3: Exotic Interest Rate Derivatives	157
12.2.1. Structured Bond, Structured Swap, Zero Structure	157
12.2.2. Bermudan Callables and Cancelable	163
12.2.3. Compound Options	166
12.2.4. Trigger Products	166
12.2.5. Structured Coupons	168
12.2.6. Shout Options	172
12.3. Product Toolbox	172

IV. Discretization and Numerical Valuation Methods	175
Motivation and Overview	177
13. Discretization of time and state space	179
13.1. Discretization of Time: The <i>Euler</i> and the <i>Milstein Scheme</i>	179
13.1.1. Definitions	179
13.1.2. Time-Discretisation of a Lognormal Process	183
13.2. Discretization of Paths (Monte-Carlo Simulation)	184
13.2.1. Monte-Carlo Simulation	184
13.2.2. Weighted Monte-Carlo Simulation	185
13.2.3. Implementation	185
13.2.4. Review	188
13.3. Discretization of State Space	189
13.3.1. Definitions	189
13.3.2. Backward-Algorithm	191
13.3.3. Review	192
13.4. <i>Path Simulation</i> through a <i>Lattice</i> : Two Layers	192
14. Pricing Options with Early Exercise by Monte Carlo Simulation	195
14.1. Introduction	195
14.2. Bermudan Options: Notation	196
14.2.1. Bermudan Callable	196
14.2.2. Relative Prices	197
14.3. Bermudan Option as Optimal Exercise Problem	198
14.3.1. Bermudan Option Value as single (unconditioned) Expecta- tion: The Optimal Exercise Value	198
14.4. Bermudan Option Pricing - The Backward Algorithm	199
14.5. Re-simulation	200
14.6. Perfect Foresight	201
14.7. Conditional Expectation as Functional Dependence	202
14.8. Binning	203
14.8.1. Binning as a Least-Square Regression	205
14.9. Foresight Bias	207
14.10. Regression Methods - Least Square Monte Carlo	207
14.10.1. Least Square Approximation of the Conditional Expectation	208
14.10.2. Example: Evaluation of a <i>Bermudan</i> Option on a Stock (Back- ward Algorithm with Conditional Expectation Estimator)	209
14.10.3. Example: Evaluation of an Bermudan Callable	210
14.10.4. Implementation	214
14.10.5. Binning as linear Least-Square Regression	214
14.11. Optimization Methods	216
14.11.1. Andersen Algorithm for Bermudan Swaptions	216

14.11.2. Review of the Threshold Optimization Method	217
14.11.3. Optimization of Exercise Strategy: A more general Formulation	219
14.11.4. Comparison of Optimization Method and Regression Method	220
14.12. Duality Method: Upper Bound for Bermudan Option Prices - The Method of Rogers	221
14.12.1. Foundations	221
14.12.2. American Option Evaluation as <i>Optimal Stopping</i> Problem . .	223
14.13. Primal-Dual Method: Upper and Lower Bound	226
15. Sensitivities (Partial Derivatives) of Monte-Carlo Prices	229
15.1. Introduction	229
15.2. Problem Description	229
15.2.1. Pricing using Monte Carlo Simulation	230
15.2.2. Sensitivities from Monte Carlo Pricing	231
15.2.3. Example: The Linear and the Discontinuous Payout	231
15.2.4. Example: Trigger Products	232
15.3. Generic Sensitivities: Bumping the Model	232
15.4. Sensitivities by Finite Differences	236
15.4.1. Example: Finite Differences applied to Smooth and Discon- tinuous Payout	237
15.5. Sensitivities by Pathwise Differentiation	238
15.5.1. Example: Delta of a European Option under a Black-Scholes Model	239
15.5.2. Pathwise Differentiation for Discontinuous Payouts	240
15.6. Sensitivities by Likelihood Ratio Weighting	240
15.6.1. Example: Delta of a European Option under a Black-Scholes Model using Pathwise Derivative	241
15.6.2. Example: Variance Increase of the Sensitivity when using Likelihood Ratio Method for Smooth Payouts	241
15.7. Sensitivities by Malliavin Weighting	242
15.8. Proxy Simulation Scheme	243
16. Proxy Simulation Schemes for Monte-Carlo Sensitivities and Im- portance Sampling	245
16.1. Full Proxy Simulation Scheme	245
16.1.1. Calculation of Monte-Carlo weights	246
16.2. Sensitivities by Finite Differences on a Proxy Simulation Scheme . . .	248
16.2.1. Localization	248
16.2.2. Object Oriented Design	249
16.3. Importance Sampling	249
16.3.1. Example	251
16.4. Partial Proxy Simulation Schemes	252

V. Pricing Models for Interest Rate Derivatives	253
17. Market Models	255
17.1. LIBOR Market Model	256
17.1.1. Derivation of the Drift Term	257
17.1.2. The Short Period Bond $P(T_{m(t)+1}; t)$	262
17.1.3. Discretization and (Monte Carlo) Simulation	263
17.1.4. Calibration - Choice of the free Parameters	264
17.1.5. Interpolation of Forward Rates in the LIBOR Market Model	271
17.2. Object Oriented Design	275
17.2.1. Reuse of Implementation	275
17.2.2. Separation of Product and Model	276
17.2.3. Abstraction of Model Parameters	276
17.2.4. Abstraction of Calibration	278
17.3. Swap Rate Market Models (Jamshidian 1997)	279
17.3.1. The <i>Swap Measure</i>	280
17.3.2. Derivation of the Drift Term	281
17.3.3. Calibration - Choice of the free Parameters	281
18. Excursion: Instantaneous Correlation and Terminal Correlation	285
18.1. Definitions	285
18.2. Terminal Correlation studied on the Example of a LIBOR Market Model	286
18.2.1. De-correlation in a One-Factor-Model	287
18.2.2. Impact of the Time Structure of the Instantaneous Volatility on Caplet and Swaption Prices	288
18.2.3. The Swaption Value as a Function of Forward Rates	289
18.3. Terminal Correlation depends on the Equivalent Martingale Measure	292
18.3.1. Dependence of the Terminal Density on the Martingale Measure	292
19. Heath-Jarrow-Morton Framework: Foundations	295
19.1. Short Rate Process in the HJM Framework	296
19.2. The HJM Drift Condition	296
20. Short Rate Models	301
20.1. Introduction	301
20.2. The <i>Market Price of Risk</i>	302
20.3. Overview: Some Common Models	304
20.4. Implementations	304
20.4.1. Monte-Carlo Implementation of Short-Rate Models	304
20.4.2. Lattice Implementation of Short-Rate Models	305

21. Heath-Jarrow-Morton Framework: Immersion of Short Rate Models and LIBOR Market Model	307
21.1. Short Rate Models in the HJM Framework	307
21.1.1. Example: The Ho-Lee Model in the HJM Framework	307
21.1.2. Example: The Hull-White Model in the HJM Framework	308
21.2. LIBOR Market Model in the HJM Framework	310
21.2.1. HJM Volatility Structure of the LIBOR Market Model	310
21.2.2. LIBOR Market Model Drift under the $\mathbb{Q}^B$ Measure	312
21.2.3. LIBOR Market Model as a Short Rate Model	313
22. Excursion: Shape of the Interest Rate Curve under Mean Reversion and a Multi-Factor Model	315
22.1. Model	315
22.2. Interpretation of the Figures	316
22.3. Mean Reversion	316
22.4. Factors	318
22.5. Exponential Volatility Function	319
22.6. Instantaneous Correlation	321
23. Markov Functional Models	323
23.1. Introduction	323
23.1.1. The Markov Functional Assumption (independent of the model considered)	324
23.1.2. Outline of this Chapter	325
23.2. Equity Markov Functional Model	325
23.2.1. Markov Functional Assumption	325
23.2.2. Example: The Black-Scholes Model	326
23.2.3. Numerical Calibration to a Full Two Dimensional European Option Smile Surface	327
23.2.4. Interest Rates	328
23.2.5. Model Dynamics	330
23.2.6. Implementation	335
23.3. LIBOR Markov Functional Model	335
23.3.1. LIBOR Markov Functional Model in Terminal Measure - Hunt, Kennedy, Pelsser	335
23.4. Implementation: Lattice	340
23.4.1. Convolution with the Normal Probability Density	341
23.4.2. State space discretization	343
VI. Extended Models	345

24. Hybrid Models	347
24.1. Cross Currency LIBOR Market Model	347
24.1.1. Derivation of the Drift Term under Spot-Measure	348
24.1.2. Implementation	352
24.2. Equity Hybrid LIBOR Market Model	352
24.2.1. Derivation of the Drift Term under Spot-Measure	352
24.2.2. Implementation	354
24.3. Equity-Hybrid Cross-Currency LIBOR Market Model	354
24.3.1. Summary	356
24.3.2. Implementation	357
 VII. Implementation	 359
25. Object-Oriented Implementation in Java™	361
English version of this part not available	
25.1. Elements of Object Oriented Programming: Class and Objects	362
25.1.1. Example: Class of a Binomial Distributed Random Variable	363
25.1.2. Constructor	364
25.1.3. Methods: Getter, Setter, Static Methods	365
25.2. Principles of Object Oriented Programming	366
25.2.1. Data Hiding and Interfaces	366
25.2.2. Abstraction and Inheritance	370
25.2.3. Polymorphism	373
25.3. Example: A Class Structure for One Dimensional Root Finders	374
25.3.1. Root Finder for General Functions	374
25.3.2. Root Finder for Functions with Analytic Derivative: Newton Method	376
25.3.3. Root Finder for Functions with Derivative Estimation: Secant Method	377
25.4. Anatomy of a Java™ Class	380
25.5. Libraries	381
25.5.1. Java™ 2 Platform, Standard Edition (j2se)	381
25.5.2. Java™ 2 Platform, Enterprise Edition (j2ee)	382
25.5.3. Colt	382
25.5.4. Commons-Math: The Jakarta Mathematics Library	382
25.6. Some Final Remarks	382
25.6.1. Object Oriented Design (OOD) / Unified Modeling Language (UML)	382
 VIII. Appendix	 385

A. Tools (Selection)	387
English version of this part not available	
A.1. Cholesky Decomposition	387
A.2. Linear Regression	388
A.3. Generation of Random Numbers	389
A.3.1. Uniform Distributed Random Variables	389
A.3.2. Transformation of the Random Number Distribution via the Inverse Distributionfunction	389
A.3.3. Normal Distributed Random Variables	389
A.3.4. Poisson Distributed Random Variables	390
A.3.5. Generation of Paths of an n -dimensional Brownian Motion . .	390
A.4. Generation of Correlated Brownian Motion	393
A.5. Factor Reduction	394
A.6. Optimization (one dimensional): Golden Section Search	396
A.7. Convolution with the Normal Density	397
B. Exercises	399
English version of this part not available	
C. List of Symbols	407
English version of this part not available	
D. Java™ Source Code (Selection)	409
D.1. Java™ Classes for Chapter 25	409
E. Technical Terms (Selection)	413
English version of this part not available	
List of Figures	415
List of Tables	419
Bibliography	423

CHAPTER 1

Introduction

1.1. How to Read this Book

The text may be read *non linearly*, i.e. dependencies between chapters have been kept as low as possible. Chapter 2 gives the foundations in the order of their dependence. The reader familiar with the concepts of stochastic processes and martingales may skip the chapter and use it as reference only. To get a feeling for the mathematical concepts one should read the special sections *Interpretation* and *Motivation*. Readers familiar with programming and implementation may prefer Chapter 13 as illustration of the basic concepts.

The appendix give a selection of the results and techniques from diverse areas (linear algebra, calculus, optimization), which are used in the text and in the implementation, but which are less important in the understanding of the essential concepts.

1.1.1. Abridged Versions

For a *crash course* focusing on particular aspects some chapters may be skipped. What follows are a few suggestions in this direction.

1.1.1.1. Abridged version “Monte-Carlo pricing”

Foundations (Ch. 2) → Replication (Ch. 3) → Black-Scholes Model (Ch. 4)
→ Discretization / Monte-Carlo Simulation (Ch. 13)

1.1.1.2. Abridged version “LIBOR Market Model”

Foundations (Ch. 2) → Replication (Ch. 3) → Interest Rate Structures (Ch. 8)
→ Black Model (Ch. 10) → LIBOR Market Model (Ch. 17)
→ Instantaneous and Terminal Correlation (Ch. 18)
→ Shape of the Interest Rate Curve (Ch. 22)

1.1.1.3. Abridged version “Markov Functional Model”

Foundations (Ch. 2) → Replication (Ch. 3) → Interest Rate Structures (Ch. 8)
→ Black Model (Ch. 10)
→ The Density of the Underlying of an European Option (Ch. 5)
→ Markov Functional Models (Ch. 23)

1.1.2. Special Sections

The text contains special sections to give with notes on interpretation, motivation and practical aspects. These are marked by the following symbols:



Interpretation: *Gives an interpretation of the preceding topic. Casts light on purposes and practical aspects.* <|



Motivation: *Gives motivation for the following topic. Sometimes notes deficiencies of the previous results.* <|



Further Reading: *Points to literature and associated topics.* <|



Experiment: *Hints to a software experiment where aspects of the preceding topic may be explored.* <|



Tip: *Hints for practical use and software implementation of the preceding topics.* <|

Part I.

Foundations: Probability Theory, Stochastic Processes and Risk Neutral Evaluation

CHAPTER 2

Foundations

2.1. Probability Theory

Definition 1 (Probability Space, σ Algebra):

Let Ω denote a set and $\mathcal{F}$ a family of subsets of Ω . $\mathcal{F}$ is a σ algebra if

1. $\emptyset \in \mathcal{F}$
2. $F \in \mathcal{F} \Rightarrow \Omega \setminus F \in \mathcal{F}$
3. $F_1, F_2, F_3, \dots \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} F_i \in \mathcal{F}$.

The pair $(\Omega, \mathcal{F})$ is a *measurable space*. A function $P : \mathcal{F} \rightarrow [0, \infty)$ is a *probability measure* if

1. $P(\emptyset) = 0, P(\Omega) = 1$,
2. For $F_1, F_2, F_3, \dots \in \mathcal{F}$ mutually disjoint (i.e. $i \neq j \Rightarrow F_i \cap F_j = \emptyset$), we have

$$P\left(\bigcup_{i=1}^{\infty} F_i\right) = \sum_{i=1}^{\infty} P(F_i).$$

The triple $(\Omega, \mathcal{F}, P)$ is called *probability space* (if instead of 1 we require only $P(\emptyset) = 0$ then P is called *measure* and $(\Omega, \mathcal{F}, P)$ is called *measure space*).



Interpretation: The set Ω may be interpreted as the set of elementary events. Only one such event may occur. The subset $F \subset \Omega$ may then be interpreted as an event configuration, e.g. as if one would ask only for a specific property of an event, a property that might be shared by more than one event. Then the complement of a set of events corresponds to the negation of the property in question, the union of two subset $F_1, F_2 \subset \Omega$ corresponds to combining the questions for the two corresponding properties with an “or”. Likewise

the intersection corresponds to an “and”: only those events that share both properties are part of the intersection. A σ -algebra may then be interpreted as a set of properties, e.g. the set of properties by which we may distinguish the events or the set of properties on which we may base decisions and answer questions. Thus the σ -algebra may be interpreted as information (on properties of events).

Thus a probability space $(\Omega, \mathcal{F}, P)$ may be interpreted as a set of elementary events, a family of properties of the events and a map that assigns a probability to each property of the events, the probability that an event with the respective properties occurs. $\triangleleft$

Since conditional expectation will be one of the central concepts we remind the reader of the notions of *conditional probability* and *independence*:

Definition 2 (Independence, Conditional Probability): $\triangleright$

Let $(\Omega, \mathcal{F}, P)$ denote a probability space and $A, B \in \mathcal{F}$.

1. We say that A and B are *independent*, if

$$P(A \cap B) = P(A) \cdot P(B).$$

2. For $P(B) > 0$ we define the *conditional probability of A under the hypothesis B* as

$$P(A|B) := \frac{P(A \cap B)}{P(B)}.$$

The Borel σ -algebra $\mathcal{B}(\mathbb{R})$ or $\mathcal{B}(\mathbb{R}^n)$ plays a special role in integration theory. We define it next.

Definition 3 (Borel σ Algebra, Lebesgue Measure): $\triangleright$

Let $n \in \mathbb{N}$ and $a_i < b_i$ ($i = 1, \dots, n$). By $\mathcal{B}(\mathbb{R}^n)$ we denote the smallest σ -algebra for which

$$(a_1, b_1) \times \dots \times (a_n, b_n) \in \mathcal{B}(\mathbb{R}^n).$$

$\mathcal{B}(\mathbb{R}^n)$ is called the *Borel σ -algebra*. The measure λ defined on $\mathcal{B}(\mathbb{R}^n)$ with

$$\lambda((a_1, b_1) \times \dots \times (a_n, b_n)) := \prod_{i=1}^n (b_i - a_i)$$

is called a *Lebesgue measure* on $\mathcal{B}(\mathbb{R}^n)$. $\triangleleft$

Remark 4 (Lebesgue Measure): Obviously the Lebesgue measure is *not* a probability measure on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ since $\lambda(\mathbb{R}^n) = \infty$. It will be needed in the discussion of Lebesgue integration and we give the definition merely for completeness.¹

Definition 5 (Measurable, Random Variable): $\triangleright$

Let $(\Omega, \mathcal{F})$ and $(S, \mathcal{S})$ denote two measurable spaces.

¹ The Lebesgue measure measures intervals ($n = 1$) according to their length, rectangles ($n = 2$) according to their area and cubes ($n = 3$) according to their volume.

1. A map $T : \Omega \mapsto S$ is called $(\mathcal{F}, \mathcal{S})$ -measurable if²

$$T^{-1}(A) \in \mathcal{F} \text{ for all } A \in \mathcal{S}.$$

If $T : \Omega \mapsto S$ is a $(\mathcal{F}, \mathcal{S})$ -measurable map we write briefly

$$T : (\Omega, \mathcal{F}) \mapsto (S, \mathcal{S}).$$

2. A measurable map $X : (\Omega, \mathcal{F}) \mapsto (S, \mathcal{S})$ is also called a *random variable*. A random variable $X : (\Omega, \mathcal{F}) \mapsto (S, \mathcal{S})$ is called a *n-dimensional real-valued random variable* if $S = \mathbb{R}^n$ and $\mathcal{S} = \mathcal{B}(\mathbb{R}^n)$.

We are interested in the probability for which a given random variable attains a certain value or range of values. This is given by

Definition 6 (Image Measure):

Let $X : (\Omega, \mathcal{F}) \mapsto (S, \mathcal{S})$ denote a random variable and P a measure on the measure space $(\Omega, \mathcal{F})$. Then

$$P_X(A) := P(X^{-1}(A)) \quad \forall A \in \mathcal{S}$$

defines a probability measure on $(S, \mathcal{S})$ which we call the *image measure* of P with respect to X .



Interpretation: A real-valued random variable assigns to each elementary event ω a real value (or vector of values). This value may be interpreted as the result of an experiment depending on the events. In our context the random variables mostly stand for payments or values of financial products depending on the *state of the world*. How random a random variable is depends on the random variable itself. The random variable that assigns the same value to all events ω exhibits no randomness at all. If we could observe only the result of such an experiment (random variable) we would not be able to say anything on the state of the world ω that led to the result.

The image measure is the probability measure induced by the probability measure P (a probability measure on $(\Omega, \mathcal{F})$) and the map X on the image space $(S, \mathcal{S})$.

The property of being measurable may be interpreted as the property that the distinguishable events in the image space $(S, \mathcal{S})$ are not finer (better distinguishable) than the events in the pre-image space $(\Omega, \mathcal{F})$. Only then it is possible to use the probability measure P on $(\Omega, \mathcal{F})$ to define a probability measure on $(S, \mathcal{S})$.

Exercise: Let X be as in Definition 6. Show that

$$\{X^{-1}(A) \mid A \in \mathcal{S}\}$$

is a σ -algebra. What would be an interpretation of $X^{-1}(A)$?

² We define $T^{-1}(A) := \{\omega \in \Omega \mid T(\omega) \in A\}$.

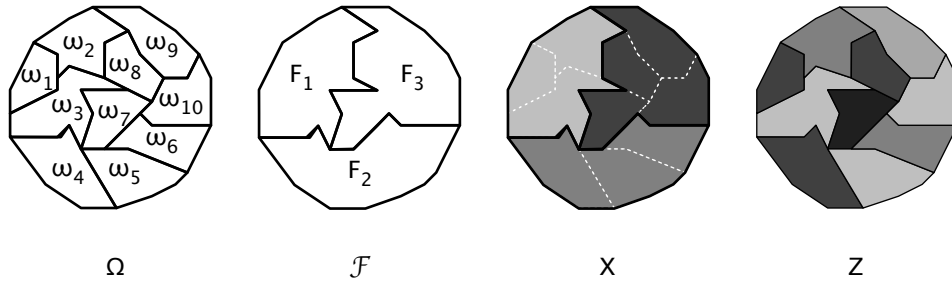


Figure 2.1.: Illustration of Measurability: The random variables X and Z assign a grey value to each elementary event $\omega_1, \dots, \omega_{10}$ as shown. The σ -algebra $\mathcal{F}$ is generated by the sets $F_1 = \{\omega_1, \omega_2, \omega_3\}$, $F_2 = \{\omega_4, \omega_5, \omega_6\}$, $F_3 = \{\omega_7, \dots, \omega_{10}\}$. The random variable X is measurable with respect to $\mathcal{F}$, the random variable Z is not measurable with respect to $\mathcal{F}$.



Motivation: We will now define the *Lebesgue integral* and give an interpretation and a comparison to the (possibly more familiar) Riemann integral.

The definition of the Lebesgue integral is not only given to prepare the definition of the conditional expectation (definition 14). The definition will also show the construction of the Lebesgue integral and we will later use similar steps to construct the Itô integral. ◁|

Definition 7 (Lebesgue Integral):

Let $(\Omega, \mathcal{A}, \lambda)$ denote a measure space. ▮

1. Let f denote a $(\mathcal{A}, \mathcal{B}(\mathbb{R}))$ -measurable real valued, non-negative map. f is called an *elementary function* if f takes on only a finite number of values $a_1, \dots, a_n$. For an elementary function we define

$$\int_{\Omega} f(\omega) d\lambda(\omega) := \sum_{i=1}^n a_i \lambda(A_i)$$

where $A_i := f^{-1}(\{a_i\})$ ($\Rightarrow A_i \in \mathcal{A}$)³ as the *(Lebesgue) Integral* of f .

2. Let f denote a non-negative map defined on Ω , such that a monotone increasing sequence $(u_k)_{k \in \mathbb{N}}$ of elementary maps with $f := \sup_{k \in \mathbb{N}} u_k$ exists. Then

$$\int_{\Omega} f(\omega) d\lambda(\omega) := \sup_{k \in \mathbb{N}} \int_{\Omega} u_k(\omega) d\lambda(\omega)$$

is unique and is called the *(Lebesgue) integral* of f .

³ We have $(a_i - \frac{1}{n}, a_i + \frac{1}{n}) \in \mathcal{B}(\mathbb{R})$ by Definition. Then $\{a_i\} = \bigcap_{n=1}^{\infty} (a_i - \frac{1}{n}, a_i + \frac{1}{n}) \in \mathcal{B}(\mathbb{R})$.

3. Let f denote a map on Ω such that we have for $f^+ := \max(f, 0)$ and $f^- := \max(-f, 0)$ respectively a monotone increasing sequence of elementary maps as in the previous definition. Furthermore we require that $\int_{\Omega} f^{\pm} d\lambda < \infty$. Then f is called *integrable* with respect to λ and we define

$$\int_{\Omega} f d\lambda := \int_{\Omega} f^+ d\lambda - \int_{\Omega} f^- d\lambda$$

as the (Lebesgue) integral of f .

┘

Theorem 8 (measurable $\Leftrightarrow$ integrable for non-negative maps): For a non-negative map f on Ω a monotone increasing sequence $(u_k)_{k \in \mathbb{N}}$ of elementary maps with $f = \sup_{k \in \mathbb{N}} u_k$ exists if and only if f is $\mathcal{A}$ -measurable.

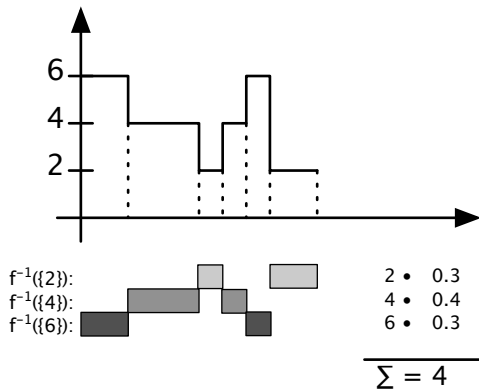
Proof: See [1], §12.

□



Interpretation: To develop a first intuition for the Lebesgue integral we consider its definition for elementary maps

$$\int_{\Omega} f(\omega) d\lambda(\omega) = \sum_{i=1}^n a_i \lambda(A_i), \quad A_i := f^{-1}(\{a_i\}).$$



The Lebesgue integral of an (elementary) map f is the weighted sum of the function values a_i of f , each weighted by the measure $\lambda(A_i)$ of the set on which this value is attained (i.e. $A_i = f^{-1}(\{a_i\})$).

If in addition we have $\lambda(\Omega) = 1$, where Ω is the domain of f , then the integral is a weighted average of the function values a_i of f .

For a real valued (elementary) function of a real valued argument, e.g. $f : [a, b] \mapsto \mathbb{R}$, and the Lebesgue measure the integral corresponds to the naive concept of an integral as being the sum all rectangles given by

base area (Lebesgue measure of the interval) $\times$ height (function value)

which is also the concept behind the Riemann integral.

Part 2 and 3 of Definition 7 extend this concept via a limit approximation to more general functions.

◁

Excursion: On the difference between Lebesgue and Riemann integrals⁴

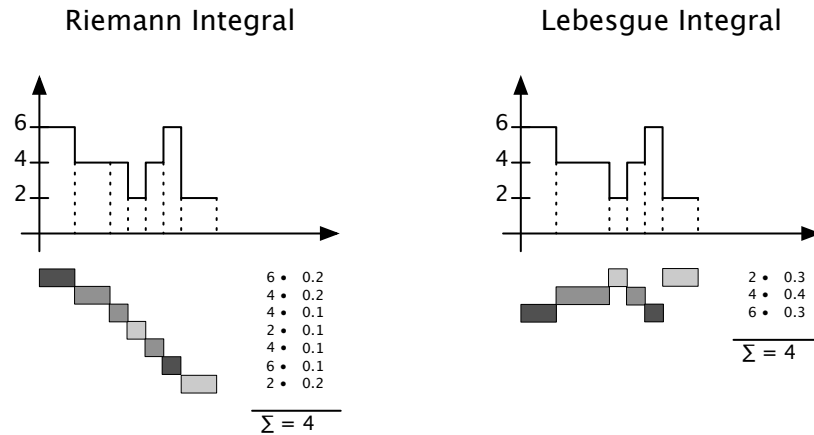


Figure 2.2.: On the difference of Lebesgue and Riemann integral

The construction of the Lebesgue integral differs from the construction of the Riemann integral (which is perhaps more familiar) in the way the sets A_i are chosen. The Riemann integral starts from a given partition of the domain and multiplies the size of each subinterval by a corresponding functional value of (some) chosen point belonging to that interval (e.g. the center point). The Lebesgue integral chooses the partition as pre-image $f^{-1}(\{a_i\})$ of given function values a_i . In short: the Riemann integral partitions the domain of f , the Lebesgue integral partitions the range of f . For elementary functions both approaches give the same integral value, see Figure 2.2. For general functions the corresponding integrals are defined as the limit of a sequence of approximating elementary functions (if it exists). Here, the two concepts are different: In the limit, all Riemann integrable functions are Lebesgue integrable and the two limits give the same value for the integral. However, there exist Lebesgue integrable functions for which the Riemann integral is not defined (its limit construction does not converge).

Definition 9 (Distribution):

Let P denote a probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ (e.g. the image measure of a random variable). The function

$$F_P(x) := P((-\infty, x))$$

is called the *distribution function*⁵ of P .

If P denotes a probability measure on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ the *n-dimensional distribution func-*

⁴ An understanding of the difference between the Lebesgue and Riemann integrals does not play a major role in the following text. The excursion may be skipped without harm. It should serve to satisfy curiosity, e.g. if the concept of a Riemann integral is more familiar.

⁵ We have $(-\infty, x) \in \mathcal{B}(\mathbb{R})$ since $(-\infty, x) = \cup_{i=1}^{\infty} (x-i, x)$ - see Definition 1.

tion is defined as

$$F_P(x_1, \dots, x_n) := P((-\infty, x_1) \times \dots \times (-\infty, x_n)).$$

The distribution function of a random variable X is defined as the distribution function of its image measure P_X (see Definition 6). $\lrcorner$

Definition 10 (Density): $\lrcorner$

Let F_P denote the distribution function of a probability measure P . If F_P is differentiable we define

$$\phi(x) := \frac{\partial}{\partial x} F_P(x)$$

as the *density* of P . If P is the image measure of some random variable X we also say that ϕ is the *density* of X . $\lrcorner$

Remark 11 (Integration using a known density / distribution): To calculate the integral of a function of a random variable (e.g. to calculate expectation or variance) it is sufficient to know the density or distribution function of the random variable. Let g denote a sufficiently smooth function and X a random variable on $(\Omega, \mathcal{F}, P)$; then we have

$$\int_{\Omega} g(X(\omega)) dP(\omega) = \int_{-\infty}^{\infty} g(x) dF_{P_X}(x) = \int_{-\infty}^{\infty} g(x) \phi(x) dx,$$

where F_{P_X} denotes the distribution function of X and ϕ the density of X (i.e. of P_X). In this case it is neither necessary to know the underlying space Ω , the measure P , nor how X is modeled (i.e. defined) on this space.

Definition 12 (Independence of Random Variables): $\lrcorner$

Let $X : (\Omega, \mathcal{F}) \mapsto (S, \mathcal{S})$ and $Y : (\Omega, \mathcal{F}) \mapsto (S, \mathcal{S})$ denote two random variables. X and Y are called *independent*, if for all $A, B \in \mathcal{S}$ the events $X^{-1}(A)$ and $Y^{-1}(B)$ are independent in the sense of Definition 2. $\lrcorner$

Remark 13 (Independence): For $i = 1, \dots, n$ let $X_i : \Omega \mapsto \mathbb{R}$ denote random variables with distribution functions F_{X_i} and let $F_{(X_1, \dots, X_n)}$ denote the distribution function of $(X_1, \dots, X_n) : \Omega \mapsto \mathbb{R}^n$. Then the X_i are pairwise independent if and only if

$$F_{(X_1, \dots, X_n)}(x_1, \dots, x_n) = F_{X_1}(x_1) \cdot \dots \cdot F_{X_n}(x_n).$$

Definition 14 (Expectation, Conditional Expectation): $\lrcorner$

Let X denote a real valued random variable on the probability space $(\Omega, \mathcal{F}, P)$.

1. If X is P -integrable we define

$$E^P(X) := \int_{\Omega} X dP$$

as the *expectation* of X .

2. Let furthermore $F_i \in \mathcal{F}$ with $P(F_i) > 0$. Then

$$E^P(X|F_i) := \frac{1}{P(F_i)} \int_{F_i} X dP$$

is called the *conditional expectation of X under (the hypothesis) F_i* .

┘

Theorem 15 (Conditional Expectation⁶): Let X denote a real valued random variable on $(\Omega, \mathcal{F}, P)$, either non-negative or integrable. Then we have for each σ -algebra $C \subset \mathcal{F}$ a non-negative or integrable real valued random variable $X|_C$ on Ω , unique in the sense of almost sure equality⁷, such that $X|_C$ is C -measurable and

$$\forall C \in \mathcal{C} : \int_C X|_C dP = \int_C X dP \quad , \text{ i.e. } E^P(X|C) = E^P(X_C|C).$$

We will discuss the interpretation of this theorem after giving a name to X_C :

Definition 16 (Conditional Expectation (cont')):

┘

Under the assumptions and with the notation of Theorem 15 we define:

1. The random variable $X|_C$ is called the *conditional expectation of X under (the hypothesis) C* and is denoted by

$$E^P(X|C) := X|_C. \quad (2.1)$$

2. Let Y denote another random variable on the same measure space. We define:

$$E^P(X|Y) := E(X|\sigma(Y)), \quad (2.2)$$

where $\sigma(Y)$ is the σ -algebra generated by Y , i.e. the smallest σ -algebra, with respect to which Y is measurable, i.e. $\sigma(Y) := \sigma(Y^{-1}(S))$.

┘



Interpretation: First note that the two concepts of expectations from Definition 14 are just special cases of the conditional expectation defined in Definition 16, namely:

- Let $C = \{\emptyset, \Omega\}$.
Then $E(X|C) = X|_C$ where $X|_C(\omega) = E(X) \quad \forall \omega \in \Omega$.

⁶ See [2], §15

⁷ A property holds *P-almost surely*, if the set of $\omega \in \Omega$ for which the property does not hold has measure zero.

- For $C = \{\emptyset, F, \Omega \setminus F, \Omega\}$ we have $X|_C(\omega) = \begin{cases} E(X|F) & \text{if } \omega \in F \\ E(X|\Omega \setminus F) & \text{if } \omega \in \Omega \setminus F \end{cases}$,

$$\text{with } E(X) = P(F) \cdot E(X|F) + (1 - P(F)) \cdot E(X|\Omega \setminus F)$$

The conditional expectation is a random variable that is derived from X such that only events (sets) in C can be distinguished. In the first case we have a very coarse C and the image of $X|_C$ contains only the expectation $E(X)$. This is the smallest piece of information on X . As C becomes finer more and more information of X becomes visible in $X|_C$. Furthermore: If X itself is C -measurable then we have that X and $X|_C$ are (P -almost surely) indistinguishable.

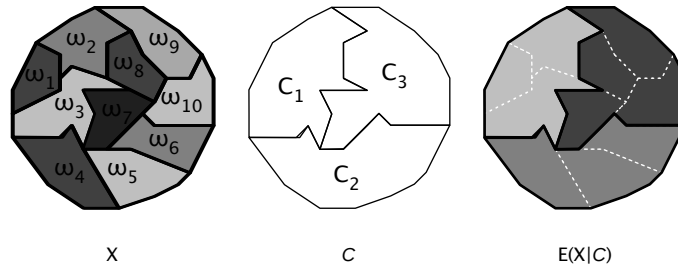


Figure 2.3.: Conditional Expectation: Let the σ -algebra C be generated by the sets $C_1 = \{\omega_1, \omega_2, \omega_3\}$, $C_2 = \{\omega_4, \omega_5, \omega_6\}$, $C_3 = \{\omega_7, \dots, \omega_{10}\}$.

In this sense C may be interpreted as an information set and $X|_C$ as a *filtered* version of X . If it is only possible to make statements upon events in C then one may only make statements about X which could also be made about $X|_C$. $\triangleleft$

2.2. Stochastic Processes

Definition 17 (Stochastic Process):

A family $X = \{X_t \mid 0 \leq t < \infty\}$ of random variables

$$X_t : (\Omega, \mathcal{F}) \rightarrow (S, \mathcal{S})$$

is called (*time continuous*) *stochastic process*. If $(S, \mathcal{S}) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, we say that X is a *d-dimensional stochastic process*. The family X may also be interpreted as a $X : [0, \infty) \times \Omega \rightarrow S$:

$$X(t, \omega) := X_t(\omega) \quad \forall (t, \omega) \in [0, \infty) \times \Omega.$$

If the range $(S, \mathcal{S})$ is not given explicitly we assume $(S, \mathcal{S}) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$.



Interpretation: The parameter t has the obvious interpretation of time. For fixed $t \in [0, \infty)$ we view $X(t)$ as the outcome of an experiment at time t . Note that all random variables $X(t)$ are modeled over the *same* measurable space $(\Omega, \mathcal{F})$. Thus we do not assume a family $(\Omega_t, \mathcal{F}_t)$ of measurable spaces, one for each X_t . The stochastic process X assigns a *path* to each $\omega \in \Omega$: For a fixed $\omega \in \Omega$ the *path* $X(\cdot, \omega) := \{(t, X(t, \omega)) \mid t \in [0, \infty)\}$ is a sequence of outcomes of the random experiments X_t (a *trajectory*), associated with a state ω . Knowledge about $\omega \in \Omega$ implies knowledge of the whole history (past, present and future) $X(\omega)$.

To model the different levels of knowledge and thus distinguish between past and future we will define in 2.3 the concept of a filtration and an adapted process. $\triangleleft$

Definition 18 (Path):

Let X denote a stochastic process. For a fixed $\omega \in \Omega$ the mapping $t \mapsto X(t, \omega)$ is called the path of X (in state ω). $\lrcorner$

2.3. Filtration

Definition 19 (Filtration):

Let $(\Omega, \mathcal{F})$ denote a measurable space. A Family of σ -algebras $\{\mathcal{F}_t \mid t \geq 0\}$, where

$$\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F} \quad \text{for } 0 \leq s \leq t,$$

is called a *Filtration* on $(\Omega, \mathcal{F})$. $\lrcorner$

Definition 20 (Generated Filtration):

Let X denote a stochastic process on $(\Omega, \mathcal{F})$. We define

$$\begin{aligned} \mathcal{F}_t^X &:= \sigma(X_s; 0 \leq s \leq t) \\ &:= \text{the smallest } \sigma\text{-algebra with respect to which } X_s \text{ is measurable } \forall s \in [0, t]. \end{aligned}$$

Definition 21 (Adapted Process):

Let X denote a stochastic process on $(\Omega, \mathcal{F})$ and $\{\mathcal{F}_t\}$ a filtration on $(\Omega, \mathcal{F})$. The process X is called $\{\mathcal{F}_t\}$ -adapted, if X_t is $\mathcal{F}_t$ -measurable for all $t \geq 0$. $\lrcorner$



Interpretation: In Figure 2.4 we depict a filtration of four σ -algebras with increasing refinement (left to right). The black borders surround the generators of the corresponding σ -algebra. If a stochastic process maps a gray value to each elementary event (or path) ω_i of Ω (left), then we have

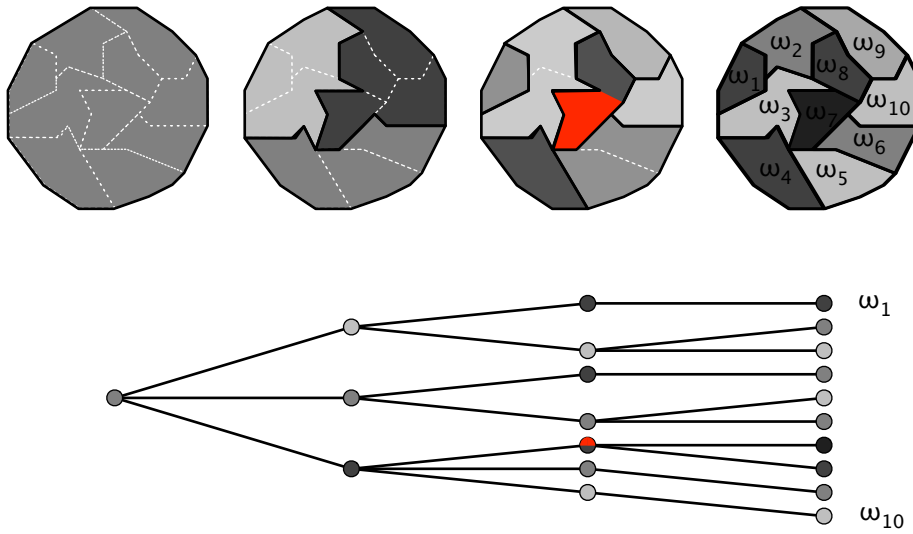


Figure 2.4.: *Illustration of a filtration and an adapted process*

that the process is adapted if it takes a constant gray value on the generators of the respective σ -algebra. If the red parts in Figure 2.4 were dark gray, the process would be adapted, otherwise it is not.

By means of the conditional expectation (see Theorem 15 and the interpretation of Figure 2.3) we may create an adapted process from a given filtration $\{\mathcal{F}_t \mid t \geq 0\}$ and an $\mathcal{F}$ -measurable random variable Z :

Lemma 22 (Process of the Conditional Expectation): Let $\{\mathcal{F}_t \mid t \geq 0\}$ denote a filtration $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ and Z an $\mathcal{F}$ measurable random variable. Then

$$X(t) := E(Z \mid \mathcal{F}_t)$$

is a $\{\mathcal{F}_t\}$ -adapted process.

This lemma shows how the filtration (and the corresponding adapted process) may be viewed as a model for *information*: The random variable $X(t)$ in Lemma 22 allows with increasing t more and more specific statements about the nature of Z . Compare this to the illustrations in Figure 2.1 and 2.3. $\triangleleft$

2.4. Brownian Motion

Definition 23 (Brownian Motion):

Let $W : [0, \infty) \times \Omega \rightarrow \mathbb{R}^n$ denote a stochastic process with the following properties:

1. $W(0) = 0$ (P -almost surely)
2. The map $t \mapsto W(t)$ is continuous (P -almost surely)
3. For given $t_0 < t_1 < \dots < t_k$ the increments $W(t_1) - W(t_0), \dots, W(t_k) - W(t_{k-1})$ are mutually independent.
4. For all $0 \leq s \leq t$ we have that $W(t) - W(s) \sim \mathcal{N}(0, (t - s)I_n)$, i.e. normally distributed with mean 0 and covariance matrix $(t - s)I_n$, where I_n denotes the $n \times n$ identity matrix.

Then W is called (n -dimensional) P -Brownian motion or a (n -dimensional) P -Wiener process. $\lrcorner$

We have not yet discussed the question whether a process with such properties exists (it does). The question for its existence is non-trivial. For example if we want to replace normally distributed by lognormally distributed in property 4 there would be no such process.⁸ If we set $s = 0$ in 4 we see that we have prescribed the distribution of $W(t)$ as well as the distribution of the increments $W(t) - W(s)$.

Remark 24 (Brownian Motion): The property 4 is less axiomatic than one might assume: The central requirement is the independence of the increments together with the requirement that increments of the same time step size $t - s$ have the same non negative variance (here $t - s$) and mean 0. That the increments are normal distributed is more a consequence than an requirement, see Theorem 25. This theorem also gives a construction of the Brownian motion.



Tip (time discrete realizations): In the following we will often consider the realizations of a stochastic process at discrete times $0 = T_0 < T_1 < \dots < T_N$ only (e.g. this will be the case when we consider the implementation). If we need only the realizations $W(T_i)$ we may generate them by the time discrete increments $\Delta W(T_i) := W(T_{i+1}) - W(T_i)$ since from Definition 23 we have $W(T_i) = \sum_{k=0}^{i-1} \Delta W(T_k)$, $W(T_0) := 0$. See Figure 2.5. $\triangleleft$

2.5. Wiener Measure, Canonical Setup

The following theorem gives a construction (or approximation) of a Brownian motion, it defines the *Wiener measure* and shows that the properties of a Brownian motion are less axiomatic than one might assume from Definition 23; rather they are consequences of independence.

Theorem 25 (Invariance Principle of Donsker (1951), see. [18] §2): Let $(\Omega, \mathcal{F}, P)$ denote a probability space and $(Y_j)_{j=1}^\infty$ a sequence of independent identi-

⁸ Note that the sum of two (independent) normal distributed random variables is normally distributed, but the sum of two lognormal distributed random variables is not lognormal.

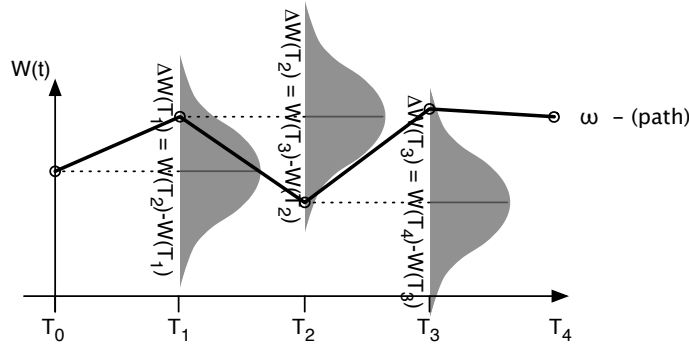


Figure 2.5.: Time discretization of a Brownian motion: The transition $\Delta W(T_i) := W(T_{i+1}) - W(T_i)$ from time T_i to T_{i+1} is normal distributed. The mean of the transition is 0, i.e. under the condition that at time T_i the state $W(T_i) = x^*$ was attained, we have that the (conditional) expectation of $W(T_{i+1})$ is x^* :

$$E(W(T_{i+1}) | W(T_i) = x^*) = x^*.$$

cally distributed random variables (not necessarily normal distributed!) with mean 0 and variance $\sigma^2 > 0$. Define $S_0 := 0$ and $S_k := \sum_{j=1}^k Y_j$. Let X^n denote a stochastic process defined as the (scaled) linear interpolation of the S_k 's at time steps of size $\frac{1}{n}$:

$$X^n(t) := \frac{1}{\sigma \sqrt{n}} ((t \cdot n - [t \cdot n])S_{[t \cdot n]+1} + ([t \cdot n] + 1 - t \cdot n)S_{[t \cdot n]}),$$

where $[x]$ denotes the largest integer number less or equal x .

A path $X^n(\omega)$, $\omega \in \Omega$, defines a continuous map $[0, \infty) \mapsto \mathbb{R}$ and X^n is $(\Omega, \mathcal{F}) \mapsto (C([0, \infty)), \mathcal{B}(C([0, \infty))))$ -measurable⁹. Let P^n denote the image measure of X^n defined on $(C([0, \infty)), \mathcal{B}(C([0, \infty))))$. Then we have:

- $\{P_n\}_{n=1}^\infty$ converges on $(C([0, \infty)), \mathcal{B}(C([0, \infty))))$ to a measure P^* in the weak sense¹⁰.
- The process W defined on $(C([0, \infty)), \mathcal{B}(C([0, \infty))))$ by

$$W(t, \omega) := \omega(t)$$

is a P^* Brownian motion.

⁹ With $C([0, \infty))$ denoting the space of continuous maps $[0, \infty) \mapsto \mathbb{R}$ endowed with the metric of equicontinuous convergence $d(f, g) = \sum_{i=1}^\infty \frac{1}{2^i} \frac{d_n(f, g)}{1 + d_n(f, g)}$ with $d_n(f, g) = \sup_{0 \leq t \leq n} |f(t) - g(t)|$ (then $(C([0, \infty)), d)$ is a complete metric space) and $\mathcal{B}(C([0, \infty)))$ denoting the Borel σ -algebra induced by that metric, i.e. the smallest σ -algebra containing the d -open sets.

¹⁰ A sequence of probability measures $\{P_n\}_{n=1}^\infty$ converges in the weak sense to a measure P^* , if $\int f dP_n \rightarrow \int f dP^*$ for all continuous bounded maps $f : \Omega \mapsto \mathbb{R}$.

Proof: See [18] §2. □

Definition 26 (Wiener Measure): ↗

The measure P^* from Theorem 25 is called *Wiener measure*. ↘

Definition 27 (Canonical Setup): ↗

The space

$$(C([0, \infty)), \mathcal{B}(C([0, \infty))), P^*)$$

(as defined in Theorem 25) is called the *canonical setup* for a Brownian motion W defined by $W(t, \omega) := \omega(t)$, $\omega \in C([0, \infty))$. ↘

Remark 28: A more detailed discussion of Theorem 25 may be found in [18]. A less formal discussion of properties of the Brownian motion may be found in [12].

2.6. Itô Calculus



Motivation: The Brownian motion W is our first encounter with an important continuous stochastic process. The Brownian motion may be viewed as the limit of a scaled *random walk*.¹¹ If we interpret the Brownian motion W in this sense as a model for the movement of a particle, then $W(T)$ denotes the position of the particle at time T and $W(T + \Delta T) - W(T)$ the position change that occurs from T to $T + \Delta T$; to be precise $W(T)$ models the probability distribution of the particle position.

The model of a Brownian motion is that position changes are normal distributed with mean 0 and standard deviation $\sqrt{\Delta T}$. Requiring mean 0 corresponds to requiring that the position change has no directional preference. The standard deviation $\sqrt{\Delta T}$ is, apart from a constant which we assume to be 1, a consequence of the requirement that position changes are independent from the position and time at which they occur.

To motivate the class of Itô processes we consider the Brownian motion at discrete times $0 = T_0 < T_1 < \dots < T_N$. The random variable $W(T_i)$ (position of the particle) may be expressed through the increments $\Delta W(T_i) := W(T_{i+1}) - W(T_i)$:

$$W(T_i) = \sum_{j=0}^{i-1} \Delta W(T_j).$$

Using the increments $\Delta W(T_j)$ we may define a whole family of discrete stochastic processes. We give an step by step introduction and use the illustrative interpretation of a particle movement: First we assume that the particle may lose energy over time

¹¹ In a (one dimensional) random walk a particle changes position at discrete time steps by a (constant) distance (say 1) in either direction with equal probability. In other words we have binomial distributed Y_i in Theorem 25.

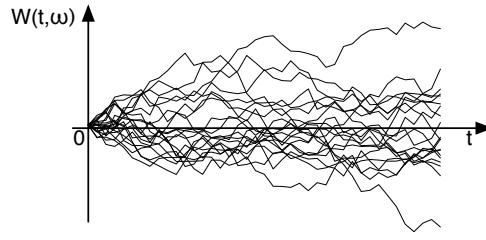


Figure 2.6.: Paths of a (discretization of a) Brownian motion

(for example). Then the increments may be still normal distributed but their standard deviation no longer will be $\sqrt{T_{j+1} - T_j}$. Instead it might be a time-dependent scaling thereof, e.g. $e^{-T_j} \cdot \sqrt{T_{j+1} - T_j}$ where the standard deviation decays exponentially. Multiplying the increments $\Delta W(T_j)$ by a factor gives normal distributed increments with arbitrary standard deviations. Thus we consider a process of the form

$$X(T_i) = \sum_{j=0}^{i-1} \sigma(T_j) \Delta W(T_j),$$

where in our example we would use $\sigma(T_j) := e^{-T_j}$.

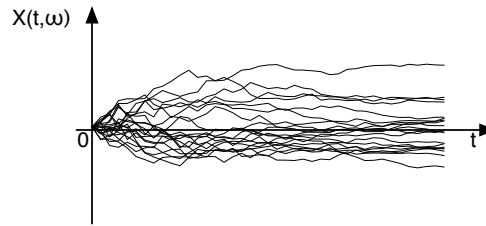


Figure 2.7.: Paths of a (discretization of a) Brownian motions with time dependent instantaneous volatility

Next we consider the case where the particle has a preference for a certain direction, i.e. a *drift*. This is modeled by increments having mean different from zero. The addition of a constant μ to a normal distributed random variable with mean zero will give a normal distributed random variable with mean μ . We want μ to be the drift per time unit and allow that μ may change over time. Thus we add $\mu(T_j) \cdot (T_{j+1} - T_j)$ to the corresponding increment over period T_j to T_{j+1} . If we also consider the starting point to be random, modeled by a random variable $X(0)$ we then consider processes of

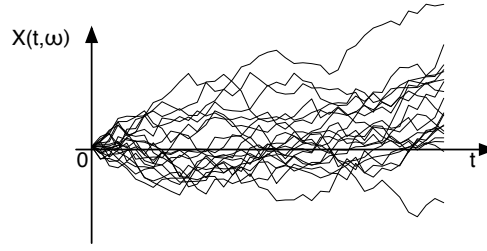


Figure 2.8.: Paths of a (discretization of a) Brownian motion with a drift

the form

$$X(T_i) = X(0) + \sum_{j=0}^{i-1} \underbrace{\mu(T_j) \underbrace{(T_{j+1} - T_j)}_{=: \Delta T_j} + \sigma(T_j) \Delta W(T_j)}_{\text{Normal distributed with mean } \mu(T_j) \Delta T_j \text{ and standard deviation } \sigma(T_j) \sqrt{\Delta T_j}},$$

i.e.

$$X(T_{j+1}) - X(T_j) = \mu(T_j) \underbrace{(T_{j+1} - T_j)}_{=: \Delta T_j} + \sigma(T_j) \Delta W(T_j).$$

Our next generalization of the process is that the parameter $\mu(T_j)$ and $\sigma(T_j)$ could depend on the paths, i.e. are assumed to be random variables. This might appear odd since then one could create *any* time discrete stochastic process.¹² However it would make sense to allow that the parameters $\mu(T_j)$ and/or $\sigma(T_j)$ (used in the increment from $X(T_j)$ to $X(T_{j+1})$) to depend on the current state of $X(T_j)$ as in

$$\underbrace{X(T_{j+1}) - X(T_j)}_{=: \Delta X(T_j)} = X(T_j) \underbrace{(T_{j+1} - T_j)}_{=: \Delta T_j} + \sigma(T_j) \Delta W(T_j).$$

Here we would have $\mu(T_j) = \mu(T_j, X(T_j)) = X(T_j)$, i.e. a drift that is a random variable. It is an important fact that the drift for the increment from T_j to T_{j+1} is known in T_j . More general we allow μ and σ to be stochastic processes if they are $\{\mathcal{F}_t\}$ -adapted.¹³

¹² If $T_i \mapsto S(T_i)$ is an arbitrary time discrete stochastic process we set $\sigma(T_j) := 0$ and $\mu(T_j) := (S(T_{j+1}) - S(T_j)) / (T_{j+1} - T_j)$ and have $X(T_i) = S(T_i)$.

¹³ The increment $\Delta W(T_j)$ is not $\mathcal{F}_{T_j}$ -measurable. It is only $\mathcal{F}_{T_{j+1}}$ -measurable. The requirements that $\mu(T_j)$ is $\mathcal{F}_{T_j}$ -measurable excludes the example in footnote 12.

The continuous analog to the time discrete processes considered above are Itô processes, i.e. processes of the form

$$X(T) = X(0) + \int_0^T \mu(t) dt + \int_0^T \sigma(t) dW(t), \quad (2.3)$$

i.e.

$$X(T) - X(0) = \int_0^T \mu(t) dt + \int_0^T \sigma(t) dW(t)$$

and as a short hand we will write

$$dX(t) = \mu(t)dt + \sigma(t)dW(t).$$

While the dt part of (2.3) may be (and will be) understood pathwise as a Lebesgue (or even Riemann) integral, we need to define the $dW(t)$ part (as we will see), the *Itô Integral*.¹⁴ ◁|

2.6.1. Itô Integral

In this section we define the Itô integral $\int_S^T f(t, \omega) dW(t, \omega)$. We do not present the mathematical theory in full detail. For a more detailed discussion of the Itô integral see, e.g. [12, 25, 18, 19]).

Definition 29 (The Filtration $\{\mathcal{F}_t\}$ generated by W): ▮

Let $(\Omega, \mathcal{F}, P)$ denote a probability space and $W(t)$ a Brownian motion defined on $(\Omega, \mathcal{F}, P)$ (e.g. by the [canonical setup](#)). We define $\mathcal{F}_t$ as the σ -algebra generated by $W(s)$, $s \leq t$, i.e. the smallest σ -algebra, which contains sets of the form

$$\{\omega; W(t_1, \omega) \in F_1, \dots, W(t_k, \omega) \in F_k, \} = \bigcap_{i=1}^k W(t_i)^{-1}(F_i)$$

for arbitrary $t_i < t$ and $F_i \subset \mathbb{R}$, $F_i \in \mathcal{B}(\mathbb{R})$ ($j \leq k$) and arbitrary $k \in \mathbb{N}$. Furthermore we assume that all sets of measure zero belong to $\mathcal{F}_t$. Then $\{\mathcal{F}_t\}$ is a filtration which we call the *filtration generated by W* . ┘

Remark 30: W is a $\{\mathcal{F}_t\}$ -adapted process.

¹⁴ The $dW(t)$ part may not be interpreted as a Lebesgue-Stieltjes integral through $\sum f(\tau_j)(W(t_{j+1}) - W(t_j))$, $\tau_j \in [t_j, t_{j+1}]$, since $t \mapsto W(t, \omega)$ is not of bounded variation. Thus the limit will depend on the specific choice of $\tau_j \in [t_j, t_{j+1}]$, see Exercise 7.

Definition 31 (Itô integral for elementary processes): ↗

A stochastic process ϕ is called elementary, if

$$\phi(t, \omega) = \sum_{j \in \mathbb{N} \cup \{0\}} e_j(\omega) \cdot 1_{(t_j, t_{j+1}]}(t),$$

where $\{t_j \mid j \in \mathbb{N} \cup \{0\}\}$ is a strictly monotone sequence in $[0, \infty)$ with $t_0 := 0$ and $\{e_j \mid j \in \mathbb{N} \cup \{0\}\}$ a sequence of $\mathcal{F}_{t_j}$ -measurable random variables and $1_{(t_j, t_{j+1}]}$ denotes the indicator function¹⁵ ¹⁶

For an elementary process we define the Itô integral as

$$\begin{aligned} \int_0^\infty \phi(t, \omega) dW(t) &:= \sum_{j \in \mathbb{N} \cup \{0\}} e_j(\omega) \cdot (W(t_{j+1}, \omega) - W(t_j, \omega)), \\ \int_S^T \phi(t, \omega) dW(t) &:= \int_0^\infty \phi(t, \omega) \cdot 1_{(S, T]}(t) dW(t). \end{aligned}$$

Remark 32 (on the one sided continuity of the integrand): In some textbooks (e.g. [25]) the elementary integral is defined using the indicator function $1_{[t_j, t_{j+1})}$ in place of $1_{(t_j, t_{j+1}]}$. For continuous integrators, as we consider here ($W(t)$), it makes no difference which variant we use. However if jump processes are considered (see e.g. [27]) and also with respect to the interpretation of the integral as trading strategy (see page 73) our definition is the better suited.

Lemma 33 (Itô Isometry): Let ϕ denote an elementary process such that $\phi(\cdot, \omega)$ is bounded. Then we have

$$\mathbb{E}[(\int_S^T \phi(t, \cdot) dW(t))^2] = \mathbb{E}[\int_S^T \phi(t, \cdot)^2 dt].$$

Definition 34 (Itô Integral): ↗

The class of integrands of the Itô integral is defined as the set of maps

$$f : [0, \infty) \times \Omega \mapsto \mathbb{R},$$

for which

1. f is a $\mathcal{B} \times \mathcal{F}$ -measurable map,
2. f is a $\mathcal{F}_t$ -adapted process,
3. f is P -most sure of finite quadratic variation, i.e. we have $\mathbb{E}[\int_S^T f(t, \omega)^2 dt] < \infty$.

¹⁵ We define $1_{(t_j, t_{j+1}]}(t) = 1$, if $t \in (t_j, t_{j+1}]$ and = 0 else.

¹⁶ Note that by this definition every path is elementary in the sense of Definition 7.

If f belongs to this class then there exists an approximating sequence $\{\phi_n\}$ of elementary processes with

$$\mathbb{E}^P \left[\int_S^T (f(t, \omega) - \phi_n(t, \omega))^2 dt \right] \rightarrow 0 \quad (\text{for } n \rightarrow \infty),$$

and the *Itô integral* is defined as the (unique) L_2 limit

$$\int_S^T f(t, \omega) dW(t) := \lim_{n \rightarrow \infty} \int_S^T \phi_n(t, \omega) dW(t).$$

┘

Remark 35: For a proof of the statements made in this definition (e.g. the existence and uniqueness of the limit) see [25].

2.6.2. Itô Process

Definition 36 (Itô Process):

┘

Let σ denote a stochastic process belonging to the class of integrands of the Itô integrals (see 34) with

$$P \left(\int_0^t \sigma(\tau, \omega)^2 d\tau < \infty \quad \forall t \geq 0 \right) = 1$$

and μ an $\{\mathcal{F}_t\}$ -adapted process with

$$P \left(\int_0^t |\mu(\tau, \omega)| d\tau < \infty \quad \forall t \geq 0 \right) = 1.$$

Then the process X defined through

$$X(t, \omega) = X(0, \omega) + \int_0^t \mu(s, \omega) ds + \int_0^t \sigma(s, \omega) dW(s, \omega),$$

where $X(0, \cdot)$ is $(\mathcal{F}_0, \mathcal{B}(\mathbb{R}))$ -measurable is called *Itô process* (Remark: X is $\mathcal{F}_t$ -adapted).

┘

This definition is generalized by the m -dimensional Brownian motion as

Definition 37 (Itô Process (m -factorial, n -dimensional))¹⁷:

┘

Let $W = (W_1, \dots, W_m)^\top$ denote an m -dimensional Brownian motion defined on $(\Omega, \mathcal{F}, P)$. Let $\sigma_{i,j}$ ($i = 1, \dots, n$, $j = 1, \dots, m$) denote stochastic processes belonging to the class of integrands of the Itô integral (compare Definition 34) with

$$P \left(\int_0^t \sigma_{i,j}(\tau, \omega)^2 d\tau < \infty \quad \forall t \geq 0 \right) = 1 \quad i = 1, \dots, n, \quad j = 1, \dots, m,$$

¹⁷ Compare [25], §4.2

and μ_i ($i = 1, \dots, n$) an $\{\mathcal{F}_t\}$ -adapted process with

$$P\left(\int_0^t |\mu_i(\tau, \omega)| d\tau < \infty \quad \forall t \geq 0\right) = 1 \quad i = 1, \dots, n.$$

Then the (n -dimensional) process $X = (X_1, \dots, X_n)^\top$ with $X(0, \cdot)$ being $(\mathcal{F}_0, \mathcal{B}(\mathbb{R}^n))$ -measurable and

$$X_i(t, \omega) = X_i(0, \omega) + \int_0^t \mu_i(s, \omega) ds + \sum_{j=1}^m \int_0^t \sigma_{i,j}(s, \omega) dW_j(s, \omega) \quad (i = 1, \dots, n),$$

is called (n -dimensional) (m -factorial) Itô process.¹⁸ We will write X in the shorter matrix notation as

$$X(t, \omega) = X(0, \omega) + \int_0^t \mu(s, \omega) ds + \int_0^t \sigma(s, \omega) dW(s, \omega),$$

with

$$X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \quad \mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} \quad \sigma = \begin{pmatrix} \sigma_{1,1} & \dots & \sigma_{1,m} \\ \vdots & & \vdots \\ \sigma_{n,1} & \dots & \sigma_{n,m} \end{pmatrix}.$$

┘

Remark 38 (Itô Process, differential notation): For an Itô process

$$X(T, \omega) = X(0, \omega) + \int_0^T \mu(t, \omega) dt + \int_0^T \sigma(t, \omega) dW(t, \omega),$$

as defined by Definition 36 and 37 we will use the shortened notation

$$dX(t, \omega) = \mu(t, \omega) dt + \sigma(t, \omega) dW(t, \omega).$$

The stochastic integral

$$\int_{t_1}^{t_2} X(t) dY(t)$$

was defined pathwise for integrators Y that are Brownian motions ($dY = dW$) and for integrands X belonging to the class of integrands of the Itô integral. We extend the definition of the stochastic integral to integrators that are Itô processes:

Definition 39 (Integral with Itô Process as Integrator):

┘

Let Y denote an Itô process of the form

$$dY = \mu dt + \sigma dW$$

¹⁸ The dimension n denotes the dimension of the image space. The factor dimension m denotes the number of (independent) Brownian motions needed to construct the process.

and X a stochastic process such that $X(t)\mu(t)$ is integrable with respect to t and $X(t)\sigma(t)$ is integrable with respect to $W(t)$. Then we define

$$\int_{t_1}^{t_2} X(t) dY(t) := \int_{t_1}^{t_2} X(t)\mu(t) dt + \int_{t_1}^{t_2} X(t)\sigma(t) dW(t).$$

As in remark 38 we will write

$$X(t) dY(t) := X(t)\mu(t) dt + X(t)\sigma(t) dW(t).$$

J

2.6.3. Itô Lemma and Product Rule

Theorem 40 (Itô Lemma (1 D)¹⁹): Let X denote an Itô process with

$$dX(t) = \mu dt + \sigma dW.$$

Let $g(t, x) \in C^2([0, \infty] \times \mathbb{R})$. Then we have that

$$Y(t) := g(t, X(t))$$

is an Itô process with

$$dY = \frac{\partial g}{\partial t}(t, X(t))dt + \frac{\partial g}{\partial x}(t, X(t))dX + \frac{1}{2} \frac{\partial^2 g}{\partial x^2}(t, X(t)) \cdot (dX)^2, \quad (2.4)$$

where $(dX)^2 = (dX) \cdot (dX)$ is given by formal expansion with

$$\begin{aligned} dt \cdot dt &= 0, & dt \cdot dW &= 0, \\ dW \cdot dt &= 0, & dW \cdot dW &= dt, \end{aligned}$$

i.e.

$$\begin{aligned} (dX)^2 &= (dX) \cdot (dX) = (\mu dt + \sigma dW) \cdot (\mu dt + \sigma dW) \\ &= \mu^2 \cdot dt \cdot dt + \mu \cdot \sigma \cdot dt \cdot dW + \mu \cdot \sigma \cdot dW \cdot dt + \sigma^2 \cdot dW \cdot dW = \sigma^2 dt \end{aligned}$$

Theorem 41 (Itô Lemma²⁰): Let X denote an n -dimensional, m -factorial Itô process with

$$dX(t) = \mu dt + \sigma dW.$$

¹⁹ Compare [25] §4.1.

²⁰ Compare [25] §4.2.

Let $g(t, x) \in C^2([0, \infty] \times \mathbb{R}^n; \mathbb{R}^d)$, $g = (g_1, \dots, g_d)^\top$. Then we have that

$$Y(t) := g(t, X(t))$$

is an d -dimensional, n -factorial Itô process with

$$\begin{aligned} dY_k(t) &= \frac{\partial g_k}{\partial t}(t, X(t))dt + \sum_{i=1}^n \frac{\partial g_k}{\partial x_i}(t, X(t))dX_i(t) \\ &\quad + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 g_k}{\partial x_i \partial x_j}(t, X(t)) \cdot (dX_i(t) \cdot dX_j(t)), \end{aligned}$$

where $dX_i(t)dX_j(t)$ is given by formal expansion with

$$\begin{aligned} dt \cdot dt &= 0, & dt \cdot dW_j &= 0, \\ dW_i \cdot dt &= 0, & dW_i \cdot dW_j &= \begin{cases} dt & i = j \\ 0 & i \neq j \end{cases}. \end{aligned}$$

Theorem 42 (Product Rule): Let X , Y and $X_1, \dots, X_n$ denote Itô processes. Then we have

1. $d(X \cdot Y) = Y \cdot dX + X \cdot dY + dX \cdot dY$
2. $d\left(\prod_{i=1}^N X_i\right) = \sum_{i=1}^N \prod_{\substack{k=1 \\ k \neq i}}^N X_k \cdot dX_i + \sum_{\substack{i,j=1 \\ j>i}}^N \prod_{\substack{k=1 \\ k \neq i,j}}^N X_k \cdot dX_i dX_j$

Proof: We prove only 1 since 2 follows from 1 by induction. We apply the Itô lemma to the map

$$g : \mathbb{R} \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}, \quad g(t, x, y) := x \cdot y.$$

We have

$$\begin{aligned} \frac{\partial g}{\partial t} &= 0 & \frac{\partial g}{\partial x} &= y & \frac{\partial g}{\partial y} &= x \\ \frac{\partial g^2}{\partial x^2} &= 0 & \frac{\partial g^2}{\partial y^2} &= 0 \\ \frac{\partial g^2}{\partial xy} &= 1 & \frac{\partial g^2}{\partial yx} &= 1 \end{aligned}$$

and thus

$$d(X \cdot Y) = d(g(X, Y)) = Y \cdot dX + X \cdot dY + dX \cdot dY.$$

□

Theorem 43 (Quotient Rule): Let X and Y and $Y_1, \dots, Y_n$ denote Itô processes where, $Y > c$ for a given $c \in \mathbb{R}$. Then we have

1. $d\left(\frac{X}{Y}\right) = \frac{X}{Y} \left(\frac{dX}{X} - \frac{dY}{Y} - \frac{dX}{X} \cdot \frac{dY}{Y} + \left(\frac{dY}{Y}\right)^2 \right)$
2. $d\left(\frac{1}{Y}\right) = -\frac{1}{Y^2}dY + \frac{1}{Y^3}dYdY$
3. $d\left(\prod_{i=1}^N \frac{1}{Y_i}\right) = \prod_{i=1}^N \frac{1}{Y_i} \cdot \sum_{i=1}^N \left(\frac{-dY_i}{Y_i} + \sum_{j \geq i}^N \frac{dY_i}{Y_i} \frac{dY_j}{Y_j} \right)$

Lemma 44 (Drift Adjustment of Lognormal Process): Let $S(t) > 0$ denote an Itô process of the form

$$dS(t) = \mu(t)S(t)dt + \sigma(t)S(t)dW(t),$$

and $Y(t) := \log(S(t))$. Then we have

$$dY(t) = (\mu(t) - \frac{1}{2}\sigma^2(t))dt + \sigma(t)dW(t).$$

Proof: Exercise. □

 **Interpretation:** The Itô lemma and its implications like Lemma 44 may appear unfamiliar. They state that a non-linear function of a stochastic process will induce a drift of the mean. This may be seen in an elementary example: Consider the time-discrete stochastic process $X(t_i)$ constructed from binomial distributed increments $\Delta B(t_i)$ (instead of Brownian increments $dW(t)$), where $\Delta B(t_i)$ are independent and attain with probability $p = \frac{1}{2}$ the value $+1$ or -1 , respectively. Assuming $X(t_0) = 10$ we draw the process

$$X(t_{i+1}) = X(t_i) + \Delta B(t_i)$$

in Figure 2.9 (left), i.e. $\Delta X(t_i) = \Delta B(t_i)$. This process does not exhibit a drift. We have

$$X(t_i) = E(X(t_{i+1}) | \mathcal{F}_{t_i}).$$

In other words: In each node in Figure 2.9 the process X attains the mean of the values from the two child nodes.

As in Figure 2.9 (right) we then consider the process $Y(t_i) = f(X(t_i)) = X(t_i)^2$. This process exhibits in each time step a drift of the mean of $+1$. One can easily check that the increments of the process Y are given by

$$Y(t_{i+1}) = Y(t_i) + 1 + 2 \cdot X(t_i) \cdot \Delta B(t_i)$$

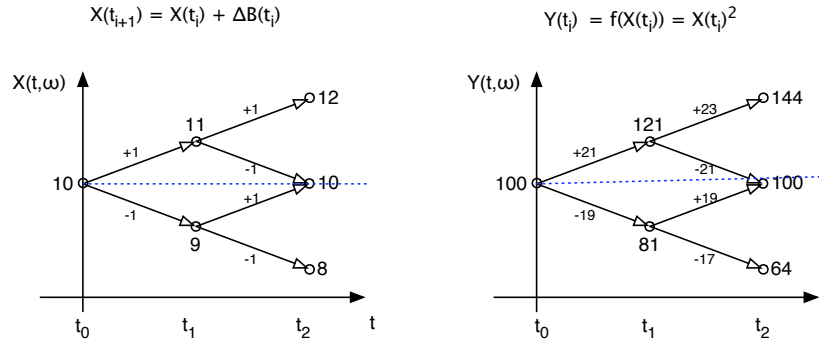


Figure 2.9.: Non-linear functions of stochastic processes induce a drift to the mean.

(check this in Figure 2.9). This corresponds to the result stated by Itô's lemma (see Theorem 40 with $g(t, x) = f(x) = x^2$). Indeed we have

$$\begin{aligned}
 Y(t_{i+1}) &= (X(t_{i+1}))^2 = (X(t_i) + \Delta X(t_i))^2 \\
 &= X(t_i)^2 + \Delta X(t_i)^2 + 2 \cdot X(t_i) \cdot \Delta X(t_i) \\
 &= Y(t_i) + \Delta B(t_i)^2 + 2 \cdot X(t_i) \cdot \Delta B(t_i) \\
 &= Y(t_i) + 1 + 2 \cdot X(t_i) \cdot \Delta B(t_i) \\
 &= Y(t_i) + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(X(t_i)) \cdot (\Delta X(t_i))^2 + \frac{\partial f}{\partial x}(X(t_i)) \Delta X(t_i).
 \end{aligned}$$

Obviously we may interpret Itô formula (2.4) in Itô's lemma as a (formal) Taylor expansion of $g(X + dX)$ up to the order $(dX)^2$. For the continuous case the higher order increments are (almost surely) 0. For the discrete case this is not the case. E.g. consider $(X(t_i) + \Delta X(t_i))^3$ in the example above. ◁|

2.7. Brownian Motion with Instantaneous Correlation

In Definition 23 Brownian motion was defined through normal distributed increments $W(t) - W(s)$, $t > s$ having covariance matrix $(t - s)I_n$. In other words we have that for $W = (W_1, \dots, W_n)$ the components are one dimensional Brownian motions with pairwise independent increments, i.e. for $i \neq j$ we have that $W_i(t) - W_i(s)$ and $W_j(t) - W_j(s)$ are independent (thus uncorrelated).

We define the *Brownian motion with instantaneously correlated increments* as a special Itô process:

Definition 45 (Brownian Motion with instantaneous correlated Increments): ▮

Let U denote an m -dimensionale Brownian motion as defined in Definition 23. Let

$f_{i,j}$ ($i = 1, \dots, n$, $j = 1, \dots, m$) denote stochastic processes belonging to the class of integrands of the Itô integral (see Definition 34 and Definition 37) with

$$P\left(\int_0^t f_{i,j}(\tau, \omega)^2 d\tau < \infty \quad \forall t \geq 0\right) = 1 \quad i = 1, \dots, n, j = 1, \dots, m,$$

furthermore let

$$F(t) := (f_1, \dots, f_m) := \begin{pmatrix} f_{1,1}(t) & \dots & f_{1,m}(t) \\ \vdots & & \vdots \\ f_{n,1}(t) & \dots & f_{n,m}(t) \end{pmatrix}$$

denote an $n \times m$ matrix with

$$\sum_{j=1}^m f_{i,j}^2(t) = 1, \quad \forall i = 1, \dots, n.$$

Then the Itô process

$$dW(t) = F(t) \cdot dU(t), \quad W(0) = 0$$

(see Definition 37) is called *m-factorial, n-dimensional Brownian motion with factors* f_j , $j = 1, \dots, m$. With $R := FF^T$ we call R the *instantaneous correlation* and W a *Brownian motion with instantaneous correlation* R . F is called the *factor matrix*. $\square$



Interpretation: For simplicity let us consider a constant matrix $F = (f_1, \dots, f_m)$. Then we have for time discrete increments

$$\Delta W(T_k) = F \cdot \Delta U(T_k) = f_1 \Delta U_1(T_k) + \dots + f_m \Delta U_m(T_k). \quad (2.5)$$

Note that $\Delta U_i(T_k)$ and $\Delta U_j(T_k)$ are independent. They may be interpreted as independent scenarios. If ω is a path with $\Delta U_i(T_k; \omega) \neq 0$ and $\Delta U_j(T_k; \omega) = 0$ for $j \neq i$, then we have from (2.5) that $\Delta W(T_k; \omega) = f_i \cdot \Delta U_i(T_k; \omega)$, i.e. on the path ω the vector W will receive increments corresponding to the scenario f_i (multiplied by the amplitude $\Delta U_i(T_k; \omega)$). If, for example, $f_1 = (1, \dots, 1)^T$ then the scenario corresponds to a parallel shift of W (by the shift size $\Delta U_1(T_k)$).

Our definition of a factor matrix does not allow arbitrary scenarios since we require that $\sum_{j=1}^m f_{i,j}^2(t) = 1$, i.e. that $R := FF^T$ is a correlation matrix. By this assumption we ensure that the components of W_i of W are one dimensional Brownian motions in the sense of Definition 23.

By means of the factor matrix F we may interpret the implied correlation structure R in a geometrical way. The calculation of F from a given R is a Cholesky decomposition.

We will make use of this construction in the modeling of interest rate curves (Chapter 17: [LIBOR Market Model](#)). Here the interpretation of the factors is given by movements of the interest rate curve. The possible shapes of an interest rate curve will then be investigated (Chapter 22)). The question on how to obtain a set of factors or reduce a given set of factors to the relevant ones is discussed in Appendix A.4 and A.5.

<|

2.8. Martingales

Definition 46 (Martingale):

⌋

The stochastic process $\{X(t), \mathcal{F}_t; 0 \leq t < \infty\}$ is called a *martingale* with respect to the filtration $\{\mathcal{F}_t\}$ and the measure P if

$$X_s = E(X(t) | \mathcal{F}_s) \quad P \text{ almost sure}, \quad \forall 0 \leq s < t < \infty. \quad (2.6)$$

If (2.6) holds for $\leq$ in place of $=$, then X is called sub-martingale. If (2.6) holds for $\geq$ in place of $=$, then X is called super-martingale. ⌋

Lemma 47 (Martingale Itô-Processes are drift free): Let X denote an Itô process of the form

$$dX = \mu dt + \sigma dW \quad \text{under } P$$

with $E^P\left(\int_0^T \sigma^2(t) dt\right)^{1/2} < \infty$. Then we have

$$X \text{ is a } P \text{ martingale} \Leftrightarrow \mu = 0 \quad (\text{i.e. } X \text{ is drift free}).$$

2.8.1. Martingale Representation Theorem

Theorem 48 (Martingale Representation Theorem²¹): Let $W(t) = (W_1(t), \dots, W_m(t))^T$ denote an m -dimensional Brownian motion, $\mathcal{F}_t$ the corresponding filtration. Let $M(t)$ denote a martingale with respect to $\mathcal{F}_t$ with $\int_0^t |M(t)|^2 dP < \infty$ ($\forall t \geq 0$).

Then there exists a stochastic process g (with $g(t)$ belonging to the class of integrands of the Itô integral) with

$$M(t) = M(0) + \int_0^t g(s) dW(s) \quad P \text{ almost sure}, \quad \forall t \geq 0.$$

²¹ See [25], §4.3 and [18], Theorem 4.15.

2.9. Change of Measure (Girsanov, Cameron, Martin)

Definition 49 (Measure with Density):

Let $(\Omega, \mathcal{F}, P)$ denote a measure space and ϕ a P integrable non-negative real valued random variable. Then

$$Q(A) := \int_A \phi \, dP$$

defines a measure on $(\Omega, \mathcal{F})$ which we call *measure with density ϕ with respect to P* .

Definition 50 (Equivalent Measure):

Let P and Q denote two measures on the same measurable space $(\Omega, \mathcal{F})$.

1. Q is called *continuous with respect to P* $\Leftrightarrow (P(A) = 0 \Rightarrow Q(A) = 0 \, \forall A \in \mathcal{F})$.
2. P and Q are called *equivalent* $\Leftrightarrow (P(A) = 0 \Leftrightarrow Q(A) = 0 \, \forall A \in \mathcal{F})$.

Theorem 51 (Radon-Nikodým Density): Let P and Q denote two measures on a measure space $(\Omega, \mathcal{F})$. Then we have

$$Q \text{ is continuous with respect to } P \Leftrightarrow Q \text{ has a density with respect to } P.$$

Proof: See [1].

Definition 52 (Radon-Nikodým Density):

If Q is continuous with respect to P then we call the density of Q with respect to P the *Radon-Nikodým density* and denote it by $\frac{dQ}{dP}$.

Theorem 53 (Change of Measure (Girsanov, Cameron, Martin)): Let W denote a (d -dimensional) P -Brownian motion and $\{\mathcal{F}_t\}$ the filtration generated by W . Let Q denote a measure equivalent to P (w.r.t. $\{\mathcal{F}_t\}$).

1. Then there exists a $\{\mathcal{F}_t\}$ -previsible process C with

$$\left. \frac{dQ}{dP} \right|_{\mathcal{F}_t} = \exp \left(\int_0^t C(s) \, dW(s) - \frac{1}{2} \int_0^t |C(s)|^2 \, ds \right). \quad (2.7)$$

2. Reversed, if ρ denotes a strictly positive P -martingale with respect to $\{\mathcal{F}_t \mid t \in [0, T]\}$ for a fixed T with $E^P(\rho) = 1$, then ρ has the representation (2.7) and defines (as Radon-Nikodým process) a measure $Q = Q^T$ which is equivalent to P with respect to $\mathcal{F}_T$.

In any case we have that

$$\tilde{W}(t) := W(t) - \int_0^t C(s)ds \quad (2.8)$$

is a Q -Brownian motion (with respect to $\{\mathcal{F}_t\}$) (and for $t \leq T$ in the second case).

Remark 54 (Change of Measure): Equation (2.8) may be written as

$$d\tilde{W}(t) = -C(t)dt + dW(t) \quad (2.9)$$

and we see that the *change of measure* (2.7) corresponds to a *change of drift* (2.9). The second case has a restriction on a finite time horizon T , which will be irrelevant in the following applications.

Exercise: (*Change of Measure* in a binomial tree): Calculate the probabilities (the measure) such that the process Y depicted in Figure 2.9 is a martingale.



Interpretation: The form of the change of measure (2.7) may be motivated via a simple calculation and derived for a time discrete Itô process

$$\Delta X(T_i) = \mu(T_i)\Delta T_i + \sigma(T_i)\Delta W(T_i)$$

by elementary calculations.

At first: Let Z denote a normal distributed random variable with mean 0 and standard deviation σ on a probability space $(\Omega, \mathcal{F}, P)$. Under which measure will Z be normal distributed with mean c and standard deviation σ ? The density of a normal distributed random variable with mean c and standard deviation σ is $z \mapsto \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{(z-c)^2}{2\sigma^2})$. Thus we seek a change of measure $\frac{dQ}{dP}$ such that

$$\underbrace{\frac{1}{2\sqrt{\pi}\sigma} \exp(-\frac{(z-c)^2}{2\sigma^2})}_{\frac{dQ}{dz} = \text{desired density under } Q} dz = dQ = \frac{dQ}{dP} dP = \underbrace{\frac{1}{2\sqrt{\pi}\sigma} \exp(-\frac{z^2}{2\sigma^2})}_{\frac{dP}{dz} = \text{known density under } P} \underbrace{\frac{dQ}{dP}}_{=?} dz.$$

With

$$\exp(-\frac{(z-c)^2}{2\sigma^2}) = \exp(-\frac{z^2 - 2cz + c^2}{2\sigma^2}) = \exp(-\frac{z^2}{2\sigma^2}) \exp(\frac{cz - \frac{1}{2}c^2}{\sigma^2})$$

it follows that the desired change of measure is

$$\frac{dQ}{dP} = \exp(\frac{cz - \frac{1}{2}c^2}{\sigma^2}). \quad (2.10)$$

This corresponds to the term in (2.7).

To illustrate this we consider the time discrete process

$$\Delta X(T_i) = \mu^P(T_i)\Delta T_i + \sigma(T_i)\Delta W(T_i) \quad \text{under } P.$$

Under which measure is X a (time discrete) martingale? We have:

$$\begin{aligned} X \text{ is a } Q\text{-martingale} &\Leftrightarrow \mu^Q(T_i) = 0 \Leftrightarrow E^Q(\Delta X(T_i) | \mathcal{F}_{T_i}) = 0 \\ &\Leftrightarrow E^Q\left(\frac{1}{\sigma(T_i)}\Delta X(T_i) | \mathcal{F}_{T_i}\right) = 0 \end{aligned}$$

First consider a single increment: The random variable

$$\frac{1}{\sigma(T_i)}\Delta X(T_i) = \frac{\mu(T_i)}{\sigma(T_i)} \cdot \Delta T_i + \Delta W(T_i)$$

is (under P) normal distributed with mean $\frac{\mu(T_i)}{\sigma(T_i)} \cdot \Delta T_i$ and standard deviation $\sqrt{\Delta T_i}$. For the conditional (!) expectation under Q we have

$$E^Q\left(\frac{1}{\sigma(T_i)}\Delta X(T_i) | \mathcal{F}_{T_i}\right) = \frac{\mu(T_i)}{\sigma(T_i)} \cdot \Delta T_i + E^Q(\Delta W(T_i) | \mathcal{F}_{T_i})$$

Then $E^Q(\frac{1}{\sigma(T_i)}\Delta X(T_i) | \mathcal{F}_{T_i}) = 0$ if $E^Q(\Delta W(T_i) | \mathcal{F}_{T_i}) = C(T_i)\Delta T_i$ where $C(T_i) = -\frac{\mu(T_i)}{\sigma(T_i)}$ (we correct for the drift). Given the considerations above we have to apply a change of measure

$$\frac{dQ}{dP}\Big|_{\Delta T_i} := \exp(C(T_i)\Delta W(T_i) - \frac{1}{2}C(T_i)^2\Delta T_i)$$

for the time step ΔT_i (apply (2.10) with $Z = \Delta W(T_i)$, $c = C(T_i)\Delta T_i$ and $\sigma = \sqrt{\Delta T_i}$).

This is just the change of measure needed to make the increment $\Delta X(T_i)$ drift free. Since the increments $\Delta W(T_i)$ are independent we get the change of measure for the process X from T_0 to T_n by multiplying the Radon-Nikodym densities, i.e.

$$\begin{aligned} \frac{dQ}{dP}\Big|_{T_n} &:= \prod_{i=0}^{n-1} \frac{dQ}{dP}\Big|_{\Delta T_i} = \prod_{i=0}^{n-1} \exp(C(T_i)\Delta W(T_i) - \frac{1}{2}C(T_i)^2\Delta T_i) \\ &= \exp\left(\sum_{i=0}^{n-1} C(T_i)\Delta W(T_i) - \frac{1}{2}\sum_{i=0}^{n-1} C(T_i)^2\Delta T_i\right). \end{aligned} \tag{2.11}$$

The change of measure $\frac{dQ}{dP}\Big|_{T_n}$ will make all increments $\Delta X(T_i)$ drift free for $i = 0, \dots, n-1$. Due to the independence of the increments $\Delta W(T_i)$ we obtain independence of the $\frac{dQ}{dP}\Big|_{\Delta T_i}$ and thus

$$E^Q(\Delta W(T_i)) = E^P(\Delta W(T_i) \cdot \underbrace{\prod_{j=0}^{n-1} \frac{dQ}{dP}\Big|_{\Delta T_j}}_{=1}) = \prod_{j \neq i} E^P\left(1 \cdot \frac{dQ}{dP}\Big|_{\Delta T_j}\right) \cdot E^P(\Delta W(T_i) \cdot \frac{dQ}{dP}\Big|_{\Delta T_i}).$$

The term (2.11) is a discrete version of (2.7). To some extent we have just proven a version of the change of measure theorem for time discrete Itô processes.

The term $-C(t)^2 dt$ in (2.7) (or $-C(T_i)^2 \Delta T_i$ in (2.11)) may also be motivated as follows: The random variable

$$Z(t) := \frac{dQ}{dP} \Big|_{\mathcal{F}_t}$$

represents a density of the measure $Q|_{\mathcal{F}_t}$ with respect to the measure $P|_{\mathcal{F}_t}$. Since $Q|_{\mathcal{F}_t}$ should be a probability measure we must have

$$Q|_{\mathcal{F}_t}(\Omega) = E^{P|_{\mathcal{F}_t}}(Z(t)) = 1 \quad (2.12)$$

and in this case we must have that Z is a martingale. Thus the *drift correction* $-C(t)^2 dt$ follows from Lemma 44, because Z is a lognormal process. $\triangleleft$

2.10. Stochastic Integration

In the previous sections the following integrals were considered:

Maps:

- $\int_{t_1}^{t_2} f(t) dt$ – Lebesgue or Riemann integral.
Integral of a real valued function with respect to t .

Random Variables:

- $\int_{\Omega} Z(\omega) dP(\omega)$ – Lebesgue integral.
Integral of a random variable Z with respect to a measure P (cf. expectation).

Stochastic Processes:

- $\int_{\Omega} X(t_1, \omega) dP(\omega)$ – Lebesgue integral.
Integral of a random variable $X(t_1)$ with respect to a measure P .
- $\int_{t_1}^{t_2} X(t) dt$ – Lebesgue Integral or Riemann integral.
The (pathwise) integral of the stochastic process X with respect to t .
- $\int_{t_1}^{t_2} X(t) dW(t)$ – Itô integral.
The (pathwise) integral of the stochastic process X with respect to a Brownian motion W .

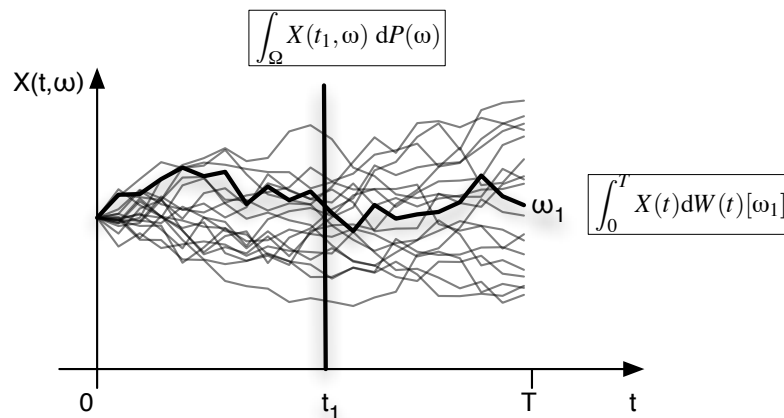


Figure 2.10.: *Integration of stochastic processes*

The notion of a stochastic integral may be extended to more general integrands and/or more general integrators. For completeness we mention:

Definition 55 (Previsible Process): ↗

Let X denote a (real valued) stochastic process on $(\Omega, \mathcal{F})$ and $\{\mathcal{F}_t\}$ a filtration on $(\Omega, \mathcal{F})$. The process X is called $\{\mathcal{F}_t\}$ -previsible, if X is $\{\mathcal{F}_t\}$ -adapted and bounded with left continuous paths. ↘

Definition 56 (Integral with respect to a semi-martingale as integrator²²): ↗

Let Y denote a semi-martingale of the form

$$Y(t) = A(t) + M(t),$$

where $A(t)$ is a process with locally bounded variation and $M(t)$ a local martingale. Let $X(t)$ denote a previsible process. Then we define

$$\int_{t_1}^{t_2} X(t) dY(t) := \int_{t_1}^{t_2} X(t) dA(t) + \int_{t_1}^{t_2} X(t) dM(t).$$
↘

Remark 57 (Stochastic Integral): The class of processes (integrands) for which we may define a stochastic integral depends on the properties of the integrators (and vice versa). For continuous integrators (as the Brownian motion) the integrands merely have to be adapted processes. For more general integrators one has to restrict to a smaller class of integrands, e.g. previsible processes. Compare Example 4.1 and Remark 4.4 in [12]. For more detailed discussion of the stochastic integral see [6], §5.5 and (especially for more general integrators) [12], §4 and [18], §3.



Further Reading: On stochastic processes: As introduction: [25]. For an in depth discussion: [18], [27] and [29]. ◀|

²² For a more detailed discussion and a definition of a semi-martingale see remark 57.

2.11. List of Symbols

The following list of symbols summarizes the most important notions from Chapter 2:

Symbol	Object	Interpretation
ω	element of Ω	State. In the context of stochastic processes: path.
Ω	set	State space.
X	random variable	Map which assigns an event / outcome (e.g. a number) to a state. Example: the payoff of a financial product (this may be interpreted as a snapshot of the financial product itself).
X	stochastic process	Sequence (in time) of random variables (e.g. the evolution of a financial product (could be its payoffs but also its value)).
$X(t), X_t$	stochastic process evaluated at time t ($\equiv$ random variable)	see above
$X(\omega)$	stochastic process evaluated in state ω	Path of X in state ω .
W	Brownian motion	Model for a continuous (random) movement of a particle with independent increments (position changes).
$\mathcal{F}$	σ -algebra (set of sets)	Set of information configurations (set of sets of states).
$\{\mathcal{F}_t \mid t \geq 0\}$ const	filtration	$\mathcal{F}_t$ is the information known at time t .

CHAPTER 3

Replication

Nowadays people know the price of everything
and the value of nothing.

Oscar Wilde
The Picture of Dorian Gray, [36]

3.1. Replication Strategies

3.1.1. Introduction

We motivate the important principle of *replication* by considering the simplest financial derivative, the *forward contract*. Consider the following products:

- A (Forward Contract on Rain):** At time $T_1 > 0$ the amount of rain fallen $R(T_1)$ is measured (in mm) at a predefined place and the \$ amount

$$A(T_1) := (R(T_1) - X) \cdot \frac{\$}{\text{mm}}$$

is paid. Here, X denotes a constant reference amount of rain.

- B (Forward Contract on IBM Stock):** At time $T_1 > 0$ the value $S(T_1)$ of an IBM stock is fixed and the \$ amount

$$B(T_1) := (S(T_1) - X)$$

is paid. Here, X denotes a constant reference value.

These products may be interpreted as an guarantee or insurance.¹ The product A is a *weather derivative*, the product B is an *equity derivative*. What is a fair value for

¹ The product B is a guarantee to buy the stock S at time T_1 for the amount X , since the product B pays the difference required to buy the stock at its value $S(T_1)$. The product A is an insurance against a change in rain quantity, it might be of sense for the operator of an irrigation system supplying water to farmers. In case of a rainy season he suffers from a loss of earnings gets paid back a quantity proportional to the rain fallen.

the product A and the product B? What do we expect to pay in T_0 (today) for such a guarantee?

Consider the product A: It appears that the determination of its fair value requires a exact assessment of the probability of rain at time T_1 . So let $R_{T_1} : \Omega \rightarrow \mathbb{R}$ denote a random variable (rain quantity at time T_1 , modeled over a suitable probability space – for $\omega \in \Omega$ the $R_{T_1}(\omega)$ denotes the quantity of rain that falls in state ω). Then, the random variable $A(T_1) := R_{T_1} - X$ defines the payoff of product A. We wish to determine the value $A(T_0)$ of this product at time T_0 .

For the trivial case of a single-point distribution, i.e.

$$R_{T_1}(\omega) = R^* = \text{const.} \quad \forall \omega \in \Omega,$$

i.e. the amount of rain at time T_1 is R^* with probability 1, in other words: it is for sure that in T_1 the amount of rain falling is R^* . Then the product A pays at time T_1 the amount $A^* := R^* - X$ with probability 1. In this case, the product corresponds to a savings account. Let $P(T_0)$ denote the value that has to be invested at time T_0 (into a savings account) to receive in T_1 (including interest) the amount 1 ($=: P(T_1)$), then the product A pays in T_1 the value A^* -times the value of P . Since A is equivalent to A^* times P , we have

$$A(T_0) = A^* \cdot P(T_0)$$

(we assumed that the interest rate paid is independent of the amount invested, i.e. the interest is proportional to the amount invested). With $P(T_1) = 1$ this may be written as

$$\frac{A(T_0)}{P(T_0)} = \frac{A^*}{P(T_1)}.$$

For the (quite unrealistic) case of a one-point distribution (i.e. a deterministic payment) we may derive the value of the product A by comparing it to the value of another product with deterministic payoff (the savings account).

In the general case, where a probability distribution of $R(T_1)$ is known, it is questionable if we may replace the value $\frac{A^*}{P(T_1)}$ by - e.g. - the expectation –

$$\frac{A(T_0)}{P(T_0)} \stackrel{?}{=} E\left(\frac{A(T_1)}{P(T_1)}\right)$$

– if so, this would imply that the risk (for example the variance of $A(T_1)$) does not influence the value. Since the product A is an insurance against the risk of variations in rain fall, this appears to be non-sense.

The product B is very similar to the product A. Instead of the amount of rain $R(T_1)$ the value of a stock $S(T_1)$ determines the payout. The considerations of the previous section apply accordingly. However, for B it is possible to determine its value in $T_0 = 0$ independent of the specific probability distribution of $S(T_1)$:

In time T_0 we take a credit with fixed interest rate such that the amount to be repaid at time T_1 is X . This credit pay us in T_0 the amount $X \cdot P(T_0)$.² In addition we acquire the stock for a it current market price $S(T_0)$. In total we have to pay in T_0 the amount

$$V(T_0) = S(T_0) - XP(T_0). \quad (3.1)$$

The essential difference between product A and product B is, that for product B the quantity that carries the insecurity (the stock) may be acquired. In other words, it is possible to buy (and sell) “risk”. We assume that we may buy and sell parts of stocks (and other traded product) is arbitrary (real valued) quantities.

It was key for the construction of the replication portfolio that:

- It is possible to express today’s value of a deterministic future payment and thus there is a vehicle to transfer a deterministic future payment to an earlier time: $P(T_1; t)$ is a traded product.
- It is possible to buy or sell the underlying (that carries the risk) at any time in any quantities: $S(t)$ is a traded product.

It was a surprising consequence of the construction of a replication portfolio that:

- The real probabilities do not enter in today value of the replication portfolio.

Our strategy was to construct a portfolio at time T_0 and wait until time T_1 (*buy-and-hold strategy*). Obviously the strategy may be refined by restructuring the portfolio at other times. Dynamic (infinitesimal) restructuring will allow the replication of arbitrary (continuous) payouts, given that the underlying random variables (*the underlyings*) are traded product. It is this condition that prevents the replication of product A: We may not buy or sell the random variable “rain”.³

Consider again the equation

$$\frac{B(T_0)}{P(T_1; T_0)} = E^{\mathbb{P}} \left(\frac{B(T_1)}{P(T_1; T_1)} \right).$$

This equation would reduce the pricing (i.e. the calculation of $B(T_0)$) to the calculation of an expectation. The equation holds, if

$$\frac{S(T_0)}{P(T_1; T_0)} = E^{\mathbb{P}} \left(\frac{S(T_1)}{P(T_1; T_1)} \right).$$

² We denote here by $P(T_0)$ the amount paid by a credit in T_0 , which has to be repaid in T_1 by the amount 1 =: $P(T_1)$ and such a credit is acquired X -times.

³ Under certain conditions it would be possible to replicate the product A: If there would exist a company for which its stock value is perfectly correlated to the amount of rain falling, then the product A may be replicated using stocks of that company. Of course, such a stock will only exist in some approximate sense, however then it might be possible to replicate the product A in an approximate sense.

It cannot be expected that this equation holds in general (if so, only the expectation of $S(T_1)$ would enter into the pricing and we would be ignorant to any risk). However it is possible to change the measure such that the corresponding equation hold, i.e. it exists a measure $\mathbb{Q}$ such that

$$\frac{S(T_0)}{P(T_1; T_0)} = E^{\mathbb{Q}}\left(\frac{S(T_1)}{P(T_1; T_1)}\right).$$

A change of measure is indeed a admissible tool when considering replication portfolios, since - as we have seen - the real probabilities (the measure $\mathbb{P}$) does not enter. To motivate this important concept we consider the following (simple) example:

3.1.2. Replication in a discrete Model

3.1.2.1. Example: two times (T_0, T_1) , two states (ω_1, ω_2) , two assets (S, N)

Let S and N denote stochastic processes defined over a filtered probability space $(\Omega, \mathcal{F}, P, \mathcal{F}_t)$ with $\Omega = \{\omega_1, \omega_2\}$. Here $S(t, \omega)$ denotes the value of a financial product “S” and $N(t, \omega)$ denote the value of a financial product “N”. We consider two points in time T_0 (present) and T_1 (future). We assume that “S” and “N” are traded at these times. Let the filtration be given by $F_{T_0} = \{\emptyset, \Omega\}$ and $F_{T_1} = \{\emptyset, \{\omega_1\}, \{\omega_2\}, \Omega\}$, i.e. in T_0 it is not possible to decide which state ω_1, ω_2 we are in, but in T_1 this information is known. Assume that the processes S and N are $\{F_t\}$ -adapted, i.e. in T_0 we have $S(T_0, \omega_1) = S(T_0, \omega_2)$ and $N(T_0, \omega_1) = N(T_0, \omega_2)$. So independent of the (unknown) state, the products have a defined value in T_0 .

Given a derivative product (with the stochastic value process V), depending in T_1 on the attained state ω_i . We seek to determine the value of V in T_0 . Our setup is illustrated in Figure 3.1.

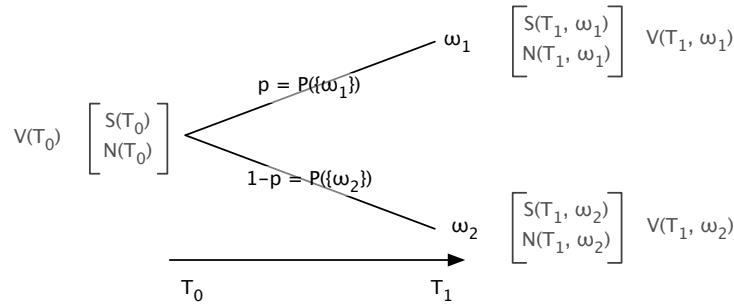


Figure 3.1.: Replication: The two-times-two-states-two-assets example

To have the derivative product V replicated by a portfolio we seek α, β such that

$$\begin{aligned} V(T_1, \omega_1) &= \alpha S(T_1, \omega_1) + \beta N(T_1, \omega_1), \\ V(T_1, \omega_2) &= \alpha S(T_1, \omega_2) + \beta N(T_1, \omega_2). \end{aligned} \quad (3.2)$$

This system of equations has a solution (α, β) , if

$$S(T_1, \omega_1)N(T_1, \omega_2) \neq S(T_1, \omega_2)N(T_1, \omega_1). \quad (3.3)$$

With this solution the value of the replication portfolio (and thus the cost of replication) is known in T_0 , and thus the “fair” value of the derivative product “V” in T_0 as

$$V(T_0) = \alpha S(T_0) + \beta N(T_0). \quad (3.4)$$

As before we have that the probabilities $P(\{\omega_1\})$, $P(\{\omega_2\})$ do not enter into the calculation of the cost of replication and thus $V(T_0)$. Let us investigate now whether $V(T_0)$ may be expressed in terms of an expectation. Assume that $N \neq 0$ and consider (3.2), (3.3), (3.4) for N -relative prices. Equivalent to (3.2), (3.3), (3.4) the portfolio satisfies in T_1

$$\frac{V(T_1, \omega_1)}{N(T_1, \omega_1)} = \alpha \frac{S(T_1, \omega_1)}{N(T_1, \omega_1)} + \beta \quad \frac{V(T_1, \omega_2)}{N(T_1, \omega_2)} = \alpha \frac{S(T_1, \omega_2)}{N(T_1, \omega_2)} + \beta \quad (3.5)$$

with the solvability condition being

$$\frac{S(T_1, \omega_1)}{N(T_1, \omega_1)} \neq \frac{S(T_1, \omega_2)}{N(T_1, \omega_2)} \quad (3.6)$$

and in T_0 we have

$$\frac{V(T_0)}{N(T_0)} = \alpha \frac{S(T_0)}{N(T_0)} + \beta. \quad (3.7)$$

Obviously we have

$$\frac{V(T_0)}{N(T_0)} = E\left(\frac{V(T_1)}{N(T_1)} | \mathcal{F}_{T_0}\right) \Leftrightarrow \frac{S(T_0)}{N(T_0)} = E\left(\frac{S(T_1)}{N(T_1)} | \mathcal{F}_{T_0}\right).$$

Now let $q \in \mathbb{R}$ such that

$$\frac{S(T_0)}{N(T_0)} = q \cdot \frac{S(T_1, \omega_1)}{N(T_1, \omega_1)} + (1 - q) \cdot \frac{S(T_1, \omega_2)}{N(T_1, \omega_2)},$$

i.e. $q = \frac{\frac{S(T_0)}{N(T_0)} - \frac{S(T_1, \omega_2)}{N(T_1, \omega_2)}}{\frac{S(T_1, \omega_1)}{N(T_1, \omega_1)} - \frac{S(T_1, \omega_2)}{N(T_1, \omega_2)}} \quad (\text{c.f. (3.6)}).$ If in addition to (3.6), we have

$$0 \leq q \leq 1, \quad (3.8)$$

then $\mathbb{Q}^N(\{\omega_1\}) := q$, $\mathbb{Q}^N(\{\omega_2\}) := 1 - q$ defines a probability measure and under $\mathbb{Q}^N$ we have

$$\frac{S(T_0)}{N(T_0)} = E^{\mathbb{Q}^N}\left(\frac{S(T_1)}{N(T_1)} | \mathcal{F}_{T_0}\right) \quad \text{d.h.} \quad \frac{V(T_0)}{N(T_0)} = E^{\mathbb{Q}^N}\left(\frac{V(T_1)}{N(T_1)} | \mathcal{F}_{T_0}\right). \quad (3.9)$$

So instead of calculating the parameters (α, β) for the replication portfolio we may alternatively calculate the measure $\mathbb{Q}^N$, i.e. the parameter q . At first it appears to be equally complex to calculate $\mathbb{Q}^N$ compared to the calculation of the replication strategy. However, determining $\mathbb{Q}^N$ has a striking advantage: The calculation of $\mathbb{Q}^N$ is independent of the derivative V , but the pricing formula (3.9) is valid for all derivatives V . If $\mathbb{Q}^N$ has been determined once, all derivatives V may be priced as a $\mathbb{Q}^N$ -expectation.

In Equations (3.5) to (3.9) we have considered N -relative prices, i.e. the value of any product V was expressed in fractions of N , i.e. by $\frac{V}{N}$. As long as $S \neq 0$ we may repeat these considerations with S -relative prices, i.e. we have

$$\frac{V(T_1, \omega_1)}{S(T_1, \omega_1)} = \alpha + \beta \frac{N(T_1, \omega_1)}{S(T_1, \omega_1)} \quad \frac{V(T_1, \omega_2)}{S(T_1, \omega_2)} = \alpha + \beta \frac{N(T_1, \omega_2)}{S(T_1, \omega_2)} \quad (3.10)$$

with the same (α, β) and we obtain the same value for the replication portfolio $V(T_0)$, namely

$$\frac{V(T_0)}{S(T_0)} = \alpha + \beta \frac{N(T_0)}{S(T_0)}. \quad (3.11)$$

If we determine the measure $\mathbb{Q}^S$, such that

$$\frac{N(T_0)}{S(T_0)} = E^{\mathbb{Q}^S} \left(\frac{N(T_1)}{S(T_1)} \middle| \mathcal{F}_{T_0} \right) \quad \text{d.h.} \quad \frac{V(T_0)}{S(T_0)} = E^{\mathbb{Q}^S} \left(\frac{V(T_1)}{S(T_1)} \middle| \mathcal{F}_{T_0} \right), \quad (3.12)$$

then the measure $\mathbb{Q}^S$ is different from $\mathbb{Q}^N$ - we have for example $\mathbb{Q}^S(\{\omega_1\}) = \frac{\frac{N(T_0)}{S(T_0)} - \frac{N(T_1, \omega_2)}{S(T_1, \omega_2)}}{\frac{N(T_1, \omega_1)}{S(T_1, \omega_1)} - \frac{N(T_1, \omega_2)}{S(T_1, \omega_2)}}$. However, the measure $\mathbb{Q}^S$ also allows to calculate the value of all derivatives V as a $\mathbb{Q}^S$ -expectation via (3.12).

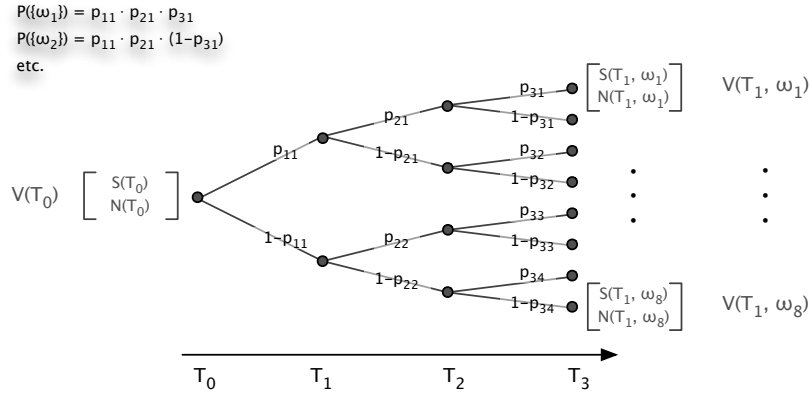


Figure 3.2.: Replication: Generalisation to multiple states

We conclude this section with some remarks:

- Under the measure $\mathbb{Q}^N$ the relative price $\frac{S}{N}$ is a martingale: This is the defining property for the measure $\mathbb{Q}^N$.
- Under the measure $\mathbb{Q}^N$ the relative price $\frac{V}{N}$ of any replication portfolio V is a martingale:⁴ This allows to calculate the price of V as $\mathbb{Q}^N$ -expectation of N -relative payout.
- The choice of the product that functions as reference (Numéraire) is arbitrary (as long as it is non-zero). The measure $\mathbb{Q}$, under which any Numéraire-relative

⁴ This follows since the replication portfolio is a linear combination of martingales and the expectation is linear.

replication portfolio becomes a martingale depends on the chosen Numéraire. This allows the change of the Numéraire-measure-pair $(N, \mathbb{Q}^N)$, e.g. if this simplifies the calculation of the expectation.

- It is necessary to consider relative prices, such that
 - $\mathbb{Q}^N$ is independent of V ,
 - $\mathbb{Q}^N$ is a probability measure, i.e. $\mathbb{Q}^N(\Omega) = 1$.

Exercise: Change of Measure

1. Reconsider the above under the Numéraire S , calculate $\mathbb{Q}^S$ and show (e.g. for an example), that $\mathbb{Q}^N \neq \mathbb{Q}^S$.
2. Instead of relative prices consider absolute prices, i.e. choose as Numéraire 1 and determine the measure $\mathbb{Q}^1$ for which

$$V(T_0) = E^{\mathbb{Q}^1}(V(T_1)|\mathcal{F}_{T_0}).$$

Show that $\mathbb{Q}^1$ depends on the specific $V(T_1)$ and is thus not universal for all replication portfolios.

In case we wish to replicate a payoff $X(T)$, which depends on multiple states $\omega_1, \dots, \omega_n$, then the above may be extended either by considering multiple time steps $T_0, T_1, \dots, T_{n-1} = T$ (dynamic replication) or multiple assets $N, S_1, \dots, S_{n-1}$ - see Figure 3.2.

3.2. Foundations: *Equivalent Martingale Measure*

3.2.1. Challenge and Solution Outline

An English version of this part is currently not available. What follows is the German version.



Motivation: Die Bewertung eines Produktes durch den Wert einer entsprechenden Replikationsstrategie lässt sich (entsprechend obigem Beispiel) mittels eines (äquivalenten) Maßes $\mathbb{Q}^N$ über den $\mathbb{Q}^N$ -Erwartungswert des N relativen Preises berechnen (wobei N eine gewählte Bezugsgröße ist - das *Numéraire*).

Wie könnte man das Maß $\mathbb{Q}^N$ bestimmen, gegeben man kennt Preis-Prozesse unter dem realen Maß $\mathbb{P}$? Dazu beachte man, dass unter dem Maß $\mathbb{Q}^N$ die N -relativen Preise Martingale, d.h. als Itô-Prozesse drift-frei (Lemma 47) sind (das ist die definierende Eigenschaft von $\mathbb{Q}^N$), und ein Maßwechsel $\mathbb{P} \rightarrow \mathbb{Q}^N$ für Itô Prozesse durch einen Driftwechsel charakterisiert wird (Satz 53).

Sind also die Preis-Prozesse unter dem realen Maß $\mathbb{P}$ bekannt (d.h. liegen als „Modell“ vor), so lassen sich die N -relativen Preis-Prozesse unter dem realen Maß $\mathbb{P}$ mittels der [Quotientenregel](#) bestimmen. Anschließend lässt sich das äquivalente Martingal Maß $\mathbb{Q}^N$ aus dem (bekannten) Driftwechsel über den [Satz von Girsanov](#) (Satz 53) ermitteln (vergleiche Abbildung 3.3).

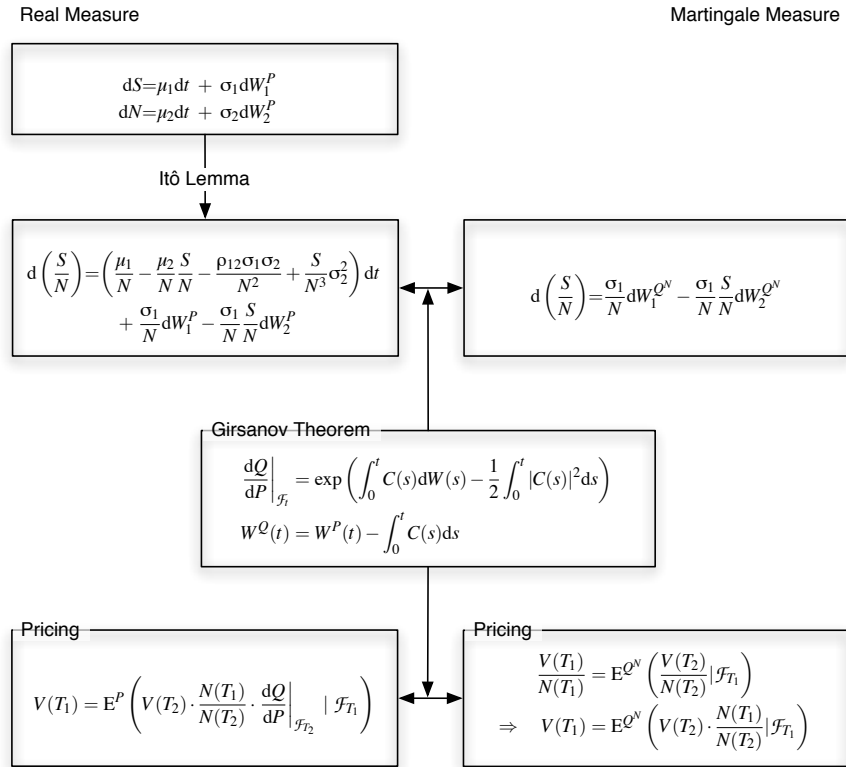
In unseren Anwendungen muss man das äquivalente Martingal-Maß hingegen *nie* bestimmen: Da man die Prozesse unter $\mathbb{Q}^N$ kennt (man kennt ihre Drift und nur die ändert sich unter dem neuen Maß), kennt man die bedingten Wahrscheinlichkeitsverteilungen unter $\mathbb{Q}^N$, und dies genügt, um den Erwartungswert zu berechnen (vgl. Definition 9).

Es bleibt nun lediglich die Frage zu klären, unter welchen Bedingungen sich eine vorgegebene Auszahlungsfunktion replizieren lässt und unter welchen Bedingungen ein äquivalentes Martingal-Maß (*equivalent martingale measure*) existiert. Wir geben in dieses Kapitel einen knappen Überblick über die entsprechenden mathematischen Grundlagen, werden in den Anwendungen jedoch die Existenz des äquivalentes Martingalmaßes nicht weiter diskutieren, sondern lediglich voraussetzen. $\triangleleft$

Problemstellung:

Gegeben: $\mathcal{M} = \{X_1, \dots, X_n\}$, wobei X_i Preis-Prozesse unter einem (realen) Maß $\mathbb{P}$ sind, sowie ein Auszahlungsprofil (*contingent claim*) $V^{(T)}$, wobei $V^{(T)}$ eine $\mathcal{F}_T$ messbare Zufallsvariable ist.

Gesucht: Preisindikation, d.h. der Wert $V(t)$ von $V^{(T)}$ zum Zeitpunkt $t < T$, insbesondere $V(0)$, wobei V ein $\{\mathcal{F}_t\}$ -adaptierter stochastischer Prozess sei.

**Figure 3.3.:** Real Measure versus Martingale Measure.

Lösung (skizziert):

- *Wahl des Numéraires:* Sei $N \in \mathcal{M}$ ein Preis-Prozess, der als Bezugsgröße (Numéraire) diene, ohne Beschränkung der Allgemeinheit sei $N = X_1$.
- *Existenz eines Martingalmaßes für N -relative Preise:* Nach Satz 65 existiere ein Maß $\mathbb{Q}^N$, so dass $\frac{X_i}{N}$ Martingale bezüglich $\mathcal{F}_t$ sind, $\forall i = 1, \dots, n$.
- *Definition des (Kandidaten) für den Wertprozess der Replikationsportfolios:* Definiere

$$V(t) := N(t) \cdot E^{\mathbb{Q}^N} \left(\frac{V(T)}{N(T)} \middle| \mathcal{F}_t \right).$$

Dann ist $\frac{V(t)}{N(t)}$ ein $\mathbb{Q}^N$ -Martingal bezüglich $\mathcal{F}_t$ (tower law).

- *Martingaldarstellungssatz liefert Handelsstrategie:* Da unter $\mathbb{Q}^N$ die Prozesse $\frac{X}{N} = (\frac{X_1}{N}, \dots, \frac{X_n}{N})$ sowie $\frac{V}{N}$ Martingale sind, existiert ein $\phi = (\phi_1, \dots, \phi_n)$ mit

$$\frac{V(t)}{N(t)} = \frac{V(0)}{N(0)} + \int_0^t \phi(s) \cdot d \left(\frac{X(s)}{N(s)} \right) \quad (3.13)$$

(Martingale Representation Theorem).

- *Der Portfolioprozess ϕ kann als selbstfinanzierend gewählt werden durch*

$$\phi_1(t) := \frac{V(0)}{N(0)} + \sum_{i=2}^n \int_0^t \phi_i(s) d \left(\frac{X_i(s)}{N(s)} \right) - \sum_{i=2}^n \phi_i(t) \frac{X_i(t)}{N(t)}$$

(beachte: $d \left(\frac{X_1}{N} \right) = d \left(\frac{N}{N} \right) = 0$, d.h. (3.13) gilt unverändert).

- *Der so gefundene Portfolioprozess ϕ beschreibt das Replikationsportfolio für $V(T)$:* Es ist

$$V(T) = \sum_{i=1}^n \phi_i(T) X_i(T) = \sum_{i=1}^n \phi_i(0) X_i(0) + \int_0^T \phi(s) \cdot dX(s).$$

- *Eine Bewertung ist ohne Bestimmung des Replikationsportfolios möglich:* $V(t)$ ist der Wert des Replikationsportfolios und es ist

$$\frac{V(t)}{N(t)} = E^{\mathbb{Q}^N} \left(\frac{V(T)}{N(T)} \middle| \mathcal{F}_t \right).$$

3.2.2. Steps towards the *Universal Pricing Theorem*

Wir geben nun die zentralen Bausteine der risikoneutralen Bewertung auf dem Weg zum *Universal Pricing Theorem*:

- *Handelsstrategien* und Charakterisierung der *Selbstfinanzierung* einer Handelsstrategie - im Folgenden in Nummer 58-62.
- *Äquivalentes Martingalmaß* um über das Martingal Representation Theorem die Existenz einer Handelsstrategie zu erhalten - im Folgenden in Nummer 63-65.
- *Replikation gegebener Auszahlungsfunktionen* und *Universal Pricing Theorem* folgen durch Definition des Martingalprozesses $t \mapsto E^{Q^N}(\frac{V^T}{N(T)} \mid \mathcal{F}_t)$ einer gegebenen Auszahlungsfunktion V^T - im Folgenden in Nummer 66-70.

Dieses Programm findet sich meist in dieser Reihenfolge in der Literatur, wobei mal weniger ([3]), mal mehr ([6]) auf technische Voraussetzungen eingegangen wird. Wir skizzieren im Folgenden diese Schritte meist ohne technische Beweise, jedoch mit entsprechenden Literaturverweisen, um möglichst zügig die Theorie zur Anwendung zu bringen.

3.2.2.1. Self-financing Trading Strategy

Definition 58 (Portfolio, Trading Strategy, Self-financing):

Es sei $\mathcal{M} = \{X_1, \dots, X_n\}$ eine Familie von (Itô) stochastischen Prozessen, definiert auf dem filtrierten Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})$, der n -dimensionale Prozess $X = (X_1, \dots, X_n)^T$ sei $\{\mathcal{F}_t\}$ -adaptiert. Die Elemente von $\mathcal{M}$ seien Preis-Prozesse von gehandelten Produkten ($\mathcal{M}$ stellt eine Gesamtheit gehandelter Produkte, einen Markt, dar).

1. Ein n -dimensionaler $\{\mathcal{F}_t\}$ -adaptierter Prozess⁵ $\phi = (\phi_1, \dots, \phi_n)$, beschränkt und aus der Klasse der Integranden des Itô-Integrals heißt *Portfolioprozess* oder *Trading Strategy*.
2. Der Wert des Portfolios ϕ zum Zeitpunkt t ist gegeben durch das Skalarprodukt

$$V_\phi(t) := \phi(t) \cdot X(t) = \sum_{i=1}^n \phi_i(t) X_i(t).$$

Der Prozess V_ϕ heißt *wealth process* des Portfolios ϕ .

⁵ Man findet diese Definition in der Literatur mit der Bedingung, dass es sich um einen $\{\mathcal{F}_t\}$ -vorhersehbaren Prozess (vergleiche Definition 55) handelt. Da wir als Integratoren nur Brownsche Bewegungen betrachten, genügt es, sich auf adaptierte Prozesse zu beziehen. Vergleiche Bemerkung 57.

3. Der *gain process* des Portfolios ϕ sei definiert durch

$$G_\phi(t) := \int_0^t \phi(\tau) \cdot dX(\tau) = \sum_{i=1}^n \int_0^t \phi_i(\tau) \cdot dX_i(\tau).$$

4. Ein Portfolio Prozess ϕ heißt *selbstfinanzierend*, falls

$$V_\phi(t) = V_\phi(0) + G_\phi(t), \quad (3.14)$$

d.h.

$$dV_\phi = \phi \cdot dX. \quad (3.15)$$

┘



Interpretation: Wir stellen uns unter $\{X_i \mid i = 1, \dots, n\}$ eine Familie von Aktienpreisprozessen vor und $\phi(t) := (\phi_1(t), \dots, \phi_n(t))$ beschreibe ein Aktienportfolio, d.h. ϕ_i ist die Anzahl der Aktien X_i im Portfolio.

Die Beziehung (3.15) kann wie folgt interpretiert werden: Die Veränderung des Portfoliowertes V_ϕ resultiert aus der Änderung der Aktien X , so *als ob* sich das Portfolio nicht ändern würde, d.h. man hält ein Portfolio von ϕ Aktien und erzielt damit über den Zeitraum dt den Gewinn $\phi \cdot dX$. Nun ist aber ebenso

$$dV_\phi = d(\phi \cdot X) = d\phi \cdot X + \phi \cdot dX + d\phi \cdot dX = \phi \cdot dX + d\phi \cdot (X + dX)$$

und daher folgt aus (3.15) als äquivalente Charakterisierung der Selbstfinanzierung

$$d\phi \cdot (X + dX) = 0$$

und mit $d\phi \cdot dX = 0$ also

$$d\phi \cdot X = 0. \quad (3.16)$$

Die Interpretation dieser Bedingungen wird besonders transparent, betrachtet man die zeitdiskrete Variante einer selbstfinanzierten Handelsstrategie:

- Zum Zeitpunkt T_i existiert ein Portfolio $\phi(T_i)$ der Produkte $X(T_i)$. Sein Wert ist $V_\phi(T_i) = \phi(T_i) \cdot X(T_i)$.
- Im Zeitpunkt T_i werden die Produkte X zu den Preisen $X(T_i)$ gehandelt und das Portfolio *selbstfinanzierend* umgeschichtet. Die Portfoliozusammensetzung ändert sich um $\Delta\phi(T_{i+1}) = \phi(T_{i+1}) - \phi(T_i)$. Das Portfolio wird nur umgeschichtet, es erfährt keine Wertänderung: $\Delta\phi(T_{i+1}) \cdot X(T_i) = 0$. Dem entspricht im kontinuierlichen Fall:

$$d\phi \cdot X = 0.$$

- Über den Zeitraum $\Delta T_i := T_{i+1} - T_i$ ändert sich der Wert der Produkte X um $\Delta X(T_i) := X(T_{i+1}) - X(T_i)$. Der Wert des Portfolios ändert sich somit um $\phi(T_{i+1}) \cdot \Delta X(T_i)$, d.h.

$$\begin{aligned} \Delta V_\phi(T_i) &= \phi(T_{i+1})X(T_{i+1}) - \phi(T_i)X(T_i) \\ &= \phi(T_{i+1})\Delta X(T_i) + \underbrace{\Delta\phi(T_i)X(T_i)}_{=0}. \end{aligned}$$

Dem entspricht im kontinuierlichen Fall:

$$dV_\phi = \phi \cdot dX.$$

Man beachte, dass diese Interpretation insbesondere konsistent mit unserer Definition des elementaren Itô-Integrals ist, siehe auch Bemerkung 32.

- Der erste Schritt der Umschichtung des Portfolios erfolgt (bedingt) unabhängig von der zukünftigen Änderung der Produktpreise, das Portfolio darf zukünftige Veränderungen der Produktpreise nicht antizipieren. Diese Information liegt ja zum Zeitpunkt der Portfolioumschichtung gar nicht vor, wir haben die beiden zeitdiskreten Änderungen $\Delta\phi(T_i)$ und $\Delta X(T_i)$ als zeitlich nacheinander dargestellt. Dieser Unabhängigkeit entspricht im kontinuierlichen Fall:

$$d\phi \cdot dX = 0.$$

- u.S.W.

Wir diskutieren diese diskrete Variante der Replikation (die dann keine vollständige Replikation mehr liefert) in Kapitel 7. <|

Definition 59 (Numéraire⁶⁷): ↗

Ein Preis-Prozess $N \in \mathcal{M}$ heißt *Numéraire* (auf $[0, T]$), falls $N(t) > 0$ ($\mathbb{P}$ -fast sicher) für alle $t \leq T$. ↘

Für den Rest des Kapitels sei nun X_1 ein Numéraire und zur Verdeutlichung $N := X_1$.

Lemma 60 (Selbstfinanzierungsbedingung ist invariant unter der Betrachtung relativer Preise): Es sei $\phi = (\phi_1, \dots, \phi_n)$ ein Portfolio. Dann ist

$$dV_\phi = \sum_{i=1}^n \phi_i dX_i \Leftrightarrow d\frac{V_\phi}{N} = \sum_{i=1}^n \phi_i d\frac{X_i}{N}.$$

Das heißt ϕ ist genau dann selbstfinanzierend, wenn

$$d\frac{V_\phi}{N} = \sum_{i=1}^n \phi_i d\frac{X_i}{N}.$$

⁶ Vgl. [6], §6.1.

⁷ Numéraire (frz.): Das Bargeld, das Hartgeld. Hier: Bezugsgröße in der Werte gemessen werden.

Proof: Es sei

$$dV_\phi = \sum_{i=1}^n \phi_i dX_i. \quad (3.17)$$

Dann ist mit $dX_i \cdot (\frac{1}{N} + d(\frac{1}{N})) + X_i \cdot d(\frac{1}{N}) = d\frac{X_i}{N}$

$$\begin{aligned} d\frac{V_\phi}{N} &= dV_\phi \cdot (\frac{1}{N} + d(\frac{1}{N})) + V_\phi \cdot d(\frac{1}{N}) \\ &\stackrel{(3.17)}{=} \sum_{i=1}^n \phi_i dX_i \cdot (\frac{1}{N} + d(\frac{1}{N})) + \sum_{i=1}^n \phi_i X_i \cdot d(\frac{1}{N}) \\ &= \sum_{i=1}^n \phi_i d\frac{X_i}{N} = \sum_{i=2}^n \phi_i d\frac{X_i}{N}. \end{aligned}$$

Umgekehrt folgt genauso aus

$$d\frac{V_\phi}{N} = \sum_{i=1}^n \phi_i d\frac{X_i}{N} \quad (3.18)$$

mit $d\frac{X_i}{N} \cdot (N + dN) + \frac{X_i}{N} \cdot dN = dX_i$ dass

$$\begin{aligned} dV_\phi &= d(\frac{V_\phi}{N} \cdot N) = d\frac{V_\phi}{N} \cdot (N + dN) + \frac{V_\phi}{N} \cdot dN \\ &\stackrel{(3.18)}{=} \sum_{i=1}^n \phi_i d\frac{X_i}{N} \cdot (N + dN) + \sum_{i=1}^n \phi_i \frac{X_i}{N} \cdot dN \\ &= \sum_{i=1}^n \phi_i dX_i. \end{aligned}$$

□

Remark 61: Man beachte, dass wegen der Wahl des Numéraire ϕ_1 nicht in die Summe über $d\frac{X_i}{N}$ eingeht. Wir nutzen dies in folgendem Lemma um ein selbstfinanzierendes Replikationsportfolio zu konstruieren.

Der Übergang zu relativen Preisen erlaubt es aus einem (Teil-)Portfolio $\phi_2, \dots, \phi_n$, welches für einen gegebenen Prozess V die Identität

$$d\frac{V}{N} = \sum_{i=2}^n \phi_i d\frac{X_i}{N} \quad (3.19)$$

erfüllt, ein selbstfinanzierendes Portfolio $\phi_1, \dots, \phi_n$ zu gewinnen. Man beachte, dass in (3.19) nicht V_ϕ (der Wertprozess des Portfolios) steht sondern ein beliebiger Prozess

V , der durch die folgende Wahl von ϕ_1 zum Wertprozess eines selbstfinanzierenden Portfolios wird (Replikation):

Lemma 62 (Selbstfinanzierende Strategie bei gegebenen Teilportfolio und Anfangswert): Seien $\phi_2, \dots, \phi_n$ gegeben, derart dass $\int_0^t \phi_i d\frac{X_i}{N}$ existiert. Dann ist $\phi = (\phi_1, \dots, \phi_n)$ mit

$$\phi_1(t) := \frac{V_0}{N(0)} + \sum_{i=2}^n \int_0^t \phi_i(s) d\left(\frac{X_i(s)}{N(s)}\right) - \sum_{i=2}^n \phi_i(t) \frac{X_i(t)}{N(t)} \quad (3.20)$$

eine selbstfinanzierende Strategie mit $V_\phi(0) = V_0$.

Proof: Zur Selbstfinanzierung von ϕ genügt es nach Lemma 60 zu zeigen, dass

$$d\frac{V_\phi}{N} = \sum_{i=1}^n \phi_i d\frac{X_i}{N}. \quad (3.21)$$

Aus Gleichung (3.20) folgt unmittelbar (beachte $\frac{X_1}{N} = 1$ und $d\frac{X_1}{N} = 0$)

$$\sum_{i=1}^n \phi_i(t) \frac{X_i(t)}{N(t)} = \frac{V_0}{N(0)} + \sum_{i=1}^n \int_0^t \phi_i(s) d\left(\frac{X_i(s)}{N(s)}\right)$$

und somit (3.21). Die Anfangsbedingung $V_\phi(0) = V_0$ folgt durch einsetzen von $t = 0$. $\square$

Die Beziehung (3.19) könnte aus dem Martingaldarstellungssatz folgen, falls die beteiligten Prozesse Martingale sind. Es drängt sich somit die Frage nach einem Maß auf, unter welchem die relativen Preise $\frac{X_i}{N}$ Martingale sind.

3.2.2.2. Equivalent Martingale Measure

Definition 63 (Äquivalentes Martingalmaß (Equivalent Martingale Measure)): $\neg$
Es sei N ein Numéraire. Ein Wahrscheinlichkeitsmaß $\mathbb{Q}^N$ definiert auf $(\Omega, \mathcal{F})$ heißt *äquivalentes Martingalmaß bzgl. N* (equivalent N -martingale measure) (auf $\mathcal{M}$), falls

1. $\mathbb{Q}^N$ und $\mathbb{P}$ sind äquivalent, und
2. Die N relativen Preis-Prozesse $\frac{X_i}{N}$ ($i = 1, \dots, n$) sind $\mathcal{F}_t$ -Martingale bezüglich $\mathbb{Q}^N$.

$\neg$

Definition 64 (Arbitrage, Arbitragefrei): $\neg$

Ein selbst finanzierender Portfolioprozess ϕ ist eine Arbitrage-Strategie, falls

$$V_\phi(0) = 0, \quad P(V_\phi(T) \geq 0) = 1, \quad P(V_\phi(T) > 0) > 0.$$

Ein Markt $\mathcal{M}$ auf dem keine Arbitrage-Strategie existiert heißt arbitragefrei. ┘

Theorem 65 (Existenz des äquivalenten Martingalmaß): Auf $\mathcal{M}$ existiert genau dann ein (zu $\mathbb{P}$) äquivalentes Martingalmaß $\mathbb{Q}^N$ (*equivalent martingale measure*) (bezüglich des Numéraires N), wenn auf $\mathcal{M}$ keine Arbitrage-Strategie existiert.

Der Satz 65 ist in dieser Form nur gültig, falls es sich bei den Preisprozessen um zeitdiskrete Prozesse über einem diskreten Wahrscheinlichkeitsraum Ω handelt. Im kontinuierlichen Fall ist die Charakterisierung des Fehlens einer Arbitragestrategie technischer als in Definition 64, siehe z.B. [6] und die dortigen Literaturhinweise. Wir sind an dieser Stelle etwas ungenauer, da zum einen der exakte Beweis kaum zur Verbesserung der Intuition und zum Verständnis der Zusammenhänge beiträgt, zum anderen da die spätere Implementierung der Modelle ohnehin immer in einem zeitdiskreten Modell resultiert, für welches der zitierte Satz so Gültigkeit hat. Die genauen technischen Voraussetzungen sind in den im Folgenden betrachteten zeitkontinuierlichen Modellen erfüllt und wir postulieren daher schlicht die Existenz des äquivalenten Martingalmaß, um uns zügig der Modellierung und Implementierung zuzuwenden.

3.2.2.3. Payoff Replication

Definition 66 (Zulässige trading strategy): ┘

Eine selbstfinanzierende Strategie ϕ heißt $\mathbb{Q}^N$ -zulässig ($\mathbb{Q}^N$ -admissible), falls

$$\int_0^t \phi(s) d\frac{X(s)}{N(s)}$$

ein $\mathbb{Q}^N$ Martingal ist. ┘

Remark 67 (Zulässige Strategie): Nach Lemma 62 ist ϕ eine $\mathbb{Q}^N$ -zulässige (selbstfinanzierende) Strategie, falls die Integrale $\int_0^t \phi_i(s) d\frac{X_i(s)}{N(s)}$ für $i = 2, \dots, n$ existieren.

Definition 68 (Replizierbare Auszahlungsfunktion, attainable contingent claim, complete): ┘

Sei $T > 0$. Eine $\mathcal{F}_T$ -messbare Zufallsvariable $V^{(T)}$ heißt replizierbare Auszahlungsfunktion (*attainable contingent claim*), falls (mindestens) eine zulässige trading strategy ϕ existiert mit

$$V_\phi(T) = V^{(T)}.$$

Der Markt $\mathcal{M}$ heißt *complete*, falls jede Auszahlungsfunktion replizierbar ist. ┘

Es folgt

Theorem 69 (selbstfinanzierendes Replikationsportfolio einer gegebenen Auszahlungsfunktion): Es sei $\mathbb{Q}^N$ ein äquivalentes Martingalmaß und $V^{(T)}$ eine

gegebene Auszahlungsfunktion. Ferner sei

$$V(t) := N(t) \cdot E^{\mathbb{Q}^N} \left(\frac{V(T)}{N(T)} \middle| \mathcal{F}_t \right).$$

Damit ist $\frac{V}{N}$ ein $\mathbb{Q}^N$ -Martingal. Seien nun $\phi_2, \dots, \phi_n$ entsprechend des Martingal Representation Theorems, d.h.

$$d \frac{V}{N} = \sum_{i=2}^n \phi_i d \frac{X_i}{N}.$$

Für das Portfolio $\phi = (\phi_1, \dots, \phi_n)$ sei ϕ_1 entsprechend Lemma 62 mit $V_\phi(0) := V(0)$ gewählt.

Dann ist ϕ ein selbstfinanzierendes Replikationsportfolio von V , d.h. es ist

$$V(t) = V_\phi = \sum_{i=1}^n \phi_i X_i.$$

Proof: Dass $\frac{V}{N}$ ein $\mathbb{Q}^N$ -Martingal ist folgt unmittelbar aus der Definition und dem *Tower Law*⁸. Nach der Definition des Martingalmaßes sind $\frac{X_i}{N}$ ebenfalls $\mathbb{Q}^N$ -Martingale und der Martingaldarstellungssatz liefert die Existenz von $\phi_2, \dots, \phi_n$ mit

$$d \frac{V}{N} = \sum_{i=2}^n \phi_i d \frac{X_i}{N}$$

und mit Lemma 62 wird ϕ selbstfinanzierend.

Für den Wertprozess V_ϕ des Portfolios gilt nach Lemma 60 ebenfalls

$$d \frac{V_\phi}{N} = \sum_{i=2}^n \phi_i d \frac{X_i}{N}$$

und somit $d \frac{V_\phi}{N} = d \frac{V}{N}$, also mit $V_\phi(0) = V(0)$ auch

$$V(t) = V_\phi(t) = \sum_{i=1}^n \phi_i d X_i.$$

□

Da die Bestimmung des Martingalmaß $\mathbb{Q}^N$ vollständig unabhängig von V war, folgt

Theorem 70 (Risk-Neutral Valuation Formula): Es sei ϕ eine zulässige *self-financing trading strategy* und $\mathbb{Q}^N$ ein equivalent martingale measure bezüglich des Numéraires N . Dann gilt

$$\frac{V_\phi(t)}{N(t)} = E^{\mathbb{Q}^N} \left(\frac{V_\phi(T)}{N(T)} \middle| \mathcal{F}_t \right). \quad (3.22)$$

⁸ Siehe Übung 2 auf Seite 399

3.3. Basic Assumptions

Wir diskutieren in den folgenden Anwendungen die Voraussetzungen für die Anwendbarkeit der in Abschnitt 2-3 vorgestellten Theorie selten oder gar nicht und machen stattdessen schlichtweg folgende

Annahme

Zu $\mathcal{M}$ existiert ein (eindeutiges) *equivalent martingale measure* und alle Auszahlungsfunktionen V^T mit $\frac{V^T}{N} \in L^1(\mathcal{F}, \mathbb{Q}^N)$, $\mathcal{F}_T$ -messbar, sind replizierbar.

(3.23)

3.4. Excursion: Relative Prices and Risk Neutral Measures

3.4.1. Why relative Prices?

Im einfachen Beispiel eines zeitdiskreten Prozesses mit diskreten Zuständen (Abschnitt 3.1.2) wurde deutlich, dass sich das so nützliche Martingal-Maß im allgemeinen nur dann unabhängig von der Auszahlungsfunktion bestimmen lässt, wenn man relative Preise betrachtet. Dennoch erscheint die Betrachtung von relativen Preisen in der kontinuierlichen Theorie (Abschnitt 3.2.2) nur wenig motiviert. Zur Motivation der *relativen Preise* versuchen wir die kontinuierliche Theorie mit absoluten Preisen zu wiederholen:

Es seien

$$dX_i = \mu_i \cdot dt + \sigma_i \cdot dW_i^{\mathbb{P}}, \quad i = 1, \dots, n$$

die Preisprozesse von n gehandelten Produkten. Wir nehmen zunächst an, es sei $\sigma_i > 0$, $\forall i$. Die Prozesse seien gegeben unter dem realen Maß $\mathbb{P}$. Da $\sigma_i > 0$ ist, existiert nun nach dem [Satz von Girsanov](#) ein Maß $\mathbb{Q}^1$, sodass die Prozesse $\mathbb{Q}^1$ -Martingale sind, d.h.

$$dX_i = \sigma_i \cdot dW_i^{\mathbb{Q}^1}, \quad i = 1, \dots, n.$$

Ist nun V^T eine Auszahlungsfunktion so ist

$$V(t) := E^{\mathbb{Q}^1}(V^T | \mathcal{F}_t) \quad (3.24)$$

ein $\mathbb{Q}^1$ -Martingal und mit dem [Martingaldarstellungssatz](#) folgt die Existenz eines (Portfolio-)Prozesses $(\phi_1, \dots, \phi_n)$ mit

$$dV(t) = \sum_{i=1}^n \phi_i \cdot dX_i.$$

An dieser Stelle taucht allerdings die Frage auf, ob $(\phi_1, \dots, \phi_n)$ ein selbstfinanzierendes Replikationsportfolio ist, d.h. ob

$$V(t) = \sum_{i=1}^n \phi_i \cdot X_i.$$

Ohne diese Identität ist $V(t) - V(0)$ lediglich der *gain process* der Handelsstrategie $(\phi_1, \dots, \phi_n)$.⁹ Allerdings existiert ein selbstfinanzierendes Replikationsportfolio. Um dies zu beweisen nehmen wir den Umweg über relative Preise und wenden Satz 69

⁹ Man beachte an dieser Stelle nochmals genau die Definition der Selbstfinanziertheit: Die geforderte Bedingung ist $d(\sum_{i=1}^n \phi_i X_i) = \sum_{i=1}^n \phi_i dX_i$.

an. Man beachte, dass sowohl die Eigenschaft der Replikation als auch die der Selbstfinanzierung unabhängig vom Maß ist. Es sei nun also $(\phi_1, \dots, \phi_n)$ dieses selbstfinanzierende Replikationsportfolio. Definieren wir dann V als den Wertprozess (statt über (3.24))

$$V(t) := \sum_{i=1}^n \phi_i \cdot X_i,$$

so folgt

$$dV(t) = \sum_{i=1}^n \phi_i \cdot dX_i$$

und somit

$$dV(t) = \sum_{i=1}^n \phi_i \cdot \sigma_i \cdot dW_i^{\mathbb{Q}^1} \quad \text{unter } \mathbb{Q}^1.$$

Der Wertprozess V ist also ein Martingal und es gilt ein *Universal Pricing Theorem*

$$V(t) = E^{\mathbb{Q}^1}(V^T \mid \mathcal{F}_t).$$

Sieht man also davon ab, dass zum Beweis der Existenz der selbstfinanzierenden Strategie die Transformation auf relative Preise nötig war, scheint es so, als könne man das *Universal Pricing Theorem* auch im Maß $\mathbb{Q}^1$ (ohne relative Preise) erhalten. Dass dies jedoch i.a. nicht geht zeigt sich nun: Wir betrachten die Preisprozesse wie oben, jedoch sei $\sigma_1 = 0$ und, zum Beispiel, $\mu_1 = r \cdot X_1$ mit $r > 0$ sowie $X_1(0) > 0$. In diesem Fall entspricht X_1 dem Wert einer risikolosen ($\sigma_1 = 0$) Anlage mit kontinuierlicher Verzinsungsrate r :

$$dX_1 = r \cdot X_1 \cdot dt, \quad \text{d.h.} \quad X_1(t) = X_1(0) \cdot \exp(r \cdot t).$$

Offensichtlich ist es in keinem Maß möglich X_1 zum Martingal zu machen. Die Anlage ist risikolos, d.h. X_1 nicht stochastisch und der Erwartungswert daher sogar völlig unabhängig vom Wahrscheinlichkeitsmaß. Um also alle Prozesse durch eine Maßtransformation auf Martingale zu bringen bleibt nur der Übergang zu relativen Preisen. Nach einer Division aller Prozesse durch $N := X_1$ ist $\frac{X_1}{N} = 1$ trivialerweise ein Martingal (unter jedem Maß). Unter den verbleibenden Prozessen $\frac{X_i}{N}$ befindet sich nun kein risikoloser Prozess der kein Martingal ist, da dies einer Arbitragemöglichkeit entspräche.¹⁰

Es ist jedoch genauso möglich durch einen beliebigen anderen Numéraire zu teilen: Entweder ist $\frac{X_1}{N}$ stochastisch (und kann durch Maßtransformation zum Martingal gemacht werden) oder $\frac{X_1}{N}$ ist drift-frei.¹¹

¹⁰ Gäbe es in einem Markt zwei risikolose Anlagen mit unterschiedlicher lokaler Rendite, so liesse sich durch Verkauf der einen und entsprechend anteiligem Kauf der anderen ohne Kosten ein Portfolio zusammenstellen, dessen zukünftiger Wert mit Wahrscheinlichkeit 1 positiv ist.

¹¹ Die Aussage gilt nicht global sondern lokal.



Interpretation: Der Übergang zu relativen Preisen entspricht der Transformation auf einen risikolosen Zins (Drift) von 0. In der Physik entspricht dies dem Wechsel in ein bewegtes Bezugssystem, in dem die risikolose Anlage keine Drift besitzt. Da es nur eine risikolose Anlage gibt, kann nun ein Maß gefunden werden unter dem *alle* Prozesse Martingale sind (und der Martingaldarstellungssatz zur Herleitung *Universal Pricing Theorem* angewendet werden. $\triangleleft$

3.4.2. Risk Neutral Measure

Als *risikoneutrales Maß* wird das äquivalente Martingalmaß $\mathbb{Q}^B$ zur (lokal) risikolosen Anlage B mit

$$dB = r \cdot B \cdot dt$$

bezeichnet. Der Name *risikoneutrales Maß* leitet sich dabei aus folgender Überlegung ab: Wäre $\mathbb{Q}^B$ das reale Maß (d.h. die Wahrscheinlichkeiten, mit denen die Größen tatsächlich die entsprechenden Werte annehmen), so würde dies bedeuten, dass ein beliebiges Finanzprodukt S relativ zur risikolosen Anlage B drift-frei ist, denn $\frac{S}{B}$ ist ein $\mathbb{Q}^B$ -Martingal. Damit besitzt S unter $\mathbb{Q}^B$ lokal die gleiche erwartete Rendite wie B , d.h. z.B. für einen Lognormalprozess

$$dS = r \cdot S \cdot dt + \sigma \cdot S \cdot dW^{\mathbb{Q}^B}$$

Die erwartete lokale Rendite eines Finanzproduktes ist in einem solchen Markt also unabhängig von seinem Risiko σ , d.h. die Marktteilnehmer stehen dem *Risiko neutral* (also weder avers noch affin) gegenüber.

Part II.

First Applications: Black-Scholes Model, Hedging and Greeks

CHAPTER 4

Pricing of an European Stock Option under the Black-Scholes Model

We assume here that interest rates are deterministic and that a bank accounts exists which may be used to invest or draw cash at a *continuously compounded* rate $r(t)$, i.e. the investment of $B(0) = 1$ in $t = 0$ evolves as

$$dB(t) = r(t)B(t)dt,$$

i.e.

$$B(t) = \exp\left(\int_0^t r(\tau)d\tau\right) \quad (4.1)$$

(here r denotes a known real valued function of time, not a random variable). Furthermore we assume that the stock $S(t)$ follows

$$dS(t) = \mu^{\mathbb{P}}(t)S(t)dt + \sigma(t)S(t)dW^{\mathbb{P}}(t) \quad \text{under the real measure } \mathbb{P}, \quad (4.2)$$

(if σ is a constant, then σ is called Black-Scholes volatility). The measure $\mathbb{P}$ denotes the *real measure*. Let $K \in \mathbb{R}$. We wish to derive the value $V(0)$ of the contract paying

$$V(T) = \max(S(T) - K, 0)$$

in $t = T$.

We use the techniques from Chapter 3 (the following steps will be repeated similarly in other applications): As a Numéraire $N(t)$, i.e. as reference quantity, we choose the bank account $N(t) := B(t)$. By assumption (3.23) we have the existence of a measure $\mathbb{Q}^N$, equivalent to $\mathbb{P}$, such that $\frac{S(t)}{N(t)}$ and $\frac{V(t)}{N(t)}$ are both martingales.¹ From Theorem 53 we have that S is given $\mathbb{Q}^N$ with a changed drift, i.e.

$$dS(t) = \mu^{\mathbb{Q}^N}(t)S(t)dt + \sigma(t)S(t)dW^{\mathbb{Q}^N}(t) \quad \text{under } \mathbb{Q}^N,$$

¹ By $V(t)$ we denote the value of the replication portfolio consisting of $S(t)$, $B(t)$ such that $V(T) = \max(S(T) - K, 0)$.

(where $W^{\mathbb{Q}^N}$ denotes a $\mathbb{Q}^N$ -Brownian motion).² From the quotient rule 43 we find

$$\begin{aligned} d\left(\frac{S(t)}{B(t)}\right) &= \frac{S(t)}{B(t)}[\mu^{\mathbb{Q}^N}(t)dt + \sigma(t)dW^{\mathbb{Q}^N}(t)] - \frac{S(t)}{B(t)}[r(t)dt] \\ &\quad - \frac{S(t)}{B(t)}[\mu^{\mathbb{Q}^N}(t)dt + \sigma(t)dW^{\mathbb{Q}^N}(t)][r(t)dt] \\ &\quad - \frac{S(t)}{B(t)}[r(t)dt][r(t)dt] \end{aligned}$$

and by formal expansion, see Theorem 40

$$= \frac{S(t)}{B(t)}[(\mu^{\mathbb{Q}^N}(t) - r(t))dt + \sigma(t)dW^{\mathbb{Q}^N}(t)] \quad \text{under } \mathbb{Q}^N.$$

From Lemma 47 it follows that $\mu^{\mathbb{Q}^N}(t) = r(t)$. For the process $Y := \log(S)$ we then have by Lemma 44

$$d(\log(S(t))) = (r(t) - \frac{1}{2}\sigma^2(t))dt + \sigma(t)dW^{\mathbb{Q}^N}(t) \quad \text{under } \mathbb{Q}^N,$$

i.e. $\log(S(T))$ has normal distribution with mean $\bar{\mu} := \log(S(0)) + \bar{r}T - \frac{1}{2}\bar{\sigma}^2T$ and standard deviation $\bar{\sigma}\sqrt{T}$, where we define $\bar{r}$ and $\bar{\sigma}$ as

$$\bar{r} = \frac{1}{T} \left(\int_0^T r(t)dt \right) \quad \bar{\sigma} = \left(\frac{1}{T} \int_0^T \sigma^2(t)dt \right)^{1/2}.^3$$

That the distribution of $\log(S(T))$ is normal follows from the definition of the Itô process: By

$$d(\log(S(t))) = (r(t) - \frac{1}{2}\sigma^2(t))dt + \sigma(t)dW^{\mathbb{Q}^N}(t)$$

we have just stated that

$$\begin{aligned} \log(S(T)) &= \log(S(0)) + \int_0^T (r(t) - \frac{1}{2}\sigma^2(t))dt + \int_0^T \sigma(t)dW^{\mathbb{Q}^N}(t) \\ &= \log(S(0)) + \bar{r}T - \frac{1}{2}\bar{\sigma}^2T + \int_0^T \sigma(t)dW^{\mathbb{Q}^N}(t) \\ &= \log(S(0)) + \bar{r}T - \frac{1}{2}\bar{\sigma}^2T + \bar{\sigma}W^{\mathbb{Q}^N}(T). \end{aligned}$$

² Since B has no stochastic component, i.e. does not depend on the path parameter $\omega \in \Omega$, we have that (4.1) holds under $\mathbb{Q}^N$ too.

We thus know the dynamics of $S(t)$ and $B(t)$ and more importantly the distribution of $S(T)$ under the measure $\mathbb{Q}^N$. Furthermore we know that the value of $V(0)$ satisfies

$$\frac{V(0)}{B(0)} = E^{\mathbb{Q}^N} \left(\frac{V(T)}{B(T)} \right),$$

thus (with $B(0) = 1$, $B(T) = \exp(\bar{r}T)$)

$$\begin{aligned} V(0) &= B(0) \cdot E^{\mathbb{Q}^N} \left(\frac{V(T)}{B(T)} \right) = 1 \cdot E^{\mathbb{Q}^N} \left(\frac{\max(S(T) - K, 0)}{\exp(\bar{r}T)} \right) \\ &= \exp(-\bar{r}T) \cdot E^{\mathbb{Q}^N} \left(\max(S(T) - K, 0) \right) \\ &= \exp(-\bar{r}T) \cdot E^{\mathbb{Q}^N} \left(\max(\exp(\log(S(T))) - K, 0) \right) \\ &= \exp(-\bar{r}T) \cdot \int_{-\infty}^{\infty} \max(\exp(y) - K, 0) \cdot \frac{1}{\bar{\sigma} \sqrt{T}} \phi \left(\frac{y - \bar{\mu}}{\bar{\sigma} \sqrt{T}} \right) dy, \end{aligned}$$

where $\phi(x) := \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)$ denotes the density of the standard normal distribution. The above integral may be represented as

$$V(0) = S(0)\Phi(d_+) - \exp(-\bar{r}T)K\Phi(d_-), \quad (4.3)$$

where $\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp(-\frac{y^2}{2}) dy$ and $d_{\pm} = \frac{1}{\bar{\sigma} \sqrt{T}} [\log(\frac{S(0)}{K}) + \bar{r}T \pm \frac{\bar{\sigma}^2 T}{2}]$.⁴

To derive this formula we consider for given $K, \mu, \sigma \in \mathbb{R}$ ($\sigma > 0$)

$$\begin{aligned} &\int_{-\infty}^{\infty} \max(\exp(y) - K, 0) \cdot \frac{1}{\sigma} \phi \left(\frac{y - \mu}{\sigma} \right) dy \\ &= \int_{\log(K)}^{\infty} (\exp(y) - K) \cdot \frac{1}{\sqrt{2\pi}\sigma} \exp \left(-\frac{(y - \mu)^2}{2\sigma^2} \right) dy \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{\log(K)}^{\infty} \exp(y) \cdot \exp \left(-\frac{(y - \mu)^2}{2\sigma^2} \right) dy \\ &\quad - \frac{1}{\sqrt{2\pi}\sigma} K \int_{\log(K)}^{\infty} \exp \left(-\frac{(y - \mu)^2}{2\sigma^2} \right) dy. \end{aligned}$$

It is

$$\begin{aligned} &\frac{1}{\sqrt{2\pi}\sigma} \int_{\log(K)}^{\infty} \exp(y) \cdot \exp \left(-\frac{(y - \mu)^2}{2\sigma^2} \right) dy \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{\log(K)}^{\infty} \exp \left(-\frac{(y - \mu)^2 - 2\sigma^2 y}{2\sigma^2} \right) dy \end{aligned}$$

⁴ We denote the cumulative normal distribution function by Φ . In some books, especially in connection with the Black-Scholes model, it is denoted by N (which we usually use for the Numéraire). Instead of d_+ , d_- one often uses the symbols d_1 , d_2 .

with $(y - \mu)^2 - 2\sigma^2 y = (y - (\mu + \sigma^2))^2 - 2\mu\sigma^2 - \sigma^4$

$$\begin{aligned} &= \frac{1}{\sqrt{2\pi}\sigma} \int_{\log(K)}^{\infty} \exp\left(-\frac{(y - (\mu + \sigma^2))^2}{2\sigma^2} + \mu + \frac{1}{2}\sigma^2\right) dy \\ &= \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\mu + \frac{1}{2}\sigma^2\right) \int_{\log(K)}^{\infty} \exp\left(-\frac{(y - (\mu + \sigma^2))^2}{2\sigma^2}\right) dy \end{aligned}$$

and with the substitution $\frac{y - (\mu + \sigma^2)}{\sigma} \mapsto y$

$$= \frac{1}{\sqrt{2\pi}} \exp\left(\mu + \frac{1}{2}\sigma^2\right) \int_{\frac{\log(K) - (\mu + \sigma^2)}{\sigma}}^{\infty} \exp\left(-\frac{y^2}{2}\right) dy.$$

Similarly we have with the substitution $\frac{y - \mu}{\sigma} \mapsto y$

$$\begin{aligned} &\frac{1}{\sqrt{2\pi}\sigma} K \int_{\log(K)}^{\infty} \exp\left(-\frac{(y - \mu)^2}{2\sigma^2}\right) dy \\ &= \frac{1}{\sqrt{2\pi}} K \int_{\frac{\log(K) - \mu}{\sigma}}^{\infty} \exp\left(-\frac{y^2}{2}\right) dy \end{aligned}$$

Defining $\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{y^2}{2}\right) dy$ we thus get from $\frac{1}{\sqrt{2\pi}} \int_x^{\infty} \exp\left(-\frac{y^2}{2}\right) dy = \Phi(-x)$ that

$$\begin{aligned} &\int_{-\infty}^{\infty} \max(\exp(y) - K, 0) \cdot \frac{1}{\sigma} \phi\left(\frac{y - \mu}{\sigma}\right) dy \\ &= \exp\left(\mu + \frac{1}{2}\sigma^2\right) \Phi\left(\frac{\mu + \sigma^2 - \log(K)}{\sigma}\right) - K \Phi\left(\frac{\mu - \log(K)}{\sigma}\right). \end{aligned}$$

Remark 71 (Implied Black-Scholes Volatility): From Equation (4.3) and the definition of $d_{\pm}$ we see that under the the model (4.2) the price of an option depends on $S(0)$, $\bar{r}$ and $\bar{\sigma}$ only (apart from product parameters like T and K). The parameter

$$\bar{\sigma} = \left(\frac{1}{T} \int_0^T \sigma^2(t) dt\right)^{1/2}$$

For fixed $S(0)$, $\bar{r}$ the pricing formula (4.3) is a bijection

$$\begin{aligned} &\bar{\sigma} \mapsto V(0) \\ &[0, \infty) \rightarrow [\max(S(0) - \exp(-\bar{r}T)K, 0), S(0)] \end{aligned}$$

and thus we may calculate the Black-Scholes volatility corresponding to a given price $V(0)$. This volatility is called the implied Black-Scholes volatility.

CHAPTER 5

Excursion: The Density of the Underlying of an European Call Option

Lemma 72 (Probability density of the underlying of an European call option): Let S denote a price process and N a Numéraire on the probability space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{Q}^N)$. Let $K \in \mathbb{R}$ and $T_0 < T_1 \leq T_2$, where T_1 denotes a [fixing date](#) and T_2 a [payment date](#) of a European option with payout

$$V(K; T_2) := \max(S(T_1) - K, 0).$$

Let us further assume that the Numéraire N has been chosen such that $N(T_2) = 1$. Then we have for the probability density¹ $\phi_{S(T_1)}(s)$ of $S(T_1)$ under $\mathbb{Q}^N$

$$\phi_{S(T_1)}(s) = \frac{1}{N(T_0)} \cdot \left. \frac{\partial^2 V(K, T_0)}{\partial K^2} \right|_{K=s},$$

i.e.

$$\frac{\partial^2 V(K, T_0)}{\partial K^2} = N(T_0) \cdot \phi_{S(T_1)}(K) = \frac{d\mathbb{Q}^N}{dS}(\{\omega : S(\omega) = K\}).$$

Remark 73 (Application): We may apply Lemma 72 to European stock options ($T_1 = T_2$ and $N(t) = \exp(r \cdot (T_1 - t))$), see Section 4, or to Caplets, see Section 10.

If some model is given (e.g. as a “black box” through some pricing software) and if the model allows to price arbitrary European options, then we may use this Lemma to infer the probability distribution of $S(T)$ created by this model. This allows for some simple model tests: If the probability density is 0 in some region, then the model should be put into question if these regions play any role for the pricing. If the model

¹ To be precise, $\phi_{S(T_1)}(s)$ is the conditional probability density, conditioned on $\mathcal{F}_{T_0}$. For simplicity we assume that $\mathcal{F}_{T_0} = \{\emptyset, \Omega\}$ and write $E^{\mathbb{Q}^N}(\cdot)$ for $E^{\mathbb{Q}^N}(\cdot | \mathcal{F}_{T_0})$.

exhibits regions with negative probability density, then the model is not arbitrage-free. It is easy to setup interpolatory models that fail if tested by this lemma. See Chapter 6.

Proof: Under the measure $\mathbb{Q}^N$ we have that N -relative prices are martingales. Thus we have

$$\frac{V(K; T_0)}{N(T_0)} = \mathbb{E}^{\mathbb{Q}^N} \left(\frac{V(K; T_2)}{N(T_2)} \right)$$

i.e.

$$\begin{aligned} V(K; T_0) &= N(T_0) \cdot \mathbb{E}^{\mathbb{Q}^N} \left(\frac{V(K; T_2)}{N(T_2)} \right) \\ &= N(T_0) \cdot \int_{\Omega} \max(S(T_1; \omega) - K, 0) d\mathbb{Q}^N(\omega) \\ &= N(T_0) \cdot \int_{-\infty}^{\infty} \max(s - K, 0) \cdot \phi_{S(T_1)}(s) ds, \end{aligned}$$

and thus

$$\begin{aligned} \frac{\partial}{\partial K} V(K; T_0) &= N(T_0) \cdot \int_{-\infty}^{\infty} -1_{[0, \infty)}(s - K) \phi_{S(T_1)}(s) ds \\ &= N(T_0) \cdot \int_K^{\infty} -\phi_{S(T_1)}(s) ds, \end{aligned}$$

i.e.

$$\frac{\partial^2}{\partial K^2} V(K; T_0) = N(T_0) \cdot \phi_{S(T_1)}(K).$$

□

As a direct consequence we have

Lemma 74 (Prices of European options are convex): Prices of European options are a convex function of the strike.

CHAPTER 6

Excursion: Interpolation of European Option Prices

Insufficient facts always invite danger.

Spock
Space Seed, stardate 3141.9, (Wikiquote).

6.1. No-Arbitrage Conditions for Interpolated Prices

We consider the price $V(K; 0)$ of an European option with payoff $\max(S(T) - K, 0)$ as a function of the strike K . We assume $S(T) \geq 0$. Let $P(T; 0)$ denote the value of the zero-coupon bond with nominal 1. From Lemma 74 we have for the mapping

$$K \mapsto V(K)$$

(we write shortly $V(K)$ for $V(K; 0)$), that maps a strike K to the corresponding option price $V(K)$ (for fixed, given maturity T and notional 1) that the following conditions hold:

- $V(0) = P(T; 0) \cdot S(0)$ (since $S(T) \geq 0$)
- $V(K) = V(0) - P(T; 0) \cdot K$ for $K < 0$ (since $S(T) \geq 0$), i.e. $V(K)$ is linear for $K < 0$,
- $\lim_{K \rightarrow \infty} V(K) = 0$,
- $\frac{d^2 V}{dK^2} \geq 0$ (from Lemma 72), i.e. $V(K)$ is convex.

Given strikes $K_1 < K_2$ and option prices $V(K_1)$ bzw. $V(K_2)$, we find geometric conditions for the possible (arbitrage-free) strike-price-points $(K, V(K))$. The price points $c = (K_1, V(K_1))$ and $d = (K_2, V(K_2))$ as well as $a = (-1, V(0) - P(T; 0))$, $b = (0, V(0))$ and $e = (0, V(K_2))$ define lines $\overline{ab}$, $\overline{bc}$, $\overline{cd}$ and $\overline{de}$ which are boundaries of the set of admissible strike-price points. As depicted in Figure 6.1 we have:

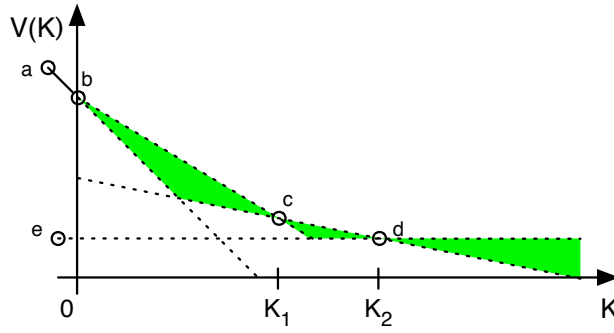


Figure 6.1.: Arbitrage-free option prices for European options (payoff $\max(S(T) - K, 0)$) with strike $K < K_1$ or $K_1 < K < K_2$ or $K > K_2$; assuming that two option prices $V(K_1) > V(K_2)$ (for $K_1 < K_2$) are given.

- For $K < K_1 < K_2$:
 - $V(K) \geq V(0) - P(T; 0) \cdot K$ $((K, V(K)) \geq \overline{ab})$.
 - $V(K) \geq V(K_1) \frac{K_2 - K}{K_2 - K_1} + V(K_2) \frac{K - K_1}{K_2 - K_1}$ $((K, V(K)) \geq \overline{cd})$.
 - $V(K) \geq V(0) + (V(K_1) - V(0)) \frac{K}{K_1}$ $((K, V(K)) \leq \overline{bc})$.
- For $K_1 < K < K_2$:
 - $V(K) \geq V(K_0) + (V(K_1) - V(0)) \frac{K}{K_1}$ $((K, V(K)) \geq \overline{bc})$.
 - $V(K) \geq V(K_2)$ $((K, V(K)) \geq \overline{de})$.
 - $V(K) \leq V(K_1) \frac{K_2 - K}{K_2 - K_1} + V(K_2) \frac{K - K_1}{K_2 - K_1}$ $((K, V(K)) \leq \overline{cd})$.
- For $K_1 < K_2 < K$:
 - $V(K) \geq V(K_1) \frac{K_2 - K}{K_2 - K_1} + V(K_2) \frac{K - K_1}{K_2 - K_1}$ $((K, V(K)) \geq \overline{cd})$.
 - $V(K) \leq V(K_2)$ $((K, V(K)) \leq \overline{de})$.

6.2. Arbitrage Violations through Interpolation

In this section we give some examples to show that simple interpolation of prices or implied volatilities may lead to arbitrage.

6.2.1. Example (1): Interpolation of four Prices

Consider a stock S with a current value of $S(0) = 1.0$ and an European option on S having maturity $T = 1.0$ and strike K , i.e. we consider the payoff

$$\max(S(1.0) - K, 0).$$

Assume that for strikes $K = 0.5, 0.75, 1.25, 2.0$ the following prices were given:

i	Strike K_i	Price $V(K_i)$	Implied Volatility $\sigma(K_i)$
1	0.50	0.5277	0.4
2	0.75	0.3237	0.4
3	1.25	0.1739	0.6
4	2.00	0.0563	0.6

The given implied volatility is the σ of a [Black-Scholes Modells](#) with assumed interest rate (short rate) of $r = 0.05$. Note however that the we start with given prices which are model independent.

We will now discuss several “obvious” interpolation methods applied to this data. One approach consists of the interpolation of prices, the other approach consists of the interpolation of implied volatilities and calculating interpolated prices from the interpolated volatilities. These interpolations are not model independent. The interpolation method itself constitutes a model. This is obvious for the second approach, where a model is involved in the calculation, but also applies to the first approach. From Lemma 72 we have that the interpolation method constitutes a model for the underlying’s probability density. It does not model the underlying’s dynamics.

6.2.1.1. Linear Interpolation of Prices

In Figure 6.2 we show the linear interpolation of the given option prices. The prices do not allow arbitrage, which is obvious from the convexity of their linear interpolation.

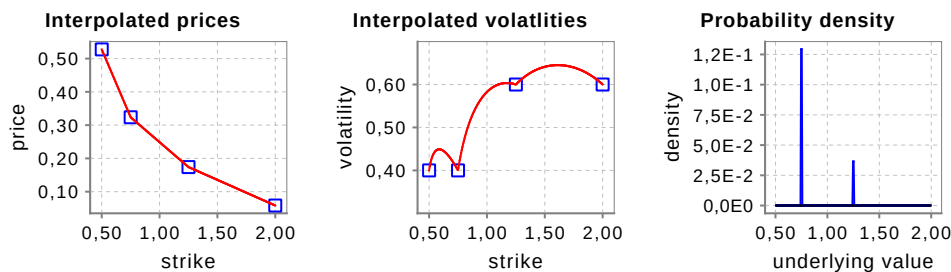


Figure 6.2.: Linear interpolation of option prices.

However, the linear interpolation of prices has severe disadvantages: The linear interpolation of option prices implies a model under which the underlying may not attain values $K \neq K_i$, the corresponding probability density is zero for these values. For values corresponding to the given strikes K_i a point measure is assigned, see Figure 6.2, right. In addition the corresponding implied volatilities look strange.

6.2.1.2. Linear Interpolation of Implied Volatilities

In Figure 6.2 we show the linear interpolation of the corresponding implied volatilities (center), while the corresponding prices are re-calculated via the Black-Scholes formula (left). The linear interpolation of implied (Black-Scholes) volatilities may lead

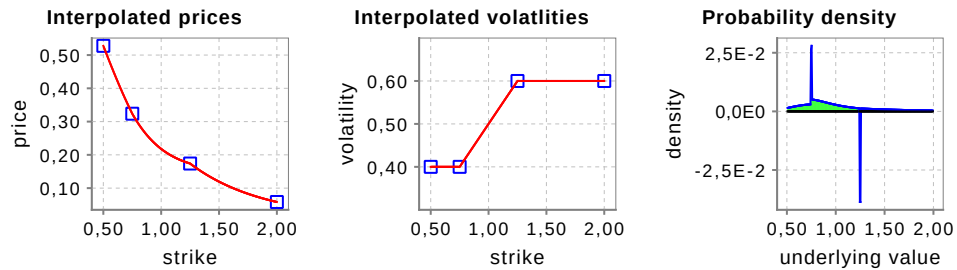


Figure 6.3.: Linear interpolation of implied volatilities.

to prices allowing for arbitrage: At the edges $K = 0.75$ and $K = 1.25$ we have a point measure, the measure for $S(T) = 1.25$ is negative. Thus the implied measure is not a probability measure, see Figure 6.2, right. From Lemma 72 we have: There is no arbitrage free pricing model that generates this interpolated price curve.

On the other hand, linear interpolation of implied volatilities also has a pleasant property: If the given prices correspond to prices from a Black-Scholes model, i.e. if the implied volatilities are constant, then, trivially, the interpolated prices correspond to the same Black-Scholes model.

6.2.1.3. Spline Interpolation of Prices respective Implied Volatilities

To remove the disadvantage of degenerated densities (i.e. the formation of point measures) as it appears for a linear interpolation, we move to a smooth (differentiable) interpolation method, e.g. spline interpolation. Figure 6.4 shows the spline interpolation of the given data. The upper row shows the spline interpolation of the option prices. The lower row shows the spline interpolation of the implied volatilities. Using (cubic) spline interpolation the probability densities are continuous functions. However it shows large regions with negative densities (Figure 6.4 right), thus the density is not a probability density and arbitrage possibilities exist. The strike-price-curve is not convex (Figure 6.4, left).

A spline interpolation of prices generates negative densities, because the spline interpolation of convex sample points is not convex.

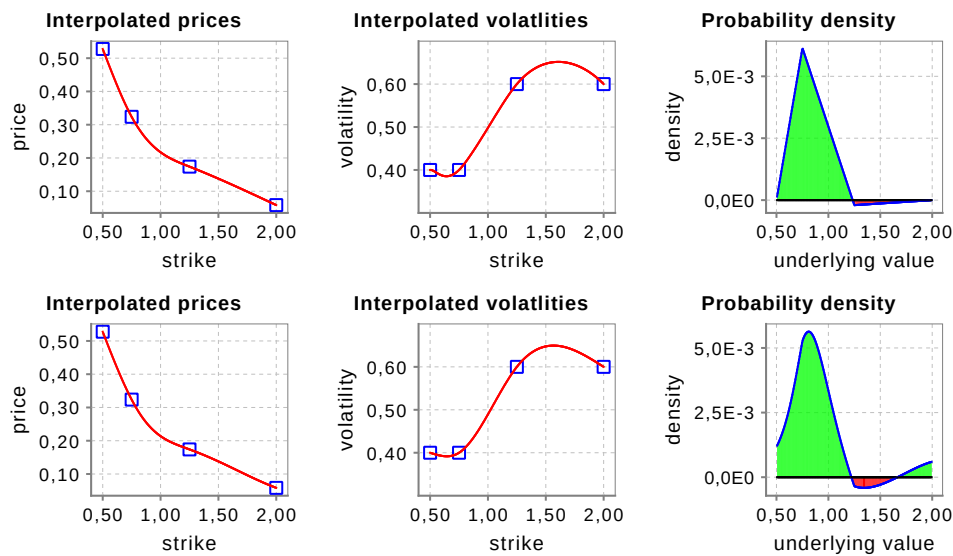


Figure 6.4.: Spline interpolation of option prices (upper row) and implied volatilities (lower row).

6.2.2. Example (2): Interpolation of two Prices

The example from Section 6.2.1 of a linear interpolation of implied volatilities may suggest that the problem arises at the joins of the linear interpolation, i.e. the behavior at $K_2 = 0.75$ and $K_3 = 1.25$. One could hope that a local smoothing would solve this problem. Instead we will give an example that a linear grow of implied volatility may be inadmissible alone, although the interpolated prices are admissible.

We consider two prices $V(K_2) > V(K_3)$. The monotony ensures that these two prices alone do not allow for an arbitrage.

6.2.2.1. Lineare Interpolation for decreasing Implied Volatilities

Given are the prices

i	Strike K_i	Price $V(K_i)$	Implied Volatility $\sigma(K_i)$
2	0.75	0.4599	0.9
3	1.25	0.0018	0.1

The implied volatility decreases with the strike K . Figure 6.5 shows the linear interpolation of the implied volatilities (center). The density (see Figure 6.5, right) shows a large region with negative values, thus it is not a probability density. The reason is a too fast decay of the implied volatility.

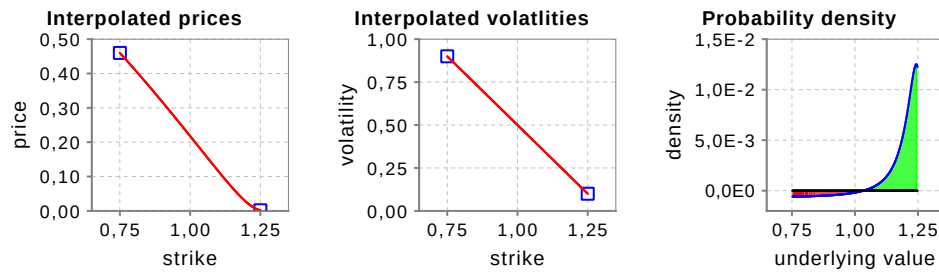


Figure 6.5.: Linear interpolation for decreasing implied volatility.

Conclusion: An arbitrarily fast decrease of implied volatility is not possible.

6.2.2.2. Linear Interpolation for increasing Implied Volatilities

For the example of increasing implied volatility we consider the prices

i	Strike K_i	Price $V(K_i)$	Implied Volatility $\sigma(K_i)$
2	0.75	0.2897	0.2
3	1.25	0.2532	0.8

Figure 6.5 shows the linear increasing interpolation of the implied volatilities.

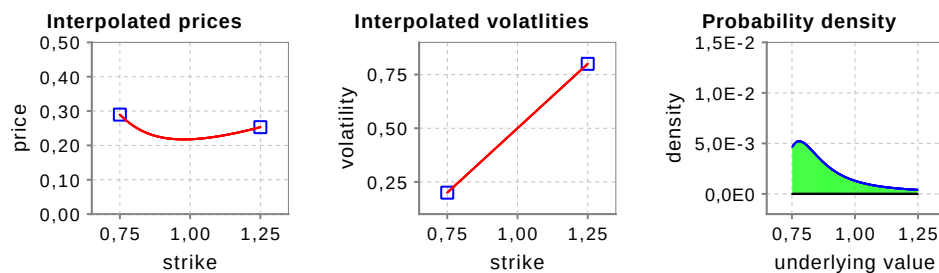


Figure 6.6.: Linear interpolation for increasing implied volatility.

At first sight the density implied by the linear interpolation of the implied volatilities exhibits no flaw (positive, no point measure). However, it is not a probability density. The integral below the segment is larger than 1. That this is inevitable is obvious from the price curve. The price curve is convex (and thus the density positive), but not monotone. The arbitrage is obvious, since the option with strike 1.0 is cheaper than the option with strike 1.25. If the prices should converge to 0 for increasing strike, then it is inevitable that the convexity will be violated. Thus it is inevitable that the

density will exhibit a region with negative values beyond the interpolation region (this however will make the integral of the density to 1).

Conclusion: An arbitrarily fast increase of implied volatility is not possible.

6.3. Arbitrage Free Interpolation of European Option Prices

The examples of the previous sections bring up the question for arbitrage free interpolation methods.

Every arbitrage free pricing model defines an arbitrage free interpolation method, if it is able to reproduce the given prices. However, this insight is almost useless, since:

- Most models are not able to reproduce arbitrarily given prices exactly.
 -
 - Extended models (e.g. models with stochastic volatility), which allow for a calibration to more than one option price per maturity, do this in an approximative way, i.e. the residual error of the fitting is minimized, but not necessarily 0. However, this may also be a desired effect, since such a fitting is more robust against errors in the input data.
- Extended models, that fit to more than one option price per maturity, usually require a great effort to find the corresponding model parameters. These models are intended for the pricing of complex derivatives (which justifies the effort), but not primarily as interpolation method to price European options. Examples of such model extensions are *stochastic volatility* or *jump-diffusion* extensions of the [LIBOR Market Model](#), [21].
- Some models even require a continuum of European option prices $K \mapsto V(T, K)$ as input. Then they require an interpolation. An example for such a model is the [Markov Functional Model](#), see Chapter 23.

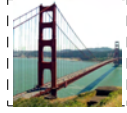
Thus, special methods and models have been developed to achieve specifically the interpolation of given European option prices. Often they reproduce given prices only approximately in the sense of a “best fit” – they are not an interpolation in the original sense. The construction of an (interpolating or approximating) price function $(T, K) \mapsto V(T, K)$ of European option prices with maturity T and strike K is called *modeling of the volatility surface*.

Definition 75 (Volatility Surface):

Given a continuum $(T, K) \mapsto V(T, K)$ of prices of European options with maturity T and strike K . Let $\sigma(T, K)$ denote the *implied volatility* of $V(T, K)$, i.e. the volatility

for which a risk neutral pricing with a log-normal model $dS = rSdt + \sigma(T, K)SdW$, $S(0) = S_0$ (Black-Scholes model) will reproduce the prices $V(T, K)$. Then $(T, K) \mapsto \sigma(T, K)$ is called *volatility surface*. $\lrcorner$

The modeling of a volatility surface may be viewed as important post processing of market data. On the other hand it may be viewed as integral part of the pricing model. A detailed discussion of volatility surface modeling is beyond the scope of this little excursion and we restrict ourself on citing a few methods:



Further Reading: The *mixture of lognormal* approach [52] uses the fact that if the probability density $\phi_S(T)$ of the underlying $S(T)$ is given as convex combination

$$\phi_{S(T)} = \sum_{i=1}^n \lambda_i \phi_{\sigma_i}, \quad \sum_{i=1}^n \lambda_i = 1, \quad \lambda_i \geq 0$$

of log-normal densities ϕ_{σ_i} , then the price of a European option is given as convex combination (using the same weights λ_i 's) of corresponding Black-Scholes formulas.¹ In some cases this allows for a choice of the ϕ_{σ_i} 's and λ_i 's that exactly reproduces given option prices. However, the method is not able to reproduce arbitrary arbitrage free prices.

The *SABR* model [66] is a four parameter model $(\sigma, \alpha, \beta, \rho)$ for European interest rate options (the underlying is an interest rate, see 8). It reproduces given prices only approximately, but has a more realistic behavior of the dependency of the modeled volatility surface from the spot value of the underlying. $\triangleleft$



Experiment: At <http://www.christian-fries.de/finmath/applets/OptionPriceInterpolation.html> several interpolation methods may be applied to a user defined configuration of European option prices. The figures in this chapter were produced with this software. $\triangleleft$

¹ This follows directly from the linearity of the integral with respect to the integrator $dQ = \phi(S)dS$.

CHAPTER 7

Hedging in Continuous and Discrete Time and the *Greeks*

An English version of this part is currently not available. What follows is the German version.

7.1. Introduction

Die vorstehende Bewertungstheorie baut darauf, dass ein Finanzderivat durch ein Portfolio aus gehandelten Produkten (Replikationsportfolio) und einer selbstfinanzierende Handelsstrategie zu replizieren ist.

Unter der Annahme einer perfekten Replikation war es durch Betrachtung des Martingalmaßes möglich ein derivatives Finanzprodukt zu bewerten, ohne explizit die replizierende Handelsstrategie zu bestimmen. Der aus dieser Theorie ermittelte Preis besitzt jedoch nur dann in der Praxis Relevanz, wenn die entsprechende Replikationsstrategie auch durchgeführt wird, also bekannt ist.

Um beliebige Auszahlungsfunktionen zu replizieren waren Bedingungen an die gehandelten Produkte und die Replikation zu stellen, insbesondere:

- Zur Replikation beliebiger Auszahlungsfunktionen über einem kontinuierlichem Zustandsraum, war es im Allgemeinen nötig kontinuierlich und in infinitesimalen Mengen in dem Portfolio zu handeln. Die Existenz einer solchen Handelsstrategie wurde theoretisch durch den Martingal Darstellungssatz (*Martingale Representation Theorem*) gesichert, siehe Satz 48, Seite 52.
- Zur Replikation beliebiger Auszahlungsfunktionen über einem endlichen, diskreten Zustandsraum, war es lediglich nötig an diskreten Zeiten zu handeln, jedoch musste die Menge der gehandelten Produkte der Menge der in einem Zeitschritt zu erreichenden Zustände entsprechen, siehe Abschnitt 3.1.2, Seite 65.

Beide Modellierungen (kontinuierliche oder diskrete Replikation) entsprechen nicht der Realität: Zum einen ist es nicht möglich Finanzprodukte kontinuierlich und in infinitesimalen Mengen zu handeln, zum anderen ist der diskrete Zustandsraum der real möglichen Preise meist bedeutend grösser als die zur Replikation zur Verfügung stehende Produkte.¹

Wir betrachten in diesem Kapitel die Frage, wie eine Handelsstrategie, d.h. das Replikationsportfolio, zu wählen ist, insbesondere, falls die Möglichkeit eines Handelns in infinitesimalen Zeiträumen nicht gegeben ist, sondern nur an diskreten Zeitpunkten stattfindet. Unter dieser Bedingung, die der Realität näher ist, existiert keine perfekte Replikation. Es bleibt ein Restrisiko.

Definition 76 (Hedge):

Ein Replikationsportfolio, welches ein derivatives Finanzprodukt repliziert, und damit zusammen mit dem derivativen Finanzprodukt eine Reduzierung des Gesamtrisikos bewirkt, wird als *Hedge* bezeichnet, die Handelsstrategie für das Replikationsportfolio als *hedging*.

Zunächst lassen sich ausgehend von der Bewertungstheorie der Delta-Hedge, Delta-Gamma-Hedge und Vega-Hedge als Handelsstrategien ableiten (Abschnitt 7.2) und auf den zeitdiskreten Fall (als dann jedoch imperfekte Replikation) übertragen (Abschnitt 7.4).

Die Methode von Bouchaud-Sornette (Abschnitt 7.5) bestimmt hingegen direkt ein Replikationsportfolio, welches das Restrisiko minimiert. Das Restrisiko ist im realen Maß $\mathbb{P}$ zu messen. Es ist ein reales Verlustrisiko in Folge von ungenügendem *Hedging*. Das reale Maß ist nun nicht länger irrelevant, da eine Replikation nicht länger möglich ist.

7.2. Deriving the Replications Strategy from Pricing Theory

Es sei $\mathcal{M} := \{S_0, S_1, \dots, S_n\}$ eine Menge von (Itô-)Preisprozessen, die einen vollständigen Markt gehandelter Produkte darstellt. Der Prozess $N = S_0$ sei ein Numéraire und $\mathbb{Q}^N$ sei das zugehörige Martingal-Maß. Es sei $t \mapsto V(t)$ der Preisprozess eines derivativen Finanzproduktes, d.h.

$$V(t) = N(t) \cdot \mathbb{E}^{\mathbb{Q}^N} \left(\frac{V(T)}{N(T)} \middle| \mathcal{F}_t \right).$$

Wir nehmen an, dass sich das Finanzprodukt V als Funktion von $(t, S_0(t), S_1(t), \dots, S_n(t))$ schreiben lässt, d.h. es ist

$$V(t) = V(t, S_0(t), S_1(t), \dots, S_n(t)). \quad (7.1)$$

¹ Die Werte, die eine Aktie an einer Börse annehmen kann sind in der Tat diskret.

Liegt $V(t)$ in dieser Weise durch eine vorhandene Bewertungstheorie als Funktion der Marktgrößen vor, so lässt sich das selbstfinanzierende Replikationsportfolio

$$\Pi(t) := \phi_0(t)S_0(t) + \phi_1(t)S_1(t) + \dots + \phi_n(t)S_n(t) \quad (7.2)$$

mit $\Pi(t) = V(t)$ (Replikation) wie folgt bestimmen:

Die Eigenschaft der Selbstfinanziertheit von Π schreibt sich als (siehe Definition 58)

$$d\Pi(t) = \phi_0(t)dS_0 + \phi_1(t)dS_1(t) + \dots + \phi_n(t)dS_n(t) = \sum_{i=0}^n \phi_i dS_i.$$

Desweiteren folgt aus Itô's Lemma (Satz 41) angewendet auf (7.1)

$$dV(t) = \frac{\partial V(t)}{\partial t}dt + \sum_{i=0}^n \frac{\partial V(t)}{\partial S_i}dS_i + \frac{1}{2} \sum_{i,j=0}^n \frac{\partial^2 V(t)}{\partial S_i \partial S_j} \underbrace{dS_i dS_j}_{=(...)dt}.$$

Mit $d\Pi(t) = dV(t)$ (Replikation), lassen sich nun $\phi_0, \dots, \phi_n$ durch Koeffizientenvergleich bestimmen. Es folgt

$$\phi_0(t) = \frac{\partial V(t)}{\partial S_0}, \phi_1(t) = \frac{\partial V(t)}{\partial S_1}, \dots, \phi_n(t) = \frac{\partial V(t)}{\partial S_n} \quad (7.3)$$

$$\frac{\partial V(t)}{\partial t}dt + \frac{1}{2} \sum_{i,j=0}^n \frac{\partial^2 V(t)}{\partial S_i \partial S_j} \underbrace{dS_i dS_j}_{=(...)dt} = 0. \quad (7.4)$$

Zusammenfassung: Unter einem Modell, welches eine Bewertung $V(t)$ mittels Martingalmaß, d.h. ohne die explizite Konstruktion des Replikationsportfolios, erlaubt, lässt sich das Replikationsportfolio *a posteriori* durch Berechnung der partiellen Ableitungen des Derivatepreises nach den Underlyings bestimmen.



Interpretation: Der Koeffizientenvergleich von dV und $d\Pi$ lieferte die Bestimmungsgleichungen (7.3) für den Portfolioprozess ϕ des Replikationsportfolios. Die *partielle Differentialgleichung* (7.4) ist eine andere Beschreibung der Tatsache, dass sich

$$V(t) = V(t, S_0(t), S_1(t), \dots, S_n(t))$$

in den *Koordinaten* $(S_0(t), S_1(t), \dots, S_n(t))$ selbstfinanzierend replizieren lässt. Diese Eigenschaft muss ja in V wiederzufinden sein. Die Gleichung (7.4) kann zusammen mit der Endbedingung

$$V(T, S_0(T), S_1(T), \dots, S_n(T)) = V^T$$

genutzt werden um bei gegebenem Auszahlungsprofil V^T den Wert $V(t)$ für $t < T$ zu bestimmen. Dies ist der Einstiegspunkt in die Derivatebewertung durch Lösen partieller Differentialgleichungen, siehe Abschnitt ??.

◁|

7.2.1. Deriving the Replication Strategy under the Assumption of a Locally Riskless Product

Der obige Ansatz $V(t) = V(t, S_0(t), S_1(t), \dots, S_n(t))$ ist nur ein möglicher. Er führt zu einer besonders natürlichen Darstellung der Portfolioparameter und der partiellen Differentialgleichung für V . In der Literatur, insbesondere wenn die S_i , $i \geq 1$ für Aktienpreisprozesse stehen, wird die obige Herleitung auf einen anderen Ansatz angewendet, den wir hier betrachten: Es sei $N = S_0$ ein lokal risikofreier Numeraire, d.h. wir nehmen an, dass der Prozess N (d.h. S_0) von der Form

$$dN = \frac{\partial N}{\partial t} dt \quad (7.5)$$

sei, d.h. keinen dW -Anteil besitze.²³⁴ Für $N > 0$ (N ist Numéraire) ist mit $r := \frac{\partial \log(N)}{\partial t}$

$$dN = \frac{\partial N}{\partial t} dt = \frac{\partial \log(N)}{\partial t} N dt = rN dt \quad (7.6)$$

Wir nehmen sodann an, dass sich $V(t)$ bei gegebenem N als Funktion in den Koordinaten $(t, S_1(t), \dots, S_n(t))$ interpretieren lässt und bezeichnen diese Funktion wiederum mit V , d.h. es ist

$$V(t) = V(t, S_1(t), \dots, S_n(t)). \quad (7.7)$$

Mit andern Worten: N wird nun nicht mehr als freie Modellgröße gesehen.⁵

Liegt $V(t)$ in dieser Form als Funktion der Marktgrößen vor, so lässt sich das selbstfinanzierende Replikationsportfolio (7.2) mit $\Pi(t) = V(t)$ wie folgt bestimmen: Aus Itô's Lemma (Satz 41) angewendet auf $V(t)$ folgt

$$dV(t) = \frac{\partial V(t)}{\partial t} dt + \sum_{i=1}^n \frac{\partial V(t)}{\partial S_i} dS_i + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 V(t)}{\partial S_i \partial S_j} \underbrace{dS_i dS_j}_{=(...)dt}.$$

Mit $d\Pi(t) = \sum_{i=0}^n \phi_i dS_i = dV(t)$ (Selbstfinanziertheit von Π , Replikation von V) ergeben sich $\phi_0, \dots, \phi_n$ durch Vergleich der Koeffizienten von dt und dS_i , $i \geq 1$ als

$$\begin{aligned} \phi_1(t) &= \frac{\partial V(t)}{\partial S_1}, \dots, \phi_n(t) = \frac{\partial V(t)}{\partial S_n} \\ \phi_0(t) \frac{\partial N}{\partial t} dt &= \frac{\partial V(t)}{\partial t} dt + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 V(t)}{\partial S_i \partial S_j} \underbrace{dS_i dS_j}_{=(...)dt}. \end{aligned} \quad (7.8)$$

² Der Prozess N ist dann *lokal* risikofrei.

³ Die Eigenschaft ist lokal, da $\frac{\partial N}{\partial t}$ durchaus noch stochastisch sein darf.

⁴ Ein Beispiel wäre $dN = r \cdot N \cdot dt$, wie es bei der Betrachtung des [Black-Scholes Modells](#) in Abschnitt 4 angenommen wurde.

⁵ Ein Beispiel ist das Black-Scholes Modell, in dem $N(t) = \exp(r \cdot t)$ ist und r ein gegebener Modellparameter (nicht aber eine Größe die selbst modelliert wird).

Die Bestimmungsgleichung für ϕ_0 erscheint nun im Vergleich zu der Bestimmungsgleichung für $\phi_1, \dots, \phi_n$ ungewöhnlich komplex. Dies liegt an der Darstellung (7.7) die wir für V gewählt haben, siehe Abschnitt 7.2.3. Da das Portfolio jedoch selbstfinanzierend ist, ist es (in der Praxis) nicht nötig ϕ_0 durch (7.8) zu bestimmen; ϕ_0 ergibt sich aus der Bedingung der Selbstfinanzierung $\sum_{i=0}^n (d\phi_i) \cdot S_i = 0$.

7.2.2. The Black-Scholes Differential Equation

Schreiben wir $\frac{\partial N}{\partial t} = \frac{\partial \log(N)}{\partial t} \cdot N =: rN$ so lautet die Gleichung für ϕ_0 in (7.8)

$$\phi_0(t)r(t)N(t)dt = \frac{\partial V(t)}{\partial t}dt + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 V(t)}{\partial S_i \partial S_j} \underbrace{dS_i dS_j}_{=(...)dt}.$$

Da nun aber $V = \phi_0 N + \sum_{i=1}^n \phi_i S_i \stackrel{(7.8)}{=} \phi_0 N + \sum_{i=1}^n \frac{\partial V(t)}{\partial S_i} S_i$ folgt $\phi_0 N = V - \sum_{i=1}^n \frac{\partial V(t)}{\partial S_i} S_i$ und mit $dS_i dS_j = \sigma_i \sigma_j \rho_{i,j} dt$ daher

$$\frac{\partial V(t)}{\partial t} + r(t) \sum_{i=1}^n \frac{\partial V(t)}{\partial S_i} S_i + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 V(t)}{\partial S_i \partial S_j} \sigma_i \sigma_j \rho_{i,j} = r(t)V(t)$$

Dies ist (eine Variante) der Black-Scholes Partiellen Differentialgleichung in den Koordinaten $(t, S_1, \dots, S_n)$.

7.2.3. The Derivative $V(t)$ as a Function of its Underlyings $S_i(t)$

Die Herleitung der Darstellung der Portfolioprozesskomponenten ϕ_i als partielle Ableitung von V sowie die von V erfüllte *partielle Differentialgleichung* hängt selbstverständlich vom gewählten Koordinatensystem der *Underlyings* ab. Statt eines Vergleiches der Koeffizienten der dS_i , $i \geq 1$ und dt (wie oben betrachtet) ist im allgemeinen Fall ebenso ein Vergleich der Koeffizienten von dt und den treibenden Brownschen Bewegungen dW_i möglich.

Wir betrachten V dargestellt in den beiden oben besprochenen Koordinaten, d.h.

$$\begin{aligned} V &= V^{n+1}(t, S_0(t), S_1(t), \dots, S_n(t)) \\ V &= V^n(t, S_1(t), \dots, S_n(t)) \end{aligned}$$

mit $S_0 = N$ und $dN(t) = r(t)N(t)dt$ (durch diese Annahme ist es erst möglich die Abhängigkeit von S_0 durch eine Abhängigkeit von t auszudrücken. Nun ist

$$\frac{\partial V^n}{\partial t} = \frac{\partial V^{n+1}}{\partial t} + \frac{\partial V^{n+1}}{\partial N} \cdot \frac{\partial N}{\partial t} = \frac{\partial V^{n+1}}{\partial t} + \phi_0 \cdot \frac{\partial N}{\partial t},$$

das heisst

$$\frac{\partial V^n}{\partial t} - \phi_0 \cdot \frac{\partial N}{\partial t} = \frac{\partial V^{n+1}}{\partial t}.$$

Dies und $dS_i dS_j = 0$ falls $i = 0$ oder $j = 0$ eingesetzt in (7.4) liefert

$$\begin{aligned} & \frac{\partial V(t)^{n+1}}{\partial t} dt + \frac{1}{2} \sum_{i,j=0}^n \frac{\partial^2 V^{n+1}(t)}{\partial S_i \partial S_j} \underbrace{dS_i dS_j}_{=(...)dt} = 0 \\ \Leftrightarrow & \frac{\partial V^n(t)}{\partial t} dt - \phi_0 \cdot \frac{\partial N}{\partial t} dt + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 V^n(t)}{\partial S_i \partial S_j} \underbrace{dS_i dS_j}_{=(...)dt} = 0 \\ \Leftrightarrow & \frac{\partial V^n(t)}{\partial t} + r(t) \sum_{i=1}^n \frac{\partial V^n(t)}{\partial S_i} S_i + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 V^n(t)}{\partial S_i \partial S_j} \sigma_i \sigma_j \rho_{i,j} = r(t) V^n(t) \end{aligned}$$

7.2.3.1. Pfadabhängige Optionen

Unsere Annahme, dass $V(t)$ eine Funktion von $S_i(t)$ ist schließt zunächst pfadabhängige Optionen aus. Allerdings kann die Betrachtung leicht auf solche übertragen werden, in dem V als Funktion der entsprechenden Pfadgrößen geschrieben wird. Zum Beispiel leitet man für eine pfadabhängige Option der Form $V(t) = V(t, S(t), \int_0^t S(\tau) d\tau)$ mit $I(t) = \int_0^t S(\tau) d\tau$ entsprechend $dV(t) = \frac{\partial V(t)}{\partial t} dt + \frac{\partial V(t)}{\partial S} dS + \frac{\partial V(t)}{\partial I} dI$ ab und setzt $dI(t) = S(t)dt$ ein. Für das Replikationsportfolio folgt in diesem Fall dann $\phi_1 = \frac{\partial V(t, S(t), I(t))}{\partial S}$.

7.2.4. Example: Replication Portfolio and Partial Differential Equation of an European Option under a Black-Scholes Model

Zur Verdeutlichung wiederholen wir die betrachteten Rechnungen an einem einfachen Modell, dem **Black-Scholes Modell** mit Preisprozessen für Sparkonto N und Aktie S

$$\begin{aligned} dN(t) &= rN(t)dt, & N(0) &= 1, \\ dS(t) &= \mu S(t)dt + \sigma S(t)dW(t), & S(0) &= S_0 \end{aligned}$$

Für eine europäische Option mit Strike K und Fälligkeit T , d.h. $\text{Payout } \max(S(T) - K, 0) \stackrel{!}{=} V(T)$, lautet die Black-Scholes-Merton Formel (4.3) für den Wert V zum Zeitpunkt t

$$V(t, S(t)) = S(t)\Phi(d_+) - \exp(-\bar{r}(T-t))K\Phi(d_-),$$

mit $d_{\pm} = \frac{1}{\sigma\sqrt{T-t}} \left[\log\left(\frac{S(t)}{K}\right) + \bar{r}(T-t) \pm \frac{\sigma^2(T-t)}{2} \right]$, siehe (4.3) auf Seite 87 in Kapitel 4.

Dies beschreibt V als Funktion von t und $S(t)$, da r als Konstant angesehen wird und somit $N(t)$ als Funktion von t bekannt ist. Mit $N(t) = \exp(r \cdot t)$ schreibt sich V aber als Funktion von $(t, N(t), S(t))$ durch

$$V(t, N(t), S(t)) = S(t)\Phi(d_+) - N(t)\frac{K}{N(T)}\Phi(d_-) \quad (7.9)$$

mit

$$d_{\pm} = \frac{1}{\bar{\sigma} \sqrt{T-t}} \left[\log\left(\frac{S(t)}{N(t)} \cdot \frac{N(T)}{K}\right) \pm \frac{\bar{\sigma}^2(T-t)}{2} \right].$$

Es ist

$$\begin{aligned} \frac{\partial V}{\partial S(t)} &= \Phi(d_+) + S(t)\phi(d_+) \frac{1}{S(t)} \frac{1}{\sigma \sqrt{T-t}} - N(t) \frac{K}{N(T)} \phi(d_-) \frac{1}{S(t)} \frac{1}{\sigma \sqrt{T-t}} \\ &= \Phi(d_+) + \frac{1}{S(t)} \frac{1}{\sigma \sqrt{T-t}} \cdot (S(t)\phi(d_+) - N(t) \frac{K}{N(T)} \phi(d_-)) \\ &= \Phi(d_+) + \frac{1}{S(t)} \frac{1}{\sigma \sqrt{T-t}} \cdot (S(t)\phi(d_+) - N(t) \frac{K}{N(T)} \phi(d_-)) \end{aligned}$$

Mit

$$\begin{aligned} \phi(d_{\pm}) &= C \cdot \exp\left(-\frac{1}{2}d_{\pm}^2\right) \\ &= C \cdot \exp\left[-\frac{\log\left(\frac{S(t)}{N(t)} \cdot \frac{N(T)}{K}\right)^2 \pm 2 \log\left(\frac{S(t)}{N(t)} \cdot \frac{N(T)}{K}\right) \cdot \frac{\bar{\sigma}^2(T-t)}{2} + \frac{\bar{\sigma}^4(T-t)^2}{4}}{2\bar{\sigma}^2(T-t)}\right] \\ &= C \cdot \exp\left[\pm \log\left(\frac{1}{2} \frac{S(t)}{N(t)} \cdot \frac{N(T)}{K}\right) - \frac{\log\left(\frac{S(t)}{N(t)} \cdot \frac{N(T)}{K}\right)^2 + \frac{\bar{\sigma}^4(T-t)^2}{4}}{2\bar{\sigma}^2(T-t)}\right] \\ &= \left(\frac{S(t)}{N(t)} \cdot \frac{N(T)}{K}\right)^{\pm 1/2} \cdot C \cdot \exp\left[-\frac{\log\left(\frac{S(t)}{N(t)} \cdot \frac{N(T)}{K}\right)^2 + \frac{\bar{\sigma}^4(T-t)^2}{4}}{2\bar{\sigma}^2(T-t)}\right] \end{aligned}$$

folgt $\phi(d_-) = \frac{S(t)}{N(t)} \cdot \frac{N(T)}{K} \cdot \phi(d_+)$ und so

$$\frac{\partial V}{\partial S(t)} = \Phi(d_+) + \frac{1}{S(t)} \frac{1}{\sigma \sqrt{T-t}} \cdot \underbrace{(S(t)\phi(d_+) - N(t) \frac{K}{N(T)} \phi(d_-))}_{=0} = \Phi(d_+)$$

Analog zeigt man

$$\frac{\partial V}{\partial N(t)} = -\frac{N(T)}{K} \cdot \Phi(d_-).$$

Das heisst es ist $V(t, N(t), S(t)) = \phi_0(t)N(t) + \phi_1(t)S(t)$ mit

$$\phi_0(t) = -\frac{N(T)}{K} \cdot \Phi(d_-), \quad \phi_1(t) = \Phi(d_+).$$

In der Tat hätten wir diese Darstellung des Replikationsportfolios einfach in (7.9) ablesen können.

Für die Funktion

$$V(t, n, s) = s\Phi(d_+) - n \frac{K}{N(T)} \Phi(d_-)$$

mit

$$d_{\pm} = \frac{1}{\bar{\sigma} \sqrt{T-t}} \left[\log\left(\frac{s}{n} \cdot \frac{N(T)}{K}\right) \pm \frac{\bar{\sigma}^2(T-t)}{2} \right].$$

gilt nun

$$\frac{\partial V(t)}{\partial t} dt + \frac{1}{2} \frac{\partial^2 V(t)}{\partial n \partial n} dN dN + \frac{\partial^2 V}{\partial n \partial s} dN dS + \frac{1}{2} \frac{\partial^2 V(t)}{\partial s \partial s} dS dS = 0$$

$$dN dN = 0$$

$$dN dS = 0$$

$$\iff$$

$$\frac{\partial V(t)}{\partial t} dt + \frac{1}{2} \frac{\partial^2 V(t)}{\partial s \partial s} \sigma^2 = 0.$$

Siehe Übung 11. Ferner gilt die Endbedingung

$$V(T, N(T), s) = \max(s - K, 0)$$

7.2.4.1. Interpretation von V als Funktion in (t, S)

Die klassische Betrachtungsweise sieht jedoch V nur als Funktion von t und $S(t)$ an. In diesem Falle erhalten wir (siehe (7.8))

$$\begin{aligned} \phi_1 &= \frac{\partial V(t)}{\partial S} = \Phi(d_+) \\ \phi_0 r N &= \left(\frac{\partial V(t)}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V(t)}{\partial S^2} \right), \end{aligned}$$

Im übrigen gilt natürlich mit $V = \phi_0 N + \phi_1 S$

$$\phi_0 N = V - \phi_1 S,$$

was uns nochmals auf die Black-Scholes Partielle Differential Gleichung bringt

$$\begin{aligned} rV &= r\phi_0(t)N(t) + r\phi_1 S(t) \\ &= \frac{\partial V(t)}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V(t)}{\partial S^2} + r \frac{\partial V(t)}{\partial S} S \end{aligned}$$

und somit

$$\frac{\partial V(t)}{\partial t} + r \frac{\partial V(t)}{\partial S} S + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V(t)}{\partial S^2} = rV.$$

Definition 77 (Delta-Hedge):

Die Wahl eines Replikationsportfolios entsprechend (7.2), (7.3) wird als *Delta-Hedge* bezeichnet.



Interpretation: Betrachte man das Derivat V als eine Funktion der Underlying(s) S , so ist das Portfolio Π , gewählt entsprechend des Delta-Hedge, die lokale lineare Approximation von V , siehe auch Abbildung 7.2 auf Seite 112. Für einen festen Pfad ω nimmt $t \mapsto S(t, \omega)$ (fast sicher) einen stetigen Verlauf (Pfade von S sind als Itô-Prozesse fast sicher stetig) und auf dem Pfad ist $V(T) = \int_0^T dV = \int_0^T d\Pi$. Diese Aussage entspricht dem Hauptsatz der Integralrechnung: $f(T) = \int_0^T df = \int_0^T f'(t)dt$ (z.B. für differenzierbares f). Dies verdeutlicht, warum das Portfolio Π nur bezüglich seiner ersten Ableitung nach S mit dem Derivat V übereinstimmen muss, also bzgl. des *Deltas*: $\frac{\partial V}{\partial S} = \frac{\partial \Pi}{\partial S}$. $\triangleleft$

7.3. Greeks

Die partiellen Ableitungen des Preises $V(t)$ eines Derivates nach Produkt- und Modellparametern sind wichtige Kenngrößen. Sie beschreiben, wie das Derivat unter dem angenommen Modell auf infinitesimale Parameteränderungen reagiert, und sind maßgeblich für die Zusammensetzung des Replikationsportfolios. Diese Kenngrößen werden meist mit griechischen Buchstaben bezeichnet und daher *Greeks* (Griechen) genannt. Wir definieren die wichtigsten:

Definition 78 (Delta):

Mit *Delta* werden die ersten partiellen Ableitung des Preises eines derivativen Produktes nach den Underlyings bezeichnet, d.h. nach den am Markt gehandelten Produkten, welche in die Bewertung des derivativen Produktes eingehen.

Definition 79 (Gamma):

Mit *Gamma* werden die zweiten partiellen Ableitungen des Preises eines derivativen Produktes nach den Underlyings bezeichnet, d.h. nach den am Markt gehandelten Produkten, welche in die Bewertung des derivativen Produktes eingehen.

Definition 80 (Vega):

Mit *Vega* werden die ersten partiellen Ableitungen des Preises eines derivativen Produktes nach der Volatilität der Log-Preisprozesse der Underlyings bezeichnet. In einem Black-Scholes Modell ist dies die partielle Ableitung nach σ .

Definition 81 (Theta):

Mit *Theta* wird die erste partielle Ableitung des Preises eines derivativen Produktes nach der Zeit bezeichnet.

Definition 82 (Rho):

Mit *Rho* wird die erste partielle Ableitung des Preises eines derivativen Produktes nach der Zinsrate bezeichnet.

Die Definition von *Rho* wird lediglich für Aktienderivate verwendet. Für Zinsderivate, wie wir sie in den folgenden Kapitel betrachten werden, kann die Zinsrate als Underlying interpretiert und *Rho* wird als (ein) *Delta* bezeichnet. $\lrcorner$

7.3.1. Greeks einer europäischen Call-Option unter dem Black-Scholes Modell

Gegeben sei das **Black-Scholes Modell** mit Preisprozessen für Sparkonto N und Aktie S

$$\begin{aligned} dN(t) &= rN(t)dt, & N(0) &= 1, \\ dS(t) &= \mu S(t)dt + \sigma S(t)dW(t), & S(0) &= S_0, \end{aligned}$$

sowie eine Europäische Option mit *Payout* $\max(S(T) - K, 0) \stackrel{!}{=} V(T)$. Entsprechend Kapitel 4 ist unter diesem Modell der Preis $V(0)$ der Option gegeben durch die Black-Scholes-Merton Formel (4.3):

$$V(0) = S_0 \Phi(d_+) - \exp(-rT)K \Phi(d_-),$$

wobei $\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp(-\frac{y^2}{2}) dy$ und $d_{\pm} = \frac{1}{\sigma \sqrt{T}} [\log(\frac{S(0)}{K}) + rT \pm \frac{\sigma^2 T}{2}]$. Der Preis ist also eine Funktion von S_0 , r , K , σ und T .

Aus der Black-Scholes-Merton Formel lassen sich die Greeks analytisch bestimmen. Sie sind in Tabelle 7.1 zusammengestellt.

Greek	Definition	im Black-Scholes Modell
Delta (Δ)	$\frac{\partial V}{\partial S_0}$	$= \Phi(d_+)$
Gamma (Γ)	$\frac{\partial^2 V}{\partial S_0^2}$	$= \frac{\Phi'(d_+)}{S_0 \sigma \sqrt{T}}$
Vega	$\frac{\partial V}{\partial \sigma}$	$= S_0 \sqrt{T} \Phi'(d_+)$
Theta (Θ)	$-\frac{\partial V}{\partial T}$	$= -\frac{S_0 \Phi'(d_+) \sigma}{2 \sqrt{T}} - rK \exp(-rT) \Phi(d_-)$
Rho (ρ)	$\frac{\partial V}{\partial r}$	$= KT \exp(-rT) \Phi(d_-)$

Table 7.1.: *Greeks im Black-Scholes Modell.*

7.4. Hedging in Discrete Time: Delta- and Delta-Gamma-Hedging

Wird ein Delta-Hedge kontinuierlich angewendet, so liefert er eine exakte Replikation. Das Portfolio ist gegenüber infinitesimalen Änderungen der *Underlyings* S_i neutral. Wird der Delta-Hedge hingegen nur an diskreten Zeitpunkten T_i vorgenommen, d.h. wird der Portfolio-Prozess $(\phi_0, \phi_1, \dots, \phi_n)$ auf Zeitintervallen $[T_i, T_{i+1})$ konstant gewählt, so ist die Replikation nicht exakt.⁶ Während im kontinuierlichen

$$\begin{aligned} dV(t) &= \frac{\partial V(t)}{\partial t} dt + \sum_{i=0}^n \frac{\partial V(t)}{\partial S_i} dS_i + \frac{1}{2} \sum_{i,j=0}^n \frac{\partial^2 V(t)}{\partial S_i \partial S_j} dS_i dS_j, \\ &= \sum_{i=0}^n \frac{\partial V(t)}{\partial S_i} dS_i + \underbrace{\frac{\partial V(t)}{\partial t} dt + \frac{1}{2} \sum_{i,j=0}^n \frac{\partial^2 V(t)}{\partial S_i \partial S_j} dS_i dS_j}_{=0}, \end{aligned}$$

nach Itô's Lemma (und (7.4)) exakt ist, gilt im zeitdiskreten Fall ($\Delta V = V(t + \Delta t) - V(t)$)

$$\Delta V(t) = \frac{\partial V(t)}{\partial t} \Delta t + \sum_{i=0}^n \frac{\partial V(t)}{\partial S_i} \Delta S_i + \frac{1}{2} \sum_{i,j=0}^n \frac{\partial^2 V(t)}{\partial S_i \partial S_j} \underbrace{\Delta S_i \Delta S_j}_{\neq (\dots) \Delta t} \quad (7.10)$$

$$= \sum_{i=0}^n \frac{\partial V(t)}{\partial S_i} \Delta S_i + \frac{1}{2} \sum_{i,j=0}^n \frac{\partial^2 V(t)}{\partial S_i \partial S_j} \underbrace{(\Delta S_i \Delta S_j - \gamma_{i,j} \Delta t)}_{\neq 0} + \text{h.o.t.}, \quad (7.11)$$

wobei $\gamma_{i,j}$ durch $dS_i dS_j = \gamma_{i,j} dt$ definiert sei und $\text{h.o.t.} = O(|\Delta t|^2, |\Delta t \Delta S_i|, |\Delta S_i|^3)$ ist.⁷ Zum Vergleich dazu gilt für das Replikationsportfolio im zeit-diskreten Fall

$$\Delta \Pi(t) = \sum_{i=0}^n \phi_i \Delta S_i. \quad (7.12)$$

Nehmen wir an, dass die Replikation zur Zeit t perfekt ist ($V(t) = \Pi(t)$), so existiert im Allgemeinen keine Wahl des Replikationsportfolios ϕ so dass $\Delta V(t) = \Delta \Pi(t)$ und somit eine perfekte Replikation zum Zeitpunkt $t + \Delta t$ garantiert.

Soll das Replikationsportfolio selbstfinanzierend sein, so muss diese Anforderung im Zeitdiskreten betrachtet werden. Wird zu Zeitpunkten t_k gehandelt so muss die Selbstfinanzierungseigenschaft

$$\sum_{i=0}^n (\phi_i(t_k) - \phi_i(t_{k-1})) S_i(t_k) = 0$$

⁶ Siehe auch die Diskussion des Begriffes *selbstfinanzierend* (Definition 58).

zum Beispiel durch die Wahl von ϕ_0

$$\phi_0(t_k) = \phi_0(t_{k-1}) - \frac{1}{S_0(t_k)} \cdot \left(\sum_{i=1}^n (\phi_i(t_k) - \phi_i(t_{k-1})) S_i(t_k) \right) \quad (7.13)$$

sichergestellt werden.⁸ Während im kontinuierlichen V selbstfinanzierend perfekt repliziert werden konnte, stellt sich hier die Frage, ob ϕ_0 derart gewählt wird, dass der Replikationsfehler verringert wird, oder derart, dass das Portfolio selbstfinanzierend bleibt.

Die Restriktion auf ein selbstfinanzierendes Portfolio führt über (7.13) zu einer Fehlerfortpflanzung des Replikationsfehlers, siehe Abschnitt 7.4.2.

7.4.1. Delta Hedging

Beim Delta Hedging wird durch die Wahl des Replikationsportfolios das Delta von Π entsprechend dem Delta von V gewählt und so das Delta des Replikationsfehlers $V - \Pi$ zu Null gemacht. Der Replikationsfehler hängt also *in erster Näherung* nicht von den Bewegungen der Underlyings ab. Berücksichtigen wir die Forderung nach der Selbstfinanzierung, so wählen wir

$$\begin{aligned} \phi_i(t_k) &= \frac{\partial V(t_k)}{\partial S_i} \\ \phi_0(t_k) &= \phi_0(t_{k-1}) - \frac{1}{S_0(t_k)} \cdot \left(\sum_{i=1}^n (\phi_i(t_k) - \phi_i(t_{k-1})) S_i(t_k) \right) \end{aligned}$$

Für den Fehler $|\Delta V(t) - \Delta \Pi(t)|$ erhalten wir mit (7.11) und (7.12)

$$\begin{aligned} &|\Delta V(t) - \Delta \Pi(t)| \\ &= \sum_{i=0}^n \left(\frac{\partial V(t)}{\partial S_i} - \phi_i \right) \Delta S_i + \frac{1}{2} \sum_{i,j=0}^n \frac{\partial^2 V(t)}{\partial S_i \partial S_j} \underbrace{(\Delta S_i \Delta S_j - \gamma_{i,j} \Delta t)}_{\neq 0} + \text{h.o.t.} \end{aligned}$$

und mit $\phi_i = \frac{\partial V(t)}{\partial S_i}$ für $i \neq 0$

$$= \left(\frac{\partial V(t)}{\partial S_0} - \phi_0 \right) \Delta S_0 + \frac{1}{2} \sum_{i,j=0}^n \frac{\partial^2 V(t)}{\partial S_i \partial S_j} \underbrace{(\Delta S_i \Delta S_j - \gamma_{i,j} \Delta t)}_{\neq 0} + \text{h.o.t.},$$

wobei $\text{h.o.t.} = O(|\Delta t|^2, |\Delta t \Delta S_i|, |\Delta S_i|^3)$.

⁸ Wobei die Gleichung nach ϕ_0 auflösbar ist da wir postulierten, dass S_0 ein Numéraire ist.

7.4.2. Error Propagation

Die Wahl von ϕ_0 bestimmt sich zum Zeitpunkt $t = 0$ aus der Anfangsbedingung $\Pi(0) = V(0)$ und in jedem weiteren Zeitschritt $t = t_k$, $k = 1, 2, \dots$ aus der Bedingung der Selbstfinanzierung der Handelsstrategie

$$\sum_{i=0}^n \phi_i(t_k) S_i(t_k) = \sum_{i=0}^n \phi_i(t_{k-1}) S_i(t_k). \quad (7.14)$$

Dies führt zu einer Fehlerfortpflanzung. Nehmen wir an, dass zum Zeitpunkt t_k aus der vorherigen Wahl $\phi_i(t_{k-1})$ der Portfolios bereits ein Replikationsfehler $V(t_k) - \sum_{i=0}^n \phi_i(t_{k-1}) S_i(t_k) =: e(t_k)$ existiert so ist mit

$$V(t_k) = \sum_{i=0}^n \frac{\partial V(t_k)}{\partial S_i} S_i(t_k), \quad \Pi(t_k) = \sum_{i=0}^n \phi_i(t_k) S_i(t_k) \quad (7.15)$$

sowie $\phi_i(t_k) = \frac{\partial V(t_k)}{\partial S_i}$ für $i \neq 0$

$$e(t_k) = V(t_k) - \sum_{i=0}^n \phi_i(t_{k-1}) S_i(t_k) \stackrel{(7.14)}{=} V(t_k) - \Pi(t_k) \stackrel{(7.15)}{=} \left(\frac{\partial V(t_k)}{\partial S_0} - \phi_0(t_k) \right) \cdot S_0.$$

Somit impliziert die Forderung nach einem selbstfinanzierten Portfolio

$$\phi_0(t) = \frac{\partial V(t)}{\partial S_0} - \frac{1}{S_0} (V(t) - \Pi(t)).$$

Der Replikationsfehler zum Zeitpunkt t_k wird also über die Bedingung der Selbstfinanziertheit durch ϕ_0 über den Zeitschritt Δt weiter gegeben.

Insgesamt folgt für den Delta Hedge

$$\begin{aligned} & |\Delta V(t) - \Delta \Pi(t)| \\ &= (V(t) - \Pi(t)) \cdot \frac{\Delta S_0}{S_0} + \frac{1}{2} \sum_{i,j=0}^n \frac{\partial^2 V(t)}{\partial S_i \partial S_j} \underbrace{(\Delta S_i \Delta S_j - \gamma_{i,j} \Delta t)}_{\neq 0} + \text{h.o.t.}, \end{aligned}$$

wobei $\text{h.o.t.} = O(|\Delta t|^2, |\Delta t \Delta S_i|, |\Delta S_i|^3)$.

7.4.2.1. Beispiel: Zeitdiskreter Delta-Hedge unter dem Black-Scholes Modell

Wir betrachten ein Black-Scholes Modell mit den obigen Bezeichnungen. Für das Replikationsportfolio $\phi_0(t)N(t) + \phi_1(t)S(t)$ folgt $\phi_1(t) = \frac{\partial V(t)}{\partial S(t)} = \Phi\left(\frac{1}{\sigma\sqrt{T-t}} \left[\log\left(\frac{S(t)}{K}\right) + r(T-t) + \frac{\sigma^2}{2} \right]\right)$.

Wir benutzen diese Wahl um an diskreten Zeitpunkten T_i , $i = 0, 1, \dots$ in einem Replikationsportfolio zu handeln. Die Größe der „Cash-Position“ wird so gewählt, dass das Portfolio selbstfinanzierend bleibt, d.h.

$$\phi_1(T_i) = \frac{\partial V(T_i)}{\partial S(T_i)} = N\left(\frac{1}{\sigma\sqrt{T-T_i}}\left[\log\left(\frac{S(T_i)}{K}\right) + r(T-T_i) + \frac{\sigma^2}{2}\right]\right)$$

$$\phi_0(T_i) = \frac{1}{N(T_i)}(\phi(T_{i-1})N(T_i) + \phi_1(T_{i-1})S(T_i) - \phi_1(T_i)S(T_i))$$

Zum Zeitpunkt $T_0 = 0$ wird das Portfolio entsprechend dem aus der Bewertungstheorie bekannten Wert der Option mit Kapital ausgestattet.⁹

$$\phi_0(T_0) = \frac{1}{N(T_0)}(V(T_0) - \phi_1(T_0)S(T_0))$$

Abbildung 7.1 zeigt das Ergebnis eines wöchentlich vorgenommenen Delta-Hedges.

7.4.3. Delta-Gamma Hedging



Motivation (Nochmals der Delta Hedge): Ein in diskreten Zeitabständen $\Delta T_j = T_{j+1} - T_j$ durchgeführter Delta-Hedge ist nicht exakt, da der Optionspreis $V(t)$ im Allgemeinen nicht-linear von der Änderung im Underlying $\Delta S_i = S_i(T_{j+1}) - S_i(T_j)$ abhängt, das Replikationsportfolio jedoch linear im Underlying ist.

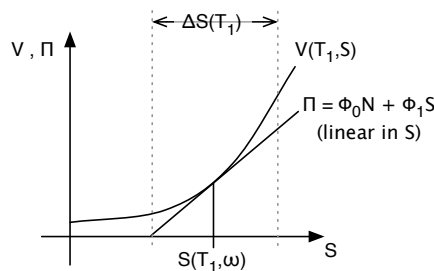


Figure 7.2.: Delta Hedge

$S(T_j, \omega)$ können also Funktionswert und erste Ableitung nach S (das Delta) gleich gewählt werden: Bei einer infinitesimalen Änderung dS stimmen die infinitesimalen Änderungen von Derivat (dS) und Replikationsportfolio $d\Pi$ überein. Ist die Änderung des Produktes S in einem grösseren (nicht-infinitesimalen) Bereich $\Delta S(T_j)$, so stimmen Portfolio und Derivat dort mitunter schlecht überein.¹⁰ <|

⁹ Eventuell ist das Anfangskapital des Replikationsportfolios jedoch nicht ausreichend, siehe Abbildung 7.1, rechts.

¹⁰ Offensichtlich stimmen Portfolio und Derivat nur dann *global* überein, wenn V selbst linear in S ist. Die bedeutet, dass Derivate die lineare Funktionen der Underlyings sind statisch durch ein Portfolio repliziert werden können.

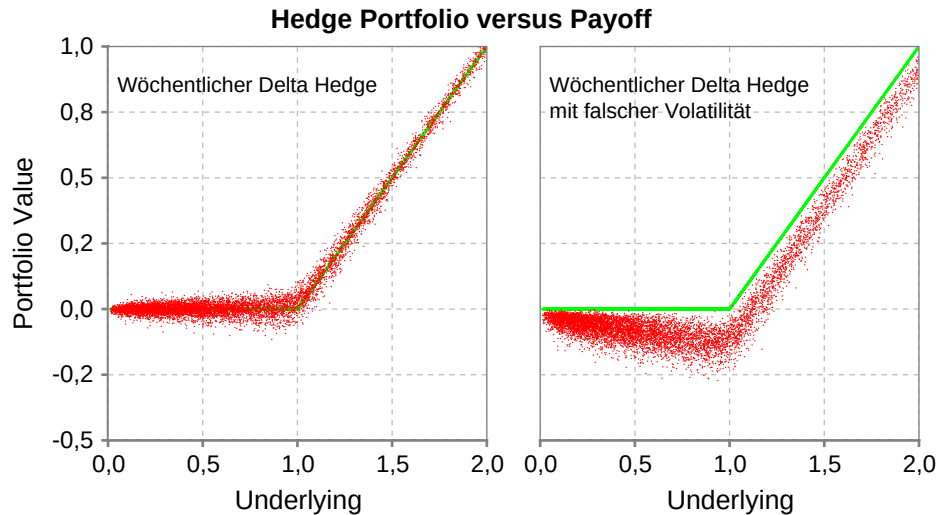


Figure 7.1.: Streuung des Wertes des Replikationsportfolios bei wöchentlichem Rehedging. Der Markt (N, S) folgt einem Black-Scholes Modell $dN(t) = rN(t)dt$, $dS(t) = \mu S(t)dt + \sigma S(t)dW(t)$ mit $(N(0), S(0)) = (1.0, 1.0)$, $r = \mu = 0.05$ und $\sigma(t) = 0.5$. Ziel ist es den Payout $X(T) = \max(S(T) - K, 0)$ mit $K = 1.0$ und $T = 6.0$ (grün) zu replizieren. Das Replikationsportfolio $\Pi = \phi_0 N + \phi_1 S$ wird zu diskreten Zeiten $T_i = i \cdot \frac{1}{52}$ entsprechend einem Delta-Hedge für das entsprechende Black-Scholes Modell gewählt. Abgetragen ist der erzielte Wert $\Pi(T)$ (rot) über $S(T)$. In der linken Abbildung wird der Delta Hedge mit den korrekten Modellparametern vorgenommen. In der rechten Abbildung wurde bei der Konstruktion des Replikationsportfolios eine Volatilität von 0.4 (statt 0.5) angenommen. Dies hat zur Folge, dass der Optionspreis unterschätzt wird. Daher wird das Replikationsportfolio mit zu wenig Kapital ausgestattet, so dass zur Maturität $T = 6.0$ der Mittelwert des Replikationsportfolios systematisch unterhalb dem Wert des zu replizierenden Payouts liegt. Zusätzlich erhöht sich die Streuung des Replikationsportfolios, da der angesetzte Delta-Hedge nicht korrekt ist.

Der durch die nur lineare Approximation des Derivates durch das Replikationsportfolio gemachte Fehler lässt sich Reduzieren, in dem Produkte in das Replikationsportfolio aufgenommen werden, welche selbst nicht-linear im Underlying sind. Falls denn solche Produkte am Markt gehandelt werden und so zur Replikation benutzt werden können. In der Praxis ist dies durchaus möglich: Einige Optionen werden in ausreichenden Mengen gehandelt. Diese können dazu genutzt werden derivative Produkte, welche nicht oder in zu geringen Mengen gehandelt werden, zu replizieren.¹¹

Wir betrachten ein Portfolio bestehend aus $S_0, S_1, \dots, S_n$ und zusätzlichen (derivativen) Produkten $C_1, \dots, C_m$

$$\Pi(t) = \phi_0(t)S_0(t) + \dots + \phi_n(t)S_n(t) + \psi_1(t)C_1(t) + \dots + \psi_m(t)C_m(t)$$

Unter der Annahme, dass sich wie schon $V(T)$ auch $C_k(t)$ als Funktion $C_k(t, S_0, \dots, S_n)$ der S_i darstellen lässt, folgt für das Replikationsportfolio $\Pi(t)$ statt (7.12) nun

$$\begin{aligned} \Delta\Pi(t) &= \sum_{i=0}^n \phi_i \Delta S_i + \sum_{k=1}^m \psi_k \Delta C_k \\ &= \sum_{i=0}^n \left(\phi_i + \sum_{k=1}^m \psi_k \frac{\partial C_k}{\partial S_i} \right) \Delta S_i + \sum_{k=1}^m \psi_k \frac{\partial C_k}{\partial t} \Delta t \\ &\quad + \sum_{i,j=1}^n \sum_{k=1}^m \psi_k \frac{\partial^2 C_k}{\partial S_i \partial S_j} \Delta S_i \Delta S_j + \text{h.o.t.} \\ &= \sum_{i=0}^n \left(\phi_i + \sum_{k=1}^m \psi_k \frac{\partial C_k}{\partial S_i} \right) \Delta S_i + \sum_{i,j=1}^n \sum_{k=1}^m \psi_k \frac{\partial^2 C_k}{\partial S_i \partial S_j} (\Delta S_i \Delta S_j - \gamma_{i,j} \Delta t) + \text{h.o.t.} \end{aligned}$$

wobei wieder $\text{h.o.t.} = O(|\Delta t|^2, |\Delta t \Delta S_i|, |\Delta S_i|^3)$ ist. Vergleichen wir nun noch mal mit der Entwicklung für V aus (7.11):

$$\Delta V(t) = \sum_{i=0}^n \frac{\partial V(t)}{\partial S_i} \Delta S_i + \frac{1}{2} \sum_{i,j=0}^n \frac{\partial^2 V(t)}{\partial S_i \partial S_j} \underbrace{(\Delta S_i \Delta S_j - \gamma_{i,j} \Delta t)}_{\neq 0} + \text{h.o.t.}$$

Löst der Portfolioprozess $(\phi_0, \dots, \phi_n, \psi_1, \dots, \psi_m)$ sodann die Gleichungen

$$\sum_{k=1}^m \frac{\partial^2 C_k}{\partial S_i \partial S_j} \cdot \psi_k = \frac{\partial^2 V}{\partial S_i \partial S_j} \quad i \leq j = 1, \dots, n \quad (7.16)$$

$$\phi_i + \sum_{k=1}^m \frac{\partial C_k}{\partial S_i} \cdot \psi_k = \frac{\partial V}{\partial S_i} \quad i = 1, \dots, n, \quad (7.17)$$

¹¹ Dies ist insbesondere sinnvoll für Optionen mit unterschiedlicher Laufzeit.

so bleibt ein Restrisiko $|\Delta V(t) - \Delta \Pi(t)| = O(|\Delta t|^2, |\Delta t \Delta S_i|, |\Delta S_i|^3)$. Um das Gamma eines Derivates V zu neutralisieren bedarf es natürlich nur höchstens so viele Hedge-Derivate C_k wie partielle Ableitungen $\frac{\partial^2 V}{\partial S_i \partial S_j}$ $i \leq j = 0, \dots, n$ (Gamma) ungleich 0 sind. Es sind daher maximal $m = n \cdot (n - 1)/2$ zusätzlichen Derivate nötig.

7.4.3.1. Beispiel: Zeiddiskreter Delta-Gamma-Hedge unter dem Black-Scholes Modell

Wir betrachten ein Black-Scholes Modell mit den obigen Bezeichnungen. Es sei C_1 ein Option mit Fälligkeit T^* und Auszahlungsprofil $\max(S(T^*) - K^*, 0)$. Zu replizieren sei eine Option V mit Fälligkeit $T < T^*$ und Auszahlungsprofil $\max(S(T) - K, 0)$. Gehandelt werde an diskreten Zeitpunkten $0 = t_0 < t_1 < \dots$

Für die zu replizierende Option V ist

$$\Delta V(t_k) = \frac{\partial V(t_k)}{\partial t} \Delta t_k + \frac{\partial V(t_k)}{\partial S} \Delta S(t_k) + \frac{1}{2} \frac{\partial^2 V(t_k)}{\partial S^2} (\Delta S(t_k))^2 + O(|\Delta t_k|^2, |\Delta t_k \Delta S(t_k)|, |\Delta S(t_k)|^3).$$

Für das Replikationsportfolio Π ist

$$\begin{aligned} \Delta \Pi(t_k) &= \phi_0(t) \Delta N(t_k) + \phi_1(t_k) \Delta S(t) + \psi_1(t_k) \Delta C(t_k) \\ &= \phi_0(t) \Delta N(t_k) + \phi_1(t_k) \Delta S(t) + \psi_1(t_k) \frac{\partial C(t_k)}{\partial t} \Delta t_k \\ &\quad + \psi_1(t_k) \frac{\partial C(t_k)}{\partial S} \Delta S(t_k) + \psi_1(t_k) \frac{1}{2} \frac{\partial^2 C(t_k)}{\partial S^2} (\Delta S(t_k))^2 \\ &\quad + O(|\Delta t_k|^2, |\Delta t_k \Delta S(t_k)|, |\Delta S(t_k)|^3) \end{aligned}$$

Für das Replikationsportfolio $\Pi(t) = \phi_0(t)N(t) + \phi_1(t)S(t) + \psi_1(t)C(t)$ folgt $\phi_1(t) = \frac{\partial V(t)}{\partial S(t)} = \Phi\left(\frac{1}{\sigma \sqrt{T-t}} \left[\log\left(\frac{S(t)}{K}\right) + r(T-t) + \frac{\sigma^2}{2} \right]\right)$.

Wir benutzen diese Wahl um an diskreten Zeitpunkten t_k in einem Replikationsportfolio zu handeln. Die Größe der „Cash-Position“ wird so gewählt, dass das Portfolio selbstfinanzierend bleibt, d.h.

$$\begin{aligned} \phi_1(t_k) &= \frac{\partial V(t_1)}{\partial S(t_1)} = N\left(\frac{1}{\sigma \sqrt{T-t_k}} \left[\log\left(\frac{S(t_k)}{K}\right) + r(T-t_k) + \frac{\sigma^2}{2} \right]\right) \\ \phi_0(t_k) &= \frac{1}{N(t_k)} (\phi(t_{k-1})N(t_k) + \phi_1(t_{k-1})S(t_k) - \phi_1(t_k)S(t_k)) \end{aligned}$$

Zum Zeitpunkt $t_0 = 0$ wird das Portfolio entsprechend dem aus der Bewertungstheorie bekannten Wert der Option mit Kapital ausgestattet.¹²

$$\phi_0(t_0) = \frac{1}{N(t_0)} (V(t_0) - \phi_1(t_0)S(t_0))$$

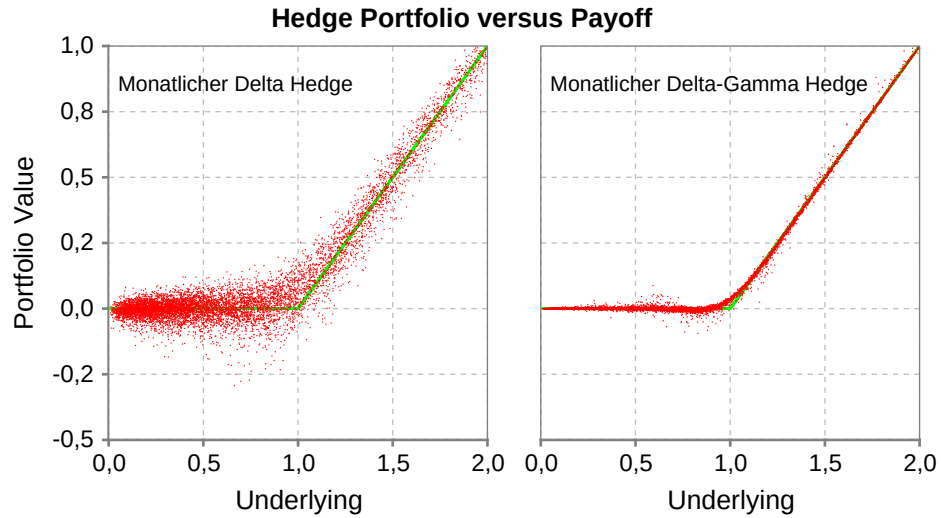


Figure 7.3.: Streuung des Wertes des Replikationsportfolios bei monatlichem Rehedging: Delta versus Delta-Gamma Hedge.



Interpretation: Unser Delta- und Delta-Gamma-Hedge könnte unvollständig anmuten: Der Delta-Hedge neutralisiert den Fehler erster Ordnung (in S), der Delta-Gamma-Hedge den Fehler erster und zweiter Ordnung. Der Restfehler des Delta-Gamma Hedge war von der Ordnung

$$O(|\Delta t|^2, |\Delta t \Delta S_i|, |\Delta S_i|^3).$$

Zwei Fragen stellen sich:

- Warum bleibt in unserer Betrachtung der Fehler in Δt unberücksichtigt?
- Sollten das Replikationsportfolio nicht entsprechend $\phi_i(t) = \frac{\partial V(t+\Delta t)}{\partial S_i}$ statt entsprechend $\phi_i(t) = \frac{\partial V(t)}{\partial S_i}$ gewählt werden? Schliesslich soll ja die Realisierung von V an $t + \Delta t$ möglichst gut getroffen werden soll?

Die Antwort gibt die Interpretation der betrachteten Hedge Strategien: Sie dienen dazu das Risiko in den Underlyings S_i zu reduzieren, also die Abhängigkeit von $V(t) - \Pi(t)$ von $S_i(t)$. Dies kann so gedeutet werden, dass von vornherein nicht klar ist, wie lange diese Replikation so beibehalten werden soll (d.h. Δt selbst ist nicht bekannt). In

¹² Eventuell ist das Anfangskapital des Replikationsportfolios jedoch nicht ausreichend, siehe Abbildung 7.1.

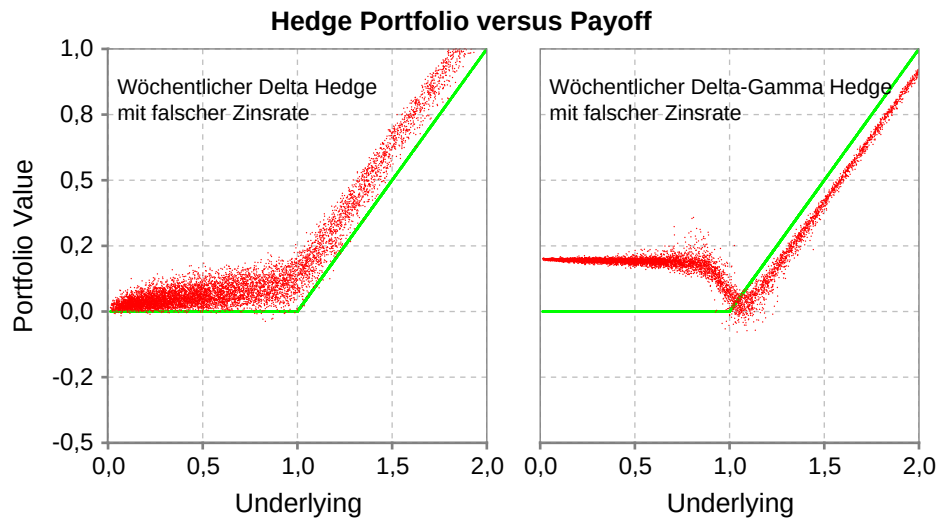


Figure 7.4.: Streuung des Wertes des Replikationsportfolios bei wöchentlichem Rehedging mit falscher Zinsrate.

unserer Betrachtung können wir daher $\Delta t \rightarrow 0$ annehmen und es bleibt ein Fehler $O(S_i^2)$ für den Delta-Hedge und $O(S_i^3)$ für den Delta-Gamma Hedge.

Selbstverständlich ist diese Wahl kein statischer Hedge und es existiert ein Fehler der Ordnung Δt . Dies wird deutlich betrachtet man beide Verfahren für sehr grosse Zeitschritte Δt , oder gar den Extremfall eines statischen Hedges, d.h. das Portfolio wird einmal bestimmt und nicht mehr geändert, siehe Abbildung 7.5. <|

7.4.4. Vega Hedging

Analog zum Delta-Gamma Hedging, in dem neben einem *Hedge* für das Options-Delta auch ein *Hedge* für das Options-Gamma berücksichtigt wurde, lassen sich weitere funktionale Abhängigkeiten der Preisfunktion berücksichtigen. Neutralisiert man mit dem Replikationsportfolio die funktionale Abhängigkeit von der (Log-)Prozess-Volatilität in erster Ordnung, d.h. betrachtet man $\frac{\partial V}{\partial \sigma_i}$ für $dS_i = \mu_i S_i dt + \sigma_i S_i dW_i$, so wird dies als Vega Hedging bezeichnet.

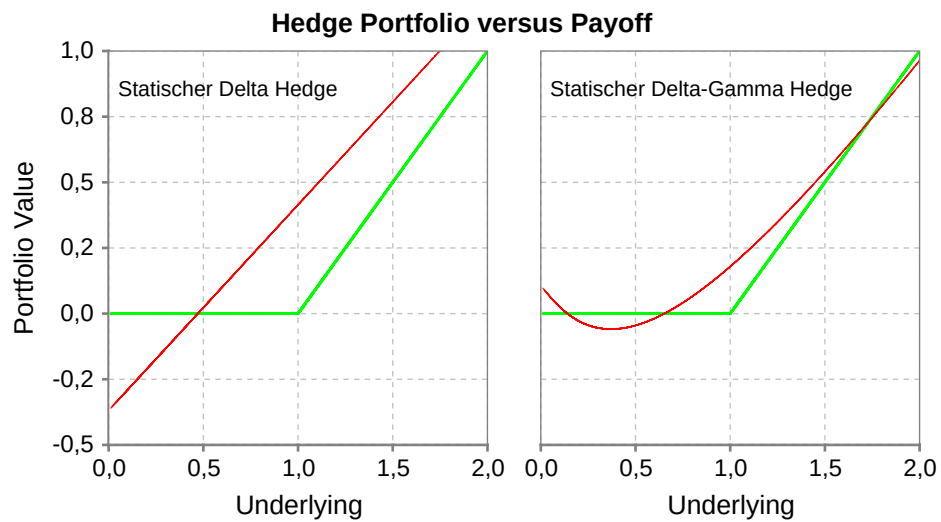


Figure 7.5.: Wert des Replikationsportfolios ohne Re-Hedging. Das Replikationsportfolio wird in $t = 0.0$ entsprechende dem Delta- bzw. Delta-Gamma Hedge gewählt und bis zur Optionsfälligkeit T nicht mehr verändert (Modell und Optionsparameter entsprechen denen in Abbildung 7.1).

7.5. Hedging in discrete Time: Minimizing the Residual Error (Bouchaud-Sornette Method)

Der Delta-Hedge überträgt die optimale Handelstrategie für kontinuierliches Handeln auf den zeit-diskreten Fall, für die die Strategie jedoch nicht notwendig optimal sein muss.¹³ Eine adäquatere Bestimmung einer risikominimierenden Handelsstrategie liefert hingegen die direkte Betrachtung des Restrisikos: Gesucht ist jene Handelsstrategie, die das Restrisiko minimiert. Dies ist der Ausgangspunkt der Methode von Bouchaud-Sornette [47]. Das mit diesem Kriterium bestimmte optimale Replikationsportfolio ist nicht notwendig identisch mit dem Replikationsportfolio des Delta-Hedge und hängt von der Norm ab, in der Restrisiko gemessen wird.

Zum Zeitpunkt T soll ein Auszahlungsprofil $V(T)$ durch ein Replikationsportfolio $t \mapsto \Pi(t)|_{t=T}$ möglichst gut repliziert werden. Dazu betrachten wir das Mittel der quadratischen Abweichung, d.h. die Varianz von $V(T) - \Pi(T)$:

$$\text{Var}(V(T) - \Pi(T)).$$

Wir möchten diese Varianz in einem bestimmten Sinne minimieren, und schieben dazu eine Definition und ein Lemma ein:

Definition 83 (Bedingte Varianz): ▮

Es seien $C \subset \mathcal{F}$ zwei σ -Algebren und X eine $\mathcal{F}$ -messbare Zufallsvariablen über dem Wahrscheinlichkeitsraum $(\Omega, \mathbb{P}, \mathcal{F})$. Dann heißt

$$\text{Var}(X|C) := E(X^2|C) - (E(X|C))^2$$

die *bedingte Varianz* von X unter C . ▮

Remark 84 (Bedingte Varianz): Die bedingte Varianz ist eine C -messbare Zufallsvariable.

Lemma 85 (Bedingte Varianz): Es seien $C \subset \mathcal{F}$ zwei σ -Algebren, X eine $\mathcal{F}$ -messbare Zufallsvariable und Y eine $\mathcal{F}$ -messbare Zufallsvariable über dem Wahrscheinlichkeitsraum $(\Omega, \mathbb{P}, \mathcal{F})$. Dann ist

$$\text{Var}(X + Y|C) = \text{Var}(X|C)$$

¹³ Da im zeit-diskreten Fall das Replikationsportfolio keine perfekte Replikation leistet, ist insbesondere zu klären, in welchem Sinne ein Hedge als optimal gesehen wird, d.h. in welcher Norm der Hedgefehler betrachtet wird.

Proof: Der Beweis ist elementar. Mit $E(X \cdot Y|C) = Y \cdot E(X|C)$ und $E(Y^2|C) = Y^2$ folgt

$$\begin{aligned} \text{Var}(X + Y|C) &\stackrel{\text{Def.}}{=} E((X + Y)^2|C) - (E(X + Y|C))^2 \\ &= E(X^2|C) + 2 \cdot E(X \cdot Y|C) + E(Y^2|C) \\ &\quad - E(X|C)^2 - 2 \cdot E(X|C) \cdot E(Y|C) - E(Y|C)^2 \\ &= E(X^2|C) + \cancel{2 \cdot Y \cdot E(X|C)} + \cancel{Y^2} - E(X|C)^2 - \cancel{2 \cdot Y \cdot E(X|C)} - \cancel{Y^2} \\ &= E(X^2|C) - E(X|C)^2 = \text{Var}(X|C) \end{aligned}$$

□

Es sei $0 = T_0 < T_1 < \dots < T_m = T$ eine Zeitdiskretisierung und $\phi(t) = (\phi_0, \dots, \phi_n)$ der **Portfolioprozess** eines Replikationsportfolios $\Pi(t) = \phi(t) \cdot S(t) = \sum_{i=0}^n \phi_i(t) S_i(t)$, mit stückweise konstanten ϕ_i , d.h. es werde in diskreter Zeit gehandelt ($\phi_i(t) = \phi_i(T_j)$ für $T_j \leq t < T_{j+1}$).

7.5.1. Minimizing the Residual Error at Maturity T

Handelsstrategie: Zu jedem Zeitpunkt T_k sei das selbstfinanzierende Portfolio $\phi(T_k)$ so gewählt, dass das Restrisiko minimal ist, d.h.

$$\text{Var}(V(T) - \Pi(T) | \mathcal{F}_{T_k}) \rightarrow \min,$$

wobei $\{\mathcal{F}_t\}$ die von S erzeugte Filtration sei.

Zur Vereinfachung der Darstellung sei $S_0 := P(T)$ der **Bond mit Maturity T** , d.h. ein Finanzprodukt welches zur Zeit T eine garantierte Auszahlung von 1 tätigt, $P(T) : t \mapsto P(T; t)$ mit $P(T; T) \equiv 1$. Dieses Finanzprodukt bildet später die Grundlage der Zinstheorie, siehe Definition 86 in Abschnitt 8. Es ist nicht notwendig, dass eines der Finanzprodukte der **Bond mit Maturity T** ist, aber es vereinfacht die folgende Darstellung, da

$$\text{Var}(V(T) - \Pi(T)) = \text{Var}\left(\frac{V(T)}{S_0(T)} - \frac{\Pi(T)}{S_0(T)}\right)$$

und statt

$$\Pi(T) = \Pi(0) + \sum_{l=0}^{m-1} \phi(T_l) \cdot \Delta S(T_l)$$

ebenso

$$\begin{aligned} \Pi(T) = \frac{\Pi(T)}{S_0(T)} &= \frac{\Pi(0)}{S_0(0)} + \sum_{l=0}^{m-1} \sum_{i=0}^{m-1} \phi_i(T_l) \cdot \Delta \frac{S_i(T_l)}{S_0(T_l)} \\ &= \frac{\Pi(0)}{S_0(0)} + \sum_{l=0}^{m-1} \sum_{i=1}^{m-1} \phi_i(T_l) \cdot \Delta \frac{S_i(T_l)}{S_0(T_l)} \end{aligned}$$

gilt – man beachte, dass ϕ_0 in der letzten Summe nicht mehr vorkommt. Somit lässt sich mit ϕ_0 die Bedingung der Selbstfinanziertheit des Portfolios erfüllen und ϕ_i , $i \geq 1$ bestimmen den „optimalen Hedge“.

Mit der Bezeichnung $\frac{\tilde{S}}{S_0}$ ist

$$\begin{aligned} V(T) - \Pi(T) &= V(T) - \Pi(0) - \sum_{l=0}^{m-1} \phi(T_l) \cdot \Delta \tilde{S}(T_l) \\ &= V(T) - \Pi(0) - \sum_{l=0}^{k-1} \phi(T_l) \cdot \Delta \tilde{S}(T_l) \\ &\quad - \sum_{l=k+1}^{m-1} \phi(T_l) \cdot \Delta \tilde{S}(T_l) - \phi(T_k) \cdot \Delta \tilde{S}(T_k) \end{aligned}$$

und mit Lemma 85

$$\begin{aligned} &\text{Var}(V(T) - \Pi(T) | \mathcal{F}_{T_k}) \\ &= \text{Var}(V(T) - \sum_{l=k+1}^{m-1} \phi(T_l) \cdot \Delta \tilde{S}(T_l) | \mathcal{F}_{T_k}) - \phi(T_k) \cdot \Delta \tilde{S}(T_k) \\ &= \text{Var}(V(T) - \sum_{l=k+1}^{m-1} \phi(T_l) \cdot \Delta \tilde{S}(T_l) | \mathcal{F}_{T_k} - \phi(T_k) \cdot \Delta \tilde{S}(T_k) | \mathcal{F}_{T_k}) \\ &= \text{Var}(V(T) - \sum_{l=k+1}^{m-1} \phi(T_l) \cdot \Delta \tilde{S}(T_l) | \mathcal{F}_{T_k} | \mathcal{F}_{T_k}) \\ &\quad - 2 \cdot \text{Cov}(V(T) - \sum_{l=k+1}^{m-1} \phi(T_l) \cdot \Delta \tilde{S}(T_l), \phi(T_k) \cdot \Delta \tilde{S}(T_k) | \mathcal{F}_{T_k}) \\ &\quad + \text{Var}(\phi(T_k) \cdot \Delta \tilde{S}(T_k) | \mathcal{F}_{T_k}) \end{aligned}$$

Nun folgt aus

$$\frac{1}{2} \frac{\partial}{\partial \phi(T_k)} \text{Var}(V(T) - \Pi(T) | \mathcal{F}_{T_k}) = 0$$

die Bestimmungsgleichung für $\phi(T_k)$

$$\begin{aligned} &\phi(T_k) \cdot \text{Cov}(\Delta \tilde{S}(T_k), \Delta \tilde{S}(T_k) | \mathcal{F}_{T_k}) \cdot \phi(T_k)^T \\ &\quad - \text{Cov}(V(T) - \sum_{l=k+1}^{m-1} \phi(T_l) \cdot \Delta \tilde{S}(T_l), \Delta \tilde{S}(T_k) | \mathcal{F}_{T_k}) = 0 \end{aligned} \quad (7.18)$$

7.5.2. Minimizing the Residual Error in each Time Step

Die im vorherigen Abschnitt 7.5.1 beschriebene Strategie fokussierte sich zu jedem Zeitpunkt T_k auf die Minimierung des Restrisikos zum Zeitpunkt T . Eine alternative

Handelsstrategie entsteht aus der Forderung im Zeitschritt T_k das Residualrisiko zur Zeit T_{k+1} durch Wahl des Portfolios $\phi(T_k)$ zu minimieren.

$$\text{Var}(V(T_{k+1}) - \Pi(T_{k+1}) | \mathcal{F}_{T_k}) \rightarrow \min.$$

Offensichtlich (betrachte den letzten Zeitschritt $k = m - 1$ in (7.18)) folgt dazu die Bestimmungsgleichung $\phi(T_k)$ als

$$\phi(T_k) \cdot \text{Cov}(\Delta\tilde{S}(T_k), \Delta\tilde{S}(T_k) | \mathcal{F}_{T_k}) \cdot \phi(T_k)^T - \text{Cov}(V(T_{k+1}), \Delta\tilde{S}(T_k) | \mathcal{F}_{T_k}) = 0$$



Experiment: Unter <http://www.christian-fries.de/finmath/applets/HedgeSimulator.html> findet sich ein *Hedge-Simulator*. Dabei werden in einer Monte-Carlo Simulation Pfade $t \mapsto (N(t, \omega), S(t, \omega))$ entsprechend einem vorgegeben Modell erzeugt. Entlang jedes Pfades wird sodann entsprechend einer vorgegebenen Handelsstrategie in einem Replikationsportfolio gehandelt. Zum Optionszeitpunkt T wird der Wert des Replikationsportfolio mit dem zu erzielenden Optionswert $C(T)$ verglichen. Es ist möglich die Modellparameter, unter deren Annahme das Replikationsportfolio konstruiert wird, und die Modellparameter, die die Evolution des Marktes (N, S) bestimmen, unabhängig zu wählen. Eine fehlende Übereinstimmung der Parameter trägt zu einem höheren Restrisiko bei (siehe Abbildung 7.1). <|

Part III.

Interest Rate Structures, Interest Rate Products and Analytic Pricing Formulas

Motivation and Overview - Part III

Part III will consider the interest rate structures and the (analytical) pricing of interest rate derivatives. The methods for pricing financial products may roughly be classified into three groups. These are

- **Model independent analytic pricing:** The financial product may be decomposed into a portfolio of traded assets. The value of the derivative product is then given by the value of the portfolio. The portfolio represents a replication portfolio, i.e. a static hedge.
- **Model dependent analytic pricing:** The replication of the financial product requires dynamic hedging, i.e. a continuous adjustment of the portfolio, and thus requires a model for the underlying stochastics. The value of the replication portfolio depends on the model. The complexity of the product and of the model, however, allows to derive an analytic pricing formula.
- **Model dependent numeric pricing:** The replication of the financial product requires dynamic hedging and thus requires a model for the underlying stochastics. The value of the replication portfolio depends on the model. The complexity of the product or of the model requires a numerical calculation of the price.

In Chapter 8 we will define some elementary products and interest rates. In Section 9.1 simple interest rate products will be presented that allow for a model independent analytic pricing. Here, the corresponding pricing is discussed right after the definition of the product. In Section 9.2 we define some simple options. In Chapter 10 and 11 we will show how to derive some analytic pricing formulas using some simple standard models. In Chapter 12 more exotic options will be presented, for which a pricing requires numerical methods in general.

Numerical pricing methods and complex pricing models will be considered in Part IV and V.

CHAPTER 8

Interest Rate Structures

8.1. Introduction

In previous sections we have considered a one dimensional stochastic process as the underlying, representing a stock for example. We will now turn to the modeling of interest rates and later the pricing of interest rate derivatives. Interest rates potentially offer a richer structure than a single stock, since at each time t we have to consider an interest rate *curve* $F(T; t)$, $T \geq t$ instead of a scalar stock price $S(t)$.

Let $F(T; t)$ denote the interest earned on an investment made over the period $[t, T]$ if the contract is written at the beginning t of the period, i.e. we assume that $N \cdot (1 + F(T; t) \cdot (T - t))$ is the amount one receives at time T if one invests the amount N at time t , and the contract is fixed at the beginning of the period, i.e. in t .¹ For different times $T_1 < T_2$ two such interest rates $F(T_1; t)$ and $F(T_2; t)$ may be different. The reason is a different interest earned on the two sub-periods $[t, T_1]$ and $[T_1, T_2]$, e.g. if it is expected that interest rates will rise or fall in future. Thus, if one decomposes the time axis into small interest rate periods $[T_i, T_{i+1}]$ and defines

$$(1 + F(T_k; t)(T_k - t)) =: (1 + L(T_0, T_1; t)(T_1 - T_0)) \cdot \dots \cdot (1 + L(T_{k-1}, T_k; t)(T_k - T_{k-1})), \quad (8.1)$$

then it is rational to represent each interest rate $L(T_i, T_{i+1}; t)$ by its own stochastic process. Here we assume $T_0 = t$ and $L(T_i, T_{i+1}; t)$ denotes the interest rate earned over the period $[T_i, T_{i+1}]$ as given by a contract which is fixed in t .² See Figure 8.1. It is common to have interest rate derivatives that depend on more than one interest rate and in this sense depend on the movement of the whole curve.

¹ We will comment below on the fact that the contract may be fixed at an earlier time than the start of the interest rate period.

² Equation (8.1) represents the relationship between interest earned over the period $[t, T_k]$ and interest earned of smaller sub periods, including compounding. Since all rates are fixed in time t , the difference is only interpretation. Thus the two sides in (8.1) must agree.

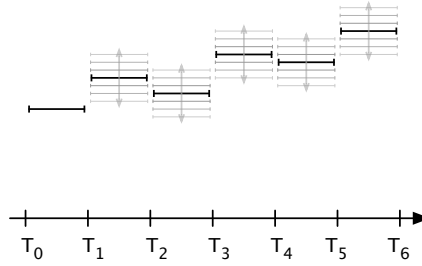


Figure 8.1.: Modeling an interest rate curve through a family of stochastic processes.

8.1.1. Fixing Times and Tenor Times

If the interest rates $F(T_i; t)$ for periods $[t, T_i]$ for $i = 1, \dots, k$ are known, then we may derive from Equation (8.1) the rates for the sub-periods $[T_{i-1}, T_i]$. It is

$$L(T_{i-1}, T_i; t) = \frac{1}{T_i - T_{i-1}} \left(\frac{1 + F(T_i; t)(T_i - t)}{1 + F(T_{i-1}; t)(T_{i-1} - t)} - 1 \right). \quad (8.2)$$

The rate $L(T_{i-1}, T_i; t)$ is the interest rate for the *future* period T_{i-1}, T_i , that has been derived from the interest rate curve as of time t . The interest rate has been fixed at time t , i.e. the stochastic process is evaluated in t . The time t is the simulation time of the stochastic process, it thus determines that the random variable $L(T_{i-1}, T_i; t)$ has to be (at least) $\mathcal{F}_t$ -measurable. The times T_i mark the start and end of the periods. The period structure $\{T_0, T_1, \dots, T_k\}$ is also called the *tenor structure*.

8.2. Definitions

All of the following random variables and stochastic processes are assumed to be defined over the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$. As the building block of all interest rates we define the *Bond*:

Definition 86 (Bond³):

Assume that a guaranteed payment of a unit currency⁴ 1 in time T_2 is a traded product at any earlier time $t < T_2$ and its value in state $\omega \in \Omega$ is uniquely determined by (t, ω) . The value of this product is called the price of the T_2 bond as seen in (t, ω) and will be denoted by $P(T_2; t, \omega)$. It defines a stochastic process on $[0, T_2]$ which we denote by $P(T_2)$:

$$P(T_2) : [0, T_2] \times \Omega \mapsto \mathbb{R}.$$

³ The term defined is the *zero-coupon bond* (there are no intermediate payments (*coupons*) until maturity). We give the trivial extension to a *coupon bond* in Definition 97.

⁴ See Remark 123.

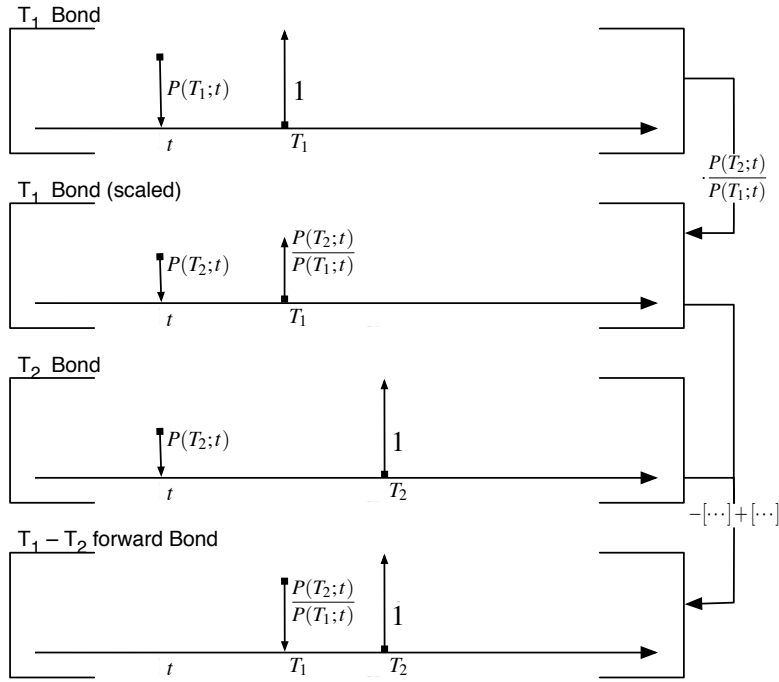


Figure 8.2.: Cashflows for a forward bond.



Interpretation: The value of the bond $P(T_2; t, \omega)$ is the amount we have to invest at a given time t and a given state ω to receive 1 in T_2 . Thus for a (riskless) investment of 1 at time t we get a guaranteed payment of $1/P(T_2; t, \omega)$ at time T_2 , because we simply buy $1/P(T_2; t, \omega)$ times the bond. Note that we assume that all products may be traded in arbitrary fractions and that the price is unique and the same for selling and buying. Thus $P(T_2; t, \omega)$ not only implies the interest earned on an investment, it also implies the interest to be paid for a corresponding credit.

Definition 87 (Forward Bond):

Let $0 \leq T_1 \leq T_2 < \infty$. As *forward bond* we denote the stochastic process

$$P(T_1, T_2) := \frac{P(T_2)}{P(T_1)},$$

defined on $[0, T_1]$.



Interpretation: The *forward bond* $P(T_1, T_2; t)$ corresponds to the amount that has to be invested at time T_1 to receive a guaranteed payment of 1 in T_2 , if that contract is finalized at time t in state ω . To receive 1 in T_2 one has to invest $P(T_2; t)$ in t . This investment is financed by borrowing until T_1 , i.e. we borrow $P(T_2; t) \cdot 1/P(T_1; t)$ -times the value of a $P(T_1)$ bond (compare Figure 8.2).

It is important to understand that $P(T_1, T_2; t)$ is not the value of a forward bond at time t . The value of that contract is simply 0. What we denoted by $P(T_1, T_2; t)$ is the amount that has to be paid in T_1 as part of a contract written in t . The subject of the contract is a bond, which lies *forward* in time.

Obviously we have $P(T_2; t) = P(t, T_2; t)$. ◁|

Often the bond (or forward bond) will be represented by a rate (or forward rate). However, the interest rate that corresponds to a bond depends on how we think of interest earned, e.g. if compounding, i.e. interest paid on interest received, is considered. Thus there are many different definitions of interest rates. For that reason, we see an interest rate as a derived quantity. We define some interest rates:

Definition 88 (Forward Rate ((Forward) LIBOR)): ⌋

Let $T_1 < T_2$. The (forward) LIBOR⁵ $L(T_1, T_2)$ is defined by

$$\frac{P(T_1)}{P(T_2)} = \frac{1}{P(T_1, T_2)} =: (1 + L(T_1, T_2) \cdot (T_2 - T_1)) \quad , \text{ i.e.}$$

$$L(T_1, T_2) := \frac{1}{T_2 - T_1} \cdot \frac{P(T_1) - P(T_2)}{P(T_2)}.$$
⌋

Definition 89 ((continuously compounded) Yield): ⌋

Let $T_1 < T_2$. The (continuously compounded) (forward) yield $r(T_1, T_2)$ is defined by

$$\frac{P(T_1)}{P(T_2)} = \frac{1}{P(T_1, T_2)} =: \exp(r(T_1, T_2) \cdot (T_2 - T_1)) \quad , \text{ i.e.}$$

$$r(T_1, T_2) := \frac{-\log(P(T_1, T_2))}{(T_2 - T_1)} = -\frac{\log(P(T_2)) - \log(P(T_1))}{(T_2 - T_1)}.$$
⌋

Definition 90 (Instantaneous Forward Rate): ⌋

The *instantaneous forward rate* is defined by

$$f(t, T) := \lim_{T_2 \searrow T} L(T, T_2; t).$$

It may be interpreted as the interest rate for the infinitesimal period $[T, T + dT]$. ⌋

⁵ London Inter-Bank Offer Rate. The acronym LIBOR is often used for a rate following the Definition 88, because the inter-bank rate follows this convention.

Remark 91 (Instantaneous Forward Rate): We have

$$\begin{aligned} f(t, T) &= \lim_{T_2 \searrow T} L(T, T_2; t) = \lim_{T_2 \searrow T} \frac{1}{P(T_2; t)} \frac{P(T; t) - P(T_2; t)}{T_2 - T} \\ &= -\frac{1}{P(T; t)} \frac{\partial P(T; t)}{\partial T} = -\frac{\partial \log(P(T; t))}{\partial T}. \end{aligned}$$

Definition 92 (Short rate):

Let $t \geq 0$. The short rate $r(t)$ is defined by

$$\begin{aligned} r(t) &:= \lim_{T \searrow t} r(t, T; t) = \frac{\partial}{\partial T} -\log(P(T; t)) \Big|_{T=t} \\ &= \frac{-\frac{\partial P(T; t)}{\partial T}}{P(T; t)} \Big|_{T=t} = -\frac{\partial P(T; t)}{\partial T} \Big|_{T=t}. \end{aligned}$$

Remark 93 (Short rate): Note that in Definitions 88, 89 and 90 we define families of stochastic processes, while the short rate from Definition 92 is a single, scalar stochastic process.

The *short rate* is a limit of the *spot forward rate* (LIBOR) and the *yield*. We have

$$r(t) = \lim_{T \searrow t} r(t, T; t) = \lim_{T \searrow t} L(t, T; t).$$

This *short rate* is the interest rate for the infinitesimal period $[t, t + dt]$, as seen in t .

Remark 94 (Forward forward): The term „forward“ is used ambiguously: In the Definition 88 and 89 we define stochastic processes for interest rates, i.e. for a given observation time t we define $L(T_1, T_2; t)$ and $r(T_1, T_2; t)$ as rates for the period $[T_1, T_2]$ as seen in t . If $T_1 > t$, then these rates are called “*forward*” (i.e. *forward LIBOR*, *forward Yield*), since the rate is associated with a future period, lying forward in time. If $T_1 = t$ one would instead use the attribute “*spot*”.

However, on the other hand, the rate $L(T_1, T_2; T_1)$ (note $t = T_1$) is often denoted as *forward rate*, since it is an interest rate for an interval up to T_2 . This is in contrast to the *short rate*, which is defined for an infinitesimal period. Being precise, terms like “forward forward rate” should be used.

There is a similar ambiguity for volatilities. The term forward volatility may be interpreted as the volatility of a forward rate or as a volatility of some process considered at a future period in time.

Remark 95 (Interest Rate Models): The various definitions of interest rates (as stochastic processes) are the starting point for the corresponding interest rate modeling. The corresponding models are

Interest Rate	Model
Forward Rate L	LIBOR Market Model → Chapter 17.1
Instantaneous Forward Rate f	HJM Framework → Chapter 19
Short Rate r	Short Rate Models → Chapter 20

Table 8.1.: *Interest Rate Models.*

Although the models above do not model the bond prices directly, we view zero-bond prices as the basic building blocks.

Remark 96 (Bond prices as a function of interest rates): The bond prices may be calculated from the interest rates. We have

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

$$P(T; t) = \exp\left(-\int_t^T f(t, \tau) d\tau\right)$$

$$P(T; t) = \prod_{i=0}^{n-1} (1 + L(T_i, T_{i+1}; t)(T_{i+1} - T_i))^{-1} \quad \text{for } t = T_0 < T_1 < \dots < T_n = T$$

The short rate represents an exception. Here the reconstruction of the bond prices is possible, only if the short rate process is known under the equivalent martingale measure $\mathbb{Q}^N$ corresponding to the Numéraire $N(t) := \exp(\int_0^t r(\tau) d\tau)$. Then we have

$$P(T; t) = E^{\mathbb{Q}^N} \left(\exp\left(-\int_t^T r(\tau) d\tau\right) \mid \mathcal{F}_t \right)$$



Tip (Discount factors as a basic market data object): We consider the price of zero bonds as given and view interest rates as derived quantities. It is natural to take this view also in the implementation. If we want to provide information on market interest rates through a [Class](#)⁶, then we store internally a discretization of the Bond price curve $j \mapsto P(T_j; 0)$ (also called *discount factors*). The class then provides the various interest rates under various conventions through [methods](#). This design reduces the errors of misinterpretation of the stored data (especially if more than one developer works on the class), since the data stored is free of market conventions and the convention used to calculate the rates is explicit in the implementation of the corresponding method. This also reduces documentation overhead for the [data model](#).

⁶ See Chapter 25.

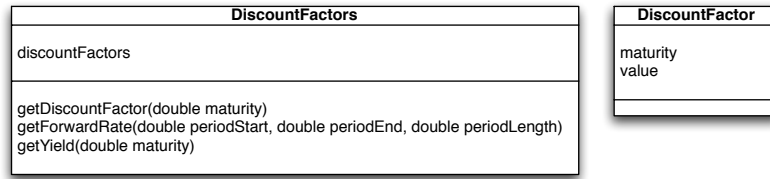


Figure 8.3.: UML diagram: The class *DiscountFactors* internally stores discount factors and provides various interest rates through corresponding methods.

A problem of this design is that discount factors are usually not the quantities that are observable in the interest rate market. The market quotes the prices of various different interest rate products (e.g. futures or [swaps](#)), from which discount factors have to be calculated. This calculation is called [bootstrapping](#). Under some conditions, it might be useful to store the original market data. One such application is the numerical calculation of partial derivatives with respect to a change in these input prices.⁷ <|

8.3. Interest Rate Curve Bootstrapping

Since bootstrapping of discount factors from market prices involves the inversion of a pricing formula, we have to discuss the interest rate products and their pricing. We will do so in Chapter 9, but give an abstract description of the bootstrap algorithm already at this point:

Let $0 = T_0 < T_1 < T_2 < \dots$ denote a time discretization. For given discount factors $df_j = P(T_j; 0)$, $j = 0, 1, \dots, i$ we assume the existence of an interpolation function $df(df_0, \dots, df_i; t)$, having the property that an additional sample point beyond T_i will not change the interpolation in $t \leq T_i$, i.e.

$$df(df_0, \dots, df_i; t) = df(df_0, \dots, df_i, df_{i+1}; t) \quad \forall t \leq T_i \quad \forall i = 1, 2, 3, \dots$$

Let V_i^{market} denote given market prices of interest rate products for which the price may be expressed as a function V_i of the discount factors in $t \leq T_i$, i.e.

$$V_i = V_i(\{df(t) \mid t \leq T_i\}).$$

We further assume that the discount factor $df(T_i)$ enters into the pricing, i.e. let $\frac{\partial}{\partial df(T_i)} V_i \neq 0$. Then the bootstrap algorithm is given by:

Induction start (T_0):

- $df_0 := df(T_0) = P(T_0; 0) = P(0; 0) = 1.0$;

⁷ The importance of the partial derivative with respect to the price of an *underlying* has been discussed in Chapter 7.

Induction step ($T_{i-1} \rightarrow T_i$):

- Calculate $df_i := df(T_i)$ such that, using the discount factor interpolation, we have

$$V_i(\{df(df_0, \dots, df_i; t) \mid t \leq T_i\}) \stackrel{!}{=} V_i^{\text{market}}. \quad (8.3)$$

In some cases the Equation (8.3) may be directly solved for df_i , especially if it does not depend on the interpolation method used. In general, a numerical solution is possible (see Appendix A.6).⁸

8.4. Interpolation of Interest Rate Curves

We consider, as before, a family of bond prices $T \mapsto P(T; 0)$, i.e. the discount factor curve, as the basic representation of the interest rate curve. If the prices $df_i := P(T_i; 0)$ are known for times $0 = T_0 < T_1 < T_2 < \dots$, we seek a meaningful interpolation method to calculate interest rates for sub periods. The interpolation method should fulfill two basic requirements:

- The interpolation method should be sufficiently smooth, at least continuously differentiable. This is desirable, because the calculation of an interest rate corresponds to a finite difference, i.e. converges to a derivative for decreasing period length.
- The interpolation method should preserve the monotonicity of discount factors, i.e. if we have monotone decreasing sample points then the interpolation should be a monotone decreasing curve.

The following additional requirement is also desirable:

- If the sample points correspond to a set of constant rates, then their interpolation should give constant rates. In other words: the interpolation of sample points from a flat interest rate curve should be flat (where flat means flat with respect to a rate).

The linear interpolation of bond prices fulfills the second, but neither the first nor the third requirement. The linear interpolation of forward rates fulfills the first and third requirement, but not necessarily the second. A simple interpolation method, which is also popular in practice, is the linear interpolation of the logarithm of the discount factors, i.e. the linear interpolation of $r(0, T_j) \cdot T_j$:

$$df(t) = \exp\left(\frac{T_{j+1} - t}{T_{j+1} - T_j} \cdot \log(df(T_j)) + \frac{t - T_j}{T_{j+1} - T_j} \cdot \log(df(T_{j+1}))\right).$$

⁸ If interest rates are positive, a simple interval bisection, like the [Golden Section Search](#), works since $df_i \in [0, df_{i-1}]$.

This interpolation fulfills the second and the third requirement.

A more complete discussion of various interpolation methods for interest rate curves may be found in [67].

8.5. Implementation

We extend the design of the `DiscountFactors` class of Figure 8.3 by an interpolation algorithm and an bootstrap algorithm.

If the interpolation method is realized as part of the `getDiscountFactor()` method, and if the methods which calculate interest rates from discount factors (like `getForwardRate()` oder `getYield()`) only use `getDiscountFactor()` (and not the internal data model), then the interpolation method is available in all derived interest rates once it has been implemented in `getDiscountFactor()`.⁹

The bootstrapper is then realized through one additional method `appendDiscountFactor(ProductSpecification productSpec, double marketPrice)`, where `productSpec` contains the description of the financial product for which an additional discount factor has to be calculated from the given market price `marketPrice`.

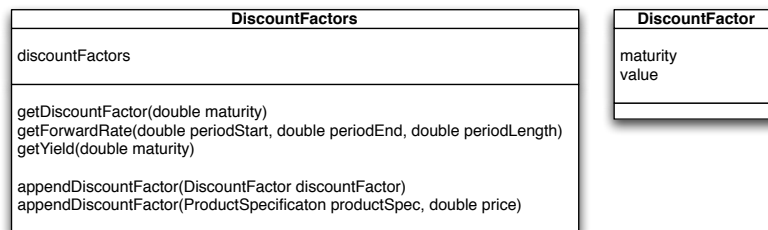


Figure 8.4.: UML Diagram: The class *DiscountFactors* internally stores discount factors and provides various interest rates through corresponding methods. The method *appendDiscountFactor(ProductSpecification productSpec, double marketPrice)* implements one induction step of a bootstrap algorithm.

⁹ This is one reason for [encapsulation](#) of the internal data model, which should only be accessible to a small set of methods (even within the same class!).

CHAPTER 9

Simple Interest Rate Products

So far we have defined a single interest rate product, the [zero coupon bond](#) $P(T)$. In the following, we give the definitions of some basic interest rate products. Many definitions use Definition [88](#) of the forward rate (which, of course, is based on the definition of the zero bond).

9.1. Interest Rate Products Part 1: Products without Optionality

9.1.1. Fix, Floating and Swap

We define a trivial generalization of the [zero bond](#):

Definition 97 (Coupon Bond): ↗

A coupon bond with coupon C_i , $i = 1, \dots, n - 1$ and *tenor structure* T_i , $i = 1, \dots, n$ and maturity T_n pays

$$C_i \cdot (T_{i+1} - T_i) + \begin{cases} 0 & \text{if } i + 1 < n \\ 1 & \text{if } i + 1 = n \end{cases} \quad \text{in } T_{i+1}$$

($n - 1$ payments). ↘

Theorem 98 (Value of a Coupon Bond): The Coupon Bond consists of $n - 1$ guaranteed payments with different payment dates. Obviously, we have that the value of the coupon bond as seen in $t < T_2$ is given by

$$\sum_{i=1}^{n-1} C_i \cdot (T_{i+1} - T_i) \cdot P(T_{i+1}; t) + P(T_n; t). \quad (9.1)$$

Remark 99 (Dirty Price, Clean Price, Accrued Interest): The value of a coupon bond as given by Equation (9.1) is called *dirty price*. The dirty price is sometimes split into two parts, called the *Clean Price* and *Accrued Interest*.

If $T_1 < t < T_2$, i.e. the bond is evaluated within the first interest rate period, then the accrued interest is defined by

$$A(T_1, T_2; t) := \frac{t - T_1}{T_2 - T_1} \cdot C_1 \cdot (T_2 - T_1).$$

Note that we assumed $t < T_2$ already in Definition 98. $A(T_1, T_2; t)$ is called *accrued interest*. The dirty price and clean price are now related through

$$P_{\text{Dirty}}(t) = \sum_{i=1}^{n-1} C_i \cdot (T_{i+1} - T_i) \cdot P(T_{i+1}; t) + P(T_n; t), \quad P_{\text{Clean}}(t) = P_{\text{Dirty}}(t) - A(T_1, T_2; t).$$

The *accrued interest* represents the fraction of the future coupon payment that relates to the past fraction of the period. This stems from the interpretation of the coupon payment as equally distributed. The price of a bond is often quoted only by its *Clean Price*.

The decomposition in *clean price* and *accrued interest* may appear useless, since upon trading their sum, i.e. the dirty price, has to be paid. However, quoting the clean price has an advantage: The clean price evolves continuously in t across period end dates, while the dirty price exhibits a jump at the end of a period, due to the paid coupon.

The zero bond $P(T_1; t)$ is the time t value of a guaranteed payment of 1 in T_1 . It thus represents a fixed interest rate payment. A product with variable interest rate payments is the *floater*.

Definition 100 (Floater): ↗

Let T_i , $i = 1, \dots, n$ denote a given time discretization (a tenor structure). A floater pays

$$N \cdot L(T_i, T_{i+1}; T_i) \cdot (T_{i+1} - T_i) \quad \text{in } T_{i+1}$$

for $i = 1, \dots, n - 1$ ($n - 1$ payments). ↘

Theorem 101 (Value of a Floater): At time $t \leq T_1$ we have that the value of a floater (as in Definition 100) is given by

$$\begin{aligned} V_{\text{Floater}}(t) &= N \cdot \sum_{i=1}^{n-1} L(T_i, T_{i+1}; t) \cdot (T_{i+1} - T_i) \cdot P(T_{i+1}; t) \\ &= N \cdot (P(T_1) - P(T_n)). \end{aligned}$$

Proof: *Variant 1 of the proof:* From the definition of the forward rate we have for a single payment of the floater in time T_{i+1}

$$V_{\text{Floater}}^i(T_{i+1}) := N \cdot L(T_i, T_{i+1}; T_i) \cdot (T_{i+1} - T_i) \stackrel{\text{Def. } L}{=} N \cdot \frac{P(T_i; T_i) - P(T_{i+1}; T_i)}{P(T_{i+1}; T_i)}.$$

This payment is a $\mathcal{F}_{T_i}$ -measurable random variable. The interest rate of V_{Floater}^i was fixed in T_i and is no longer stochastic when observed on $[T_i, T_{i+1}]$. As seen in T_i the value of this payment is thus a multiple of $P(T_{i+1}; T_i)$, namely

$$V_{\text{Floater}}^i(T_i) = V_{\text{Floater}}^i(T_{i+1}) \cdot P(T_{i+1}; T_i) = N \cdot (P(T_i; T_i) - P(T_{i+1}; T_i)).$$

Thus we see that the time T_i value of the floater is given by a portfolio of bonds. The time t value of this portfolio is known. Thus we have

$$V_{\text{Floater}}^i(t) = N \cdot (P(T_i; t) - P(T_{i+1}; t)) = N \cdot L(T_i, T_{i+1}; t) \cdot (T_{i+1} - T_i) \cdot P(T_{i+1}; t).$$

This is the value of a single floater payment. The claim follows by summation over i .

Variant 2 of the proof: We have to value $V_{\text{Floater}}^i(T_{i+1})$. Choose $N = P(T_{i+1})$ as Numéraire and let $\mathbb{Q}^N$ denote a corresponding martingal measure. Then we have that $N(T_{i+1}) = 1$ and

$$\begin{aligned} V_{\text{Floater}}^i(t) &= N(t) \cdot \mathbb{E}^{\mathbb{Q}^N} \left(\frac{V_{\text{Floater}}^i(T_{i+1})}{N(T_{i+1})} \middle| \mathcal{F}_t \right) = N(t) \cdot \mathbb{E}^{\mathbb{Q}^N} (V_{\text{Floater}}^i(T_{i+1}) | \mathcal{F}_t) \\ &= N(t) \cdot \mathbb{E}^{\mathbb{Q}^N} \left(N \cdot \frac{P(T_i; T_i) - P(T_{i+1}; T_i)}{P(T_{i+1}; T_i)} \middle| \mathcal{F}_t \right) \\ &= N(t) \cdot \mathbb{E}^{\mathbb{Q}^N} \left(N \cdot \frac{P(T_i; T_i) - P(T_{i+1}; T_i)}{N(T_i)} \middle| \mathcal{F}_t \right) \\ &= N(t) \cdot N \cdot \frac{P(T_i; t) - P(T_{i+1}; t)}{N(t)} \\ &= N \cdot (P(T_i; t) - P(T_{i+1}; t)) = N \cdot L(T_i, T_{i+1}; t) \cdot (T_{i+1} - T_i) \cdot P(T_{i+1}; t). \end{aligned}$$

This is the value of a single floater payment. The claim follows by summation over i . □

Definition 102 (Floating Rate Bond): ↗

Let T_i , $i = 1, \dots, n$ denote a given time discretization (a tenor structure). A floating rate bond pays

$$N \cdot L(T_i, T_{i+1}; T_i) \cdot (T_{i+1} - T_i) + N \cdot \begin{cases} 0 & \text{if } i+1 < n \\ 1 & \text{if } i+1 = n \end{cases} \quad \text{in } T_{i+1}$$

for $i = 1, \dots, n-1$ ($n-1$ payments). ↘

The value of a floating rate bond is $N \cdot P(T_1)$, because it is just the sum of a floater (value $N \cdot P(T_1) - N \cdot P(T_2)$) and a zero coupon bond with maturity T_2 (value $N \cdot P(T_2)$).

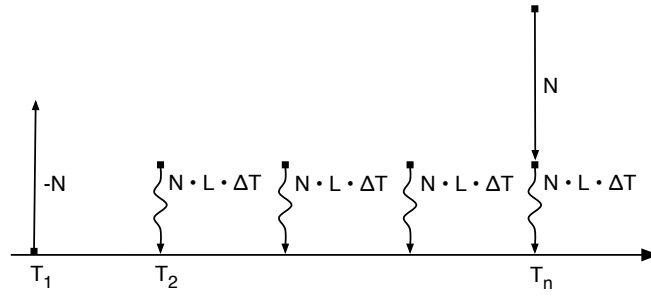


Figure 9.1.: Cashflow of a floater with exchange of notional N .



Interpretation: Definition 100 considers only the coupon payments of a floater. It is common that an exchange of notional takes place at the beginning and end of the product: A payment of $-N$ is made in T_1 (receive notional) and a payment of N is made in T_n (pay notional). Since from Theorem 101 the value of the pure coupon payments is $N \cdot P(T_1) - N \cdot P(T_n)$, we have that the value of a floater with exchange of notional is 0.

Figure 9.1 shows a cashflow diagram for a floater with exchange of notional. At time T_1 the notional N is invested over the period $[T_1, T_2]$ with an interest rate $L(T_1, T_2; T_1)$, fixed at the beginning of the period. In T_2 the interest is paid and the notional N is reinvested over the following period (with a newly fixed rate). At the end T_n the interest for the last period is paid together with the notional. $\triangleleft$

We derived the value of the floater using two different ways. The first method used the fact that the payment $N \cdot L(T_i, T_{i+1}; T_i) \cdot (T_{i+1} - T_i)$ is a $\mathcal{F}_{T_i}$ -measurable random variable paid in T_{i+1} . Thus, its value as of time T_i is given by multiplication with $P(T_{i+1}; T_i)$. Since this value could be expressed as a portfolio of bonds we know its time t value. Essentially, we have derived a replication portfolio for each cashflow. The second method considered relative prices and applied Theorem 70.

In this context, the time T_i is called *fixing date* and the time T_{i+1} is called *payment date*.

Definition 103 (Fixing Date, Payment Date): $\triangleright$

Let $T_2 \geq T_1$ and $V_{T_1}(T_2)$ be a $\mathcal{F}_{T_1}$ -measurable random variable defining a payment made in T_2 . Then T_1 is called *Fixing Date* and T_2 is called *Payment Date* of $V_{T_1}(T_2)$. See also Figure 9.2. $\lrcorner$

Lemma 104 (Moving the payment date): Let $t \geq T_1$. The value of a payment $V_{T_1}(T_2)$ with *fixing date* T_1 and *payment date* T_2 corresponds to the value of a payment of $V_{T_1}(T_2) \cdot P(T_2; t)$ in t for $t < T_2$ and the value of a payment of $V_{T_1}(T_2) \cdot \frac{1}{P(t; T_2)}$ in t for $t > T_2$.

Proof: The first part follows as in the proof of Theorem 101, the second part follows

from exchanging t and T_2 . □

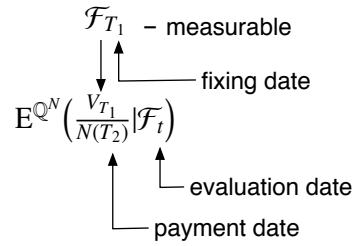


Figure 9.2.: Fixing date, payment date and evaluation date

Remark 105 (On the additivity of cash flows with different payment dates): The value of a financial product or a single cash flow depends on its evaluation time, the time upon which we observe it. Two payments with the same payment date may be added to a single one. For two payments with different payment dates a summation is not meaningful. To calculate the total value of several products or cash flows at time t we have to move all cash flows as in Lemma 104. However, Lemma 104 applies only to times t larger than the fixing dates. The lemma is not applicable for times before the fixing date. Here, a (risk neutral) evaluation has to be performed.

Relative prices behave differently: *Relative prices are additive (independent of the fixing date)*. Let V_{T_1} denote a $\mathcal{F}_{T_1}$ -measurable random variable defining the value of a financial product in time T_2 . Then we have:

- For $t < T_1$: $E^{Q^N}(\frac{V_{T_1}}{N(T_2)} | \mathcal{F}_t)$ is the $N(t)$ -relative value of V_{T_1} at time t . This follows from Theorem 70.
- For $t > T_1$: $E^{Q^N}(\frac{V_{T_1}}{N(T_2)} | \mathcal{F}_t) = V_{T_1} \cdot E^{Q^N}(\frac{1}{N(T_2)} | \mathcal{F}_t) = \frac{V_{T_1} \cdot P(T_2; t)}{N(t)}$ is the $N(t)$ -relative value of V_{T_1} at time t . This follows from Lemma 104.

The additivity of relative prices follows from the linearity of the expectation operator.

Definition 106 (Swap (Payer Swap)): □

A swap is an exchange payment of fixed rate for a floating rate. Let $0 = T_0 < T_1 < T_2 < \dots < T_n$ denote a given tenor structure. A swap pays

$$N \cdot (L(T_i, T_{i+1}; T_i) - S_i) \cdot (T_{i+1} - T_i) \quad \text{in } T_{i+1}$$

for $i = 1, \dots, n-1$ ($n-1$ payments), where $S_i \in \mathbb{R}$ denotes the fixed *swap rate* and $L(T_i, T_{i+1}; T_i)$ denotes the forward rate from Definition 88, and N denotes the *notional*. The swap defined here is a *payer swap*, see Definition 107. □

Definition 107 (Payer Swap, Receiver Swap): ↗

The swap defined in Definition 106 with payments

$$N \cdot (L(T_i, T_{i+1}; T_i) - S_i) \cdot (T_{i+1} - T_i) \quad \text{in } T_{i+1}$$

is called *payer swap*. In contrast the swap with reversed payments

$$N \cdot (S_i - L(T_i, T_{i+1}; T_i)) \cdot (T_{i+1} - T_i) \quad \text{in } T_{i+1}$$

is called *receiver swap*. The term payer / receiver points to whether the holder of the swap has to pay the fixed coupon (it enters negative) or receives the fixed coupon (it enters positive). ↘

Definition 108 (Floating Leg, Fixed Leg): ↗

The payments of a swap may be decomposed into

$$N \cdot L(T_i, T_{i+1}; T_i) \cdot (T_{i+1} - T_i) \quad \text{in } T_{i+1} \tag{9.2}$$

and

$$N \cdot S_i \cdot (T_{i+1} - T_i) \quad \text{in } T_{i+1}. \tag{9.3}$$

The payments (9.2) of the variable rates are called *floating leg*, the payments of the constant rates (9.3) are called *fixed leg*. ↘

Theorem 109 (Value of a Swap): At time $t \leq T_1$ the value of a *swap* is given by

$$V_{\text{Swap}}(t) = N \cdot \sum_{i=1}^{n-1} (L(T_i, T_{i+1}; t) - S_i) \cdot (T_{i+1} - T_i) \cdot P(T_{i+1}; t).$$

Proof: The swap consists of a floater (floating leg, (9.2)) and fixed payments $-N \cdot S_i \cdot (T_{i+1} - T_i)$ in T_{i+1} for which their time t value is the corresponding multiple of $P(T_{i+1}; t)$. The claim follows by applying Theorem 101 to the floating leg. □

Remark 110 (Swap rate): Let $T_1, \dots, T_n$ be a given tenor structure. Consider a swap as in Definition 106. The *par swap rate* S (in t) is the unique rate for which a swap with $S_i := S$ has the time t value 0, i.e. the total time t value of the payments

$$N \cdot (L(T_i, T_{i+1}; T_i) - S) \cdot (T_{i+1} - T_i) \quad \text{in } T_{i+1}$$

is 0. Since the time t value of such a swap is given by

$$N \cdot \sum_{i=1}^{n-1} (L(T_i, T_{i+1}; t) - S) \cdot (T_{i+1} - T_i) \cdot P(T_{i+1}; t),$$

we have that

$$\begin{aligned} S &= \frac{\sum_{i=1}^{n-1} L(T_i, T_{i+1}; t)(T_{i+1} - T_i) \cdot P(T_{i+1}; t)}{\sum_{i=1}^{n-1} (T_{i+1} - T_i) \cdot P(T_{i+1}; t)} \\ &= \frac{P(T_1; t) - P(T_n; t)}{\sum_{i=1}^{n-1} (T_{i+1} - T_i) \cdot P(T_{i+1}; t)}. \end{aligned}$$

Definition 111 (Par swap rate): ↗

Let $T_1 < T_2 < \dots < T_n$. The *par swap rate* (often just called *swap rate*) $S(T_1, \dots, T_n)$ is defined by

$$S(T_1, \dots, T_n; t) := \frac{P(T_1; t) - P(T_n; t)}{\sum_{i=1}^{n-1} (T_{i+1} - T_i) \cdot P(T_{i+1}; t)}.$$



Interpretation: Since the par swap rate is the rate for which a corresponding swap has value 0, we may see the par swap rate $S_{i,j} := S(T_i, \dots, T_j)$ as some mean of the forward rates L_k , $k = i, \dots, j-1$. Indeed, the par swap rate $S_{i,j}$ is a convex combination (and thus a weighted average) of the forward rates L_k , $k = i, \dots, j-1$ as shown in the following lemma. ↘

Lemma 112 (Swap rate as convex combination of the forward rates): Let $T_i < T_{i+1} < \dots < T_j$ denote a given tenor structure. Then we have

$$S_{i,j} = \sum_{k=i}^{j-1} \alpha_k L_k, \quad \text{with } \alpha_k \geq 0, \quad \sum_{k=i}^{j-1} \alpha_k = 1.$$

The weights α_k are given by $\alpha_k := \frac{P(T_{k+1})(T_{k+1} - T_k)}{\sum_{l=i}^{j-1} P(T_{l+1})(T_{l+1} - T_l)}$. The weights are stochastic.

Proof: With $P(T_{k+1})(T_{k+1} - T_k) \cdot L_k = P(T_k) - P(T_{k+1})$ we have

$$\sum_{k=i}^{j-1} P(T_{k+1})(T_{k+1} - T_k) \cdot L_k = P(T_i) - P(T_j),$$

and thus

$$\sum_{k=i}^{j-1} \frac{P(T_{k+1})(T_{k+1} - T_k)}{\sum_{l=i}^{j-1} P(T_{l+1})(T_{l+1} - T_l)} \cdot L_k = \frac{P(T_i) - P(T_j)}{\sum_{l=i}^{j-1} P(T_{l+1})(T_{l+1} - T_l)} = S_{i,j}.$$

□



Interpretation (Usage of the term *bond* and *swap*): The terms *bond* and *swap* are used in a much broader sense, compared to what has been defined in the previous definitions. A financial product with coupon payments and final notional payments at maturity is called a *bond*. A financial product where coupon payments are exchanged (and no notional is paid) is called a *swap*. The terms are used independently of the specific structure of the coupons. A coupon may be a constant (*fix*), a variable rate (*float*) or a complex function of one or more interest rates (*structured*). In the last case, the coupon is called a *structured coupon* and the corresponding bond and swap are called *structured bond* and *structured swap*.

A swap may be interpreted as a portfolio of a bond long (i.e. with positive cashflows) and a bond short (i.e. with negative cashflows), where the two notional payments cancel. In the following Section 12.2.1 we will consider the relation of bonds and swaps.

<|

9.1.2. Money-Market Account

If we invest at time $T_0 = 0$ a unit currency over the period $[T_0, T_1]$, then we receive at T_1 the amount $1 + L(T_0, T_1; T_0) \cdot (T_1 - T_0)$. If this amount is reinvested for another period and if this process is continued for periods $[T_j, T_{j+1}]$ with $j = 1, 2, \dots$, then we have at time T_i a value of

$$\prod_{j=0}^{i-1} (1 + L(T_j, T_{j+1}; T_j) \cdot (T_{j+1} - T_j)). \quad (9.4)$$

Equivalently, we may write this with the instantaneous forward rate as

$$\prod_{j=0}^{i-1} \exp \left(\int_{T_j}^{T_{j+1}} f(T_j, \tau) d\tau \right) = \exp \left(\sum_{j=0}^{i-1} \int_{T_j}^{T_{j+1}} f(T_j, \tau) d\tau \right). \quad (9.5)$$

If we consider a continuum of infinitesimal periods, i.e. we consider continuously compounding with the short rate, then the corresponding value will evolve as

$$B(t) = \exp \left(\int_{T_0}^t r(\tau) d\tau \right). \quad (9.6)$$

Definition 113 (Rolling Bond):

The financial product

$$R(t) := P(T_{m(t)+1}; t) \cdot \prod_{j=0}^{m(t)} (1 + L(T_j, T_{j+1}; T_j) \cdot (T_{j+1} - T_j)),$$

7

where $m(t) := \max\{i : T_i \leq t\}$, is called (*single period*) *rolling bond*. $\lrcorner$

Definition 114 (Savings Account, Money Market Account): $\lrcorner$

The financial product B in (9.6) is called *savings account* or *money market account*. $\lrcorner$



Interpretation: The financial product B has to be interpreted as an idealization (like the short rate itself), since infinitesimal periods are an idealization.

Note that (9.4) and (9.5) are equivalent, whereas $B(T_i)$ does not coincide with the value of (9.4) in general. The expressions (9.4) and (9.5) depend on the choice of the periods and if evaluated in T_i , they are $\mathcal{F}_{T_{i-1}}$ measurable random variables, whereas $B(T_i)$ is $\mathcal{F}_{T_i}$ measurable only. $\triangleleft$

9.2. Interest Rate Products Part 2: Simple Options

9.2.1. Cap, Floor, Swaption

Definition 115 (Caplet): $\lrcorner$

A *Caplet* is an option on the *forward rate* (LIBOR) and pays

$$V_{\text{caplet}}(T_2) = N \cdot \max(L(T_1, T_2; T_1) - K, 0) \cdot (T_2 - T_1) \quad \text{in } T_2 \quad (9.7)$$

where K is the *strike rate*, $L(T_1, T_2; t)$ the LIBOR and N the *notional*. Payment date and fixing date $0 < T_1 < T_2$ coincide with the LIBOR period $[T_1, T_2]$. $\lrcorner$

Definition 116 (Cap): $\lrcorner$

A *cap* is a portfolio of caplets. Let $0 = T_0 < T_1 < T_2 < \dots < T_n$ denote a given tenor structure. A cap pays

$$N \cdot \max(L(T_i, T_{i+1}; T_i) - K_i, 0) \cdot (T_{i+1} - T_i) \quad \text{in } T_{i+1} \quad (9.8)$$

for $i = 1, \dots, n-1$ ($n-1$ payments), where K_i are the *strike rates*, $L(T_i, T_{i+1}; T_i)$ are the LIBOR rates and N denotes the *notional*. $\lrcorner$

Remark 117 (Floorlet, Floor): If in (9.8) or correspondingly in (9.7) the payoff is

$$N \cdot \max(K_i - L(T_i, T_{i+1}; T_i), 0) \cdot (T_{i+1} - T_i) \quad \text{in } T_{i+1},$$

then the product is called *Floor* or *Floorlet*, respectively.

Remark 118 (Caplet, Cap): The name caplet (and thus cap) seems contra-intuitive.

A cap is usually an upper bound, a floor a lower bound. Indeed, the payoff

$$[L]^K := \min(L, K)$$

is called *capped* and the payoff

$$[L]_K := \max(L, K)$$

is called *floored*. The contra-intuitive name caplet for (9.7) stems from its application as a swap that exchanges a floating rate L against a capped coupon $[L]^K$:

$$L - [L]^K = \max(L - K, 0),$$

i.e.

$$-[L]^K = -L + \max(L - K, 0).$$

If we have the obligation to pay a variable interest rate ($-L$), buying a cap ($+\max(L - K, 0)$) will cover the risk of an increasing interest rate, i.e. the payment is capped ($-[L]^K$). The cap is the product one has to buy to have floating payments capped.¹

Definition 119 (Swaption):

A *swaption* is an option on a *swap*. Let $V_{\text{swap}}(t)$ denote the time t value of a swap as defined by Definition 106. Then the value of a *swaption* (with *underlying* V_{swap}) is given by the payout

$$V_{\text{swaption}}(T_1) := \max(V_{\text{swap}}(T_1), 0) \quad \text{in } T_1.$$

Definition 120 (Digital Caplet):

A *digital caplet* pays

$$V_{\text{digital}}(T_2) = N \cdot \mathbf{1}(L(T_1, T_2; T_1) - K) \cdot (T_2 - T_1) \quad \text{in } T_2,$$

where K is the *strike* rate, $L(T_1, T_2; t)$ is the LIBOR, N denotes the *notional* and $\mathbf{1}$ denotes the indicator function (with $\mathbf{1}(x) := 1$ for $x > 0$ and $\mathbf{1}(x) := 0$ else).

Lemma 121 (Digital Caplet valuation, Call-Spread): For the value $V_{\text{digital}}(K, 0)$ of a digital caplet with strike K we have

$$V_{\text{digital}}(K; 0) = -\frac{\partial}{\partial K} V_{\text{caplet}}(K; 0).$$

The approximation of the differential using finite differences

$$\tilde{V}_{\text{digital}}(K; 0) = -\frac{V_{\text{caplet}}(K + \epsilon; 0) - V_{\text{caplet}}(K - \epsilon; 0)}{2\epsilon}$$

is called *call spread*.

Proof: The proof follows the lines of the proof in Lemma 72.

¹ See [8], p. 12.

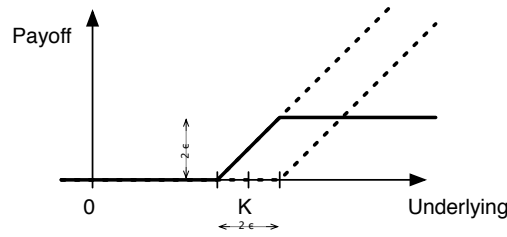


Figure 9.3: Call spread approximation of a digital option by two call options.

9.2.2. Foreign Caplet, Quanto

Definition 122 (Foreign Caplet):

A *foreign caplet* is a caplet in a foreign currency. From the domestic investor's view it pays

$$\tilde{N} \cdot \max(\tilde{L}(T_1, T_2; T_1) - K, 0) \cdot (T_2 - T_1) \cdot FX(T_2) \quad \text{in } T_2,$$

where $K \in \mathbb{R}$ is the strike rate, $\tilde{L}(T_1, T_2; t)$ is the *foreign* LIBOR, $FX(T_2)$ is the exchange rate and $\tilde{N}$ denotes the notional in foreign currency.

Remark 123 (Units): It is useful to consider units, like one would do in physics. A domestic bond P has the unit of one domestic currency, $[P] = \text{dom}$. Interest rates have the dimension $\frac{1}{\text{time}}$, e.g. for the forward rate (LIBOR) we have $[L(T_i, T_{i+1}) \cdot (T_{i+1} - T_i)] = 1$, since it is the quotient of two bonds. The unit of the stochastic process FX is $[FX] = \frac{\text{dom}}{\text{for}}$, i.e. $FX(t)$ is the time t value of a foreign currency unit in domestic currency. In Definition 122 we have $[\tilde{N}] = \text{for}$. For the following product it is crucial to consider units.

Definition 124 (Quanto, Quanto Caplet):

A *quanto* is a financial product, for which a payout will be converted from a foreign currency to a domestic currency without use of the exchange rate. Instead of $FX(t)$ it uses $1 \frac{\text{dom}}{\text{for}}$ or another conversion factor (the *quanto rate*) fixed a-priori.

Let $0 < T_1 < T_2$ denote *fixing* and *payment date*, respectively. A *Quanto Caplet* pays

$$\tilde{N} \cdot \max(\tilde{L}(T_1, T_2; T_1) - K, 0) \cdot (T_2 - T_1) \cdot 1 \frac{\text{dom}}{\text{for}} \quad \text{in } T_2,$$

where K is the strike rate (dimension $\frac{1}{\text{time}}$), $\tilde{L}(T_1, T_2; t)$ is the *foreign* LIBOR (dimensionless) and $\tilde{N}$ denotes the notional in *foreign* currency.



Further Reading: An introduction to the basics of interest rates products may be found in [5] (in German).

CHAPTER 10

The Black Model for a Caplet

We consider a caplet as defined by Definition 115, an option on the forward rate $L_1 := L(T_1, T_2)$ for given times $0 < T_1 < T_2$.

The Black model for the valuation of a caplet postulates a log-normal dynamic of the underlying LIBOR¹

$$dL_1(t) = \mu^{\mathbb{P}}(t)L_1(t)dt + \sigma(t)L_1(t)dW^{\mathbb{P}}(t), \quad \sigma(t) \geq 0, \text{ under } \mathbb{P}. \quad (10.1)$$

We seek the price $V(0)$ of the payoff profile

$$V(T_2) := \max((L_1(T_1) - K) \cdot (T_2 - T_1), 0),$$

where $L_1 \cdot (T_2 - T_1) := P(T_1)/P(T_2) - 1$, i.e. $L_1 = L_1(T_1, T_2)$ denotes the forward rate (the (forward) LIBOR) of the period $[T_1, T_2]$. Without loss of generality we assumed that the notional is 1. We choose the T_2 -Bond as Numéraire

$$N(t) := P(T_2; t).$$

This choice of the Numéraire is the crucial trick in the derivation of a risk neutral pricing formula. Since

$$\begin{aligned} L_1 &= \frac{1}{(T_2 - T_1)} \cdot (P(T_1)/P(T_2) - 1) = \frac{1}{(T_2 - T_1)} \cdot \frac{P(T_1) - P(T_2)}{P(T_2)} \\ &= \frac{\frac{1}{(T_2 - T_1)} \cdot (P(T_1) - P(T_2))}{N}, \end{aligned}$$

we have that L_1 is the N -relative price of a traded asset.² From assumption (3.23) we have the existence of an *equivalent martingale measure* $\mathbb{Q}^N$ such that all N -relative prices of traded assets are martingales. Thus L_1 is drift-free (see Lemma 47)), i.e.

$$dL_1(t) = \sigma(t)L_1(t)dW^{\mathbb{Q}^N}(t), \quad \text{under } \mathbb{Q}^N.$$

¹ The log-normal Process is often written in the form $\frac{dL_1(t)}{L_1(t)} = \mu^{\mathbb{P}}(t)dt + \sigma(t)dW^{\mathbb{P}}(t)$.

² The traded asset is the portfolio $\frac{1}{(T_2 - T_1)} \cdot (P(T_1) - P(T_2))$ consisting of $\frac{1}{(T_2 - T_1)}$ fractions of a T_1 -bond (long) and T_2 -bond (short).

For the process $Y := \log(L_1)$ we have from Lemma 44 that

$$d(\log(L_1(t))) = -\frac{1}{2}\sigma(t)^2 dt + \sigma(t)dW^{\mathbb{Q}^N}(t),$$

i.e. $\log(L_1(T))$ is normal distributed with mean $\log(L_1(0)) - \frac{1}{2}\bar{\sigma}^2 T$ and standard deviation $\bar{\sigma}\sqrt{T}$, with $\bar{\sigma} := (\frac{1}{T} \int_0^T \sigma^2(t) dt)^{1/2}$, see Section 4.

For the option value we now have

$$V(T_2) = \max((L_1(T_1) - K) \cdot (T_2 - T_1), 0) \quad \text{in } T_2$$

and from $N(T_2) = 1$ we have³

$$\begin{aligned} \frac{V(0)}{N(0)} &= E^{\mathbb{Q}^N} \left(\frac{V(T_2)}{N(T_2)} \mid \mathcal{F}_0 \right) = E^{\mathbb{Q}^N} (V(T_2) \mid \mathcal{F}_0) \\ &= E^{\mathbb{Q}^N} (\max((L_1(T_1) - K) \cdot (T_2 - T_1), 0) \mid \mathcal{F}_0) \end{aligned}$$

i.e.

$$V(0) = P(T_2; 0) \cdot E^{\mathbb{Q}^N} (\max((L_1(T_1) - K), 0)) \cdot (T_2 - T_1).$$

Knowing the distribution of L_1 under $\mathbb{Q}^N$ this expectation may be represented as

$$V(0) = P(T_2; 0) \cdot [L_1(0)\Phi(d_+) - K\Phi(d_-)] \cdot (T_2 - T_1), \quad (10.2)$$

where $\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp(-\frac{y^2}{2}) dy$ and $d_{\pm} = \frac{\ln(\frac{L_1(0)}{K}) \pm \frac{1}{2}\bar{\sigma}^2 T_1}{\bar{\sigma}\sqrt{T_1}}$, see Section 4.

Remark 125 (Implied Black Volatility): Similarly to Remark 71 on page 88 we have: Equation (4.3) gives us the price of the option under the model (10.1) as a function of the model parameter $\bar{\sigma}$. In this context $\bar{\sigma}$ is called the Black volatility. Taking the other model and product parameters (r, K, T_1, T_2) as constants, the pricing formula (10.2) represents a bijection

$$\begin{aligned} \bar{\sigma} &\mapsto V(0) \\ [0, \infty) &\rightarrow [P(T_2; 0) \cdot \max(L(0) - K, 0), P(T_2; 0) \cdot L(0)]. \end{aligned}$$

The $\bar{\sigma}$ calculated for a given price $V(0)$ through inversion of the pricing formula is called the *implied Black volatility*.

Lemma 126 (Price of a digital caplet under the Black model dynamics): For the price of a digital caplet we have under the Black model that

$$V_{\text{digital}}(K; 0) = N \cdot P(T_2; 0) \cdot \Phi(d_-) \cdot (T_2 - T_1),$$

where $\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp(-\frac{y^2}{2}) dy$ and $d_- = \frac{\ln(\frac{L_1(0)}{K}) - \frac{1}{2}\bar{\sigma}^2 T_1}{\bar{\sigma}\sqrt{T_1}}$.

³ This is the point where the specific choice of Numéraire comes in handy for the second time.

CHAPTER 11

Pricing of a Quanto Caplet (Modeling the FFX)

In this section all quantities related to a foreign currency are marked with a tilde. Let $0 < T_1 < T_2$ denote *fixing* and *payment date* respectively. The payoff profile of a quanto caplet is given by

$$V(T_2) = \max\left(\tilde{L}(T_1, T_2; T_1) - K, 0\right) \cdot (T_2 - T_1) \cdot 1 \frac{\text{dom}}{\text{for}} \quad \text{in } T_2,$$

where K is the given strike rate and $\tilde{L}(T_1, T_2; t)$ is the foreign forward rate. The notional and quanto rate are assumed to be 1.

We assume a log-normal dynamic for the foreign LIBOR, i.e. we model it as¹

$$d\tilde{L}(t) = \mu^{\mathbb{P}}(t)\tilde{L}(t) dt + \sigma_{\tilde{L}}(t)\tilde{L}(t) dW_3^{\mathbb{P}}(t).$$

11.1. Choice of Numéraire

If we would choose the foreign T_2 -bond converted to domestic currency, i.e. $\tilde{P}(T_2; t)FX(t)$, as Numéraire, then we have from $\tilde{L}(T_1, T_2; t) = \frac{1}{T_2 - T_1} \frac{\tilde{P}(T_1) - \tilde{P}(T_2)}{\tilde{P}(T_2)} = \frac{1}{T_2 - T_1} \frac{\tilde{P}(T_1)FX(t) - \tilde{P}(T_2)FX(t)}{\tilde{P}(T_2)FX(t)}$ (see Section 10) that

$$d\tilde{L}(t) = \sigma_{\tilde{L}}(t)\tilde{L}(t) dW_3(t) \quad \text{under } \mathbb{Q}^{\tilde{P}(T_2) \cdot FX}.$$

Remark 127 (Foreign Market, Cross Currency Change of Numéraire): Note that the foreign LIBOR is *not* a martingale with respect to $\mathbb{Q}^{\tilde{P}(T_2)}$, since we are based in the domestic market. For the domestic investor the foreign bond $\tilde{P}(T_i)$ is not a traded

¹ For short we write $\tilde{L}$ for $\tilde{L}(T_1, T_2)$.

asset, but the foreign bond converted to domestic currency $\tilde{P}(T_i) \cdot FX$ is a traded asset. Although we have

$$d\tilde{L}(t) = \sigma_{\tilde{L}}(t)\tilde{L}(t) dW_3(t) \quad \text{under } \mathbb{Q}^{\tilde{P}(T_2)},$$

we cannot use this change of measure, since $\tilde{P}(T_2)$ is not a traded asset and thus not a Numéraire in the domestic market. Choosing the domestic bond $P(T_2)$ (a traded asset) as Numéraire, we have in general

$$d\tilde{L}(t) \neq \sigma_{\tilde{L}}(t)\tilde{L}(t) dW_3(t) \quad \text{under } \mathbb{Q}^{P(T_2)}.$$

Since the payoff profile of the quanto caplet is

$$V(T_2) = \max(\tilde{L}(T_1, T_2; T_1) - K, 0) \cdot (T_2 - T_1) \cdot 1 \frac{\text{dom}}{\text{for}} \quad \text{in } T_2,$$

it is advantageous to know the dynamics of $\tilde{L}(T_1, T_2)$ under the measure $\mathbb{Q}^{P(T_2)}$, the *domestic T_2 terminal measure*. Choosing $P(T_2)$ as Numéraire, we have that the Numéraire is 1 at payment date and will not show up in the expectation operator above. This trick was already used in Chapter 10, there we were lucky that additionally the underlying was a martingale under this measure.

By the change of measure from $\mathbb{Q}^{\tilde{P}(T;t) \cdot FX(t)}$ to $\mathbb{Q}^{P(T_2)}$ we have a change of the drift, see Theorem 53 (Girsanov, Cameron, Martin), i.e.

$$d\tilde{L}(t) = \mu^{P(T_2)}(t)\tilde{L}(t) dt + \sigma_{\tilde{L}}(t)\tilde{L}(t) dW_3^{P(T_2)}(t) \quad \text{under } \mathbb{Q}^{P(T_2)}.$$

With other words: The dynamics of the underlying is known under the measure $\mathbb{Q}^{\tilde{P}(T;t) \cdot FX(t)}$. From the shape of the payoff function a change of Numéraire from $\tilde{P}(T;t) \cdot FX(t)$ to $P(T;t)$, thus a change of measure from $\mathbb{Q}^{\tilde{P}(T;t) \cdot FX(t)}$ to $\mathbb{Q}^{P(T;t)}$ is desirable. We thus define:

Definition 128 (Forward FX Rate): ↗

Let $0 < t < T$. The *forward FX rate* $FFX(T)$ is defined as

$$FFX(T; t) := \frac{\tilde{P}(T; t)}{P(T; t)} FX(t).$$

↘

Remark 129 (Forward FX Rate): The forward FX rate (also known as FX forward) is a relative price of two domestic traded assets. It is dimensionless, since $[\tilde{P}(T; t)] = 1_{\text{for}}$, $[P(T; t)] = 1_{\text{dom}}$ and $[FX(t)] = 1 \frac{\text{dom}}{\text{for}}$. It is a $\mathbb{Q}^{P(T)}$ martingale.

We assume log-normal dynamics for $FFX(T_2)$, i.e.

$$dFFX(T_2; t) = \sigma_{FFX}(t)FFX(T_2; t) dW_2^{P(T_2)}(t) \quad \text{under } \mathbb{Q}^{P(T_2)}.$$

Since

$$\begin{aligned}\tilde{L}(T_1, T_2) \cdot FFX(T_2) &= \frac{1}{T_2 - T_1} \frac{\tilde{P}(T_1) - \tilde{P}(T_2)}{\tilde{P}(T_2)} \cdot \frac{\tilde{P}(T_2)}{P(T_2)} FFX(t) \\ &= \frac{1}{T_2 - T_1} \frac{(\tilde{P}(T_1) - \tilde{P}(T_2)) FFX(t)}{P(T_2)}\end{aligned}$$

is a $P(T_2)$ -relative price of domestic traded assets (namely a portfolio of foreign bonds), we have that $\tilde{L}(T_1, T_2) \cdot FFX(T_2)$ is a martingale under $\mathbb{Q}^{P(T_2)}$, i.e.

$$\text{Drift}^{\mathbb{Q}^{P(T_2)}}(\tilde{L}(T_1, T_2) \cdot FFX(T_2)) = 0. \quad (11.1)$$

On the other hand

$$\begin{aligned}d(\tilde{L} \cdot FFX) &= d\tilde{L} \cdot FFX + \tilde{L} \cdot dFFX + d\tilde{L} \cdot dFFX \\ &= FFX \cdot \tilde{L} \cdot \mu^{P(T_2)} dt + FFX \cdot \tilde{L} \cdot \sigma_{\tilde{L}} dW_3^{P(T_2)} \\ &\quad + FFX \cdot \tilde{L} \cdot \sigma_{FFX} dW_2^{P(T_2)} + FFX \sigma_{FFX} dW_2^{P(T_2)} \cdot \sigma_{\tilde{L}} \tilde{L} dW_3^{P(T_2)},\end{aligned}$$

and assuming an instantaneous correlation $\rho(t)$ for $dW_2^{P(T_2)}$ and $dW_3^{P(T_2)}$

$$= FFX \cdot \tilde{L} \left((\mu^{P(T_2)} + \rho \sigma_{FFX} \sigma_{\tilde{L}}) dt + \sigma_{\tilde{L}} dW_3^{P(T_2)} + \sigma_{FFX} dW_2^{P(T_2)} \right).$$

From (11.1) we thus find

$$\mu^{P(T_2)}(t) = -\rho(t) \sigma_{FFX}(t) \sigma_{\tilde{L}}(t).$$

We now know the dynamics of $\tilde{L}$ under $\mathbb{Q}^{P(T_2)}$

$$d\tilde{L} = -\rho(t) \sigma_{FFX}(t) \sigma_{\tilde{L}}(t) \tilde{L}(t) dt + \sigma_{\tilde{L}}(t) \tilde{L}(t) dW_3^{P(T_2)}(t)$$

and (as in Chapter 10) we know the distribution of $\tilde{L}(T_1, T_2)$ (under $\mathbb{Q}^{P(T_2)}$) as being log-normal with

$$\begin{aligned}\log(\tilde{L}(T)) &= \mathcal{N} \left(\log(\tilde{L}(0)) - \int_0^T \rho(t) \sigma_{FFX}(t) \sigma_{\tilde{L}}(t) dt - \frac{1}{2} \bar{\sigma}_{\tilde{L}}^2 \cdot T, \bar{\sigma}_{\tilde{L}}^2 \cdot T \right) \\ &= \mathcal{N} \left(\log(\tilde{L}(0) \cdot e^{-\int_0^T \rho(t) \sigma_{FFX}(t) \sigma_{\tilde{L}}(t) dt}) - \frac{1}{2} \bar{\sigma}_{\tilde{L}}^2 \cdot T, \bar{\sigma}_{\tilde{L}}^2 \cdot T \right),\end{aligned}$$

where

$$\bar{\sigma}_{\tilde{L}}^2 = \frac{1}{T} \int_0^T \sigma_{\tilde{L}}^2(t) dt \quad (\text{mean variance}).$$

Altogether it follows an (adjusted) Black formula, where in contrast to the Black formula from Chapter 10 $\tilde{L}(0)$ is replaced by $\tilde{L}(0) \cdot e^{-\int_0^T \rho(t) \sigma_{FFX}(t) \sigma_{\tilde{L}}(t) dt}$. The factor

$$e^{-\int_0^T \rho(t) \sigma_{FFX}(t) \sigma_{\tilde{L}}(t) dt}$$

is called the *quanto adjustment*.

CHAPTER 12

Exotic Derivatives

We have already introduced some simple interest rate derivatives. In this section we give a selection of so called *exotic derivatives*. The name “exotic” does not mean that these derivative products are of less importance. With respect to evaluation models the converse is true: The value of exotic derivatives usually depends on a multitude of model properties, which may not even play a role in the pricing of simple derivatives. An example is given by the time structure $t \mapsto \sigma(t)$ of the volatility in the Black-(Scholes-)Model (4.2), (10.1): Its distribution over time does not play a role for the pricing of a [European option](#), only the integrated variance enters into the pricing formula. It will however play a role for the pricing of a [Bermudan option](#).

Thus, in this sense we may view certain prototypical properties of exotic derivatives (e.g. having more than one exercise date) as test functions for prototypical properties of (complex) models (e.g. the term structure of volatility).¹

12.1. Prototypical Product Properties

The list of exotic interest rate derivatives we give in Section 12.2 does not claim to be complete or representative. It is exemplary for *prototypical product properties* and for applications of the pricing models and methodologies which we will discuss later. We focus a bit on more recent products, where we will discuss the relation of *prototypical product properties* to models and their implementation.

Some product properties, like *path dependency* or *early exercise* are characterizing a whole class of products. To evaluate a product of the respective class, the object *path*, i.e. the history, for *path dependency* and *conditional expectation* for *early exercise* are central. The path dependency may be represented best in a *path simulation* ([forward algorithm](#)). The conditional expectation may be represented best in a *state lattice* ([backward algorithm](#)). See Table 12.1.

¹ To make the meaning of this sentence more vivid we like to remind the reader that a digital option may be used to extract the model implied terminal distribution function of it underlying, see Chapter 5. Thus digital options may be viewed as test functions for a model’s terminal distributions.

Furthermore some models have a preferred way of implementation, i.e. as *path simulation* or as *state lattice*. Whether a model may be implemented on a *state lattice* is often decided by it's Markov dimension, see Table 12.2.

Thus, prototypical product properties impose requirements on model and implementation.

Prototypical Product Property	Model Requirement / Implementation
Path Dependency, high Markov dimension	forward algorithm → path simulation, Chapter 13.1
Early Exercise, Bermudan, low Markov dimension	backward algorithm, coarse time discretization → state lattice / tree, Chapter 13.3
Early Exercise, American, low Markov dimension	backward algorithm, fine time discretization → PDE, Chapter ??
Pfad Dependency, low Markov dimension	backward algorithm with extension of state space → Chapter 14
Early Exercise, high Markov dimension	forward algorithm with estimator of conditional expectation → Chapter ??

Table 12.1.: Prototypical product properties and corresponding model requirements and implementation techniques.

Model	Property
Short Rate Models (→ Chapter 20)	low Markov dimension
Market Models (→ Chapter 17)	high Markov dimension
Markov Functional Models (→ Chapter 23)	low Markov dimension

Table 12.2.: Markov dimension of some models.

12.2. Interest Rate Products Part 3: Exotic Interest Rate Derivatives



Motivation (“Why exotic derivatives?”): A simple European option with payoff $\max(L(T) - K, 0)$ may be interpreted as an insurance against an increase of the interest rate $L(T)$. It pays in case of an increasing interest rate the corresponding compensation.

To interpret an exotic derivatives, e.g. one with payoff

$$\begin{array}{lll} C_i & \text{if } L(T_j) < K \forall j \leq i & \text{and } i + 1 < n \\ C_i + 1 & \text{if } L(T_{i-1}) < K \leq L(T_i) & \text{or } i + 1 = n \\ 0 & \text{else.} & \end{array}$$

as an insurance is not intuitive. The payoff above constitutes a [coupon bond](#) which matures if $L(T_i)$ exceeds the rate K . Such a structure is usually offered with an above average coupon C_1 and below average coupons C_i , $i > 1$.² Thus, this product is appealing if the investor would like to receive a high initial coupon (this could be done with a standard coupon bond) and at the same time expects that the interest rate will rise faster than the market predicts. Since the coupon will mature early if interest rates rise, the lower than average coupons C_i , $i > 1$ do not come to effect. Thus, in this case, the investor would have a coupon bond with an above average coupon. Since the investor takes the risk that he receives lower than average coupons if interest rates do not rise, the product will be much cheaper than a standard coupon bond paying a coupon C_1 and maturing early. The investor is financing the initial above average coupon by taking the risk of losing his bet on rising interest rates. He is a risk taker. The product is appealing to the investor since he has a different view on the future as the market (i.e. the average).

Exotic derivatives interpreted as an investment usually link a guaranteed high initial payment with a risky structure, which extracts the favorable case of the investor’s market view.

An exotic derivative may both reward for taking risk as well as cover risk (in the sense of an insurance). Both interpretations jointly exist. For example the structure above has an insurance against total loss of investment. The worst case scenario is a coupon bond with below average coupons. <|

12.2.1. Structured Bond, Structured Swap, Zero Structure

Exotic interest rate options mainly come in two different forms: as a (structured) bond or a (structured) swap. The products bond and swap are closely related. For a given

² A similar structure is given by the [target redemption note](#).

(coupon) bond we may define a swap, such that the swap together with a floating rate bond replicate the coupon. We consider this relationship first for the trivial case of a simple coupon bond, then for the more tricky case of a zero coupon bond. All coupons may be structured coupons.

A structured bond is a bond for which the coupons C_i are arbitrary complex functions of interest rates or other market observables.³ In this case the coupons are called *structured coupons*. A corresponding swap exchanges the structured coupon payments of the bond by coupon payments of a corresponding floating rate bond, see Figure 12.1 and 12.2. Taking the swap and the floating rate bond (which just pays the current market rate on the notional) together, we may hedge the structured bond. For the values of the products in T_0 we thus have $N_1 + V_{\text{swap}} = V_{\text{bond}}$.

The structured bond and its (hedge-)swap are separate products, since they are often offered by separate institutions.⁴

Zero Structures

Besides (structured) coupon paying bonds, another common type of bonds are those for which the coupon is accrued instead of paid. Then, as for the zero coupon bond, there will be a single payment at maturity, see Figure 12.3, right. An accruing product is called a *zero structure*. A bond with accruing coupons is sometimes called a zero coupon bond.

For an accruing zero coupon bond we may define a corresponding swap too. To do so we consider the following (equivalent) representation of the bond: At the end of each coupon period the notional and the period's coupon is paid. The amount defines the new notional for the following period and is reinvested (this corresponds to accruing the coupon), see Figure 12.3, left. The swap is then defined such that it exchanges the structured coupon (on the various notionals) by a corresponding coupon with a given market rate.

The swap will then allow to build up the bonds payment at maturity from the starting notional and investing at market rates, see Figure 12.4.

For the valuation in T_0 we again have

$$N_1 + V_{\text{swap}} = V_{\text{bond}}, \quad (12.1)$$

where the swap is interpreted as payer swap, i.e. it pays $C_i - L_i$. That Equation (12.1) holds follows iteratively by considering a single period: The nominal N_1 is invested at the market rate like a floating rate bond. The floating rate bond pays $N_1 + N_1 \cdot L_1$ in T_2 . The swap exchanges the coupon $N_1 \cdot L_1$ for the structured coupon. Thus we have

³ Common coupons are options on interest rates, e.g. a guaranteed minimum rate in the form of $C_i = \max(L_i(T_i), K)$, or even coupons which depend on the performance of one or more stocks, in this case the bond would be a hybrid interest rate product.

⁴ E.g. mortgage bank and investment bank.

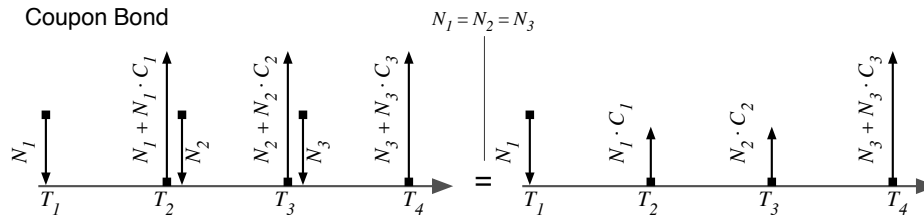


Figure 12.1.: Cashflows for a coupon bond. Left with imaginary exchange of notionals at the end of each period, right with effective cashflows only.

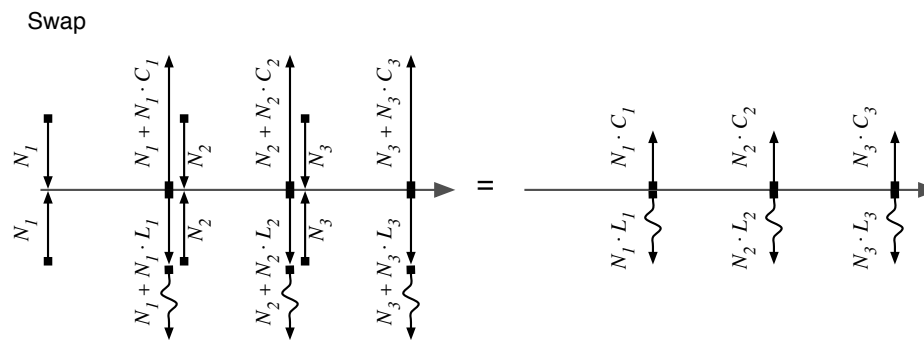


Figure 12.2.: Cashflows for a swap who's fixed leg corresponds to the coupon bond in Figure 12.1. Left with imaginary exchange of notionals at the end of each period, right with effective cashflows only.

$N_1 + N_1 \cdot C_1 =: N_2$ which is the notional N_2 which is used for the same construction for the following period.

Such a construction is especially meaningful, if the zero coupon bond may be canceled at the end of a coupon period. In this case it would pay the accrued notional up to this period. For a cancelable bond the swap has the same cancellation right and is canceled simultaneously.⁵

We repeat the notions *structured bond* and *structured swap* in a definition, although these definitions differ only in one minor aspect from the ones for coupon bond or swap

⁵ Other arguments for this construction are reduction of default and market risk.

respectively: the coupon C_i may be an arbitrary $\mathcal{F}_{T_{i+1}}$ -measurable random variable:

Definition 130 (Structured Bond): ↗

Let $0 = T_0 < T_1 < T_2 < \dots < T_n$ denote a given tenor structure.

For $i = 1, \dots, n-1$ let C_i denote a (generalized) “interest rate” for the periods $[T_i, T_{i+1}]$, respectively. Let C_i be a $\mathcal{F}_{T_{i+1}}$ -measurable random variable.

Furthermore let N_i denote a constant value (*notional*). The *structured bond* pays

$$N_i \cdot C_i \cdot (T_{i+1} - T_i) + N_i \cdot \begin{cases} 0 & (i \neq n) \\ 1 & (i = n-1) \end{cases}$$

in T_{i+1} . The value of the *structured bond* seen in $t < T_2$ is

$$V_{\text{Bond}}(T_1, \dots, T_n; t) = N(t) \cdot \mathbb{E}^{\mathbb{Q}^N} \left(\sum_{k=1}^{n-1} \frac{N_k \cdot C_k \cdot (T_{k+1} - T_k)}{N(T_{k+1})} \middle| \mathcal{F}_t \right) + N_{n-1} \cdot P(T_n; t).$$

Definition 131 (Structured Swap (Structured Receiver Swap)): ↗

Let $0 = T_0 < T_1 < T_2 < \dots < T_n$ denote a given tenor structure.

For $i = 1, \dots, n-1$ let C_i denote a (generalized) “interest rate” for the periods $[T_i, T_{i+1}]$, respectively. Let C_i be a $\mathcal{F}_{T_{i+1}}$ -measurable random variable.

Furthermore let s_i denote a constant interest rate (*spread*) and N_i a constant value (*notional*). The *structured swap* pays

$$X_i := N_i \cdot (C_i - (L(T_i, T_{i+1}; T_i) + s_i)) \cdot (T_{i+1} - T_i).$$

in T_{i+1} . The value of the *structured swap* seen in $t < T_2$ is

$$V_{\text{Swap}}(T_1, \dots, T_n; t) = N(t) \cdot \mathbb{E}^{\mathbb{Q}^N} \left(\sum_{k=1}^{n-1} \frac{X_k}{N(T_{k+1})} \middle| \mathcal{F}_t \right).$$

Remark 132 (Structured Coupon): By C_i we denote an arbitrary, generalized interest rate. In general it will be a function of $L(T_k, T_{k+1})$ with fixing date T_i , and thus even $\mathcal{F}_{T_i}$ -measurable. We allow that the generalized interest rate C_i depends on events within the period $[T_i, T_{i+1}]$ and require $\mathcal{F}_{T_{i+1}}$ -measurability only. An example for C_i is a constant rate $C_i = \text{const.}$, the forward rate $C_i = L(T_i, T_{i+1}; T_i)$, or a [swap rate](#) $C_i = S(T_i, \dots, T_k; T_i)$.

Remark 133 (Zero Structure): The swap in Definition 131 is called *zero structure*, if the notional N_i is given by

$$N_{i+1} := N_i \cdot C_i \cdot (T_{i+1} - T_i) \quad i = 1, \dots, n-1.$$

Remark 134 (Structured Payer Swap / Structured Receiver Swap): The swap defined in Definition 131 is a *receiver swap*. A swap with reversed payments

$$X_i := N_i \cdot ((L(T_i, T_{i+1}; T_i) + s_i) - C_i) \cdot (T_{i+1} - T_i)$$

is called *structured payer swap*. See Definition 107.

Zero Coupon Bond (Accruing Structured Coupon)

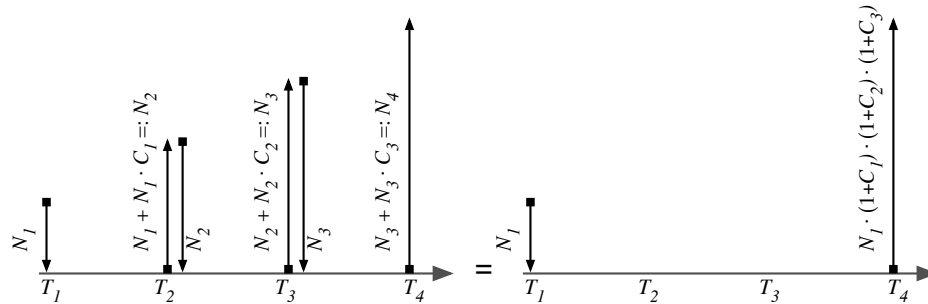


Figure 12.3.: Cashflows for a zero coupon bond. Left with imaginary exchange of notionals at the end of each period, right with effective cashflows only.

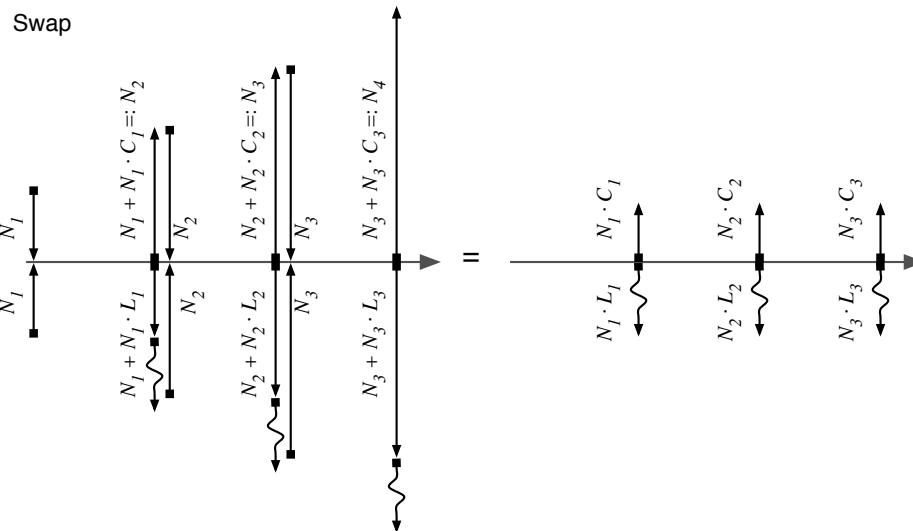


Figure 12.4.: Cashflows for a swap who's fixed leg corresponds to the zero coupon bond in Figure 12.3. Left with imaginary exchange of notionals at the end of each period, right with effective cashflows only.

12.2.2. Bermudan Callables and Cancelable

Definition 135 (Bermudan):

A financial product is called *Bermudan*, if it has multiple exercise dates (options), i.e. there are times T_i at which the holder of a Bermudan may choose between different payments or financial products (underlyings).

An example is given by the Bermudan swaption. Here the option holder has the right to enter a swap at several different times. The optimal exercise strategy chooses the maximal value from either the swap or the Bermudan with the remaining exercise dates.

Definition 136 (Bermudan Swaption):

Let $0 = T_0 < T_1 < T_2 < \dots < T_n$ denote a given tenor structure. The value $V_{\text{BermSwpt}}(T_1, \dots, T_n; T_0)$ of a Bermudan swaption seen at time T_0 is defined recursively by

$$\begin{aligned} V_{\text{BermSwpt}}(T_i, \dots, T_n; T_i) \\ := \max(V_{\text{BermSwpt}}(T_{i+1}, \dots, T_n; T_i), V_{\text{Swap}}(T_i, \dots, T_n; T_i)), \end{aligned} \quad (12.2)$$

where $V_{\text{BermSwpt}}(T_n; T_n) := 0$ and $V_{\text{Swap}}(T_i, \dots, T_n; T_i)$ denotes the value of a swap with *fixing dates* $T_i, \dots, T_{n-1}$ and *payment dates* $T_{i+1}, \dots, T_n$, seen in T_i . Furthermore we have that with a given Numéraire N and a corresponding equivalent martingale measure $\mathbb{Q}^N$

$$V_{\text{BermSwpt}}(T_{i+1}, \dots, T_n; T_i) = N(T_i) \cdot \mathbb{E}^{\mathbb{Q}^N} \left(\frac{V_{\text{BermSwpt}}(T_{i+1}, \dots, T_n; T_{i+1})}{N(T_{i+1})} \middle| \mathcal{F}_{T_i} \right). \quad (12.3)$$



Interpretation: The *Bermudan swaption* $V_{\text{BermSwpt}}(T_{n-1}, T_n)$ is simply a *swaption* (option on a swap) and since the *swap* has only a single period $[T_{n-1}, T_n]$ it even is a *Caplet*. The *Bermudan swaption* $V_{\text{BermSwpt}}(T_{n-2}, T_{n-1}, T_n)$ is an option which allows in time T_{n-2} to choose between a *swaption* (with later exercise date) or a (longer) *swap*. Thus it is an option on an option. Iteratively the Bermudan swaption is an option on an option (on an option, etc.).

Taking the underlying *swap* as the defining object, we see that the *Bermudan swaption* $V_{\text{BermSwpt}}(T_1, \dots, T_n)$ may also be interpreted as an option to enter at times $T_1, \dots, T_{n-1}$ a swap with remaining periods up to T_n , or to wait. Options with multiple exercise times are called *Bermudan*. Options with a single exercise time are called *European*.

Remark 137 (Bermudan Swaption): It is key for the evaluation of the Bermudan swaption that by Equation (12.3) we have at each exercise date T_i an evaluation of a

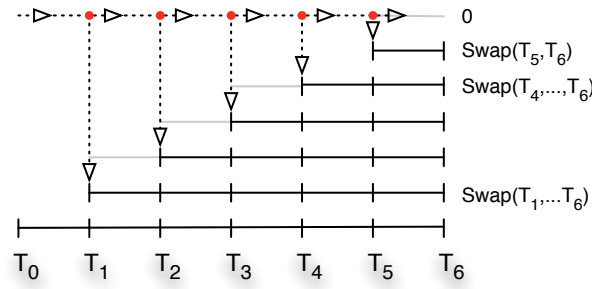


Figure 12.5.: Bermudan Swaption

derivative product. For this the *conditional expectation* has to be calculated. Depending on the model and the implementation, the calculation of conditional expectations may be non-trivial. In Chapter 14 we give an in depth discussion on how to calculate a conditional expectation in a path simulation (Monte-Carlo simulation).

We will now define a product class that generalizes the structure of a *Bermudan swaption*.

Definition 138 (Bermudan Callable Structured Swap⁶):

Let $0 = T_0 < T_1 < T_2 < \dots < T_n$ denote a given tenor structure.

For $i = 1, \dots, n-1$ let C_i denote a (generalized) “interest rate” for the periods $[T_i, T_{i+1}]$, respectively. Let C_i be a $\mathcal{F}_{T_{i+1}}$ -measurable random variable.

Furthermore let s_i denote a constant interest rate (*spread*), N_i a constant value (*notional*) and

$$X_i := N_i \cdot (C_i - (L(T_i, T_{i+1}) + s_i)) \cdot (T_{i+1} - T_i). \quad (12.4)$$

Let $V_{\text{underl}}(T_i, \dots, T_n; T_i)$ denote the value of the product paying X_k in T_{k+1} for $k = i, \dots, n-1$, seen in T_i . If N denotes a Numéraire and $\mathbb{Q}^N$ a corresponding equivalent martingale measure, then

$$V_{\text{underl}}(T_i, \dots, T_n; T_i) = N(T_i) \cdot \mathbb{E}^{\mathbb{Q}^N} \left(\sum_{k=i}^{n-1} \frac{X_k}{N(T_{k+1})} \mid \mathcal{F}_{T_i} \right). \quad (12.5)$$

The value of a *Bermudan callable swap with structured leg* C_i is recursively defined by

$$\begin{aligned} V_{\text{BermCall}}(T_i, \dots, T_n; T_i) \\ := \max(V_{\text{BermCall}}(T_{i+1}, \dots, T_n; T_i), V_{\text{underl}}(T_i, \dots, T_n; T_i)), \end{aligned} \quad (12.6)$$

⁶ Compare [79].

where $V_{\text{BermCall}}(T_n; T_n) := 0$ and

$$V_{\text{BermCall}}(T_{i+1}, \dots, T_n; T_i) = N(T_i) \cdot E^{\mathbb{Q}^N} \left(\frac{V_{\text{BermCall}}(T_{i+1}, \dots, T_n; T_{i+1})}{N(T_{i+1})} \middle| \mathcal{F}_{T_i} \right). \quad (12.7)$$

Remark 139 (Bermudan Callable, structured Leg): For $C_i = S_i = \text{const.}$ the product defined in Definition 138 is a **Bermudan swaption**.

The payments (cash flows) X_i consist of a part $N_i \cdot C_i \cdot (T_{i+1} - T_i)$ which is called the *structured leg* and a part $N_i \cdot (L(T_i, T_{i+1}) + s_i) \cdot (T_{i+1} - T_i)$ which is called the *floating leg*.

The *underlying* $V_{\text{underl}}(T_i, \dots, T_n)$ is a swap swapping the rate C_i against $L(T_i, T_{i+1}) + s_i$, with *fixing dates* $T_i, \dots, T_{n-1}$ and *payment dates* $T_{i+1}, \dots, T_n$.

To evaluate a Bermudan Callable we need to calculate at most two conditional expectations (12.5), (12.7) at each exercise time T_i . Under the assumption of optimal exercise these two values are linked by (12.6), going backward from exercise date to exercise date. In Chapter 14 we give an in depth discussion of the calculation of conditional expectation in a path simulation (Monte-Carlo simulation).

The Bermudan swaption (or a Bermudan callable) allows to enter into a (eventually structured) swap at one time of predefined times T_i . In contrast to this, the *Bermudan cancelable* swap allows to cancel a swap at one time of predefined times T_i .

Definition 140 (Bermudan Cancelable Swap):

With the notation from Definition 138 let

$$V_{\text{coupons}}(T_i; T_i) = N(T_i) \cdot E^{\mathbb{Q}^N} \left(\frac{X_i}{N(T_{i+1})} \middle| \mathcal{F}_{T_i} \right).$$

the value of the coupon payments (cash flows) for the period $[T_i, T_{i+1}]$, seen in T_i .

The value of a *Bermudan Cancelable with structured leg* C_i is recursively defined by

$$\begin{aligned} V_{\text{BermCancel}}(T_i, \dots, T_n; T_i) \\ := \max(V_{\text{coupons}}(T_i; T_i) + V_{\text{BermCancel}}(T_{i+1}, \dots, T_n; T_i), 0), \end{aligned}$$

where $V_{\text{BermCancel}}(T_n; T_n) := 0$ and

$$V_{\text{BermCancel}}(T_{i+1}, \dots, T_n; T_i) = N(T_i) \cdot E^{\mathbb{Q}^N} \left(\frac{V_{\text{BermCancel}}(T_{i+1}, \dots, T_n; T_{i+1})}{N(T_{i+1})} \middle| \mathcal{F}_{T_i} \right).$$

Remark 141 (Bermudan Cancelable): With the notation from Definition 138 and 140 we have

$$V_{\text{BermCancel}}(T_i, \dots, T_n; T_i) = V_{\text{underl}}(T_i, \dots, T_n; T_i) + V_{\text{BermCallPayer}}(T_i, \dots, T_n; T_i),$$

where $V_{\text{BermCallPayer}}$ denotes a Bermudan callable with reversed sign in the underlying. The right to cancel the structured swap V_{underl} corresponds to the right to enter such a structured swap with reversed cashflows.

From this we conclude that the problem to evaluate a *cancelable* corresponds to the problem to evaluate a *callable*. We will thus restrict ourself to the case of a Bermudan callable swap structure. The notions defined in the following exist for Bermudan callables and Bermudan cancelables alike.

12.2.3. Compound Options

A *compound option* is an option on an option. Common are a European call option on a European call option, a call on a put, a put on a call and a put on a put. The compound option is closely related to the Bermudan option with two exercise dates. The compound option may be viewed as a limit case of a Bermudan option.

As for the Bermudan option the evaluation of a compound option requires the evaluation of an option at a future time (the exercise date of the first option). The methods used for pricing of Bermudan options may thus be applied to the evaluation of compound options.

12.2.4. Trigger Products

For a Bermudan callable or Bermudan cancelable the exercise criteria is given by optimal exercise: The option holder chooses the maximum value. Thus the recursive definition of the product value uses the maximum function on the values of the two alternatives non-exercise and exercise. For a trigger product the exercise is given by some criteria, the *trigger*, which does not necessarily represent an optimal exercise. An example for such a product is the auto cap, which we will define in Section 12.2.5.4 below, or the following target redemption note.

12.2.4.1. Target Redemption Note

Definition 142 (Target Redemption Note):

Let $0 = T_0 < T_1 < T_2 < \dots < T_n$ denote a given tenor structure.

For $i = 1, \dots, n - 1$ let C_i denote a (generalized) “interest rate” for the periods $[T_i, T_{i+1}]$, respectively. Let C_i be a $\mathcal{F}_{T_{i+1}}$ -measurable random variable.

Furthermore let N_i denote a constant value (*notional*). A *target redemption note* pays

$$N_i \cdot X_i \quad \text{in} \quad T_{i+1}$$

with

$$X_i := \begin{cases} C_1 & \text{for } i = 1, \\ \min(C_i, K - \sum_{k=1}^{i-1} C_k) & \text{for } i > 1 \end{cases} \quad \text{structured coupon}$$

$$+ \begin{cases} 1 & \text{for } \sum_{k=1}^{i-1} C_k < K \leq \sum_{k=1}^i C_k \text{ or } i = n, \\ 0 & \text{else.} \end{cases} \quad (\text{redemption})$$

$$+ \begin{cases} \max(0, K - \sum_{k=1}^i C_k) & \text{for } i = n \\ 0 & \text{else.} \end{cases} \quad (\text{target coupon guarantee}).$$

┘

Remark 143 (Target Redemption Note Coupon): The rule, that defines the coupon C_i in Definition 142 may vary and Definition 142 is only the framework of the product. With a coupon, e.g., like

$$C_i = K - a \cdot L_i \cdot (T_{i+1} - T_i) \quad \text{for } i > 1. \quad (12.8)$$

the product is also called *variable maturity inverse floater (VMIF)*. We will consider structured coupons in Section 12.2.5.



Interpretation: The holder of a *target redemption note* receives the coupon C_i as long as the sum of the coupons has reached K , the target coupon. If the cumulated coupon would exceed the target coupon ($\sum_{k=1}^i C_k > K$), then the difference to the target coupon and the notional is paid. After this no coupon payments will be made. The structure is canceled if the target coupon has been reached. If the target coupon has not been reached over the full life time of the product, then the difference to the target coupon and the notional is paid at maturity.

Thus, the target redemption note guarantees the payment of a coupon K and the redemption of the notional. What is unsecure is the time of payment and the maturity, and thus the yield of the product. The yield of the product depends on when the condition

$$\sum_{k=1}^i C_k = K \quad \text{and} \quad \sum_{k=1}^{i-1} C_k < K \quad (12.9)$$

is fulfilled.

It is common to have an above average initial coupon C_1 . The life time of the product varies between T_2 and T_n , giving yields between K/T_2 and K/T_n . An investor would buy this product if he assumes that the condition (12.9) will be fulfilled early, such that the yield of the product (is expected to be) above average.

An example of a target redemption note that had been offered in 2004 is

$$C_i = 7.5\% - 2 \cdot L_i \cdot (T_{i+1} - T_i) \quad \text{for } i > 1.$$

Here the option holder profits from the product if the interest rates L_i decline faster than expected (and thus the product will redeem early). $\triangleleft$

12.2.5. Structured Coupons

In the previous definitions we have defined the [structured bond](#) (Definition 130), the [structured swap](#) (Definition 131), the [Bermudan callable](#) (Definition 138), the [Bermudan cancelable](#) (Definition 140) and the [Target Redemption Note](#) (Definition 142) without specifying the coupons C_i . We now refine our definitions by defining some of the most common structured coupons C_i . In the respective definition we only give the notion that describes the coupon. For example, we define a coupon of a *CMS spread* product and the name of the corresponding Bermudan callable swap is *Bermudan callable CMS spread swap*.

12.2.5.1. Capped, Floored, Inverse, Spread, CMS

Definition 144 (Capped Floater): $\triangleright$

A product as in Definition 130, 131, 138 or 140 with

$$C_i := \min(L(T_i, T_{i+1}) + s_i, c_i)$$

is called *Capped Floater*. Here c_i denotes a constant (*cap*). $\triangleleft$

Definition 145 (Floored Floater): $\triangleright$

A product as in Definition 130, 131, 138 or 140 with

$$C_i := \max(L(T_i, T_{i+1}) + s_i, f_i)$$

is called *Floored Floater*. Here f_i denotes a constant (*floor*). $\triangleleft$

Definition 146 (Inverse Floater): $\triangleright$

A product as in Definition 130, 131, 138 or 140 with

$$C_i = \min(\max(k_i - L(T_i, T_{i+1}), f_i), c_i)$$

is called *Inverse Floater*. Here $f_i < c_i$ (*Floor, Cap*) and k_i are constants. $\triangleleft$

Definition 147 (Capped CMS Floater): $\triangleright$

A product as in Definition 130, 131, 138 or 140 with

$$C_i := \min(S_{i,i+m} + s_i, c_i),$$

where s_i, c_i denote constants (*Spread, Cap*) and $S_{i,i+m} = S(T_i, \dots, T_{i+m})$ denotes a swap rate as in Definition 111, is called *Capped CMS⁷ Floater*. The rate $S_{i,i+m}$ is called

⁷ CMS stands for *Constant Maturity Swap*, i.e. the maturity of the underlying swap relative to the swap start ($T_{i+m} - T_i$) is constant. For simplicity we assume here that the tenor structure is equidistant.

constant maturity swap rate (CMS) since the maturity relative to the period start T_i , is constant ($T_{i+m} - T_i$ is constant, i.e. independent of i), assuming an equidistant tenor structure. $\lrcorner$

Definition 148 (Inverse CMS Floater): $\rceil$

A product as in Definition 130, 131, 138 or 140 with

$$C_i = \min(\max(k_i - S_{i,i+m}, f_i), c_i),$$

where k_i, f_i, c_i denote constants (*Strike, Floor, Cap*) and $S_{i,i+m} = S(T_i, \dots, T_{i+m})$ denotes a swap rate as in Definition 111, is called *Inverse CMS Floater*. $\lrcorner$

Definition 149 (CMS Spread): $\rceil$

A product as in Definition 130, 131, 138 or 140 with

$$C_i = \min(\max(S_{i,i+m_1}^1 - S_{i,i+m_2}^2, f_i), c_i),$$

where f_i, c_i denote constants (*Floor, Cap*) and $S_{i,i+m_1} = S(T_i, \dots, T_{i+m_1}), S_{i,i+m_2} = S(T_i, \dots, T_{i+m_2})$ denotes a swap rate as in Definition 111, is called *CMS Spread*. $\lrcorner$

12.2.5.2. Range Accruals

Definition 150 (Range Accrual): $\rceil$

Let $t_{i,k} \in [T_i, T_{i+1})$ denote given observation points for the period $[T_i, T_{i+1})$. Furthermore let K_i and $b_i^l < b_i^h$ denote given interest rates (constants). A product as in Definition 130, 131, 138 or 140 with

$$C_i := K_i \cdot \frac{1}{n_i} \sum_{k=1}^{n_i} \mathbf{1}_{[b_i^l, b_i^h]}(L(t_{i,k}, t_{i,k} + \Delta T_i; t_{i,k})) \quad \text{in } T_{i+1}.$$

is called *Range Accrual*. Here $\Delta T_i := T_{i+1} - T_i$. $\lrcorner$



Interpretation: The interval $[b_i^l, b_i^h]$ describes an interest rate corridor. It is then counted how often the reference rate $L(t, t + \Delta T_i; t)$ stays within this corridor at the times $t_{i,k}$ and the product pay the corresponding fractional amount of the rate K_i at the end of the period.

Since the interest rate corridor may be chosen as a function of the period i , the product allows to profit from a specific evolution of the rate. $\triangleleft$

12.2.5.3. Path Dependent Coupons

Definition 151 (Snowball):

A product as in Definition 130, 131, 138 or 140 with

$$C_i = \min(\max(C_{i-1} + K_i - L_i(T_i), f_i), c_i),$$

is called *Snowball*, where f_i, c_i denote constants (*Floor*, *Cap*) and C_{i-1} denotes the coupon of the previous periods, and $C_0 := 0$.

Remark 152 (Snowball, Path Dependency): The snowball is called a path dependent product, since its coupon depends on the previous coupon, i.e. on the history.

Definition 153 (Ratchet-Cap):

The *Ratchet Cap* pays in T_{i+1}

$$X_i = N \cdot \max(L(T_i, T_{i+1}; T_i) - K_i, 0) \cdot (T_{i+1} - T_i)$$

for times $T_1, \dots, T_n$, where

$$K_i := \min(K_{i-1} + R, L(T_i, T_{i+1}; T_i)) \quad \text{for } i > 1$$

and K_1, R denote constants (*strike* and *Ratchet*), $L(T_i, T_{i+1}; T_i)$ denote the forward rate (LIBOR) and N denotes the *notional*.

Remark 154 (Ratchet Cap): The ratchet cap has an automatic adjustment of the strike K_i . Since the adjustments depends on past realizations, the ratchet cap is a path dependent product.

12.2.5.4. Flexi Cap

The *flexi cap* comes in two variants: As an auto cap with a simple (automatic) exercise criteria and as an chooser cap with an assumed optimal exercise (like for the Bermudan). Both caps have in common that the maximum number of exercises is limited.

Definition 155 (Auto Cap):

Let $n_{\max\text{Ex}} \in \mathbb{N}$. An *Auto Cap* pays in T_{i+1}

If $|\{j : j < i \text{ und } X_j > 0\}| < n_{\max\text{Ex}}$:

$$X_i := N \cdot \max(L(T_i, T_{i+1}; T_i) - K_i, 0) \cdot (T_{i+1} - T_i)$$

and else:

$$X_i := 0$$

for times $T_1, \dots, T_n$, where K_i denotes the *strike rate*, $L(T_i, T_{i+1}; T_i)$ denotes the forward rate (LIBOR) and N denotes the *notional*. The auto cap pays like a normal cap as long as the total number of (positive) payments has not exceeded $n_{\max\text{Ex}}$. $\lrcorner$

Definition 156 (Chooser-Cap): $\rceil$

Let $n_{\max\text{Ex}} \in \mathbb{N}$. A *Chooser Cap* pays in T_{i+1}

If $|\{j : j < i \text{ and } X_j > 0\}| < n_{\max\text{Ex}}$ and the payment is desired:

$$X_i := N \cdot \max(L(T_i, T_{i+1}; T_i) - K_i, 0) \cdot (T_{i+1} - T_i)$$

and else:

$$X_i := 0$$

for times $T_1, \dots, T_n$, where K_i denotes the *strike rate*, $L(T_i, T_{i+1}; T_i)$ denotes the forward rate (LIBOR) and N denotes the *notional*. $\lrcorner$

Remark 157 (Auto Cap): The auto cap is a path dependent products, since at a future time the number of exercises allowed depends on the history. It is also a trigger product, since the exercise is not optimal, but triggered by a simple trigger. The corresponding optimal exercise is given by the chooser cap.

Remark 158 (Chooser Cap (Backward Algorithm)): Since the option holder chooses to exercise optimal, we have that the value $V_{\text{Chooser}}^{(n_{\max\text{Ex}}, T_1, T_n)}(T_0)$ of the chooser cap is given by

$$V_{\text{Chooser}}^{(n_{\text{Ex}}, T_i, T_n)}(T_i) = \max(X_i + V_{\text{Chooser}}^{(n_{\text{Ex}}-1, T_{i+1}, T_n)}(T_i), V_{\text{Chooser}}^{(n_{\text{Ex}}, T_{i+1}, T_n)}(T_i)),$$

where

$$X_i := N \cdot \max(L(T_i, T_{i+1}; T_i) - K_i, 0) \cdot (T_{i+1} - T_i)$$

and

$$V_{\text{Chooser}}^{(n_{\text{Ex}}, T_n, T_n)}(T_n) = 0, \quad V_{\text{Chooser}}^{(0, T_i, T_n)}(T_i) = 0,$$

where $V_{\text{Chooser}}^{(n_{\text{Ex}}, T_i, T_n)}(T_k)$ denotes the value of a chooser cap with at most n_{Ex} exercises in $T_i, \dots, T_n$, seen in T_k .

Remark 159 (Chooser Cap as Bermudan): A chooser cap with $n_{\max\text{Ex}} = 1$ exercises is a *Bermudan cap*.

12.2.6. Shout Options

Definition 160 (Shout Option): ↗

Assume that a financial product pays an Underlying $S(t^*)$ at time T , i.e. it is t^* the fixing date and T the payment date. The owner of the financial product has a *shout right* on the underlying, if the holder of the right may determine once at any time t with $T_1 \leq t \leq T_2 \leq T$ the fixing date t^* as $t^* := t$. The holder determines by *shouting* that the underlying should be fixed. ↘

Remark 161: While for an American option or an Bermudan option, e.g. on $S(t) - K$, the fixing date and the payment date are both determined upon exercise, i.e. $S(t^*) - K$ is paid in t^* , we have that for a Shout option the fixing date is determined upon exercise, while the payment date stays predefined.

Theorem 162 (Value of a Shout Right): A shout right on a convex function of a sub martingale (under terminal measure) is worthless.

Proof: Let T denote the payment date of $f(S(t))$, where t is fixed as $t = t^*$ by shouting. For the chosen Numéraire N we assume $N(T) = 1$ (terminal measure). Let f be convex and S a sub-martingale under $\mathbb{Q}^N$. Then we have:

$$\begin{aligned}
 S(t) &\leq E(S(T_2)|\mathcal{F}_t) && \text{(sub-martingale property)} \\
 f(S(t)) &\leq E(f(S(T_2))|\mathcal{F}_t) && \text{(Jensen's inequality)} \\
 E(f(S(t))|\mathcal{F}_{T_0}) &\leq E(f(S(T_2))|\mathcal{F}_{T_0}) && \text{(tower law)} \\
 E\left(\frac{f(S(t))}{N(T)}|\mathcal{F}_{T_0}\right) &\leq E\left(\frac{f(S(T_2))}{N(T)}|\mathcal{F}_{T_0}\right) && \text{(terminal measure)}
 \end{aligned}$$

Thus, to maximize the value the option holder will exercise her shout right at $t^* = T_2$ always. □

12.3. Product Toolbox

The notions used in the previous section, like *Capped*, *Inverse*, *Ratchet*, etc., describe properties of the payoff function. In practice the notions are used less strict and the name of a product corresponds to its mathematical definition only loosely. Here marketing aspects are more important. However we may still extract key product features from the name of a product.

The following Table 12.3 tries to give a rough mapping of some of the most common notions to properties of the product or payoff function.

Attribute	Product Property
Bond (aka. Note)	Receive coupons. At maturity receive the notional. See Definition 130
Swap	Exchange coupons (usually against <i>float</i>). See Definition 131
Bermudan Callable	Receive an underlying at one of multiple exercise dates. See Definition 138
Bermudan Cancelable	Cancel produkt (e.g. bond or swap) at one of multiple cancellation dates. See Definition 140
Target Redemption	$\min(C_i, K - \sum_{k=1}^{i-1} C_k)$ with cancelation and notional redemption at (12.9) (<i>trigger</i>). See Definition 142
Chooser	Receive an underlying at some of multiple exercise dates. See Definition 156

Attribute	Payoff Function
LIBOR / Floater	$C_i = L(T_i, T_{i+1}; T_i)$
CMS	C_i is swap rate with constant time to maturity
Capped	$\min(C_i - K_i, c_i)$
Floored	$\max(C_i - K_i, f_i)$
Inverse	$K_i - C_i$
Spread	$C_i^1 - C_i^2$
Ratchet	$K_i = \min(K_{i-1} + R, C_i)$
Snowball	$C(T_i) = C(T_{i-1}) + K_i - L_i(T_i)$
Range Accrual	$K_i \cdot \frac{1}{n_i} \sum_{k=1}^{n_i} \mathbf{1}_{[b_i^l, b_i^h]}(L(t_{i,k}, t_{i,k} + \Delta T_i; t_{i,k}))$

Table 12.3.: *Product Toolbox: Common attributes and their representation in product property and payoff function.*



tion.

Experiment: At <http://www.christian-fries.de/finmath/applets/LMMPricing.html> several interest rate products may be priced. Among them also a *cancelable swap*. The model used is a **LIBOR Market Model** implemented in a Monte-Carlo simulation.

<|



as a short discussion of product properties.

The book of Zhang, [41], presents some of the most important exotic options, especially equity options.

<|

Part IV.

Discretization of Stochastic Differential Equations and Numerical Valuation Methods

Motivation and Overview - Part IV

In Chapter 4 we presented the [Black-Scholes Model](#) of a stock S and a risk less account B

$$\begin{aligned} dS(t) &= \mu^{\mathbb{P}}(t)S(t)dt + \sigma(t)S(t)dW^{\mathbb{P}}(t), \\ dB(t) &= r(t)B(t)dt. \end{aligned}$$

In Chapter 10 we presented the corresponding [Black Model](#) for a forward interest rate $L_1 = L(T_1, T_2)$

$$dL_1(t) = \mu^{\mathbb{P}}(t)L_1(t)dt + \sigma(t)L_1(t)dW^{\mathbb{P}}(t), \quad \sigma(t) \geq 0.$$

Under these models we could derive analytic pricing formulas for the corresponding European options. An obvious generalization of these models consists in modeling multiple stocks, e.g.

$$\begin{aligned} dS_i(t) &= \mu_i^{\mathbb{P}}(t)S_i(t)dt + \sigma_i(t)S_i(t)dW_i^{\mathbb{P}}(t), \\ dB(t) &= r(t)B(t)dt, \end{aligned}$$

respectively a model for multiple forward rates, e.g. $L_i = L(T_i, T_{i+1})$ for $T_1 < T_2 < \dots$ with e.g.

$$dL_i(t) = \mu_i^{\mathbb{P}}(t)L_i(t)dt + \sigma_i(t)L_i(t)dW_i^{\mathbb{P}}(t), \quad \sigma(t) \geq 0. \quad (12.10)$$

The model (12.10) is called the [LIBOR Market Model](#).

Such models may then be used for the evaluation of complex derivatives, e.g. a spread option, where the payout depends on two or more forward rates. The pricing, i.e. the calculation of the expectation $E^{\mathbb{Q}^N}(\frac{V(T)}{N(T)}|\mathcal{F}_0)$, where N is a chosen Numéraire and $\mathbb{Q}^N$ is a corresponding martingal measure, often requires a numerical method. E.g. a Monte-Carlo integration

$$E^{\mathbb{Q}^N}(\frac{V(T)}{N(T)}|\mathcal{F}_0) \approx \frac{1}{n} \sum_{i=1}^n \frac{\tilde{V}(T, \omega_i)}{\tilde{N}(T, \omega_i)}$$

for a sampling $\omega_1, \dots, \omega_n$. Here, $\tilde{V}$ and $\tilde{N}$ denote approximations of V and N respectively, since the corresponding Monte-Carlo samples in general are generated for an approximating model, namely a time-discrete model.

In Chapter 13 we start by consider the approximation of time-continuous stochastic processes through time discrete stochastic processes. We then consider the approximation of the random variables by a Monte-Carlo simulation or a discretization of the state space.

These discretizations allow to calculate (approximate) expectations and thus derivative prices. It turns out that within the discrete set up some calculations are difficult. In a Monte-Carlo simulation the calculation of a conditional expectation is non-trivial. A conditional expectation may be required in the pricing of Bermudan products, see Chapter 14. In a discretization of a state space the calculation of path dependent quantities is non-trivial. Path dependent quantities appear in path dependent options, see Chapter ??.

The treatment of complex models, like the [LIBOR Market Model](#), will be given in Part V.

CHAPTER 13

Discretization of time and state space

In this chapter we consider methods for discretization and implementation of Itô stochastic processes. We give an integrated presentation of “path simulation” (Monte-Carlo simulation) and “lattice methods” (e.g. trees). Finally, both methods may be combined, see Section 13.4.

Throughout our discussion of the discretization and implementation we will repeat some of the notions from Chapter 2, e.g. path, σ -algebra, filtration, process and $\mathcal{F}_t$ -adapted. Thus, this chapter may also serve as an illustration of some of the mathematical concepts from Chapter 2.

The discretization and implementation should not be seen as a minor additional step after the mathematical analysis and it should not be underestimated. The discretization and implementation may give a second view and thus insights on a model.¹

In Figure 13.1 we give an overview of the steps involved in the discretization and implementation of Itô processes.

13.1. Discretization of Time: The *Euler* and the *Milstein Scheme*

As a first step we consider the discretization of time and present the Euler scheme and Milstein scheme.

13.1.1. Definitions

Definition 163 (Euler Scheme):

Given is a Itô process

$$dX(t) = \mu(t, X(t))dt + \sigma(t, X(t))dW(t),$$

¹ Indeed, it is common in mathematics to prove analytic results as a limit of a numerical, i.e. discrete, procedure.

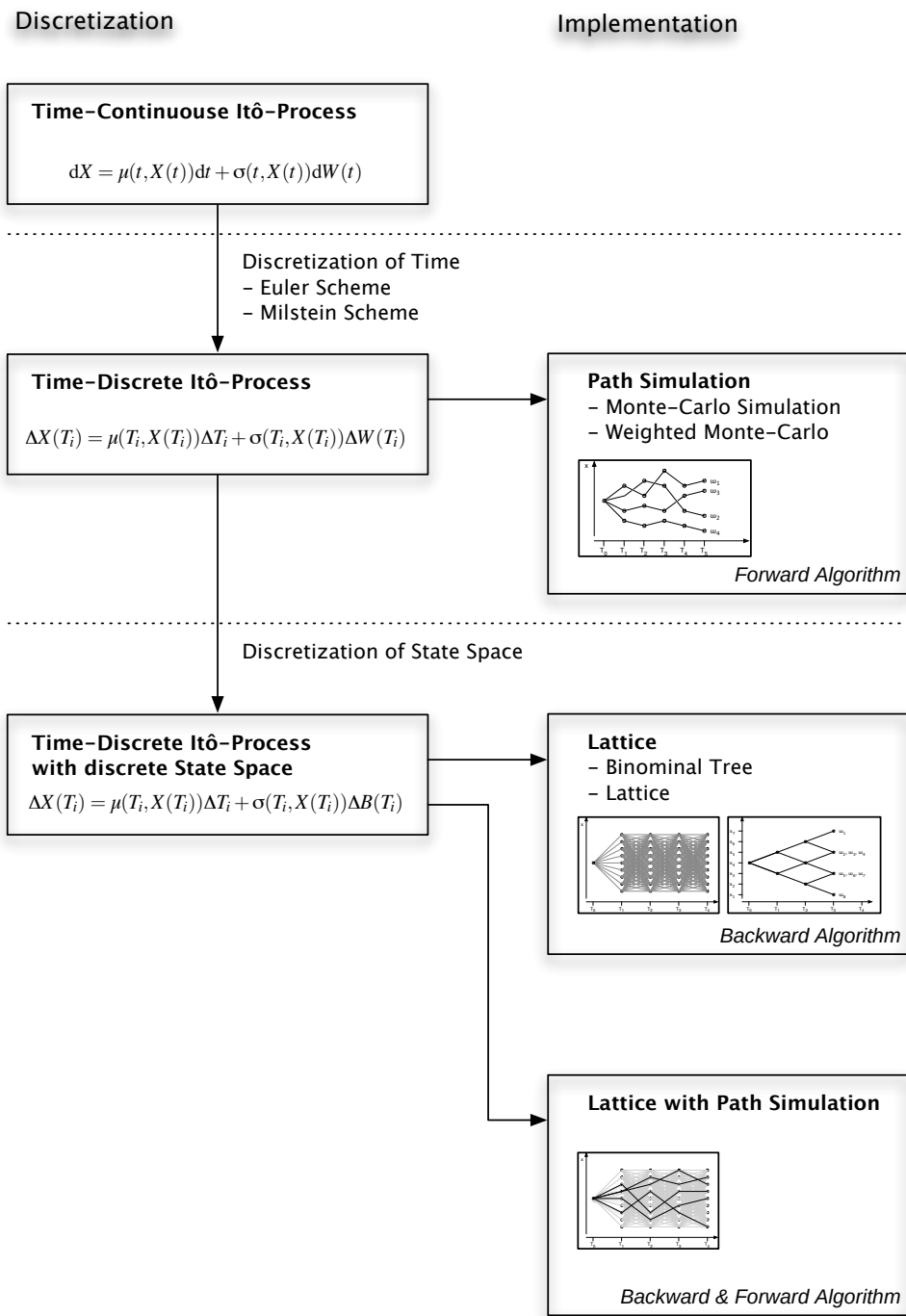


Figure 13.1.: *Discretization and Implementation of Itô Processes*

and a time discretization $\{t_i \mid i = 0, \dots, n\}$ with $0 = t_0 < \dots < t_n$. Then the time-discrete stochastic process $\tilde{X}$ defined by

$$\tilde{X}(t_{i+1}) = \tilde{X}(t_i) + \mu(t_i, \tilde{X}(t_i))\Delta t_i + \sigma(t_i, \tilde{X}(t_i))\Delta W(t_i)$$

Euler scheme of the process X (where $\Delta t_i := t_{i+1} - t_i$ and $\Delta W(t_i) := W(t_{i+1}) - W(t_i)$). $\lrcorner$



Interpretation (Euler discretization): The Euler scheme derives from a simple integration rule. From the definition of the Itô process we have

$$X(t_{i+1}) = X(t_i) + \int_{t_i}^{t_{i+1}} \mu(t, X(t))dt + \int_{t_i}^{t_{i+1}} \sigma(t, X(t))dW(t).$$

Obviously, the Euler scheme is given by the approximation of the integrals

$$\begin{aligned} \int_{t_i}^{t_{i+1}} \mu(t, X(t))dt &\approx \int_{t_i}^{t_{i+1}} \mu(t_i, X(t_i))dt = \mu(t_i, X(t_i))\Delta t_i \\ \int_{t_i}^{t_{i+1}} \sigma(t, X(t))dW(t) &\approx \int_{t_i}^{t_{i+1}} \sigma(t_i, X(t_i))dW(t) = \sigma(t_i, X(t_i))\Delta W(t_i). \end{aligned}$$

$\triangleleft$

The following Milstein scheme improves the approximation of the stochastic integral $\int dW$.

Definition 164 (Milstein Scheme): $\rceil$

Given is a Itô process

$$dX(t) = \mu(t, X(t))dt + \sigma(t, X(t))dW(t),$$

and a time discretization $\{t_i \mid i = 0, \dots, n\}$ with $0 = t_0 < \dots < t_n$. Then the time-discrete stochastic process $\tilde{X}$ defined by

$$\begin{aligned} \tilde{X}(t_{i+1}) &= \tilde{X}(t_i) + \mu(t_i, \tilde{X}(t_i))\Delta t_i + \sigma(t_i, \tilde{X}(t_i))\Delta W(t_i) \\ &\quad + \frac{1}{2}\sigma(t_i, \tilde{X}(t_i))\sigma'(t_i, \tilde{X}(t_i))(\Delta W(t_i)^2 - \Delta t_i) \end{aligned}$$

Milstein Scheme of the process X (where $\Delta t_i := t_{i+1} - t_i$ and $\Delta W(t_i) := W(t_{i+1}) - W(t_i)$ and $\sigma' := \frac{\partial}{\partial X}\sigma$). $\lrcorner$

Remark 165 (Milstein Scheme): The Milstein scheme gives an “improvement” only if σ depends on X .

We give another discretization scheme:

Definition 166 (Euler Scheme with Predictor-Corrector Step): ↗

Given is a Itô process

$$dX(t) = \mu(t, X(t))dt + \sigma(t, X(t))dW(t),$$

and a time discretization $\{t_i \mid i = 0, \dots, n\}$ with $0 = t_0 < \dots < t_n$. Then the time-discrete stochastic process $\tilde{X}$ defined by

$$\tilde{X}(t_{i+1}) = \tilde{X}(t_i) + \frac{1}{2}(\mu(t_i, \tilde{X}(t_i)) + \mu(t_{i+1}, \tilde{X}^*(t_{i+1})))\Delta t_i + \sigma(t_i, \tilde{X}(t_i))\Delta W(t_i)$$

with

$$\tilde{X}^*(t_{i+1}) = \tilde{X}(t_i) + \mu(t_i, \tilde{X}(t_i))\Delta t_i + \sigma(t_i, \tilde{X}(t_i))\Delta W(t_i)$$

Euler Scheme with Predictor-Corrector Step of the process X (where $\Delta t_i := t_{i+1} - t_i$ and $\Delta W(t_i) := W(t_{i+1}) - W(t_i)$). ↘



Interpretation (Predictor-corrector scheme): The predictor-corrector scheme improves the integration of the drift term $\int dt$, not of the stochastic integral $\int dW$. Instead of approximating the integral $\int_{t_i}^{t_{i+1}} \mu(t, X(t))dt$ by a rectangular rule $\mu(t_i, X(t_i)) \cdot \Delta t_i$ the method aims to use a trapezoidal rule. With a trapezoidal rule the integral $\int_{t_i}^{t_{i+1}} \mu(t, X(t))dt$ would be approximated as $\frac{1}{2}(\mu(t_i, X(t_i)) + \mu(t_{i+1}, X(t_{i+1}))) \cdot \Delta t_i$. Since the realization $X(t_{i+1})$ and thus $\mu(t_{i+1}, X(t_{i+1}))$ is unknown, it is approximated by an Euler step $\tilde{X}^*(t_{i+1})$ (*predictor step*) and the trapezoidal rule is applied with this approximation. This corresponds to correcting $\tilde{X}^*(t_{i+1})$ (*corrector step*). We have:

$$\begin{aligned} \tilde{X}(t_{i+1}) &= \tilde{X}^*(t_{i+1}) - \mu(t_i, \tilde{X}(t_i)) \cdot \Delta t_i + \frac{1}{2}(\mu(t_i, \tilde{X}(t_i)) + \mu(t_{i+1}, \tilde{X}^*(t_{i+1}))) \cdot \Delta t_i \\ &= \tilde{X}^*(t_{i+1}) + \underbrace{\frac{1}{2}(\mu(t_{i+1}, \tilde{X}^*(t_{i+1})) - \mu(t_i, \tilde{X}(t_i))) \cdot \Delta t_i}_{\text{correction term}}. \end{aligned} \quad (13.1)$$

<|



Tip (Implementation of the predictor-corrector scheme): Note that for an implementation the Formula (13.1) is more efficient than the two Euler steps in the original Definition (166) of the scheme. The second Euler step is replaced by a correction term applied to $\tilde{X}^*(t_{i+1})$ and requiring only the additional calculation of $\mu(t_{i+1}, \tilde{X}^*(t_{i+1}))$. <|

The schemes presented give a time discrete stochastic process $\tilde{X}$ such that $\tilde{X}(t_i)$ is an approximation of $X(t_i)$. An in depth discussion of numerical methods to approximate stochastic processes may be found in [19].

13.1.2. Time-Discretisation of a Lognormal Process

We consider the process

$$dX = \mu(t, X(t))X(t)dt + \sigma(t)X(t)dW(t), \quad (13.2)$$

where $(t, x) \mapsto \mu(t, x)$ and $t \mapsto \sigma(t)$ are given deterministic functions. With Lemma 44 we have

$$d \log(X) = (\mu(t, X(t)) - \frac{1}{2}\sigma^2(t))dt + \sigma(t)dW(t). \quad (13.3)$$

in the following we discuss several possible time-discretizations of the process X . The discussion is of special importance since the [Black-Scholes model](#), the [Black model](#) and the [LIBOR Market Model](#) are all of the form (13.2).

13.1.2.1. Discretization via Euler scheme

The Euler scheme for the stochastic differential equation (13.2) is given by

$$\tilde{X}(t_{i+1}) = \tilde{X}(t_i) + \mu(t_i, \tilde{X}(t_i))\tilde{X}(t_i)\Delta t_i + \sigma(t_i)\tilde{X}(t_i)\Delta W(t_i).$$

The random variables $\tilde{X}(t)$ generated by this scheme differ from the random variables $X(t)$ of the time-continuous process by a discretization error $X(t) - \tilde{X}(t)$. This discretization error might be relatively large. Take for example the even simpler case of a vanishing drift $\mu = 0$. Then $\tilde{X}(t_1)$ normal distributed, while $X(t_1)$ is lognormal distributed. Note that $\tilde{X}$ may attain negative values, while X may not (this follows from (13.3)).

13.1.2.2. Discretization via Milstein scheme

A possibility to reduce the discretization error is to use the Milstein scheme (Definition 164):

$$\begin{aligned} \tilde{X}(t_{i+1}) = \tilde{X}(t_i) &+ (\mu(t_i, \tilde{X}(t_i)) - \frac{1}{2}\sigma(t_i)^2)\tilde{X}(t_i)\Delta t_i \\ &+ \sigma(t_i)\tilde{X}(t_i)\Delta W(t_i) + \frac{1}{2}\sigma(t_i)^2\tilde{X}(t_i)\Delta W(t_i)^2. \end{aligned}$$

13.1.2.3. Discretization of the Log-Process

A much better discretization than the two previous schemes is given by the Euler discretization of the Itô process of $\log(X)$. The Euler scheme of (13.3) is given by

$$\log(\tilde{X}(t_{i+1})) = \log(\tilde{X}(t_i)) + (\mu(t_i, \tilde{X}(t_i)) - \frac{1}{2}\sigma(t_i)^2)\Delta t_i + \sigma(t_i)\Delta W(t_i), \quad (13.4)$$

and applying the exponential we have

$$\tilde{X}(t_{i+1}) = \tilde{X}(t_i) \cdot \exp((\mu(t_i, \tilde{X}(t_i)) - \frac{1}{2}\sigma(t_i)^2)\Delta t_i + \sigma(t_i)\Delta W(t_i)).$$

Using this scheme will give a lognormal random variable $\tilde{X}(t_1)$.

13.1.2.4. Exact Discretization

For the special case where μ does not depend on X , e.g. if X is a relative price under the corresponding martingale measure and thus even drift free, then we may take the exact solution as a discretization scheme. We then have:

$$X(t_{i+1}) = X(t_i) \cdot \exp\left((\mu_i - \frac{1}{2}\sigma_i^2)\Delta t_i + \sigma_i\Delta W(t_i)\right), \quad (13.5)$$

where

$$\mu_i := \frac{1}{\Delta t_i} \int_{t_i}^{t_{i+1}} \mu(\tau) d\tau, \quad \sigma_i := \sqrt{\frac{1}{\Delta t_i} \int_{t_i}^{t_{i+1}} \sigma^2(\tau) d\tau}.$$

13.2. Discretization of Paths (Monte-Carlo Simulation)

We consider the time-discrete stochastic process

$$X(t_{i+1}) = X(t_i) + \mu(t_i, X(t_i))\Delta t_i + \sigma(t_i, X(t_i))\Delta W(t_i). \quad (13.6)$$

We consider here an Euler scheme. The considerations below carry over to any other discretization scheme. Furthermore we do not apply a tilde to the process X since we will only consider the time-discrete process, thus do not have to distinguish it from the original time-continuous process.

13.2.1. Monte-Carlo Simulation

The random variables $\Delta W(t_i)$ of the respective time steps are mutually independent, see Definition 23. In every time step t_i a random number is drawn according to the distribution of $\Delta W(t_i)$, (i.e. a vector of random numbers if $\Delta W(t_i)$ is vector valued), which we denote by $\Delta W(t_i, \omega_j)$. Then

$$X(t_{i+1}, \omega_j) = X(t_i, \omega_j) + \mu(t_i, X(t_i, \omega_j))\Delta t_i + \sigma(t_i, X(t_i, \omega_j))\Delta W(t_i, \omega_j)$$

determines the process X on a path, which we denote by ω_j . Here $\Delta W(t_i, \omega_j)$ and $\Delta W(t_k, \omega_j)$ ($i \neq k$) are independent random numbers, following the definition of the Brownian motion. If we follow this rule to generate paths $\omega_1, \dots, \omega_{n_{\text{paths}}}$, where $\Delta W(t_i, \omega_j)$ and $\Delta W(t_i, \omega_k)$ ($j \neq k$) are independent, then we say that the set

$$\{X(t_i, \omega_k) \mid i = 0, 1, \dots, n_{\text{times}}; k = 0, 1, \dots, n_{\text{paths}}\}$$

is a *Monte-Carlo simulation* of the process X . The generation of random numbers is discussed in Section A.3.

An approximation of the expectation of some function f of the $X(t_i)$'s is then given by

$$\mathbb{E}^{\mathbb{P}}(f(X(t_0), \dots, X(t_{n_{\text{times}}})) \mid \mathcal{F}_{t_0}) \approx \sum_{j=1}^{n_{\text{paths}}} f(X(t_0, \omega_j), \dots, X(t_{n_{\text{times}}, \omega_j})) \cdot \frac{1}{n_{\text{paths}}}.$$

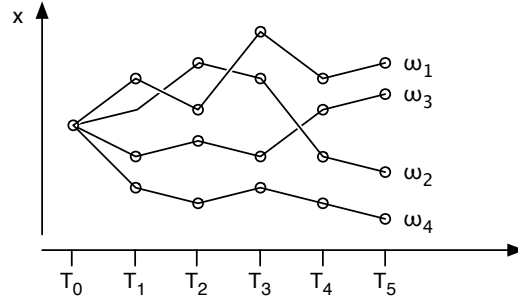


Figure 13.2.: Monte Carlo Simulation

13.2.2. Weighted Monte-Carlo Simulation

A generalization of the procedure is to generate the random numbers $\Delta W(t_i, \omega_j)$ not according to the distribution $\Delta W(t_i)$, which means that all paths ω_j are generated with non-uniform weights p_j ($\sum_{j=1}^{n_{\text{paths}}} p_j = 1$). In this case we call the simulation *weighted Monte-Carlo simulation*. For the expectation we have

$$\mathbb{E}^{\mathbb{P}}(f(X(t_0), \dots, X(t_{n_{\text{times}}})) \mid \mathcal{F}_{t_0}) \approx \sum_{j=1}^{n_{\text{paths}}} p_j f(X(t_0, \omega_j), \dots, X(t_{n_{\text{times}}, \omega_j})).$$

To summarize, the Monte-Carlo simulation consists of the time-discrete process X in (13.6), represented over a discrete probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$, where

$$\tilde{\Omega} = \{\omega_1, \dots, \omega_{n_{\text{paths}}}\} \subset \Omega, \quad \tilde{\mathcal{F}} = \sigma(\{\omega_j \mid j = 1, \dots, n_{\text{paths}}\}), \quad \tilde{P}(\{\omega_j\}) = p_j.$$

13.2.3. Implementation

We give an example for an object oriented design. The Figures 13.3, 13.4 follow the *Unified Modeling Language* (UML) 1.3, see [26].

The generation of a Monte-Carlo simulation of a lognormal process is realized through the abstract base class `LOGNORMALPROCESS`. The class defines abstract methods for initial conditions, drift and volatility. A specific model has to derive from this class and implement the three methods. The abstract base class `LOGNORMALPROCESS` gives the implementation of the discretization scheme, using the methods for initial conditions, drift and volatility.

The generation of the Brownian increments, i.e. the random numbers, is given by an additional class `BROWNIANMOTION`.

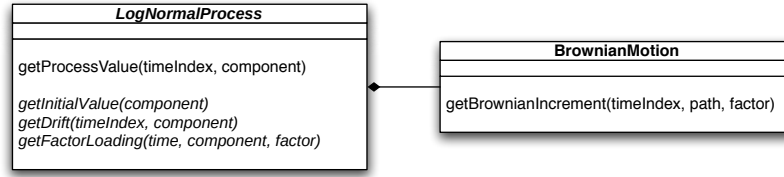


Figure 13.3.: UML Diagram: Monte Carlo Simulation / Lognormal Process.

13.2.3.1. Example: Valuation of a Stock Option under the Black-Scholes Model using Monte-Carlo Simulation

We consider the model from Section 4, the Black-Scholes Modell: We have to simulate the process

$$dS(t) = r(t)S(t)dt + \sigma(t)S(t)dW^{\mathbb{P}}(t) \quad \text{under the measure } \mathbb{Q}^N \quad S(0) = S_0$$

together with the Numéraire

$$dN(t) = r(t)N(t)dt \quad N(0) = 1,$$

which is not stochastic here. For this example we have to set $X = (X_1, X_2) = (S, N)$ in the previous section. We choose r and σ constant and apply the Euler scheme to $\log(S)$, following Section 13.1.2.3:

$$\begin{aligned} S(t_{i+1}) &= S(t_i) \cdot \exp\left((r - \frac{1}{2}\sigma^2)\Delta t_i + \sigma\Delta W(t_i)\right) & S(0) &= S_0 \\ N(t_{i+1}) &= N(t_i) \cdot \exp(r\Delta t_i). & N(0) &= 1. \end{aligned}$$

In this example the time discretization does not introduce an approximation error, we are in the special situation of Section 13.1.2.4. If $\omega_1, \dots, \omega_{n_{\text{paths}}}$ are paths of a Monte-Carlo simulation, then we have for price V of an European option with maturity t_k and strike K

$$V(0) \approx N(0) \frac{1}{n_{\text{paths}}} \sum_{j=1}^{n_{\text{paths}}} \frac{\max(S(t_k, \omega_j) - K, 0)}{N(t_k)}.$$

We extend the object oriented design from Figure 13.3 and derive the class `BLACKSCHOLESMODEL` from the abstract base class `LOGNORMALPROCESS`. The class `BLACKSCHOLESMODEL` implements the methods for the initial value (returning $S(0)$), the drift (returning r) and the factor loading (returning σ). This context the factor loading is identical to the volatility.² In addition the class implements a method that returns the corresponding Numéraire.

² In a multi-factor model the factor loading are given by the square root of the covariance matrix.

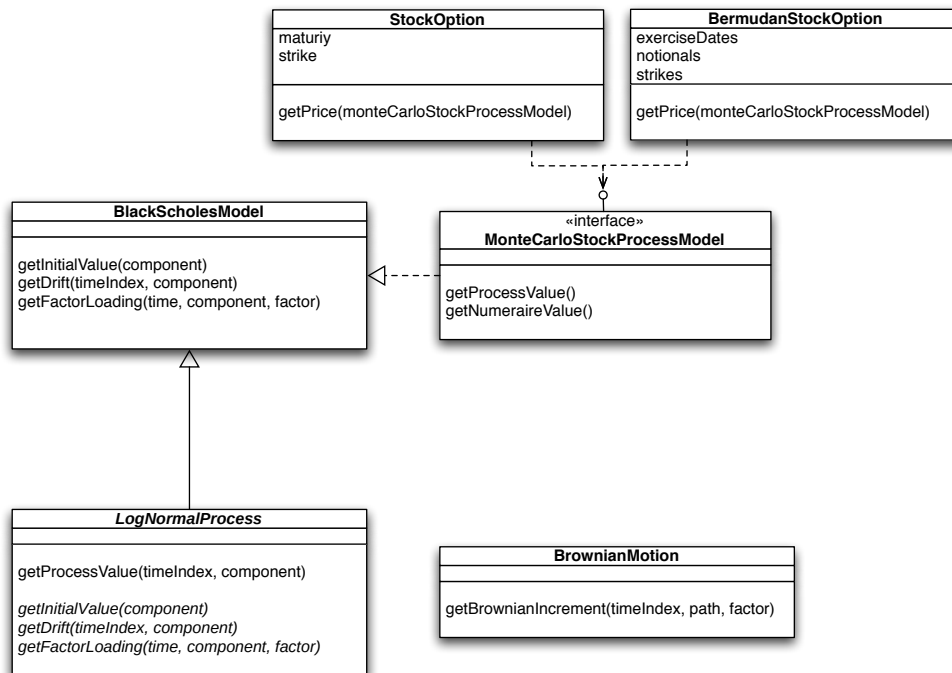


Figure 13.4.: UML Diagram: Evaluation under a Black-Scholes Model via Monte Carlo Simulation. This class hierarchy is realized in the Java™ code.

13.2.3.2. Separation of Product and Model

The evaluation of a derivative product, in our case a simple European option, is realized in its own class `StockOption`. This class communicates not directly with the `BlackScholesModel`. Instead it expects an *interface* `MonteCarloStockProcessModel` and the model implement this *interface*.

The *interface* `MonteCarloStockProcessModel` defines, that the stock model offers the stock process and the Numéraire as a Monte Carlo simulation to the stock product. All corresponding Monte-Carlo evaluation of stock products expect this interface only. All corresponding Monte-Carlo stock models implement this interface. This realizes a separation of product and model. The model used to evaluate the products may be exchanged, if the interface is respected.

We will reuse this principle in the object oriented design of the LIBOR Market Model, a multi-dimensional interest rate model. There we will reuse the classes `BrownianMotion` and `LogNormalProcess`, see Section 17.2.

13.2.4. Review

Through Monte-Carlo simulation we may evaluate simple and pure path dependent products. Non-trivial is the Monte-Carlo evaluation of derivatives, where an expectation has to be calculated, that is conditional to a future time t_i , e.g.

$$E(X(t_j) | \mathcal{F}_{t_i})$$

for $t_j > t_i > t_0$.³ Why this is non-trivial becomes apparent when we consider the filtration:

In a path simulation in general no two path will have a common past. The reason is that the number of possible states $X(t_i)$ is much higher than the number of paths that are simulated. If no two paths have a common future, then we have the following filtration

$$\begin{aligned}\tilde{\mathcal{F}}_{t_i} &= \tilde{\mathcal{F}} \quad \text{for all } t_i > t_0 \text{ and} \\ \tilde{\mathcal{F}}_{t_0} &= \tilde{\Omega}.\end{aligned}$$

A calculation of a condition expectation in a Monte-Carlo simulation is not straight forward, since no suitable time discretization of the filtration $\tilde{\mathcal{F}}_{t_i}$ of X is given.⁴ The filtration of the Monte-Carlo simulation depicted in Figure 13.2 is given by

$$\begin{aligned}\tilde{\mathcal{F}}_{t_0} &= \{\emptyset, \tilde{\Omega}\} \\ \tilde{\mathcal{F}}_{t_1} &= \sigma(\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}\}) \\ \tilde{\mathcal{F}}_{t_2} &= \sigma(\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}\}) \\ \tilde{\mathcal{F}}_{t_3} &= \sigma(\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}\}).\end{aligned}$$

³ Such a conditional expectation would be necessary for the evaluation of a Bermudan option.

⁴ In Section 14 we will present special methods for the evaluation of conditional expectation in a Monte-Carlo simulation.

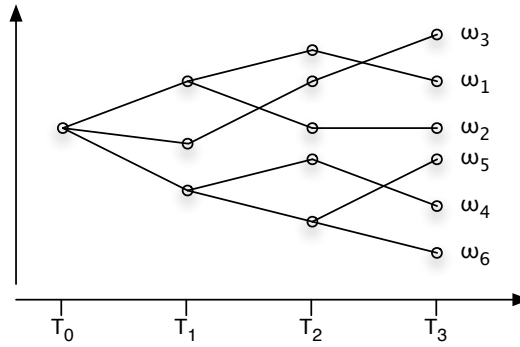


Figure 13.5.: Monte Carlo Simulation

For the path simulation depicted in Figure 13.5 the filtrations are not all the same. The filtration is given by

$$\begin{aligned}\tilde{\mathcal{F}}_{t_0} &= \{\emptyset, \Omega\} \\ \tilde{\mathcal{F}}_{t_1} &= \sigma(\{\{\omega_1, \omega_2\}, \{\omega_3\}, \{\omega_4, \omega_5, \omega_6\}\}) \\ \tilde{\mathcal{F}}_{t_2} &= \sigma(\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}, \{\omega_5, \omega_6\}\}) \\ \tilde{\mathcal{F}}_{t_3} &= \sigma(\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}, \{\omega_5\}, \{\omega_6\}\}).\end{aligned}$$

To achieve a time-discretization of the filtration, we restrict the possible values of X in each simulation time step. We thus assume a discretization of state space.

13.3. Discretization of State Space

13.3.1. Definitions

Instead of the time-discrete process

$$X(t_{i+1}) = X(t_i) + \mu(t_i, X(t_i))\Delta t_i + \sigma(t_i, X(t_i))\Delta W(t_i)$$

we consider the time-discrete process

$$X(t_{i+1}) = X(t_i) + \mu(t_i, X(t_i))\Delta t_i + \sigma(t_i, X(t_i))\Delta B(t_i), \quad (13.7)$$

where the increments $\Delta B(t_i)$ are random variable that take only a finite number of values and are mutually independent. According to Theorem 25 we may choose the

$\Delta B(t_i)$ such that in the limit $\Delta t_i \rightarrow 0$ we recover a brownian motion, i.e. for X we recover the original time-continuous process.

Using the increments $\Delta B(t_i)$ the process X from (13.7) may take only a finite number of values too

$$X(t_i) \in \{x_{t_i}^j \mid j = 0, \dots, n_{t_i} ; i = 0, \dots, n_{\text{times}}\}.$$

we denote this set as *lattice*. Let

$$\{p_i^{j_1, j_2} \mid j_1 = 0, \dots, n_i ; j_2 = 0, \dots, n_{i+1} ; i = 0, \dots, n_{\text{times}} - 1\},$$

denote the transition probabilities of the $\Delta B(t_i)$'s i.e. $p_i^{j_1, j_2} := P(X(t_{i+1}) = x_{t_{i+1}}^{j_2} \mid X(t_i) = x_{t_i}^{j_1})$. Depending on the probability distribution assumed for the $\Delta B(t_i)$ we have that the matrix of transitional probabilities $(p_i^{j_1, j_2})_{j_1, j_2}$ from t_i to t_{i+1} is sparse. For a binomial tree, i.e. binomial distributed $\Delta B(t_i)$'s, we have that in each row only two entries are non zero.

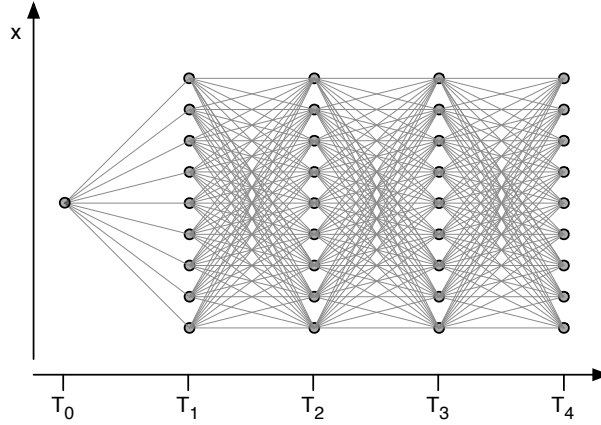


Figure 13.6.: Lattice

In the given situation of a discretized state space it is far more efficient to store transitions probabilities than to simulate all possible paths. For example the lattice in Figure 13.6 exhibits 9 states at each of the 4 time steps. In this lattice there exist $9^4 = 6561$ paths, however there are only $9 + 3 \cdot 9^2 = 252$ transition probabilities. The amount of memory needed to store all paths grows exponentially with time. The amount of memory needed to store all states and transition probabilities grows only linear in time.

If we use binomial distributed increments $\Delta B(t_i)$, then we obtain a binomial tree. Figure 13.8 shows the paths that may be distinguished in the binomial tree depicted in Figure 13.7. For the binomial tree the filtration is no longer trivial. The exhibit a

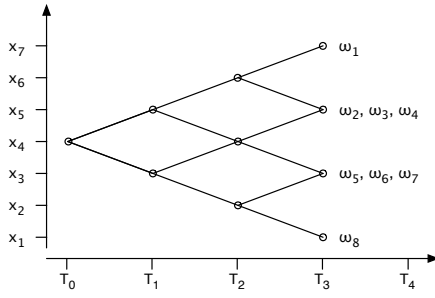


Figure 13.7.: Binomial Tree

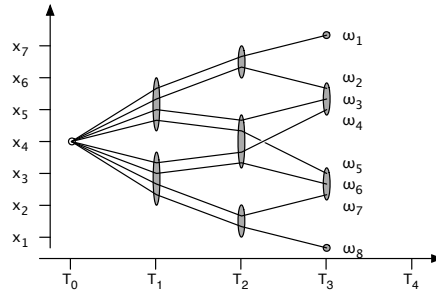


Figure 13.8.: Pfade des Binomial Tree

hierarchy of refinements. For the binomial tree in Figure 13.7 the filtration is given by

$$\begin{aligned}\mathcal{F}_{t_0} &= \{\emptyset, \Omega\} \\ \mathcal{F}_{t_1} &= \sigma(\{\{\omega_1, \omega_2, \omega_3, \omega_5\}, \{\omega_4, \omega_6, \omega_7, \omega_8\}\}) \\ \mathcal{F}_{t_2} &= \sigma(\{\{\omega_1, \omega_2\}, \{\omega_3, \omega_5\}, \{\omega_4, \omega_6\}, \{\omega_7, \omega_8\}\}) \\ \mathcal{F}_{t_3} &= \sigma(\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}, \{\omega_5\}, \{\omega_6\}, \{\omega_7\}, \{\omega_8\}\}),\end{aligned}$$

i.e. in addition we have a suitable time-discretization of the filtration. This allows for a calculation of conditional expectations.

13.3.2. Backward-Algorithm

Given a *lattice* we may calculate the conditional expectation of a function f of $X(t_{i+1})$, condition to $X(t_i) = x_{t_i}^{j_1}$ as

$$\mathbb{E}^{\mathbb{P}}(f(X(t_{i+1})) \mid \{X(t_i) = x_{t_i}^{j_1}\}) \approx \sum_{j_2=1}^{n_{t_{i+1}}} p_i^{j_1, j_2} \cdot f(x_{t_{i+1}}^{j_2}).$$

A step by step application finally gives the expectation $\mathbb{E}^{\mathbb{P}}(f(X(t_{i+1})) \mid \{X(t_0) = x_{t_0}^1\})$. Assuming that $X(t_0) = \text{const.} = x_{t_0}^1$, i.e. that the lattice has a single state in t_0 , this corresponds to

$$\mathbb{E}^{\mathbb{P}}(f(X(t_{i+1})) \mid \mathcal{F}(t_0)).$$

This procedure is called *Backward Algorithm*. If we consider the case of a Numéraire N and a derivative product V zur Zeit t_{i+1} given as a function of the states $X(t_{i+1})$, then we find

$$\frac{V(t_i, x_{t_i}^{j_1})}{N(t_i, x_{t_i}^{j_1})} \approx \sum_{j_2=1}^{n_{t_{i+1}}} p_i^{j_1, j_2} \cdot \frac{V(t_{i+1}, x_{t_{i+1}}^{j_2})}{N(t_{i+1}, x_{t_{i+1}}^{j_2})}. \quad (13.8)$$

Thus, financial products which are functions of the states $X(t_i)$, are evaluated in a lattice by storing the Numéraire-relative prices at the nodes $x_{t_{i+1}}^{j_2}$ and calculating the N -relative value in node $x_{t_i}^{j_1}$ via (13.8). The transition (13.8) is also called *rollback*.

13.3.3. Review

13.3.3.1. Path Dependencies

If a lattice with states $x_{t_i}^j$ and transition probabilities $p_i^{j_1, j_2}$ is set up, we are able to calculate (certain) conditional expectations, however it is non-trivial to calculate path dependents products, i.e. financial products that not only depend on the current state of the underlying, but also on the history (the paths). The *backward algorithm* carries information from the future backward in time and cannot consider information from the past. In contrast the Monte-Carlo simulation is a *forward algorithm* that carries information forward in time.

13.3.3.2. Course of Dimension

A further problem of lattices becomes apparent for vector values processes X . The amount of possible transitions (and thus the amount of transition probabilities to store) grows exponentially with the dimension of the vector X and already for dimensions like 3 or higher the numerical calculation of the rollback (13.8) is critical with respect to required resources cpu time and memory.

13.4. Path Simulation through a Lattice: Two Layers

To calculate path dependent product in a lattice we may create a Monte-Carlo simulations according to the state-discretized process (13.7). This may be depicted as a “Monte-Carlo simulation through a lattice” or a second Monte-Carlo layer laid over the lattice. See Figure 13.9.

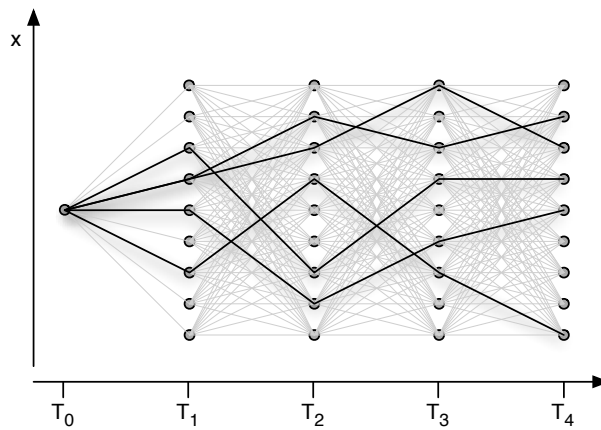


Figure 13.9.: Lattice with overlaid Monte Carlo Simulation



Further Reading: An in depth discussion of numerical methods to approximate stochastic processes may be found in [19]. <|

CHAPTER 14

Pricing Options with Early Exercise by Monte Carlo Simulation

14.1. Introduction

Let us first consider the simple case of a Bermudan option $V_{\text{berm}}(T_1, T_2)$ with two exercise dates only: The option holder has the right to receive an underlying $V_{\text{underl},1}$ in T_1 or wait and retain the right to either receive an underlying $V_{\text{underl},2}$ in T_2 or receive nothing. Put differently, the option holder has the choice of receiving the underlying value $V_{\text{underl},1}(T_1)$ or the value of an option $V_{\text{option}}(T_1)$ on $V_{\text{underl},2}(T_2)$. The Bermudan may be interpreted as an option on an option. In T_1 the optimal exercise is given by choosing the maximum value

$$\max(V_{\text{underl}}(T_1), V_{\text{option}}(T_1)), \quad (14.1)$$

where (having chosen a Numéraire N)

$$\begin{aligned} V_{\text{option}}(T_1) &= N(T_1)E^{\mathbb{Q}^N}\left(\frac{V_{\text{option}}(T_2)}{N(T_2)} \mid \mathcal{F}_{T_1}\right) \\ &= N(T_1)E^{\mathbb{Q}^N}\left(\frac{\max(V_{\text{underl},2}(T_2), 0)}{N(T_2)} \mid \mathcal{F}_{T_1}\right). \end{aligned} \quad (14.2)$$

is the value of the option with exercise in T_2 , evaluated in T_1 .

Thus, to evaluate the exercise criteria (14.1) it is in general necessary to calculate a conditional expectation. The calculation of a conditional expectation within a Monte-Carlo simulation is a non trivial problem. The two main issues are *complexity* and *foresight bias*, which we will illustrate in Section 14.5 and 14.6. In the then following section we will present methods to efficiently estimate conditional expectations and/or the Bermudan exercise criteria within Monte-Carlo simulation. The application to the pricing of Bermudan options will always be present, but apart from the description of the backward algorithm in Section 14.4 the methods presented are not limited to Bermudan option pricing.

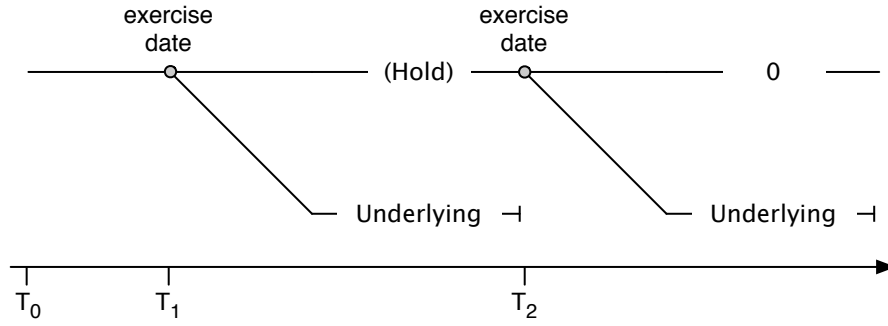


Figure 14.1.: A simple Bermudan option with two exercise dates.

14.2. Bermudan Options: Notation

We now give a fairly general definition of an Bermudan option and fix notation. Let $\{T_i\}_{i=1,\dots,n}$ denote a set of exercise dates and $\{V_{\text{underl},i}\}_{i=1,\dots,n}$ a corresponding set of underlyings. The Bermudan option is the right to receive at one and only one time T_i the corresponding underlying $V_{\text{underl},i}$ (with $i = 1, \dots, n$) or receive nothing.

At each exercise date T_i , the optimal strategy compares the value of the product upon exercise with the value of the product upon non-exercise and chooses the larger one. Thus the value of the bermudan is given recursively

$$\underbrace{V_{\text{berm}}(T_i, \dots, T_n; T_i)}_{\substack{\text{Bermudan with} \\ \text{exercise dates} \\ T_i, \dots, T_n}} := \max \left(\underbrace{V_{\text{berm}}(T_{i+1}, \dots, T_n; T_i)}_{\substack{\text{Bermudan with} \\ \text{exercise dates} \\ T_{i+1}, \dots, T_n}}, \underbrace{V_{\text{underl},i}(T_i)}_{\substack{\text{Product} \\ \text{received upon} \\ \text{exercise in } T_i}} \right), \quad (14.3)$$

where $V_{\text{berm}}(T_n; T_n) := 0$ and $V_{\text{underl},i}(T_i)$ denotes the value of the underlying $V_{\text{underl},i}$ at exercise date T_i .

14.2.1. Bermudan Callable

The most common Bermudan option is the Bermudan Callable. For a Bermudan Callable the underlyings consists of periodic payments X_k and differ only in the beginning of the periodic payments. The value of the underlying then becomes

$$V_{\text{underl},i}(T_i) = V_{\text{underl}}(T_i, \dots, T_n; T_i) = N(T_i) E^{\mathbb{Q}^N} \left(\sum_{k=i}^{n-1} \frac{X_k}{N(T_{k+1})} \middle| \mathcal{F}_{T_i} \right)$$

Here X_k denotes a payment fixed in T_k (i.e. $\mathcal{F}_{T_k}$ measurable) and paid in T_{k+1} . This is the common setup for interest rate bermudan callables. Other payment dates are a minor modification, they simply change the time argument of the Numéraire. For the value upon non exercise we have as before

$$V_{\text{berm}}(T_{i+1}, \dots, T_n; T_i) = N(T_i) E^{\mathbb{Q}^N} \left(\frac{V_{\text{berm}}(T_{i+1}, \dots, T_n; T_{i+1})}{N(T_{i+1})} \middle| \mathcal{F}_{T_i} \right).$$

If the value of the underlying may not be expressed by means of an analytical formula two conditional expectations have to be evaluated to calculate the exercise strategy (14.3).

14.2.2. Relative Prices

Since the conditional expectation of a Numéraire relative price is a Numéraire relative price the presentation will be simplified by considering the Numéraire relative quantities. We will therefore define:

$$\tilde{V}_{\text{underl},i}(T_j) := \frac{V_{\text{underl},i}(T_j)}{N(T_j)} \quad \text{and} \quad \tilde{V}_{\text{berm},i}(T_j) := \frac{V_{\text{berm}}(T_i, \dots, T_n; T_j)}{N(T_j)},$$

thus we have

$$\begin{aligned} \tilde{V}_{\text{berm},n} &\equiv 0, \\ \tilde{V}_{\text{berm},i+1}(T_i) &= E^{\mathbb{Q}^N}(\tilde{V}_{\text{berm},i+1}(T_{i+1}) \mid \mathcal{F}_{T_i}), \\ \tilde{V}_{\text{berm},i}(T_i) &= \max(\tilde{V}_{\text{berm},i+1}(T_i), \tilde{V}_{\text{underl},i}(T_i)), \end{aligned}$$

and in the case of a Bermudan callable

$$\tilde{V}_{\text{underl},i}(T_j) = E^{\mathbb{Q}^N} \left(\sum_{k=i}^{n-1} \tilde{X}_k(T_{k+1}) \mid \mathcal{F}_{T_j} \right) \quad \text{where} \quad \tilde{X}_i(T_{i+1}) := \frac{X_i}{N(T_i)}.$$

The relative prices are marked by a tilde.

Remark 167 (Notation): The processes $t \mapsto \tilde{V}_{\text{underl},i}(t)$ and $t \mapsto \tilde{V}_{\text{berm},i}(t)$ are $\mathcal{F}_t$ conditional expectations of $\tilde{V}_{\text{underl},i}(T_i)$ and $\tilde{V}_{\text{berm},i}(T_i)$ respectively and thus martingales by definition. The time-discrete processes $i \mapsto \tilde{V}_{\text{underl},i}(T_i)$, $i \mapsto \tilde{V}_{\text{berm},i}(T_i)$ consist of different products at different times and are thus not time-discrete martingales in general.

14.3. Bermudan Option as Optimal Exercise Problem

A Bermudan option consists of the right to receive one (and only one) of the underlyings $V_{\text{underl},i}$ at the corresponding exercise date T_i . The recursive definition (14.3) represents the optimal exercise strategy in each exercise time. We formalize this optimal exercise strategy:

For a given path $\omega \in \Omega$ let

$$T(\omega) := \min\{T_i : V_{\text{berm},i+1}(T_i, \omega) < V_{\text{underl},i}(T_i, \omega)\},$$

and

$$\eta : \{1, \dots, n-1\} \times \Omega \rightarrow \{0, 1\}, \quad \eta(i, \omega) := \begin{cases} 1 & \text{if } T_i \geq T(\omega) \\ 0 & \text{else.} \end{cases}$$

The definitions of T and η give equivalent descriptions of the exercise strategy: $T(\omega)$ is the optimal exercise time on a given path ω , $\eta(\cdot, \omega)$ is an indicator function which changes from 0 to 1 at the time index i corresponding to $T_i = T(\omega)$. The boundary $\partial\{\eta = 1\}$ of the set $\{\eta = 1\}$ is the so called exercise boundary. It should be noted that $\eta(k)$ is $\mathcal{F}_{T_k}$ measurable.

14.3.1. Bermudan Option Value as single (unconditioned) Expectation: The Optimal Exercise Value

With the definition of the optimal exercise strategy T (or η) it is possible to define a random variable which allows to express the Bermudan option value as a single (unconditioned) expectation. With

$$\tilde{U}(T_i) := \tilde{V}_{\text{underl},i}(T_i) \quad i = 1, \dots, n$$

denoting the relative price of the i -th underlying upon its exercise date T_i we have for the Bermudan value

$$\tilde{V}_{\text{berm}}(T_0) = E^{\mathbb{Q}}(\tilde{U}(T) \mid \mathcal{F}_{T_0}).$$

For the Bermudan Callable we may alternatively write

$$\tilde{V}_{\text{berm}}(T_0) = E^{\mathbb{Q}}\left(\sum_{k=1}^{n-1} \tilde{X}_k(T_{k+1}) \cdot \eta(k) \mid \mathcal{F}_{T_0}\right).$$

The random variable $\tilde{U}(T)$ may be calculated directly using the *backward algorithm*. We will consider this in the next section and conclude by giving $\tilde{U}(T)$ a name:

Definition 168 (Option Value upon Optimal Exercise): ▮

Let $\tilde{U}$ be the stochastic process whose time t value $U(t)$ is the (Numéraire relative)

option value received upon exercise in t . Let T be the optimal exercise strategy. The random variable $\tilde{U}(T)$, where

$$\tilde{U}(T)[\omega] := U(T(\omega), \omega)$$

is the (Numéraire relative) *option value received upon optimal exercise*. The (Numéraire relative) Bermudan option value is given by $E^{\mathbb{Q}}(\tilde{U}(T) \mid \mathcal{F}_{T_0})$. $\lrcorner$

Thus the value of $V_{\text{berm}}(T_1, \dots, T_n)$ may be expressed through a single expectation conditioned to T_0 and does not need any calculation of a conditional expectation at later times, if we have the optimal exercise date $T(\omega)$ (and thus $\eta(\cdot, \omega)$) for any path ω .

Remark 169 (Stopped Prozess): The random variable $\tilde{U}(T)$ is a so called *stopped process*. $\tilde{U}$ is a stochastic process and T is a random variable with the interpretation of a (stochastic) time. Furthermore T is a stopping time, see Definition 181. Here the stochastic process $\tilde{U}$ is the family of underlyings received upon exercise, parametrized by exercise time, and T is the optimal exercise time. Thus $\tilde{U}(T)$ is the underlying received upon optimal exercise. All quantities are stochastic, i.e. depend on the path.

14.4. Bermudan Option Pricing - The Backward Algorithm

The random variable $\tilde{U}(T)$ may be derived in a Monte-Carlo simulation through the backward algorithm, *given* the exercise criteria (14.3), i.e. the conditional expectation. The algorithm consists of the application of the recursive definition of the Bermudan value in (14.3) with a slight modification. Let:

Induction start:

$$\tilde{U}_{n+1} \equiv 0$$

Induction step $i + 1 \rightarrow i$ for $i = n, \dots, 1$:

$$\tilde{U}_i = \begin{cases} \tilde{U}_{i+1} & \text{if } \tilde{V}_{\text{underl},i}(T_i) < E^{\mathbb{Q}}(\tilde{U}_{i+1} \mid \mathcal{F}_{T_i}) \\ \tilde{V}_{\text{underl},i}(T_i) & \text{else.} \end{cases}$$

From the [Tower Law](#)¹ we have by induction $E^{\mathbb{Q}}(\tilde{U}_{i+1} \mid \mathcal{F}_{T_i}) = E^{\mathbb{Q}}(\tilde{V}_{\text{berm},i+1}(T_i) \mid \mathcal{F}_{T_i})$ and thus

$$\tilde{V}_{\text{berm}}(T_1, \dots, T_n, T_0) = E^{\mathbb{Q}}(\tilde{U}_1 \mid \mathcal{F}_{T_0}) \quad (14.4)$$

and $\tilde{U}_1 = \tilde{U}(T)$ with the notation from the previous section.

¹ See Exercise 2.



Interpretation: The recursive definition of $\tilde{U}_i$ differs from the recursive definition of $\tilde{V}_{\text{berm},i}(T_i)$. We have

$$\tilde{U}_i = \begin{cases} \tilde{U}_{i+1} & \text{if } \tilde{V}_{\text{underl},i}(T_i) < E^{\mathbb{Q}}(\tilde{U}_{i+1}|\mathcal{F}_{T_i}) \\ \tilde{V}_{\text{underl},i}(T_i) & \text{else,} \end{cases}$$

and

$$\tilde{V}_{\text{berm},i}(T_i) = \begin{cases} E^{\mathbb{Q}}(\tilde{V}_{\text{berm},i+1}(T_{i+1})|\mathcal{F}_{T_i}) & \text{if } \tilde{V}_{\text{underl},i}(T_i) < E^{\mathbb{Q}}(\tilde{U}_{i+1}|\mathcal{F}_{T_i}) \\ \tilde{V}_{\text{underl},i}(T_i) & \text{else.} \end{cases}$$

This is a subtle but crucial difference. While both definitions give the Bermudan option value (through application of (14.4)), we have that the definition of $\tilde{U}_i$ requires the conditional expectation operator only to calculate the exercise criteria.

Since a Monte-Carlo simulation requires advanced methods to obtain an (often not very accurate) estimate for the conditional expectation it is important to reduce their use.

Note that $\tilde{V}_{\text{berm},i}(T_i)$ is $\mathcal{F}_{T_i}$ measurable by definition as a $\mathcal{F}_{T_i}$ conditional expectation, while all $\tilde{U}_i$ are at most $\mathcal{F}_{T_n}$ measurable since they are defined pathwise from $\mathcal{F}_{T_k}$ measurable random variables $\tilde{V}_{\text{underl},k}(T_k)$ for $i \leq k \leq n$. $\triangleleft$

The pricing of a Bermudan option may thus be reduced to either the calculation of conditional expectations or to the calculation of the optimal exercise strategy T .

As motivation we consider in Sections 14.5 and 14.6 two methods which are not suitable to calculate conditional expectations.

14.5. Re-simulation

Let us consider the simplified example of a Bermudan option as given in Section 14.1. If no analytic calculation of the conditional expectation (14.2) is possible and if Monte-Carlo is the numerical tool to calculate expectations the straight forward way to calculate the conditional expectation is to create in T_1 a new Monte-Carlo simulation (conditioned) on each path - see Figure 14.2. This leads to a much higher number of total simulation paths needed. Especially if one considers more than one exercise date (option on option on option...) this method becomes impractical. The required number of paths, i.e. the complexity of the algorithm and thus the calculation time, grows exponentially with the number of exercise dates. This creates the need for efficient alternatives.



Interpretation: The calculation of conditional expectation in a path simulation requires further measures since the path simulation does not offer a suitable discretization of the filtration. $\triangleleft$

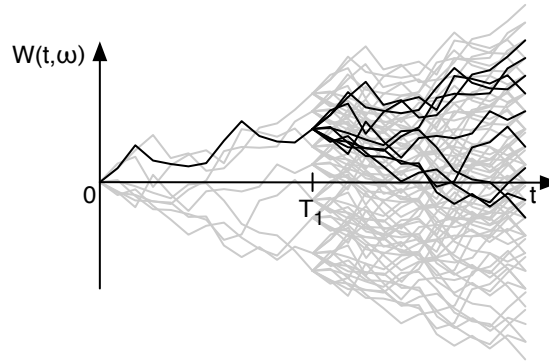


Figure 14.2.: Brute-force calculation of the conditional expectation by (path-wise) re-simulation.

14.6. Perfect Foresight

If one refuses to use a full resimulation and sticks to the paths generated in the original simulation then one effectively estimates the conditional expectation by a single path, namely by

$$\mathbb{E}^Q \left(\frac{V(T_2)}{N(T_2)} \mid \mathcal{F}_{T_1} \right) \approx \frac{V(T_2)}{N(T_2)}.$$

Put differently this is a limit case of the resimulation where each resimulation consists of a single path only, namely the one of the original simulation. If this estimate is used in the exercise criteria the exercise will be *super optimal* since it is based on future information that would be unknown otherwise.

The exercise criteria at time T_1 may only depend on information available in T_1 , i.e. on $\mathcal{F}_{T_1}$ measurable random variables. The estimate $\frac{V(T_2)}{N(T_2)}$ is *not* $\mathcal{F}_{T_1}$ measurable.

To illustrate the super optimality consider a simulation consisting of two paths, see Figure 14.3. Both paths are identical on $[0, T_1]$, i.e. $\mathcal{F}_{T_1} = \{\emptyset, \Omega\} = \{\emptyset, \{\omega_1, \omega_2\}\}$. We consider the option V to receive either $S(T_1) = 2$ at time T_1 or $S(T_2) \in \{1, 4\}$ at later time T_2 . The random variable $\eta : \Omega \rightarrow \{0, 1\}$ denotes the exercise strategy for T_1 : It is 1 on paths that exercise in T_1 , else 0. With a *perfect foresight* the super optimal exercise strategy is $T(\omega_1) = T_2$, $T(\omega_2) = T_1$, i.e. $\eta(\omega_1) = 0$, $\eta(\omega_2) = 1$ and an average value of $V(T_0) = \frac{1}{2}(4 + 2) = \frac{6}{2}$ will be received. Note that then η is not $\mathcal{F}_{T_1}$ measurable. The exercise decision is made in T_1 with knowledge of the future outcome. If we restrict the exercise strategy to the set of $\mathcal{F}_{T_1}$ measurable random variables we either get $\frac{1}{2}(4 + 1) = \frac{5}{2}$ using $\eta \equiv 0$ or $\frac{1}{2}(2 + 2) = \frac{4}{2}$ using $\eta \equiv 1$. Thus the optimal, $\mathcal{F}_{T_1}$ measurable (and thus admissible) exercise strategy is $\eta(\omega_1) = \eta(\omega_2) = 0$.

Perfect foresight is not a suitable method to estimate conditional expectation and through this the exercise criteria.

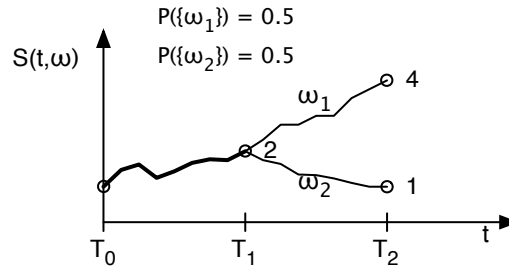


Figure 14.3.: Illustration of perfect foresight.

14.7. Conditional Expectation as Functional Dependence

Let us reconsider the calculation of the conditional expectation through brute force re-simulation as described in Section 14.5 and depicted in Figure 14.2. On each path of the original simulation a re-simulation has to be created. These re-simulations differ in their initial conditions (e.g. the value $S(T_1)$ in a simulation of a stock price following a Black-Scholes Modell, or the values $L_i(T_1)$ in a simulation of forward rates following a LIBOR Market Model). The initial conditions are $\mathcal{F}_{T_1}$ measurable random variables (known as of T_1). Thus the conditional expectation is a function of these initial conditions (and possibly other model parameters known in T_1). If it is known that the conditional expectation is a function of a $\mathcal{F}_{T_1}$ measurable random variable Z (we assume here that $Z : \Omega \rightarrow \mathbb{R}^d$ with some d) we have

$$E^{Q^N} \left(\frac{V(T_2)}{N(T_2)} \mid \mathcal{F}_{T_1} \right) = E^{Q^N} \left(\frac{V(T_2)}{N(T_2)} \mid Z \right). \quad (14.5)$$



Interpretation: If the random variable Z is such that $\mathcal{F}_{T_1}$ is the smallest σ field with respect to which Z is measurable (i.e. we have $Z^{-1}(\mathcal{B}(\mathbb{R}^d)) = \mathcal{F}_{T_1}$), then equation (14.5) is merely the definition of an expectation conditioned on a random variable. If, however, the conditional expectation on the left hand side (i.e. $\frac{V(T_1)}{N(T_1)}$) is known to be measurable with respect to a smaller σ field (e.g. because its functional depends on a smaller set of random variable), then it might be advantageous to use the right hand side representation. This representation is also useful to derive an approximation, e.g. if the functional dependence with respect to one component of Z is known to be weak and thus neglectable.

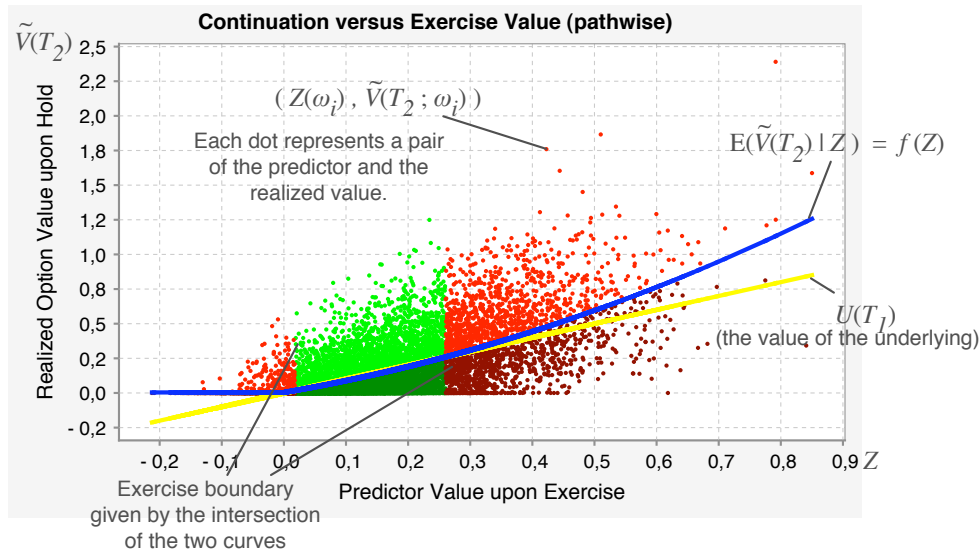


Figure 14.4.: Predictor variable versus realized value (continuation value): A diagram showing the path value of the predictor variable $Z(\omega_i)$ and the path value of $\tilde{V}(T_2; \omega_i) = \frac{V(T_2; \omega_i)}{N(T_2; \omega_i)}$. The conditional expectation is a function of Z dividing the cloud of dots.

Example: Consider a LIBOR Market Model with stochastic processes for the forward rates $L_1, L_2, \dots, L_n$. In T_1 we wish to calculate the conditional expectation of a derivative with a Numéraire relative payoff that depends on $L_2, \dots, L_k$ only (e.g. on a swap rate). While the filtration $\mathcal{F}_{T_1}$ is generated by the full set of forward rates $L_1(T_1), L_2(T_1), \dots, L_n(T_1)$ it is sufficient to know $L_2(T_1), \dots, L_k(T_1)$ to describe the conditional expectation (i.e. the conditional value of the product). $\triangleleft$

We will now describe methods that derive the functional dependance of the conditional expectation from a given set of random variables.

14.8. Binning

In a path simulation the approximation of $E^Q\left(\frac{V(T_2)}{N(T_2)} \mid Z\right)$ will be given by averaging over all paths for which Z attains the same value. For the simple example in Figure 14.3 this would remove the *perfect foresight* since $S(T_1)^{-1}(2) = \Omega$. In general the situation will be such that there are no two or more paths for which Z attains the same value - apart from the construction of the unfeasible re-simulation. Thus this approximation will show a perfect foresight.

An improvement is given by a *binning*, where the averaging will be done over those paths for which Z lies in a neighborhood (*bin*). If the quantities are continuous we

have:

$$\mathbb{E}^{\mathbb{Q}} \left(\frac{V(T_2)}{N(T_2)} \mid Z \right) [\omega] \approx \mathbb{E}^{\mathbb{Q}} \left(\frac{V(T_2)}{N(T_2)} \mid Z \in U_\epsilon(Z(\omega)) \right),$$

where $U_\epsilon(Z(\omega)) := \{y \mid \|Z(\omega) - y\| < \epsilon\}$.

Instead of defining a bin $U_\epsilon(Z(\omega))$ for each path ω it is more efficient to start with a partition of $Z(\Omega)$ into a finite set of disjoint bins $U_i \subset Z(\Omega)$. The approximation of the conditional expectation

$$\mathbb{E}^{\mathbb{Q}} \left(\frac{V(T_2)}{N(T_2)} \mid Z(\omega) \right)$$

will then be given by

$$H_i := \mathbb{E}^{\mathbb{Q}} \left(\frac{V(T_2)}{N(T_2)} \mid Z \in U_i \right)$$

where U_i denote the set with $Z(\omega) \in U_i$.

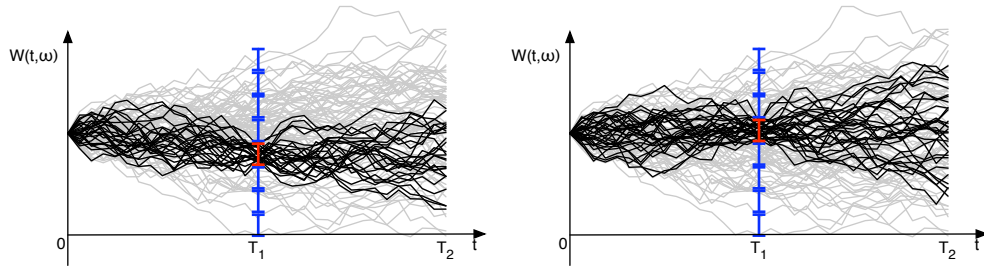


Figure 14.5.: Calculation of the conditional expectation by binning: Paths, which are nearby, i.e. paths, which belong to the same bin, are bundled. The bins are defined through the $\mathcal{F}_{T_1}$ -measurable predictor variable Z . The figure shows the special case $Z = W(T_1)$.

Example: Pricing of a simple Bermudan Option on a Stock

We illustrate the method in a simple **Black-Scholes model** for a stock S . In T_1 we wish to evaluate the option of receiving $N_1 \cdot (S(T_1) - K_1)$ in T_1 or to receive $N_2 \cdot \max(S(T_2) - K_2, 0)$ at later time T_2 (where N_1, N_2 (notional), K_1, K_2 (strike) are given). The optimal exercise in T_1 compares the exercise value with the value of the T_2 option, i.e.

$$\mathbb{E}^{\mathbb{Q}} \left(\frac{N_2 \cdot \max(S(T_2) - K_2, 0)}{N(T_2)} \mid \mathcal{F}_{T_1} \right).$$

From the model specification, e.g. here a **Black-Scholes model**

$$dS(t) = r \cdot S(t)dt + \sigma S(t)dW^{\mathbb{Q}}(t), \quad N(t) = \exp(rt)$$

it is obvious that the price of the T_2 option seen in T_1 is a given function $S(T_1)$ and the given model parameters (r, σ) . Thus it is sufficient to calculate

$$\mathbb{E}^{\mathbb{Q}} \left(\frac{N_2 \cdot \max(S(T_2) - K_2, 0)}{N(T_2)} \mid S(T_1) \right).$$

In this example the functional dependence is known analytically. It is given by the Black-Scholes formula (4.3). Nevertheless we use the binning to calculate an approximation to the conditional expectation. If we plot

$$\underbrace{\frac{N_2 \cdot \max(S(T_2, \omega_i) - K_2, 0)}{N(T_2)}}_{\text{Continuation Value}} \quad \text{as a function of} \quad \underbrace{S(T_1, \omega_i)}_{\text{Underlying}}$$

we obtain the scatter plot in Figure 14.6, left. For a given $S(T_1)$ none or very few values of the *continuation values* exists. An estimate is not possible or exhibits a foresight bias. For an interval $[S_1 - \epsilon, S_1 + \epsilon]$ with sufficiently large ϵ we have enough values to calculate an estimate of

$$\mathbb{E}^{\mathbb{Q}} \left(\frac{N_2 \cdot \max(S(T_2) - K_2, 0)}{N(T_2)} \mid S(T_1) \in [S_1 - \epsilon, S_1 + \epsilon] \right)$$

which in turn may be used as an estimate of

$$\mathbb{E}^{\mathbb{Q}} \left(\frac{N_2 \cdot \max(S(T_2) - K_2, 0)}{N(T_2)} \mid S(T_1) = S_1 \right).$$

In Figure 14.6, right, we calculate this estimate for $S_1 = 1$ and $\epsilon = 0.05$.

14.8.1. Binning as a Least-Square Regression

Consider again the binning: As an estimate for the conditional expectation $\mathbb{E}^{\mathbb{Q}} \left(\frac{V(T_2)}{N(T_2)} \mid Z(\omega) \right)$ we calculated the conditional expectation

$$H_i := \mathbb{E}^{\mathbb{Q}} \left(\frac{V(T_2)}{N(T_2)} \mid Z \in U_i \right) \quad (14.6)$$

given a *bin* U_i with $Z(\omega) \in U_i$.

For the expectation operator $\mathbb{E}^{\mathbb{Q}}$ an alternative characterization may be used:

Lemma 170 (Characterization of the Expectation as Least-Square Approximation): The expectation of a random variable X is the number h for which $X - h$ has the smallest variance (i.e. $L_2(\Omega)$ norm).

Proof: Let X be a real valued random variable. Then we have for any $h \in \mathbb{R}$

$$E((X - h)^2) = E(X^2) - 2 \cdot E(X) \cdot h + h^2 =: f(h).$$

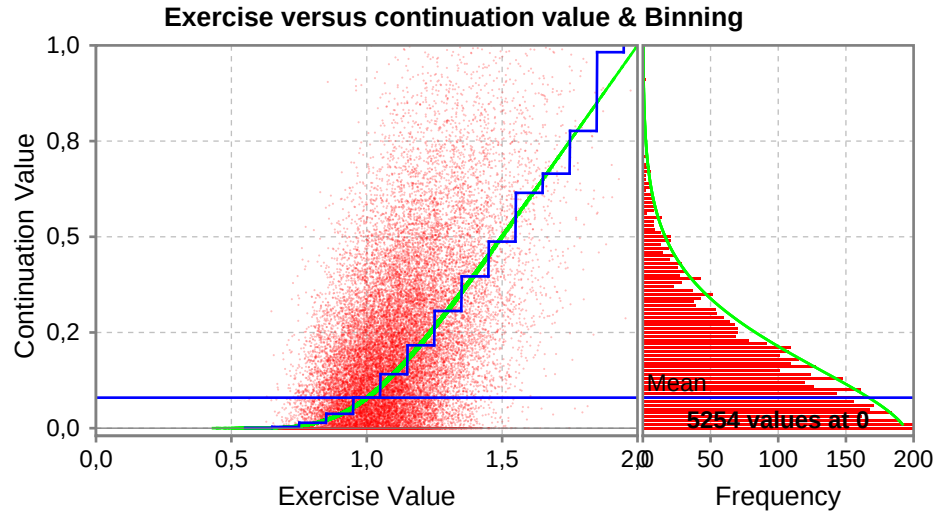


Figure 14.6.: The Continuation Value as a function of the Underlying (Spot Value) and the calculation of the conditional expectation through a binning

Since $f' = 0 \Leftrightarrow h = E(X)$ and $f'' = 1 \geq 0$ we have that f attains its minimum in $h = E(X)$. For vector valued random variables this follows componentwise. The same result holds for conditional expectations. $\square$

Using Lemma 170 we may write (14.6) as a minimization problem:

$$E^Q \left(\left(\frac{V(T_2)}{N(T_2)} - H_i \right)^2 \mid Z \in U_i \right) = \min_{G \in \mathbb{R}} E^Q \left(\left(\frac{V(T_2)}{N(T_2)} - G \right)^2 \mid Z \in U_i \right).$$

For disjoint bins U_i this may be written in a single minimization problem for the vector $(H_i)_{i=1, \dots, n}$:

$$\sum_i E^Q \left(\left(\frac{V(T_2)}{N(T_2)} - H_i \right)^2 \mid Z \in U_i \right) = \min_{G_i \in \mathbb{R}} \sum_i E^Q \left(\left(\frac{V(T_2)}{N(T_2)} - G_i \right)^2 \mid Z \in U_i \right). \quad (14.7)$$

This condition admits an alternative interpretation: H_i represents the piecewise constant function (constant on U_i) with the minimal distance to $\frac{V(T_2)}{N(T_2)}$ in the least square sense.

Let $\mathcal{H}$ be the space of functions $H : \Omega \rightarrow \mathbb{R}$ being constant on the bins $Z^{-1}(U_i)$.² Let $H \in \mathcal{H}$ with $H(\omega) := H_i$ for $\omega \in Z^{-1}(U_i)$. Then (14.7) is equivalent to

$$E^Q \left(\left(\frac{V(T_2)}{N(T_2)} - H \right)^2 \mid Z \right) = \min_{H \in \mathcal{H}} E^Q \left(\left(\frac{V(T_2)}{N(T_2)} - H \right)^2 \mid Z \right). \quad (14.8)$$

² Note that the bins U_i were defined as subsets of $Z(\Omega)$, here we consider H as a function on Ω .

Equation (14.8) is the definition of a regression: Find the function H from a function space $\mathcal{H}$ with minimum distance to $\frac{V(T_2)}{N(T_2)}$ in the L_2 norm. Binning is just a special choice of functional space:

Lemma 171 (Binning as L_2 Regression): Binning is an L_2 regression on the space of functions piecewise constant on U_i .

14.9. Foresight Bias

Definition 172 (Foresight Bias (Definition 1)):

A *foresight bias* is a super optimal exercise strategy.

A foresight bias arises due to a violation of the measurability requirements: If the exercise decision in T_i is based on a random variable which is not $\mathcal{F}_{T_i}$ measurable the exercise may be super optimal, i.e. better than if based on the information theoretically available ($\mathcal{F}_{T_i}$). If we use the same Monte-Carlo simulation to first estimate the exercise criteria and then use this criteria to price the derivative we most certain generate a foresight bias. In this case the foresight bias is created by the Monte-Carlo error of the estimate, which is in general not $\mathcal{F}_{T_i}$ measurable. The existence of this problem becomes obvious if we consider a limit case of binning where each bin contains a single path only. Here we would have a [perfect foresight](#).

If our exercise criteria at time T_i uses only $\mathcal{F}_{T_i}$ measurable random variables then there is - in theory - no foresight bias. If however the exercise criteria is calculated within a Monte Carlo simulation the Monte-Carlo error of the calculation represents a non $\mathcal{F}_{T_i}$ measurable random variable, thus it induces a foresight bias. In this case we may give an alternative definition for the foresight bias:

Definition 173 (Foresight Bias (Definition 2)):

The *foresight bias* is the value of the option on the Monte Carlo error.

With increasing number of paths the foresight bias introduced by binning tends to zero since the Monte-Carlo error with respect to a bin tends to zero.

A general solution of the problem of a foresight bias is given by using two independent Monte-Carlo simulations: One to estimate the exercise criteria (for binning this is given by the H_i corresponding to the U_i 's), the other to apply the criteria in pricing. This is a numerical removal of the foresight bias. In [59] an analytic formula for the (Monte Carlo error induced) foresight bias is derived. It may be used to correct the foresight bias analytically.

14.10. Regression Methods - Least Square Monte Carlo



Motivation Disadvantage of Binning: The partition of the state space $Z(\Omega)$ into a finite number of bins results in a piecewise constant approximation of the conditional expectation. An obvious improvement would be to approximate the conditional expectation by some smooth function of the state variable Z .

The considerations in Section 14.8.1 suggest a simple yet powerful improvement to the binning: The function giving our estimate for the conditional expectation is defined by a least square approximation (regression). $\triangleleft$

14.10.1. Least Square Approximation of the Conditional Expectation

Let us start with a fairly general definition of the *least square approximation* of the conditional expectation of random variable U .

Definition 174 (Least Square Approximation of the Conditional Expectation): ∇

Let $(\Omega, \mathcal{F}, \mathbb{Q}, \{\mathcal{F}_t\})$ be a filtered probability space and V a $\mathcal{F}_{T_1}$ measurable random variable defined as the conditional expectation of U

$$V = E^{\mathbb{Q}}(U \mid \mathcal{F}_{T_1}),$$

where U is at least $\mathcal{F}$ measurable. Furthermore let $Y := (Y_1, \dots, Y_p)$ be a given $\mathcal{F}_{T_1}$ measurable random variable and $f : \mathbb{R}^p \times \mathbb{R}^q$ a given function. Let $\Omega^* = \{\omega_1, \dots, \omega_N\}$ be a drawing from Ω (e.g. a Monte-Carlo simulation corresponding to $\mathbb{Q}$) and $\alpha^* := (\alpha_1, \dots, \alpha_q)$ such that

$$\|U - f(Y, \alpha^*)\|_{L_2(\Omega^*)} = \min_{\alpha} \|U - f(Y, \alpha)\|_{L_2(\Omega^*)}$$

where $\|U - f(Y, \alpha^*)\|_{L_2(\Omega^*)}^2 = \sum_{j=1}^N (U(\omega_j) - f(Y(\omega_j), \alpha^*))^2$. We set

$$V^{\text{LS}} := f(Y, \alpha^*).$$

The random variable V^{LS} is $\mathcal{F}_{T_1}$ measurable. It is defined over Ω and a *least square approximation* of V on Ω^* . $\lrcorner$

The method of Carriere, Longstaff and Schwartz³ uses a function f with $q = p$ and

$$f(y_1, \dots, y_p, \alpha_1, \dots, \alpha_p) := \sum_{i=1}^p \alpha_i \cdot y_i,$$

³ See [55, 76, 79].

such that α^* may be calculated analytically as a linear regression.

Lemma 175 (Linear Regression): Let $\Omega^* = \{\omega_1, \dots, \omega_n\}$ be a given sample space, $V : \Omega^* \rightarrow \mathbb{R}$ and $Y := (Y_1, \dots, Y_p) : \Omega^* \rightarrow \mathbb{R}^p$ given random variables. Furthermore let

$$f(y_1, \dots, y_p, \alpha_1, \dots, \alpha_p) := \sum \alpha_i y_i.$$

Then we have for any α^* with $X^T X \alpha^* = X^T v$

$$\|V - f(Y, \alpha^*)\|_{L_2(\Omega^*)} = \min_{\alpha} \|V - f(Y, \alpha)\|_{L_2(\Omega^*)},$$

where

$$X := \begin{pmatrix} Y_1(\omega_1) & \dots & Y_p(\omega_1) \\ \vdots & & \vdots \\ Y_1(\omega_n) & \dots & Y_p(\omega_n) \end{pmatrix}, \quad v := \begin{pmatrix} V(\omega_1) \\ \vdots \\ V(\omega_n) \end{pmatrix}.$$

If $(X^T X)^{-1}$ exists then $\alpha^* := (X^T X)^{-1} X^T v$.

Proof: See Appendix A.2. □

Definition 176 (Basis Functions): ⌋

The random variables $Y_1, \dots, Y_p$ of Lemma 175 are called *Basis Functions (explanatory variables)*. ⌋

14.10.2. Example: Evaluation of a *Bermudan* Option on a Stock (Backward Algorithm with Conditional Expectation Estimator)

We consider a simple Bermudan option on a Stock. The Bermudan should allow exercise at times $T_1 < T_2 < \dots < T_n$. Upon exercise in T_i the holder of the option will receive

$$N_i \cdot (S(T_i) - K_i)$$

once. If no exercise is made he will receive nothing.

We will apply the *backward algorithm* to derive the optimal exercise strategy. All payments will be considered in their Numéraire relative form. Thus the exercise criteria is given by a comparison of the conditional expectation of the payments received upon non-exercise with the payments received upon exercise.

Induction start: $t > T_n$ Beyond the last exercise we have:

- The value of the (future) payments is $\tilde{U}_{n+1} = 0$

Induction step: $t = T_i$, $i = n, n-1, n-2, \dots, 1$ In T_i we have:

- In the case of exercise in T_i the value is

$$\tilde{V}_{\text{underl},i}(T_i) := \frac{N_i(S(T_i) - K_i)}{N(T_i)}. \quad (14.9)$$

- In the case of non-exercise in T_i the value is $\tilde{V}_{\text{hold},i}(T_i) = \mathbb{E}^{\mathbb{Q}}(\tilde{U}_{i+1} | \mathcal{F}_{T_i})$. This value is estimated through a regression for given paths $\omega_1, \dots, \omega_m$:

- Let B_j be given ($\mathcal{F}_{T_i}$ measurable) basis functions.⁴ Let the matrix X consist of the column vectors $B_j(\omega_k)$, $k = 1, \dots, m$. Then we have

$$\begin{pmatrix} \tilde{V}_{\text{hold},i}(T_i, \omega_1) \\ \vdots \\ \tilde{V}_{\text{hold},i}(T_i, \omega_m) \end{pmatrix} \approx X \cdot (X^T \cdot X)^{-1} \cdot X^T \cdot \begin{pmatrix} \tilde{U}_{i+1}(\omega_1) \\ \vdots \\ \tilde{U}_{i+1}(\omega_m) \end{pmatrix}. \quad (14.10)$$

- The value of the payments of the product in T_i under optimal exercise is given by

$$\tilde{U}_i := \begin{cases} \tilde{V}_{\text{underl},i}(T_i) & \text{if } \tilde{V}_{\text{hold},i}(T_i) < \tilde{V}_{\text{underl},i}(T_i) \\ \tilde{U}_{i+1} & \text{else.} \end{cases}$$

Remark 177: Our example is of course just the backward algorithm with an explicit specification of an underlying (14.9) and an explicit specification of an exercise criteria, here given by the estimator of the conditional expectation (14.10).

14.10.3. Example: Evaluation of an Bermudan Callable

We consider a Bermudan Callable. The Bermudan should allow exercise at times $T_1 < T_2 < \dots < T_n$. Upon exercise in T_i the holder of the option will receive a payment of X_i in T_{i+1} , i.e. the relative value $\tilde{X}_i(T_{i+1}) := \frac{X_i}{N(T_{i+1})}$.

We will apply the *backward algorithm* to derive the optimal exercise strategy. All payments will be considered in their Numéraire relative form.

Induction start: $t > T_n$ Beyond the last exercise we have:

- The value of the (future) payments is $\tilde{U}_{n+1} = 0$

⁴ Suitable basis functions for this example are 1 (constant), $S(T_i)$, $S(T_i)^2$, $S(T_i)^3$, etc., such that the regression function f will be a polynomial in $S(T_i)$.

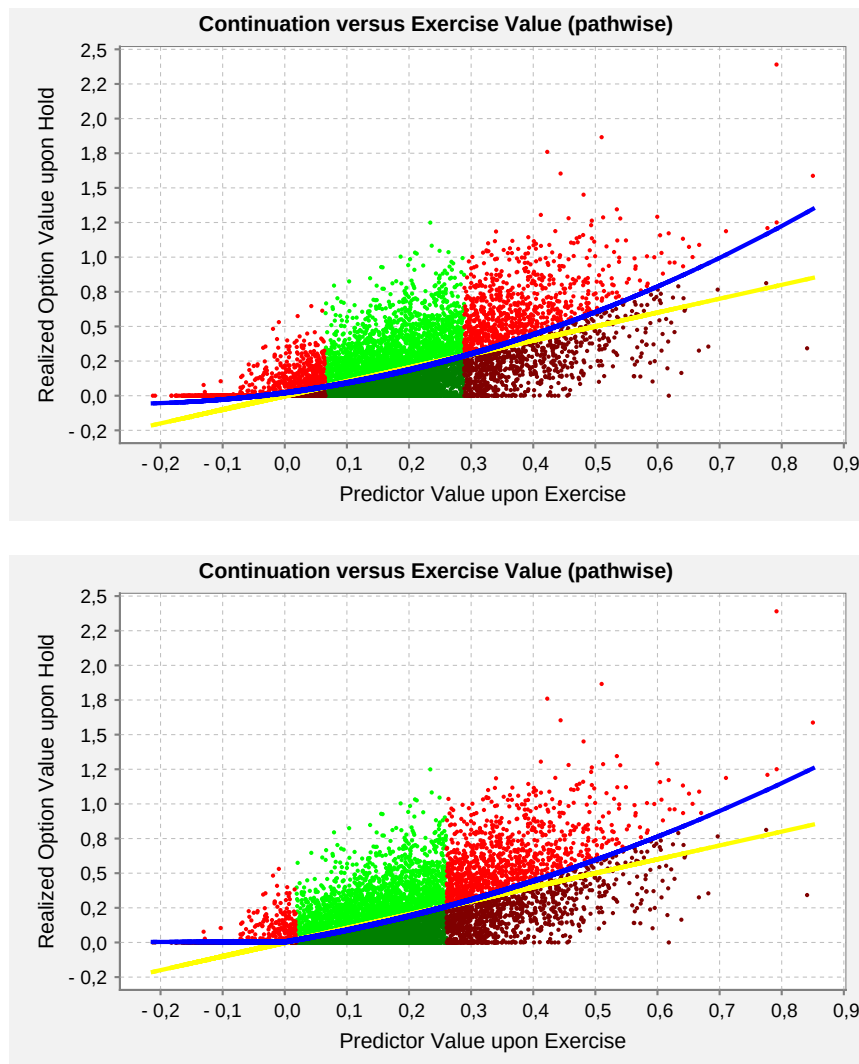


Figure 14.7.: Regression of the conditional expectation estimator with and without restriction of the regression domain: We consider a Bermudan option with two exercise dates $T_1 = 1.0$, $T_1 = 2.0$. Notional and strike are as follows: $N_1 = 0.7$, $N_2 = 1.0$, $K_1 = 0.82$, $K_2 = 1.0$. The model for the underlying S is a Black-Scholes model with $r = 0.05$ and $\sigma = 20\%$. The plot shows the values received upon exercise depending on the values received upon non-exercise in T_1 . Each dot corresponds to a path. The regression polynomial gives the estimator for the expectation of the value upon non-exercise. It is optimal to exercise if this estimate lies above the value received upon exercise. In the upper diagram the regression polynomial is a second order polynomial in $S(T_1)$. In the lower diagram it is a second order polynomial in $\max(S(T_1) - K_1, 0)$. By this values where $S(T_1) - K_1 \leq 0$ are aggregated into a single point. For the product under consideration this is advantageous since for $S(T_1) - K_1 \leq 0$ exercise is not optimal with probability 1. This restriction of the regression domain increases the regression accuracy over the remaining regression domain.

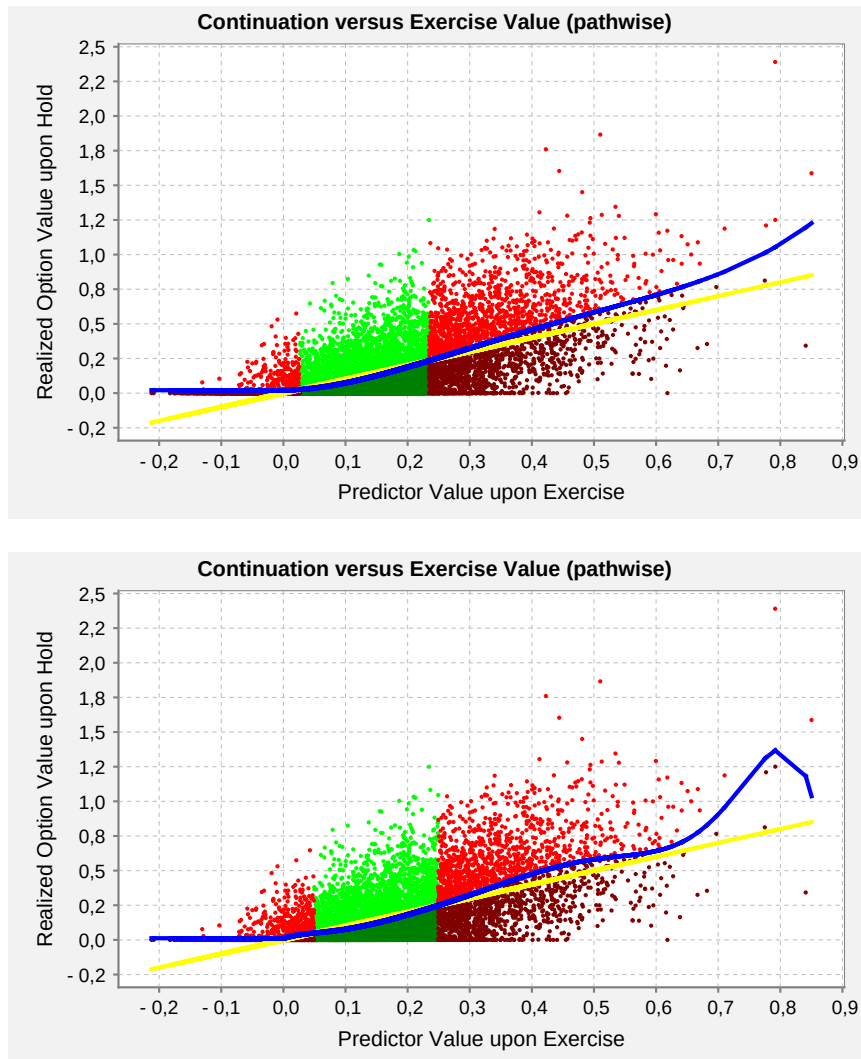


Figure 14.8.: Regression of the conditional expectation estimator - polynomial of fourth (above) and eighth (below) order in $\max(S(T_1) - K_1, 0)$. Parameters as in Figure 14.7. A polynomial of higher order shows wiggles at the boundary of the regression domain. However only few paths are affected by the wrong estimate. Restricting the regression domain may reduce the errors (compare the left end of the regression domain with the right end).

Induction step: $t = T_i$, $i = n, n-1, n-2, \dots, 1$ In T_i we have:

- In the case of exercise in T_i the value is

$$\tilde{V}_{\text{underl},i}(T_i) := E^{\mathbb{Q}^N} \left(\sum_{k=i}^{n-1} \frac{X_k}{N(T_{k+1})} \middle| \mathcal{F}_{T_i} \right).$$

This value is estimated through a regression for given paths $\omega_1, \dots, \omega_m$:

Let B_j^1 be given ($\mathcal{F}_{T_i}$ measurable) basis functions. Let the matrix X^1 consist of the column vectors $B_j^1(\omega_k)$, $k = 1, \dots, m$. Then we have

$$\begin{pmatrix} \tilde{V}_{\text{underl},i}(T_i, \omega_1) \\ \vdots \\ \tilde{V}_{\text{underl},i}(T_i, \omega_m) \end{pmatrix} \approx X^1 \cdot (X^{1,\top} \cdot X^1)^{-1} \cdot X^{1,\top} \cdot \begin{pmatrix} \sum_{k=i}^{n-1} \frac{X_k(\omega_1)}{N(T_{k+1}, \omega_1)} \\ \vdots \\ \sum_{k=i}^{n-1} \frac{X_k(\omega_m)}{N(T_{k+1}, \omega_m)} \end{pmatrix}. \quad (14.11)$$

- In the case of non-exercise in T_i the value is $\tilde{V}_{\text{hold},i}(T_i) = E^{\mathbb{Q}}(\tilde{U}_{i+1} | \mathcal{F}_{T_i})$. This value is estimated through a regression for given paths $\omega_1, \dots, \omega_m$:

Let B_j^0 be given ($\mathcal{F}_{T_i}$ measurable) basis functions. Let the matrix X^0 consist of the column vectors $B_j^0(\omega_k)$, $k = 1, \dots, m$. Then we have

$$\begin{pmatrix} \tilde{V}_{\text{hold},i}(T_i, \omega_1) \\ \vdots \\ \tilde{V}_{\text{hold},i}(T_i, \omega_m) \end{pmatrix} \approx X^0 \cdot (X^{0,\top} \cdot X^0)^{-1} \cdot X^{0,\top} \cdot \begin{pmatrix} \tilde{U}_{i+1}(\omega_1) \\ \vdots \\ \tilde{U}_{i+1}(\omega_m) \end{pmatrix}.$$

- The value of the payments of the product in T_i under optimal exercise is given by

$$\tilde{U}_i := \begin{cases} \sum_{k=i}^{n-1} \frac{X_k}{N(T_{k+1})} & \text{if } \tilde{V}_{\text{hold},i}(T_i) < \tilde{V}_{\text{underl},i}(T_i) \\ \tilde{U}_{i+1} & \text{else.} \end{cases}$$

Remark 178 (Bermudan Callable): The modification to the backward algorithm to price a Bermudan callable consists of the inductions of two conditional expectation estimators: one for the continuation value and (additionally) one for the underlying. As before, the conditional expectation estimators are used only for the exercise criteria (and not for the payment).

Remark 179 (Longstaff-Schwartz):

- The estimator of the conditional expectation is used in the estimation of the exercise strategy only.
- The choice of basis functions is crucial to the quality of the estimate.

Clément, Lamberton and Protter showed convergence of the Longstaff-Schwartz regression method to the exact solution, see [56].

14.10.4. Implementation

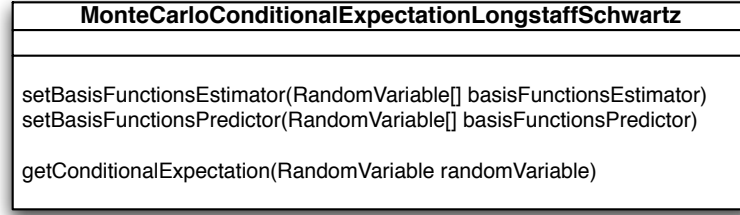


Figure 14.9.: UML Diagram: Conditional Expectation Estimator. The method *setBasisFunctionsEstimator* sets the basis functions which form the matrix X . The method *setBasisFunctionsPredictor* sets the basis functions which form the matrix X^* . These are the same basis functions as for X , but possibly evaluated in an independent Monte-Carlo simulation (to avoid foresight bias). The method *getConditionalExpectation* calculates the regression parameter $\alpha^* = (X^T X)^{-1} X^T \cdot v$ from a given vector v and returns the conditional expectation estimator $X^* \cdot \alpha^*$ of v . See Lemma 175.

The Longstaff-Schwartz conditional expectation estimator may easily be implemented in a corresponding class, independent from the given model or Monte-Carlo simulation - see Figure 14.9. This class contains nothing more than a [linear regression](#) however the methodology may be replaced by alternative algorithms (e.g. non parametric regressions).

As pointed out in the discussion of the backward algorithm it is in general not necessary to explicitly calculate the exercise strategy in form of T to η . It is sufficient to calculate the random variables $\tilde{U}_i$ in a backward recursion. Since finally only $\tilde{U}_1$ is needed to calculate the price of the Bermudan option the $\tilde{U}_i$'s may be stored in the same vector of Monte-Carlo realizations.

14.10.5. Binning as linear Least-Square Regression

We return once again to the binning. In Section 14.8.1 it turned out that *Binning* may be interpreted as least square regression with a specific set of basis functions: The indicator variables of the bins U_j which we denote by

$$h_j(\omega) := \begin{cases} 1 & \text{for } \omega \in U_j \\ 0 & \text{else.} \end{cases} \quad (14.12)$$

We now give an explicit calculation using the linear regression algorithm with the bin indicator variables as basis functions.

Let ω_k denote the paths of a Monte-Carlo simulation and X the matrix $(h_j(\omega_k))$, j column index, k row index. Since the U_j 's are disjoint we have $X^T X =$

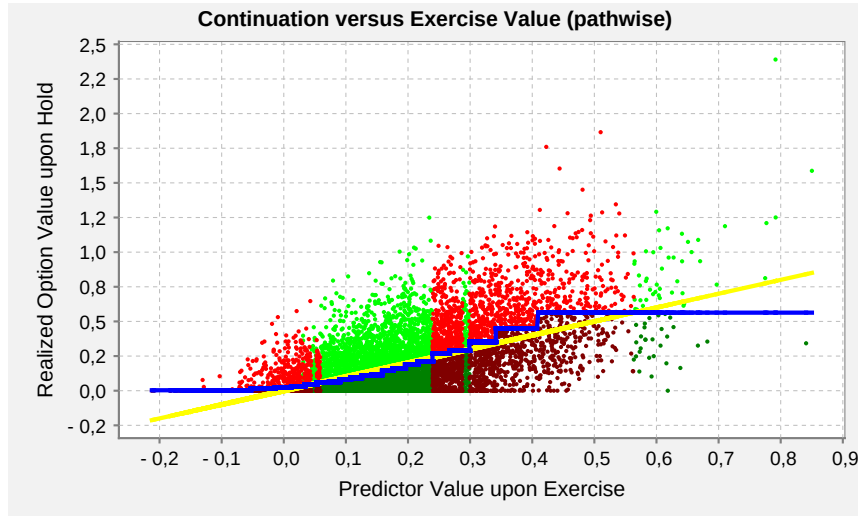


Figure 14.10.: *Binning using the linear regression algorithm with piecewise constant basis functions: We use 20 bins (basis functions). Each bin consist of approximately the same number of paths. Model- und product parameters are as in Figure 14.7.*

$\text{diag}(n_1, \dots, n_m)$, where n_j is the number of paths for which $h_j(\omega_k) = 1$. Thus we have for the regression parameter

$$\alpha^* = (X^T X)^{-1} X^T \cdot v = \text{diag}\left(\frac{1}{n_1}, \dots, \frac{1}{n_m}\right) \cdot X^T \cdot v.$$

It follows that the regression parameter gives the expectation on the corresponding bin

$$\alpha_j^* = \frac{1}{n_j} \sum_{v_k \in U_j} v_k.$$

14.11. Optimization Methods



Motivation: In the discussion of the *backward algorithm* it has become obvious that the conditional expectation estimator is needed to derive the optimal exercise strategy only. Since a sub-optimal exercise will lead to a lower Bermudan price the optimal exercise has an alternative characterization: it maximizes the Bermudan value. A solution to the pricing problem of the Bermudan thus consists of maximizing the Bermudan value over a suitable, sufficiently large space of admissible⁵ exercise strategies. $\triangleleft$

14.11.1. Andersen Algorithm for Bermudan Swaptions

The following method was proposed for the valuation of Bermudan swaptions by L. Andersen in [42]. We thus restrict our presentation to the evaluation of the Bermudan Callable and use the notation of Section 14.10.3. In [42] the method appears less generic than the Longstaff & Schwartz regression. However one might reformulate the optimization method in a fairly generic way. Since then the optimization is a high dimensional one the method becomes less useful in practice.

The exercise strategy is given by a parametrized function of the underlyings

$$I_i(\lambda) := f(V_{\text{underl},i}(T_i, \omega), \dots, V_{\text{underl},n-1}(T_i, \omega), \lambda)$$

where we replace the optimal exercise

$$\tilde{V}_{\text{underl},i}(T_i) < E^{\mathbb{Q}}(\tilde{U}_{i+1} | \mathcal{F}_{T_i})$$

by

$$I_i(\lambda) > 0.$$

Here the function f may represent a variety of exercise criterias, e.g.

$$I_i(\omega, \lambda) = V_{\text{underl},i}(T_i, \omega) - \lambda. \quad (14.13)$$

We assume that I_i is such that it may be calculated without resimulation, i.e. we assume that the underlyings $V_{\text{underl},j}(T_i)$ are either given by an analytic formula or a suitable approximation. E.g. this is the case for a swap within a LIBOR Market Model. Using the optimization method within the backward algorithm now looks as follows:

Induction start: $t > T_n$ Beyond the last exercise we have:

- The value of the (future) payments is $\tilde{U}_{n+1} = 0$

⁵ By an admissible exercise strategy we denote one that respects the measurability requirements. As we noted, a violation of measurability requirements, i.e. a foresight bias or even perfect foresight, will result in a super optimal strategy. The Bermudan value with a super optimal strategy is higher than the Bermudan value with the optimal strategy, however super optimal exercise is impossible.

Induction step: $t = T_i$, $i = n, n-1, n-2, \dots, 1$ In T_i we have:

- In the case of exercise is in T_i the value is $\tilde{V}_{\text{underl},i}(T_i)$
- In the case of non-exercise in T_i the value is $\tilde{V}_{\text{hold},i}(T_i) = E^{\mathbb{Q}}(\tilde{U}_{i+1} | \mathcal{F}_{T_i})$. This value is estimated through an optimization for given paths $\omega_1, \dots, \omega_m$:
 - $I_i(\lambda, \omega) = f(V_{\text{underl},i}(T_i, \omega), \dots, V_{\text{underl},n-1}(T_i, \omega), \lambda)$.
 - $\tilde{U}_i(\lambda, \omega) := \begin{cases} \tilde{V}_{\text{underl},i}(T_i) & \text{if } I_i(\lambda, \omega) > 0 \\ \tilde{U}_{i+1} & \text{else.} \end{cases}$
 - $\tilde{V}_{\text{berm},i}(T_0, \lambda) = E^{\mathbb{Q}}(\tilde{U}_i(\lambda) | \mathcal{F}_{T_0}) \approx \frac{1}{m} \sum_k \tilde{U}_i(\lambda, \omega_k)$
 - $\lambda^* = \arg \max_{\lambda} \left(\frac{1}{m} \sum_k \tilde{U}_i(\lambda, \omega_k) \right)$
- The value of the payments of the product in T_i under optimal exercise is given by

$$\tilde{U}_i(\omega) := \begin{cases} \tilde{V}_{\text{underl},i}(T_i) & \text{if } I_i(\lambda^*, \omega) > 0 \\ \tilde{U}_{i+1} & \text{else.} \end{cases} = \tilde{U}_i(\lambda^*, \omega)$$

The exercise strategy is estimated in T_i by choosing the λ^* for which I_i gives the maximal Bermudan option value. This is done by going backward in time, from exercise date to exercise date.

14.11.2. Review of the Threshold Optimization Method

14.11.2.1. Fitting the exercise strategy to the product

We apply the optimization method to the pricing of a simple Bermudan option on a stock following a Black-Scholes model. We show that choosing the exercise strategy too simple will give surprisingly unreliable results.

The simple strategy (14.13) fails for the most simple type of a Bermudan option. We consider the option to receive

$$N_1 \cdot (S(T_1) - K_1)$$

in T_1 or receive

$$\max(N_2 \cdot (S(T_2) - K_2), 0)$$

in T_2 , where - as before - N_i and K_i denote notional and strike and S follows a Black-Scholes model. This gives us an analytic formula for the option in T_2 and thus the true optimal exercise. Figure 14.11 shows an example where the optimization of the simple strategy (14.13) gives the value of the Bermudan option.

A small change of notional N_1 and strike K_1 changes the picture. If both are smaller than N_2 and K_2 respectively we may obtain two intersection points of the exercise and

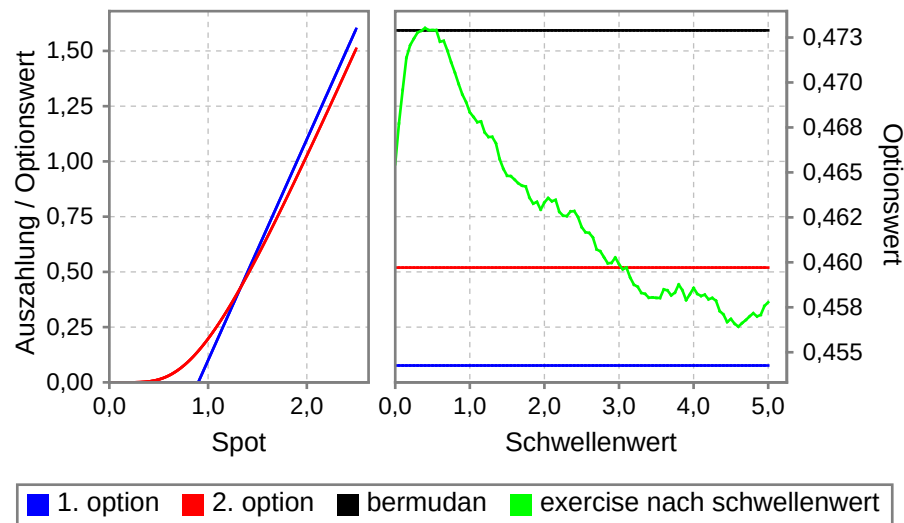


Figure 14.11.: Example of the successful optimization of the exercise criteria (intersection of the two price curves (left)). The graph on the right shows the Bermudan option value as a function of the exercise threshold λ .

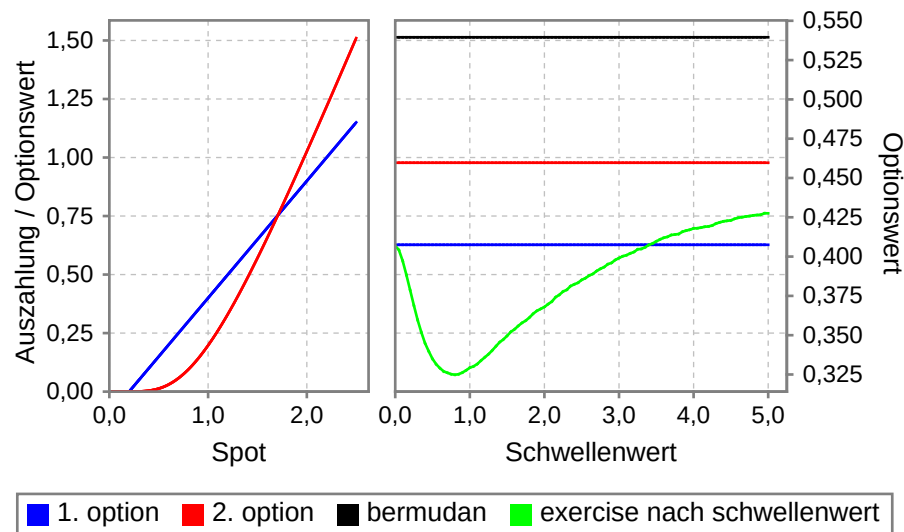


Figure 14.12.: Example of a failing optimization of the exercise criteria (intersection of the two price curves (left)). The graph on the right shows the Bermudan option value as a function of the exercise threshold λ .

continuation value. In T_1 it is optimal to exercise in between these two intersection points. Our simple exercise criteria cannot render this case. Optimizing the threshold parameter λ shows two maxima: the value of the two european options “exercise never” and “exercise always”. Both values are below the true Bermudan option value, see Figure 14.12, right.

The conclusion of this is example is that the choice of the exercise strategy has to be made carefully in accordance with the product. But this remark applies to any method in varying extends.

14.11.2.2. Disturbance of the Optimizer through Discontinuities and local Minima

The Monte Carlo Bermudan price calculated from the backward algorithm is a discontinuous function of the exercise criteria. The Bermudan price jumps if the exercise criteria $I(\omega, \lambda)$ changes sign for a given path ω . The price jumps by the difference of *exercise value* and *continuation value*. Even in the case of an optimal exercise criteria (i.e. $\lambda = \lambda_{\max}$) we see a jump in price since even then *exercise value* und *continuation value* will be different in general (at optimal exercise, only the expected (!) continuation value equals the exercise value).

As a function of λ , the price will not only exhibit discontinuities, but also small local maxima induced by them, see Figure 14.11 (right). These may prevent the optimizing algorithm from finding the global maxima.

However, for sufficiently many path the local maxima appear only on a small scale. The jumps in price will be of the order $O(\frac{1}{n})$, where n denotes the number of paths. Thus with a robust minimizer one would rarely encounter this problem. As example consider the case that the limit function (for $n \rightarrow \infty$) would satisfy an estimate of the form $V_{\max} - V(\lambda) > C(\lambda - \lambda_{\max})^2$, i.e. without the Monte-Carlo discontinuities no other local maxima or saddle point would exist. Then a bisection search on the disturbed function will miss the true maxima only by the order of a jump $O(\frac{1}{n})$.

14.11.3. Optimization of Exercise Strategy: A more general Formulation

There is a trivial generalization of the optimization method considered in Section 14.11.1:

- The exercise criteria will be given as a function of arbitrary $\mathcal{F}_{T_i}$ measurable random variables.
- The exercise criteria will be given as a function of a parameter vector $\lambda \in \mathbb{R}^k$.

Thus we replace the “true” exercise criteria used in the backward algorithm

$$\tilde{V}_{\text{underl},i}(T_i) < E^{\mathbb{Q}}(\tilde{U}_{i+1} | \mathcal{F}_{T_i})$$

by a function

$$I_i(\lambda, \omega) := f(B_{i,1}(\omega), \dots, B_{i,m}(\omega), \lambda),$$

where $B_{i,j}$ is a set of $\mathcal{F}_{T_i}$ measurable random variables and $\lambda \in \mathbb{R}^k$.

14.11.4. Comparison of Optimization Method and Regression Method

The difference between the optimization method and the regression method becomes apparent in Figure 14.7-14.8. While the regression method requires that the regression functions give a good fit to the conditional expectations across the whole domain of the independent variable, the optimization method only requires that the functional $I_i(\lambda)$ captures the exercise boundary. In Figures 14.7-14.8 the conditional expectation estimator is a curve, but the exercise boundary is given by two points only (the intersection of the conditional expectation estimator with the bisector).

Thus the optimization method can cope with far less parameters than the regression. On the other hand, as noted in the example above, it is far more important that the functional is adapted to the Bermudan product under consideration.

It is trivial to choose the map $I_i(\lambda)$ such that the optimal λ^* give the same or a better value than the least-square regression. If $B_{i,j}$ denote the basis functions used in the least-square regression for exercise date T_i we set

$$I_i(\lambda_1, \dots, \lambda_n) = (\lambda_1 B_{i,1} + \dots + \lambda_n B_{i,n}) - \tilde{V}_{\text{underl},i}(T_i).$$

Then we have that for $\lambda = \alpha^*$, where α^* is the regression parameter from the least-square regression, the exercise criteria agrees with the one of the least-square regression. This result however is rarely an advantage of the optimization method since in practice a high dimensional optimization does not represent an alternative.

14.12. Duality Method: Upper Bound for Bermudan Option Prices - The Method of Rogers

In diesem Abschnitt stellen wir eine Methode zur Bewertung von Bermudan Optionen vor, die, für verschwindenden Monte Carlo Fehler, gegen eine obere Preisschranke konvergiert, welche mit Ausschöpfung eines Testraumes von oben gegen den theoretischen Preis der Bermudan Option konvergiert. Zusammen mit den *Lower Bound Methods* (14.8 - 14.11.1) besteht so die Möglichkeit, modulo des Monte Carlo Fehlers, ein Intervall anzugeben, indem sich der theoretische Preis der Bermudan Option befindet.

Die Methode geht zurück auf Rogers [81], Haugh & Kogan [68], sowie ursprünglich Davis & Karatzas [57].

14.12.1. Foundations

Die in diesem Abschnitt vorgestellten Grundlagen werden erst in Abschnitt 14.12 bedeutsam. Zwar werden schon vorher einige der hier definierten Begriffe, wie Supermartingal oder Snell-Envelope, benutzt, diese sind jedoch für das Verständnis der Abschnitte 14.5 bis 14.11.1 nicht wesentlich.

Definition 180 (Supermartingal):

Der stochastische Prozess $\{X_t, \mathcal{F}_t ; 0 \leq t < \infty\}$ heißt *Supermartingal* bezüglich der Filtration $\mathcal{F}_t$ und des Maßes $\mathbb{P}$, falls

$$X_s \geq E(X_t | \mathcal{F}_s), \quad \mathbb{P}\text{-fast sicher,} \quad \forall 0 \leq s < t < \infty.$$

Definition 181 (Stoppzeit (zeit-diskrete Prozesse)):

Es sei $T : \Omega \rightarrow \{T_0, \dots, T_n\}$ eine Zufallsvariable. T heißt Stoppzeit (*Stopping Time*), falls

$$\{T = T_i\} \in \mathcal{F}_{T_i}. \quad (14.14)$$

Remark 182 (Stoppzeit): Äquivalent zu (14.14) ist

$$\{T \leq T_i\} \in \mathcal{F}_{T_i}, \quad (14.15)$$

da $\mathcal{F}_{T_j} \subset \mathcal{F}_{T_i}$ für $T_j < T_i$.



Interpretation: Eine Stoppzeit kann als mathematische Repräsentation einer Ausübestrategie für eine Bermudan Option interpretiert werden. Wird $T(\omega) = T_i$ als Zeitpunkt interpretiert, indem, gegeben den

Pfad ω , ein Ereignis eintritt (z.B. die Entscheidung bezüglich einer (optimalen) Ausübung einer Option), so stellt (14.15) die Forderung dar, dass die Informationsgrundlage für dieses Ereignis spätestens in T_i vorliegt (ausgedrückt durch die $\mathcal{F}_{T_i}$ -Messbarkeit).

Die Bedingung (14.15) verhindert also den *Foresightfehler*. Man vergleiche Abschnitt 14.6⁶. ◁|

Gegeben ein stochastischer Prozess und eine Stoppzeit, lässt sich der sogenannte gestoppte Prozess konstruieren:

Definition 183 (Gestoppter Prozess): ↗

Es sei X ein zeit-diskreter stochastischer Prozess und T eine Stoppzeit. Der in T gestoppte Prozess X^T von X ist definiert als

$$X^T(t, \omega) = \begin{cases} X(t, \omega) & t < T(\omega) \\ X(T(\omega), \omega) & t \geq T(\omega). \end{cases}$$

┘

Definition 184 (Snell Envelope): ↗

Es sei Z ein zeit-diskreter adaptierter Prozess und U definiert durch

$$\begin{aligned} U(T_n) &= Z(T_n) \\ U(T_i) &= \max(Z(T_i), E(U(T_{i+1}) | \mathcal{F}_{T_i})). \end{aligned}$$

Der Prozess U heißt *Snell Envelope* von Z . ┘

Lemma 185 (Snell Envelope): Der *Snell Envelope* ist das kleinste Supermartingal, dass Z dominiert.

Theorem 186 (Doob-Meyer Zerlegung): Es sei U ein (zeit-diskretes) Supermartingal. Dann existiert eine (eindeutige) Zerlegung

$$U(T_i) = M(T_i) - A(T_i),$$

wobei $M(T_i)$ ein Martingal und $A(T_i)$ ein vorhersehbarer nicht-fallender Prozess (d.h. $A(T_{i-1}, \omega) \leq A(T_i, \omega)$) mit $A(T_0) = 0$ ist.

Proof:

$$\begin{aligned} M(T_i) &:= M(T_{i-1}) + U_i - E(U_i | \mathcal{F}_{T_{i-1}}) & M(T_0) &= U(T_0) \\ A(T_i) &:= A(T_{i-1}) + U_{i-1} - E(U_i | \mathcal{F}_{T_{i-1}}) & A(T_0) &= 0 \end{aligned}$$

□|

⁶ Für das Beispiel aus Abschnitt 14.6 sind $T(\omega_1) = T(\omega_2) = T_1$ und $T(\omega_1) = T(\omega_2) = T_2$ Stoppzeiten, hingegen ist die super-optimale Ausübestrategie $T(\omega_1) = T_1, T(\omega_2) = T_2$ keine Stoppzeit.

14.12.2. American Option Evaluation as *Optimal Stopping Problem*

Wir betrachten, wie schon zuvor, eine *Bermudan Option*: Es seien $0 = T_0 < T_1 < \dots < T_n$. Es sei $V_{\text{underl}}(T_i)$, $i = 1, \dots, n-1$ eine Folge von $\mathcal{F}_{T_i}$ -messbaren Zufallsvariablen, die den Wert eines Finanzproduktes zum Zeitpunkt T_i darstellen, sowie rekursiv

$$V_{\text{berm}}(T_i) := \max(V_{\text{underl}}(T_i), E^{\mathbb{Q}^N}(V_{\text{berm}}(T_{i+1}) \cdot \frac{N(T_i)}{N(T_{i+1})} \mid \mathcal{F}_{T_i})),$$

wobei $(N, \mathbb{Q}^N)$ ein geeignetes Numéraire-Martingalmaß Paar sei. Ferner bezeichnen $\tilde{V}_U(T_i) := \frac{V_{\text{underl}}(T_i)}{N(T_i)}$, $\tilde{V}_B(T_i) := \frac{V_{\text{berm}}(T_i)}{N(T_i)}$ die zugehörigen N -relativen Preise, d.h. es ist

$$\tilde{V}_B(T_i) := \max(\tilde{V}_U(T_i), E^{\mathbb{Q}^N}(\tilde{V}_B(T_{i+1}) \mid \mathcal{F}_{T_i})).$$

Theorem 187 (Rogers (zeit-kontinuierliche Version)): Es sei

$$V_{\text{amer}}(0) := \sup_{T \text{ Stoppzeit}} E(\tilde{V}_U(T) \mid \mathcal{F}_0).$$

Dann gilt

$$V_{\text{amer}}(0) = \inf_{M \in H_0^1} E(\sup_{0 \leq t \leq T} (\tilde{V}_U(t) - M(t)) \mid \mathcal{F}_0),$$

wobei H_0^1 die Menge aller Martingale M mit $\sup_{0 \leq t \leq T} |M(t)| \in L_1(\Omega)$ und $M(0) = 0$ sei.

Da wir nur die Bewertung von *Bermudan* Optionen betrachten, d.h. Optionen mit endlich vielen Ausübungszeiten, geben wir eine zeitdiskrete Version dieses Satzes:

Theorem 188 (Rogers (zeit-diskrete Version)): Es sei

$$\tilde{V}_B(T_i) := \max(\tilde{V}_U(T_i), E(\tilde{V}_B(T_{i+1}) \mid \mathcal{F}_{T_i})).$$

Dann gilt

$$\tilde{V}_B(T_0) = \inf_{M \in H_0^1} E(\max_{T_i} (\tilde{V}_U(T_i) - M(T_i)) \mid \mathcal{F}_0), \quad (14.16)$$

wobei H_0^1 die Menge aller (zeit-diskreten) Martingale M mit

$$\begin{aligned} M(T_0) &= 0 \\ M(T_i) &= E(M(T_{i+1}) \mid \mathcal{F}_{T_i}) \quad M(T_i) \in L_1(\Omega) \end{aligned}$$

ist und $\max_{T_i} (\tilde{V}_U(T_i) - M(T_i))$ pfadweise zu verstehen ist, d.h.

$$\max_{T_i} (\tilde{V}_U(T_i) - M(T_i))[\omega] = \max_{T_i} (\tilde{V}_U(T_i, \omega) - M(T_i, \omega))$$

(und *nicht* im Sinne eines Maximums über alle Stoppzeiten).

Remark 189 (zu Satz 188): Bemerkenswert an Satz 188 ist zunächst, dass das Maximum pfadweise angewendet wird. Angewendet auf V_{underl} würde $\max_{T_i}(V_{\text{underl}}(T_i))$ einem *foresight* Fehler entsprechen.⁷ Ein *foresight* Fehler wird durch das Martingale M verhindert, wie in Lemma 190 und dessen Beweis klar wird.

Der Erwartungswert in (14.16) wird nur auf $\mathcal{F}_{T_0} = \mathcal{F}_0$ bedingt, d.h. es ist nicht nötig, eine (spätere) bedingte Erwartung zu berechnen. Hingegen muss über alle Martingale M minimiert werden, was zunächst ebenso komplex ist wie die Berechnung der optimalen Ausübestrategie (durch eine Maximierung über alle Stoppzeiten). Allerdings liefert Satz 188 eine obere Schranke für den Wert der Bermudan Option:

$$V_{\text{berm}}(T_0) \leq E^{\mathbb{Q}^N} \left(\max_{T_i} (V_{\text{underl}}(T_i) \cdot \frac{N(0)}{N(T_i)} - M(T_i) \mid \mathcal{F}_0) \right), \quad M \in H_0^1,$$

die mit geeigneter Wahl von M dem Optionspreis $V_{\text{berm}}(T_0)$ beliebig nahe kommt.

Lemma 190 (Eliminierung des Foresight): Es seien $\tilde{V}_U(T_i)$, $\tilde{V}_B(T_i)$ wie zuvor, d.h.

$$\tilde{V}_B(T_i) := \max_{T_i} (\tilde{V}_U(T_i), E(\tilde{V}_B(T_{i+1}) \mid \mathcal{F}_{T_i})), \quad (14.17)$$

sowie M entsprechend der Doob-Meyer Zerlegung von $\tilde{V}_B(T_i)$

$$\begin{aligned} M(T_0) &:= 0, \\ M(T_i) &:= M(T_{i-1}) + \tilde{V}_B(T_i) - E(\tilde{V}_B(T_{i+1}) \mid \mathcal{F}_{T_i}). \end{aligned} \quad (14.18)$$

Es sei $T^{\text{opt}} : \Omega \rightarrow \{T_1, \dots, T_n\}$ die durch

$$T^{\text{opt}}(\omega) := \min\{T_j : \tilde{V}_U(T_j, \omega) \geq E(\tilde{V}_B(T_{j+1}) \mid \mathcal{F}_{T_j})[\omega]\}$$

definierte (optimale) Stoppzeit (es ist dann $\tilde{V}_B(T_0) = E(\tilde{V}_U(T^{\text{opt}}) \mid \mathcal{F}_{T_0})$). Dann gilt (pfadweise!)

$$\max_{T_j} (\tilde{V}_U(T_j, \omega) - M(T_j, \omega)) = \tilde{V}_U(T^{\text{opt}}(\omega), \omega) - M(T^{\text{opt}}, \omega).$$

Proof: Wir führen den Beweis elementar nur mit (14.17) und (14.18). Es sei $\omega \in \Omega$. Per Definition ist

$$\tilde{V}_B(T_j, \omega) = E(\tilde{V}_B(T_{j+1}) \mid \mathcal{F}_{T_j})[\omega] \quad \text{für } T_j < T^{\text{opt}}(\omega) \quad (14.19)$$

$$\tilde{V}_B(T_j, \omega) \geq E(\tilde{V}_B(T_{j+1}) \mid \mathcal{F}_{T_j})[\omega] \quad \text{für } T_j \geq T^{\text{opt}}(\omega) \quad (14.20)$$

sowie

$$\tilde{V}_B(T_i, \omega) = \tilde{V}_U(T_i, \omega) \quad \text{für } T_i = T^{\text{opt}}(\omega). \quad (14.21)$$

⁷ Vergleiche Abschnitt 14.6

1. Für $T_i = T^{\text{opt}}(\omega)$ und $T_j < T_i$ ist

- $\tilde{V}_U(T_i, \omega) - M(T_i, \omega) = \tilde{V}_B(T_i, \omega) - M(T_i, \omega)$
- $\begin{aligned} \tilde{V}_B(T_{j+1}, \omega) - M(T_{j+1}, \omega) \\ &= \tilde{V}_B(\overline{T_{j+1}, \omega}) - M(T_j, \omega) - \tilde{V}_B(\overline{T_{j+1}, \omega}) + E(\tilde{V}_B(T_{j+1}, \omega) \mid \mathcal{F}_{T_j})[\omega] \\ &= \tilde{V}_B(T_j, \omega) - M(T_j, \omega) \end{aligned}$
- $\tilde{V}_B(T_j, \omega) - M(T_j, \omega) \geq \tilde{V}_U(T_j, \omega) - M(T_j, \omega)$

und somit insgesamt für $T_j \leq T_i$:

$$\tilde{V}_U(T_j, \omega) - M(T_j, \omega) \leq \tilde{V}_U(T_i, \omega) - M(T_i, \omega)$$

2. Für $T_i = T^{\text{opt}}(\omega)$ und $T_j > T_i$ ist

- $\begin{aligned} \tilde{V}_U(T_i, \omega) - M(T_i, \omega) \\ &= \tilde{V}_U(T_i, \omega) - M(T_{i+1}, \omega) + \tilde{V}_B(T_{j+1}, \omega) - E(\tilde{V}_B(T_{j+1}) \mid \mathcal{F}_{T_j})[\omega] \\ &= \tilde{V}_B(T_i, \omega) - M(T_{i+1}, \omega) + \tilde{V}_B(T_{j+1}, \omega) - E(\tilde{V}_B(T_{j+1}) \mid \mathcal{F}_{T_j})[\omega] \\ &\geq \tilde{V}_B(T_{i+1}, \omega) - M(T_{i+1}, \omega) \end{aligned}$
- $\begin{aligned} \tilde{V}_B(T_j, \omega) - M(T_j, \omega) \\ &= \tilde{V}_B(T_j, \omega) - M(T_{j+1}, \omega) + \tilde{V}_B(T_{j+1}, \omega) - E(\tilde{V}_B(T_{j+1}, \cdot) \mid \mathcal{F}_{T_j})[\omega] \\ &\geq \tilde{V}_B(T_{j+1}, \omega) - M(T_{j+1}, \omega) \end{aligned}$
- $\tilde{V}_B(T_{j+1}, \omega) - M(T_{j+1}, \omega) \geq \tilde{V}_U(T_{j+1}, \omega) - M(T_{j+1}, \omega)$

und somit insgesamt für $T_j \geq T_i$

$$\tilde{V}_U(T_i, \omega) - M(T_i, \omega) \geq \tilde{V}_U(T_j, \omega) - M(T_j, \omega)$$

d.h. es gilt für alle T_j

$$\tilde{V}_U(T_i, \omega) - M(T_i, \omega) \geq \tilde{V}_U(T_j, \omega) - M(T_j, \omega).$$

□

Remark 191 (Foresight Fehler): Das Lemma 190 zeigt, dass das Martingal M den Foresight Fehler eliminiert. Es gilt

$$\tilde{V}_U(T_i) - M(T_i) \geq \tilde{V}_U(T_j) - M(T_j) \quad \forall T_j$$

jedoch ist im allgemeinen

$$\tilde{V}_U(T_i) \not\geq \tilde{V}_U(T_j) \quad \forall T_j.$$

Lemma 192 (Optimal Stopping): Es ist

$$E(\tilde{V}_U(T^{\text{opt}}) | \mathcal{F}_{T_0}) = E(\max_{T_j} (\tilde{V}_U(T_j) - M(T_j)) | \mathcal{F}_{T_0}).$$

Proof: Es ist

$$\begin{aligned} E(\tilde{V}_U(T^{\text{opt}}) | \mathcal{F}_{T_0}) &\stackrel{\text{Def. } T^{\text{opt}}}{=} E\left(\sup_{T \text{ stoppingtime}} \tilde{V}_U(T) | \mathcal{F}_{T_0}\right) \\ &= E\left(\sup_{T \text{ stoppingtime}} \tilde{V}_U(T) - M(T) | \mathcal{F}_{T_0}\right) \\ &\stackrel{\text{La. 190}}{=} E(\max_{T_i} \tilde{V}_U(T_i) - M(T_i) | \mathcal{F}_{T_0}) \end{aligned}$$

□

14.13. Primal-Dual Method: Upper and Lower Bound

Die Bestimmung des Ausübekriteriums und damit des Preises einer Bermudan Option durch eine (approximative) Bestimmung der bedingten Erwartung wird als *primal* bezeichnet. Die Bestimmung des Ausübekriteriums durch die Bestimmung der Stoppzeit nach 14.12 wird als *dual* bezeichnet. Aufgrund des Ergebnisses von Lemma 190 lassen sich nun beide Methoden kombinieren. Nach Satz 188 und Lemma 190 ist

$$\tilde{V}_B(T_0) = E(\max_{T_i} (\tilde{V}_U(T_i) - M(T_i)) | \mathcal{F}_0), \quad (14.22)$$

mit

$$M(T_i) := M(T_{i-1}) + U_i - E(U_i | \mathcal{F}_{T_{i-1}}) \quad M(T_0) = U(T_0). \quad (14.23)$$

Die bedingte Erwartung in (14.23) lässt sich nun mit einem beliebigen (*primalen*) Verfahren schätzen, z.B. bei der Bestimmung einer unteren Schranke des Wertes der Bermudan Option. Hiermit liefert (14.22) sofort eine zugehörige obere Schranke.

Man vergleiche hierzu auch [43].



Further Reading: Die Bücher [14] und [16] und der Artikel [51] behandeln Monte-Carlo Methoden für Bewertungsprobleme. In [56] werden theoretische Aspekte der Regressionsmethoden beleuchtet (z.B. Konvergenz). Die Optimierungsmethode für Bermudan Swaptions findet sich in [42], eine Primal-Dual Methode in [43]. <



Experiment: Unter <http://www.christian-fries.de/finmath/applets/BermudanStockOptionPricing.html> kann die Bewertung einer Bermudan Option auf eine Aktie die einem Black-Scholes Modell folgt studiert werden. Verschiedene Bewertungsverfahren sind wählbar. <|

CHAPTER 15

Sensitivities (Partial Derivatives) of Monte-Carlo Prices

section under construction.

15.1. Introduction

The technique of risk neutral pricing, i.e. the change towards the martingale measure, allows to calculate the cost of a (selffinancing) replication portfolio to be expressed as a expectation. It is not required to determine the replication portfolio itself. However, once a pricing formula or pricing algorithm (e.g. a Monte-Carlo simulation) has been derived, the replication portfolio may be given in terms of the partial derivatives of the price with respect to current model parameters (like the initial values of the underlyings).¹ The partial derivatives of the price with respect to the model parameters are also called sensitivities or *Greeks*. They are important to assess the risk of a financial product, see also Chapter 7.

For complex products, like Bermudan options, an analytic pricing formula is usually not available. The pricing has to be done numerically. Under a high dimensional model, like the LIBOR market model, the numerical method of choice usually is Monte-Carlo simulation. Given that, we investigate the numerical calculation of sensitivities (partial derivatives) of Monte-Carlo prices.

The most simple way to calculate a derivative is finite differences. Applying finite differences the result of a Monte-Carlo algorithm may give unstable or wrong results.

15.2. Problem Description

We consider a pricing algorithm that uses Monte-Carlo simulation to calculate a price of a financial product as the expectation of the Numéraire-relative value under an

¹ Note that all market parameters enter into model parameters via the calibration of the model.

equivalent martingale measure $\mathbb{Q}$

$$V(t_0) = N(t_0) \cdot \mathbb{E}^{\mathbb{Q}} \left(\frac{V(T)}{N(T)} \mid \mathcal{F}_{t_0} \right).$$

We are interested in the calculation of a partial derivative of $V(t_0)$ with respect to some model parameter, e.g. the initial values of the underlying ($\rightarrow$ delta), the volatility ($\rightarrow$ vega), etc.

Since we treat this problem as a general numerical problem, not necessarily related to derivative pricing, we do not consider a specific model and use a notation that is slightly more general. To fix notation, let us restate Monte-Carlo pricing first:

15.2.1. Pricing using Monte Carlo Simulation

Assume that our model is given as a stochastic process X , for example an Itô process

$$dX = \mu dt + \sigma \cdot dW(t)$$

modeling our model primitives, like functions of the the underlyings (e.g. financial products (stocks) or rates (forward rates, swap rates, fx rates)). For example, for the Black-Scholes model we would have $X = (\log(S), \log(B))$. Let $X^*(t_i)$ denote an approximation of $X(t_i)$ generated by some (time-)discretization scheme, e.g. an Euler scheme

$$X^*(t_{i+1}) = X^*(t_i) + \mu(t_i)\Delta t_i + \sigma(t_i) \cdot \Delta W(t_i)$$

or one of the more advanced schemes². We assume that our financial product depends only on realizations of X as some finitely many time points, i.e. we assume that the risk neutral pricing of the financial product may be expressed as the expectation (with respect to the pricing measure) of a function f of some realizations $Y := (X(t_0), X(t_1), \dots, X(t_m))$. This is true for many products (e.g. Bermudan options). If these are approximated through the realizations of the numerical scheme we have:

$$\mathbb{E}(f(Y) \mid \mathcal{F}_{t_0}) \approx \mathbb{E}(f(Y^*) \mid \mathcal{F}_{t_0}) = \mathbb{E}(f((X^*(t_0), X^*(t_1), \dots, X^*(t_m))) \mid \mathcal{F}_{t_0}).$$

Here f denotes the Numéraire relative payoff function.

The Monte-Carlo pricing consists of the averaging over (in general) equidistributed sample path ω_i , $i = 1, \dots, n$

$$\mathbb{E}(f(Y^*) \mid \mathcal{F}_{t_0}) \approx \hat{\mathbb{E}}(f(Y^*) \mid \mathcal{F}_{t_0}) := \frac{1}{n} \sum_{i=1}^n f(Y^*(\omega_i)).$$

To summarize: we have two approximation steps involved: The first approximated the time-continuous process by a time-discrete process. The second approximated the expectation by a Monte-Carlo simulation of n sample paths. This is somehow the minimum requirement to have the pricing implemented as a Monte-Carlo simulation.

² For alternative schemes see e.g. [61, 19]

15.2.2. Sensitivities from Monte Carlo Pricing

Assume that θ denotes some model parameter³ or a parametrization of a generic market data movement and let Y_θ denote the model realizations depending on that parameter. Let us further assume that ϕ_{Y_θ} denote the probability density of Y_θ . Then the analytic calculation of the sensitivity is given by

$$\frac{\partial}{\partial \theta} E(f(Y_\theta) | \mathcal{F}_{t_0}) = \frac{\partial}{\partial \theta} \int_{\mathbb{R}^m} f(y) \phi_{Y_\theta}(y) dy.$$

While the payoff f may be discontinuous the density in general is a smooth function of θ . In that case the expectation $E(f(Y_\theta) | \mathcal{F}_{t_0})$ (the price) is a smooth function of θ too. The price inherits the smoothness of ϕ_{Y_θ} .

The calculation of sensitivities using finite differences on a Monte-Carlo based pricing algorithm is known to exhibit instabilities if the payoff function is not smooth enough, e.g. if the payoff exhibits discontinuities as for a digital option. The difficulties arise when we consider the Monte Carlo approximation. It inherits the regularity of the payoff f not that of the density ϕ :

$$\hat{E}(f(Y_\theta) | \mathcal{F}_{t_0}) = \frac{1}{n} \sum_{i=1}^n f(Y_\theta(\omega_i)).$$

So while $E(f(Y_\theta) | \mathcal{F}_{t_0})$ may be smooth in θ the Monte Carlo approximation $\hat{E}(f(Y_\theta) | \mathcal{F}_{t_0})$ may have discontinuities. In that case a finite difference approximation of the derivative applied to the Monte Carlo pricing will perform poorly.

15.2.3. Example: The Linear and the Discontinuous Payout

The challenge in calculating Monte-Carlo sensitivities becomes obvious if we consider two very simple examples:

15.2.3.1. Linear Payout

First consider a linear payout, say

$$f(X(T)) = a \cdot X(T) + b.$$

The (discounted) payout depends only on the time T realization of X (as one would have for an European option). Let $Y_\theta(\omega) := X(T, \omega, \theta)$, where θ denotes some model parameter. The partial derivative of the Monte-Carlo value of the payout with respect to θ is

$$\frac{\partial}{\partial \theta} \hat{E}(f(Y_\theta) | \mathcal{F}_{t_0}) = \frac{1}{n} \sum_{i=1}^n \frac{\partial}{\partial \theta} f(Y_\theta(\omega_i)) = \frac{1}{n} \sum_{i=1}^n a \cdot \frac{\partial}{\partial \theta} Y_\theta(\omega_i)$$

³ So for *delta* θ is an initial value $X(0)$, for *vega* θ denotes a volatility, etc.

Obviously the accuracy of the Monte-Carlo approximation depends on the variance of $\frac{\partial Y_\theta}{\partial \theta}$ only. For the case, when $\frac{\partial}{\partial \theta} Y_\theta(\omega_i)$ does not depend on ω_i , then the Monte-Carlo approximation gives the exact value of the partial derivative, even if we use only a single path.

15.2.3.2. Discontinuous Payout

Next, consider a discontinuous payout, say

$$f(X(T)) = \begin{cases} 1 & \text{if } X(T) > K \\ 0 & \text{else.} \end{cases}$$

Analytically we have from $Y_{\theta+h} = Y_\theta + \frac{\partial Y_\theta}{\partial \theta} \cdot h + O(h^2)$ and

$$\mathbb{E}^\mathbb{Q}(f(Y_{\theta+h}) | \mathcal{F}_{t_0}) = \mathbb{Q}(\{Y_\theta > K - \frac{\partial Y_\theta}{\partial \theta} \cdot h - O(h^2)\}) = \int_{K - \frac{\partial Y_\theta}{\partial \theta} \cdot h - O(h^2)}^{\infty} \phi_{Y_\theta}(y) dy$$

that

$$\lim_{h \rightarrow 0} \frac{1}{2h} (\mathbb{E}^\mathbb{Q}(f(Y_{\theta+h}) | \mathcal{F}_{t_0}) - \mathbb{E}^\mathbb{Q}(f(Y_{\theta-h}) | \mathcal{F}_{t_0})) = \phi_{Y_\theta}(K) \cdot \frac{\partial Y_\theta}{\partial \theta}.$$

However, the partial derivative of the Monte-Carlo value of the payout is

$$\frac{\partial}{\partial \theta} \hat{\mathbb{E}}(f(Y_\theta) | \mathcal{F}_{t_0}) = \frac{1}{n} \sum_{i=1}^n \frac{\partial}{\partial \theta} f(Y_\theta(\omega_i)) = 0 \text{ assuming that } Y_\theta(\omega_i) \neq K \text{ for all } i.$$

Thus, here, the partial derivative of the Monte-Carlo value is always wrong.

15.2.4. Example: Trigger Products

The two simple examples above suggest that a finite difference approximation of a Monte-Carlo price works well if the payout is smooth, but fails if the payout exhibits discontinuities. The problem becomes a bit more subtle if we consider products where the discontinuous behaviour is just one part of the payout which, in addition, may also be of more complex nature. Consider for example the [auto cap](#). For given times $T_1, \dots, T_n$ the auto cap pays at each payment date T_{i+1} the payout of a caplet $\max(L(T_i, T_{i+1}; T_i) - K_i, 0) \cdot (T_{i+1} - T_i)$, but does so only if the number of non-zero payments up to T_i is less than some $n_{\max \text{Ex}}$. This latter condition represents a trigger which makes the otherwise continuous payoff discontinuous, see Figure [15.1](#).

15.3. Generic Sensitivities: Bumping the Model

The finite difference approximation calculates the sensitivity by

$$\frac{\partial}{\partial \theta} \mathbb{E}(f(Y_\theta) | \mathcal{F}_{t_0}) \approx \frac{\mathbb{E}(f(Y_{\theta+h}) | \mathcal{F}_{t_0}) - \mathbb{E}(f(Y_{\theta-h}) | \mathcal{F}_{t_0})}{2h}.$$

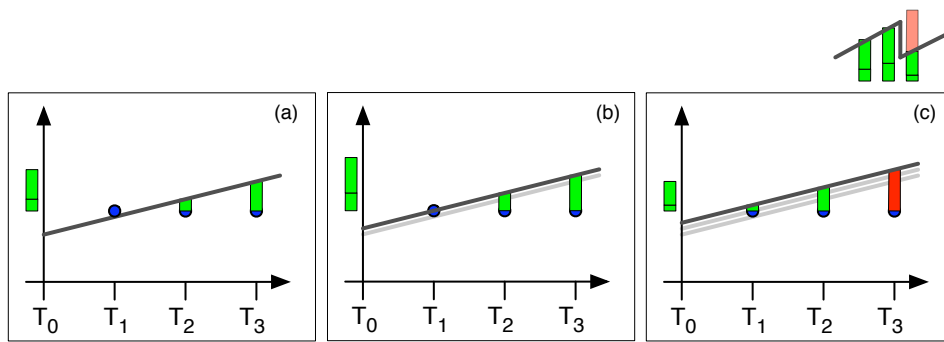


Figure 15.1.: The payoff of an auto cap paying a maximum of two out of three caplets, considered under a parallel shift of the interest rate curve (black line). The strike rate is depicted by a blue dot, a positive payout is marked in green: In scenario a) and b) the first caplet does not lead to a positive payout while the second and third caplet generate a positive payout. The shift of the interest rate curve from a) to b) changes the payout continuously. In scenario c) the first caplet leads to a positive payout. Since the auto cap is limited to two positive payouts the payout of third caplet is lost as soon as the first caplet pays a positive amount. Thus, from scenario b) to c) the payout of the auto cap changes discontinuously.

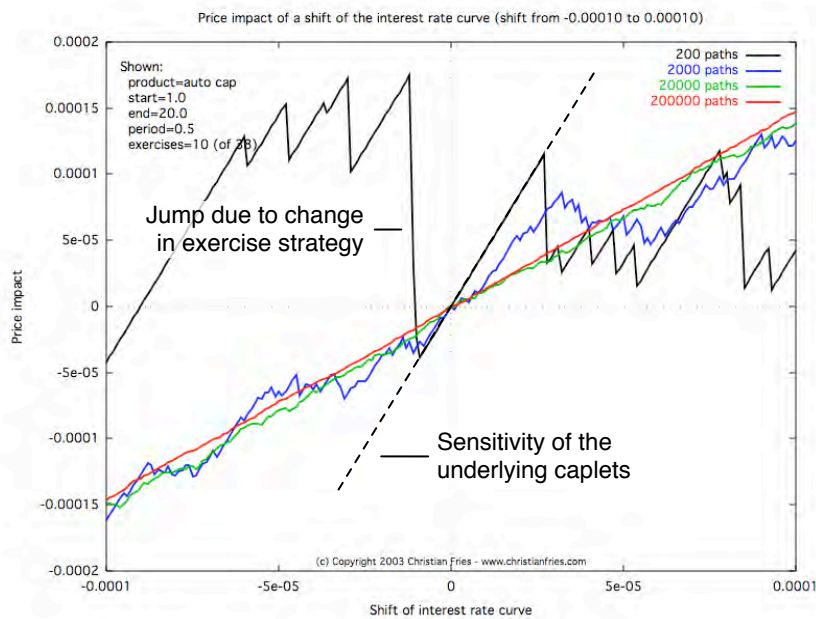


Figure 15.2.: The value of an auto cap as a function of the shift size of a parallel shift of the interest rate curve. Using only a small number of paths, a small shift does not lead to a change of the exercise strategy. The price change is driven by the sensitivity of the underlying caplets. Thus, for small shifts one might be tempted to call the sensitivity stable. For a larger shift the exercise strategy changes on some path, leading to a jump in payoff.

This *brute force finite difference* calculation of sensitivities is sometimes referred to as *bumping the model*. Bumping the model has a charming advantage: If you keep your model and your pricing code separated (a design pattern one should always consider) then you may implement a generic code for generating sensitivities by feeding the pricing code with differently bumped models. In other words:

Once the pricing code is written, all sensitivities are available. (15.1)

It seems as if you get sensitivities almost for free (i.e. without any effort in modeling and implementation) and the only price you pay is a doubling of calculation time compared to pricing. However it is known that applying such a finite difference approximation to a Monte-Carlo implementation will often result in extremely large Monte-Carlo errors. Especially if the payout function of the derivative is discontinuous this Monte-Carlo error even tends to infinity as h tends to zero. And discontinuous payout are present whenever a trigger feature is present.

In general sensitivities in Monte-Carlo is known as a challenge. Numerous methods have been proposed to handle sensitivities in Monte Carlo, among them the Likelihood Ratio [53] and the application of Malliavin calculus, which has drawn more attention recently, [58]. These methods improve the robustness of sensitivities, but require more information.

It appears as if the measures you have to take to improve Monte-Carlo sensitivities will lose the advantage (15.1) of *bumping the model*. Later, we will present a method (which is also an implementation design pattern) that gives you the possibility to calculate sensitivities through *bumping the model* while providing the accuracy and robustness achieved by the likelihood ratio or Malliavin calculus approach. The method is essentially a likelihood ratio reconsidered on the level of the numerical scheme.

There are basically two different method to calculate sensitivities in Monte-Carlo:

- the pathwise method, which differentiates the payout on every simulation path, see 15.5
- the likelihood ratio method, which differentiates the probability density, see 15.6.

Numerically these two methods may be realized as

- (traditional) finite differences, see 15.4
- finite differences applied to a proxy simulation scheme.

However, a proxy simulation scheme is much more powerful design, see Chapter 16. It is also possible to mix the two approaches by considering a partial proxy simulation scheme [60].

In the following we will present the different methods for calculating sensitivities in Monte-Carlo simulations. Each section starts with a short description in term of the approximating formula and gives the methods requirements and properties in terms of bullet points. We assume that a Monte-Carlo pricing algorithm has been implemented and we mention only the additional requirements compared to the pricing.

15.4. Sensitivities by Finite Differences

The finite difference approximation is given by

$$\begin{aligned}\frac{\partial}{\partial \theta} \mathbb{E}^{\mathbb{Q}}(f(Y_{\theta}) | \mathcal{F}_{t_0}) &\approx \frac{1}{2h} (\mathbb{E}^{\mathbb{Q}}(f(Y_{\theta+h}) | \mathcal{F}_{t_0}) - \mathbb{E}^{\mathbb{Q}}(f(Y_{\theta-h}) | \mathcal{F}_{t_0})) \\ &\approx \frac{1}{2h} (\hat{\mathbb{E}}^{\mathbb{Q}}(f(Y_{\theta+h}) | \mathcal{F}_{t_0}) - \hat{\mathbb{E}}^{\mathbb{Q}}(f(Y_{\theta-h}) | \mathcal{F}_{t_0})) \\ &= \frac{1}{n} \sum_{i=1}^n \frac{1}{2h} (f(Y_{\theta+h}(\omega_i)) - f(Y_{\theta-h}(\omega_i)))\end{aligned}$$

Requirements

- No additional information from the model sde X
- No additional information from the simulation scheme $X^*(t_{i+1})$
- No additional information from the payout f
- No additional information on the nature of θ ($\Rightarrow$ generic sensitivities)

Properties

- Biased derivative for *large* h due to finite difference of order h
- Extremely large variance for discontinuous payouts and *small* h (order h^{-1})

The most important feature of finite differences is there genericity. Once the the pricing code has been written all kinds of sensitivities may be calculated.

For smooth payouts we have that the finite difference approximation converges to the derivative for $h \rightarrow 0$. Thus, if the payout is smooth, small shift sizes h are favorable. Using large h the approximation of the derivative is biased.

For discontinuous payouts we have that for $h \rightarrow 0$ the finite difference of the Monte-Carlo price does not converge to the derivative of the Monte-Carlo price. The reason is that for discontinuous payouts the Monte-Carlo approximation ($n \rightarrow \infty$) and the approximation of the derivative ($h \rightarrow 0$) are not interchangeable.

For discontinuous payouts finite differences with a fixed, small shift size h perform poorly. The contribution of a discontinuity to the sensitivity may be calculated analytically. It is the jump size multiplied by the probability density at the discontinuity. Finite differences resolve this contribution only through those sample paths which fall at a neighborhood around the discontinuity, having the width of the shift size. Thus, if the shift size is small the discontinuity is resolved by few points, ultimately resulting in a large Monte-Carlo error. For discontinuous payouts large shift sizes are favorable. However, if the shift size is large the derivative becomes biased by second order effects (if present).

Since finite difference do not require anything more than a given pricing algorithm, we are tempted to apply it to any product for which a Monte-Carlo pricing may be calculated. If the product exhibits discontinuities in the payout the finite difference approximation tends to be unreliable and a careful analysis of the Monte-Carlo error for a given shift size h has to be performed.

15.4.1. Example: Finite Differences applied to Smooth and Discontinuous Payout

We consider a finite difference approximation of the partial derivative for the case of the linear payout $f(X(T)) = a \cdot X(T) + b$ from Section 15.2.3.1. We have

$$\begin{aligned} \frac{\partial}{\partial \theta} \mathbb{E}^{\mathbb{Q}}(f(Y_{\theta}) | \mathcal{F}_{t_0}) &\approx \frac{1}{2h} (\mathbb{E}^{\mathbb{Q}}(f(Y_{\theta+h}) | \mathcal{F}_{t_0}) - \mathbb{E}^{\mathbb{Q}}(f(Y_{\theta-h}) | \mathcal{F}_{t_0})) \\ &\approx \frac{1}{2h} (\hat{\mathbb{E}}^{\mathbb{Q}}(f(Y_{\theta+h}) | \mathcal{F}_{t_0}) - \hat{\mathbb{E}}^{\mathbb{Q}}(f(Y_{\theta-h}) | \mathcal{F}_{t_0})) \\ &= \frac{1}{n} \sum_{i=1}^n \frac{1}{2h} (f(Y_{\theta+h}(\omega_i)) - f(Y_{\theta-h}(\omega_i))) \\ &= \frac{1}{n} \sum_{i=1}^n a \cdot \frac{1}{2h} (Y_{\theta+h}(\omega_i) - Y_{\theta-h}(\omega_i)), \end{aligned}$$

which is a good approximation if $\frac{\partial}{\partial \theta} Y_{\theta}(\omega_i) \approx \frac{1}{2h} (Y_{\theta+h}(\omega_i) - Y_{\theta-h}(\omega_i))$. This is usually the case and throughout this chapter we assume that the model is such that its realizations $Y_{\theta}(\omega_i)$ are smooth in the model parameters θ .

For the discontinuous payout $f(X(T)) = 1$ if $X(T) > K$ and $f(X(T)) = 0$ else, considered in Section 15.2.3.2, we have

$$\begin{aligned} \frac{\partial}{\partial \theta} \mathbb{E}^{\mathbb{Q}}(f(Y_{\theta}) | \mathcal{F}_{t_0}) &\approx \frac{1}{2h} (\mathbb{E}^{\mathbb{Q}}(f(Y_{\theta+h}) | \mathcal{F}_{t_0}) - \mathbb{E}^{\mathbb{Q}}(f(Y_{\theta-h}) | \mathcal{F}_{t_0})) \\ &\approx \frac{1}{2h} (\hat{\mathbb{E}}^{\mathbb{Q}}(f(Y_{\theta+h}) | \mathcal{F}_{t_0}) - \hat{\mathbb{E}}^{\mathbb{Q}}(f(Y_{\theta-h}) | \mathcal{F}_{t_0})) \\ &= \frac{1}{n} \sum_{i=1}^n \frac{1}{2h} (f(Y_{\theta+h}(\omega_i)) - f(Y_{\theta-h}(\omega_i))) \\ &= \frac{1}{n} \sum_{i=1}^n \frac{1}{2h} \begin{cases} 1 & \text{if } Y_{\theta-h}(\omega_i) < K < Y_{\theta+h}(\omega_i) \\ -1 & \text{if } Y_{\theta-h}(\omega_i) > K > Y_{\theta+h}(\omega_i) \\ 0 & \text{else.} \end{cases} \end{aligned}$$

This is a valid approximation, but it has a large Monte-Carlo variance, since the true value is sampled by 0 and $\frac{1}{2h}$ occurring in the appropriate frequency. If h gets smaller then we have to represent true value by a sampling of 0 and a very large constant.

Simplified Example: Assume for simplicity that Y_θ is linear in θ , i.e. we have $Y_{\theta+h} = Y_\theta + \frac{\partial Y_\theta}{\partial \theta} \cdot h$ and thus

$$\begin{aligned} \frac{f(Y_{\theta+h}(\omega_i)) - f(Y_{\theta-h}(\omega_i))}{2h} &= \begin{cases} \frac{1}{2h} & \text{if } Y_{\theta-h}(\omega_i) < K < Y_{\theta+h}(\omega_i) \\ \frac{-1}{2h} & \text{if } Y_{\theta-h}(\omega_i) > K > Y_{\theta+h}(\omega_i) \\ 0 & \text{else.} \end{cases} \\ &= \begin{cases} \frac{\text{sign } \frac{\partial Y_\theta}{\partial \theta}}{2h} & \text{if } Y_\theta(\omega_i) \in [K - \epsilon, K + \epsilon] \\ 0 & \text{else.} \end{cases} \end{aligned}$$

where $\epsilon := \left| \frac{\partial Y_\theta}{\partial \theta} \right| \cdot h$. For the probability we have

$$q := \mathbb{Q}(Y_\theta \in [K - \epsilon, K + \epsilon]) \approx \phi_{Y_\theta}(K) \cdot 2\epsilon = \phi_{Y_\theta}(K) \cdot \left| \frac{\partial Y_\theta}{\partial \theta} \right| \cdot 2h.$$

In other words: We are sampling the partial derivative of the expectation by a binomial experiment:

$$\frac{\text{sign } \frac{\partial Y_\theta}{\partial \theta}}{2h} \quad \text{with probability } q \quad \text{and} \quad 0 \quad \text{with probability } 1 - q.$$

The expectation of this binomial experiment is

$$\frac{\text{sign } \frac{\partial Y_\theta}{\partial \theta}}{2h} \cdot q + 0 \cdot (1 - q) \approx \phi_{Y_\theta}(K) \cdot \frac{\partial Y_\theta}{\partial \theta},$$

which is the desired analytic value for the finite difference approximation as $h \rightarrow 0$. The variance of the binomial experiment is

$$\left(\frac{1}{2h} \right)^2 \cdot q \cdot (1 - q) \approx \phi_{Y_\theta}(K) \cdot \frac{\partial Y_\theta}{\partial \theta} \cdot (1 - q) \cdot \frac{1}{2h} = O\left(\frac{1}{2h} \right),$$

which explodes as $h \rightarrow 0$.

15.5. Sensitivities by Pathwise Differentiation

The pathwise differentiation method is given by

$$\begin{aligned} \frac{\partial}{\partial \theta} \mathbb{E}^\mathbb{Q}(f(Y(\theta)) | \mathcal{F}_{t_0}) &= \frac{\partial}{\partial \theta} \int_\Omega f(Y(\omega, \theta)) d\mathbb{Q}(\omega) = \int_\Omega \frac{\partial}{\partial \theta} f(Y(\omega, \theta)) d\mathbb{Q}(\omega) \\ &= \int_\Omega f'(Y(\omega, \theta)) \cdot \frac{\partial Y(\omega, \theta)}{\partial \theta} d\mathbb{Q}(\omega) = \mathbb{E}^\mathbb{Q}(f'(Y(\theta)) \cdot \frac{\partial Y(\theta)}{\partial \theta} | \mathcal{F}_{t_0}) \\ &\stackrel{f \text{ smooth}}{\approx} \hat{\mathbb{E}}^\mathbb{Q}(f'(Y(\theta)) \cdot \frac{\partial Y(\theta)}{\partial \theta} | \mathcal{F}_{t_0}) = \frac{1}{n} \sum_{i=1}^n f'(Y(\omega_i, \theta)) \cdot \frac{\partial Y(\omega_i, \theta)}{\partial \theta} \end{aligned}$$

Requirements

- Additional information on the model sde X
- No additional information on the simulation scheme $X(t_{i+1})$
- Additional information on the payout f (derivative of f must be known)
- Additional information on the nature of θ ($\Rightarrow$ no generic sensitivities)

Properties

- Unbiased derivative
- Discontinuous payouts may be handled (interpret f' as distribution, see below)

It seems as if a discontinuity in the payout may not be handled, since we require f' to exist. However the impact of a discontinuity be calculated analytically. See Joshi & Kainth [74] or Rott & Fries [82] for an example on how use pathwise differentiation with discontinuous payouts (there in the context of Default Swaps, CDOs).

The pathwise method requires the knowledge of the derivative of the payout f' and the derivative of the process realizations with respect to the parameter θ , i.e. $\frac{\partial Y(\omega_i, \theta)}{\partial \theta}$. It is thus only applicable for a restricted class of models and model parameters, where $\frac{\partial Y(\omega_i, \theta)}{\partial \theta}$ may be calculated analytically.

It is a major disadvantage of the method that it requires special knowledge of the payout function and of model realizations.

15.5.1. Example: Delta of a European Option under a Black-Scholes Model

We consider a Black-Scholes Model

$$\begin{aligned} S(t) &= S(0) \cdot \exp(\bar{r}t - \frac{1}{2}\bar{\sigma}^2 t + \bar{\sigma}W(t)), & S(0) &= S_0 \\ B(t) &= B(0) \cdot \exp(\bar{r}t) \end{aligned} \quad (15.2)$$

In this case we have, e.g.

$$\frac{\partial}{\partial S_0} S(T) = \frac{S(T)}{S_0}.$$

Using the notation above, our model primitive are $X = (S, B)$. We assume that the payout of our derivative depends on $Y = X(T) = (S(T), B(T))$ only, i.e. we consider an European option. Then we have

$$\begin{aligned} \frac{\partial}{\partial S_0} \mathbb{E}^{\mathbb{Q}}(f(S(T)) | \mathcal{F}_{t_0}) &= \mathbb{E}^{\mathbb{Q}}(f'(S(T)) \cdot \frac{S(T)}{S_0} | \mathcal{F}_{t_0}) \\ &\approx \hat{\mathbb{E}}^{\mathbb{Q}}(f'(S(T)) \cdot \frac{S(T)}{S_0} | \mathcal{F}_{t_0}) \\ &= \frac{1}{n} \sum_{i=1}^n f'(S(T, \omega_i)) \cdot \frac{S(T, \omega_i)}{S_0} \end{aligned}$$

15.5.2. Pathwise Differentiation for Discontinuous Payouts

In case that the payout f exhibits a discontinuities the pathwise method may be applied, provided that f allows for a decomposition

$$f = g + \sum_i \alpha_i \mathbf{1}(\{Y(\theta) > y_i\}),$$

with g being smooth. In this case we have

$$\begin{aligned} \frac{\partial}{\partial \theta} \mathbb{E}^{\mathbb{Q}}(f(Y(\theta)) | \mathcal{F}_{T_0}) &= \frac{\partial}{\partial \theta} \int_{\Omega} f(Y(\omega, \theta)) d\mathbb{Q}(\omega) = \int_{\Omega} \frac{\partial}{\partial \theta} f(Y(\omega, \theta)) d\mathbb{Q}(\omega) \\ &= \int_{\Omega} f'(Y(\omega, \theta)) \cdot \frac{\partial Y(\omega, \theta)}{\partial \theta} d\mathbb{Q}(\omega) = \mathbb{E}^{\mathbb{Q}}(f'(Y(\theta)) \cdot \frac{\partial Y(\theta)}{\partial \theta} | \mathcal{F}_{T_0}) \\ &\stackrel{g \text{ smooth}}{\approx} \hat{\mathbb{E}}^{\mathbb{Q}}(g'(Y(\theta)) \cdot \frac{\partial Y(\theta)}{\partial \theta} | \mathcal{F}_{T_0}) + \sum_i \alpha_i \cdot \phi(y_i) \cdot \frac{\partial Y(\theta)}{\partial \theta} \Big|_{Y(\theta)=y_i} \end{aligned}$$

See Joshi & Kainth [JK] or Rott & Fries [RF] for an example on how use pathwise differentiation with discontinuous payouts (there in the context of Default Swaps, CDOs).

15.6. Sensitivities by Likelihood Ratio Weighting

The pathwise method differentiates the path value $Y(\theta)$ of the underlying process realizations Y . Provided there is an probability density $\phi_{Y(\theta)}$ of $Y(\theta)$ we may write the expectation as a convolution with the density. The likelihood ratio weighting [51, 53, 14] is then given by

$$\begin{aligned} \frac{\partial}{\partial \theta} \mathbb{E}^{\mathbb{Q}}(f(Y(\theta)) | \mathcal{F}_{t_0}) &= \frac{\partial}{\partial \theta} \int_{\Omega} f(Y(\omega, \theta)) d\mathbb{Q}(\omega) = \frac{\partial}{\partial \theta} \int_{\mathbb{R}^m} f(y) \cdot \phi_{Y(\theta)}(y) dy \\ &= \int_{\mathbb{R}^m} f(y) \cdot \frac{\frac{\partial}{\partial \theta} \phi_{Y(\theta)}(y)}{\phi_{Y(\theta)}(y)} \cdot \phi_{Y(\theta)}(y) dy = \mathbb{E}^{\mathbb{Q}}(f(Y) \cdot w(\theta) | \mathcal{F}_{t_0}) \\ &\approx \hat{\mathbb{E}}^{\mathbb{Q}}(f(Y) \cdot w(\theta) | \mathcal{F}_{t_0}) = \frac{1}{n} \sum_{i=1}^n f(Y(\omega_i)) \cdot w(\theta, \omega_i), \end{aligned}$$

where $w(\theta) := \frac{\frac{\partial}{\partial \theta} \phi_{Y(\theta)}(Y(\theta))}{\phi_{Y(\theta)}(Y(\theta))}$.

Requirements

- Additional information on the model sde $X \rightarrow \phi_{Y(\theta)}$
- No additional information on the simulation scheme $X(t_{i+1})$
- No additional information on the payout f
- Additional information on the nature of θ ($\Rightarrow$ no generic sensitivities)

Properties

- Unbiased derivative
- Discontinuous payouts may be handled.

The Likelihood Ratio method requires no additional information on the payout function. This is an advantage compared to the pathwise differentiation. However it requires that the density of the model's realizations $X(t)$ is known and furthermore that its derivative with respect to the parameter θ is known. This is rarely the case and this is viewed as being the major drawback of the method.

The Likelihood Ratio method does not require the payout to be smooth. The method works very well for calculating the impact of a discontinuity in the payout. However, the method has its problems with smooth payouts: the Monte-Carlo error of the approximation using Likelihood Ratio is larger than the Monte-Carlo error of the finite-difference approximation. We give a simple example of this effect next.

15.6.1. Example: Delta of a European Option under a Black-Scholes Model using Pathwise Derivative

We consider again a European Option under the [Black-Scholes model \(15.2\)](#). Since B is deterministic we need to consider the probability density of S . Since $\log(S(T))$ is normal distributed, see [Chapter 4](#), we have for the density of $S(T)$

$$\phi_{S(T)}(s) = \frac{\phi_{\text{std.norm.}}\left(\frac{1}{\bar{\sigma}\sqrt{T}}(\log(s) + \bar{r}(T) - \frac{1}{2}\bar{\sigma}(T)^2 - \log(S_0))\right)}{s},$$

where $\phi_{\text{std.norm.}}(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)$ is the density of the standard normal distribution.

Thus, the delta of a European option with (Numeraire-rebased) payout $f(S(T), B(T))$, calculated by the likelihood ratio method, is given by

$$\mathbb{E}^{\mathbb{Q}} \left(f(S(T), B(T)) \cdot \frac{\frac{\partial}{\partial S_0} \phi_{S(T)}(S(T))}{\phi_{S(T)}(S(T))} \mid \mathcal{F}_0 \right)$$

15.6.2. Example: Variance Increase of the Sensitivity when using Likelihood Ratio Method for Smooth Payouts

For some smooth payouts, the Likelihood Ratio method may perform less accurate than the pathwise method [15.5](#) or its finite difference approximation [15.4](#). A simple examples illustrates this effect: Consider constant payout $f(S(T), B(T)) = b$. Then,

the likelihood ratio method gives the delta of this option as

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}} \left(f(S(T), B(T)) \cdot \frac{\frac{\partial}{\partial S_0} \phi_{S(T)}(S(T))}{\phi_{S(T)}(S(T))} \mid \mathcal{F}_0 \right) &= \int_{\mathbb{R}} f(s, B(T)) \cdot \frac{\partial}{\partial S_0} \phi_{S(T)}(s) \, ds \\ &\stackrel{f=b=\text{const.}}{=} b \cdot \int_{\mathbb{R}} \frac{\partial}{\partial S_0} \phi_{S(T)}(s) \, ds \end{aligned}$$

and indeed (using substitution $y = \log(s)$, $dy = \frac{1}{s} ds$) we see that the delta is zero:

$$= \int_{\mathbb{R}} \frac{\partial}{\partial S_0} \phi(y) dy = 0$$

The Monte-Carlo approximation is

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}} \left(f(S(T), B(T)) \cdot \frac{\frac{\partial}{\partial S_0} \phi_{S(T)}(S(T))}{\phi_{S(T)}(S(T))} \mid \mathcal{F}_0 \right) \\ \approx \frac{1}{n} \sum_{i=1}^n f(S(T, \omega_i), B(T)) \cdot \frac{\frac{\partial}{\partial S_0} \phi(S(T, \omega_i))}{\phi(S(T, \omega_i))} = \frac{1}{n} \sum_{i=1}^n b \cdot \frac{\frac{\partial}{\partial S_0} \phi(S(T, \omega_i))}{\phi(S(T, \omega_i))}, \end{aligned}$$

which is in general non zero. It is an approximation of zero, having some variance.

On the other hand, note that the pathwise method and even a finite difference approximation thereof would give a delta of zero with zero Monte-Carlo variance.

15.7. Sensitivities by Malliavin Weighting

The Malliavin weighting is similar to the Likelihood Ratio method: the sensitivity is expressed as the expectation of a weighted payout function.

$$\begin{aligned} \frac{\partial}{\partial \theta} \mathbb{E}^{\mathbb{Q}}(f(Y(\theta)) \mid \mathcal{F}_{t_0}) &= \mathbb{E}^{\mathbb{Q}}(f(Y(\theta)) \cdot w(\theta) \mid \mathcal{F}_{t_0}) \\ &\approx \hat{\mathbb{E}}^{\mathbb{Q}}(f(Y(\theta)) \cdot w(\theta) \mid \mathcal{F}_{t_0}) = \frac{1}{n} \sum_{i=1}^n f(Y(\theta, \omega_i)) \cdot w(\theta, \omega_i) \end{aligned}$$

Requirements

- Additional information on the model sde $X \rightarrow w$
- No additional information on the simulation scheme $X(t_{i+1})$
- No additional information on the payout f
- Additional information on the nature of θ ($\Rightarrow$ no generic sensitivities)

Properties

- Unbiased derivative
- Discontinuous payouts may be handled.

Benhamou [45] showed that the Likelihood Ratio corresponds to the Malliavin weights with minimal variance and may be expressed as a conditional expectation of all corresponding Malliavin weights (we thus view the Likelihood Ratio as an example for the Malliavin weighting method).

However here the weights are derived directly through Malliavin calculus which makes this method more general and applicable even if the density is not known. The derivation of the Malliavin weights requires in depth knowledge of the underlying continuous process X and heavily depends on the nature of θ .

15.8. Proxy Simulation Scheme

The proxy simulation scheme defines a design of a Monte Carlo pricing engine that has the remarkable properties that the application of finite differences to the pricing will result in Likelihood Ratio weighted sensitivities without actually the need to know the density ϕ analytically. Thus it combines the robustness of Likelihood Ratio or Malliavin weighting with the genericity of finite differences.

Since the proxy simulation scheme method is not solely devoted to the calculation of sensitivities it will be discussed in its own chapter in 16. For a detailed discussion of the proxy simulation scheme see also [60, 61]. We will state here shortly the key properties.

The Monte-Carlo sensitivity under a proxy simulation scheme is given by

$$\begin{aligned}
 \frac{\partial}{\partial \theta} \mathbb{E}^{\mathbb{Q}}(f(Y^*(\theta)) | \mathcal{F}_{t_0}) &\approx \frac{1}{2h} (\mathbb{E}^{\mathbb{Q}}(f(Y^*(\theta + h)) | \mathcal{F}_{t_0}) - \mathbb{E}^{\mathbb{Q}}(f(Y^*(\theta - h)) | \mathcal{F}_{t_0})) \\
 &= \frac{\partial}{\partial \theta} \int_{\mathbb{R}^m} f(y) \cdot \frac{1}{2h} (\phi_{Y^*(\theta+h)}(y) - \phi_{Y^*(\theta-h)}(y)) dy \\
 &= \int_{\mathbb{R}^m} f(y) \cdot \frac{\frac{1}{2h} (\phi_{Y^*(\theta+h)}(y) - \phi_{Y^*(\theta-h)}(y))}{\phi_{Y^\circ}(y)} \cdot \phi_{Y^\circ}(y) dy \\
 &\approx \frac{1}{n} \sum_{i=1}^n f(Y^\circ(\omega_i)) \cdot \frac{1}{2h} (w(\theta + h, \omega_i) - w(\theta - h, \omega_i))
 \end{aligned}$$

Requirements

- No additional information on the model sde X
- Additional information on the simulation scheme $X^*(t_{i+1})$, $X^\circ(t_{i+1})$
- No additional information on the payout f
- No additional information on the nature of θ ($\Rightarrow$ generic sensitivities)

Properties

- Biased derivative (but small shift h possible!)
- Discontinuous payouts may be handled.

CHAPTER 16

Proxy Simulation Schemes for Monte-Carlo Sensitivities and Importance Sampling

section under construction.

16.1. Full Proxy Simulation Scheme

In this section we describe the proxy simulation scheme technique as it is given in [61] and [60]. We take the notation of the previous chapter, see 15.2 and 15.3 and consider two time discrete schemes for the stochastic process X :

X^* $t_i \mapsto X^*(t_i)$ $i = 0, 1, 2, \dots$ time discretization scheme of $X \rightarrow \text{target scheme}$

X° $t_i \mapsto X^\circ(t_i)$ $i = 0, 1, 2, \dots$ any other time discrete stochastic process
(assumed to be *close* to X^*) $\rightarrow \text{proxy scheme}$

Let $\phi_{Y^\circ}(y)$ denote the *density* of Y° and $\phi_{Y^*}(y)$ the *density* of Y^* . We require

$$\forall y : \phi^{Y^\circ}(y) = 0 \Rightarrow \phi^{Y^*}(y) = 0. \quad (16.1)$$

Using the additional scheme X° the pricing of a payout function f is now performed in the following way: Let $Y = (X(t_1), \dots, X(t_m))$, $Y^* = (X^*(t_1), \dots, X^*(t_m))$, $Y^\circ = (X^\circ(t_1), \dots, X^\circ(t_m))$. We have $E^{\mathbb{Q}}(f(Y(\theta)) | \mathcal{F}_{t_0}) \approx E^{\mathbb{Q}}(f(Y^*(\theta)) | \mathcal{F}_{t_0})$ and furthermore

$$\begin{aligned} E^{\mathbb{Q}}(f(Y^*(\theta)) | \mathcal{F}_{t_0}) &= \int_{\Omega} f(Y^*(\omega, \theta)) d\mathbb{Q}(\omega) = \int_{\mathbb{R}^m} f(y) \cdot \phi_{Y^*}(\theta)(y) dy \\ &= \int_{\mathbb{R}^m} f(y) \cdot \frac{\phi_{Y^*}(\theta)(y)}{\phi_{Y^\circ}(y)} \cdot \phi_{Y^\circ}(y) dy = E^{\mathbb{Q}}(f(Y^\circ) \cdot w(\theta) | \mathcal{F}_{t_0}), \end{aligned}$$

where $w(\theta) = \frac{\phi_{Y^*}(\theta)(y)}{\phi_{Y^\circ}(y)}$.

For the Monte-Carlo approximation this implies that the sample paths are generated from the scheme X° while the probability densities are corrected towards the target scheme X^* .

Notes

- For $X^\circ = X^*$ we have $w(\theta) = 1$ and in this case the proxy simulation scheme corresponds to the ordinary Monte Carlo simulation of X^* .
- The proxy scheme X° and thus its realization vector Y° is seen as being independent of θ . This has important implications on the calculation of sensitivities, see Section 16.2.
- The requirement $\forall y : \phi^{Y^\circ}(y) = 0 \Rightarrow \phi^{Y^*}(y) = 0$ corresponds to the non-degeneracy condition of the diffusion matrix as it appears in the application of the likelihood ratio and Malliavin weights. However, here this requirement is by far less restrictive since we are free to choose the proxy scheme X° .

16.1.1. Calculation of Monte-Carlo weights

For the most common numerical scheme the densities ϕ^{Y° , ϕ^{Y^*} and thus the Monte-Carlo weights may be calculated numerically. Consider for example the schemes

$$\text{target scheme: } X^*(t_{i+1}) = X^*(t_i) + \mu^{X^*}(t_i)\Delta t_i + \Sigma(t_i) \cdot \Gamma(t_i) \cdot \Delta U(t_i)$$

$$\text{proxy scheme: } X^\circ(t_{i+1}) = X^\circ(t_i) + \mu^{X^\circ}(t_i)\Delta t_i + \Sigma^\circ(t_i) \cdot \Gamma^\circ(t_i) \cdot \Delta U(t_i)$$

where Σ denotes an invertible volatility matrix and Γ denotes a projection matrix, the factor matrix which defines the correlation structure $R = \Gamma\Gamma^\top$.

Assume for simplicity that $\mu^{X^*}(t_i)$ depends on $X^*(t_i)$, $X^*(t_{i+1})$ only (and similar for $\mu^{X^\circ}(t_i)$) (this holds for, e.g. Euler Scheme, Predictor Corrector), then we have for the transition probability densities

$$\begin{aligned} \phi^{X^*}(t_i, X_i^*; t_{i+1}, X_{i+1}^*) = \\ \frac{1}{(2\pi\Delta t_i)^{n/2}} \exp\left(-\frac{1}{2\Delta t_i} (\Lambda^{-1/2} F^\top \Sigma^{-1} (X_{i+1}^* - X_i^* - \mu^{X^*}(t_i)\Delta t_i))^2\right) \end{aligned}$$

$$\begin{aligned} \phi^{X^\circ}(t_i, X_i^\circ; t_{i+1}, X_{i+1}^\circ) = \\ \frac{1}{(2\pi\Delta t_i)^{n/2}} \exp\left(-\frac{1}{2\Delta t_i} (\Lambda^{\circ-1/2} F^{\circ\top} \Sigma^{\circ-1} (X_{i+1}^\circ - X_i^\circ - \mu^{X^\circ}(t_i)\Delta t_i))^2\right). \end{aligned}$$

And the proxy scheme weights are given by

$$w(t_{i+1}) | \mathcal{F}_{t_k} = \prod_{j=k}^i \frac{\phi^{X^*}(t_j, X_j^\circ; t_{j+1}, X_{j+1}^\circ)}{\phi^{X^\circ}(t_j, X_j^\circ; t_{j+1}, X_{j+1}^\circ)},$$

see [61] for details.

Notes

- We used the factor decomposition (PCA) $\Gamma = F \cdot \sqrt{\Lambda}$ where $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_m)$ are the non-zero Eigenvalues of $\Gamma \cdot \Gamma^\top$.
- A [change of market data / calibration](#) enters into transition probabilities only.

16.2. Sensitivities by Finite Differences on a Proxy Simulation Scheme

$$\begin{aligned}
\frac{\partial}{\partial \theta} \mathbb{E}^{\mathbb{Q}}(f(Y^*(\theta)) | \mathcal{F}_{t_0}) &\approx \frac{1}{2h} (\mathbb{E}^{\mathbb{Q}}(f(Y^*(\theta + h)) | \mathcal{F}_{t_0}) - \mathbb{E}^{\mathbb{Q}}(f(Y^*(\theta - h)) | \mathcal{F}_{t_0})) \\
&= \frac{\partial}{\partial \theta} \int_{\mathbb{R}^m} f(y) \cdot \frac{1}{2h} (\phi_{Y^*(\theta+h)}(y) - \phi_{Y^*(\theta-h)}(y)) dy \\
&= \int_{\mathbb{R}^m} f(y) \cdot \frac{\frac{1}{2h} (\phi_{Y^*(\theta+h)}(y) - \phi_{Y^*(\theta-h)}(y))}{\phi_{Y^\circ}(y)} \cdot \phi_{Y^\circ}(y) dy \\
&\approx \frac{1}{n} \sum_{i=1}^n f(Y^\circ(\omega_i)) \cdot \frac{1}{2h} (w(\theta + h, \omega_i) - w(\theta - h, \omega_i))
\end{aligned}$$

Requirements

- No additional information on the model sde X
- Additional information on the simulation scheme $X^*(t_{i+1}), X^\circ(t_{i+1})$
- No additional information on the payout f
- No additional information on the nature of θ ($\Rightarrow$ generic sensitivities)

Properties

- Biased derivative (but small shift h possible!)
- Discontinuous payouts may be handled.

Notes

We noted above that additional information on the simulation scheme is required, which are the densities of the two schemes. Note however that we require these densities to setup the pricing algorithm. For the sensitivity calculation no additional information is needed. Note also that the required densities are densities of numerical schemes which may usually calculated numerical from known transition probability densities (see Section 16.1.1).

16.2.1. Localization

If the payout function f is smooth then ordinary finite differences perform better than the weighting techniques. The latter shows an increase in Monte-Carlo variance of the sensitivity. This effect is not only visible for smooth payouts f , but also for large finite difference shifts, see Figure ??.

A solution that has been proposed in [58] is localization. Here the weighting is applied only to a region where the payoff is discontinuous.

Let g denote the localization function, i.e. a smooth function $0 \leq g \leq 1$ such that $g = 1$ at discontinuities of f . Consider the decomposition

$$f = (1 - g) \cdot f + g \cdot f.$$

We define the pricing of the payout f as

$$E(f(Y^*)|\mathcal{F}_{t_0}) = E((1 - g(Y^*)) \cdot f(Y^*)|\mathcal{F}_{t_0}) + E(g(Y^\circ) \cdot f(Y^\circ) \cdot \frac{\phi_{Y^*}}{\phi_{Y^\circ}}|\mathcal{F}_{t_0}).$$

In other words: We use a pricing based on a proxy simulation scheme for $g \cdot f$ and a pricing based on direct simulation for $(1 - g) \cdot f$.

It should be noted that localization is carried out by a redefinition of the payout. The product is split in two parts, where one is priced by a direct simulation scheme and the other is priced by a proxy simulation scheme method. This allows to implement localization on the product level, completely independent of the actual simulation properties. In addition, localization does not reduce the ability to calculate generic sensitivities.

16.2.2. Object Oriented Design

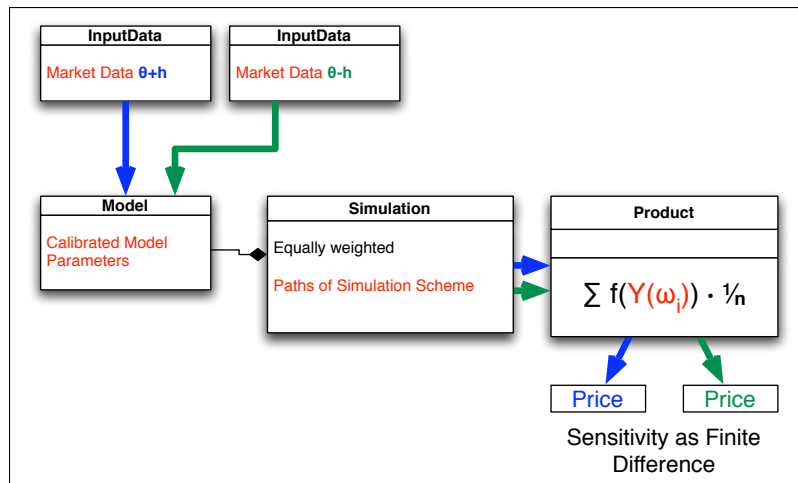
The proxy scheme simulation method may in part also be viewed as an implementation design. In Figure 16.1(a) we depict the object oriented design of a standard Monte-Carlo simulation where a change in market data results in a change of simulation path. In Figure 16.1(b) we contrast the proxy scheme simulation method where a change in market data results in a change of Monte-Carlo weights. In practice we propose that the model driving the proxy schemes paths generation is calibrated to market data used for pricing while a market data scenario used for sensitivity calculation, i.e. bumping the model, impacts the Monte-Carlo weights only. A method should be offered to reset the proxy simulation's market data to the target simulation's market data.

16.3. Importance Sampling

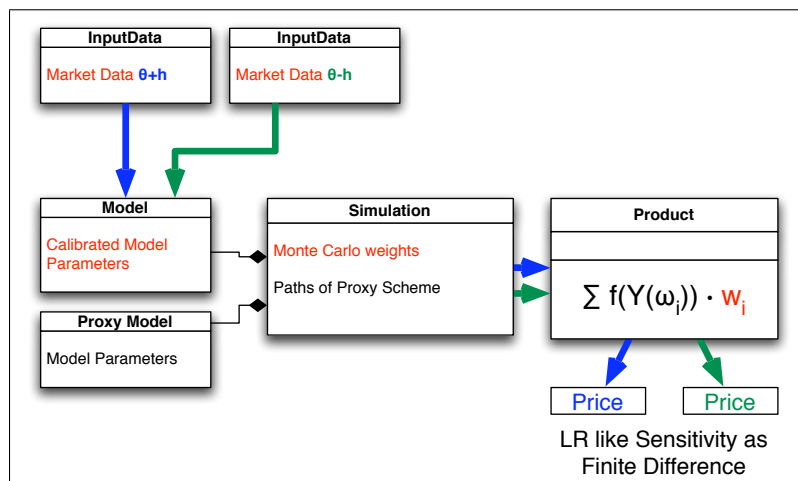
The key idea of importance sampling is to generate the paths according to their importance to the application, not according to their probability law, and doing so adjust towards their probability by a suitable Monte-Carlo weight (the change of measure).

Using a proxy simulation scheme, the paths are generated according to the proxy scheme while a Monte-Carlo weight adjusts their probability towards the target scheme. Actually, once the proxy simulation scheme framework has been established, the Monte-Carlo weights are calculated automatically from the two numerical schemes.

Thus, choosing the proxy scheme such that it creates paths according to their importance to the application is a form of importance sampling. It has the advantage that



(a) Standard Monte Carlo Simulation



(b) Proxy Scheme Monte Carlo Simulation

Figure 16.1.: Object Oriented Design of the Monte Carlo Pricing Engine: We depict the impact of a change of different market data scenarios $\theta + h$ and $\theta - h$ on the pricing code of a standard Monte-Carlo simulation and a proxy scheme simulation.

specifying a suitable process might come easier than calculating the optimal sampling and the corresponding Monte-Carlo weights.

16.3.1. Example

We consider the pricing of a far out of the money option under a log-normal model (like the Black-Scholes model or the LIBOR market model):

$$\begin{aligned} \text{Log Euler Scheme:} \quad & d \log(X)(t_{j+1}) = \log(X)(t_{j+1}) + \mu(t)dt + \sigma dW \\ \text{OTM option:} \quad & \max(X(T) - K, 0), \end{aligned}$$

where $X(0) = X_0$ and $K \gg X_0$. The drift of the model is determined by the specific pricing measure. However, in our application we would prefer that the mean of $X(T)$ is close to the option strike K rather than being $\exp(\log X_0 + \int_0^T \mu(t)dt)$. To achieve this, simply use a proxy scheme with artificial drift:

$$\begin{aligned} \text{Proxy Scheme:} \quad & d \log(X)(t_{j+1}) = \log(X)(t_{j+1}) + \frac{\log(K) - \log(X_0)}{T}dt + \sigma_i dW \\ \text{Target Scheme:} \quad & d \log(X)(t_{j+1}) = \log(X)(t_{j+1}) + \mu(t)dt + \sigma dW \end{aligned}$$

This will bring the paths to the region important for the pricing of the option, while the proxy simulation scheme framework automatically adjust probabilities accordingly. Figure 16.2 shows a comparison of distribution of Monte-Carlo prices obtained from direct simulation and the importance adjusted proxy scheme simulation.

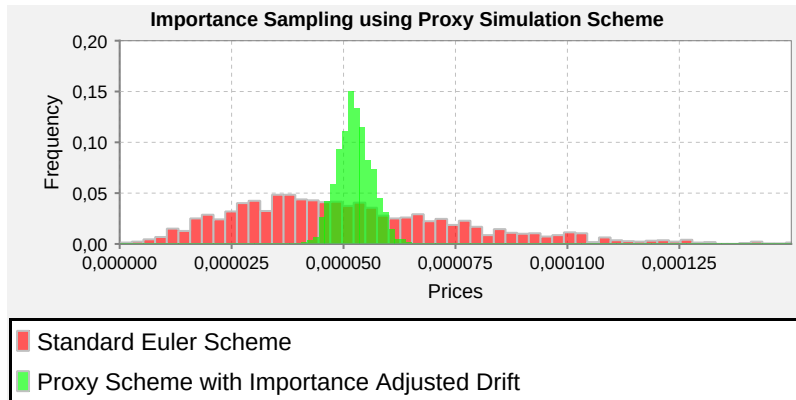


Figure 16.2.: Importance sampling using a drift adjusted proxy scheme. The example was created using a LIBOR market model to price a caplet with strike $K = 0.3$, the initial forward rate being $X_0 = L_i(0) = 0.1$.

16.4. Partial Proxy Simulation Schemes

The (full) proxy simulation scheme method requires that the density of the target scheme realization is zero if the density of the proxy scheme is zero, (16.1). In other words it is required that the paths simulated under the proxy scheme comprise all paths possible under the target scheme. If the property is violated then the Monte-Carlo expectation using the weighted paths of the proxy scheme will leave out some mass. This limits the application of the full proxy simulation scheme. For the calculation of sensitivities the limitation means that we may not calculate the sensitivity with respect to all possible perturbations.

However, in order to improve the calculation of sensitivities of trigger products it is not required to keep all underlying quantities rigid (as for a full proxy simulation); it is sufficient to keep the quantity rigid that induces the discontinuity. This gives rise to the notion of a partial proxy simulation scheme, [60].

Part V.

Pricing Models for Interest Rate Derivatives

CHAPTER 17

Market Models

The pure and simple truth
is rarely pure and never simple.

Oscar Wilde
The Importance of Being Earnest, [37].

Up to this point now we have considered Models of a single scalar stochastic process and options on it: The [Black-Scholes model](#) for a stock S , or the [Black model](#) for a forward rate L . The true challenge in evaluation of interest rates products lies in the modeling of the whole interest rate curve (instead of a scalar) and in the evaluation of complex derivatives, which depend on the whole curve.

Historically the path to modeling the interest rate curve started with the modeling of the short rate, from which we may calculate the whole interest rate curve, see Remark [93](#). The motivation to initially consider the short rate derived from the wish to model a scalar quantity, thus being able to apply the well known numerical methods from stock models, e.g. binomial trees.

For didactic reasons we do not follow the historical paths. Instead we consider first the LIBOR Market Model. It is a high dimensional model, which discretizes the interest rate curve into a finite number of forward rates. It is highly flexibly due to its huge number of free parameters. It will allow us to study model properties like *mean reversion*, *number of factors* (Chapter [22](#)) and *instantaneous volatility*, instantaneous and terminal correlation (Chapter [18](#)). Despite its presumed complexity, the LIBOR Market Model is essentially a very simple model: It is nothing more than the simultaneous consideration of multiple [Black models](#) under a common measure. We thus continue from Chapter [10](#).

For the [Short Rate models](#) the modeled quantity is the *short rate*, a quantity not directly observable. Here we model quantities which are observable as market quotes, like the *LIBOR* or the *Swaprte*. The class of models that model quantities which are directly observable on the market are called “Market Models”. We consider first the *LIBOR Market Model*:

17.1. LIBOR Market Model

We assume a time discretization (*tenor structure*)

$$0 = T_0 < T_1 < \dots < T_n.$$

We model the **forward rates** $L_i := L(T_i, T_{i+1})$ for $i = 0, \dots, n-1$, see Definition 88. This represents a discretization of the interest rate curve, where the continuum of maturities has been discretized.¹

The *LIBOR market model* assumes a lognormal dynamic for LIBORs $L_i := L(T_i, T_{i+1})$, i.e.²

$$\frac{dL_i(t)}{L_i(t)} = \mu_i^{\mathbb{P}}(t)dt + \sigma_i(t)dW_i^{\mathbb{P}}(t) \quad \text{für } i = 0, \dots, n-1, \text{ under } \mathbb{P}, \quad (17.1)$$

with initial conditions

$$L_i(0) = L_{i,0}, \quad \text{mit } L_{i,0} \in [0, \infty), i = 0, \dots, n-1,$$

where $W_i^{\mathbb{P}}$ denote (possibly instantaneously correlated) **$\mathbb{P}$ -Brownian motions** with

$$dW_i^{\mathbb{P}}(t)dW_j^{\mathbb{P}}(t) = \rho_{i,j}(t)dt.$$

Let $\sigma_i : [0, T] \mapsto \mathbb{R}$ and $\rho_{i,j} : [0, T] \mapsto \mathbb{R}$ be deterministic functions and μ_i the Drift as $\mathcal{F}_t$ -adapted process. By $R(t) := (\rho_{i,j}(t))_{i,j=0,\dots,n-1}$ we denote the correlation matrix.



Motivation: Equation (17.1) is a lognormal model for the **forward rates** L_i . If we consider only a single equation, i.e. fix $i \in \{1, \dots, n-1\}$, it represents the **Black model** considered in Section 10: Equation (17.1) is identical with Equation (10.1). If we change the measure such that L_i is drift free (see Section 10), we see that the terminal distribution of L_i is lognormal.

Thus, the LIBOR market model is equivalent to the consideration of n **Black models** under a unified measure.

As was discussed in Section 10, to evaluate a caplet under this model it is not relevant that σ_i is time dependent (we have assumed time-dependency of σ_i in Section 10 for didactical reasons). For the value of complex derivatives the time dependency matters. A further degree of freedom that is introduced in (17.1) is the **instantaneous correlation** $\rho_{i,j}$ of the driving Brownian motions. For the value of a caplet the instantaneous correlation is insignificant (indeed, it does not enter in the Black model). For the evaluation of **Swaptions** the correlation of the forward rates is significant.

¹ In practice it is common to model semi-annual or quarterly rates $T_{i+1} - T_i = 0.25$ and to consider these up to a maturity of 20 or 30 years, resulting in 80 or 120 interest rates to model.

² We denote the simulation time parameter of the stochastic process by t .

Further generalization of the model consider non deterministic σ_i , i.e. stochastic volatility models. In this case the terminal LIBOR distribution no longer correspond to the ones of the Black model, which is, of course, intended. Equation (17.1) is to be seen as a starting point (the umbilic point) of a whole model family. The starting point has been chosen in the form of (17.1), because (historically) the lognormal (Black-)model is well understood, especially by traders.³ $\triangleleft$

Remark 193 (Interest Rate Structure): Equation (17.1) models the evolution of the LIBOR $L(T_i, T_{i+1})$. Without further interpolation assumption, these are the shortest forward rates that may be considered in our time discretization (*tenor structure*). The equation system (17.1) thus determines the evolution of all bond prices with maturities T_i and all forward rates for the periods $[T_i, T_k]$, since

$$\begin{aligned} 1 + L(T_i, T_k)(T_k - T_i) &= \frac{P(T_i)}{P(T_k)} = \prod_{j=i}^{k-1} \frac{P(T_j)}{P(T_{j+1})} \\ &= \prod_{j=i}^{k-1} (1 + L(T_j, T_{j+1})(T_{j+1} - T_j)). \end{aligned}$$

To shorten notation we write $\delta_i := T_{i+1} - T_i$, $i = 0, \dots, n-1$ for the period length.

17.1.1. Derivation of the Drift Term

As in Chapter 10 and 11, our first step is to choose some Numéraire N and derive the drift under a martingale measure $\mathbb{Q}^N$. If the processes have been derived under the martingale measure $\mathbb{Q}^N$, then the (discretized) interest rate curve may be simulated numerically and a derivative V may be priced through $V(0) = N(0)E^{\mathbb{Q}^N}(\frac{V}{N}|\mathcal{F}_{T_0})$ (siehe Kapitel 13).

We fix a Numéraire N . From assumption (3.23) there exists a corresponding equivalent martingale measure $\mathbb{Q}^N$ such that N -relative prices are martingales. From Theorem 53 we have that under $\mathbb{Q}^N$ the process (17.1) is given with a changed drift, namely as

$$\frac{dL_i(t)}{L_i(t)} = \mu_i^{\mathbb{Q}^N}(t)dt + \sigma_i(t)dW_i^{\mathbb{Q}^N}(t) \quad \text{für } i = 0, \dots, n-1. \quad (17.2)$$

17.1.1.1. Derivation of the Drift Term under the Terminal Measure

$$\frac{dL_i(t)}{L_i(t)} = \mu_i^{\mathbb{Q}^{P(T_n)}}(t)dt + \sigma_i(t)dW_i^{\mathbb{Q}^{P(T_n)}}(t) \quad \text{for } i = 0, \dots, n-1. \quad (17.3)$$

³ Caplet prices are quoted by traders by the implied Black volatility. This is of course just another unit of the price, since the Black model is a one-to-one map from price to implied volatility.

To determine are $\mu_i^{\mathbb{Q}^{P(T_n)}}$. The martingale measure $\mathbb{Q}^{P(T_n)}$ corresponding to $N(t) = P(T_n; t)$ is also called *terminal measure* (since T_n is the time horizon of our time discretization).

As in Chapter 10, we will construct relative prices with respect to $P(T_n)$ and obtain equations from which we will derive the drifts μ_i . From Definition 88 we have for $i = 0, \dots, n-1$

$$\prod_{k=i}^{n-1} \underbrace{(1 + \delta_k L_k)}_{= \frac{P(T_k)}{P(T_{k+1})}} = \prod_{k=i}^{n-1} \frac{P(T_k)}{P(T_{k+1})} = \frac{P(T_i)}{P(T_n)}. \quad (17.4)$$

Since we have a $P(T_n)$ -relative price of a traded product on the right hand side in (17.4), we have for the drifts:

$$\text{Drift}_{\mathbb{Q}^{P(T_n)}} \left[\prod_{k=i}^{n-1} (1 + \delta_k L_k) \right] = 0 \quad i = 0, \dots, n-1 \text{ in } \mathbb{Q}^{P(T_n)}.$$

We apply Theorem 42 and obtain $\forall i = 0, \dots, n-1$

$$\begin{aligned} d \left(\prod_{k=i}^{n-1} (1 + \delta_k L_k) \right) &= \sum_{j=i}^{n-1} \prod_{\substack{k=i \\ k \neq j}}^{n-1} (1 + \delta_k L_k) \cdot \delta_j dL_j + \sum_{\substack{j,l=i \\ l > j}}^{n-1} \prod_{\substack{k=i \\ k \neq j,l}}^{n-1} (1 + \delta_k L_k) \cdot \delta_j dL_j \delta_l dL_l \\ &= \prod_{k=i}^{n-1} (1 + \delta_k L_k) \cdot \left(\sum_{j=i}^{n-1} \frac{\delta_j dL_j}{(1 + \delta_j L_j)} + \sum_{\substack{j,l=i \\ l > j}}^{n-1} \frac{\delta_j dL_j}{(1 + \delta_j L_j)} \cdot \frac{\delta_l dL_l}{(1 + \delta_l L_l)} \right) \\ &= \prod_{k=i}^{n-1} (1 + \delta_k L_k) \cdot \sum_{j=i}^{n-1} \left(\frac{\delta_j dL_j}{(1 + \delta_j L_j)} + \sum_{\substack{l \geq j+1 \\ l \leq n-1}} \frac{\delta_j dL_j}{(1 + \delta_j L_j)} \cdot \frac{\delta_l dL_l}{(1 + \delta_l L_l)} \right). \end{aligned}$$

Since $\forall i = 0, \dots, n-1$

$$\text{Drift}_{\mathbb{Q}^{P(T_n)}} \left[\prod_{k=i}^{n-1} (1 + \delta_k L_k) \right] = 0 \quad (17.5)$$

it follows $\forall i = 0, \dots, n-1$

$$\sum_{j=i}^{n-1} \text{Drift}_{\mathbb{Q}^{P(T_n)}} \left[\frac{\delta_j dL_j}{(1 + \delta_j L_j)} + \sum_{\substack{l \geq j+1 \\ l \leq n-1}} \frac{\delta_j dL_j}{(1 + \delta_j L_j)} \cdot \frac{\delta_l dL_l}{(1 + \delta_l L_l)} \right] = 0 \quad (17.6)$$

and thus $\forall j = 0, \dots, n-1$

$$\text{Drift}_{\mathbb{Q}^{P(T_n)}} \left[\frac{\delta_j dL_j}{(1 + \delta_j L_j)} + \sum_{\substack{l \geq j+1 \\ l \leq n-1}} \frac{\delta_j dL_j}{(1 + \delta_j L_j)} \cdot \frac{\delta_l dL_l}{(1 + \delta_l L_l)} \right] = 0 \quad (17.7)$$

If we now use

$$dL_j = L_j \mu_j dt + L_j \sigma_j dW_j^{\mathbb{Q}^{P(T_n)}} \quad \text{and} \quad dL_j \cdot dL_l = L_j L_l \sigma_j \sigma_l \rho_{j,l} dt$$

in (17.7) (we write shortly $\mu_j(t)$ for $\mu_j^{\mathbb{Q}^{P(T_n)}}(t)$), then we have

$$\mu_j \frac{\delta_j L_j}{(1 + \delta_j L_j)} + \sum_{\substack{l \geq j+1 \\ l \leq n-1}} \frac{\delta_j L_j}{(1 + \delta_j L_j)} \cdot \frac{\delta_l L_l}{(1 + \delta_l L_l)} \cdot \sigma_j \sigma_l \rho_{j,l} = 0,$$

i.e.

$$\mu_j^{\mathbb{Q}^{P(T_n)}}(t) = - \sum_{\substack{l \geq j+1 \\ l \leq n-1}} \frac{\delta_l L_l(t)}{(1 + \delta_l L_l(t))} \cdot \sigma_j(t) \sigma_l(t) \rho_{j,l}(t). \quad (17.8)$$

The procedure above may be summarized as follows: To derive the n drifts we write down n independent traded assets as a function of the model quantities. By considering the drifts of their relative prices we obtain n equation for the drifts of the modeled quantities.

17.1.1.2. Derivation of the Drift Term under the Spot LIBOR Measure

We fix the *rolled over one period bond* as Numéraire, i.e. the investment of 1 at time T_0 into the T_1 -Bond and after its maturity the reinvestment of the proceeds into the bond of the next period, i.e. in T_j the reinvestment into the T_{j+1} -Bond. It is

$$\begin{aligned} N(t) &:= P(T_{m(t)+1}; t) \prod_{j=1}^{m(t)+1} \frac{1}{P(T_j; T_{j-1})} = P(T_{m(t)+1}; t) \prod_{j=0}^{m(t)} (1 + L_j(T_j) \cdot \delta_j), \quad (17.9) \\ &= \frac{P(T_{j-1}; T_{j-1})}{P(T_j; T_{j-1})} = (1 + L_{j-1}(T_{j-1}) \delta_{j-1}) \end{aligned}$$

$m(t) := \max\{i : T_i \leq t\}$ and $\delta_j := T_{j+1} - T_j$. The corresponding equivalent martingale measure $\mathbb{Q}^N$ is called the *spot measure*.

As before we consider the processes of N -relative prices of traded products (from which we know that they have drift 0 under $\mathbb{Q}^N$). We consider the N -relative prices of the bonds $P(T_i)$. It is

$$\begin{aligned} \frac{P(T_i; t)}{N(t)} &= \frac{P(T_i; t)}{P(T_{m(t)+1}; t)} \cdot \prod_{j=0}^{m(t)} (1 + L_j(T_j) \delta_j)^{-1} \\ &= \prod_{j=m(t)+1}^{i-1} (1 + L_j(t) \delta_j)^{-1} \cdot \prod_{j=0}^{m(t)} (1 + L_j(T_j) \delta_j)^{-1}, \quad (17.10) \end{aligned}$$

thus

$$\text{Drift}_{\mathbb{Q}^N} \left[\prod_{k=m(t)+1}^{i-1} (1 + L_k \delta_k)^{-1} \right] = 0. \quad (17.11)$$

Since

$$\begin{aligned} d \left(\prod_{j=m(t)+1}^{i-1} (1 + L_j(t) \delta_j)^{-1} \cdot \prod_{j=0}^{m(t)} (1 + L_j(T_j) \delta_j)^{-1} \right) \\ = d \left(\prod_{j=m(t)+1}^{i-1} (1 + L_j(t) \delta_j)^{-1} \right) \cdot \prod_{j=0}^{m(t)} (1 + L_j(T_j) \delta_j)^{-1} \end{aligned}$$

we consider

$$\begin{aligned} & d \left(\prod_{k=m(t)+1}^{i-1} \frac{1}{(1 + \delta_k L_k)} \right) \\ &= \sum_{j=m(t)+1}^{i-1} \prod_{\substack{k=m(t)+1 \\ k \neq j}}^{i-1} \frac{1}{(1 + \delta_k L_k)} \cdot \left(\frac{-\delta_j dL_j}{(1 + \delta_j L_j)^2} + \frac{\delta_j^2 dL_j dL_j}{(1 + \delta_j L_j)^3} \right) \\ &\quad + \sum_{\substack{j,l=m(t)+1 \\ l < j}}^{i-1} \prod_{\substack{k=m(t)+1 \\ k \neq j,l}}^{i-1} \frac{1}{(1 + \delta_k L_k)} \cdot \left(\frac{\delta_j dL_j}{(1 + \delta_j L_j)^2} + \frac{\delta_j^2 dL_j dL_j}{(1 + \delta_j L_j)^3} \right) \left(\frac{\delta_l dL_l}{(1 + \delta_l L_l)^2} + \frac{\delta_l^2 dL_l dL_l}{(1 + \delta_l L_l)^3} \right) \\ &= \prod_{k=m(t)+1}^{i-1} \frac{1}{(1 + \delta_k L_k)} \cdot \left(\sum_{j=m(t)+1}^{i-1} \frac{-\delta_j dL_j}{(1 + \delta_j L_j)} + \frac{\delta_j^2 dL_j dL_j}{(1 + \delta_j L_j)^2} + \sum_{\substack{j,l=m(t)+1 \\ l < j}}^{i-1} \frac{-\delta_j dL_j}{(1 + \delta_j L_j)} \cdot \frac{-\delta_l dL_l}{(1 + \delta_l L_l)} \right) \\ &= \prod_{k=m(t)+1}^{i-1} \frac{1}{(1 + \delta_k L_k)} \cdot \sum_{j=m(t)+1}^{i-1} \left(-\frac{\delta_j dL_j}{(1 + \delta_j L_j)} + \sum_{l=m(t)+1}^j \frac{\delta_j dL_j}{(1 + \delta_j L_j)} \cdot \frac{\delta_l dL_l}{(1 + \delta_l L_l)} \right). \end{aligned}$$

With (17.11) we have $\forall i = 0, \dots, n-1$

$$\sum_{j=m(t)+1}^{i-1} \text{Drift}_{\mathbb{Q}^N} \left[-\frac{\delta_j dL_j}{(1 + \delta_j L_j)} + \sum_{l=m(t)+1}^j \frac{\delta_j dL_j}{(1 + \delta_j L_j)} \cdot \frac{\delta_l dL_l}{(1 + \delta_l L_l)} \right] = 0$$

and thus $\forall j = 0, \dots, n-1$

$$\text{Drift}_{\mathbb{Q}^N} \left[\frac{-\delta_j dL_j}{(1 + \delta_j L_j)} + \sum_{l=m(t)+1}^j \frac{\delta_j dL_j}{(1 + \delta_j L_j)} \cdot \frac{\delta_l dL_l}{(1 + \delta_l L_l)} \right] = 0 \quad (17.12)$$

If we now use

$$dL_j = L_j \mu_j dt + L_j \sigma_j dW_j \quad \text{and} \quad dL_j \cdot dL_l = L_j L_l \sigma_j \sigma_l \rho_{j,l} dt$$

in (17.7) (we write shortly $\mu_j(t)$ for $\mu_j^{\mathbb{Q}^N}$), then we have⁴

$$-\mu_j \frac{\delta_j L_j}{(1 + \delta_j L_j)} + \sum_{l=m(t)+1}^j \frac{\delta_j L_j}{(1 + \delta_j L_j)} \cdot \frac{\delta_l L_l}{(1 + \delta_l L_l)} \cdot \sigma_j \sigma_l \rho_{j,l} = 0,$$

i.e.

$$\mu_j^{\mathbb{Q}^N}(t) = \sum_{l=m(t)+1}^j \frac{\delta_l L_l(t)}{1 + \delta_l L_l(t)} \sigma_j(t) \sigma_l(t) \rho_{j,l}(t).$$

17.1.1.3. Derivation of the Drift Term under the T_k -Forward Measure

Exercise: (Drift under the T_k -forward measure) We consider

$$N(t) := \begin{cases} P(T_k; t), & t \leq T_k \\ P(T_{m(t)+1}; t) \prod_{j=k+1}^{m(t)+1} \frac{1}{P(T_j; T_{j-1})}, & t > T_k, \end{cases}$$

where $m(t) := \max\{i : T_i \leq t\}$.

1. Give an interpretation of $N(t)$ as traded product.
2. Derive the drift of the model (17.1) under the $\mathbb{Q}^N$ measure with the Numéraire N .

Solution:

$$\begin{aligned} \mu_j(t) &= - \sum_{l=j+1}^{k-1} \frac{\delta_l L_l}{(1 + \delta_l L_l)} \sigma_j \sigma_l \rho_{j,l} && \text{for } j < k-1 \quad \text{and } t \leq T_k \\ \mu_j(t) &= 0 && \text{for } j = k-1 \quad \text{and } t \leq T_k \\ \mu_j(t) &= \sum_{l=k}^j \frac{\delta_l L_l}{(1 + \delta_l L_l)} \sigma_j \sigma_l \rho_{j,l} && \text{for } j \geq k \quad \text{and } t \leq T_k \\ \mu_j(t) &= \sum_{l=m(t)+1}^j \frac{\delta_l L_l}{(1 + \delta_l L_l)} \sigma_j \sigma_l \rho_{j,l} && \text{for } t > T_k \end{aligned}$$

⁴ Since the coefficient of dt equals 0.

17.1.2. The Short Period Bond $P(T_{m(t)+1}; t)$

For $t \notin \{T_1, \dots, T_n\}$ the Numéraire $N(t)$ of the terminal measure as well as the Numéraire of the spot measure is not fully described though the processes $L_i(t)$. In both Numéraires the unspecified bond $P(T_{m(t)+1}; t)$ enters. We will now discuss the relevance of $P(T_{m(t)+1}; t)$.

17.1.2.1. Role of the short bond in a LIBOR Market Model

For the modeling of the forward rates $L_i(t) := L(T_i, T_{i+1}; t)$ on the tenor periods $[T_i, T_{i+1}]$, $i = 0, \dots, n$ the specification of $P(m(t) + 1; t)$ is not relevant. For the derivation of the corresponding drift terms it was not relevant to specify the stochastic of $P(T_{m(t)+1}; t)$, since the term canceled for the relative prices considered.

Conversely, the LIBOR Market Model does not describe the stochastic of the short bond $P(T_{m(t)+1}; t)$, since it is not given as a function of the processes $L_i(t)$.

17.1.2.2. Link to continuous time tenors

The specification of the short bond $P(T_{m(t)+1}; t)$ will become relevant, if the model should describe interest rates of interest rate periods which are not part of the tenor structure. The specification of $P(m(t) + 1; t)$ will determine how the fractional forward rates $L(T_s, T_e; t)$ with $T_s \notin \{T_1, \dots, T_n\}$ and/or $T_e \notin \{T_1, \dots, T_n\}$ will evolve, see Section 17.1.5. It is the link from a model with discrete tenors (LIBOR Market Model) to a model with continuous time tenors (Heath-Jarrow-Morton Framework). In the special case where $P(m(t) + 1; t)$ has zero volatility, the LIBOR Market Model under spot measure coincides with a Heath-Jarrow-Morton Framework with a special volatility structure under the (so called) risk neutral measure, see Section 21.2.

17.1.2.3. Drift of the short bond in a LIBOR Market Model

Within the LIBOR Market Model there is no constraint on the drift of $P(m(t) + 1; t)$, because in $\frac{P(m(t)+1;t)}{N(t)}$ the term cancels out. The relative price $\frac{P(m(t)+1;t)}{N(t)}$ is always a martingale for any choice of $P(m(t) + 1; t)$. This might come as a surprise, but we already encountered this behavior: In the Black-Scholes model the drift r of $B(t)$ is a free parameter, because it is the drift of the Numéraire. The parameter r is determined by calibration to a market interest rate. In a short rate model the drift is a free parameter. It is determined by calibration to the market interest rate curve, see Section 20. Here, similarly, $P(m(t) + 1; t)$ determines the interpolation of the initial interest rate curve.

The trivial fact that the Numéraire relative price of the Numéraire, i.e. $\frac{N(t)}{N(t)}$, is always a martingale plays a role in Markov functional models. There, the Numéraire is postulated to be a functional of some Markov process.

17.1.3. Discretization and (Monte Carlo) Simulation

After the models drift has been derived we assume for a while that the remaining free parameters σ_i , $\rho_{i,j}$ and $L_{i,0}$ ($i, j = 1, \dots, n$) are given. In this section we will discuss the discretization and implementation of the model. Section 17.1.4 will then discuss how the parameters $L_{i,0}$, σ_i , $\rho_{i,j}$ are obtained.

17.1.3.1. Generation of the (time-discrete) Forward Rate Process

As it has been discussed in Chapter 13, we choose the Euler discretization of the Itô process of $\log(L_i)$. From Lemma 44 we have

$$d(\log(L_i(t))) = (\mu_i^{\mathbb{Q}^N}(t) - \frac{1}{2}\sigma_i^2(t))dt + \sigma_i(t)dW_i^{\mathbb{Q}^N}(t) \quad (17.13)$$

and the corresponding Euler scheme of (17.13) is

$$\log(\tilde{L}_i)(t + \Delta t) = \log(\tilde{L}_i)(t) + (\mu_i(t) - \frac{1}{2}\sigma_i^2(t))\Delta t + \sigma_i(t)\Delta W_i(t), \quad (17.14)$$

Applying the exponential gives us the discretization scheme of L_i as

$$\tilde{L}_i(t + \Delta t) = \tilde{L}_i(t) \cdot \exp((\mu_i(t) - \frac{1}{2}\sigma_i^2(t))\Delta t + \sigma_i(t)\Delta W_i(t)). \quad (17.15)$$

In the special case, that the process L_i is considered under the measure $\mathbb{Q}^{P(T_{i+1})}$, i.e. $\mu_i^{\mathbb{Q}^{P(T_{i+1})}}(t) = 0$, and that the given $\sigma_i(t)$ is a known deterministic function, then we may use the exact solution for a discretization scheme:

$$L_i(t + \Delta t) = L_i(t) \cdot \exp(-\frac{1}{2}\bar{\sigma}_i^2(t, t + \Delta t)\Delta t + \bar{\sigma}_i(t, t + \Delta t)\Delta W_i),$$

where

$$\bar{\sigma}_i(t, t + \Delta t) := \sqrt{\frac{1}{\Delta t} \int_t^{t+\Delta t} \sigma_i^2(\tau) d\tau}.$$

In the case where L_i is not drift free, we choose instead of (17.15) the discretization scheme

$$L_i(t + \Delta t) = L_i(t) \cdot \exp((\mu_i(t) - \frac{1}{2}\bar{\sigma}_i^2(t, t + \Delta t)^2)\Delta t + \bar{\sigma}_i(t, t + \Delta t)\Delta W_i(t)) \quad (17.16)$$

(we write L in place of $\tilde{L}$, although (17.16) is an approximation of (17.1)). The diffusion dW is discretized by exact solution, the drift dt is discretized by an Euler scheme. The discretization error of this scheme stems from the discretization of the stochastic drift μ_i only. This discretization error results in a violation of the *no-arbitrage*

requirement of the model (the discretized model does not have the correct, arbitrage free drift). Methods which do not exhibit an arbitrage due to a discretization error are called arbitrage-free discretization, see [64]).

The volatility functions σ_i are usually assumed to be piecewise constant functions on $[T_j, T_{j+1})$, such that $\bar{\sigma}_i(t, t + \Delta t)$ may be calculated analytically. It is $\bar{\sigma}_i(t, t + \Delta t) = \sigma_i(t)$.

17.1.3.2. Generation of the Sample Paths

Equipped with the time discretization (17.16), realizations of the process are calculated for a given number of paths $\omega_1, \omega_2, \omega_3, \dots$. To do so, normal distributed random numbers $\Delta W_i(t_j)(\omega_k)$, correlated according to $R = (\rho_{i,j})$, are generated (see Appendix A.3 and A.4). These are used in the scheme (17.16). The result is a three dimensional tensor $L_i(t_j, \omega_k)$ parametrized by

i : Index of the interest rate period (tenor structure),

j : Index of the simulation time,

k : Index of the simulation path.

17.1.3.3. Generation of the Numéraire

Given a simulated interest rate curve $L_i(t_j, \omega_k)$, we may calculate the Numéraire. Of course we have to use the Numéraire that was chosen for the martingale measure under which the process was simulated (form of the drift in (17.13)). For the *terminal measure* we would calculate

$$N(T_i, \omega_k) = \prod_{j=i}^{n-1} (1 + L_j(T_i, \omega_k) \cdot (T_{j+1} - T_j))^{-1}.$$

Note: The Numéraire is given only at the tenor times $t = T_i$, since for $t \neq T_i$ we did not define the [short period bond](#) $P(T_{m(t)+1}; t)$.⁵ An interpolation is possible, see Section 17.1.5 and [84].

17.1.4. Calibration - Choice of the free Parameters

We describe now how the free parameters of the model may be chosen. The free parameters are

- the initial conditions $L_{i,0}$, $i = 0, \dots, n-1$,

⁵ See Section 17.1.2.

- the volatility functions or volatility processes⁶ σ_i , $i = 1, \dots, n-1$,
- the (instantaneous) correlation $\rho_{i,j}$, $i, j = 1, \dots, n-1$.

The determination of the free parameters is also called *calibration of the model*.



Motivation (Reproduction of Market Prices versus Historical Estimation):

- **Reproduction of Market Prices:** The parameters are chosen such that the model reproduces given market prices.
- **Historical Estimation:** The parameter are estimated from historical data, e.g. time series of interest rate fixings.

It may be surprise at first, but the second approach is not meaningful, being in the context of risk neutral evaluation. The model is considered under the martingale measure $\mathbb{Q}^N$ and its aim is to the evaluation and *hedging* (!) of derivatives. An expectation of the relative value under the martingale measure corresponds to the relative value of the replication portfolio. This replication portfolio has to be setup from traded products, traded at current (!) market prices. If the model would not replicate current market prices, then it would not possible to buy the replication portfolio of a derivative at the model price of the derivative. The model price will be wrong in general.

This remark applies to all free model parameters. In practice however, it may be difficult or impossible to derive all parameters from market prices. This may be because for a specific product is no reliable price known (low liquidity). It may also be that a corresponding product does not exist. This is often the case for correlation sensitive products from which one would like to derive the correlation parameters. If a parameter may not derived from a market price a historical estimate becomes an option. If in such a case complete hedge is not possible the residual risk has to be considered, e.g. by a conservative estimate of the parameter.

For the LIBOR market model it is common to first apply a parameter reduction, based on historical estimates of rough market assessment. An example for such a parameter reduction is the assumption of a family of functional forms for the volatility $\sigma_i(t)$ or the correlation $\rho_{i,j}(t)$. The remaining degrees of freedoms are then derived from market prices. <|

17.1.4.1. Choice of the Initial Conditions

Reproduction of Bond Market Prices Let $P^{\text{Market}}(T_i) \in [0, 1]$ denote a market observed (i.e. given) price of a T_i -bond. If we set

$$L_{i,0} := \frac{P^{\text{Market}}(T_i) - P^{\text{Market}}(T_{i+1})}{P^{\text{Market}}(T_{i+1})(T_{i+1} - T_i)},$$

⁶ The parameters σ_i may well be stochastic processes. In this cases σ_i is called a stochastic volatility model.

then the model reproduces the given market prices of the bonds P^{Market} . This is ensured by the model having the “right” drift and it is independent of the other parameters.

17.1.4.2. Choice of the Volatilities

Reproduction of Caplet Market Prices We assume here, that the σ_i ’s are deterministic functions (i.e. not random variables or stochastic processes). The forward rate L_i follows the Itô process

$$dL_i(t) = \mu_i^{\mathbb{Q}}(t)L_i(t)dt + \sigma_i(t)L_i(t)dW_i^{\mathbb{Q}}(t) \quad \text{unter } \mathbb{Q} := \mathbb{Q}^N.$$

Thus the model corresponds to the **Black modell** discussed in Chapter 10. Under $\mathbb{Q}^{P(T_{i+1})}$ we have $\mu_i^{\mathbb{Q}^{P(T_{i+1})}} = 0$, the distribution of $L_i(T_i)$ is lognormal and there exists an analytic evaluation formula for Caplets. The only model parameters that enter the caplet price are $L_0(T_i)$ and

$$\sigma_i^{\text{Black,Model}} := \left(\frac{1}{T_i} \int_0^{T_i} \sigma_i^2(t)dt \right)^{1/2}. \quad (17.17)$$

If the market price $V_{\text{Caplet},i}^{\text{Market}}$ of a caplet on the forward rate $L_i(T_i)$ is given, then the corresponding implied Black volatility $\sigma_i^{\text{Black,Market}}$ may be calculated by inverting⁷ Equation (10.2). If then $\sigma_i(t)$ is chosen such that

$$\sigma_i^{\text{Black,Model}} = \sigma_i^{\text{Black,Market}}, \quad (17.18)$$

then the model reproduces the given caplet price $V_{\text{Caplet},i}^{\text{Market}}$. A possible trivial choice is, e.g., $\sigma_i(t) = \sigma_i^{\text{Black,Market}} \forall t$.

Remark 194 (Caplet Smile Modeling): The fact that the LIBOR Market Model calibrates to the cap market by a simple boundary condition is one reason for its initial popularity. However, since the model restricted to a single LIBOR is a Black model, we have that the implied volatility does not depend on the strike of an option. Thus, in this form, the model may calibrate to a single caplet per maturity only. It cannot render a caplet smile, yet.

For extensions of the model that try to remove this restriction see [21].

Reproduction of Swaption Market Prices If the correlation $R = (\rho_{i,j})$ is given and fixed, then we influence swaption prices through the time structure of the volatility function $t \mapsto \sigma_i$. We consider swaptions that correspond to our tenor structure, i.e. option on the **Swaprates**

$$S(T_i, \dots, T_j; T_i), \quad 0 < i < j \leq n.$$

⁷ For inversion of pricing formula we may use a simple numerical algorithm. For the Black formula (10.2) we have that the price is increasing strictly monotone in the volatility.

From the definition of the swap and swap rate it is obvious that the price of a corresponding swaption with *exercise date* on or before T_i and periods $[T_i, T_{i+1}], \dots, [T_{j-1}, T_j]$ depends only on the behavior of the forward rates $L_i(t), \dots, L_{j-1}(t)$ until the *fixing* $t \leq T_i$, see Figure 17.1.

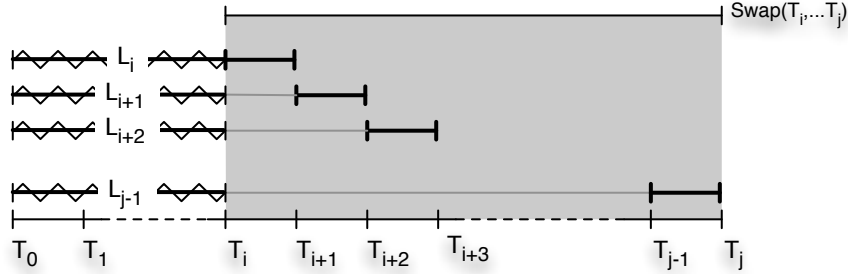


Figure 17.1.: Swaption as a function of the forward rates: The swap with periods $[T_i, T_{i+1}], \dots, [T_{j-1}, T_j]$ is a function of the forward rates $L(T_i, T_{i+1}; T_i), \dots, L(T_{j-1}, T_j; T_i)$ (all with fixing in T_i). The corresponding swaption depends only on the joint distribution of these forward rates. Under our model, with given initial conditions $L_{i,0}$ and correlation $R = (\rho_{i,j})$, the swaption price depends on $\sigma_i(t), \dots, \sigma_{j-1}(t)$, $t \in [0, T_i]$ only. The dynamic of these forward rates beyond the $t > T_i$ and all other forward rates do not influence the swaption price.

If we discretize the volatility function corresponding to the tenor structure and define

$$\sigma_{k,l} := \left(\frac{1}{T_{l+1} - T_l} \int_{T_l}^{T_{l+1}} \sigma_k^2(t) dt \right)^{1/2},$$

the price of an option on the swap rate $S(T_i, \dots, T_j; T_i)$ depends only on $\sigma_{k,l}$ for $k = i, \dots, j-1$ and $l = 0, \dots, i-1$. This allows an iterative calculation of $\sigma_{k,l}$ from given swaption market prices:

For $i = 1, \dots, n-1$:

For $j = i+1, \dots, n$:

Calculate $\sigma_{j-1,i-1}$ from the price of an option on $S(T_i, \dots, T_j; T_i)$ by considering the already calculated $\sigma_{k,l}$ with $k = i, \dots, j-1$ and $l = 0, \dots, i-2$ from the previous iterations.

To derive $\sigma_{j-1,i-1}$ from the market price $V_{\text{swaption}}^{\text{market}}(T_i, \dots, T_j)$ we have to invert the mapping

$$\sigma_{j-1,i-1} \mapsto V_{\text{swaption}}^{\text{model}}(T_i, \dots, T_j).$$

In principle this mapping may be realized by a Monte-Carlo evaluation of the swaption. To allow for a faster pricing, and thus a faster calibration, analytic approximation

formulas for swaption prices within a LIBOR Market Model have been derived, see also Section 17.1.4.4.

Remark 195 (Bootstrapping): The above procedure of calculating a piecewise constant instantaneous volatility from swaption prices is called *volatility bootstrapping*.

Remark 196 (Review: Over fitting): The calculation of a piecewise constant volatility function $\sigma_{i,j}$ from swaption prices bears the risk of an *over fitting* of the model. Note that if one applies this procedure then one trusts the validity of every single given swaption price, i.e. that the prices given show a sufficiently good quality with respect to *topicality* (fixing time of the price) and *liquidity*. If not all prices are given in the same quality, then some have to be interpolated or corrected by hand. In this case, the calibration problem has been replaced by an interpolation problem. If the interpolation and maintenance of the market data is not done with care, the calibration exactly fitting to these prices may be useless. See for example Chapter 6.

In addition, the bootstrapping of the instantaneous volatility from swaption prices does not allow for a weighting of the swaption prices according to their importance.

A solution to this is the reduction of the free parameters by a reduction of the family of admissible volatility functions. This will induce a loss of the ability of a perfect fit to market prices which is a desired feature.

Functional Forms To reduce the risk of an *over fitting* the admissible volatility functions may be restricted to a parametrized family of volatility functions. For example, a functional which is empirically motivated by the historical shapes of the volatility and which is common in practice is

$$\sigma_i(t) := (a + b \cdot (T_i - t)) \cdot \exp(-c \cdot (T_i - t)) + d.$$

Given a functional form, the calibration of the model consists of a selection of (liquid) market prices of caps and swaptions and an optimization of the remaining parameters (e.g. a, b, c, d above) to fit the model prices to the market prices.

For a detailed discussions of a robust calibration to cap and swaption prices we refer to the literature, especially [8, 28].

17.1.4.3. Choice of the Correlations

Factors We assumed in (17.1) a models in which (potentially) each forward rate L_i **Brownian motion** W_i . The model is driven by an n -dimensional⁸ Brownian motion

$$W = \begin{pmatrix} W_0 \\ W_1 \\ \dots \\ W_{n-1} \end{pmatrix}.$$

The effective number of *factors*, i.e. the number of independent Brownian motions, that are driving the model is determines by the correlation

$$\rho_{i,j}(t)dt = dW_i(t) \cdot dW_j(t).$$

By an eigenvector decomposition (PCA) of the correlation matrix $R = (\rho_{i,j})_{i,j=1\dots n}$ we may represent dW as

$$dW(t) = F(t)dU(t),$$

where $U := (U_1, \dots, U_m)^T$ und $U_1, \dots, U_m$ denote independent Brownian motions and $F = (f_{i,j})$ denotes a $n \times m$ -matrix. In other words we have

$$dW_i(t) = \sum_{j=1}^m f_{i,j}(t)dU_j(t), \quad \rho_{i,j}(t) = \sum_{k=1}^m f_{i,k}(t) \cdot f_{j,k}(t).$$

A proof of this representation may be found in the Appendix A.4. Note that we may have $m < n$. The columns of the matrix F are called *factors*.

Functional Forms A full rank correlation matrix R is hard to derive from market instruments. As before a common procedure is to reduce the family of admissible correlation matrixes R . One ansatz consists of functional forms, for example

$$\rho_{i,j}(t) := \exp(-\alpha \cdot |i - j|). \quad (17.19)$$

Factor Reduction The specification of the correlation matrix as a functional form is usually followed by a reduction of the number of factors. This is done in a so called factor reduction (PCA, *Principal Component Analysis*). There only the eigenvectors corresponding to the largest eigenvalues of R are considered and a new correlation matrix is build from these selected factors. For a discussion of the factor reduction see Appendix A.5.

The advantage of the factor reduction is that after it only an m -dimensional Brownian motion has to be simulated (and not an n -dimensional Brownian motion). Often

⁸ An $n - 1$ -dimensional Brownian motion is sufficient here, since we may choose $W_0 = 0$, because the forward rate L_0 is not stochastic. It is fixed in $T_0 = 0$. Formally we achieve this by setting $\sigma_0 \equiv 0$.

$n \geq 40$ is required (e.g. a 20 year interest rate curve with semi-annual periods), however often $m \leq 5$ is sufficient. The choice of the actual number m of factors depends on the application, see Chapter 22.

Calibration The correlation model, e.g. the free parameter a in (17.19), may be chosen such that the fit of model prices to given market prices is improved. Alternatively it may be chosen to give more realistic interest rate correlations. See also Remark 196.

It should be stressed that we calibrate the *instantaneous* correlation, i.e. the correlation of the Brownian increments dW , and not the terminal correlation, i.e. the correlations of the distribution of the interest rates at a fixed time. We will consider the relation of the two in Chapter 18.

17.1.4.4. Covariance Matrix, Calibration by Parameter Optimization

In the previous sections we considered volatility and correlation separately. This is not necessary, both may be considered together in the form of the correlation matrix $(\sigma_i \sigma_j \rho_{i,j})$. Thus the calibration problem consists of the calculation of the (market implied) covariance matrix (or covariance matrix function).

Restricting to parametrized functional forms for volatility and correlation, e.g. as

$$\sigma_i(t) := (a + b \cdot (T_i - t)) \cdot \exp(-c \cdot (T_i - t)) + d, \quad \rho_{i,j}(t) := \exp(-\alpha \cdot |i - j|)$$

reduces the number of degrees of freedoms of the covariance model and thus the possible number of products for which an exact fit is possible. This might be a desired feature, e.g. to avoid an over fittings. A disadvantage is the lack of transparency of the parameters. To derive the parameters a numerical optimization has to be used in general, e.g. the minimization of a suitable norm of the error vector of some selected product prices as a function of the model parameters. The optimization of volatility parameters and correlation parameters may occur jointly, i.e. we consider a functional form of the covariance structure.

17.1.4.5. Analytic Evaluation of Caplets and Swaptions

To calculate the calibration error we need to calculate the corresponding model prices. Since a numerical calculation of the model price (e.g. by a full Monte-Carlo simulation) is time consuming, and since the optimization requires many calculation of model prices, there is a need for fast analytical pricing formulas of specific calibration products.

Analytic Evaluation of a Caplet in the LIBOR Market Model The analytic evaluation of caplets in the LIBOR market model by the Black formula (10.2) using (17.17) to calculate the Black volatility.

Analytic Evaluation of a Swaption in the LIBOR Market Model The analytic evaluation of swaption in the LIBOR market model is possible by a an approximation formula only. An approximation formula may be derived by expressing the volatility of the swap rate as a function of the volatility and correlation of the forward rates. Assuming a lognormal model for the swap rate, which is an already an approximation, one may then apply the Black formula for Swaptions.⁹ Corresponding approximation formulas may be found in [8].

17.1.5. Interpolation of Forward Rates in the LIBOR Market Model



Motivation: An implementation of the LIBOR Market Model (e.g. as Monte-Carlo simulation) allows to calculate the forward rate $L(T_j, T_k; t_i)$ for interest rate periods $[T_j, T_k]$, $T_j < T_k$ and fixing times t_i .¹⁰ Using these rates we may evaluate almost all interest rate derivatives that may be represented as a function of these rates.

The discretization of the simulation time $\{t_i\}$ determines at which times we may have interest rates fixings. The discretization of the tenor structure $\{T_j\}$ determines for which periods forward rates are available and, since the Numéraire is only defined at $t = T_j$, it determines at which times we may have payment dates. The tenor structure imposes a significant restriction, since a change of the tenor structure is essentially a change of the model.

In practice it is desired to calculate as many financial products as possible with the same model. Two reasons should be mentioned: First, the aggregation of risk measures, i.e. of sensitivities¹¹ of products to the sensitivity of a portfolio, is correct only if the product sensitivities have been calculated under the same model. Second, the setup of a pricing model (calibration, generation of Monte-Carlo paths) requires usually much more calculation time as the evaluation of a product, i.e., eventually, it is very efficient to reuse a model.

Thus, it is desirable to know how to calculate from a given LIBOR market model the quantities $L(T^s, T^e; t)$ for $T^s, T^e \notin \{T_0, T_1, \dots, T_n\}$ (unaligned period) and / or $t \notin \{t_0, t_1, \dots, t_n\}$ (unaligned fixing). <|

17.1.5.1. Interpolation of the Tenor Structure $\{T_i\}$

We consider the question on how to interpolate the tenor structure, i.e. we derive an expression for $L(T^s, T^e; t)$ for $T^s \notin \{T_0, T_1, \dots, T_n\}$ and/or $T^e \notin \{T_0, T_1, \dots, T_n\}$. Let

⁹ Assuming lognormal processes for the forward rates, the swap rate is not a lognormal process in general.

¹⁰ In a Monte-Carlo Simulation the rates carry, of course, some approximation error of the time discretization scheme.

¹¹ A sensitivity is a partial derivative of the product price by a model- or product-parameter (e.g. volatility or strike).

$T^s < T^e$.

The forward rates $L(T^s, T^e; t)$ may be derived from corresponding bonds $P(T; t)$. We have

$$1 + L(T^s, T^e; t) \cdot (T^e - T^s) = \frac{P(T^s; t)}{P(T^e; t)}$$

For arbitrary $T > t$ the bond $P(T; t)$ is given by

$$P(T; t) = N(t) \cdot \mathbb{E}^{\mathbb{Q}^N} \left(\frac{1}{N(T)} \middle| \mathcal{F}_t \right).$$

The definition of the Numéraire of the LIBOR Market Model shows that the specification of the short period bond $P(T_{m(t)+1}; t)$ is sufficient (and necessary) to determine all bonds $P(T; t)$ and thus all forward rates.¹²

Assumption 1: No stochastic shortly before maturity. We assume that $P(T_{m(t)+1}; t)$ is a $\mathcal{F}_{T_{m(t)}}$ -measurable random variable, i.e. the bond has no volatility at time t with $T_{m(t)} \leq t \leq T_{m(t)+1}$, i.e. shortly before its maturity. With other words: $t \mapsto P(T_{m(t)+1}; t)$ is a deterministic interpolation functions on quantities known at the period start $T_{m(t)}$. In this case we have for the bond with maturity t , seen at time s with $T_{m(t)} \leq s \leq t \leq T_{m(t)+1}$:

$$\begin{aligned} P(t; s) &= N(s) \mathbb{E}^{\mathbb{Q}^N} \left(\frac{1}{N(t)} \middle| \mathcal{F}_s \right) \\ &= N(s) \mathbb{E}^{\mathbb{Q}^N} \left(\frac{P(T_{m(t)+1}; t)}{N(t)} \cdot \frac{1}{P(T_{m(t)+1}; t)} \middle| \mathcal{F}_s \right) \\ &= N(s) \mathbb{E}^{\mathbb{Q}^N} \left(\frac{P(T_{m(t)+1}; t)}{N(t)} \middle| \mathcal{F}_s \right) \cdot \frac{1}{P(T_{m(t)+1}; t)} = \frac{P(T_{m(t)+1}; s)}{P(T_{m(t)+1}; t)} \end{aligned} \quad (17.20)$$

Especially for $s = T_{m(t)}$

$$P(t; T_{m(t)}) = \frac{P(T_{m(t)+1}; T_{m(t)})}{P(T_{m(t)+1}; t)}. \quad (17.21)$$

Thus we see that (under Assumption 1), the interpolation function $t \mapsto P(t; T_{m(t)})$ (interpolation of the maturity t) may be derived directly from the interpolation function $t \mapsto P(T_{m(t)+1}; t)$ (interpolation of the evaluation time t) and vice versa. The functions are reciprocal.

¹² Note that the considerations on interpolation given in this section do not assume a LIBOR Market Model. They are valid in general.

Assumption 2: Linearity shortly before maturity. If the chosen interpolation function $T \mapsto P(T; T_{m(t)})$ is linear, then the interpolation of bond prices $P(T; s)$ seen in $s < T_{m(t)}$ is linear too. This follows directly from the linearity of the expectation:

$$P(T; s) = N(s) \mathbb{E}^{\mathbb{Q}^N} \left(\frac{P(T; T_{m(t)})}{N(T_{m(t)})} \middle| \mathcal{F}_s \right). \quad (17.22)$$

By this, the linear interpolation takes a distinct role. With $P(T_{m(t)}; T_{m(t)}) = 1$ and $P(T_{m(t)+1}; T_{m(t)}) = \frac{1}{1 + L_{T_{m(t)}}(T_{m(t)}) \cdot (T_{m(t)+1} - T_{m(t)})}$ the linear interpolation of $t \mapsto P(t; T_{m(t)})$ follows as

$$\begin{aligned} P(t; T_{m(t)}) &= \frac{T_{m(t)+1} - t}{T_{m(t)+1} - T_{m(t)}} P(T_{m(t)}; T_{m(t)}) + \frac{t - T_{m(t)}}{T_{m(t)+1} - T_{m(t)}} P(T_{m(t)+1}; T_{m(t)}) \\ &= \frac{1 + L_{T_{m(t)}}(T_{m(t)}) \cdot (T_{m(t)+1} - t)}{1 + L_{T_{m(t)}}(T_{m(t)}) \cdot (T_{m(t)+1} - T_{m(t)})}. \end{aligned} \quad (17.23)$$

The corresponding interpolation for the short period bond $P(T_{m(t)+1}; t)$ thus with (17.21)

$$P(T_{m(t)+1}; t) := \frac{1}{1 + L_{T_{m(t)}}(T_{m(t)}) \cdot (T_{m(t)+1} - t)}.$$

Applying (17.22) to the (in t) linear interpolation (17.23) we find for all $s \leq T_{m(t)}$

$$\frac{P(t; s)}{P(T_{m(t)}; s)} = \frac{1 + L_{T_{m(t)}}(s) \cdot (T_{m(t)+1} - t)}{1 + L_{T_{m(t)}}(s) \cdot (T_{m(t)+1} - T_{m(t)})},$$

and thus

$$\frac{P(t; s)}{P(T_{m(t)+1}; s)} = 1 + L_{T_{m(t)}}(s) \cdot (T_{m(t)+1} - t).$$

From (17.20) we have for $T_{m(t)} \leq s \leq t$

$$P(t; s) = \frac{1 + L_{T_{m(t)}}(T_{m(t)}) \cdot (T_{m(t)+1} - t)}{1 + L_{T_{m(t)}}(T_{m(t)}) \cdot (T_{m(t)+1} - s)}.$$

We summarize this result in a theorem:

Theorem 197 (Interpolation of Forward Rates on unaligned Periods): Given a tenor structure $T_0 < T_1 < T_2 < \dots$. For all t let the short bond $P(T_{m(t)+1}; t)$ be given by the interpolation

$$P(T_{m(t)+1}; t) := \frac{1}{1 + L_{T_{m(t)}}(T_{m(t)}) \cdot (T_{m(t)+1} - t)}. \quad (17.24)$$

Then we have for arbitrary $t \leq T$ with $k = m(t)$, $l = m(T)$

$$\begin{aligned} P(T; t) &= P(T_{k+1}; t) \cdot \prod_{j=k+1}^l \frac{P(T_{j+1})}{P(T_j)} \cdot \frac{P(T; t)}{P(T_{l+1}; t)} \\ &= (1 + L_k(T_k) \cdot (T_{k+1} - t))^{-1} \cdot \prod_{j=k+1}^l (1 + L_j(t) \cdot (T_{j+1} - T_j))^{-1} \cdot (1 + L_l(t^*) \cdot (T_{l+1} - T)), \end{aligned}$$

where $t^* = \min(t, T_l)$. For $T_s \leq T_e$ with $T_k < T_s \leq T_{k+1}$ and $T_l \leq T_e < T_{l+1}$ we have

$$\begin{aligned}
 & 1 + L(T_s, T_e; t) \cdot (T_e - T_s) \\
 &= \frac{P(T_s; t)}{P(T_e; t)} = \frac{P(T_s; t)}{P(T_{k+1}; t)} \cdot \prod_{i=k+1}^l \frac{P(T_i; t)}{P(T_{i+1}; t)} \cdot \frac{P(T_{l+1}; t)}{P(T_e; t)} \\
 &= (1 + L_k(t^*) \cdot (T_{k+1} - T_s)) \cdot \prod_{i=k+1}^l (1 + L_i(t) \cdot (T_{i+1} - T_i)) \cdot (1 + L_l(t^*) \cdot (T_{l+1} - T_e))^{-1}.
 \end{aligned} \tag{17.25}$$



Experiment: At <http://www.christian-fries.de/finmath/applets/LMMPricing.html> several interest rate products may be priced using a LIBOR market model. <|

17.2. Object Oriented Design

Figure 17.2 and 17.3 shows an object oriented design of a *Monte-Carlo LIBOR Market Model*. The following important aspects are considered in the design:

- **Reuse of Implementation**
- **Separation of Product and Model**
- **Abstraction of Model Parameters**
- **Abstraction of Calibration**

We will describe these aspects in the following.

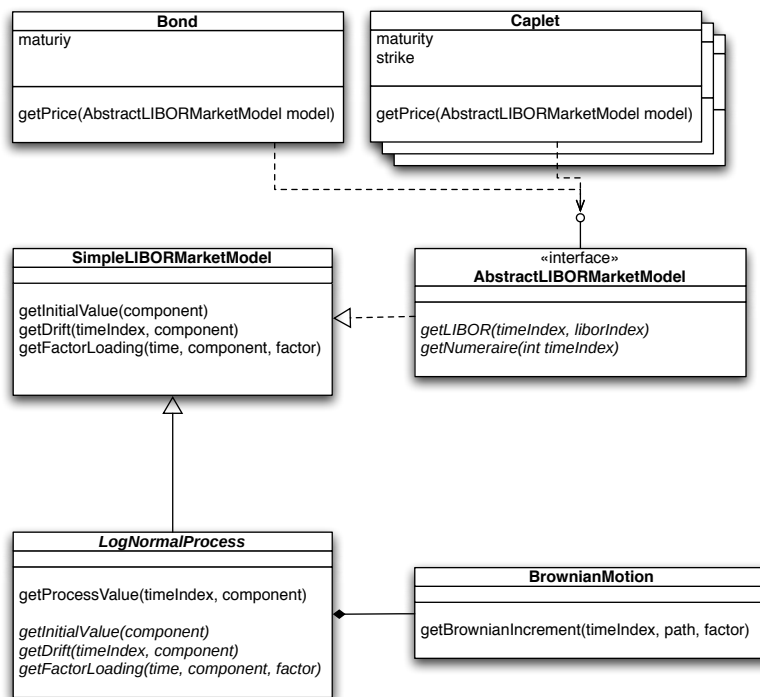


Figure 17.2.: UML Diagram: Evaluation of LIBOR related products in a LIBOR Market Model via Monte Carlo Simulation.

17.2.1. Reuse of Implementation

For the Monte-Carlo simulation of the lognormal process we use the same classes as in the example of the Black-Scholes model, see Figure 13.4 on Page 187. To do so

the classes `BROWNIANMOTION` and `LOGNORMALPROCESS` were from the beginning designed for vector-valued, i.e. multi-factorial processes, although the Black-Scholes model does not require it. Improvements to the classes `BROWNIANMOTION` and `LOGNORMALPROCESS` will result in improvement of both applications.

17.2.2. Separation of Product and Model

The interface `ABSTRACTLIBORMARKETMODEL` defines how LIBOR related products communicate with a Monte Carlo LIBOR model. Through this interface the model serves to the product the process of the underlyings (the forward rates) and the Numéraire as a Monte Carlo simulation. All corresponding Monte-Carlo evaluations of interest rate products expect this interface. All corresponding Monte-Carlo LIBOR models implement this interface. This realizes a separation of product and model. The specific *LIBOR Market Model* is realized through the class `SIMPLELIBORMARKETMODEL`. Model extensions may be introduced here, without a need to changed classes that realize LIBOR related Monte-Carlo products.

17.2.3. Abstraction of Model Parameters

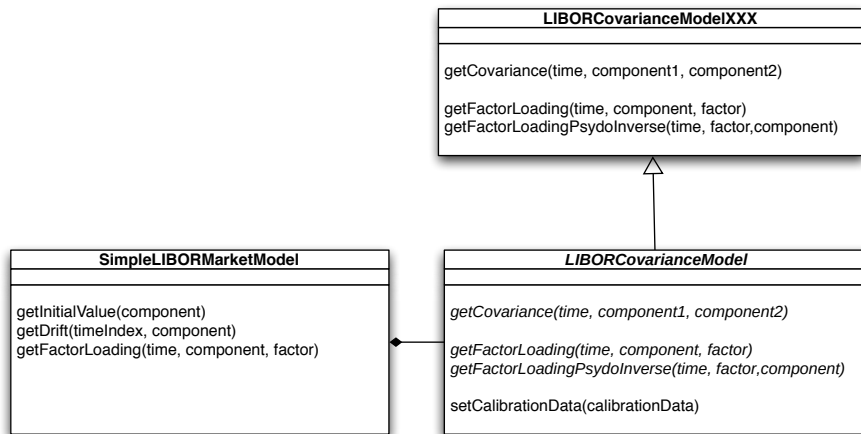


Figure 17.3.: UML Diagram: LIBOR Market Model: Abstraction of model parameters.

The model parameters, i.e. the covariance structure, are encapsulated in their own classes. The model parameter classes implement a simple *Interface* `LIBORCOVARIANCEMODEL`. A specific covariance model $(i, j, t) \mapsto \gamma_{i,j}(t) = \sigma_i(t)\sigma_j(t)\rho_{i,j}(t)$ is realized through a class that implements the interface `LIBORCOVARIANCEMODEL`. This class is then served to the model. The interfaces are designed such that $(i, j, t) \mapsto \gamma_{i,j}(t)$ may

be stochastic.¹³ See Figure 17.3.

This abstraction of model parameters allows to easily exchange different modelings of covariance, i.e. volatility and correlation.

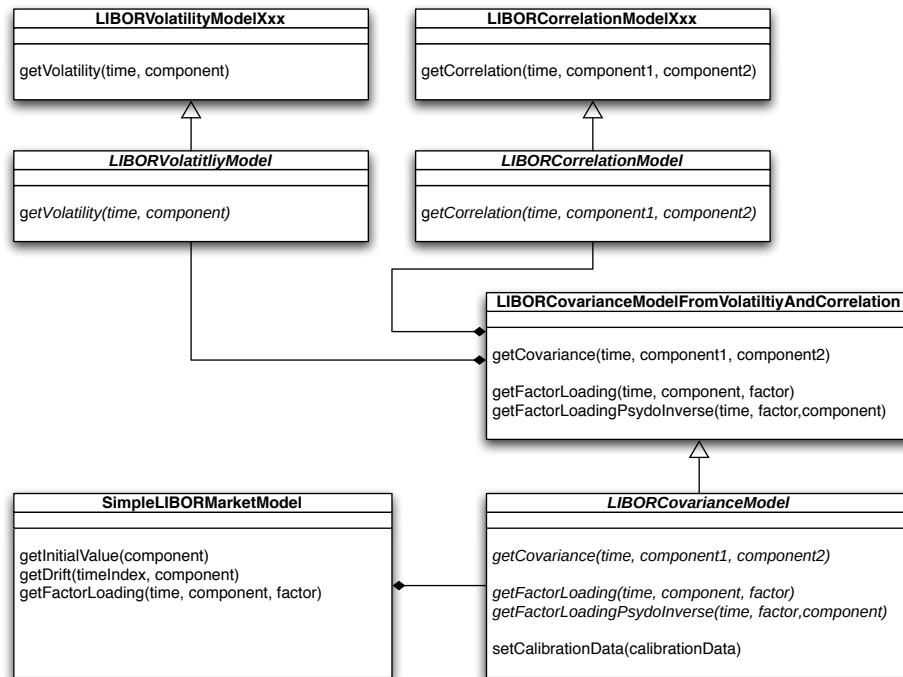


Figure 17.4.: UML Diagram: LIBOR Market Model: Abstraction of model parameters as volatility and correlation. Introducing separate classes for volatility and correlation will bring some disadvantages for joint calibration and general covariance modeling. The design above might be considered over-designed.



Warning: In the case where the covariance structure is modeled by volatility and correlation it seems reasonable to define corresponding interfaces `LIBORVOLATILITYMODEL` and `LIBORCORRELATIONMODEL`. A simple class `LIBORCOVARIANCEMODELFROMVOLATILITYANDCORRELATION` calculates the factor loadings and covariances from given volatility and correlation models. See Figure 17.4. However, the separation of volatility and correlation in their own classes will bring some disadvantages for a joint calibration and general covariance modeling. The corresponding code may become over-designed. The design in Figure 17.4 would make sense if one would like to explore many combinations of different volatility and correlation models.

¹³ A stochastic volatility model would result in a stochastic covariance model.

17.2.4. Abstraction of Calibration

The abstraction of model parameters allows for the abstraction of calibration. The algorithm calibrating the covariance model is clearly a part of the covariance model. Thus each covariance model object may carry calibration data (e.g. market data) that, once set, is used to calibrate the model. The calibration data itself may be anything from given correlation and volatility parameters to a list of product with associated target values.

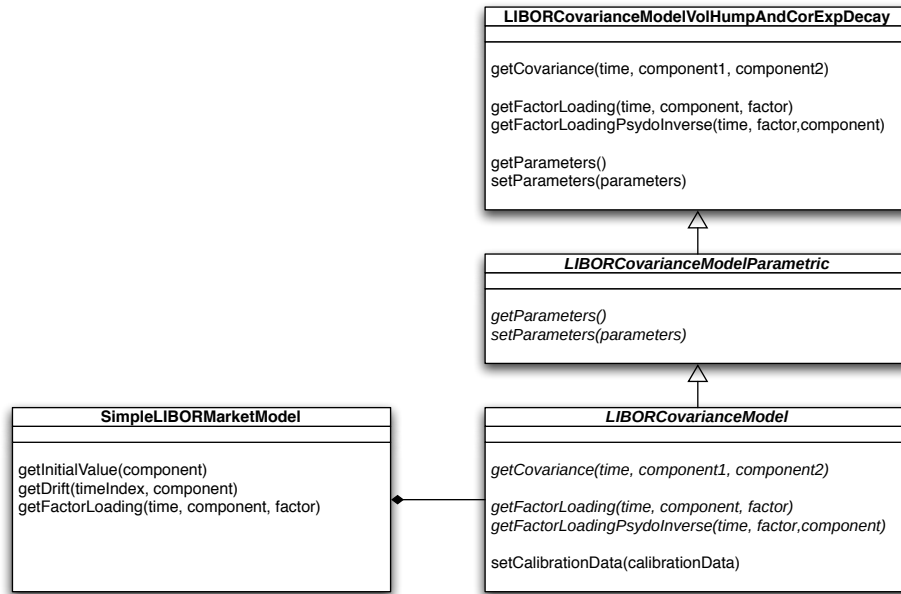
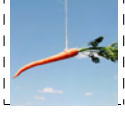


Figure 17.5.: UML Diagram: LIBOR Market Model: Abstraction of model parameters: Parametric covariance models.

17.3. Swap Rate Market Models (Jamshidian 1997)



Motivation: The [LIBOR Market Model](#) postulates as log-normal dynamic for the [forward rate](#) $L_i := L(T_i, T_{i+1})$. In other words, each single forward rate follows a [Black model](#). This allowed for an easy calibration of the IBOR market model to Caplet prices. We only have to fulfill condition (17.18).

If, however, swap options (i.e. swaption or swaption related products like Bermudan swaptions) are in the focus, then a model that simulates the swap rate directly appears to be the first choice.¹⁴ If, for example, the swap rate follows a log-normal process, then the corresponding swaptions may be calibrated by a simple condition involving the implied Black-Volatility of the swap rate. $\triangleleft$

$$S_{i,k}(t) := S(T_i, T_k; t) := \frac{P(T_i; t) - P(T_k; t)}{\sum_{j=i}^{k-1} (T_{j+1} - T_j) \cdot P(T_{j+1}; t)}, \quad k > i. \quad (17.26)$$

Since the set of swap rates defined for a given tenor structure $T_0 < T_1 < \dots < T_n$ is a two parametric family of $\frac{(n-1)n}{2}$ rates which are related by functional dependencies, a meaningful dynamic may be given only for a subset of swap rates.¹⁵

Upon choosing a system of swap rates $S_{i,k}$, for which we like to specify the dynamics, we have to take care that the system is neither over-determined or, with respect to the given tenor structure, under-determined. The system of rates has to consist of n swap rates since on the tenor structure $0 = T_0 < T_1 < \dots < T_n$ we have n degrees of freedoms in terms of bond prices.

Two common variants are given by the set of *co-sliding* swap rates and *co-terminal* swap rates, see Table 17.1. In the specification of co-sliding swap rates it is necessary

$$\text{Co-Sliding:} \quad S_{i, \min(i+k, n)} \quad i = 0, 1, \dots, n-1$$

$$\text{Co-Terminal:} \quad S_{i, n} \quad i = 0, 2, \dots, n-1$$

Table 17.1.: *co-sliding and co-terminal swap rates*

to close the system. In our definition we achieve this by first considering the swap rates $S_{i,i+k}$ with k periods (co-sliding, $i < n-1-k$) and starting with $i = n-k$ we consider co-terminal swap rates.

¹⁴ Later, we will give an argument why a forward rate based model might be the choice even for swap rate related products, see Remark 199.

¹⁵ For example, the swap rate $S_{i,i+4}$ is a function of the swap rates $S_{i,i+2}$ and $S_{i+2,i+4}$, which in turn are functional for the swap rates $S_{i,i+1}, \dots, S_{i+3,i+4}$. The swap rates with one period are forward rates $L_i = S_{i,i+1}$.

If the selection of swap rates is made, we model each $S_{i,k}$ from the selection as a log-normal process:

$$dS_{i,k}(t) = \mu_{i,k}^{\mathbb{P}}(t)S_{i,k}(t)dt + \sigma_{i,k}(t)S_{i,k}(t)dW_{i,k}^{\mathbb{P}}(t) \quad \text{unter } \mathbb{P}, \quad (17.27)$$

with initial conditions

$$S_{i,k}(0) = S_{i,k,0}.$$



Interpretation: The modeling of co-terminal swap rates is a suitable choice if, e.g., we have to price Bermudan swaptions, which have these swap rates as underlying. The modeling of co-sliding swap rates is a suitable choice if we have to price products in which swap rates with constant time to maturity (CMS rate¹⁶) enter. $\triangleleft$

17.3.1. The Swap Measure

If we consider the Definition of the swap rate in (17.26), then it is apparent that $S_{i,k}$ is a martingale under the martingale measure $\mathbb{Q}^N$ corresponding to the Numéraire

$$N(t) := A(T_i, \dots, T_k; t) := \sum_{j=i}^{k-1} (T_{j+1} - T_j) \cdot P(T_{j+1}; t), \quad k > i, t \leq T_{i+1}. \quad (17.28)$$

The right hand side in (17.28) is a portfolio of zero bonds and thus a traded product and the swap rate is the N -relative price of a traded product.

Definition 198 (Swap Measure, Swap Annuity):

The equivalent martingale measure $\mathbb{Q}^N$ corresponding to the Numéraire N in (17.28) is called *swap measure* corresponding to the swap rate $S(T_1, \dots, T_k)$. The expression on the right hand side in (17.28) is called *swap annuity*.

The Numéraire is, so far, defined for $t \leq T_{i+1}$ only, since at $t = T_{i+1}$ the first bond $P(T_{i+1})$ is at its maturity and we have to specify how its payment has to be reinvested.¹⁷ A continuation of the Numéraire definition to $t > T_{i+1}$ may be given by a re-investment into the next swap annuity. This is the analog to the Numéraire (17.9) of the *Spot Measures*. For $i = 1, \dots, k-1$ we have

$$N(t) = A(T_i, \dots, T_k; t) \cdot \prod_{j=1}^{i-1} \frac{A(T_j, \dots, T_k; T_{j+1})}{A(T_{j+1}, \dots, T_k; T_{j+1})}, \quad T_{i-1} \leq t < T_i$$

¹⁶ See Definition 147.

¹⁷ The reinvestment determines the evolution of the Numéraire for $t > T_{i+1}$: For example, if we compare the investment of the paid 1 in $\frac{1}{P(T_k; T_{i+1})}$ parts of a T_k -bond with the investment in $\frac{1}{P(T_{k+1}; T_{i+1})}$ parts of a T_{k+1} -bonds, then the evolution of the Numéraire will differ by the evolution of the T_k forward rate, i.e. by the factor $\frac{P(T_k; t)}{P(T_k; T_{i+1})} / \frac{P(T_{k+1}; t)}{P(T_{k+1}; T_{i+1})} = \frac{1+L(T_k, T_{k+1}; t)(T_{k+1}-T_k)}{1+L(T_k, T_{k+1}; T_{i+1})(T_{k+1}-T_k)}$.

where $T_0 := 0$. The swap rates we consider here are co-terminal. Of course we may consider co-sliding swaps in a similar way, using the swap annuities $A(T_j, \dots, T_{j+k}; t)$. The corresponding Numéraire of re-investment in co-sliding swap annuities, i.e. a *rolling co-sliding swap annuity* then is

$$N(t) = A(T_i, \dots, T_{i+k}; t) \cdot \prod_{j=1}^{i-1} \frac{A(T_j, \dots, T_{j+k}; T_{j+1})}{A(T_{j+1}, \dots, T_{j+1+k}; T_{j+1})}, \quad T_{i-1} \leq t < T_i.$$

For $k = i + 1$ this corresponds to (17.9).

17.3.2. Derivation of the Drift Term

For the swap rate market model we have multiple sets of swap rates, which may be modeled and (as in the LIBOR Market Model) multiple possible choices of Numéraires. In this section we do not present a detailed derivation of the drift terms. The derivation is done similar to the derivation of the drift in the LIBOR Market Model by expressing a martingale through the elementary swap rate processes $S_{i,j}$. If for example $A_{k,l}$ is the Numéraire we consider the $\mathbb{Q}^{A_{k,l}}$ -martingale $(S_{i,j} \cdot \frac{A_{i,j}}{A_{k,l}})$.

17.3.3. Calibration - Choice of the free Parameters

17.3.3.1. Choice of the Initial Conditions

Reproduction of Bond Market Prices or Swap Market Prices If we set in the definition of the swap rate (17.26) t to the preset time, i.e. $t = 0$ following our convention, then we get an equation relating today's bond prices with today's swap $S_{i,k}(0)$ and the latter are just the initial conditions of the chosen swap rate processes. Thus, the initial conditions of the processes are given by (17.26) with $t = 0$ and today's bond prices, i.e. today's interest rate curve.

Although we regard the family of zero bonds as the natural description of the interest rate curve and we see swap rates and swap prices as derived quantities, it is in this case natural to calculate today's swap rates directly from today's swap prices (assuming they are given). In this case the initial conditions are given by today's swap prices. With this choice, the model will reproduce these prices.

17.3.3.2. Choice of the Volatilities

Reproduction of Swaption Market Prices The calibration of the model to swaption prices is analog to the calibration of the LIBOR Market Model to caplet prices. Let the dynamic of the swap rate $S_{i,k}$ be given by (17.27). Furthermore let

$\sigma_{i,k}^{\text{Black,Market}}$ denote the market prices of an option on $S_{i,k}$ given as implied Black-volatility. If we calculate

$$\sigma_{i,k}^{\text{Black,Model}} := \left(\frac{1}{T_i} \int_0^{T_i} \sigma_{i,k}^2(t) dt \right)^{1/2},$$

then the model reproduces the given swaption market prices if

$$\sigma_{i,k}^{\text{Black,Model}} = \sigma_{i,k}^{\text{Black,Market}}.$$

This statement is trivial, since, if we consider only a single swap rate $S_{i,k}$, then (17.27) is a Black model for this swap rate and under this model the implied volatility is defined by inverting the pricing formula. The inversion of the pricing formula is what a calibration should achieve.

Remark 199 (LIBOR Market Model versus Swaprate Market Model): The question, whether one should choose a LIBOR Market Model or a Swaprate Market Model seems to depend on the application only, to be precise, on the question if the model should calibrate to caplet or swaptions - and if one sees log-normale forward rate or a log-normalen swap rate as realistic model.¹⁸

The criteria that defines the choice of the model thus is the quality of the model calibration to the specific application.

However, the swap rate market model has a disadvantage compared to the LIBOR market model: If we calculate a forward rate L_i in a swap rate market model, then the forward rate tends to suffer from numerical instabilities. Conversely the calculation of a swap rate from forward rates models in a LIBOR Market Model is in general much more stable.



Interpretation: The reason lies in the representation of the swap rate as convex combination of the forward rates. By Lemma 112 we have

$$S_{i,j} = \sum_{k=i}^{j-1} \alpha_k^{i,j} L_k, \quad \text{mit } \alpha_k^{i,j} \geq 0 \quad \text{und} \quad \sum_{k=i}^{j-1} \alpha_k^{i,j} = 1,$$

with $\alpha_k^{i,j} := \frac{P(T_{k+1}) \cdot (T_{k+1} - T_k)}{\sum_{k=i}^{j-1} P(T_{k+1}) \cdot (T_{k+1} - T_k)}$.

If we calculate a forward rate L_i froms (e.g. co-terminal) swap rates $S_{j,n}$ we have

$$\begin{aligned} L_i &= \frac{1}{\alpha_i^{i,n}} S_{i,n} - \frac{1}{\alpha_{i+1,n}} S_{i+1,n+1} \\ &= \frac{1}{\alpha_i^{i,n}} (S_{i,n} - S_{i+1,n}) + \left(\frac{1}{\alpha_i^{i,n}} - \frac{1}{\alpha_{i+1,n}} \right) S_{i+1,n}. \end{aligned}$$

¹⁸ In general both assumptions do not hold and it is necessary to modify the models with respect to their distribution assumption. Such a modification of the model is called *smile modeling*.

Assuming for simplicity $\alpha_j^{i,n} = \frac{1}{n-i-1}$, which is with $\sum_{j=i}^{n-1} \alpha_j^{i,n} = 1$ plausible¹⁹, then we have

$$S_{i,n} = \frac{1}{n-i-1} \sum_{k=i}^{n-1} L_k$$

$$L_i = (n-i-1) \cdot (S_{i,n} - S_{i+1,n}) + S_{i+1,n}.$$

This shows:

- The calculation of a swap rate $S_{i,n}$ from forward rates L_k corresponds to the calculation of an average (rate) - the swap rate may be interpreted as integral of the forward rates. Errors in L_k are averaged and thus smoothed. The variance of an unsystematic error is reduced.
- The calculation of a forward rate L_i from swap rates $S_{i,n}, S_{i+1,n}$ consists of a finite difference term - this part of the forward rate may be interpreted as a derivative. The calculation of a difference is sensible to errors in the swap rates (e.g. small jumps) and the error is scaled up by the factor $(n-i-1)$ for n large and i small. Thus forward rates for short periods in a model of long period swap rates have a tendency to numerical instability.

◁|



Tip: If there is no strong reason for a swap rate market model, one should prefer a *generic* LIBOR Market Model and calculate the corresponding swap rates from forward rates. This allows to have a single, thus consistent model for multiple applications (produces), which allows the aggregation of risk parameters (**Delta**, **Gamma**). The difference in the distributional properties is often to be negligible (see [8]).

◁|



Further Reading: The original articles on the LIBOR market model are [48] and [78]. On the swap rate market model see [72]. On the calibration of the LIBOR Market Model see [8, 28]. On the arbitrage free discretization see [64]. On the interpolation of forward rates see [84]. The evaluation of Bermudan options in Monte-Carlo is considered in Chapter 14, see also [42, 43].

We will use the LIBOR Market model as foundation for further investigations of general interest rate model properties. In Chapter 18 we will investigate the instantaneous correlation $\rho_{i,j}$ and volatility σ_i and their effect on terminal correlation. In Chapter 22 we will investigate the mean reversion and multi factorality on the shape of interest rate curve.

◁|

¹⁹ Indeed we have $\frac{1}{\alpha_i^{i,n}} = \frac{P(T_{i+2})(T_{i+2}-T_{i+1})}{P(T_{i+1})(T_{i+1}-T_i)} \left(\frac{1}{\alpha_i^{i+1,n}} + 1 \right)$.

CHAPTER 18

Excursion: Instantaneous Correlation and Terminal Correlation

In this chapter we use the [LIBOR Market Model](#) to discuss the influence of instantaneous volatility and instantaneous correlation on option prices. Although our study is based on the LIBOR market model, the intuition gained from our experiments is of general importance.

We experiment with different (extreme) parameter configurations and we will see that a single-factor model, in which all interest rates $L(T_i, T_{i+1})$ move (instantaneously) perfectly correlated, however, may exhibit at time $t > 0$ (terminal) perfectly decorrelated random variables $L(T_j, T_{j+1}; t)$, $L(T_k, T_{k+1}; t)$.

We start by fixing some notions:

18.1. Definitions

Definition 200 (Covariance, Correlation):

Let X, Y denote two (numeric) random variables, $\bar{X} = E(X)$, $\bar{Y} = E(Y)$. Then

$$\text{Cov}(X, Y) := E((X - \bar{X}) \cdot (Y - \bar{Y}))$$

is called the *covariance* of X and Y , $\text{Var}(X) := \text{Cov}(X, X)$ is called the variance of X and

$$\text{Cor}(X, Y) := \frac{E((X - \bar{X}) \cdot (Y - \bar{Y}))}{\sqrt{\text{Var}(X)} \cdot \sqrt{\text{Var}(Y)}}$$

is called the *correlation* of X and Y .

Let $L = (L_1, \dots, L_n)$ denote an n -dimensional m -factorial Itô-Prozess of the form

$$dL_i = \mu_i dt + \sigma_i dW_i, \quad \text{where} \quad dW_i = \sum_{k=1}^m f_{i,k} dU_k \quad (18.1)$$

and U_k denote independent Brownian motions. Furthermore, let $f_{i,k}$ such that

$$R := (\rho_{i,j}(t))_{i,j=1,\dots,n} = \left(\sum_{k=1}^m f_{i,k} f_{j,k} \right)_{i,j=1,\dots,n}$$

is a correlation matrix (i.e. $\sum_{k=1}^m f_{i,k}^2 = 1$). We have

$$\langle dW(t), dW(t) \rangle = R dt.$$

Definition 201 (Instantaneous Covariance, Instantaneous Correlation): ↗

With the notation above we call

$$\rho_{i,j}, \quad \text{defined by} \quad \rho_{i,j}(t) := \left(\sum_{k=1}^m f_{i,k} f_{j,k} \right)_{i,j=1,\dots,n},$$

the *instantaneous correlation* of the processes L_i and L_j , and we call $\sigma_i \sigma_j \rho_{i,j}$ the *instantaneous covariance* of the processes L_i and L_j . ↘

Definition 202 (Terminale Kovarianz, Terminale Korrelation): ↗

With the notation above we call

$$\rho_{i,j}^{\text{Term}}, \quad \text{defined by} \quad \rho_{i,j}^{\text{Term}}(t) := \text{Cor}(L_i(t), L_j(t)),$$

the *terminal correlation* of the processes L_i and L_j . Correspondingly we call $t \mapsto \text{Cov}(L_i(t), L_j(t))$ the *terminal covariance* of the processes L_i and L_j . ↘

18.2. Terminal Correlation studied on the Example of a LIBOR Market Model

We consider a [LIBOR Market Model](#) with semi-annual tenor structure $T_i := 0.5 \cdot i$ and investigate the behaviour of the two rates $L_{10} = L(5.0, 5.5)$ and $L_{11} = L(5.5, 6.0)$. Under the Numéraire $N = P(T_{12}) = P(6.0)$ we have for the dynamic of these rates (see (17.3), (17.8))

$$\begin{aligned} dL_i(t) &= \mu_i(t)L_i(t)dt + \sigma_i(t)L_i(t)dW_i^{\mathbb{Q}^N}(t) & (i = 10, 11) & \quad (18.2) \\ \mu_{10} &= -\frac{\delta_{11}L_{11}(t)}{(1 + \delta_{11}L_{11}(t))} \cdot \sigma_{10}(t) \cdot \sigma_{11}(t) \cdot \rho_{10,11}(t), & \delta_{11} &:= T_{11} - T_{10} \\ \mu_{11} &= 0. \end{aligned}$$

If we neglect the drift (i.e. set $\mu_{10} = 0$) and assume a constant instantaneous covariance $\sigma_{10}\sigma_{11}\rho_{10,11} = \text{const.}$, then it follows from (18.1) that the terminal correlation is

$$\rho_{i,j}^{\text{Term}}(t) = \rho_{10,11} \quad \forall t.$$

As one might have expected, the terminal correlation is given by the choice of the instantaneous correlation. In this case, to achieve a terminal correlation different from zero we need at least a two factor model. Figure 18.1 shows a scatter plot for a one and five factor model¹ of the interest rates $L_{10}(t)$, $L_{11}(t)$ at time $t = T_{10} = 5.0$.

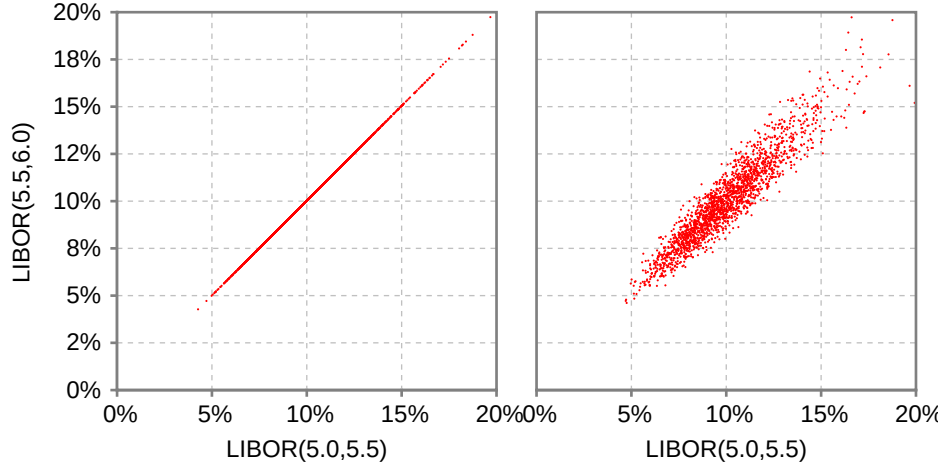


Figure 18.1.: The two (adjacent) rates $L_{10} = L(5.0, 5.5)$ and $L_{11} = L(5.5, 6.0)$ in a one- and multi-factor model for constant instantaneous volatility $\sigma_{10}(t) = \sigma_{11}(t) = \text{const}$. In a one-factor model both random variables are perfectly correlated (left). In a five-factor model both random variables show a correlation different from 1. This is a consequence of the instantaneous correlation $\rho_{10,11}$ being different from 1.

18.2.1. De-correlation in a One-Factor-Model

It is possible to achieve a terminal decorrelation for processes which have perfect instantaneous correlation. Consider

$$\sigma_{10}(t) \begin{cases} > 0 & \text{für } t < 2.5, \\ = 0 & \text{für } t \geq 2.5, \end{cases} \quad \sigma_{11}(t) \begin{cases} = 0 & \text{für } t < 2.5, \\ > 0 & \text{für } t \geq 2.5, \end{cases} \quad (18.3)$$

i.e. the processes receive the Brownian increment $dW(t)$ at different times t , thus the increments received are independent. Since in the case we have $\mu_{10} = \mu_{11} = 0$ in (18.2), the two random variables $L_{10}(5.0)$, $L_{11}(5.0)$ are given by

$$\begin{aligned} \log(L_{10})(5.0) &= -\frac{1}{2}\bar{\sigma}_{10}^2 \cdot 2.5 + \bar{\sigma}_{10}(W_{10}(2.5) - W_{10}(0)) \\ \log(L_{11})(5.0) &= -\frac{1}{2}\bar{\sigma}_{11}^2 \cdot 2.5 + \bar{\sigma}_{11}(W_{11}(5.0) - W_{11}(2.5)), \end{aligned}$$

¹ The exact model specification is: $L_{i,0} = 0.1$, $\sigma_i = 0.1$, and $\rho_{i,j} = \exp(-0.5|i - j|)$, followed by a factor reduction as given in Section A.5. For the five factor model we have $\rho_{10,11} = 0.94$.

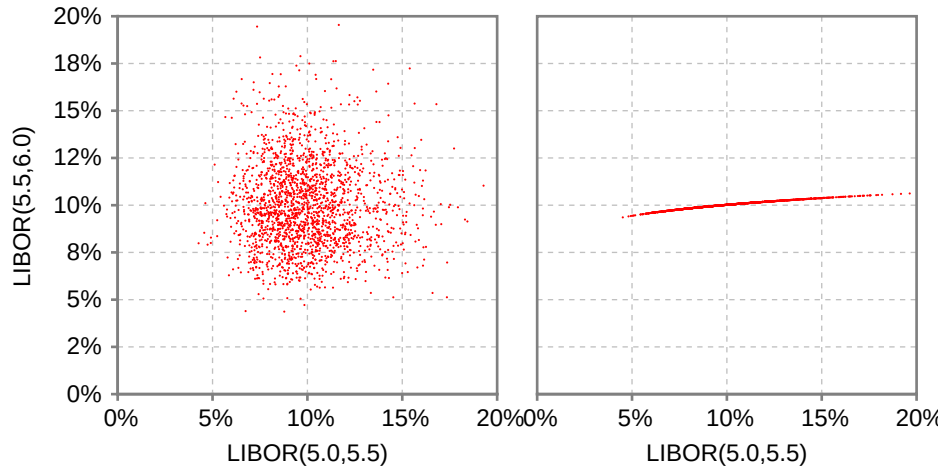


Figure 18.2.: The two (adjacent) rates $L_{10} = L(5.0, 5.5)$ and $L_{11} = L(5.5, 6.0)$ in a one-factor model. Left: The two random variables exhibit a correlation close to 0 (perfect decorrelation). Right: The two random variables exhibit a very different variances. The covariance is close to zero since the variance of L_{11} is close to 0. Both scenarios are the consequence of a very special choice for the instantaneous volatility.

where $\bar{\sigma}_{10}^2 = \frac{1}{2.5} \int_0^{2.5} \sigma_{10}^2(t) dt$ und $\bar{\sigma}_{11}^2 = \frac{1}{2.5} \int_{2.5}^{5.0} \sigma_{11}^2(t) dt$.

Since, even for a one-factor model, the increments $(W(2.5) - W(0))$, $(W(5) - W(2.5))$ are independent, we have that $L_{10}(5.0)$, $L_{11}(5.0)$ are independent two, see Figure 18.2, left.

18.2.2. Impact of the Time Structure of the Instantaneous Volatility on Caplet and Swaption Prices

The previous example of the decorrelation of the rates L_{10} , L_{11} in a one-factor model shows the importance of the time structure of the instantaneous volatility for the (terminal) distribution of (L_{10}, L_{11}) at time $t = 5.0$. We consider now the corresponding caplets and a swaption with maturity 5.0 and payment dates 5.5, 6.0, which depends on L_{10} and L_{11} :

In all scenarios we have $\int_0^{T_i} \sigma_i(t)^2 dt = 0.05$ for $i = 10$, thus all caplet prices are the same.²

The Figures 18.1, 18.2 were generated with these parameters.

² We have $\int_0^{T_i} (b \cdot \exp(-c \cdot (T_i - t)))^2 dt = \frac{b^2}{2c} (1 - \exp(-2 \cdot c \cdot T_i))$. For $T_i = 5.0$, $b = 0.7$, $c = 4.9$ we thus have $\int_0^{T_i} \sigma_i(t)^2 dt = 0.05(1 - \exp(-49)) = 0.05(1 - 5 \cdot 10^{-22})$.

Scenario	$\sigma_i(t)$	$\rho_{10,11}$	Caplet 5.0-5.5	Caplet 5.5-6.0	Swaption 5.0-6.0
1	0.1	1.0	0.26%	0.26%	0.51%
2	0.1	0.94	0.26%	0.26%	0.50%
3	as in (18.3)	1.0	0.26%	0.26%	0.36%
4	$0.7 \exp(4.9(T_i - t))$	1.0	0.26%	0.26%	0.27%

Table 18.1.: Caplet and swaption prices for different instantaneous correlations and volatilities.

18.2.3. The Swaption Value as a Function of Forward Rates

To interpret these results we analyze the dependency of the swaption value from the rates L_{10}, L_{11} .

For the value of a swaption $V_{\text{Swaption}}(T_0)$ with fixed swap rate (strike) K we have

$$V_{\text{Swaption}}(T_0) = N(0)E^{\mathbb{Q}^N} \left(\frac{\max(S(T_i) - K, 0) \cdot A(T_i)}{N(T_i)} \mid \mathcal{F}_{T_0} \right)$$

with

$$A(T_i) = \sum_{j=i}^{n-1} (T_{j+1} - T_j) P(T_{j+1}; T_i) \quad (\text{Swap Annuity})$$

$$S(T_i) = \frac{1 - P(T_n; T_i)}{A(T_i)} \quad (\text{Par Swap Rate}).$$

With the Numéraire $N = P(T_n)$ we have

$$\begin{aligned} \frac{A(T_i)}{N(T_i)} &= \frac{A(T_i)}{P(T_n; T_i)} = \sum_{j=i}^{n-1} (T_{j+1} - T_j) \frac{P(T_{j+1}; T_i)}{P(T_n; T_i)} \\ &= \sum_{j=i}^{n-1} (T_{j+1} - T_j) \prod_{k=j+1}^{n-1} (1 + L_k(T_i)(T_{k+1} - T_k)) \end{aligned}$$

and

$$S(T_i) \frac{A(T_i)}{N(T_i)} = \frac{1 - P(T_n; T_i)}{P(T_n; T_i)} = \prod_{j=i}^{n-1} (1 + L_j(T_i)(T_{j+1} - T_j)) - 1,$$

i.e. it is

$$V_{\text{Swaption}}(T_0) = P(T_n; T_0) E^{\mathbb{Q}^{P(T_n)}} \left(\max \left(\frac{(S(T_i) - K) \cdot A(T_i)}{P(T_n; T_i)}, 0 \right) \mid \mathcal{F}_{T_0} \right)$$

with

$$\frac{(S(T_i) - K) \cdot A(T_i)}{P(T_n; T_i)} = \prod_{j=i}^{n-1} (1 + L_j(T_i)(T_{j+1} - T_j)) - 1 - K \cdot \sum_{j=i}^{n-1} (T_{j+1} - T_j) \prod_{k=j+1}^{n-1} (1 + L_k(T_i)(T_{k+1} - T_k)).$$

If we apply this to the special case of a swaption with a two period tenor $\{T_i, \dots, T_n\} = \{T_{10}, T_{11}, T_{12}\} = \{5.0, 5.5, 6.0\}$ we get

$$\begin{aligned} & \max\left(\frac{S(T_i) - K \cdot A(T_i)}{P(T_n; T_i)}, 0\right) \\ &= \max((1 + L_{10}\Delta T)(1 + L_{11}\Delta T) - K(\Delta T(1 + L_{11}\Delta T) + \Delta T), 0) \\ &= \max((L_{10} - K)\Delta T + (L_{11} - K)\Delta T + L_{11}(L_{10} - K)(\Delta T)^2, 0). \end{aligned} \quad (18.4)$$

From (18.4) we may derive the following observations for the value of the swaption:

- If $L_{11}(T_{10}) = K$, then the value of the swaption corresponds to the value of a caplet paying $\max(L_{10} - K, 0)$. If L_{11} has at time T_{10} no or small variance and if $L_{11}(T_{10})$ is close to K , then the value of the swaption is close to the value of a caplet with payoff $\max(L_{10} - K, 0)$.
- Neglecting the term $L_{11}(T_{10})(L_{10}(T_{10}) - K)(\Delta T)^2$, which is justified for small rates and short periods ΔT , and considering thus only

$$\begin{aligned} (L_{10}(T_{10}) - K)\Delta T + (L_{11}(T_{10}) - K)\Delta T \\ = (L_{10}(T_{10}) + L_{11}(T_{10}) - 2K)\Delta T, \end{aligned}$$

we see that the option price is determined by the variance of $L_{10}(T_{10}) + L_{11}(T_{10})$. For this we have

$$\begin{aligned} & \text{Var}(L_{10}(T_{10}) + L_{11}(T_{10})) \\ &= \text{Var}(L_{10}(T_{10})) + \text{Var}(L_{11}(T_{10})) + 2 \cdot \text{Cov}(L_{10}(T_{10}), L_{11}(T_{10})) \end{aligned}$$

- From the previous we have that the option value is maximal for $\rho_{10,11}^{\text{Term}}(T_{10}) = 1$ and minimal (even 0) for $\rho_{10,11}^{\text{Term}}(T_{10}) = -1$ (still neglecting the term $L_{11}(T_{10})(L_{10}(T_{10}) - K)\Delta T^2$).

From these remarks the results in table 18.1 become plausible. In Scenario 4 the rate $L_{11}(T_{10})$ has a negligible small variance (compare Figure 18.2, right). The swaption value is close to the caplet value. The caplet on the period $[T_{11}, T_{12}]$ however has the same price as in the other scenarios, since the high instantaneous volatility for $t \in [T_{10}, T_{11}]$ will give the rate $L_{11}(T_{11})$ the required (terminal) variance.

While for the swaption the rate $L_{11}(T_{10})$ is relevant, for the caplet it is the rate $L_{11}(T_{10})$.



Experiment: The influence of the instantaneous volatility and instantaneous correlation on terminal correlation, caplet and swaption prices may be investigated at

<http://www.christian-fries.de/finmath/applets/LMMCorrelation.html>.



18.3. Terminal Correlation depends on the Equivalent Martingale Measure

The terminal correlation depends on the martingale measure and thus the Numéraire used. The whole (terminal) probability density is of course measure dependent, see also Lemma 72 in Chapter 5. Thus an interpretation of terminal correlation and other terminal quantities should be made with cautiousness.

How the chosen martingale measure influences the terminal distribution, especially the terminal correlation, may be easily seen in a LIBOR Market Model. Consider the processes $L_i = L(T_i, T_{i+1})$ and $L_{i+1} = L(T_{i+1}, T_{i+2})$, i.e. two adjacent forward rates, under the martingale measure $\mathbb{Q}^{P(T_n)}$ corresponding to the Numéraire $P(T_n)$ (*terminal measure*). It is

$$d \log(L_i) = - \sum_{i < j < n} \frac{\delta_j L_j(t)}{(1 + \delta_j L_j(t))} \cdot \sigma_i(t) \sigma_j(t) \rho_{i,j}(t) dt + \frac{1}{2} \sigma_i(t)^2 dt + \sigma_i(t) dW_i^{\mathbb{Q}^{P(T_n)}}.$$

With $dK_j(t) := \frac{\delta_j L_j(t)}{(1 + \delta_j L_j(t))} dt$ we thus have

$$\begin{aligned} d \log(L_i) &= \sigma_i \cdot \left(-\sigma_{i+1} \rho_{i,i+1} dK_{i+1} - \sum_{i+1 < j < n} \sigma_j \rho_{i+1,j} dK_j + \frac{1}{2} \sigma_i(t) dt + dW_i^{\mathbb{Q}^{P(T_n)}} \right) \\ d \log(L_{i+1}) &= \sigma_{i+1} \cdot \left(-\sum_{i+1 < j < n} \sigma_j \rho_{i+1,j} dK_j + \frac{1}{2} \sigma_{i+1} dt + dW_{i+1}^{\mathbb{Q}^{P(T_n)}} \right). \end{aligned}$$

The terminal correlation is influenced through common drift term $\sum_{i+1 < j < n} dK_j$ and this influence may be increased arbitrarily through the factor σ_j in front of K_j . If and how this term is present depends on the chosen martingale measure: For $n = i + 2$ the sum is empty and the term is $= 0$, for $n > i + 2$ the term is > 0 . In theory it might be possible that L_i and L_{i+1} appear almost perfectly correlated under $\mathbb{Q}^{P(T_{i+3})}$ and perfectly uncorrelated under $\mathbb{Q}^{P(T_{i+2})}$.

18.3.1. Dependence of the Terminal Density on the Martingale Measure

How the chosen martingale measure influences the terminal distribution function is shown in the following examples. We consider in Figure 18.3 the density of a forward rate und a one factor LIBOR Market model with constant instantaneous volatility, equal for all rates. Under different martingale measure (spot measure, terminal measure) the distribution is slightly different. If, however, the volatility of the *other* rates is increased, then, depending on the chosen martingale measure, the distribution will change, see Figure 18.4. As before, the change in the distribution function stems from the drift of the LIBOR Market Model.

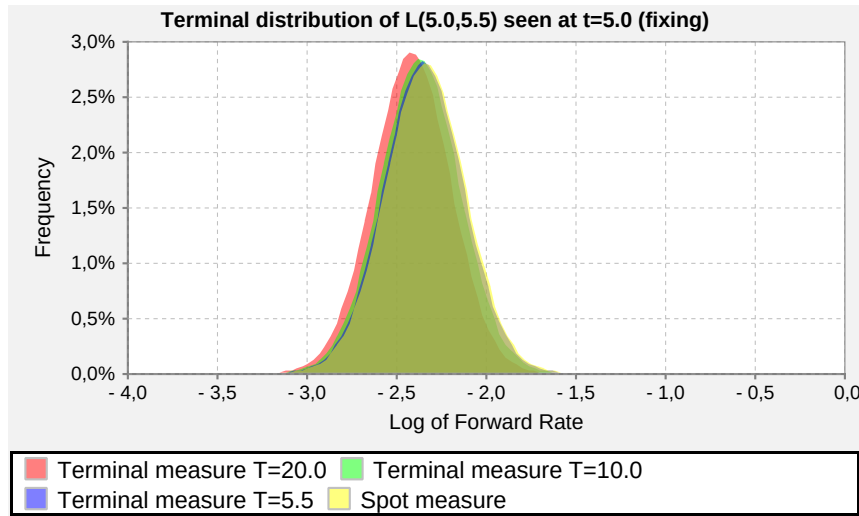


Figure 18.3.: The terminal distribution function of a forward rate under different martingale measures. Shown is the rate $L(5.0, 5.5)$ upon it fixing $t = 5.0$. All other rates are simulated in a one factor LIBOR market model with constant instantaneous volatility $\sigma = 10\%$.



Tip (Martingale Measure independent terminal Quantities):

In place of martingale measure dependent quantities, like the terminal distribution of the terminal correlation, we may define meaningful alternatives. The implied (Black-)volatility is an example of an martingale measure independent quantity. Apart from the scaling with the square root of the maturity $\sqrt{T_k}$ it corresponds to the terminal standard deviation under the T_{k+1} -forward measure. If, for example,

$$\begin{aligned} d \log(L_i(t)) &= (\dots)dt + \sigma_i(t)dW_i(t) \\ d \log(L_j(t)) &= (\dots)dt + \sigma_j(t)dW_j(t) \end{aligned}$$

are given processes, then the *integrated instantaneous covariance*

$$\int_0^T d \log(L_i(t)) d \log(L_j(t)) = \int_0^T \sigma_i(t) \sigma_j(t) \rho_{i,j}(t) dt$$

is independent of the chosen martingale measure.³ It would correspond the the covariance of $\log(L_i(t))$ and $\log(L_j(t))$, of both would be martingales. $\triangleleft$

³ This is clear, because a change of martingale measure is a change of drift only.

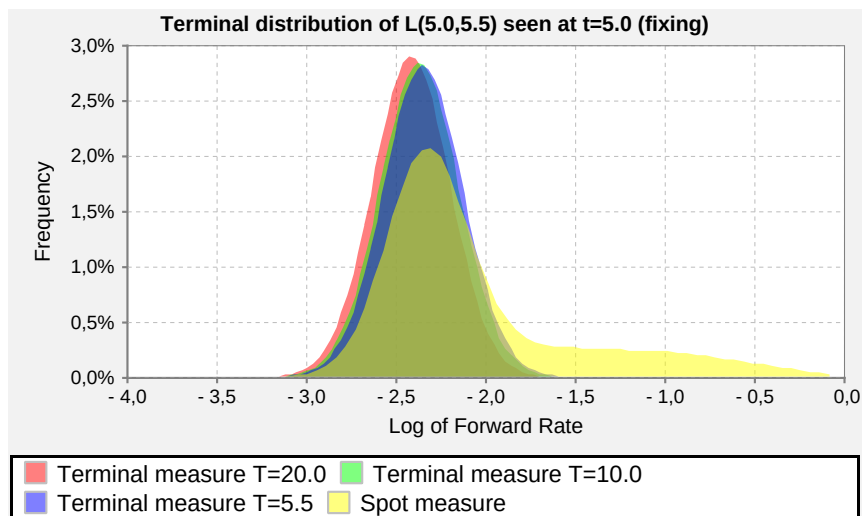


Figure 18.4.: The terminal distribution function of a forward rate under different martingale measures. Shown is the rate $L(5.0, 5.5)$ upon the fixing $t = 5.0$. In contrast to Figure 18.3 the rates $L(T_i, T_{i+1})$ for $T_i < 5.0$ are simulated differently. They are simulated with a high volatility of 150%. All other rates (and the rate shown) is simulated as in Figure 18.3 with volatility $\sigma = 10\%$. The change of the simulation of the other rates has a significant impact on the distribution of $L(5.0, 5.5)$ under the spot measure.

CHAPTER 19

Heath-Jarrow-Morton Framework: Foundations

The Heath-Jarrow-Morton Framework [69] postulates an Itô process as a model for the [instantaneous forward rate](#)¹:

$$\begin{aligned}df(t, T) &= \alpha^{\mathbb{P}}(t, T)dt + \sigma(t, T) \cdot dW^{\mathbb{P}}(t) \\ f(0, T) &= f_0(T)\end{aligned}\tag{19.1}$$

for $0 \leq t < T$, where $W^{\mathbb{P}} = (W_1^{\mathbb{P}}, \dots, W_m^{\mathbb{P}})$ is an m -dimensional $\mathbb{P}$ -Brownian motion with instantaneously uncorrelated components.² Furthermore we assume that $\sigma(t, T) = (\sigma_1(t, T), \dots, \sigma_m(t, T))$ and $\alpha^{\mathbb{P}}(t, T)$ are adapted processes.

In case of its existence, let $\mathbb{Q}$ denote the [risk neutral measure](#), i.e. the martingale measure $\mathbb{Q} = \mathbb{Q}^B$ corresponding to the Numéraire B with

$$B(t) := \exp\left(\int_0^t f(\tau, \tau)d\tau\right) = \exp\left(\int_0^t r(\tau)d\tau\right),\tag{19.2}$$

where r denotes the [short rate](#) - see Definition 92.

By the [Theorem of Girsanov](#) (Theorem 53) the process (19.1) is given under $\mathbb{Q}$ as

$$\begin{aligned}df(t, T) &= \alpha^{\mathbb{Q}}(t, T)dt + \sigma(t, T) \cdot dW^{\mathbb{Q}}(t) \\ f(0, T) &= f_0(T)\end{aligned}\tag{19.3}$$

Equation (19.3) represents a family of stochastic processes parametrized by T , which give a complete description of the interest rate curve: From Definition (90) we have $f(t, T) = -\frac{\partial \log(P(T; t))}{\partial T}$, i.e.

$$P(T; t) = \exp\left(-\int_0^T f(t, \tau)d\tau\right).$$

¹ Definition 90 on Page 130.

² I.e. that $dW^T \cdot dW = Idt$, see Section 2.7.

Apart from the requirement that the processes are Itô processes, we do not consider a specific model or its implementation. A *specific model* would be given, if we would have specified the form of $(t, T) \mapsto \sigma(t, T)$. With a specific choice of $\sigma(t, T)$ (19.3) may become a known short rate model or the LIBOR market model, see Chapter 21.

In this chapter we will discuss the no-arbitrage conditions of (19.3) and discuss how other models fit into this *frame work*.

19.1. Short Rate Process in the HJM Framework

The specification of the families of processes $f(\cdot, T)$ implies a process for the [short rate](#) r . We write Equation (19.3) in integral form³

$$f(t, T) = f_0(T) + \int_0^t \alpha(s, T) ds + \int_0^t \sigma(s, T) \cdot dW(s). \quad (19.4)$$

With $T \rightarrow t$ we find for the short rate $r(t) := \lim_{T \searrow t} f(t, T)$ that

$$r(t) = f(t, t) = f_0(t) + \int_0^t \alpha(s, t) ds + \int_0^t \sigma(s, t) \cdot dW(s) \quad (19.5)$$

and thus the short rate process is in differential notation given as

$$dr(t) = \left(\frac{\partial f_0}{\partial T}(t) + \alpha(t, t) + \int_0^t \frac{\partial \alpha}{\partial T}(s, t) ds + \int_0^t \frac{\partial \sigma}{\partial T}(s, t) \cdot dW(s) \right) dt + \sigma(t, t) \cdot dW(t). \quad (19.6)$$

Remark 203 (Notation): Equation (19.6) follows from (19.5) by differentiating with respect to t . Since t enters into the second argument of α and σ , we have to calculate the partial derivative of α and σ with respect to their second argument. In accordance with the notation in (19.1) we denote the partial derivative of α with respect to its second argument by $\frac{\partial \alpha}{\partial T}$ and the partial derivative of σ with respect to its second argument by $\frac{\partial \sigma}{\partial T}$. Likewise we denote the (partial) derivative of f_0 with respect to its argument by $\frac{\partial f_0}{\partial T}$.

19.2. The HJM Drift Condition

Theorem 204 (Heath, Jarrow, Morton - HJM Drift Condition): For the family of bond price processes $P(T)$ the following holds: The B -relative price $\frac{P(T)}{B}$ is a $\mathbb{Q}^B$ -

³ We drop the superscript $\mathbb{Q}^B$ on the drift α and the diffusion W for a while.

martingale, if and only if

$$\int_s^T \alpha(s, S) dS = \frac{1}{2} \int_s^T \sigma(s, S) dS \cdot \int_s^T \sigma(s, S) dS.$$

From this we have: All bond price processes of the bond curve $T \mapsto P(T)$ are $\mathbb{Q}^B$ -martingales, i.e. the model is arbitrage free, if and only if

$$\alpha^{\mathbb{Q}^B}(t, T) = \sigma(t, T) \cdot \int_t^T \sigma(t, \tau) d\tau \quad \forall T. \quad (19.7)$$

Equation (19.7) is called the HJM drift condition.

Proof (of the HJM drift condition): Let T denote a fixed maturity. With $B(t) = \exp(\int_0^t r(s) ds)$ and $P(T; t) = \exp(-\int_t^T f(t, S) dS)$ we find for the B -relative price of the bonde $P(T)$:

$$\frac{P(T; t)}{B(t)} = \exp(X(t)) \quad \text{with} \quad X(t) = -\int_t^T f(t, S) dS - \int_0^t r(s) ds.$$

From (19.4) and (19.5) follows

$$\begin{aligned} X(t) &= -\int_t^T f(t, S) dS - \int_0^t r(s) ds \\ &= -\int_t^T f_0(S) dS - \int_t^T \int_0^t \alpha(s, S) ds dS - \int_t^T \int_0^t \sigma(s, S) dW(s) dS \\ &\quad - \int_0^t f_0(S) dS - \int_0^t \int_0^u \alpha(s, u) ds du - \int_0^t \int_0^u \sigma(s, u) dW(s) du. \end{aligned}$$

With $\int_0^t \int_0^u dW(s) du = \int_0^t \int_s^t du dW(s)$ and the interchange of the integrals this is

$$\begin{aligned} &= -\int_t^T f_0(S) dS - \int_0^t \int_t^T \alpha(s, S) dS ds - \int_0^t \int_t^T \sigma(s, S) dS dW(s) \\ &\quad - \int_0^t f_0(S) dS - \int_0^t \int_s^t \alpha(s, u) du ds - \int_0^t \int_s^t \sigma(s, u) du dW(s) \\ &= -\int_0^T f_0(S) dS - \int_0^t \int_s^T \alpha(s, S) dS ds - \int_0^t \int_s^T \sigma(s, S) dS dW(s) \\ &= X(0) + \int_0^t A(s) ds + \int_0^t \Sigma(s) dW(s), \end{aligned}$$

thus

$$dX(t) = A(t)dt + \Sigma(t)dW(t),$$

where

$$\begin{aligned} X(0) &= - \int_0^T f_0(S) dS \\ A(s) &= - \int_s^T \alpha(s, S) dS \\ \Sigma(s) &= - \int_s^T \sigma(s, S) dS. \end{aligned}$$

Let the B -relative price of $P(T)$ be a martingal under $\mathbb{Q}^B$, i.e. the process $\exp(X(t))$ is drift free. From Itô's Lemma we have $d \exp(X(t)) = X(t) dX(t) + \frac{1}{2} X(t) dX(t) dX(t)$, i.e.

$$d \exp(X(t)) = X(t) \cdot ((A(s) + \frac{1}{2} \Sigma(s) \Sigma(s)) dt + \Sigma(t) dW(t)).$$

That $\exp(X(t))$ is drift free thus implies $A(s) + \frac{1}{2} \Sigma(s) \Sigma(s) = 0$, i.e.

$$\int_s^T \alpha(s, S) dS = \frac{1}{2} \int_s^T \sigma(s, S) dS \cdot \int_s^T \sigma(s, S) dS.$$

If this equation is valid for all T , we get by differentiation $\frac{\partial}{\partial T}$ the HJM drift condition

$$\alpha(t, T) = \sigma(t, T) \cdot \int_t^T \sigma(t, S) dS.$$

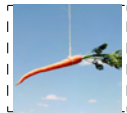
□



Interpretation (Bond volatility): The expression $\Sigma(t) = - \int_t^T \sigma(t, S) dS$ corresponds to the volatility of the bond price process $P(T)$ at time t (bond volatility), since we have

$$\begin{aligned} dP(T; t) &= d(B(t) \exp(X(t))) = B(t) d \exp(X(t)) \\ &= B(t) \exp(X(t)) \cdot (dX(t) + \frac{1}{2} dX(t) \cdot dX(t)) \\ &= P(T; t) \cdot ((\dots) dt + \Sigma(t) \cdot dW(t)) \\ &= P(T; t) \cdot (\dots) dt + P(T; t) \cdot \Sigma(t) \cdot dW(t). \end{aligned}$$

<



Motivation (Embedding other Models): If an interest rate model is arbitrage free and if the processes of the instantaneous forward rates $f(\cdot, T)$ are Itô processes, then the model has to fulfill the HJM drift condition (19.7). Thus, these interest rate models may be derived as a special case of the HJM framework. Since the volatility structure $(t, T) \mapsto \sigma(t, T)$ and

the initial conditions $f(0, T)$ are the only free parameter of the HJM framework, this embedding of arbitrage free interest rate models may be achieved through a choice of the HJM volatility structure and the initial interest rate curve. We will show in Chapter 21 how [models of the short rate](#) and the [LIBOR market model](#) may be interpreted as special HJM volatility structures. $\triangleleft$

CHAPTER 20

Short Rate Models

20.1. Introduction

At a fixed point t in time the *short rate* is given by

$$r(t) := -\frac{\partial P(T; t)}{\partial T} \Big|_{T=t},$$

see Definition 92. Thus $r : t \mapsto r(t)$ is a real valued stochastic process. We make the following assumptions:

1. Given is a model for r (*short rate model*), e.g. in the form of an Itô process

$$dr = \mu^{\mathbb{P}}(t, r)dt + \sigma(t, r)dW^{\mathbb{P}}(t), \quad r(0) = r_0, \quad (20.1)$$

where $\mathbb{P}$ denote the real probability measure.

2. The *continuously compounding money market account* $B(t)$,

$$dB(t) = r(t)B(t)dt, \quad B(0) = 1, \quad \text{d.h.} \quad B(t) = \exp\left(\int_0^t r(\tau)d\tau\right),$$

is a traded asset.¹

3. Corresponding to the Numéraire $N(t) = B(t)$ there exists a martingale measure $\mathbb{Q} = \mathbb{Q}^B$ equivalent to $\mathbb{P}$.

From the [Theorem of Girsanov](#)² the process of r under $\mathbb{Q}$ is

$$dr = \mu^{\mathbb{Q}}(t, r)dt + \sigma(t, r)dW^{\mathbb{Q}}(t), \quad r(0) = r_0. \quad (20.2)$$

¹ The short rate r is, as an interest rate for an infinitesimal period dt , an idealized quantity. Correspondingly the product B is an idealized quantity: The continuous re-investment of an initial value of 1 over infinitesimal periods $[t, t + dt]$ with rate $r(t)$.

² Theorem 53 on page 53.

with $\mu^{\mathbb{Q}}(t, r) = \mu^{\mathbb{P}}(t, r) + C(t)$. Since under $\mathbb{Q}$ all B -relative prices of traded assets are martingales, we have that all bond prices are given by

$$P(T; t) = B(t)E^{\mathbb{Q}}\left(\frac{1}{B(T)} \mid \mathcal{F}_t\right) = E^{\mathbb{Q}}\left(\exp\left(-\int_t^T r(\tau)d\tau\right) \mid \mathcal{F}_t\right).$$

From the bond prices $P(T; t)$ we may calculate all derived quantities like forward rates or swap rates, see Section 8.2. Thus, the *short rate model* (20.2) gives a complete description of the interest rate curve dynamic.

Short rate models were and are popular, since the underlying stochastic process r is one-dimensional (i.e. scalar valued). Thus many techniques that are known from the modeling of (also one dimensional) stock price processes may be used out of the box (e.g. finite difference implementations). Depending on the specific model (i.e. the form of $\mu^{\mathbb{Q}}$ and σ) analytic formulas for bond prices or simple European interest rate options may be derived, similar to the Black-Scholes formula for european stock options under a Black-Scholes model.

Instead of specifying the model (20.1) of the short process under the real measure $\mathbb{P}$ and applying the measure transformation to $\mathbb{Q}$, it is usual to specify the model (20.2) directly under $\mathbb{Q}$ and calibrate given model parameters.

20.2. The Market Price of Risk

We consider the bond with maturity T . Under a short rate model its price process $P(T) : t \mapsto P(T; t)$ is a function of $(t, r(t))$ and if Itô's Lemma is applicable we have³

$$dP(T) = \alpha_T^{\mathbb{P}}(t, r)P(T)dt + \sigma_T(t, r)P(T)dW^{\mathbb{P}}(t), \quad (20.3)$$

where the price process is considered under the real measure $\mathbb{P}$.

Let $P(T_1)$ and $P(T_2)$ denote two bonds with different maturities $T_1 \neq T_2$. We construct a portfolio process (ϕ_0, ϕ_1) for a self financing portfolio of B and $P(T_1)$, which replicates $P(T_2)$. The portfolio process (ϕ_0, ϕ_1) has to satisfy the following equations:

$$\phi_0 B + \phi_1 P(T_1) = P(T_2) \quad (\text{"replicating"}) \quad (20.4)$$

$$d(\phi_0 B + \phi_1 P(T_1)) = \phi_0 dB + \phi_1 dP(T_1) \quad (\text{"self financing"}). \quad (20.5)$$

From (20.4) we find $dP(T_2) = d(\phi_0 B + \phi_1 P(T_1))$ and with

$$\begin{aligned} d(\phi_0 B + \phi_1 P(T_1)) &\stackrel{(20.5)}{=} (\phi_0 r B + \phi_1 \alpha_{T_1}^{\mathbb{P}} P(T_1))dt + \phi_1 \sigma_{T_1} P(T_1) dW^{\mathbb{P}}(t) \\ dP(T_2) &\stackrel{(20.3)}{=} \alpha_{T_2}^{\mathbb{P}} P(T_2)dt + \sigma_{T_2} P(T_2) dW^{\mathbb{P}}(t). \end{aligned}$$

³ At this point, it is not obvious that Itô's Lemma is applicable, especially if the functional dependence of $P(T; t)$ from $r(t)$ is sufficiently smooth. However, for the short rate models presented this is the case.

From Itô's Lemma we then have $\alpha_T = \frac{\frac{\partial}{\partial t} P(T) + \mu \frac{\partial}{\partial r} P(T) + \frac{1}{2} \sigma^2 \frac{\partial^2}{\partial r^2} P(T)}{P(T)}$ and $\sigma_T = \frac{\sigma \frac{\partial}{\partial r} P(T)}{P(T)} = \sigma \frac{\partial}{\partial r} \log(P(T))$.

we have by comparing coefficients

$$\alpha_{T_2}^{\mathbb{P}} P(T_2) = \phi_0 r B + \phi_1 \alpha_{T_1}^{\mathbb{P}} P(T_1) \quad (20.6)$$

$$\sigma_{T_2} P(T_2) = \phi_1 \sigma_{T_1} P(T_1). \quad (20.7)$$

While (20.7) and (20.4) uniquely determine the portfolio process (ϕ_0, ϕ_1) , (20.6) is a consistency condition for r , α_{T_1} and α_{T_2} . If (20.6) would be violated, then the model would not be arbitrage free. We rewrite the consistency condition (20.6) as:

$$\begin{aligned} \Leftrightarrow \quad & \alpha_{T_2}^{\mathbb{P}} \cdot P(T_2) = \phi_0 \cdot r \cdot B + \phi_1 \cdot \alpha_{T_1}^{\mathbb{P}} \cdot P(T_1) \\ \Leftrightarrow \quad & \alpha_{T_2}^{\mathbb{P}} \cdot P(T_2) = \phi_0 \cdot r \cdot B + \phi_1 \cdot r \cdot P(T_1) + \phi_1 \cdot (\alpha_{T_1}^{\mathbb{P}} - r) \cdot P(T_1) \\ (20.4) \quad \Leftrightarrow \quad & \alpha_{T_2}^{\mathbb{P}} \cdot P(T_2) = r \cdot P(T_2) + \phi_1 \cdot (\alpha_{T_1}^{\mathbb{P}} - r) \cdot P(T_1) \\ \Leftrightarrow \quad & (\alpha_{T_2}^{\mathbb{P}} - r) \cdot P(T_2) = \phi_1 \cdot (\alpha_{T_1}^{\mathbb{P}} - r) \cdot P(T_1) \\ (20.7) \quad \Leftrightarrow \quad & \frac{\alpha_{T_2}^{\mathbb{P}} - r}{\sigma_{T_2}} = \frac{\alpha_{T_1}^{\mathbb{P}} - r}{\sigma_{T_1}}. \end{aligned}$$

It follows that there exists a $\lambda^{\mathbb{P}}$, such that for *all* bond price processes

$$dP(T) = \alpha_T^{\mathbb{P}}(t, r) \cdot P(T) \cdot dt + \sigma_T(t, r) \cdot P(T) \cdot dW^{\mathbb{P}}(t)$$

we have

$$\frac{\alpha_T^{\mathbb{P}} - r}{\sigma_T} =: \lambda^{\mathbb{P}}.$$

Since $\alpha^{\mathbb{P}}$ is the local rate of return of the bond, we may interpret $\lambda^{\mathbb{P}}$ as the local excess return rate over r per risk unit σ_T .

Definition 205 (Market Price of Risk): ▮

The quantity $\lambda^{\mathbb{P}} := \frac{\alpha_T^{\mathbb{P}} - r}{\sigma_T}$, which is independent of T , is called the *Market Price of Risk*. ▮

$$dP(T) = \alpha_T^{\mathbb{Q}}(t)P(T)dt + \sigma_T(t)P(T)dW^{\mathbb{Q}}(t)$$

under the measure $\mathbb{Q}$, it is obvious that $\alpha_T^{\mathbb{Q}} = r$ for all T , since all B -relative prices are $\mathbb{Q}$ -martingales. Thus, under $\mathbb{Q}$ the *Price of Risk* $\lambda^{\mathbb{Q}} = 0$. It follows

$$\mu_T^{\mathbb{Q}} = \lambda^{\mathbb{Q}} \cdot \sigma_T + r = 0 + r = \lambda^{\mathbb{P}} \cdot \sigma_T + r - \lambda^{\mathbb{P}} \cdot \sigma_T = \mu_T^{\mathbb{P}} - \lambda^{\mathbb{P}} \cdot \sigma_T,$$

and we find that *Market Price of Risk* $\lambda^{\mathbb{P}}$ appears in the change of drift to the measure $\mathbb{Q}$, i.e. we have $C(t) = -\lambda^{\mathbb{P}} \cdot \sigma_T$ in Theorem 53.

Definition 206 (Risk Neutral Measure): ▮

Let $r(t)$ denote the short rate. The martingale measure $\mathbb{Q}^B$ corresponding to the Numéraire $B(t) = \exp(\int_0^t r(\tau)d\tau)$ is called the *risk neutral measure* bezeichnet. ▮

Remark 207 (Risk Neutral Measure): The *continuously compounding money market account* B is locally risk free, since the process $dB(t) = r(t)B(t)dt$ does not exhibit a $dW(t)$ term. However $r(t)$ may be stochastic. If r would not be stochastic, then B would be globally risk free.

20.3. Overview: Some Common Models

Table 20.3 gives a selection of the most common short rate models.

Name	Model
Vasicek Model	$dr = (b - ar)dt + \sigma dW^{\mathbb{Q}}$
Hull-White Model	$dr = (\phi(t) - ar)dt + \sigma(t)dW^{\mathbb{Q}}$
Ho-Lee Model	$dr = a(t)dt + \sigma(t)dW^{\mathbb{Q}}$
Dothan Model	$dr = ardt + \sigma r dW^{\mathbb{Q}}$
Black-Derman-Toy Model	$d \ln(r) = \phi(t)dt + \sigma(t)dW^{\mathbb{Q}}$
Black-Karasinski Model	$d \ln(r) = (\phi(t) - a \ln(r))dt + \sigma(t)dW^{\mathbb{Q}}$
Cox-Ingersoll-Ross Model	$dr = (b - ar)dt + \sigma(t) \sqrt{r} dW^{\mathbb{Q}}$

Table 20.1.: Selection of Short Rate Models

The Hull-White model is sometimes called *extended Vasicek model*. The Vasicek, Hull-White and Ho-Lee model allow for negative short rates. The Black-Derman-Toy (BDT) and Black-Karasinski (BK) model use a log-normal process and the Cox-Ingersoll-Ross model uses a square-root process. Both processes do not allow for negative rates.⁴

20.4. Implementations

20.4.1. Monte-Carlo Implementation of Short-Rate Models

A short-rate model gives a description of the dynamic of the short rate. To obtain a complete interest rate curve at a given simulation time t , we have to calculate the bond prices from (20.1) as *conditional expectation*. To calculate a conditional expectation in a Monte-Carlo simulation numerically requires additional, numerically expensive methods, see Chapter 14. To obtain a Monte-Carlo simulation of the full interest rate curve from a Monte-Carlo simulation of the short rate, analytic formulas for bond-prices are indispensably. The popularity of some short rate models is thus partly due to the need for a simple and efficient implementation.

The possibility to calculate bond prices analytically in a specific short rate model is also required for a fast calibration to a give interest rate curve.

⁴ This result holds for the time-continuous process. A time-discretization of the process may allow for negative rates. See for example Section 13.1.2.

20.4.2. Lattice Implementation of Short-Rate Models

If the short rate model is Markovsch in low dimensions, then it is best to implement the short rate model on a lattice, allowing for a backward algorithm.⁵ Depending on the model, implementations using binomial or trinomial trees or general finite differences for pde's are used. See [33] for a detailed discussion.



Further Reading: The book of Björg [7] contains a discussion of short rate model with affine term structure. The book of Tavella and Randal [33] gives an introduction to finite difference methods, also with applications to interest rate models. <|

⁵ See Section 13.3.2.

CHAPTER 21

Heath-Jarrow-Morton Framework: Immersion of Short Rate Models and LIBOR Market Model

You're going to find that many of the truths we cling to depend greatly on our own point of view.

*Obi-Wan Kenobi / George Lucas
Star Wars: Episode VI, (Wikiquote).*

21.1. Short Rate Models in the HJM Framework

The Heath-Jarrow-Morton framework

$$\begin{aligned}df(t, T) &= \alpha(t, T)dt + \sigma(t, T)dW(t) \\ f(0, T) &= f_0(T)\end{aligned}\quad (\text{see } 19.3)$$

implies the short rate process

$$\begin{aligned}dr(t) &= \left(\frac{\partial f_0}{\partial T}(t) + \alpha(t, t) + \int_0^t \frac{\partial \alpha}{\partial T}(s, t)ds + \int_0^t \frac{\partial \sigma}{\partial T}(s, t)dW(s) \right) dt \\ &\quad + \sigma(t, t)dW(t),\end{aligned}\quad (\text{see } 19.6)$$

both under the measure $\mathbb{Q}^B$ - see (19.3), (19.4), (19.5), (19.6) on Page 296. The short rate model is thus given by the specific choice of the HJM volatility structure $\sigma(t, t)$ ($\rightarrow$ short rate volatility) and initial conditions f_0 ($\rightarrow$ short rate drift).

21.1.1. Example: The Ho-Lee Model in the HJM Framework

We consider the simple case of a constant volatility function

$$\sigma(t, T) = \sigma = \text{const.}$$

From (19.7) we have $\alpha(t, T) = \sigma \cdot \int_t^T \sigma \, d\tau = \sigma^2 \cdot (T - t)$, i.e.

$$df(t, T) = \sigma^2 \cdot (T - t)dt + \sigma dW(t), \quad f(0, T) = f_0(T).$$

For the short rate it follows

$$\begin{aligned} r(t) = f(t, t) &= f_0(t) + \int_0^t \sigma^2 \cdot (T - s)ds \Big|_{T=t} + \int_0^t \sigma dW(s) \\ &= f_0(t) + \frac{1}{2}\sigma^2 t^2 + \sigma W(t), \end{aligned}$$

i.e.

$$dr(t) = \left(\frac{df_0}{dT}(t) + \sigma^2 t \right) dt + \sigma dW(t).$$

Using the notation from the Ho-Lee model, $dr(t) = \phi(t)dt + \sigma dW(t)$, it is

$$\phi(t) = \frac{df_0}{dT}(t) + \sigma^2 t.$$



Interpretation: The Equation (21.1.1) allows a calibration of the Ho-Lee model to a given curve of bond prices $P(T)$ by setting $\phi(t) = -\frac{d^2 \log(P(T))}{dT^2}(t) + \sigma^2 t$. With this choice the model reproduces the given bond prices.

If we consider the interest rate curve $f_{T_1}(T) := f(T_1, T_1 + T)$, $T \geq 0$ at a later time $T_1 > 0$, then we have from

$$\frac{df_{T_1}}{dT}(t) + \sigma^2 t = \phi(T_1 + t) = \frac{df_0}{dT}(T_1 + t) + \sigma^2(T_1 + t),$$

that $f_{T_1}(t) = f_{T_1}(0) + f_0(T_1 + t) - f_0(t) + \sigma^2(T_1 \cdot t + \frac{1}{2}t^2)$.

So to summarize, the model reproduces all bond prices, but in the evolution the interest rate curve get steeper and steeper - a rather unrealistic behavior. $\triangleleft$

21.1.2. Example: The Hull-White Model in the HJM Framework

We consider the case of an exponentially volatility function

$$\sigma(t, T) = \sigma \cdot e^{-a \cdot (T-t)}, \quad (a > 0).$$

Then we have $\frac{\partial \sigma}{\partial T}(t, T) = -a \cdot \sigma \cdot e^{-a \cdot (T-t)} = -a\sigma(t, T)$. For the drift $\mu(t)$ of the short rate process $dr(t) = \mu(t)dt + \sigma(t, t)dW(t)$ we get

$$\begin{aligned} \mu(t) &\stackrel{(19.6)}{=} \frac{\partial f_0}{\partial T}(t) + \alpha(t, t) + \int_0^t \frac{\partial \alpha}{\partial T}(s, t)ds + \int_0^t \frac{\partial \sigma}{\partial T}(s, t)dW(s) \\ &= \frac{\partial f_0}{\partial T}(t) + \alpha(t, t) + \int_0^t \frac{\partial \alpha}{\partial T}(s, t)ds - \int_0^t a \cdot \sigma(s, t)dW(s) \\ &\stackrel{(19.5)}{=} \frac{\partial f_0}{\partial T}(t) + \alpha(t, t) + \int_0^t \frac{\partial \alpha}{\partial T}(s, t)ds - a \cdot r(t) + a \cdot f_0(t) + \int_0^t a\alpha(s, t)ds, \end{aligned}$$

i.e.

$$dr(t) = (\phi(t) - ar(t)) dt + \sigma dW(t)$$

with

$$\phi(t) = \frac{\partial f_0}{\partial T}(t) + a \cdot f_0(t) + \alpha(t, t) + \int_0^t \frac{\partial \alpha}{\partial T}(s, t) ds + \int_0^t a \cdot \alpha(s, t) ds.$$

With the HJM drift condition (19.7) it follows that $\alpha(t, T) = \sigma^2 \cdot e^{-a \cdot (T-t)} \cdot \frac{1}{a} (1 - e^{-a \cdot (T-t)}) = \sigma^2 \cdot \frac{1}{a} (e^{-a \cdot (T-t)} - e^{-2a \cdot (T-t)})$ and thus

$$\begin{aligned} \phi(t) &= \frac{\partial f_0}{\partial T}(t) + a \cdot f_0(t) + \alpha(t, t) + \int_0^t \frac{\partial \alpha}{\partial T}(s, t) ds + \int_0^t a \cdot \alpha(s, t) ds \\ &= \frac{\partial f_0}{\partial T}(t) + a \cdot f_0(t) + \int_0^t \sigma^2 e^{-2a \cdot (t-s)} ds \\ &= \frac{\partial f_0}{\partial T}(t) + a \cdot f_0(t) + \frac{\sigma^2}{2a} (1 - e^{-2a \cdot t}). \end{aligned}$$

Altogether we have

$$dr(t) = \left(\frac{\partial f_0}{\partial T}(t) + a \cdot f_0(t) + \frac{\sigma^2}{2a} (1 - e^{-2a \cdot t}) - ar(t) \right) dt + \sigma dW(t).$$

Note that this equation allows a calibration of the Hull-White model to a given curve of bond prices. From the bond price curve we may calculate $\frac{\partial f_0}{\partial T}(t) + a \cdot f_0(t)$.



Interpretation (Mean Reversion): The derivation of a Hull-White model from a Heath-Jarrow-Morton model gives an important insight to the relevance of the timer structure of the volatility function:

A volatility function of the instantaneous forward rate $f(t, T)$, which is exponentially decaying in $(T - t)$ (*time to maturity*), i.e. $\sigma(t, T) = \exp(-a(T - t))$, corresponds to a *mean reversion* term for the short rate process $r(t)$ with mean reversion speed a .

Correspondingly, this effect is visible in the LIBOR Markt Modell, see Chapter 22.

<|

21.2. LIBOR Market Model in the HJM Framework

21.2.1. HJM Volatility Structure of the LIBOR Market Model

In the specification (17.1) of the **LIBOR Market Model** dW denoted the increment of a n -dimensional Brownian motion with instantaneous correlation R . In the specification (19.3) of the **HJM Framework** dW denoted the increment of a m -dimensional Brownian motion with instantaneous uncorrelated components. To resolve this conflict we use the notation of Section 2.7: Let U denote a m -dimensional Brownian motion with instantaneous uncorrelated components and W denote a n -dimensional Brownian motion with $dW(t) = F(t) \cdot dU(t)$, i.e. the instantaneous correlation of W is $R := FF^T$. We consider the HJM model

$$\begin{aligned} df(t, T) &= \alpha^Q(t, T)dt + \sigma(t, T)dU^Q(t) \\ f(0, T) &= f_0(T) \end{aligned} \quad (21.1)$$

with $dU = (dU_1, \dots, dU_m)$. From

$$P(T; t) = \exp\left(-\int_t^T f(t, \tau)d\tau\right).$$

(see Remark 91) it follows that the forward Rate $L_i(t) := L(T_i, T_{i+1}; t)$ is given by

$$1 + L_i(t) \cdot \Delta T_i = \frac{P(T_i; t)}{P(T_{i+1}; t)} = \exp\left(\int_{T_i}^{T_{i+1}} f(t, \tau)d\tau\right).$$

Note that for $X(t) := \int_{T_i}^{T_{i+1}} f(t, \tau)d\tau$ we have by the linearity of the integral that $dX = \int_{T_i}^{T_{i+1}} df(t, \tau)d\tau$, thus we find from **Itô's Lemma** that within the HJM framework the process of the forward rate $L_i(t)$ is

$$\begin{aligned} dL_i(t) \cdot \Delta T_i &= d \exp(X) = \exp(X) \cdot (dX + \frac{1}{2}dX \cdot dX) \\ &= \exp\left(\int_{T_i}^{T_{i+1}} f(t, \tau)d\tau\right) \cdot \left[\int_{T_i}^{T_{i+1}} (df(t, \tau))d\tau \right. \\ &\quad \left. + \frac{1}{2} \int_{T_i}^{T_{i+1}} (df(t, \tau))d\tau \cdot \int_{T_i}^{T_{i+1}} (df(t, \tau))d\tau\right] \\ &= (1 + L_i(t) \cdot \Delta T_i) \cdot \left[(A(t) + \frac{1}{2}\Sigma(t) \cdot \Sigma(t))dt + \Sigma \cdot dU^Q\right] \end{aligned}$$

where $A(t) = \int_{T_i}^{T_{i+1}} \alpha^Q(t, \tau)d\tau$ und $\Sigma(t) = \int_{T_i}^{T_{i+1}} \sigma(t, \tau)d\tau$.

$$dL_i(t) = \frac{1 + L_i(t) \cdot \Delta T_i}{\Delta T_i} \cdot \left((A(t) + \frac{1}{2}\Sigma(t) \cdot \Sigma(t))dt + \int_{T_i}^{T_{i+1}} \sigma(t, \tau)d\tau dU^Q(t)\right) \quad (21.2)$$

We will now choose the volatility structure such that (21.2) corresponds to the process of a LIBOR market model: Let $W = (W_1, \dots, W_n)$ denote an n -dimensional Brownian motion as given in Section 2.7.

$$dW(t) = F(t) \cdot dU(t), \quad \text{with correlation matrix } R := FF^T,$$

i.e.

$$dW_i(t) = F_i(t) \cdot dU(t), \quad \text{mit } F = \begin{pmatrix} F_1 \\ \vdots \\ F_n \end{pmatrix}.$$

Let the volatility structure be chosen as

$$\sigma(t, \tau) = \begin{cases} \frac{L_i(t)}{1+L_i(t)\Delta T_i} \sigma_i(t) \cdot F_i(t) & \text{für } t \leq T_i \\ (0, \dots, 0) & \text{für } t > T_i, \end{cases} \quad (21.3)$$

where i is such that $\tau \in [T_i, T_{i+1})$. Then we have

$$\frac{1 + L_i(t) \cdot \Delta T_i}{\Delta T_i} \cdot \int_{T_i}^{T_{i+1}} \sigma(t, \tau) d\tau dU = \begin{cases} L_i(t) \sigma_i(t) dW(t) & \text{für } t \leq T_i, \\ 0 & \text{für } t > T_i, \end{cases}$$

The forward rate then follows the process

$$dL_i = \mu_i^{\mathbb{Q}^B}(t) L_i(t) dt + \sigma_i(t) L_i(t) dW_i(t),$$

where

$$\mu_i^{\mathbb{Q}^B} = \frac{1 + L_i(t) \cdot \Delta T_i}{L_i(t) \Delta T_i} \cdot \left(\int_{T_i}^{T_{i+1}} \alpha^{\mathbb{Q}}(t, \tau) d\tau + \frac{1}{2} \int_{T_i}^{T_{i+1}} \sigma(t, \tau) d\tau \cdot \int_{T_i}^{T_{i+1}} \sigma(t, \tau) d\tau \right).$$



Interpretation (LIBOR Market Model as HJM Framework with discrete Tenor Structure): Apart from the factor $\frac{L_i(t)}{1+L_i(t)\Delta T_i}$, (21.3) gives the volatility structure $\sigma(t, T)$ of $f(t, T)$ as piecewise constant in T .

The factor $L_i(t)$ results from the requirement to have a log-normal process for L_i . The factor $\frac{1}{1+L_i(t)\Delta T_i}$ results from the discretization of the tenor structure. This illustrates that the LIBOR Market Model may be interpreted as *HJM framework with discrete tenor structure*. In the limit $\Delta T_i \rightarrow 0$ the factor $\frac{1}{1+L_i(t)\Delta T_i}$ vanishes and we obtain (apart from the restriction to a log-normal model) the HJM framework. $\triangleleft$

21.2.2. LIBOR Market Model Drift under the $\mathbb{Q}^B$ Measure

The HJM drift condition states that

$$\alpha^{\mathbb{Q}^B}(t, T) = \sigma(t, T) \cdot \int_t^T \sigma(t, \tau) d\tau.$$

Since for fixed t $\sigma(t, T)$ is a piecewise constant function in T - namely constant on $[T_i, T_{i+1})$ -, we have for $T \in [T_i, T_{i+1})$

$$\alpha^{\mathbb{Q}^B}(t, T) = \sigma(t, T_i) \cdot \underbrace{(\sigma(t, T)(T_{m(t)+1} - t))}_{=0} + \sum_{j=m(t)+1}^{i-1} \sigma(t, T_j) \Delta T_j + \sigma(t, T_i)(T - T_i)$$

where $m(t) := \max\{i : T_i \leq t\}$. Thus we have

$$\int_{T_i}^{T_{i+1}} \alpha^{\mathbb{Q}^B}(t, \tau) d\tau = \sigma(t, T_i) \Delta T_i \cdot \left(\sum_{j=m(t)+1}^{i-1} \sigma(t, T_j) \Delta T_j + \frac{1}{2} \sigma(t, T_i) \Delta T_i \right).$$

With $\sigma(t, T_i) \cdot \sigma(t, T_j) = \frac{\sigma_i L_i}{1 + L_i \Delta T_i} \frac{\sigma_j L_j}{1 + L_j \Delta T_j} \rho_{i,j}$ we find

$$\begin{aligned} \mu_i^{\mathbb{Q}^B} &= \frac{1 + L_i(t) \cdot \Delta T_i}{L_i(t) \Delta T_i} \cdot \left(\int_{T_i}^{T_{i+1}} \alpha^{\mathbb{Q}^B}(t, \tau) d\tau + \frac{1}{2} \int_{T_i}^{T_{i+1}} \sigma(t, \tau) d\tau \cdot \int_{T_i}^{T_{i+1}} \sigma(t, \tau) d\tau \right) \\ &= \frac{1 + L_i(t) \cdot \Delta T_i}{L_i(t) \Delta T_i} \cdot \frac{\sigma_i L_i \Delta T_i}{1 + L_i \Delta T_i} \cdot \left(\sum_{j=m(t)+1}^{i-1} \frac{\sigma_j L_j \Delta T_j}{1 + L_j \Delta T_j} \rho_{i,j} + \frac{\sigma_i L_i \Delta T_i}{1 + L_i \Delta T_i} \right) \\ &= \frac{1 + L_i(t) \cdot \Delta T_i}{L_i(t) \Delta T_i} \cdot \frac{\sigma_i L_i \Delta T_i}{1 + L_i \Delta T_i} \cdot \left(\sum_{j=m(t)+1}^i \frac{\sigma_j L_j \Delta T_j}{1 + L_j \Delta T_j} \rho_{i,j} \right) \\ &= \sigma_i \cdot \left(\sum_{j=m(t)+1}^i \frac{\sigma_j L_j \Delta T_j}{1 + L_j \Delta T_j} \rho_{i,j} \right). \end{aligned}$$



Interpretation: Surprisingly, we find that the drift under $\mathbb{Q}^B$ is identical to the drift under the **spot LIBOR measure** (see Section 17.1.1.2))

$$N(t) = P(T_{m(t)+1}; t) \prod_{j=0}^{m(t)} (1 + L_j(T_j) \cdot \Delta T_j),$$

(Spot LIBOR Measure, s. 17.9)

The reason is simple: Under the assumed volatility structure the Numéraires $B(t)$ and $N(t)$ are identical. To be precise, it is the assumption

$$\sigma(t, T) = 0 \quad \text{für } T_{m(t)} \leq t \leq T < T_{m(t)+1} \quad (21.4)$$

which implies that the two Numéraires coincide. By this the HJM drift implies

$$\alpha^{\mathbb{Q}^B}(t, T) = 0 \quad \text{for } T_{m(t)} \leq t \leq T < T_{m(t)+1}$$

and thus for $T_{m(t)} \leq T < T_{m(t)+1}$:

$$f(t, T) = f(T_{m(t)}, T) + \underbrace{\int_{T_i}^t \alpha^{\mathbb{Q}^B}(\tau, T) d\tau}_{=0} + \underbrace{\int_{T_i}^t \sigma(\tau, T) dW^{\mathbb{Q}^B}(\tau)}_{=0}.$$

From $f(t, T) = f(T_{m(t)}, T)$ we have

$$\begin{aligned} \frac{B(t)}{B(T_{m(t)})} &= \exp\left(\int_{T_{m(t)}}^t f(\tau, \tau) d\tau\right) = \exp\left(\int_{T_{m(t)}}^t f(T_i, \tau) d\tau\right) \\ &= \frac{P(T_{m(t)+1}; t)}{P(T_{m(t)+1}; T_{m(t)})} = \frac{N(t)}{N(T_{m(t)})}, \end{aligned}$$

with $B(0) = N(0) = 1$ i.e. $B(t) = N(t)$. ◁|

We summarize this result in a theorem:

Theorem 208 (Equivalence of the risk neutral measure and the spot LIBOR measure): Given a tenor structure $0 = T_0 < T_1 < \dots < T_n$. Under the assumption that for all $i = 0, 1, 2, \dots$ we have that the T_{i+1} -bond $P(T_{i+1}; t)$ has volatility 0 on $t \in [T_i, T_{i+1}]$, we have

$$B(t) = N(t)$$

for $B(t)$ as in (19.2) and $N(t)$ as in (17.9).

Proof: The claim follows from the considerations above, since the assumption in the theorem is equivalent to (21.4). ◻|

21.2.3. LIBOR Market Model as a Short Rate Model

In 21.2.1 we have given the volatility structure for $(t, T) \mapsto f(t, T)$ under which the forward rates L_i evolve as in a LIBOR market model. Since the short rate is given as $r(t) := \lim_{T \searrow t} f(t, T)$ we have that the volatility structure also implies a short rate model. Furthermore, the Numéraire $B(t) = \exp(\int_0^t r(\tau) d\tau)$ is fully determined by the short rate, thus the short rate process under $\mathbb{Q}^B$ gives a complete description of all bond prices (and all derivatives):

$$P(T; t) = B(t) E^{\mathbb{Q}^B} \left(\frac{1}{B(T)} \middle| \mathcal{F}_t \right).$$

The short rate process r implied by the volatility structure (21.3) generates a LIBOR market model. The short rate process under $\mathbb{Q}^B$ is given by (19.6):

$$dr(t) = \left(\frac{\partial f_0}{\partial T}(t) + \alpha^{\mathbb{Q}^B}(t, t) + \int_0^t \frac{\partial \alpha^{\mathbb{Q}^B}}{\partial T}(s, t) ds + \int_0^t \frac{\partial \sigma}{\partial T}(s, t) dW^{\mathbb{Q}^B}(s) \right) dt + \sigma(t, t) dW^{\mathbb{Q}^B}(t), \quad (19.6)$$

The drift of this short rate model is, as a function of $\{r(s) | 0 \leq s \leq t\}$, path dependent. Only in high dimensions, namely as a function of $\{L_i(t) | i = 0, \dots, n\}$, the model will be Markovian (i.e. the drift is no longer path dependent).

CHAPTER 22

Excursion: Shape of the Interest Rate Curve under Mean Reversion and a Multi-Factor Model

In this chapter we consider the influence of model properties like *mean reversion*, *number of factors*, *instantaneous correlation* und *instantaneous volatility* on the possible future shapes of the interest rate curve.

As in Chapter 18, which illustrated the relation of *instantaneous correlation* and *instantaneous volatility* to the *terminal correlation*, our goal is to develop an intuition for the significance of the model properties rather than looking at them rigorously in abstract mathematical terms. We thus pose the question how the interest rate curve differs *qualitatively* under different model configurations.

22.1. Model

As a model framework we use the [LIBOR Market Model](#). Due to its richness in parameters it gives us enough freedom to play with. We restrict the set of parameters and concentrate on three (important) parameters that are sufficient to create the interesting phenomenas we are interested in.

We restrict the model to a simple volatility structure, namely

$$\sigma_i(t) = \sigma^* \cdot \exp(-a \cdot (T_i - t)) \quad (22.1)$$

with $\sigma^* = 0.1$ and $a = 0.05$. We choose an equally simple correlation model, namely $dW_i dW_j = \rho_{i,j} dt$ with

$$\rho_{i,j} = \exp(-r \cdot |T_i - T_j|). \quad (22.2)$$

To this correlation model we apply a factor reduction (Principal Component Analysis), see Appendix A.5. The number of factors is the number of independent Brownian motions (effectively) entering the model, see Definition 45. Upon a factor reduction

the m largest eigenvalues of the correlation matrix are determined. Together with the corresponding eigenvectors a new correlation matrix constructed, having at most m non-zero eigenvalues. Through this process it may be guaranteed that the resulting correlation model defines a valid correlation matrix.

We simulate under the terminal measure and start with an initially flat interest rate curve $L_i(0) = 0.1, i = 0, 1, 2, \dots$

To summarize, our model framework consists of three degrees of freedoms which will be varied in our analysis:

Parameter	Effect
a	Damping of the exponentially decaying, time-homogenous volatility
r	Damping of the exponentially decaying instantaneous correlation
m	Number of factors extracted from the correlation matrix

Table 22.1.: Free parameters of the LIBOR Market Model considered.

22.2. Interpretation of the Figures

The Figures 22.1, 22.2, 22.3 and 22.4 show 100 paths of a Monte-Carlo simulation of the interest rate curve. The simulation was freezed at a fixed point in time ($t = 7.5$ in 22.2, 22.3 and 22.4 and $t = 17.5$ in 22.1). Left from this point the forward rates $L_i(T_i)$, each upon their individual maturity, are shown - this is a discrete analog to the [Short Rate](#). To the right of this point the future forward rate curve $L_j(t)$ is drawn.

The figures differ only in the parameters used to generate the paths. The same random numbers were used, thus the simulated paths depend smoothly on a and r .

To improve the visibility of the individual paths each path is given a different color, where the hue of the color depends smoothly on the level of the last rate $L_n(t)$.¹ This makes it very easy to check if the interest rate curves are parallel or exhibit some regular structure, see Figure 22.2.

22.3. Mean Reversion

We consider the example of a simple one factor Brownian motion ($\rho_{i,j} = 1$, i.e. $r = 0$). Figure 22.1 shows the simulated forward rates for different parameters a in (22.1).

¹ We choice of the last rate is arbitrary.

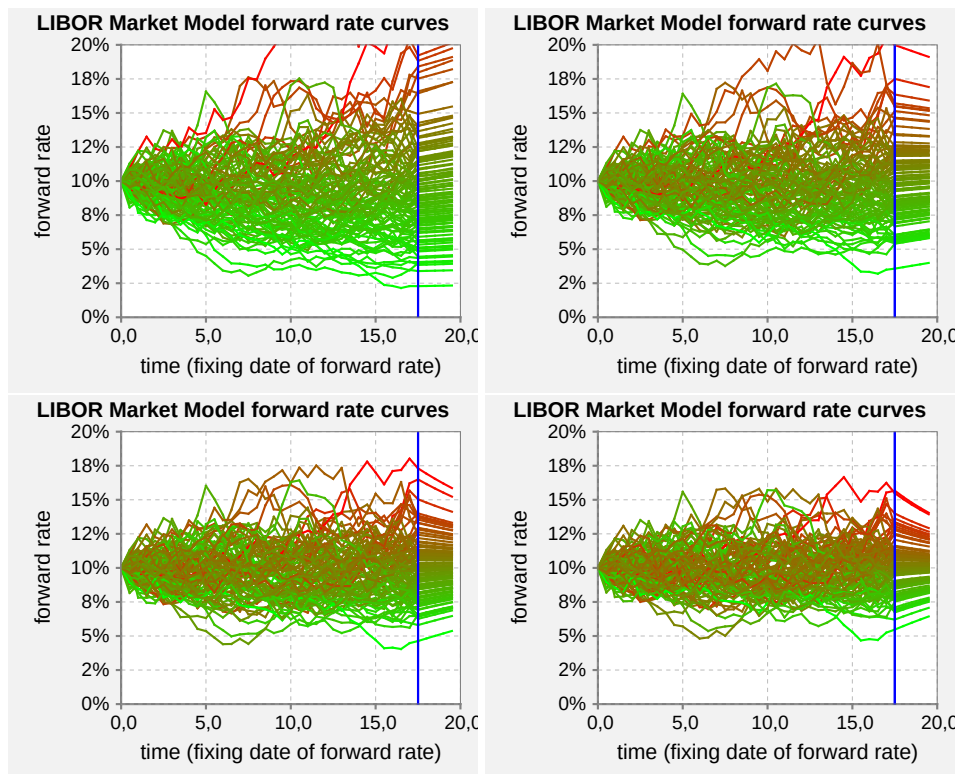


Figure 22.1.: Shape of the fixed rates $L_i(T_i)$ and the interest rate curve for different instantaneous volatilities (corresponds to different mean reversion) frozen at time $t = 17.5$ using a one factor model. Used were $a = 0$ (upper left), $a = 0.05$ (upper right), $a = 0.10$ (lower left) and $a = 0.15$ (lower right). For interpretation see Section 22.3.

From the derivation of the Hull-White-Model from the HJM-Framework it became obvious that an exponentially decreasing volatility structure of the forward rate corresponds to a *mean reversion* of the *short rate*, see page 309. We rediscover this property qualitatively here. Figure 22.1 shows 100 paths of a Monte-Carlo simulation of a LIBOR market model with different values for the parameter a : $a = 0$, $a = 0.05$ in the upper and $a = 0.1$, $a = 0.15$ in the lower row (left to right). Observe the fixed rates $L_i(T_i)$ left from the simulation time. They may be interpreted as a direct analog to the short rate. In Figure 22.1 it becomes obvious that with an increasing parameter a the paths develop a tendency to revert to the mean (*mean reversion*).

22.4. Factors

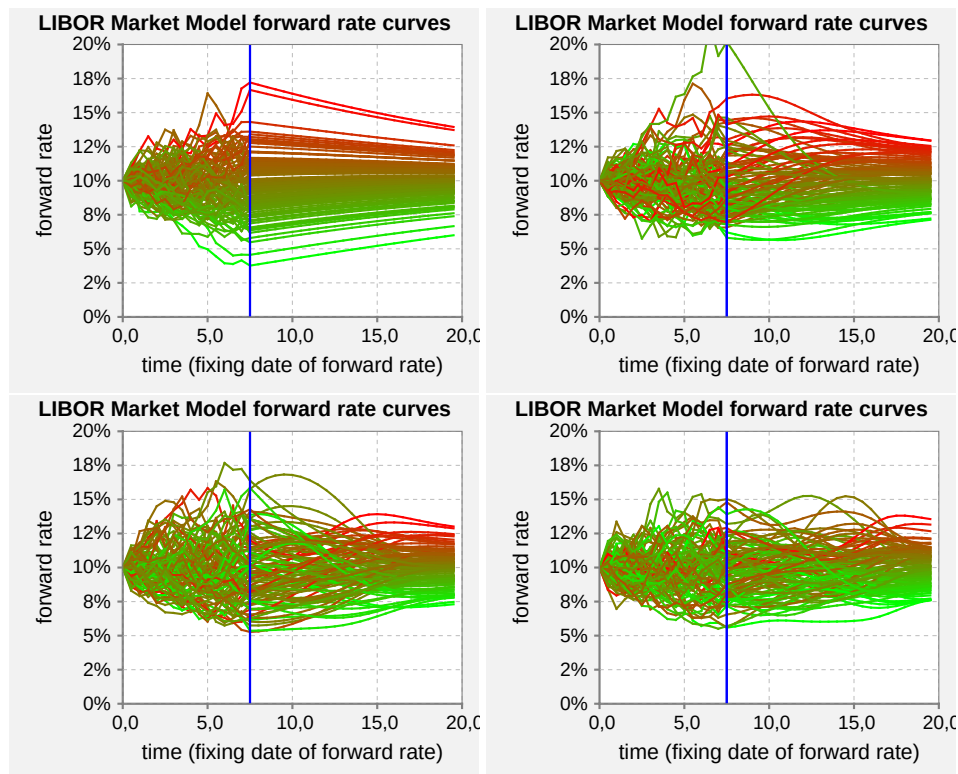


Figure 22.2.: Shape of the interest rate curve under different factor configurations, seen at time $t = 7.5$: One, two, three and five factors (from upper left to lower right). For interpretation see Section 22.4.

Figure 22.2 shows a Monte-Carlo simulation with the parameters above and different number of factors m . The possible shapes of the interest rate curve are given by combinations of the factors *parallel shift*, *tilt*, *bend* and oscillations with increasing frequencies, see also Figure A.1 on Page 395.

22.5. Exponential Volatility Function

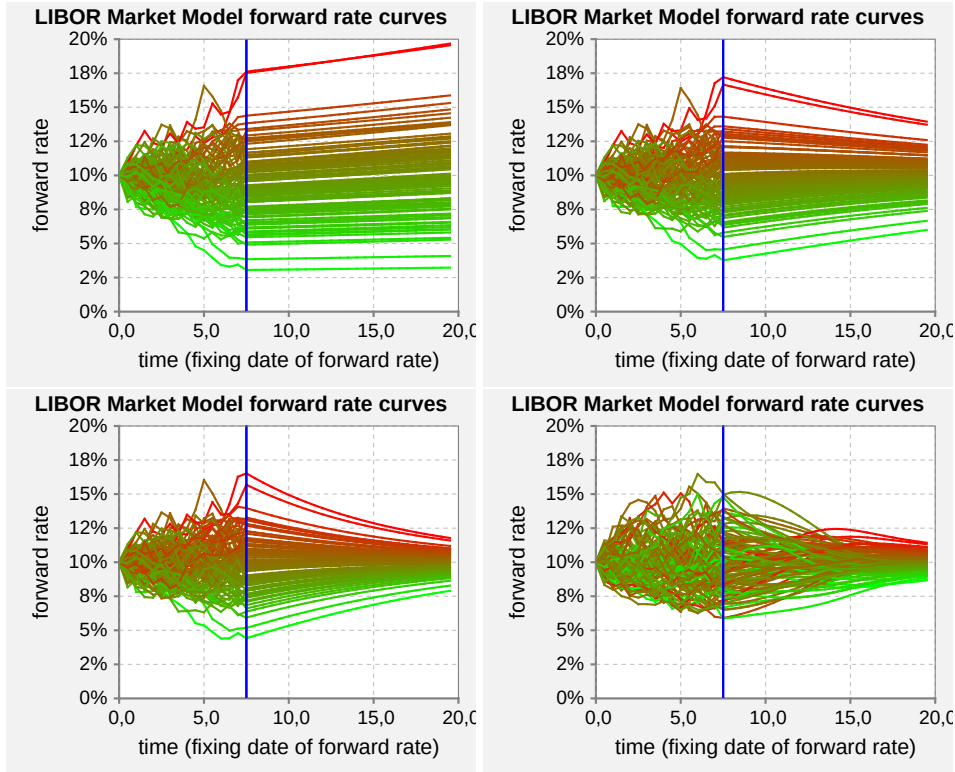


Figure 22.3.: Shape of the fixed rates $L_i(T_i)$ and the interest rate curve under different instantaneous volatilities (corresponds to mean reversion) at time $t = 7.5$ in a one factor model (upper row and lower left) with $a = 0.0$, $a = 0.05$ and $a = 0.1$ and a three factor model (lower right) with $a = 0.1$. For interpretation see Section 22.5.

As in Figure 22.1 we consider the Monte-Carlo simulation under different parameters a . First $a = 0$, $a = 0.05$ and $a = 0.1$ in a one factor model ($r = 0, m = 1$), and last $a = 0.1$ in a three factor model. We consider this at simulation time $t = 7.5$ and concentrate here on the part right of the simulation time, i.e. the interest rate curve $L_j(t)$ for $j > m(t)$.

It becomes apparent that the curve $\{L_j(t) \mid j > m(t)\}$ shows a shape similar to an exponential in j , depending on the parameter a , see Figure 22.3, lower left ($a = 0.1$) to the right of the simulation time. If we concise a one factor model (as it was used in the figure), we have:

$$L_j(t) = L_j(0) \cdot \exp \left(\int_0^t \mu_j(\tau) d\tau + \sqrt{\int_0^t \sigma_j(\tau) d\tau} \cdot W(t) \right).$$

For a fixed point in time t (and a state (path) ω) the interest rate curve shows the following dependence on j

$$j \mapsto L_j(0) \cdot \exp\left(\int_0^t \mu_j(\tau, \omega) d\tau\right) \cdot \exp\left(k \cdot \sqrt{\int_0^t \sigma_j^2(\tau) d\tau}\right)$$

where $k := W(t, \omega)$.

Especially for the volatility structure (22.1) we find

$$\begin{aligned} \int_0^t \sigma_j^2(\tau) d\tau &= \int_0^t \exp(-2a(T_j - \tau)) d\tau \\ &= \frac{1}{2a} (\exp(-2a(T_j - t)) - \exp(-2a(T_j - 0))) \\ &= \frac{1}{2a} \cdot (\exp(2a \cdot t) - 1) \cdot \exp(-2aT_j) \end{aligned}$$

$$i \mapsto L_j(0) \cdot \exp\left(\int_0^t \mu_j(\tau, \omega) d\tau\right) \cdot \exp(\tilde{k} \cdot \exp(-aT_j)), \quad (22.3)$$

where $\tilde{k} = k \cdot \sqrt{\frac{1}{2a}(\exp(2a \cdot t) - 1)}$.

The drift $\int_0^t \mu_j(\tau, \omega) d\tau$ is monotone increasing in j , see (17.8). This explains the shape of the interest rate curve in Figure 22.3 upper left. With increasing parameter a the interest rate curve is multiplied by the double exponential (22.3) with increasing steepness. This explains the shape of the interest rate curve in Figure 22.3 upper right and lower left. Only the addition of more driving factors allows for a richer family of possible curves. If the parallel movement (level) stays the dominant factor, then the shape (22.3) still dominates the interest rate curve, Figure 22.3 lower right.

22.6. Instantaneous Correlation

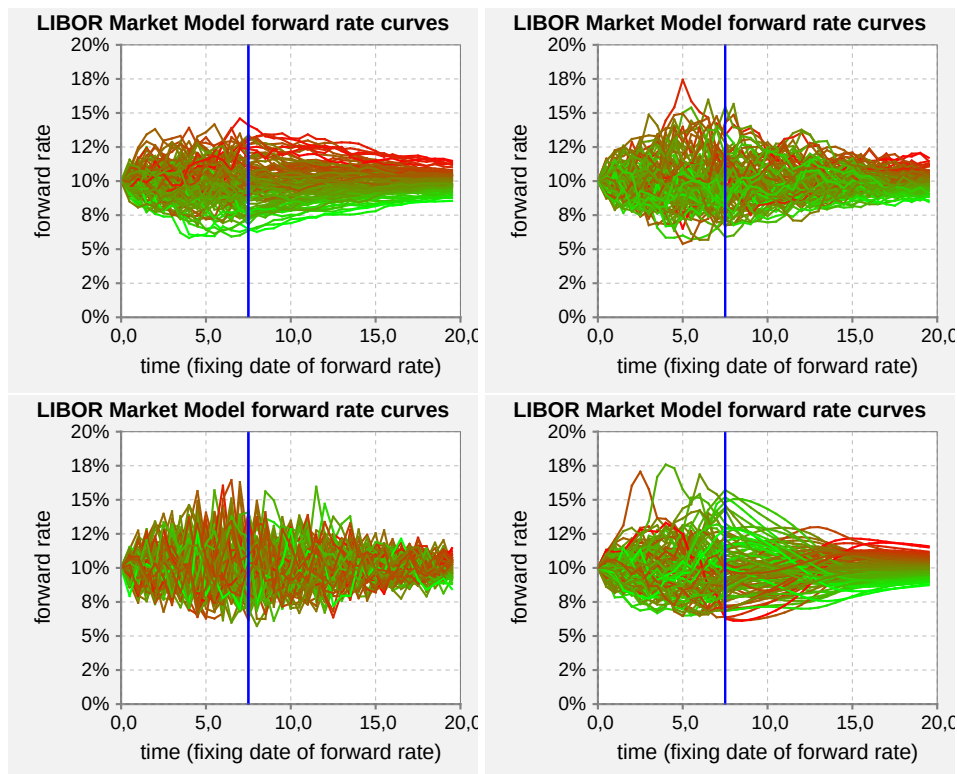


Figure 22.4.: Shape of the fixed rates $L_i(T_i)$ and the interest rate curve under different instantaneous correlations, seen at time $t = 7.5$. Used was a correlation matrix with (all) 40 factors and $r = 0.01$ (upper left, high correlation), $r = 0.1$ (upper right) and $r = 1.0$ (lower left, high decorrelation). In the lower right we used a correlation matrix with $r = 1.0$ (the same as in lower left), but reduced the number of factors to three. For interpretation see Section 22.6.

We fix a slightly decreasing volatility structure (22.1) with $a = 0.1$ and vary the parameter r of the correlation function (22.2). We do not apply a factor reduction, thus keep all 40 factors. The parameter $r = 0.01$ corresponds to an almost perfect correlation of the processes. Thus the possible shapes of the curve are almost parallel, the curve is very smooth since we started from a smooth (namely flat) curve. If the correlation parameter is increased to $r = 1.0$, then the distribution of rates within the curve is almost independent. See Figure 22.4 upper left, upper right and lower left.

It should be noted that this (terminal) decorrelation is also achievable under $r = 0.01$ by a proper choice of the volatility structure, see Chapter 18. The instantaneous decorrelation introduces an *additional* terminal decorrelation. The statement, that a model with perfect instantaneous correlation exhibits perfect terminal decorrelation of

the forward rates is wrong.

Finally we choose in Figure 22.4 lower right the parameter $r = 1.0$ again (as for the lower left with strong decorrelation), but apply a reduction to the three largest factors. It is obvious that this strongly reduces the possibilities for a decorrelation. The three factors only allow that the beginning, the middle and the end of the curve attain different values. Adjacent rates are still on similar levels.



rate curve.

Experiment: At <http://www.christian-fries.de/finmath/applets/LMMSimulation.html> a simulation of an interest rate curve with the model framework above may be found. The parameters may be chosen freely to study the different shaped of the interest

<|

CHAPTER 23

Markov Functional Models

23.1. Introduction



Motivation: From Chapter 5 we have a relation between the prices of European options and the probability distribution function (or probability density) of the underlying (under the martingale measure). If we consider an European option on some underlying, say the forward rate $L_i := L(T_i, T_{i+1}; T_i)$ (i.e. a caplet), then Lemma 72 allows to calculate the probability density of L_i from the given market prices. It seems as if this allows for a perfect calibration of a “model” to a continuum of given market prices. However, the terminal distribution alone does not determine a pricing model. What is missing is the specification of the dynamics, i.e. the specification of transition probabilities, and, of course, the specification of the Numeraire. This is the motivation for the *Markov functional modeling*. There we postulate a simple Markov process, e.g. $dx = \sigma(t)dW(t)$ for which the distribution function $\xi \mapsto P(x(T) \leq \xi)$ is known analytically. Then we require that the underlying L_i is a function of $x(T_i)$. Let us denote this function by g_i , i.e. let $L_i(\omega) = g_i(x(T_i, \omega))$ for all paths ω . If the functional g_i is strictly monotone, then we have with $K = g_i(\xi)$

$$F_{L_i}(K) := P(L_i \leq K) = P(g_i(x(T_i)) \leq K) = P(x(T_i) \leq \xi) =: F_{x(T_i)}(\xi) = F_{x(T_i)}(g_i^{-1}(K)).$$

With a given distribution function F_{L_i} of L_i (e.g. extracted from market prices through Lemma 72), the choice of the functional g_i allows the calibration while the process x (and the sequence of functionals $\{g_i\}$) describe the dynamics. To have a fully specified pricing model we further require the specification of the Numeraire as function of the Markov process x . To achieve this we may use Theorem 70, if

- x is given under the equivalent martingale measure $\mathbb{Q}$, and
- x generate the filtration.

Given a filtered probability space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{Q}, \{\mathcal{F}_t\})$. Consider a time-discretization $0 = T_0 < T_1 < \dots < T_n$. Financial products beyond T_n are not considered.

Let $t \mapsto N(t)$ denote the price process of a traded asset, which we choose a Numéraire and let $\mathbb{Q}$ denote the corresponding equivalent martingale measure. Then we have for any replicable asset price process $V(t)$ (see Definition 63 and Theorem 70) that

$$\frac{V(T_i)}{N(T_i)} = E^{\mathbb{Q}^N} \left(\frac{V(T_k)}{N(T_k)} \middle| \mathcal{F}_{T_i} \right).$$

Especially we have for every zero coupon bond $P(T_k)$, paying 1 in T_k , that

$$\begin{aligned} \frac{P(T_k; T_i)}{N(T_i)} &= E^{\mathbb{Q}^N} \left(\frac{1}{N(T_k)} \middle| \mathcal{F}_{T_i} \right), \text{ d.h.} \\ P(T_k; T_i) &= N(T_i) E^{\mathbb{Q}^N} \left(\frac{1}{N(T_k)} \middle| \mathcal{F}_{T_i} \right). \end{aligned}$$

Let x denote a $\mathcal{F}_t$ -adapted stochastic process with

$$dx(t) = \sigma(t) dW(t) \quad , \quad x(0) = x_0.$$

The filtration should be generated by x . On this space we consider time-discrete stochastic process, namely those for which their T_i -realization is a function of $(x(T_0), \dots, x(T_i))$, for all i . Especially we consider process for which their time T_i -realization is a function of $x(T_i)$ alone (i.e. independent of the process's history).

23.1.1. The Markov Functional Assumption (independent of the model considered)

We assume that the time T_i -realization of the numéraire process is a function of $x(T_i)$, i.e.

$$N(T_i, \omega) = N(T_i, x(T_i, \omega)), \quad (23.1)$$

where we use the same letter N for the (deterministic) functional $\xi \mapsto N(T_i, \xi)$.

Then, for any payoff $V(T_k)$ that is itself a function of $x(T_k)$ for some k , we have that the value process $V(T_i)$ for $i \leq k$ is

$$V(T_i) = N(T_i) E \left(\frac{V(T_k)}{N(T_k)} \middle| \mathcal{F}_{T_i} \right) = N(T_i, x(T_i)) E \left(\frac{V(T_k, x(T_k))}{N(T_k, x(T_k))} \middle| \mathcal{F}_{T_i} \right).$$

Thus, the time T_i realization of the value process $V(T_i)$ is also a functional of $x(T_i)$ which we denote with the same letter V . The functional $\xi \mapsto V(T_i, \xi)$ of the value process is

$$\xi \mapsto N(T_i, \xi) E \left(\frac{V(T_k, x(T_k))}{N(T_k, x(T_k))} \middle| \{x(T_i) = \xi\} \right).$$

Note: The markov functional assumption (23.1) may be relaxed such that the Numéraire is allowed to depend on $x(T_0), \dots, x(T_i)$. This relaxation is used in the LIBOR Markov Functional Model in spot measure.

23.1.2. Outline of this Chapter

In 23.2 we consider a Markov Functional Model for a stock (or any other non-interest rate related (single) asset). In 23.3 we will then consider a Markov Functional Model for the forward rate $L(T_i, T_{i+1}; T_i)$, which may be viewed as a time-discrete analog to the short rate. Both sections are essentially independent of each other. In Section 23.4 we will discuss how to implement a Markov functional model using a lattice in the state space.

23.2. Equity Markov Functional Model

23.2.1. Markov Functional Assumption

Consider a simple one dimensional Markov process, e.g.

$$dx(t) = \sigma(t) \cdot dW^{\mathbb{Q}}(t), \quad x(0) = x_0, \quad (23.2)$$

where σ is a deterministic function and $W^{\mathbb{Q}}$ denotes a $\mathbb{Q}$ -Brownian motion. Without loss of generality we may assume $x_0 = 0$. Equation (23.2) is the most simplest choice of a Markovian driver process. We will consider the addition of a drift term to (23.2) in our discussion on model dynamics in Section 23.2.5.

Let $S(t)$ denote the time t value of some asset for which we assume to have a continuum of European option prices. Let x and S be adapted stochastic processes defined on $(\Omega, \mathbb{Q}, \mathcal{F}_t)$, where $\{\mathcal{F}_t\}$ denotes the filtration generated by $W^{\mathbb{Q}}$.

We assume that the time t value of the asset S is a function of $x(t)$, i.e. we assume the existence of a functional $(t, \xi) \mapsto S(t, \xi)$ such that

$$S(t, \omega) = S(t, x(t, \omega)),$$

where the left hand side denotes our asset value at time t on path ω and the right hand side denotes some functional of our Markovian driver x , which we ambiguously name S . We allow for some ambiguity in notation here. From here on S will also denote a deterministic mapping (the functional)

$$(t, \xi) \mapsto S(t, \xi).$$

It will be clear from the arguments of S if we speak of the functional $(t, \xi) \mapsto S(t, \xi)$ or of the process $t \mapsto S(t)$.

For $t_1 < t_2$ we trivially have that

$$\frac{S(t_1)}{S(t_2)} = E^{\mathbb{Q}}\left(\frac{S(t_2)}{S(t_2)} \mid \mathcal{F}_{t_1}\right).$$

We now postulate that $\mathbb{Q}$ is the equivalent martingale measure with respect to the Numéraire S and that a universal pricing theorem holds for all other traded products, i.e. that their S relative price is a $\mathbb{Q}$ -martingale.

This implies that the zero coupon bond $P(T; t)$ having maturity T and being observed in $t < T$ fulfills

$$\frac{P(T; t)}{S(t)} = E^{\mathbb{Q}}\left(\frac{1}{S(T)} \mid \mathcal{F}_t\right).$$

Using the functional representation of S we find that $P(T; t)$ is represented as a functional of $x(t)$ too, namely

$$(t, \xi) \mapsto P(T; t)$$

with

$$\frac{P(T; t, \xi)}{S(t, \xi)} = E^{\mathbb{Q}}\left(\frac{1}{S(T, x(T))} \mid \{x(t) = \xi\}\right). \quad (23.3)$$

23.2.2. Example: The Black-Scholes Model

Let us assume a Markovian driver with constant instantaneous volatility $\sigma(t) = \sigma$. For the Black-Scholes Model we have

$$S(t, \xi) = S(0) \cdot \exp\left(r \cdot t + \frac{1}{2}\sigma_{\text{BS}}^2 t + \frac{\sigma_{\text{BS}}}{\sigma} \cdot \xi\right), \quad (23.4)$$

where σ_{BS} denotes the (constant) Black-Scholes volatility. Plugging this into (23.3) we find

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

so interest rates are indeed deterministic here.

This is indeed the Black-Scholes model: From the definition of the Markovian driver we have that $\frac{1}{\sigma} \cdot x(t) = W(t)$ and thus

$$S(t, x(t)) = S(0) \cdot \exp\left(r \cdot t + \frac{1}{2}\sigma_{\text{BS}}^2 t + \sigma_{\text{BS}} \cdot W(t)\right).$$

In other words the $\mathbb{Q}$ dynamics of S is¹

$$dS(t) = rS(t)dt + \sigma_{\text{BS}}^2 S(t)dt + \sigma_{\text{BS}} S(t) \cdot dW^{\mathbb{Q}}(t).$$

Introducing a new Numéraire

$$dB(t) = rB(t)dt, \quad B(0) = 1$$

we find for the change of Numéraire process $\frac{S}{B}$ that

$$d\frac{S}{B} = \sigma_{\text{BS}}^2 \frac{S}{B} dt + \sigma_{\text{BS}} \frac{S}{B} \cdot dW^{\mathbb{Q}}(t).$$

¹ Note that $\mathbb{Q}$ is the equivalent martingale measure with respect to the Numéraire S .

For $\frac{S}{B}$ to be a martingale under $\mathbb{Q}^B$ we have $dW^{\mathbb{Q}}(t) = dW^{\mathbb{Q}^B} - \sigma_{BS}^2 dt$ and thus we have

$$\begin{aligned} dS(t) &= rS(t)dt + \sigma_{BS} \cdot S(t) \cdot dW^{\mathbb{Q}}(t) \\ dB(t) &= rB(t)dt. \end{aligned}$$

Note: $dW^{\mathbb{Q}}(t)$ is a $\mathbb{Q}$ Brownian motion, where $\mathbb{Q}$ is the equivalent martingale measure with respect to the Numéraire S , while $dW^{\mathbb{Q}^B}(t)$ is a $\mathbb{Q}^B$ Brownian motion, where $\mathbb{Q}^B$ is the equivalent martingale measure with respect to the Numéraire B .

23.2.3. Numerical Calibration to a Full Two Dimensional European Option Smile Surface

As for the interest rate Markov functional model we are able to calculate the functionals numerically from a given two dimensional smile surface. Our approach here is similar the approach for the one-dimensional LIBOR Markov functional model under spot measure, [62]. Consider the following time T payout:

$$V(T, K; T) := \begin{cases} S(T) & \text{if } S(T) > K \\ 0 & \text{else.} \end{cases} \quad (23.5)$$

Obviously we have

$$V(T, K; T) = \max(S(T) - K, 0) + K \cdot \begin{cases} 1 & \text{if } S(T) > K \\ 0 & \text{else,} \end{cases}$$

i.e. the value of V is given by the value of a portfolio of one call option and K digital options, all having strike K . This is our calibration product.

23.2.3.1. Market Price

Let $\bar{\sigma}_{BS}(T, K)$ denote the Black-Scholes implied volatility surface given from market prices. Then we have that the market price of V is

$$\begin{aligned} V^{\text{market}}(T, K; 0) &= \underbrace{S(0)\Phi(d_+) - \exp(-rT)K\Phi(d_-)}_{\text{call option part}} + \underbrace{K \exp(-rT)(\Phi(d_-) + S(0)\sqrt{T}\Phi'(d_+)\frac{\partial \bar{\sigma}_{BS}(T, K)}{\partial K})}_{\text{digital part}} \\ &= S(0)\Phi(d_+) + KS(0)\sqrt{T}\Phi'(d_+)\frac{\partial \bar{\sigma}_{BS}(T, K)}{\partial K}, \end{aligned}$$

$$\text{where } \Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp(-\frac{y^2}{2}) dy \text{ and } d_{\pm} = \frac{\ln(\frac{\exp(rT)S(0)}{K}) \pm \frac{1}{2}\bar{\sigma}_{BS}^2(T, K)T}{\bar{\sigma}_{BS}(T, K)\sqrt{T}}.$$

23.2.3.2. Model Price

Within our model the price of the product (23.5) is

$$\begin{aligned} V^{\text{model}}(T, K; 0) &= S(0) \cdot \mathbb{E}^{\mathbb{Q}}\left(\frac{S(T, x(T)) \cdot \mathbf{1}_{\{S(T, x(T)) > K\}}}{S(T)} \mid \{x(0) = x_0\}\right) \\ &= S(0) \cdot \mathbb{E}^{\mathbb{Q}}(\mathbf{1}_{\{S(T, x(T)) > K\}} \mid \{x(0) = x_0\}) \end{aligned}$$

Assuming that our functional $(T, \xi) \mapsto S(T, \xi)$ is monotone increasing in ξ we may write

$$V^{\text{model}}(T, K; 0) = S(0) \cdot \mathbb{E}^{\mathbb{Q}}(\mathbf{1}_{\{x(T) > x^*\}} \mid \{x(0) = x_0\}), \quad (23.6)$$

where x^* is the (unique) solution of $S(T, x^*) = K$. Note that (23.6) depends on x^* and the probability distribution of $x(T)$ only and that $x(T)$ is known due to the simple form of our Markovian driver. It does not depend on the functional S ! Thus for given x^* we may calculate

$$V^{\text{model}}(T, x^*; 0) := S(0) \cdot \mathbb{E}^{\mathbb{Q}}(\mathbf{1}_{\{x(T) > x^*\}} \mid \{x(0) = x_0\}).$$

23.2.3.3. Solving for the Functional

For given x^* we now solve the equation

$$V^{\text{market}}(T, K^*; 0) = V^{\text{model}}(T, x^*; 0)$$

to find $S(T, x^*) = K^*$ and thus the functional form $(T, \xi) \mapsto S(T, \xi)$. This can be done very efficiently using fast one dimensional root finders, e.g. Newton's method.

23.2.4. Interest Rates

23.2.4.1. A Note on Interest Rates and the No-Arbitrage Requirement

Functional models for equity option pricing have been investigated before, see e.g. [?] and [?]. However, their approach chooses deterministic interest rates and chooses the bank account as Numéraire. As suggested in Section 23.2.2 this will impose a very strong self similarity requirement on the functionals (which is fulfilled by the Black-Scholes model). Such models may calibrate only to a one dimensional sub-manifold of a given implied volatility surface, see [?]. For the Markov functional model this follows directly from (23.3). Assuming that the Markovian driver x is given and that the interest rate dynamic $P(T; t, \xi)$ is given, we find from (23.3) that

$$S(t, \xi) = \frac{P(T; t, \xi)}{\mathbb{E}^{\mathbb{Q}}\left(\frac{1}{S(T, x(T))} \mid \{x(t) = \xi\}\right)}.$$

So once a terminal time T functional $\xi \mapsto S(T, \xi)$ has been defined, all other functionals are implied by the interest rate dynamics P and the dynamics of the Markovian driver.

Sticking to prescribed interest rates, the only way to allow for more general functional is to violate the no-arbitrage requirement (23.3) or change the Markovian driver. The latter will be considered in Section 23.2.5.

23.2.4.2. Where are the Interest Rates?

Our model calibrates to a continuum of options on S . We do not even specify interest rates. This is not necessary, since the specification of the interest rates is already contained in the specification of a continuum of options on S . Consider options on $S(T)$, i.e. options with maturity T . First note that from a continuum $K \mapsto V_{\text{call}}^{\text{market}}(T, K; 0)$ of market prices for call option payouts

$$V_{\text{call}}^{\text{market}}(T, K; T) = \max(S(T) - K, 0)$$

we obtain prices for the corresponding digital payouts

$$V_{\text{digital}}^{\text{market}}(T, K; T) = \begin{cases} 1 & S(T) > K \\ 0 & \text{else} \end{cases}$$

by

$$V_{\text{digital}}^{\text{market}}(T, K; 0) = -\frac{\partial}{\partial K} V_{\text{call}}^{\text{market}}(T, K; 0).$$

Thus the value of the zero coupon bond with maturity T is

$$P(T; 0) = \lim_{K \searrow 0} V_{\text{digital}}^{\text{market}}(T, K; 0) = -\lim_{K \searrow 0} \frac{\partial}{\partial K} V_{\text{call}}^{\text{market}}(T, K; 0).$$

Note that this argument is model independent.

Within the functional model, Equation (23.2.4.2) hold locally in each state. Given we are at time t in state $x(t) = \xi$ we have for the corresponding bond that

$$P(T; t, \xi) = \lim_{K \searrow 0} V_{\text{digital}}^{\text{model}}(T, K; t, \xi).$$

From this it is clear why specifying interest rates would represent a violation of the no-arbitrage requirement. This is precisely the reason why the models in [?, ?] allow for arbitrage.

In the next section we show that the model implied interest rate dynamics will likely be undesirable. However, as it is known from interest rate hybrid Markov functional models [62], it is possible to calibrate to different model dynamics by changing the Markovian driver x .

23.2.5. Model Dynamics

23.2.5.1. Introduction

Markov functional models calibrate perfectly to a continuum of option prices, i.e. to the market implied probability density of the underlying, cf. [50]. Indeed, the functional $(t, \xi) \mapsto S(t, \xi)$ is nothing more than the measure transformation from the probability density of $x(t)$ to the market implied probability density of the underlying $S(t)$.

While the calibration to terminal probability densities is a desirable feature, it is not the only requirement on a model, specifically if the model is used to price complex derivatives like Bermudan options. Here the transition probabilities play a role, i.e. the model dynamics. The most prominent aspects of model dynamics are

- **Interest Rate Dynamics:** For an equity Markov functional model the joint movement of the interest rate and the asset has to be analysed. It is possible to calibrate to a given interest rate dynamics by adding a drift to the Markovian driver, see Section 23.2.5.2.
- **Forward volatility:** This is the implied volatility of an option with maturity T and strike K , given we are in state (t, ξ) , i.e.

$$S(t, \xi) \cdot E^Q\left(\frac{\max(S(T, x(T)) - K, 0)}{S(T)} \mid \{x(t) = \xi\}\right).$$

Obviously it will play an important role for compound options and Bermudan options. The forward volatility may be calibrated by changing the instantaneous volatility of the Markovian driver, see Section 23.2.5.3.

- **Auto correlation / Forward spread volatility:** The auto correlation of the process S impacts the forward spread volatility. This is the implied volatility of an option on $S(T_2) - S(T_1)$ with maturity T_2 , given we are in state (t, ξ) , i.e.

$$S(t, \xi) \cdot E^Q\left(\frac{\max(S(T_2, x(T_2)) - S(T_1, x(T_1)), 0)}{S(T_2, x(T_2))} \mid \{x(t) = \xi\}\right).$$

Markov functional models allow a limited calibration to different model dynamics by changing the dynamics of the Markovian driver x . For our choice

$$dx = \sigma(t) \cdot dW(t)$$

we may change the auto correlation of x by choosing different instantaneous volatility functions σ . Since the calibration of the functionals is scale invariant with respect to the terminal standard deviation $\bar{\sigma}(t)$ of $x(t)$, the calibration to the terminal probability densities is independent of the choice of σ . See the Black-Scholes example in Section 23.2.2 for an example of this invariance.

Time Copula

The specification of the auto correlation of x (through σ) is sometime called *time copula*, [?], since it may be specified through the joint distribution of $(x(t_1), x(t_2))$. For this reasons similar functional models are sometimes called *copula models*, a term that is more associated with credit models, where joint default distribution is constructed from marginal default distributions.

In addition to a specification of the instantaneous volatility, the Markovian driver may be endowed with a drift.

Time-Discrete Markovian Driver

We assume a given time discretization $\{0 = t_0 < t_1 < t_2 < \dots\}$ and consider the realizations $x(t_i)$ of the Markovian driver x given trough increments $\Delta x(t_i) = x(t_{i+1}) - x(t_i)$. It is natural that a practical implementation of the model will feature a certain time discretization. Thus, speaking of calibration of a specific time discretized implementation, it is best to consider the Markovian driver given through an Euler scheme (as in (23.7), (23.8))

$$x(t_{i+1}) = x(t_i) + \mu(t_i, x(t_i))\Delta t_i + \sigma(t_i, x(t_i))\Delta W(t_i).$$

23.2.5.2. Interest Rate Dynamics

Example: Black-Scholes Model with a Term Structure of Volatility

Let us first assume that the Markovian driver is given by

$$x(t_{i+1}) = x(t_i) + \sigma_i \Delta W(t_i). \quad (23.7)$$

Consider a term structure of Black-Scholes implied volatilities, i.e. let $\bar{\sigma}_{BS}(t_i)$ denote the implied volatility of an option with maturity t_i . Assuming the simple Markovian driver (23.7) the corresponding functionals that calibrate to these options are

$$S(t_i, \xi) = S(0) \cdot \exp\left(r \cdot t_i + \frac{1}{2} \bar{\sigma}_{BS}(t_i)^2 t + \frac{\bar{\sigma}_{BS}(t_i)}{\bar{\sigma}_i} \cdot \xi\right),$$

where $\bar{\sigma}_i^2 := \frac{1}{t_i} \sum_{j=0}^{i-1} \sigma_j^2 \Delta t_j$.

Already within this model a stochastic interest rate dynamic is implied. From (23.3) we find

$$\begin{aligned} P(t_{i+1}; t_i, \xi) &= S(t_i, \xi) \cdot \mathbb{E}^{\mathbb{Q}}\left(\frac{1}{S(t_{i+1}, x(t_{i+1}))} \mid \{x(t_i) = \xi\}\right) \\ &= \exp\left(-r \cdot (t_{i+1} - t_i) - \frac{1}{2} (\bar{\sigma}_{BS}(t_{i+1})^2 \cdot (t_{i+1} - \frac{\sigma_i^2}{\bar{\sigma}_{i+1}^2} \Delta t_i) - \bar{\sigma}_{BS}(t_i)^2 \cdot t_i) \right. \\ &\quad \left. - \left(\frac{\bar{\sigma}_{BS}(t_{i+1})}{\bar{\sigma}_{i+1}} - \frac{\bar{\sigma}_{BS}(t_i)}{\bar{\sigma}_i}\right) \cdot \xi\right), \end{aligned}$$

If the volatility of the Markovian driver x decays faster than the implied Black-Scholes volatility, then interest rate will move positively correlated with the stock. If the volatility of the Markovian driver x decays slower than the implied Black-Scholes volatility, then interest rate will move negatively correlated.² If we choose the instantaneous volatility of x such that $\bar{\sigma}_i = \bar{\sigma}_{BS}(t_i)$ we have

$$P(t_{i+1}; t_i, \xi) = \exp(-r \cdot (t_{i+1} - t_i))$$

i.e. we recovered a model with deterministic interest rates. Reconsidering the case of a continuous driver $dx = \sigma(t)dW(t)$, we see that for $\frac{1}{t} \int_0^t \sigma(\tau)^2 d\tau = \bar{\sigma}_{BS}(t)^2$ the functionals above define a Black-Scholes model with instantaneous volatility σ .

However, we do not need to sacrifice the instantaneous volatility of x to match the interest rate dynamics. A much more natural choice is to add a suitable drift to the Markovian driver x . Consider a Markovian driver x such that

$$x(t_{i+1}) = x(t_i) + \alpha_i x(t_i) \Delta t_i + \sigma_i \Delta W(t_i), \quad x(t_0) = x_0 = 0. \quad (23.8)$$

Note that the $x(t_i)$'s are normal distributed with mean 0 (assuming $x_0 = 0$) and standard deviations $\gamma_i \sqrt{t_i}$ where

$$\gamma_{i+1}^2 t_{i+1} = \gamma_i^2 t_i \cdot (1 + \alpha_i \Delta t_i) + \sigma_i^2 \Delta t_i.$$

Together with the the functionals

$$\begin{aligned} S(t_i, \xi) &= S(0) \cdot \exp(r \cdot t_i + \frac{1}{2} \bar{\sigma}_{BS}(t_i)^2 + \frac{\bar{\sigma}_{BS}(t_i)}{\gamma_i} \cdot \xi) \\ S(t_{i+1}, \xi) &= S(0) \cdot \exp(r \cdot t_{i+1} + \frac{1}{2} \bar{\sigma}_{BS}(t_{i+1})^2 + \frac{\bar{\sigma}_{BS}(t_{i+1})}{\gamma_{i+1}} \cdot \xi) \end{aligned}$$

We have

$$\begin{aligned} P(t_{i+1}; t_i, \xi) &= S(t_i, \xi) \cdot E^Q\left(\frac{1}{S(t_{i+1}, x(t_{i+1}))} \mid \{x(t_i) = \xi\}\right) \\ &= \exp(-r \cdot \Delta t_i - \frac{1}{2} (\bar{\sigma}_{BS}(t_{i+1})^2 \cdot (t_{i+1} - \frac{\sigma_i^2}{\gamma_{i+1}^2} \Delta t_i) - \bar{\sigma}_{BS}(t_i)^2 \cdot t_i) \\ &\quad - (\frac{\bar{\sigma}_{BS}(t_{i+1})}{\gamma_{i+1}} (1 + \alpha_i \Delta t_i) - \frac{\bar{\sigma}_{BS}(t_i)}{\gamma_i}) \cdot \xi) \\ &= \exp(-r \cdot \Delta t_i - \frac{1}{2} (\bar{\sigma}_{BS}(t_{i+1})^2 \cdot \frac{\gamma_i^2}{\gamma_{i+1}^2} (1 + \alpha_i \Delta t_i)^2 - \bar{\sigma}_{BS}(t_i)^2 \cdot t_i) \\ &\quad - (\frac{\bar{\sigma}_{BS}(t_{i+1})}{\gamma_{i+1}} (1 + \alpha_i \Delta t_i) - \frac{\bar{\sigma}_{BS}(t_i)}{\gamma_i}) \cdot \xi). \end{aligned}$$

² Note than in average the interest rate is still r .

Choosing α_i such that

$$\left(\frac{\bar{\sigma}_{BS}(t_{i+1})}{\gamma_{i+1}} (1 + \alpha_i \Delta t_i) - \frac{\bar{\sigma}_{BS}(t_i)}{\gamma_i} \right) = 0 \quad (23.9)$$

we have

$$P(t_{i+1}; t_i, \xi) = \exp(-r \cdot (t_{i+1} - t_i))$$

Interestingly, the Markov functional model (23.8)–(23.9) does necessarily need to be a Black-Scholes model having the $\mathbb{Q}^B$ -dynamics

$$\begin{aligned} dS(t) &= rS(t)dt + \sigma_{BS}(t) \cdot S(t) \cdot dW^{\mathbb{Q}^B}(t) \\ dB(t) &= rB(t)dt, \end{aligned} \quad (23.10)$$

where $\bar{\sigma}_{BS}(t_i)^2 = \frac{1}{t_i} \int_0^{t_i} \sigma_{BS}(t)^2 dt$. The two models are not the same, although their terminal probability densities (European option prices) and their interest rate dynamics agree. The difference lies in the forward volatility, which may be changed for the Markov functional model through the instantaneous volatility of x . Only for $\sigma(t) = \sigma_{BS}(t)$ we will have the dynamics (23.10). In this case the α_i in (23.9) will be zero, i.e. we are in the situation of the previous example.

Calibration to arbitrary Interest Rate Dynamics

Within the no-arbitrage constraints, it is possible to calibrate the model to a given arbitrary interest rate dynamics by choosing the appropriate drift. To do so, we have to find $\mu(t_i, \xi)$ such that

$$P(t_{i+1}; t_i, \xi) = E^{\mathbb{Q}}\left(\frac{S(t_i, \xi)}{S(t_{i+1}, x(t_{i+1}))} \mid \{x(t_i) = \xi - \mu(t_i, \xi)\Delta t_i\}\right).$$

This can be done numerically through a one-dimensional root finder. The functional $S(t_{i+1})$ has to be re-calibrated in every iteration.³

23.2.5.3. Forward Volatility

The calibration to European option prices (Section 23.2.3) and joint movements of asset and interest rates (Section 23.2.5.2) still leaves the instantaneous volatility σ of the Markovian driver x a free parameter. It may be used to calibrate the forward volatility, i.e. the volatility of an option, conditional we are at time $t > 0$ in state ξ .

³ The procedure is the same as for done for the calibration of the FX forward within the cross currency Markov functional model, [62].

Example: Black-Scholes Model with a Term Structure of Volatility

Consider the simple Black-Scholes like example from Section 23.2.5.2. For simplicity we consider a Markovian driver without drift, i.e.

$$x(t_{i+1}) = x(t_i) + \sigma_i \Delta W(t_i).$$

together with functionals

$$S(t_i, \xi) = S(0) \cdot \exp\left(r \cdot t_i + \frac{1}{2} \bar{\sigma}_{BS}(t_i)^2 t + \frac{\bar{\sigma}_{BS}(t_i)}{\bar{\sigma}_i} \cdot \xi\right)$$

calibrating to European options with implied volatility $\bar{\sigma}_{BS}(t_i)$.⁴

Then we have that the standard deviation of the increment $x(t_k) - x(t_i)$ is $\bar{\sigma}_{t_i, t_k} \cdot \sqrt{(t_k - t_i)}$, where $\frac{1}{t_k - t_i} \sum_{j=i}^{k-1} \sigma_j^2 \Delta t_j$. It follows that the implied volatility of an option with maturity t_k , given we are in state (t_i, ξ) , is

$$\bar{\sigma}_{BS}(t_k) \cdot \frac{\bar{\sigma}_{t_i, t_k}}{\bar{\sigma}_{t_k}}.$$

Thus a decay in the instantaneous volatility of the driver process will result in a forward volatility decaying with simulation time t_i (for fixed maturity t_k).

Example: Exponential Decaying Instantaneous Volatility

Let us consider the case of a time continuous Markovian driver $dx = \sigma(t)dW$ with a decaying instantaneous volatility

$$\sigma(t) = \exp(-a \cdot t), \quad a \neq 0.$$

Then we have $\bar{\sigma}_{t_1, t_2} := \sqrt{\frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \sigma(t) dt} = \sqrt{\frac{-1}{2a(t_2 - t_1)} (\exp(-2at_2) - \exp(-2at_1))}$. Assuming functionals

$$S(t, \xi) = S(0) \cdot \exp\left(r \cdot t + \frac{1}{2} \bar{\sigma}_{BS}(t)^2 t + \frac{\bar{\sigma}_{BS}(t)}{\bar{\sigma}(0, t)} \cdot \xi\right)$$

we have that the forward volatility for an option with maturity T , given we are in t , is

$$\bar{\sigma}_{BS}(T) \cdot \frac{\bar{\sigma}(t, T)}{\bar{\sigma}(0, T)} = \bar{\sigma}_{BS}(T) \cdot \sqrt{\frac{T}{T-t} \frac{\exp(-2at) - \exp(-2aT)}{1 - \exp(-2aT)}}.$$

⁴ As before we use the notation $\bar{\sigma}_i^2 := \frac{1}{t_i} \sum_{j=0}^{i-1} \sigma_j^2 \Delta t_j$.

23.2.6. Implementation

The model may be implemented in the same way as it is done for a one dimensional Markov functional LIBOR model. The basic steps for a numerical implementation are, see Figure ??:

- Choose a suitable discretization of $(t, x(t))$, i.e. set up a grid $(t_i, x_{i,j})$.
- For each $x^* = x_{i,j}$:
 - Calculate $V^{\text{model}}(T, x^*; 0)$.
 - Find K^* such that $V^{\text{market}}(T, K^*; 0) = V^{\text{model}}(T, x^*; 0)$.
 - Set $S(x_{i,j}) := K^*$.

23.3. LIBOR Markov Functional Model

We postulate that the forward rate viewed upon its reset date (fixing) may be given as a function of the realization of the underlying Markov process x :

$$\begin{aligned} \text{The forward rate (LIBOR)} L(T_k) &= L_k(T_k) = \frac{1 - P(T_{k+1}; T_k)}{P(T_{k+1}; T_k)(T_{k+1} - T_k)} \\ \text{(seen upon its reset date } T_k) &\text{ is a (deterministic) function of } x(T_k), \\ \text{where } x &\text{ is a Markov process of the form} \\ dx &= \sigma(t)dW \quad \text{under } \mathbb{Q}^N, \quad x(0) = x_0. \end{aligned} \tag{23.11}$$

At that point we left the choice of the Numéraires N and the corresponding martingale measure $\mathbb{Q}$ open. We will make this choice now, and depending on it we obtain a Markov functional model in *terminal measure* (Section 23.3.1) or in *spot measure* (Section ??).

23.3.1. LIBOR Markov Functional Model in Terminal Measure - Hunt, Kennedy, Pelsser

We choose as Numéraire the T_n -bond $N := P(T_n)$. The measure $\mathbb{Q}$ should denote the corresponding martingale measure (*terminal measure*). From assumption 23.11 we have:

Lemma 209 (Numéraire of the Markov Functional Model under Terminal Measure): The Numéraire $N(T_i) = P(T_n; T_i)$ is a (deterministic) function of $x(T_i)$,

i.e. $N(T_i) = N(T_i, x(T_i))$. For the functional $\xi \mapsto N(T_i, \xi)$ we have the recursion

$$N(T_n, \xi) = 1$$

$$N(T_i, \xi) = \left(\mathbb{E}^{\mathbb{Q}} \left(\frac{1}{N(T_{i+1}, x(T_{i+1}))} | \{x(T_i) = \xi\} \right) \cdot (1 + L_i(T_i, \xi)(T_{i+1} - T_i)) \right)^{-1}. \quad (23.12)$$

Remark 210 (Notation for the Functionals): Here and in the following we denote the functionals with the same symbol as the corresponding random variables they are representing. In the equation

$$N(T_i) = N(T_i, x(T_i))$$

the $N(T_i)$ on the left hand side denotes the random variable, on the right hand side $\xi \mapsto N(T_i, \xi)$ denotes a function. If we would use different symbols it would rather reduce the readability of the following text. The difference of the two will be obvious from the additional argument ξ or $x(T_i)$ in its place.

Proof: Since $\mathbb{Q}$ is the corresponding martingale measure we have may use the *pricing theorem*. Thus we have for the zero coupon bond $P(T_k)$ (maturity in T_k)

$$\frac{P(T_k; T_i)}{N(T_i)} = \mathbb{E}^{\mathbb{Q}} \left(\frac{1}{N(T_k)} | \mathcal{F}_{T_i} \right), \text{ d.h.}$$

$$P(T_k; T_i) = N(T_i) \mathbb{E}^{\mathbb{Q}} \left(\frac{1}{N(T_k)} | \mathcal{F}_{T_i} \right).$$

Since the process x generates the filtration $\{\mathcal{F}_{T_i}\}$ and since x is Markovian, it is sufficient to know the $\mathcal{F}_{T_i}$ -mesurable random variable $P(T_k; T_i)$ on the set $\{x(T_i) = \xi\}$. Thus we have that the bond $P(T_k)$ seen at time T_i may given as a function of $x(T_i)$, namely as $P(T_k; T_i) = P(T_k; T_i, x(T_i))$ with

$$P(T_k; T_i, \xi) := N(T_i, \xi) \mathbb{E}^{\mathbb{Q}} \left(\frac{1}{N(T_k, x(T_k))} | \{x(T_i) = \xi\} \right) \quad (23.13)$$

Since (compare Definition 88)

$$1 + L_i(T_i)(T_{i+1} - T_i) := \frac{P(T_i; T_i)}{P(T_{i+1}; T_i)} = \frac{1}{P(T_{i+1}; T_i)},$$

we have

$$1 + L_i(T_i, \xi)(T_{i+1} - T_i) = \left(N(T_i, \xi) \mathbb{E}^{\mathbb{Q}} \left(\frac{1}{N(T_{i+1}, x(T_{i+1}))} | \{x(T_i) = \xi\} \right) \right)^{-1},$$

and thus (23.12). □

With $N(T_n) = P(T_n; T_n) = 1$, we have from (23.12) a recursion to determine the functionals $N(T_i, \xi)$ from the functionals $L_i(T_i, \xi)$.

With the specification of the numéraire the model is fully described. Die functionals $\xi \mapsto L_i(T_i, \xi)$ are the free quantities that may be used to calibrate the model.

23.3.1.1. Evaluation within the LIBOR Markov Functional Model

As preparation for the discussion of the calibration of the model we consider the evaluation of a [caplet](#) and a [digital caplet](#) with the LIBOR Markov functional model.

Valuation of a Caplets within the LIBOR Markov Functional Model

Let T_i denote the fixing and T_{i+1} the payment date of caplet with notional 1 and strike K . Then we have for its value $V_{\text{caplet}}(T_0)$ in the LIBOR Markov functional model

$$\begin{aligned} V_{\text{caplet}}(T_0) &= N(T_0) \cdot E^{\mathbb{Q}}\left(\frac{\max(L(T_i, x(T_i)) - K, 0) \cdot (T_{i+1} - T_i)}{N(T_{i+1}, x(T_{i+1}))} \mid \{x(T_0) = x_0\}\right) \\ &= N(T_0) \cdot E^{\mathbb{Q}}(\max(L(T_i, x(T_i)) - K, 0) \cdot (T_{i+1} - T_i) \cdot \bar{P}(T_{i+1}; T_i, x(T_i)) \mid \{x(T_0) = x_0\}), \end{aligned} \quad (23.14)$$

where $\bar{P}(T_{i+1}; T_i, \xi) = E^{\mathbb{Q}}\left(\frac{1}{N(T_{i+1}, x(T_{i+1}))} \mid \{x(T_i) = \xi\}\right)$.

Valuation of a Digital Caplets within the LIBOR Markov functional model

Let T_i denote the fixing and T_{i+1} the payment date of digital caplet with notional 1 and strike K . Then we have for its value $V_{\text{digital}}(T_0)$ in the LIBOR Markov functional model

$$\begin{aligned} V_{\text{digital}}(T_0) &= N(T_0) \cdot E^{\mathbb{Q}}\left(\frac{\mathbf{1}(L(T_i, x(T_i)) - K) \cdot (T_{i+1} - T_i)}{N(T_{i+1}, x(T_{i+1}))} \mid \{x(T_0) = x_0\}\right) \\ &= N(T_0) \cdot E^{\mathbb{Q}}(\mathbf{1}(L(T_i, x(T_i)) - K) \cdot (T_{i+1} - T_i) \cdot \bar{P}(T_{i+1}; T_i, x(T_i)) \mid \{x(T_0) = x_0\}), \end{aligned} \quad (23.15)$$

where $\bar{P}(T_{i+1}; T_i, \xi) = E^{\mathbb{Q}}\left(\frac{1}{N(T_{i+1}, x(T_{i+1}))} \mid \{x(T_i) = \xi\}\right)$. A detailed description of how the given expectations may be numerically calculated, i.e. how to implement the model, is given in Section [23.4](#).

23.3.1.2. Calibration of the LIBOR Functional

The LIBOR functionals may be derived from market prices of [caplets](#)⁵. One possibility is to give a parametrization of the functional and optimize the parameters by comparing the model prices to given market prices. Another possibility is to derive the functional point-wise from a continuum of given market prices. For this we have to know the market prices of the caplets with periods $[T_i, T_{i+1}]$ and strikes $K \in [0, \infty)$. If the prices are known only for a finite number of strikes K_j then we require a corresponding interpolation of the caplet prices. This is non-trivial, see Chapter [6](#).

⁵ Definition [115](#).

Calibration of parametrized LIBOR Functionals

If we assume a parametrized functional form for $L(T_i)$, then we may calculate the model prices of caplet on $L(T_i)$ from a known functional for $N(T_{i+1})$. This allows to optimize the parameters of a parametrized functional form to achieve the replication of given market prices. Thus we have the following backward induction:

Induction start:

- $N(T_n) = 1$.

Induction step ($T_{i+1} \rightarrow T_i$):

- Calculate $\bar{P}(T_{i+1}; T_i, \xi) = E^{\mathbb{Q}}\left(\frac{1}{N(T_{i+1}, x(T_{i+1}))} \mid \{x(T_i) = \xi\}\right)$.
- “Optimize” $L(T_i)$ by comparing (23.14) with given market prices.
- Calculate $N(T_i)$ from $L(T_i)$ and $\bar{P}(T_{i+1}; T_i)$ via (23.12).

If $L(T_i)$ should be almost lognormal distributed, then a good starting point for the parametrization is an exponential. In [24] the following parametrization is discussed:

$$L_i(\xi) = \exp(a + b\xi + c(\xi - d)^2).$$

Calibration of discretized LIBOR functionals (point-wise)

The evaluation of a digital caplet allows to calculate the value $L(T_i, x^*)$ of the functional for an arbitrary point x^* . We make the following additional assumption:

The LIBOR functional $\xi \mapsto L(T_i, \xi)$ is
strictly monotone increasing in ξ .

(23.16)

With this assumption we may give a simple algorithm to derive the functional $L(T_i)$ from a given curve of caplet market prices.⁶ We assume that for any T_i and any strike K we have a market caplet price $V_{\text{caplet}}(T_i, K; T_0)$ given.⁷ Then we may calculate all digital caplet prices $V_{\text{digital}}^{\text{market}}(T_i, K^*; T_0)$. We have

$$V_{\text{digital}}^{\text{market}}(T_i, K; T_0) = \frac{\partial}{\partial K} V_{\text{caplet}}^{\text{market}}(T_i, K; T_0). \quad (23.17)$$

⁶ See also the considerations in Section 5 from Page 89, where we showed how to derive the probability density of the underlying from market prices of European options.

⁷ A complete price curve is usually not available. It has to be constructed by interpolating on given market prices. This interpolation procedure had to be understood as part of the model, see Section 6.

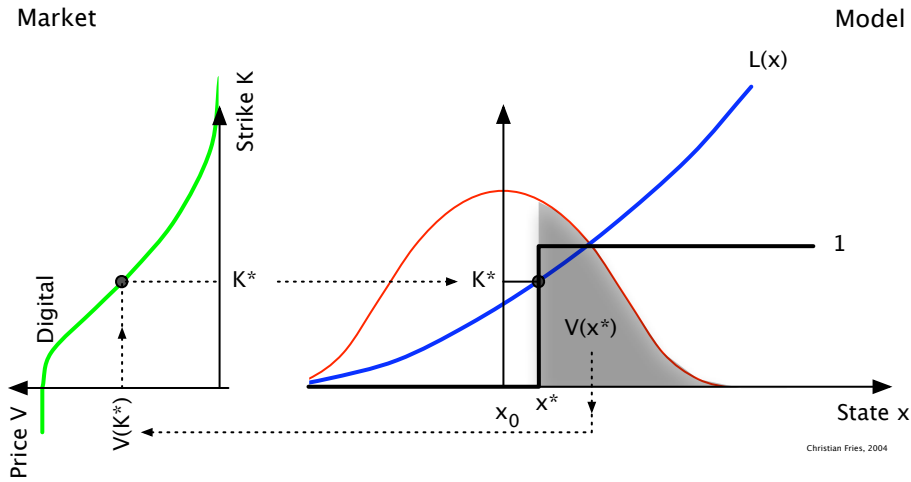


Figure 23.1.: Calibration of the LIBOR functional $L(x)$ within the Markov functional model: For a given x^* we calculate the model price of a digital caplet with strike $K^* = L(x^*)$ (payoff profile in black). This is possible without knowing the functional L . For the given model price $V(x^*)$ (grey surface) we find the corresponding strike K^* by looking it up on the (inverse) of the market price curve (green). This determines the LIBOR functional (blue).

Let x^* denote a given state point. We would like to derive the functional $\xi \mapsto L(T_i, \xi)$, i.e. $K^* := L(T_i, x^*)$. The model price $V_{\text{digital}}^{\text{model}}(T_i, K^*; T_0)$ for a digital caplet with strike K^* is given by

$$\begin{aligned} V_{\text{digital}}(T_0) &= N(T_0) \cdot E^{\mathbb{Q}}(\mathbf{1}(L(T_i, x(T_i)) - K) \cdot (T_{i+1} - T_i) \\ &\quad \cdot \bar{P}(T_{i+1}; T_i, x(T_i)) \mid \{x(T_0) = 0\}), \\ &= N(T_0) \cdot E^{\mathbb{Q}}(\mathbf{1}(x(T_i) - x^*) \cdot (T_{i+1} - T_i) \\ &\quad \cdot \bar{P}(T_{i+1}; T_i, x(T_i)) \mid \{x(T_0) = 0\}), \end{aligned} \quad (23.18)$$

thus, we have from the monotony assumption ($L(x(T_i)) > K \Leftrightarrow x(T_i) > x^*$) that the product may be evaluated without the knowledge of the LIBOR functional $L(T_i)$. We thus write shortly $V_{\text{digital}}^{\text{model}}(T_i, x^*; T_0)$. By solving the equation

$$V_{\text{digital}}^{\text{market}}(T_i, K; T_0) = V_{\text{digital}}^{\text{model}}(T_i, x^*; T_0) \quad (23.19)$$

to K we obtain K^* , i.e. $L(x^*)$, for the given x^* . Compare Figure 23.1.

The LIBOR functional obtained from this procedure replicates the given market price curve of the digital caplets and thus the market price curve of the caplets. Using backward induction we may calibrate the model to caplets of different maturities:

Induction start:

- $N(T_n) = 1$.

Induction step ($T_{i+1} \rightarrow T_i$):

- Calculate $P(T_{i+1}; T_i, \xi) = E^{\mathbb{Q}}(\frac{1}{N(T_{i+1}, x(T_{i+1}))} \mid \{x(T_i) = \xi\})$.
- For any given $\{x_k^*\}$:
 - Calculate the model price $V_{\text{digital}}^{\text{model}}(T_i, x^*; T_0)$ from (23.18).
 - Calculate $K^* = L(x^*)$ from (23.19) and (23.17).
- If required, calculate an interpolation from sample points $x^*, L(x^*)$ obtained in the previous step.
- Calculate $N(T_i)$ from $L(T_i)$ and $P(T_{i+1}; T_i)$ through (23.12).

23.4. Implementation: Lattice

We consider an implementation which relies on functionals discretized in the state space (in contrast to parametrized functionals). In addition to the time discretization $\{0 = T_0 < \dots < T_n\}$ we consider a discretization of the state space

$$x_{i,j} \in x(T_i)(\Omega) = \mathbb{R} \quad j = 1, \dots, m_i, \quad i = 0, \dots, n. \quad (23.20)$$

For any given time discretization point T_i and any given space discretization point $x^* = x_{i,j}$ we calculate the model price of the digital caplet $V_{\text{digital}}^{\text{model}}(T_i, x^*; T_0)$ and calculate from it the value of the LIBOR functional $L(T_i, x_{i,j})$ and from there the numéraire functional $N(T_i, x_{i,j})$. It remains to specify how we interpolate the functionals $\xi \mapsto L(T_i, \xi)$, $\xi \mapsto N(T_i, \xi)$ for $\xi \neq x_{i,j}$ and calculate the conditional expectations, i.e. how we perform the numerical integration

$$E^{\mathbb{Q}}\left(\frac{1}{N(T_{i+1}, x(T_{i+1}))} \mid \{x(T_i) = x_{i,j}\}\right) = \int_{-\infty}^{\infty} \frac{1}{N(T_{i+1}, \xi)} \cdot \phi(\xi - x_{i,j}, \sigma_i) d\xi$$

where $\phi(\xi - \mu, \sigma_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp(-\frac{(\xi - \mu)^2}{2\sigma_i^2})$ denotes the (transition) probability density of $x(T_{i+1}) - x(T_i) \mid x(T_i) = \mu$ with $\sigma_i := (\int_{T_i}^{T_{i+1}} \sigma^2(t) dt)^{1/2}$.

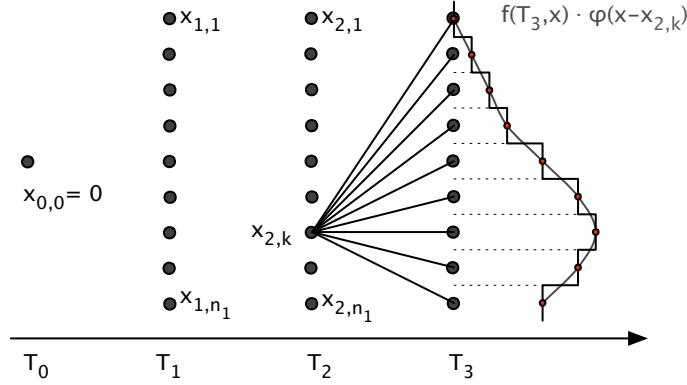


Figure 23.2.: Berechnung der bedingten Erwartung (numerische Integration) im „Lattice“.

23.4.1. Convolution with the Normal Probability Density

A main part of the implementation of the model is the numerical integration, i.e. the convolution of a functional with the density of the normal probability

$$\mathbb{E}^{\mathbb{Q}}\left(f(x(T_{i+1})) \mid \{x(T_i) = x_{i,j}\}\right) = \int_{-\infty}^{\infty} f(\xi) \cdot \phi(\xi - x_{i,j}, \sigma_i) d\xi =: I \quad (23.21)$$

with $\phi(\xi - \mu, \sigma_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp(-\frac{(\xi-\mu)^2}{2\sigma_i^2})$. In the literature one finds many methods for numerical integration. Since f is given only at the state space sample points (lattice) $\{x_{i+1,k} \mid j = 1, \dots, m_i\}$, but the density ϕ is given analytically, it is natural to make use of this.

23.4.1.1. Piecewise constant Approximation

Let $x_{i+1,k+\frac{1}{2}} = \frac{x_{i+1,k+1} + x_{i+1,k}}{2}$ denote the center point of the intervall $[x_{i+1,k}, x_{i+1,k+1}]$ and let $f_k := f(x_{i+1,k})$. Then an approximation of the integral I from (23.21) is given by

$$I \approx \sum_{k=1}^{k=n} f_k \cdot (\Phi_{x_{i,j}, \sigma_i}(x_{i+1,k+\frac{1}{2}}) - \Phi_{x_{i,j}, \sigma_i}(x_{i+1,k-\frac{1}{2}})),$$

where $\Phi_{\mu, \sigma}(x) = \int_{-\infty}^x \phi(\xi - \mu, \sigma) d\xi$. For there cumulative normal distribution function $\Phi_{\mu, \sigma}$ there exist very accurate approximations through rational functions.

23.4.1.2. Piecewise polynomial Approximation

If we approximate the integrand f by a piecewise polynomial, then we have to calculate integrals of the form

$$\int_{h_1}^{h_2} \xi^m \cdot \phi(\xi - x_{i,j}, \sigma_i) d\xi.$$

These may be given recursively in terms of the cumulative distribution function Φ . We have

Lemma 211 (Convolution of a monomial with the normal distribution function⁸):

$$\begin{aligned} \int_{h_1}^{h_2} \xi^m \cdot \phi(\xi - \mu, \sigma) d\xi &= \mu \cdot \int_{h_1}^{h_2} \xi^{m-1} \cdot \phi(\xi - \mu, \sigma) d\xi \\ &+ (m-1) \cdot \int_{h_1}^{h_2} \xi^{m-2} \cdot \phi(\xi - \mu, \sigma) d\xi \\ &- \sigma^2 \left[\xi^{m-1} \cdot \phi(\xi - \mu, \sigma) \right]_{h_1}^{h_2} \end{aligned}$$

Proof: For $m \in \mathbb{N}$ we have

$$\begin{aligned} \frac{\partial}{\partial \xi} \xi^{m-1} \cdot \exp\left(-\frac{1}{2}\left(\frac{\xi-\mu}{\sigma}\right)^2\right) \\ &= (m-1) \cdot \xi^{m-2} \cdot \exp\left(-\frac{1}{2}\left(\frac{\xi-\mu}{\sigma}\right)^2\right) + \xi^{m-1} \exp\left(-\frac{1}{2}\left(\frac{\xi-\mu}{\sigma}\right)^2\right) \frac{-(\xi-\mu)}{\sigma^2} \\ &= (m-1) \cdot \xi^{m-2} \cdot \exp\left(-\frac{1}{2}\left(\frac{\xi-\mu}{\sigma}\right)^2\right) - \frac{\xi^m}{\sigma^2} \exp\left(-\frac{1}{2}\left(\frac{\xi-\mu}{\sigma}\right)^2\right) \\ &\quad + \frac{\mu}{\sigma^2} \xi^{m-1} \exp\left(-\frac{1}{2}\left(\frac{\xi-\mu}{\sigma}\right)^2\right), \end{aligned}$$

i.e.

$$\begin{aligned} \xi^m \exp\left(-\frac{1}{2}\left(\frac{\xi-\mu}{\sigma}\right)^2\right) \\ &= (m-1)\sigma^2 \cdot \xi^{m-2} \cdot \exp\left(-\frac{1}{2}\left(\frac{\xi-\mu}{\sigma}\right)^2\right) + \mu \xi^{m-1} \exp\left(-\frac{1}{2}\left(\frac{\xi-\mu}{\sigma}\right)^2\right) \\ &\quad - \frac{\partial}{\partial \xi} \sigma^2 \xi^{m-1} \cdot \exp\left(-\frac{1}{2}\left(\frac{\xi-\mu}{\sigma}\right)^2\right). \end{aligned}$$

The claim follows after multiplication with $\frac{1}{\sqrt{2\pi}\sigma}$ and integration $\int_{h_1}^{h_2}$. □

To calculate the interpolation polynomials many different methods may be used (e.g. *Cubic Spline Interpolation*). The Neville algorithm gives a simple recursion by which we may calculate a piecewise polynomial interpolation functions:

⁸ Compare [24].

Lemma 212 (Neville Algorithm - Piecewise polynomial interpolation function):

Let x_i denote given sample points and f_i the corresponding values. If $F_{i-k_1, i+k_2}$ is a polynomial of degree $k_1 + k_2$ interpolating the points (x_j, f_j) , $j = i - k_1, \dots, i + k_2$, i.e.

$$F_{i-k_1, i+k_2}(x_j) = f_j \quad \text{für } j = i - k_1, \dots, i + k_2,$$

then

$$F_{i-k_1, i+k_2+1}(\xi) = \frac{(\xi - x_{i-k_1}) \cdot F_{i-k_1+1, i+k_2+1}(\xi) + (x_{i+k_2+1} - \xi) \cdot F_{i-k_1, i+k_2}}{x_{i+k_2+1} - x_{i-k_1}}$$

defines an interpolation polynomial of degree $k_1 + k_2 + 1$ for the points (x_j, f_j) , $j = i - k_1, \dots, i + k_2 + 1$. With the trivial interpolation polynomial $F_{i,i} \equiv f_i$ having degree 0, induction gives a construction of the desired interpolation polynomial.

Proof: A polynomial of degree $k_1 + k_2 + 1$ is uniquely determined by $k_1 + k_2 + 2$ points. The claim follows then by inductions, evaluating the approximations polynomial $F_{i-k_1, i+k_2+1}$ at the sample points x_{i+k_2+1} , x_{i-k_1} and the sample points which are common to $F_{i-k_1+1, i+k_2+1}$ and $F_{i-k_1, i+k_2}$. $\square$

The algorithm is called *Neville algorithm*. The polynomial $F_{i-k_1, i+k_2}(x_j)$ will then be used as an interpolation polynomial on the interval x_i, x_{i+1} .⁹ Using Lemma 211 the corresponding integral part is calculated.

The interpolation function is (for degree ≥ 0) continuous, but it not differentiable at the interval bounds, as it would be the case with a cubic spline interpolation.

23.4.2. State space discretization

The rule to choose the state space discretization points $x_{i,j}$, $j = 1, \dots, m_i$ (the lattice) for the realization of the Markovian driver $x(T_i)$ has a great impact on the accuracy of the model.

23.4.2.1. Equidistant discretization

A simple rule gives an an equidistant discretization of the interval $[x_{i,min}, x_{i,max}]$, where the interval has been chosen such that the neglected area $(-\infty, x_{i,min}) \cup (x_{i,max}, \infty)$ supports only a small probability measure $\mathbb{Q}(x(T_i) \notin [x_{i,min}, x_{i,max}]) = \epsilon$, see Figure 23.3. The bounds $x_{i,min}$, $x_{i,max}$ may be derived from the definition of the process x , e.g. by in terms of the standard deviation of $x(T_i)$:

$$x_{i,max} = -x_{i,min} = k \cdot \bar{\sigma}_i, \quad \bar{\sigma}_i := \left(\int_0^{T_i} \sigma^2(t) dt \right)^{1/2}.$$

The factor k has to be chosen sufficiently large (e.g. $k > 3$).

⁹ It is natural to choose the sample point symmetrically, i.e. use $k_2 = k_1 + 1$.

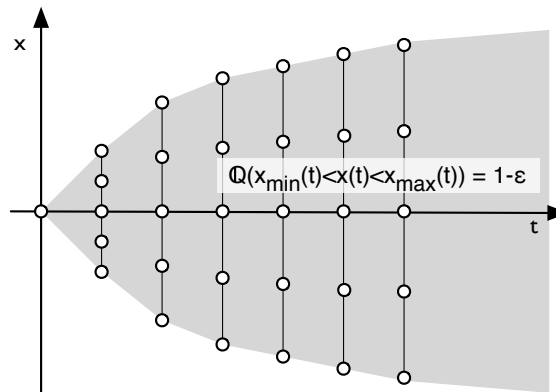


Figure 23.3.: State space discretization for the Markov functional model: A simple rule to choose the discretization points $x_{i,j}$.

$$\mathbb{Q}(x_{i,j} \leq x(T_i) \leq x_{i,j+1}) = \frac{1 - \epsilon}{m_i - 1}.$$



Further Reading: The LIBOR and Swaprate Markov Functional Model in terminal measure is discussed in the books [12, 24] as well as in the original article [70]. A LIBOR Markov functional model in spot measure and a hybrid Markov functional model is discussed in [62]. <|

Part VI.

Extended Models

CHAPTER 24

Hybrid Models

In this chapter we introduce several kinds of hybrid models. A hybrid model is a model that models multiple (different) assets in a single unified model. In general one combines several well known models to a single unified model. Since different models usually come under different pricing measures, the essence of a hybrid model is the question “How do these models look under a common pricing measure?”. So apart from the prerequisite that the models are compatible, the construction of a hybrid model is just a change of measure.

In this sense, we already encountered the basic technique required for a hybrid model in the discussion of the [LIBOR Market Model](#): There we wrote down multiple individual [Black models](#) (Chapter 10) and asked ourselves how they look under a common pricing measure. The result was the [LIBOR Market Model](#).

24.1. Cross Currency LIBOR Market Model

For two currencies, “domestic” and “foreign”, we model the interest rate curves, each with a [LIBOR Market Model](#), as it had been discussed for an interest rate curve in a single currency in Section 17.1.

In addition we model the foreign exchange rate $FX(t)$. In Chapter 11 we presented the pricing of a [quanto caplets](#) and therefor modeled that FX forward as a lognormal process. Here the spot exchange rate $FX(t)$ is modeled directly, also as a lognormal process. Let $FX(t)$ denote the amount (in domestic currency) that has to be paid by an *domestic* investor at time t for one unit of foreign currency (for). Thus, $FX(t)$ has the (physical) unit $[FX(t)] = \frac{1 \text{ dom}}{1 \text{ for}}$.

We assume that for the chosen numéraire N that there exists a corresponding equivalent martingale measure $\mathbb{Q}^N$. By the [change of measure theorem](#) (Girsanov theorem,

53) the modelled quantities are again log-normal processes under $\mathbb{Q}^N$, i.e. we have

$$\begin{aligned} dL_i(t) &= L_i(t) \cdot \mu_i(t)dt + L_i(t)\sigma_i(t)dW_i^{\mathbb{Q}^N}(t) \quad (0 \leq i \leq n-1) \\ dFX(t) &= FX(t) \cdot \mu^{FX}(t)dt + FX(t)\sigma^{FX}(t)dW_{FX}^{\mathbb{Q}^N}(t) \\ d\tilde{L}_i(t) &= \tilde{L}_i(t) \cdot \tilde{\mu}_i(t)dt + \tilde{L}_i(t)\tilde{\sigma}_i(t)d\tilde{W}_i^{\mathbb{Q}^N}(t) \quad (0 \leq i \leq n-1) \end{aligned}$$

with initial conditions

$$L_i(0) = L_{i,0}, \quad FX(0) = FX_0, \quad \tilde{L}_i(0) = \tilde{L}_{i,0}.$$

As before, this is the starting point for

- Determination of the drift terms μ_i , μ^{FX} and $\tilde{\mu}_i$ for a chosen numéraire $N(t)$ using the $\mathbb{Q}^N$ -martingale property of N -relative prices.
- Determination of the initial conditions $L_{i,0}$, FX_0 , $\tilde{L}_{i,0}$ using the bond and foreign-bond prices observed at time $t = 0$.
- Determination/choice of the volatility and correlation to reproduce given option prices.

24.1.1. Derivation of the Drift Term under Spot-Measure

As numéraire we chose the *rolled over one period bond* we had already made use of in Section 17.1.1.2

$$N(t) := P(T_{m(t)+1}; t) \prod_{j=0}^{m(t)} (1 + L_j(T_j) \cdot \delta_j).$$

where $m(t) := \max\{i : T_i \leq t\}$, $\delta_j := T_{j+1} - T_j$. We will now derive the corresponding processes (i.e. the drifts) under the corresponding equivalent martingale measure $\mathbb{Q}^N$, the *spot measure*.

24.1.1.1. Dynamic of the *domestic* LIBOR under Spot Measure

For the drift μ_i we have, exactly as in Section 17.1.1.2

$$\mu_i(t) = \sum_{l=m(t)+1}^i \frac{\delta_l L_l(t)}{(1 + \delta_l L_l(t))} \sigma_i(t) \sigma_l(t) \rho_{i,l}(t).$$

This is the already known LIBOR Market Model in *domestic* currency.

24.1.1.2. Dynamic of the *foreign* LIBOR under Spot Measure

We derive the drift $\tilde{\mu}_i(t)$ of the foreign LIBOR by considering financial product from the foreign market. The foreign bond $\tilde{P}(T_{i+1})$, converted to domestic currency, i.e. $\tilde{P}(T_{i+1}) \cdot FX$ is a traded asset for the domestic investor. Thus, we have that $\frac{\tilde{P}(T_{i+1};t) \cdot FX(t)}{N(t)}$ is a $\mathbb{Q}^N$ -martingale.¹

$$\text{Drift}_{\mathbb{Q}^N} \left(\frac{\tilde{P}(T_{i+1};t) \cdot FX(t)}{N(t)} \right) = 0. \quad (24.1)$$

Likewise we have that

$$\tilde{L}_i \cdot \tilde{P}(T_{i+1}) \cdot FX = \frac{\tilde{P}(T_i) - \tilde{P}(T_{i+1})}{T_{i+1} - T_i} \cdot FX \quad (24.2)$$

is a traded asset for the domestic investor, because it is a portfolio of (foreign) bonds converted to domestic currency. Thus,

$$\text{Drift}_{\mathbb{Q}^N} \left(\tilde{L}_i \cdot \frac{\tilde{P}(T_{i+1};t) \cdot FX(t)}{N(t)} \right) = 0.$$

From the [product rule](#) we find

$$d(\tilde{L}_i(t) \cdot \frac{\tilde{P}(T_{i+1};t) \cdot FX(t)}{N(t)}) \quad (24.3)$$

$$= \frac{\tilde{P}(T_{i+1};t) \cdot FX(t)}{N(t)} \cdot d\tilde{L}_i + \tilde{L}_i \cdot d\left(\frac{\tilde{P}(T_{i+1};t) \cdot FX(t)}{N(t)}\right) + d\tilde{L}_i \cdot d\left(\frac{\tilde{P}(T_{i+1};t) \cdot FX(t)}{N(t)}\right) \quad (24.4)$$

from which we derive $\tilde{\mu}_i$ after calculation of $d\frac{\tilde{P}(T_{i+1};t) \cdot FX(t)}{N(t)}$ by comparing drift terms.

We have

$$d\left(\frac{\tilde{P}(T_{i+1};t) \cdot FX(t)}{N(t)}\right) = d\left(\frac{\tilde{P}(T_{m(t)+1};t)}{P(T_{m(t)+1};t)} \frac{\prod_{j=m(t)+1}^i (1 + \tilde{L}_j(t)\delta_j)^{-1}}{\prod_{j=0}^{m(t)} (1 + L_j(T_j)\delta_j)} \cdot FX(t)\right).$$

Remark 213 (Interpolation der Bondpreise): At this point we encounter an interesting difference to the single currency LIBOR Market Modell (see Section 17.1): While for the case of the domestic currency the corresponding term $\frac{P(T_{m(t)+1};t)}{P(T_{m(t)+1};t)}$ vanishes, it is required to give an interpretation to $\frac{\tilde{P}(T_{m(t)+1};t)}{P(T_{m(t)+1};t)}$. See also Section 17.1.2.3.

¹ In the view of the domestic investor the foreign bond is not a traded asset. Only after conversion to the domestic currency by the applicable conversion rate $FX(t)$ the product becomes tradable for the domestic investor. This is also appeared from the fact that relative prices are dimensionless: While $\frac{\tilde{P}(T_{i+1};t) \cdot FX(t)}{N(t)}$ is dimensionless we find that $\frac{\tilde{P}(T_{i+1};t)}{N(t)}$ has the (physical) unit $\frac{1}{1 \text{ dom}}$.

We have not yet defined the value of the short period bonds $P(T_{i+1}; t)$ and $\tilde{P}(T_{i+1}; t)$ for $t \neq T_j$, $j = 0, 1, 2, \dots$. We do this now and define for $T_j < t \leq T_{j+1}$

$$\begin{aligned} P(T_{m(t)+1}; t) &:= (1 + L_{m(t)}(T_{m(t)}) \cdot (T_{m(t)+1} - t))^{-1}, \\ \tilde{P}(T_{m(t)+1}; t) &:= (1 + \tilde{L}_{m(t)}(T_{m(t)}) \cdot (T_{m(t)+1} - t))^{-1}. \end{aligned}$$

This concludes the definition of the numéraire.

From this we find

$$d \frac{\tilde{P}(T_{m(t)+1}; t)}{P(T_{m(t)+1}; t)} = \frac{\partial}{\partial t} \left(\frac{1 + L_{m(t)}(T_{m(t)}) \cdot (T_{m(t)+1} - t)}{1 + \tilde{L}_{m(t)}(T_{m(t)}) \cdot (T_{m(t)+1} - t)} \right) dt,$$

i.e. the term $\frac{\tilde{P}(T_{i+1}; t)}{P(T_{i+1}; t)}$ has no diffusion part dW for $T_j < t \leq T_{j+1}$. This is sufficient for the following derivation, the specific form of the drift does not need to be known.

Indeed, it would be sufficient to require that $P(T_{i+1}; t)$ and $\tilde{P}(T_{i+1}; t)$ have zero volatility, i.e. no diffusion part, in the short period $t \in (T_i, T_{i+1}]$. A specific interpolation is not required. The zero volatility assumption for $P(T_{i+1}; t)$ and $\tilde{P}(T_{i+1}; t)$ on $t \in (T_i, T_{i+1}]$ closes the definition of the LIBOR Market Model for all $t \in \{T_0, T_1, \dots\}$.²

Continuing in the derivation of the drift, we have

$$\begin{aligned} & d \left(\frac{\tilde{P}(T_{i+1}; t) \cdot FX(t)}{N(t)} \right) \\ &= d \left(\frac{\tilde{P}(T_{i+1}; t)}{P(T_{i+1}; t)} \frac{\prod_{j=0}^i (1 + \tilde{L}_j(t) \delta_j)^{-1}}{\prod_{j=0}^{-1} (1 + L_j(T_j) \delta_j)} \cdot FX(t) \right) \\ &= (\dots) dt - \frac{\tilde{P}(T_{i+1}; t) \cdot FX(t)}{N(t)} \cdot \sum_{j=m(t)+1}^i \frac{\delta_j \tilde{L}_j}{(1 + \delta_j \tilde{L}_j)} \tilde{\sigma}_j d\tilde{W}_j^{\mathbb{Q}^N}(t) \\ &\quad + \frac{\tilde{P}(T_{i+1}; t) \cdot FX(t)}{N(t)} \cdot \sigma^{FX}(t) dW_{FX}^{\mathbb{Q}^N}(t) \end{aligned}$$

If we plug this into (24.4) and compare drift terms (the coefficients of dt) we get together with (24.1) and (24.2)

$$\begin{aligned} 0 &= \frac{\tilde{P}(T_{i+1}; t) \cdot FX(t)}{N(t)} \cdot \tilde{L}_i(t) \cdot \tilde{\mu}_i(t) + 0 \\ &\quad + \frac{\tilde{P}(T_{i+1}; t) \cdot FX(t)}{N(t)} \cdot \tilde{L}_i(t) \left(- \sum_{j=m(t)+1}^i \frac{\delta_j \tilde{L}_j}{(1 + \delta_j \tilde{L}_j)} \tilde{\sigma}_i \tilde{\sigma}_j d\tilde{W}_i^{\mathbb{Q}^N}(t) d\tilde{W}_j^{\mathbb{Q}^N}(t) \right. \\ &\quad \left. + \tilde{\sigma}_i \sigma^{FX}(t) d\tilde{W}_j^{\mathbb{Q}^N}(t) dW_{FX}^{\mathbb{Q}^N}(t) \right). \end{aligned}$$

² See [22].

Denoting the interest rate correlation within the foreign currency by $\tilde{\rho}_{i,j}$, i.e.

$$d\tilde{W}_i^{\mathbb{Q}^N}(t)d\tilde{W}_j^{\mathbb{Q}^N}(t) = \tilde{\rho}_{i,j}(t)dt$$

and denoting the correlation of the foreign currency and the foreign exchange rate by $\rho_{i,FX}$, i.e.

$$d\tilde{W}_i^{\mathbb{Q}^N}(t)dW_{FX}^{\mathbb{Q}^N}(t) = \rho_{i,FX}(t)dt,$$

gives

$$\tilde{\mu}_i(t) = \sum_{j=m(t)+1}^i \frac{\delta_j \tilde{L}_j(t)}{(1 + \delta_j \tilde{L}_j(t))} \tilde{\sigma}_i(t) \tilde{\sigma}_j(t) \tilde{\rho}_{i,j}(t) - \tilde{\sigma}_i(t) \sigma^{FX}(t) \rho_{i,FX}.$$

24.1.1.3. Dynamic of the FX Rate under Spot Measure

For given $t \in [T_{m(t)}, T_{m(t)+1})$ consider the foreign bond with maturity $T_{m(t)+1}$ – the next possible maturity in our tenor discretization. Converting it to domestic currency its N -relative value $\frac{FX(t) \cdot \tilde{P}(T_{m(t)+1}; t)}{N(t)}$ is a martingale. Furthermore we have

$$\begin{aligned} & d\left(FX(t) \cdot \frac{\tilde{P}(T_{m(t)+1}; t)}{N(t)}\right) \\ \stackrel{\text{Def. } N}{=} & \prod_{j=0}^{m(t)} (1 + L_j(T_j) \delta_j) \cdot d\left(FX(t) \cdot \frac{\tilde{P}(T_{m(t)+1}; t)}{P(T_{m(t)+1}; t)}\right) \\ = & dFX(t) \cdot \frac{\tilde{P}(T_{m(t)+1}; t)}{N(t)} \\ & + FX(t) \prod_{j=0}^{m(t)} (1 + L_j(T_j) \delta_j) \frac{\partial}{\partial t} \left(\frac{\tilde{P}(T_{m(t)+1}; t)}{P(T_{m(t)+1}; t)} \right) dt \\ \stackrel{\text{Def. } N}{=} & dFX(t) \cdot \frac{\tilde{P}(T_{m(t)+1}; t)}{N(t)} + FX(t) \cdot \frac{\tilde{P}(T_{m(t)+1}; t)}{N(t)} \frac{\partial}{\partial t} \left(\frac{\tilde{P}(T_{m(t)+1}; t)}{P(T_{m(t)+1}; t)} \right) dt \end{aligned}$$

and thus

$$\begin{aligned} \mu^{FX}(t) &= - \frac{\frac{\partial}{\partial t} \left(\frac{\tilde{P}(T_{m(t)+1}; t)}{P(T_{m(t)+1}; t)} \right)}{\frac{\tilde{P}(T_{m(t)+1}; t)}{P(T_{m(t)+1}; t)}} = - \frac{\partial}{\partial t} \log \left(\frac{\tilde{P}(T_{m(t)+1}; t)}{P(T_{m(t)+1}; t)} \right) \\ &= \frac{\partial}{\partial t} \log \left(\frac{P(T_{m(t)+1}; t)}{\tilde{P}(T_{m(t)+1}; t)} \right). \end{aligned}$$

The drift $\mu^{FX}(t)$ thus depends on the chosen interpolation of the bond prices, see Remark 213. However, for its use in an implementation via a time-discretization scheme

it is not necessary to calculate the corresponding derivative after t , since the only the integrated drift enters the time-discretization scheme. For the integral $\int_{T_i}^{T_{i+1}} \mu^{FX}(t) dt$ we have

$$\begin{aligned} \int_{T_i}^{T_{i+1}} \mu^{FX}(t) dt &= \log \left(\frac{P(T_{i+1}; T_{i+1})}{\tilde{P}(T_{i+1}; T_{i+1})} \right) - \log \left(\frac{P(T_{i+1}; T_i)}{\tilde{P}(T_{i+1}; T_i)} \right) = -\log \left(\frac{P(T_{i+1}; T_i)}{\tilde{P}(T_{i+1}; T_i)} \right) \\ &= \log \left(\frac{1 + \tilde{L}(T_i, T_{i+1}; T_i)}{1 + L(T_i, T_{i+1}; T_i)} \right), \end{aligned}$$

i.e. it is independent of the interpolation of the bond prices. See also Section 24.3.2.

24.1.2. Implementation

We will discuss the implementation together with the *Equity Hybrid LIBOR Market Model* in Section 24.3.2.

24.2. Equity Hybrid LIBOR Market Model

In Section 4 we introduced the Black-Scholes Modell for a (single) stock. The stock S was modeled as a lognormal process

$$dS(t) = \mu^{S, \mathbb{P}}(t) S(t) dt + \sigma^S(t) S(t) dW^{S, \mathbb{P}}(t) \quad \text{im realen Maß } \mathbb{P}.$$

We had assumed interest rates as non-stochastic and constant and the chosen numéraire then was $B(t) := \exp(r \cdot t)$. Under the corresponding martingale measure $\mathbb{Q}^B$ we could then derive an analytic formula for the price of a European stock option.

The pricing of products exhibiting optionalities related to both, stock and interest rates, requires a joint stochastic modelling of the stock and the interest rates. If we chose as the interest rate model the *LIBOR Market Model* and as the stock process model the *Black-Scholes Modell*, then the construction of the joint model is just the derivation of the drifts under a common measure.

24.2.1. Derivation of the Drift Term under Spot-Measure

As before, we choose as numéraire the *rolled over one period bond*

$$N(t) := P(T_{m(t)+1}; t) \prod_{j=0}^{m(t)} (1 + L_j(T_j) \cdot \delta_j),$$

where $m(t) := \max\{i : T_i \leq t\}$ and $\delta_j := T_{j+1} - T_j$.

24.2.1.1. Dynamic of the Stock Process under Spot Measure

The stock S is a traded asset, thus $\frac{S}{N}$ is a martingale under the equivalent martingale measure $\mathbb{Q}^N$. From [Produktregel](#) we have

$$d\left(\frac{S}{N}\right) = dS \cdot \frac{1}{N} + S d\left(\frac{1}{N}\right) + dS d\left(\frac{1}{N}\right).$$

For $T_i < t < T_{i+1}$ let the bond price $P(T_{i+1}; t)$ defined (interpolated) as in [Remark 213](#), i.e.

$$P(T_{i+1}; t) := (1 + L_i(T_i) \cdot (T_{i+1} - t))^{-1} \quad \text{für } T_i < t < T_{i+1}.$$

In this case we have

$$\begin{aligned} d\left(\frac{1}{N}\right) &= \left(\frac{1}{P(T_{m(t)+1}; t)} \cdot \prod_{j=0}^{m(t)} (1 + L_j(T_j) \delta_j)\right) \\ &= d\left(\frac{1}{P(T_{m(t)+1}; t)} \cdot \prod_{j=0}^{m(t)} (1 + L_j(T_j) \delta_j)\right) \\ &= d\left((1 + L_{m(t)}(T_{m(t)}) \cdot (T_{m(t)+1} - t)) \cdot \prod_{j=0}^{m(t)} (1 + L_j(T_j) \delta_j)\right) \\ &= -L_{m(t)}(T_{m(t)}) dt \cdot \prod_{j=0}^{m(t)} (1 + L_j(T_j) \delta_j) \\ &= -\frac{L_{m(t)}(T_{m(t)})}{1 + L_{m(t)}(T_{m(t)}) \cdot (T_{m(t)+1} - t)} dt \cdot \frac{1}{N} \\ &= \frac{\partial}{\partial t} \log(1 + L_{m(t)}(T_{m(t)}) \cdot (T_{m(t)+1} - t)) dt \frac{1}{N}. \end{aligned}$$

Remark 214 (on the numéraire process $\frac{1}{N}$): Under the assumption that $N(t)$ does not have a diffusion (i.e. dW) term, which is the case for the above definition of the short period bond, then we have $d\left(\frac{1}{N}\right) = \left(\frac{\partial}{\partial t} \frac{1}{N}\right) dt$. Nothing else has been calculated above. See also [remark 213](#).

Under the assumption that N does not have a diffusion (i.e. dW) term, we have $dS d\left(\frac{1}{N}\right) = 0$ and from $\text{Drift}_{\mathbb{Q}^N}\left(\frac{S}{N}\right) = 0$ we find

$$0 = \mu^S \frac{S}{N} + \frac{\partial}{\partial t} \log(1 + L_{m(t)}(T_{m(t)}) \cdot (T_{m(t)+1} - t)) + 0,$$

thus

$$\begin{aligned} \mu^S(t) &= \frac{\partial}{\partial t} \log(1 + L_{m(t)}(T_{m(t)}) \cdot (T_{m(t)+1} - t)) \\ &= -\frac{L_{m(t)}(T_{m(t)})}{1 + L_{m(t)}(T_{m(t)}) \cdot (T_{m(t)+1} - t)} \end{aligned} \tag{24.5}$$



Interpretation Comparison to the Dynamic of the Black-Scholes Model: Under the numéraire $B(t) := \exp(r \cdot t)$ used in the Black-Scholes model (see Section 4 we had derived the drift of the stock process under the martingale measure $\mathbb{Q}^B$ as $\mu^{S, \mathbb{Q}^B} = r$. Equation (24.5) is just a discrete (and stochastic) version of this drift: For the average drift over one period $[T_i, T_{i+1}]$ we have

$$\begin{aligned} \frac{1}{T_{i+1} - T_i} \int_{T_i}^{T_{i+1}} \mu^{S, \mathbb{Q}^N}(t) dt &= \log(0) - \log(1 + L_i(T_i)) \cdot (T_{i+1} - T_i) \\ &= - \frac{\log(1 + L_i(T_i)) \cdot (T_{i+1} - T_i)}{T_{i+1} - T_i} \end{aligned}$$

and for infinitesimal period lengths $T_{i+1} \rightarrow T_i$ we find – see Definition 92 –

$$\frac{-\log(1 + L_i(T_i)) \cdot (T_{i+1} - T_i)}{T_{i+1} - T_i} \xrightarrow{T_{i+1} \rightarrow T_i} r(T_i) = \mu^{S, \mathbb{Q}^B},$$

i.e.

$$\mu^{S, \mathbb{Q}^N}(t) \xrightarrow{T_{i+1} \rightarrow T_i} \mu^{S, \mathbb{Q}^B}.$$

<|

24.2.2. Implementation

We discuss the implementation together with the previously discussed *Cross Currency LIBOR Market Model* in Section 24.3.2.

24.3. Equity-Hybrid Cross-Currency LIBOR Market Model

The models given in Section 24.1 and 24.2 may be combined. This is now trivial since the numéraire and thus the martingale measure is the same in both models. Thus we have a unified model for interest rates, foreign exchange rates, foreign interest rates and equity. We will now add the model of a foreign stock. Let $\tilde{S}$ denote another stock process, modeling a stock from the foreign market, i.e. the process $\tilde{S}$ has the dimension (currency unit) 1 for.

24.3.0.1. Dynamic of the Foreign Stock under Spot-Measure

We assume that the foreign stock $\tilde{S}$ follows a lognormal process

$$d\tilde{S}(t) = \mu^{\tilde{S}, \mathbb{P}}(t) \tilde{S}(t) dt + \sigma^{\tilde{S}}(t) \tilde{S}(t) dW^{\tilde{S}, \mathbb{P}}(t) \quad \text{im realen Maß } \mathbb{P}.$$

As in the previous Sections 24.1 and 24.2, we chose as numéraire $N(t) := P(T_{m(t)+1}; t) \prod_{j=0}^{m(t)} (1 + L_j(T_j) \cdot \delta_j)$. As for the **Cross-Currency LIBOR Market Model** we have that the foreign stock $\tilde{S}$ has to be converted to domestic currency to be a traded asset for the domestic investor. I.e. $FX \cdot \tilde{S}$ is a traded asset and the N -relative price $\frac{FX \cdot \tilde{S}}{N}$ ein $\mathbb{Q}^N$ -martingale. From the **product rule** we have

$$d\left(\frac{FX \cdot \tilde{S}}{N}\right) = d(FX \cdot \tilde{S}) \cdot \frac{1}{N} + FX \cdot \tilde{S} \cdot d\left(\frac{1}{N}\right) + d(FX \cdot \tilde{S}) \cdot d\left(\frac{1}{N}\right)$$

and

$$d(FX \cdot \tilde{S}) = dFX \cdot \tilde{S} + FX \cdot d\tilde{S} + dFX \cdot d\tilde{S}.$$

With the definition of the numéraire from Remark 213 und 214 we have

$$d\left(\frac{1}{N}\right) = \left(\frac{\partial}{\partial t} \frac{1}{N}\right)dt = \left(\frac{1}{N} \cdot \frac{\partial}{\partial t} \log\left(\frac{1}{N}\right)\right)dt$$

and thus

$$\begin{aligned} \text{Drift}_{\mathbb{Q}^N}\left(\frac{FX \cdot \tilde{S}}{N}\right) \\ = \frac{FX \cdot \tilde{S}}{N} \cdot (\mu^{FX} + \mu^{\tilde{S}} + \sigma^{FX}(t)dW_{FX}^{\mathbb{Q}^N}(t) \cdot \sigma^{\tilde{S}}(t)dW_{\tilde{S}}^{\mathbb{Q}^N}(t) + \frac{\partial}{\partial t} \log\left(\frac{1}{N}\right)) \end{aligned}$$

If we denote the instantaneous correlation of the stock process $\tilde{S}$ and the foreign exchange rate process FX by $\rho_{FX, \tilde{S}}$, i.e. we have $dW_{FX}^{\mathbb{Q}^N}(t) \cdot dW_{\tilde{S}}^{\mathbb{Q}^N}(t) = \rho_{FX, \tilde{S}} dt$, then we get together with $\text{Drift}_{\mathbb{Q}^N}\left(\frac{FX \cdot \tilde{S}}{N}\right) = 0$

$$\mu^{\tilde{S}} = -\mu^{FX} - \frac{\partial}{\partial t} \log\left(\frac{1}{N}\right) - \sigma^{FX}(t)\sigma^{\tilde{S}}(t)\rho_{FX, \tilde{S}}(t).$$

With

$$\mu^{FX}(t) = \frac{\partial}{\partial t} \log\left(\frac{P(T_{m(t)+1}; t)}{\tilde{P}(T_{m(t)+1}; t)}\right)$$

and

$$\frac{\partial}{\partial t} \log\left(\frac{1}{N}\right) = \frac{\partial}{\partial t} \log\left(\frac{1}{P(T_{m(t)+1}; t)}\right)$$

we get

$$\begin{aligned} \mu^{\tilde{S}} &= \frac{\partial}{\partial t} \log(\tilde{P}(T_{m(t)+1}; t)) - \sigma^{FX}(t)\sigma^{\tilde{S}}(t)\rho_{FX, \tilde{S}}(t) \\ &= -\frac{\partial}{\partial t} \log\left(1 + \tilde{L}_{m(t)}(T_{m(t)}) \cdot (T_{m(t)+1} - t)\right) - \sigma^{FX}(t)\sigma^{\tilde{S}}(t)\rho_{FX, \tilde{S}}(t). \end{aligned}$$

24.3.1. Summary

Under the numéraire

$$N(t) := P(T_{m(t)+1}; t) \prod_{j=0}^{m(t)} (1 + L_j(T_j) \cdot \delta_j),$$

where $m(t) := \max\{i : T_i \leq t\}$, and the assumption that $N(t)$ does not have a diffusion part, as it would be the case for

$$\left. \begin{aligned} P(T_{i+1}; t) &:= (1 + L_i(T_i) \cdot (T_{i+1} - t))^{-1} \\ \tilde{P}(T_{i+1}; t) &:= (1 + \tilde{L}_i(T_i) \cdot (T_{i+1} - t))^{-1} \end{aligned} \right\} \quad \text{für } T_i < t < T_{i+1}, \quad (24.6)$$

the dynamic of the *Equity-Hybrid Cross-Currency LIBOR Market Model* under the corresponding martingale measure $\mathbb{Q}^N$ (spot measure) is given by

$$\begin{aligned} dL_i(t) &= L_i(t) \cdot \mu_i(t) dt + L_i(t) \sigma_i(t) dW_i^{\mathbb{Q}^N}(t) \quad \text{für } i = 0, \dots, n-1, \\ dFX(t) &= FX(t) \cdot \mu^{FX}(t) dt + FX(t) \sigma^{FX}(t) dW_{FX}^{\mathbb{Q}^N}(t) \\ d\tilde{L}_i(t) &= \tilde{L}_i(t) \cdot \tilde{\mu}_i(t) dt + \tilde{L}_i(t) \tilde{\sigma}_i(t) d\tilde{W}_i^{\mathbb{Q}^N}(t) \quad \text{für } i = 0, \dots, n-1, \\ dS(t) &= S(t) \mu^S(t) dt + S(t) \sigma^S(t) dW_S^{\mathbb{Q}^N}(t) \\ d\tilde{S}(t) &= \tilde{S}(t) \mu^{\tilde{S}}(t) dt + \tilde{S}(t) \sigma^{\tilde{S}}(t) dW_{\tilde{S}}^{\mathbb{Q}^N}(t) \end{aligned}$$

where

$$\begin{aligned} \mu_i &= \sum_{l=m(t)+1}^i \frac{\delta_l L_l}{(1 + \delta_l L_l)} \sigma_i \sigma_l \rho_{i,l}, \\ \tilde{\mu}_i(t) &= \sum_{j=m(t)+1}^i \frac{\delta_j \tilde{L}_j}{(1 + \delta_j \tilde{L}_j)} \tilde{\sigma}_i \tilde{\sigma}_j \tilde{\rho}_{i,j} - \tilde{\sigma}_i \sigma^{FX}(t) \rho_{i,FX}, \\ \mu^{FX}(t) &= \frac{\partial}{\partial t} \log \left(\frac{P(T_{m(t)+1}; t)}{\tilde{P}(T_{m(t)+1}; t)} \right) \\ \mu^S(t) &= \frac{\partial}{\partial t} \log(P(T_{m(t)+1}; t)) \\ \mu^{\tilde{S}}(t) &= \frac{\partial}{\partial t} \log(\tilde{P}(T_{m(t)+1}; t)) - \sigma^{FX}(t) \sigma^{\tilde{S}}(t) \rho_{FX, \tilde{S}}(t). \end{aligned}$$

If we chose for $P(T_{m(t)+1}; t)$ and $\tilde{P}(T_{m(t)+1}; t)$ the interpolation given in (24.6), then

it follows that

$$\begin{aligned}\mu^{FX}(t) &= \frac{\partial}{\partial t} \log \left(\frac{1 + \tilde{L}_{m(t)}(T_{m(t)}) \cdot (T_{m(t)+1} - t)}{1 + L_{m(t)}(T_{m(t)}) \cdot (T_{m(t)+1} - t)} \right) \\ \mu^S(t) &= \frac{\partial}{\partial t} \log \left(\frac{1}{1 + L_{m(t)}(T_{m(t)}) \cdot (T_{m(t)+1} - t)} \right) \\ \mu^{\tilde{S}}(t) &= \frac{\partial}{\partial t} \log \left(\frac{1}{1 + \tilde{L}_{m(t)}(T_{m(t)}) \cdot (T_{m(t)+1} - t)} \right) - \sigma^{FX}(t) \sigma^{\tilde{S}}(t) \rho_{FX, \tilde{S}}(t).\end{aligned}$$

24.3.2. Implementation

Due to the many state variables, i.e. the high Markov dimension, it is natural to consider an implementation as a path simulation (Monte-Carlo simulation). The first steps towards in implementation is the discretization of the simulation time t by suitable discretization scheme. Since the processes are lognormal processes, we use the Euler-scheme for the log-process, see Section 13.1, 13.1.2.3.

The discretization of the interest rate processes L_i has been presented in Section 17.1.3, the interest rate processes of the foreign currency rates $\tilde{L}_i$ are discretized likewise. For FX we find

$$\begin{aligned}FX(t + \Delta t) &= FX(t) \cdot \exp \left(\int_t^{t+\Delta t} \mu^{FX}(\tau) d\tau - \frac{1}{2} \int_t^{t+\Delta t} \sigma^{FX}(\tau)^2 d\tau + \int_t^{t+\Delta t} \sigma^{FX}(\tau) dW^{FX}(\tau) \right) \\ &= FX(t) \cdot \exp \left((\bar{\mu}^{FX}(t) - \frac{1}{2} \bar{\sigma}^{FX}(t)^2) \Delta t + \bar{\sigma}^{FX}(t) \Delta W^{FX}(t) \right)\end{aligned}$$

with $\Delta W^{FX}(t) := W^{FX}(t + \Delta t) - W^{FX}(t)$ and

$$\begin{aligned}\bar{\mu}^{FX}(t) &:= \frac{1}{\Delta t} \int_t^{t+\Delta t} \mu^{FX}(\tau) d\tau \\ \bar{\sigma}^{FX}(t) &:= \sqrt{\frac{1}{\Delta t} \int_t^{t+\Delta t} \sigma^{FX}(\tau)^2 d\tau}\end{aligned}$$

If we especially chose the time discretization of the Monte-Carlo Simulation to match the *tenor* structure $T_0, T_1, T_2, \dots$, then we have

$$FX(T_{i+1}) = FX(T_i) \cdot \exp \left(\bar{\mu}^{FX}(T_i) \Delta T_i - \frac{1}{2} \bar{\sigma}^{FX}(T_i)^2 \Delta T_i + \bar{\sigma}^{FX}(T_i) \Delta W^{FX}(T_i) \right),$$

with $\Delta T_i := T_{i+1} - T_i$ and $\Delta W^{FX}(T_i) := W^{FX}(T_{i+1}) - W^{FX}(T_i)$ and

$$\begin{aligned}\bar{\mu}^{FX}(T_i) &= \frac{1}{\Delta T_i} \int_{T_i}^{T_{i+1}} \mu^{FX}(\tau) d\tau \\ &= \frac{1}{\Delta T_i} \log \left(\frac{1 + \tilde{L}_i(T_i) \cdot (T_{i+1} - t)}{1 + L_i(T_i) \cdot (T_{i+1} - t)} \right) \Bigg|_{t=T_i}^{T_{i+1}} \\ &= \frac{1}{\Delta T_i} \log \left(\frac{1 + L_i(T_i) \cdot (T_{i+1} - T_i)}{1 + \tilde{L}_i(T_i) \cdot (T_{i+1} - T_i)} \right) \\ \bar{\sigma}^{FX}(T_i) &:= \sqrt{\frac{1}{\Delta T_i} \int_{T_i}^{T_{i+1}} \sigma^{FX}(\tau)^2 d\tau}.\end{aligned}$$

This Euler discretization is exact. The discretization of the processes S and $\tilde{S}$ follows likewise. With “discrete drift term”

$$\begin{aligned}\bar{\mu}^S(T_i) &= \frac{1}{\Delta T_i} \log(1 + L_i(T_i) \cdot (T_{i+1} - T_i)), \\ \bar{\mu}^{\tilde{S}}(T_i) &= \frac{1}{\Delta T_i} \log(1 + \tilde{L}_i(T_i) \cdot (T_{i+1} - T_i)) - \int_{T_i}^{T_{i+1}} \sigma^{FX}(t) \sigma^{\tilde{S}}(t) \rho_{FX, \tilde{S}}(t) dt.\end{aligned}$$

Part VII.

Implementation

CHAPTER 25

Object-Oriented Implementation in Java™

An English version of this part is currently not available. What follows is the German version.

From early on, we wanted a product that would seem so natural and so inevitable and so simple, you almost wouldn't think of it as having been designed.

*Jonathan Ive
iPod™ Design Team, Apple Computer, [35].*

Wir können in diesem Buch keine umfassende Einführung in die objektorientierte Programmierung oder speziell die Programmiersprache Java™ geben. Hierzu sei auf die weiterführende Literatur [34, 11] verwiesen. Warum dann überhaupt dieses Kapitel? Die Intention dieses Kapitels ist ähnlich der Intention des Kapitel 2. Dort bestand der Anspruch nicht in einer umfassenden Behandlung der Wahrscheinlichkeitstheorie und der Theorie der stochastischen Prozesse. Vielmehr sollten die wesentlichen Grundlagen mit der späteren Anwendung im Blick entwickelt werden und insbesondere eine Intuition und Interpretation für die abstrakten Begriffen und deren Bedeutung in und für die Anwendung kreiert werden.¹ Unsere Zielsetzung ist hier ähnlich: Wir geben einen kurzen Überblick über Elemente der objektorientierten Programmierung (Kapitel 25) sowie ihre Anwendung im Kontext der Implementierung von Bewertungsmodellen.

¹ Beispiel: Warum wird ein so abstraktes Konstrukt wie eine σ -Algebra definiert? Warum ist in unserem Kontext die Betrachtung des Itô-Integrals sinnvoller als die Betrachtung des Stratonovich-Integral?

25.1. Elements of Object Oriented Programming: Class and Objects

Wir definieren zuerst die beiden Begriffe *Klasse* und *Objekt*.

Definition 215 (Klasse):

Eine Klasse besteht aus

- Einer Beschreibung einer *Datenstruktur*.
- Einer Beschreibung eines Satzes von Funktionen, den *Methoden*, die auf der Datenstruktur sowie weiteren (als Argument übergebenen) Daten operiert, d.h. die Elemente der Datenstruktur (ggf. mit als Argument übergebenen Daten) logisch verknüpft. Die Beschreibung der Methoden besteht ihrerseits aus
 - Einer Beschreibung der Aufrufkonventionen der Methoden, dem *Interface*, siehe Abschnitt 25.1.3 und Definition 218.
 - Einer Beschreibung der logischen Verknüpfung der Daten, die die Methode (Funktion) leistet, der *Implementierung*.

Definition 216 (Objekt):

Wir sagen *X* ist ein *Objekt* der Klasse *K*, wenn

- *X* Speicher bereitstellt um Daten entsprechend der in *K* beschriebene Datenstruktur zu speichern und
- die in *K* beschriebenen Methoden auf die in *X* enthaltenen Daten angewendet werden können.

In diesem Fall heißt auch *K* der *Typ* von *X*.



Interpretation: Die Klasse ist der Bauplan für ein Objekt, während ein Objekt ein konkretes Exemplar (*instance*²) der Klasse darstellt.

Betrachtet man wie Klasse und Objekt im Rechner realisiert werden, so stellt man fest, dass das Objekt lediglich die Daten entsprechend des in der Klasse *K* beschriebenen Datenlayouts speichert, während die auf diese Daten operierenden Funktionen alle in *K* beschrieben werden. Die Klasse ist somit die Beschreibung des Datenlayouts und der Funktionalitäten während das Objekt den jeweiligen Datensatz repräsentiert. Die Definition einer Klasse existiert nur einmal, während Objekte einer Klasse (Datensätze) mehrfach existieren. *Klasse und Objekt trennen Logik (Code) und Daten*. Selbstverständlich muss die Logik, also die Klasse, das Layout (die Struktur) der Daten kennen.

² (the object) *X* is an *instance of* (the class) *C*.

Oft wird zur Verdeutlichung des Verhältnisses von Klasse zu Objekt eine Analogie angeführt, wie z.B. *der Mensch* ist eine Klasse und *das konkrete Individuum „Christian Fries“* ist ein Objekt der Klasse *Mensch*. Solche Analogien sind jedoch nicht bis ins letzte tragfähig. Es wird dabei kaum deutlich, dass ein Objekt nur ein Datenspeicher ist, der Code, d.h. die Logik der Datenverarbeitung jedoch nur einmal existiert, nämlich in der Klasse.³ Hingegen besitzt jedes Individuum eigene Erfahrungen (als Analogie zu den Daten) und eigene Handlungsmuster (als Analogie zu Code der Daten und Umgebungsparameter verknüpft).

Neben den in der Klasse beschriebenen Daten trägt das Objekt noch ein weiteres Datum, nämlich seinen Typ. Dieser Typ bezeichnet die Klasse von der das Objekt ist. Es entsteht somit eine feste Verknüpfung von Objekt und zugehöriger Klasse, d.h. den auf die Daten anwendbaren Methoden. ◁|

25.1.1. Example: Class of a Binomial Distributed Random Variable

Es sei B eine binomialverteilte Zufallsvariable definiert über dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ mit $\Omega = \{\omega_1, \omega_2\}$. Wahrscheinlichkeitsraum und Zufallsvariable lassen sich durch drei Werte $b_1, b_2, p \in \mathbb{R}$ charakterisieren:

$$\begin{aligned} B(\omega_1) &= b_1, & P(\omega_1) &= p \\ B(\omega_2) &= b_2, & P(\omega_2) &= 1 - p. \end{aligned} \quad (25.1)$$

Gleichung (25.1) beschreibt die Klasse „Binomialverteilte Zufallsvariable“ während die Zufallsvariable C mit

$$\begin{aligned} C(\omega_1) &= 1, & P(\omega_1) &= 0.5 \\ C(\omega_2) &= -1, & P(\omega_2) &= 0.5. \end{aligned} \quad (25.2)$$

ein (konkretes) Objekt, d.h. ein Exemplar der Klasse „Binomialverteilte Zufallsvariable“ darstellt. Selbstverständlich müssen Operationen, die auf diese Zufallsvariable angewendet, werden nur auf Klassenebene definiert werden. Zum Beispiel ist die Berechnung des Mittelwertes definiert als

$$E^{\mathbb{P}}(B) = \int_{\Omega} B \, dP = p \cdot B(\omega_1) + (1 - p) \cdot B(\omega_2) = p \cdot b_1 + (1 - p) \cdot b_2. \quad (25.3)$$

Dass $E(C) = 0.5 \cdot 1.0 + 0.5 \cdot (-1.0) = 0.0$ ist, folgt aus (25.3).

³ Dass der Code nur in der Klasse gespeichert ist zeigt sich am Speicherplatzbedarf von Methoden: Fügt man einer Klasse einen Datenfeld hinzu und legt von dieser Klasse 100 Objekte an, so wird für das neue Datenfeld natürlich das 100-fache seines Speicherplatzes benötigt. Fügt man einer Klasse eine Methode hinzu, so wird der Code nur einmal gespeichert und der Speicherbedarf ist völlig unabhängig von der Zahl der erzeugten Objekte.

In Java™ sieht eine entsprechende Klasse wie folgt aus: Die Klassendefinition beginnt (nach einem Kommentar) mit der Beschreibung des Datenlayouts, hier `value1`, `value2` und `probabilityOfState1` für b_1 , b_2 und p respektive,

Listing 25.1: *BinomialDistributedRandomVariable: Binomialverteilte Zufallsvariable*

```

package com.christianfries.finmath.tutorial.example1;
/**
 * @author Christian Fries
 * @version 1.0
 */
class BinomialDistributedRandomVariable {
    private double value1;
    private double value2;
    private double probabilityOfState1;

```

gefolgt vom Konstruktor

```

    /**
     * This class implements a binominal distributed random variable
     * @param value1 The value in state 1.
     * @param value2 The value in state 2.
     * @param probabilityOfState1 The probability of state 1.
     */
    public BinomialDistributedRandomVariable(double value1, double value2,
        double probabilityOfState1) {
        this.value1 = value1;
        this.value2 = value2;
        this.probabilityOfState1 = probabilityOfState1;
    }

```

und der Beschreibung der Methode `getExpectation`.

```

    /**
     * @return The expectation of the random variable
     */
    public double getExpectation() {
        double expectation = probabilityOfState1 * value1
            + (1-probabilityOfState1) * value2;
        return expectation;
    }
}

```

25.1.2. Constructor

Der Konstruktor einer Klasse ist eine Methode die bei Erzeugung eines jeden Objektes aufgerufen wird (es kann verschiedene Konstruktoren geben und dann ist es möglich bei Erzeugung des Objektes den entsprechenden Kontruktor zu wählen). Mittels des Konstruktors lassen sich also über die Reservierung des Speicherplatzes für das Datenlayout hinausgehende, vorbereitende (initialisierende) Maßnahmen treffen - in unserem Fall das setzen der entsprechenden Werte b_1 , b_2 und p .

Der obige Code des Konstruktors der Klasse `BinomialDistributedRandomVariable` mag verwundern: Die Argumente des Konstruktors haben die gleichen Namen wie die Datenfelder der Klasse. Dies ist erlaubt, wird oft verwendet, kann aber bei längerem Code verwirrend wirken. In einem solchen Fall bezeichnet der entsprechende Name immer das Argument des Konstruktors. Um auf das gleichnamige Datenfeld der Klasse zuzugreifen muss

`this.` vorangestellt werden. Der obige Konstruktor setzt also die Datenfelder eines Objektes der Klasse auf die Werte der Parameter.

25.1.3. Methods: Getter, Setter, Static Methods

25.1.3.1. Aufrufkonvention, Signatur

Unter der Aufrufkonvention einer Methode verstehen wir den Namen der Methode zusammen mit der Liste der Argumenttypen, d.h. die Aufrufkonvention besagt unter welchem Namen die Methode mit welchen Argumenten angesprochen werden muss. Die Liste der Argumenttypen wird *Signatur* der Methode genannt. Zwei Methoden gleichen Namens, jedoch unterschiedlicher Signatur werden als unterschiedliche Methoden interpretiert. Das Bereitstellen einer Methode gleichen Namens und unterschiedlicher Signatur wird als *Überladen* der Methode bezeichnet.⁴

25.1.3.2. Getter, Setter

Sind die Datenfelder eines Objektes öffentlich zugänglich, so kann über `objektname.datenfeldname` auf sie zugegriffen werden, d.h. sie können verändert oder ausgelesen werden. Dieser Zugriff auf die Datenfelder kann erlaubt oder verboten werden, siehe dazu *Kapselung* in Abschnitt 25.2.1.

Datenfelder können natürlich auch nach der Objekterzeugung über Methoden geändert (gesetzt, engl. *set*) oder gelesen (engl. *get*) werden. Diese Methoden werden *Setter* bzw. *Getter* genannt

Wir fügen einen Setter für den Wert b_1 in die Klassendefinition ein:

```

/**
 * @param value1 The value of state 1.
 */
public void setValue1(double value1) {
    this.value1 = value1;
}

```

45
46
47
48
49
50

Die Methode ändert lediglich den Zustand des Objektes und besitzt daher keinen Rückgabewert. Dies wird durch das Schlüsselwort `void` signalisiert.

25.1.3.3. Statische Methoden

Methoden, welche keine Datenfelder des Objektes benötigen, d.h. lesen oder verändern, heißen *statische Methoden*.

Definition 217 (statische Methode):

Eine Methode einer Klasse K , die Objekte der Klasse K invariant lässt und von deren Daten unabhängig ist heißt *statisch*. Eine statische Methode wird auch als *Klassenmethode* bezeichnet.

⁴ Es kann keine zwei Methoden gleichen Namens und gleicher Signatur mit unterschiedlichen Rückgabewerttypen geben. Der Rückgabewerttyp (allein) kann durch überladen nicht verändert werden.

Eine Methode wird in Java™ durch das Schlüsselwort `static` als statisch deklariert.

Um die Methoden auf Daten anzuwenden werden Objekte angelegt. Entsprechenden Code, der das Arbeiten mit Objekten der obigen Klasse demonstriert und unsere Klasse testet, schreiben wir in die `main`-Methode⁵.

```

    public static void main( String args[] ) {
        /*
         * Test of class BinomialDistributedRandomVariable
         */
        System.out.println("Creating_random_variable.");
        BinomialDistributedRandomVariable randomVariable =
            new BinomialDistributedRandomVariable(1.0, -1.0, 0.5);    // Create object
        double expectation = randomVariable.getExpectation();
        System.out.println("Expectation_is:_" + expectation);
        System.out.println("Changing_value_of_state_1_to_0.4.");
        randomVariable.setValue1(0.4);
        expectation = randomVariable.getExpectation();
        System.out.println("Expectation_is:_" + expectation);
    }

```

Mit `new` wird in Zeile ?? ein neues Objekt des Typs `BinomialDistributedRandomVariable` erzeugt (Reservierung des Speicherplatzes entsprechend Datenlayout, Aufruf des Konstruktors mit den angegebenen Argumenten) und sodann in eine Objektreferenz des Typs `BinomialDistributedRandomVariable` gespeichert.



Further Reading: In [11, 34]: primitive Typen, Objektreferenzen (Referenzen), statische Methoden (das Schlüsselwort `static`), Rückgabewerte (das Schlüsselwort `void`), die `main`-Methode, Kommentare und der JavaDoc-Standard. <|

25.2. Principles of Object Oriented Programming: Data Hiding, Abstraction, Inheritance, Polymorphism

25.2.1. Data Hiding and Interfaces

Für den Zugriff auf Daten eines Objektes sind zwei Wege denkbar: Eine Möglichkeit ist die Bereitstellung zweier Methoden, die es erlauben das entsprechende Datenfeld zu lesen oder zu beschreiben, d.h. es werden *Getter* und *Setter* implementiert.⁶ Im Beispiel der Klasse `BinomialDistributedRandomVariable` exemplarisch für das Datenfeld `probabilityOfState1`:

⁵ Die `main`-Methode kann von außen aufgerufen werden ohne das ein entsprechende Objekt existiert. Sie ist *statisch*. Sie ist damit (insbesonder auch, da sie kein Objekt bedarf) ein möglicher Eintrittspunkt in ein Programm.

⁶ Von englisch *to get* und *to set*.

Listing 25.2: Getter und Setter

```

/**
 * @return Returns the probability of state 1.
 */
public double getProbabilityOfState1() {
    return probabilityOfState1;
}

/**
 * @param probabilityOfState1 The probability of state 1.
 */
public void setProbabilityOfState1(double probabilityOfState1) {
    this.probabilityOfState1 = probabilityOfState1;
}

```

Die Verwendung dieser Methoden zum Zugriff auf das Datenfeld `probabilityOfState1` könnte wie folgt aussehen:

```

BinomialDistributedRandomVariable randomVariable =
    new BinomialDistributedRandomVariable(1.0, -1.0, 0.5);

double p = randomVariable.getProbabilityOfState1(); // Get probability of state 1
System.out.println( "Current_value_of_p_is:_" + p);

randomVariable.setProbabilityOfState1(0.3); // Change probability of state 1 to
0.3

```

Eine andere Möglichkeit bietet der direkte Zugriff auf das entsprechende Datenfeld:

```

BinomialDistributedRandomVariable randomVariable =
    new BinomialDistributedRandomVariable(1.0, -1.0, 0.5);

double p = randomVariable.probabilityOfState1; // Get probability of state 1
System.out.println( "Current_value_of_p_is:_" + p);

randomVariable.probabilityOfState1 = 0.3; // Change probability of state 1 to
0.3

```

Diese letzte Variante funktioniert ohne das entsprechende Methoden existieren.⁷ Es ist der direkte Zugriff auf die internen Daten des Objektes.

Zunächst scheint der direkte Zugriff auf die Datenstruktur wesentlich bequemer, sowohl für den Entwickler der Klasse, der keine *Getter* und *Setter* implementieren muss, als auch für den Nutzer der Klasse. Auch wäre es denkbar, dass der direkte Zugriff lediglich eine abkürzende Schreibweise für die trivialen *Getter* und *Setter* aus Listing 25.2 sind. Im Allgemeinen ist jedoch der direkte Zugriff auf die Datenstruktur einer Klasse nicht möglich, ja die Datenstruktur an sich muss noch nicht einmal dem Nutzer bekannt sein. Um einen direkten Zugriff auf ein Datenfeld erst zu ermöglichen muss das Schlüsselwort `public` verwendet werden:

```

class BinomialDistributedRandomVariable {

    public double value1;
    public double value2;
    public double probabilityOfState1;
}

```

25.2.1.1. Kapselung

Das Verbergen von Datenstruktur und interner Datenverarbeitung wird als *Kapselung* bezeichnet. Der grundlegende Vorteil der Kapselung liegt darin, dass die genannten

⁷ Sofern der Zugriff auf die Daten freigegeben wurde, in Java™ durch das Voranstellen des Schlüsselwortes `public`.

Elemente, insbesondere die Datenstruktur, nachträglich geändert werden können. Benutzer der Klasse (die nur auf die Methoden der Objekte Zugriff haben) können von diesen Änderung unberührt gehalten werden. Von „Aussen“ verhält sich die Klasse wie zuvor.

Die Vorteile der Kapselung lassen sich schon am einfachen Beispiel der binomialverteilten Zufallsvariable verdeutlichen. Wir geben zwei Beispiele:

Beispiel für Kapselung: Anbieten alternativer Methoden Ganz entsprechend zu *Getter* und *Setter* für die Wahrscheinlichkeit $P(\{\omega_1\})$ des Zustands ω_1 bieten wir *Getter* und *Setter* für die Wahrscheinlichkeit $P(\{\omega_2\})$ des Zustands ω_2 an:

```

/**
 * @return Returns the probability of state 2.
 */
public double getProbabilityOfState2() {
    return 1.0 - probabilityOfState1;
}

/**
 * @param probabilityOfState2 The probability of state 2.
 */
public void setProbabilityOfState2(double probabilityOfState2) {
    this.probabilityOfState1 = 1.0 - probabilityOfState2;
}

```

Während zuvor der Zustand ω_1 durch die angebotenen Methoden ausgezeichnet wurde (nur seine Wahrscheinlichkeit p konnte gesetzt werden) und die Wahrscheinlichkeit des Zustands ω_2 nur als abgeleitete Größe $P(\{\omega_2\}) = 1 - p$ existierte, werden nun die Zustände in den angebotenen Methoden völlig gleichberechtigt repräsentiert. Die Art und Weise wie die Merkmale einer binomialverteilten Zufallsvariable intern repräsentiert werden, d.h. wie die Daten des Objektes gespeichert werden, ist von aussen nicht einsehbar. Insbesondere ist es möglich das Datenlayout zu verändern ohne die bereitgestellten Methoden in ihrer Spezifikation zu ändern.

Beispiel für Kapselung: Geschwindigkeitssteigerung durch Erweitern des internen Datenmodells um einen Cache Das in Listing 25.1 beschriebene Datenmodell besteht aus $B(\omega_1)$ (`value1`), $B(\omega_2)$ (`value2`) und $P(\{\omega_1\})$ (`probabilityOfState1`). Dies hat zur Folge, dass der Erwartungswert entsprechend

$$E(B) := p \cdot (B(\omega_1) + (1 - p) \cdot B(\omega_2)) = p \cdot (B(\omega_1) - B(\omega_2)) + B(\omega_2)$$

berechnet wird. Dies leistet die Methode `getExpectation()`. Wird in einem gewissen Kontext diese Methode relativ häufig aufgerufen, so trägt es zur Geschwindigkeitssteigerung bei den Erwartungswert einmalig zu berechnen und das Ergebnis in einem Speicher (*Cache*) vorzuhalten. Wir fügen dazu das Datenfeld `mean` als *Cache* hinzu

```
class BinomialDistributedRandomVariable {
```

```
private double value1;
private double value2;
private double probabilityOfState1;

private double mean;           // Cache for the mean
```

welches durch die Methode `updateMean` mit der Berechnung des Erwartungswertes befüllt wird.

```
private void updateMean() {
    mean = probabilityOfState1 * value1 + (1-probabilityOfState1) * value2;
}
```

Zusätzlich wird am Ende des Konstruktors sowie eines jeden Setters der den Zustand eines Objektes der Klasse, d.h. die Daten, ändert, ein Aufruf von `updateMean()` angefügt. Dies stellt sicher, dass `mean` immer den aktuell gültigen Mittelwert enthält. Die Methode `getExpectation` liefert nun lediglich den Wert aus dem Zwischenspeicher `mean`:

```
/**
 * @return The expectation of the random variable
 */
public double getExpectation() {
    return mean;
}
```

Ein mehrfacher Aufruf von `getExpectation` führt nun nicht mehr zur mehrfachen Berechnung des (gleichen) Mittelwertes.

Selbstverständlich darf dem Benutzer der Klasse kein Zugriff auf das Speicherfeld `mean` eines entsprechenden Objektes gewährt werden. Dies wäre fatal. Die Zeilen

```
System.out.println("Creating_random_variable.");
BinomialDistributedRandomVariable randomVariable =
    new BinomialDistributedRandomVariable(1.0, -1.0, 0.5);    // Create object

randomVariable.mean = 0.2;
```

würden das Objekt `randomVariable` in einen inkonsistenten Zustand versetzen. Der Benutzer soll weder auf das Vorhandensein des Caches bauen noch diesen manipulieren können. Aus diesem Grunde sind `mean` und `updateMean` unbedingt als *private* zu deklarieren. Auch die übrigen Datenfelder müssen nun unbedingt als *private* deklariert werden. Werden sie geändert, so muss der `mean` Neuberechnet werden. Dies ist gewährleistet in dem jeder *Setter* zusätzlich `updateMean` veranlasst. Eine direkte Manipulation der Datenfelder würde dies umgehen.

25.2.1.2. Interfaces

Der Vorteil der Kapselung liegt also darin, dass das interne Datenlayout bei Bedarf geändert werden kann. Durch eine Modifikation der ebenfalls verborgenen Implementierung der Methoden wird sichergestellt, dass die Methoden die gleiche Funktionalität wie zuvor anbieten. Für den Nutzer der (Objekte der) Klasse sind nur die Aufrufkonventionen der Methoden, das *Interface*, relevant.

Definition 218 (Interface):

Die Beschreibung der Aufrufkonventionen von Methoden wird als *Interface* bezeichnet. Analog zur Definition 216 besteht ein Interface aus

- Einer Beschreibung der Aufrufkonventionen eines Satzes von Funktionen, den *Methoden*.

Definition 219 (Kapselung):

Bietet eine Klasse ihre Funktionalität nur in Form eines *Interface* an, so heißt die Klasse *gekapselt*. Man spricht von *Kapselung*.

25.2.1.3. Beispiel: Diskrete reellwertige Zufallsvariable

Ein Beispiel für ein Interface einer diskrete reellwertige Zufallsvariable, d.h. eine reelle Zufallsvariable definiert über einem endlichen Raum $\Omega = \{\omega_1, \dots, \omega_n\}$ ist gegeben durch

Listing 25.3: *DiscreteRandomVariableInterface: Interfacebeschreibung für eine diskrete Zufallsvariable*

```
/*
 * Created on 04.12.2004
 *
 * (c) Copyright Christian P. Fries, Germany. All rights reserved. Contact: email@christian-fries.de.
 */
package com.christianfries.finmath.tutorial.randomVariables;

/**
 * This is the interface for a discrete real valued random variable.
 *
 * @author Christian Fries
 */
public interface DiscreteRandomVariableInterface {
    int getNumberOfStates();
    double getProbabilityOfState(int stateIndex);
    double getValueOfState(int stateIndex);

    double getExpectation();
    double getVariance();
}
```



Further Reading: In [11, 34]: Die Schlüsselwörter `public`, `private` und `protected` für Datenfelder, Interfaces.

25.2.2. Abstraction and Inheritance

Interface und Klasse sind zwei Extreme zwischen denen Zwischenzustände denkbar sind, nämlich Klassen in denen einige Methoden eine Implementierung besitzen, andere jedoch nur durch die Beschreibung ihrer Aufrufkonvention gegeben sind (also als *Interface*). Methoden deren Implementierung noch nicht spezifiziert ist heißen *abstrakte Methoden*. Eine Klasse mit mindestens einer abstrakten Methode heißt *abstrakte Klasse*.

Ein Anwendungsbeispiel legt die Betrachtung der obigen Interface-Beschreibung *DiscreteRandomVariableInterface* nahe: Die Implementierung für Er-

wartungswert und Varianz lassen sich erstellen ohne Kenntnis über das interne Datenlayout der Klasse zu besitzen, sie bieten sich für eine Teilimplementierung an.⁸

Listing 25.4: *DiscreteRandomVariable: Abstrakte Basisklasse für diskrete Zufallsvariable*

```

/*
 * Created on 05.12.2004
 *
 * (c) Copyright Christian P. Fries, Germany. All rights reserved. Contact: email@christian-fries.de.
 */
package com.christianfries.finmath.tutorial.randomVariables;

/**
 * @author Christian Fries
 */
public abstract class DiscreteRandomVariable implements DiscreteRandomVariableInterface {
    public double getExpectation() {
        double expectation = 0.0;
        for(int stateIndex=0; stateIndex<this.getNumberOfStates(); stateIndex++) {
            expectation += this.getValueOfState(stateIndex) * this.getProbabilityOfState(stateIndex);
        }

        return expectation;
    }

    public double getVariance() {
        // Calculate second moment
        double secondMoment = 0.0;
        for(int stateIndex=0; stateIndex<this.getNumberOfStates(); stateIndex++) {
            double value = this.getValueOfState(stateIndex);
            secondMoment += value * value * this.getProbabilityOfState(stateIndex);
        }

        // Calculate expectation
        double expectation = this.getExpectation();

        // Return variance
        return secondMoment - expectation*expectation;
    }
}

```

Um nun eine Klasse zu definieren, welche das Interface `DiscreteRandomVariableInterface` bereitstellt ist es lediglich nötig die Klasse `DiscreteRandomVariable` um die Implementierung der noch abstrakten Methoden zu erweitern (engl.: *to extend*). Die ist in eleganter Weise möglich in dem die neue Klasse die schon definierten Eigenschaften von `DiscreteRandomVariable` *erbt*. Wir definieren auf diese Weise eine (neue) Klasse für die binomialverteilte Zufallsvariable:

Listing 25.5: *BinomialDistributedRandomVariable: Binomialverteilte Zufallsvariable*

```

/*
 * Created on 05.12.2004
 *
 * (c) Copyright Christian P. Fries, Germany. All rights reserved. Contact: email@christian-fries.de.
 */
package com.christianfries.finmath.tutorial.randomVariables;

/**
 * @author Christian Fries
 * @version 1.0
 */
class BinomialDistributedRandomVariable extends DiscreteRandomVariable {

```

⁸ Eine abstrakte Klasse muss nicht notwendigerweise keine Beschreibung eines Datenlayouts besitzen.

```

double value1;
double value2;
double probabilityOfState1;

class ChangeOfMeasureException extends Exception {
    String cause;

    /**
     * @param cause The reason for this exception.
     */
    public ChangeOfMeasureException(String cause) {
        super(cause);
        this.cause = cause;
    }
}

/**
 * This class implements a binominal distributed random variable
 *
 * @param value1 The value in state 1.
 * @param value2 The value in state 2.
 * @param probabilityOfState1 The probability of state 1.
 */
public BinomialDistributedRandomVariable(double value1, double value2, double probabilityOfState1) {
    this.value1 = value1;
    this.value2 = value2;
    this.probabilityOfState1 = probabilityOfState1;
}

public int getNumberOfStates() {
    return 2;
}

public double getProbabilityOfState(int stateIndex) {
    if (stateIndex == 0)
        return probabilityOfState1;
    else
        return 1.0 - probabilityOfState1;
}

public double getValueOfState(int stateIndex) {
    if (stateIndex == 0)
        return value1;
    else
        return value2;
}

public void changeMeasureToMatchGivenExpectation(double expectation)
    throws ArithmeticException, ChangeOfMeasureException {

    // Check if we have anything to do
    if (expectation == this.getExpectation()) return; // Nothing to do here

    // Check if change of measure is possible: check if random variable is stochastic
    if (value1 == value2)
        throw new ArithmeticException("Random_variable_is_not_stochastic.");

    // Calculate candidate for new measure
    double quasiProbability = (expectation - value2) / (value1 - value2);

    if (quasiProbability < 0 || quasiProbability > 1)
        throw new ChangeOfMeasureException("Given_expectation_out_of_possible_range.");

    // Check if change of measure is possible
    probabilityOfState1 = quasiProbability;
}
}

```

Vererbung bezieht sich nicht (nur) auf die Implementierungen von Interfaces, d.h. dem Erben von abstrakten Klassen. Es ist ebenso möglich von einer Klasse (nicht notw. abstrakt) zu Erben und diese um einen neuen Teil des Datenlayouts, neue Methoden oder neue Implementierungen bestehender Methoden zu erweitern.

Definition 220 (Geerbte Klasse):

Es seien A und B Klassen. B heißt von A geerbt, wenn B mindestens das durch A

gegeben Interface implementiert. B heißt auch *abgeleitete Klasse*. A heißt *Basisklasse*, auch *Oberklasse*, *Superklasse* (*super class*).

Ist die Klasse B von A geerbt so sind Objekte des Typs B gleichzeitig auch Objekte des Typs A , sie sind polymorph, siehe Definition 221. ┘

Ein bequemes Element der Vererbung ist die Möglichkeit die Implementierung der Basisklasse *per default* zu übernehmen. Wird für eine Methode der Basisklasse in der abgeleiteten Klasse keine Angabe gemacht, so wird diese Methode geerbt.

25.2.2.1. Methoden: Überschreiben und überladen

25.2.3. Polymorphism

Bedeutender ist jedoch die Eigenschaft, dass die Objekte einer abgeleiteten Klasse von mehrfachem Typ sind. Da die abgeleitete Klasse das der Oberklasse entsprechende Interface realisiert sind ihre Objekte in allen Anwendungsbereichen der Oberklasse gleichwertig einsetzbar. Daher ist es möglich und sinnvoll, dass sie gleichzeitig als vom Typ A (Typ der Oberklasse) und vom Typ B (Typ der abgeleiteten Klasse selbst) gesehen werden können. Man sagt die Objekte sind *polymorph*.⁹ Ist die Basisklasse selbst eine abgeleitete Klasse, so besitzen die Objekte der abgeleiteten Klasse alle Typen der Basisklassen.¹⁰

Definition 221 (Polymorph):

Eine Objekt heißt *polymorph*, wenn es von mehrfachem Typ ist und sich entsprechend seinem abgeleiteten Typ verhält, selbst wenn es mit einem Typ der Basisklasse angesprochen wird, d.h. Methoden aufgerufen werden. ┘

Die Bedeutung der Polymorphie wird deutlich beim Methodenaufruf eines polymorphen Objektes. Methodenaufrufe auf polymorphe Objekte nutzen die sogenannte späte Bindung (*late binding*). Dabei wird bei einem Methodenaufruf eines polymorphen Objektes die Methode des abgeleiteten Typs aufgerufen, selbst wenn das Objekt in seinem Kontext als Objekt der Basisklasse angesprochen wird.

Remark 222 (Interface, Das Nachrichten Paradigma): Das Konzept des Interface ist eines der zentralen Konzepte der objektorientierten Programmierung. Die Vererbung ist im gewissen Sinne lediglich eine abkürzende Schreibweise für den Sachverhalt das eine Klasse eine Obermenge des Interface einer andern Klasse anbietet, wobei die Abkürzung darin besteht, dass für Methoden die in der abgeleiteten Klasse keine Implementierung besitzen die Implementierung der Basisklasse übernommen wird.

Für die Vererbung (in Java™) kommt nun noch das Konzept des Typs eines Objektes hinzu (und so das Konzept der Polymorphie). So ist es in Java™ möglich, dass

⁹ Objekte sind polymorph, d.h. von mehrfachem Typ. Nicht Klassen.

¹⁰ Da eine Klasse mehrere Interfaces implementieren kann liesse sich die Polymorphie der Vererbung rein über Interfaces realisieren.

die Aufrufkonventionen der Methoden zweier Objekte von zwei unterschiedlichen Klassen exakt gleich sind, die Objekte in ihrer Verwendung jedoch nicht ausgetauscht werden können, da sie von unterschiedlichem Typ sind. Blendet man diese Zusatzeigenschaft (*Typsicherheit*) aus, so bleibt als einziges Charakteristikum einer Klasse/eines Objektes das angebotene Interface. Die Aufrufkonvention einer Methode wird oft als „Nachricht, die ein Objekt empfangen kann“ interpretiert. Einige Programmiersprachen verzichten auf das Konzept des Typs von Objekten und beurteilen ein Objekt lediglich nach den angebotenen Methoden (den Nachrichten, auf die das Objekt reagiert). Sehr schöne Beispiele sind Smalltalk und Objective-C.

25.3. Example: A Class Structure for One Dimensional Root Finders

Gegeben sei das Problem eine Nullstelle einer Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ zu finden. Der Algorithmus zur Suche der Nullstelle werde in einer Klasse realisiert, die die spezielle Gestalt der Funktion nicht kennt, sondern ein Frage-Antwort-Verfahren realisiert: Die Klasse schlägt in jeder Iteration einen Wert x vor zu dem sie den entsprechenden Funktionswert mitgeteilt bekommt, d.h. sie erfragt den Funktionswert zu einem (selbstgewählten) x und entwickelt auf den Antworten eine Strategie sich der Lösung zu nähern.

25.3.1. Root Finder for General Functions

25.3.1.1. Interface

Eine solche Klasse muss eine Methode bereitstellen die den vorgeschlagenen Punkt x liefert (`double getNextPoint()`) und eine Methode über die der zugehörige Funktionswert gesetzt werden kann (`void setValue(double)`). Zusammen mit einigen Methoden zur Handhabung der Iteration und des Abbruchkriteriums sind folgende Methoden bereitzustellen:

Listing 25.6: *RootFinder: Interface eines eindimensionalen Nullstelllösers*

```
/*
 * Created on 30.05.2004
 *
 * (c) Copyright Christian P. Fries, Germany. All rights reserved. Contact: email@christian-fries.de.
 */
package net.finmath.fdm1.rootFinder;

/**
 * This is the interface for a one dimensional root finder
 * implemented as an question-and-answer algorithm.
 *
 * @author Christian Fries
 */
public interface RootFinder {

    /**
     * @return Next point for which a value should be set
     * using <code>setValue</code>.
     */
}
```

```

    */
    public double getNextPoint();

    /**
     * @param value Value corresponding to point returned
     * by previous <code>getNextPoint</code> call.
     */
    public void setValue(double value);

    /**
     * @return Returns the numberOfIterations.
     */
    public int getNumberOfIterations();

    /**
     * @return Best point obtained so far
     */
    public double getBestPoint();

    /**
     * @return Returns the accuracy.
     */
    public double getAccuracy();

    /**
     * @return Returns the isDone.
     */
    public boolean isDone();
}

```

Es fehlt sowohl das Datenlayout, als auch die konkrete Implementierung. `RootFinder` beschreibt nur das *Interface*. Offensichtlich muss ein Objekt einer Klasse, welches einen Nullstellensucher entsprechend diesem Interface definiert Informationen über den aktuellen Zustand der Suche speichern können, z.B. den zuletzt erfragten Wert oder gar eine längere Historie um daraus die Strategie abzuleiten. Welche Strategie verwendet wird (die Implementierung) und welche Informationen dazu wie gespeichert werden (das Datenlayout) ist für die Benutzung einer solchen Klasse unbedeutend. Hierzu genügt die Kenntnis des Interfaces. Es ist somit möglich schon jetzt eine Methode zu schreiben die einen gegebenen `RootFinder` an einer Funktion testet:

Listing 25.7: Testfunktion für Nullstellenlöser die dem Interface *RootFinder* genügen

```

public static void testRootFinder(RootFinder rootFinder) {
    System.out.println("Testing_" + rootFinder.getClass().getName() + ":");

    // Find a solution to  $x^3 + x^2 + x + 1 = 0$ 
    while(rootFinder.getAccuracy() > 1E-11 && !rootFinder.isDone()) {
        double x = rootFinder.getNextPoint();

        double y = x*x*x + x*x + x + 1;    // Our test function. Analytic solution is -1.0.
        rootFinder.setValue(y);
    }

    // Print result:
}

```

25.3.1.2. Bisection Search

Ein einfacher Nullstellenlöser ist die *Bisektionssuche* (*bisection search*). Die Klasse `BisectionSearch` realisiert dieses Verfahren in dem sie das Interface `RootFinder` implementiert. Ein entsprechender Code findet sich im Anhang D.1.

25.3.2. Root Finder for Functions with Analytic Derivative: Newton Method

Einige Nullstellenlöser, wie das Newtonverfahren verwenden für ihre Strategie der Punktsuche die Ableitung $f' = \frac{\partial f}{\partial x}$. Die „Suchstrategie“ des Newtonverfahrens ist

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}.$$

25.3.2.1. Interface

Eine zugehörige Klasse muss folglich ein leicht modifiziertes Interface implementieren, statt der `setValue(double value)`-Methode wird im Interface `RootFinderWithDerivative` eine `setValueAndDerivative(double value, double derivative)` angeboten.

25.3.2.2. Newtonverfahren

Die Klasse `NewtonsMethod` implementiert das Interface `RootFinderWithDerivative` mit einem Newtonverfahren. Für das Newtonverfahren sieht die entsprechende Methode wie folgt aus:

Listing 25.8: *NewtonsMethod: Nullstellenlöser nach dem Newtonverfahren der dem RootFinderWithDerivative-Interface genügt.*

```
/**
 * @param value The value corresponding to the point returned by previous <code>getNextPoint</code> call.
 * @param derivative The derivative corresponding to the point returned by previous <code>getNextPoint</code> call.
 */
public void setValueAndDerivative(double value, double derivative) {
    if(numberOfIterations == 0 || Math.abs(value) < accuracy)
    {
        accuracy = Math.abs(value);
        bestPoint = nextPoint;
    }

    // Calculate next point
    nextPoint = nextPoint - value/derivative;

    numberOfIterations++;
    expectingValue = false;
    return;
}
```

25.3.3. Root Finder for Functions with Derivative Estimation: Secant Method

25.3.3.1. Vererbung

Welche Möglichkeiten Vererbung und Interfaces bieten zeigt folgende Realisierung des *Sekantenverfahrens*. Die Suchstrategie des Sekantenverfahrens ist

$$x_{i+1} = x_i - \frac{f(x_i)}{\frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}}}.$$

Daraus werden zwei Punkte deutlich:

- Das Sekantenverfahren *ist* ein Newtonverfahren mit einem Schätzwert für die Ableitung: $f'(x_i) = \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}}$.
- Das Sekantenverfahren benötigt in jeder Iteration nur den Funktionswert $f(x_i)$ für den vorgeschlagenen Punkt x_i .

Für die Klasse übersetzen sich diese Eigenschaften zu:

- Das Sekantenverfahren erweitert (*extends*) die Klasse `NewtonMethod` um eine Schätzung der Ableitung.
- Das Sekantenverfahren implementiert (*implements*) das `RootFinder`-Interface.

Eine entsprechende Klasse sieht demnach wie folgt aus:

Listing 25.9: *SecantMethod*: Nullstellenlöser nach dem Sekantenverfahren der dem *RootFinderWithDerivative*-Interface genügt.

```
/*
 * Created on 19.02.2004
 *
 * (c) Copyright Christian P. Fries, Germany. All rights reserved. Contact: email@christian-fries.de.
 */
package net.finmath.fdm1.rootFinder;

/**
 * This class implements a root finder as a question-and-answer algorithm using
 * the secant method.
 *
 * @author Christian Fries
 * @version 1.0
 */
public class SecantMethod extends NewtonMethod implements RootFinder {

    double secondGuess;
    double lastPoint;
    double lastValue;

    /**
     * @param guess
     *        initial guess where the solver will start
     */
    public SecantMethod(double firstGuess, double secondGuess) {
        super(firstGuess);
        this.secondGuess = secondGuess;
    }
}
```

```

/**
 * @param value
 *      The value corresponding to the point returned by previous
 *      <code>getNextPoint</code> call.
 * @param derivative
 *      The derivative corresponding to the point returned by previous
 *      <code>getNextPoint</code> call.
 */
public void setValue(double value) {
    // Calculate approximation for derivative
    double derivative;
    if (getNumberOfIterations() == 0) {
        derivative = value / (secondGuess - nextPoint);
    } else {
        derivative = (value - lastValue) / (nextPoint - lastPoint);
    }

    // Remember last point
    lastPoint = nextPoint;
    lastValue = value;

    super.setValueAndDerivative(value, derivative);
}

return;
}

/**
 * @param value
 *      The value corresponding to the point returned by previous
 *      <code>getNextPoint</code> call.
 * @param derivative
 *      The derivative corresponding to the point returned by previous
 *      <code>getNextPoint</code> call.
 */
public void setValueAndDerivative(double value, double derivative) {
    // Remember last point
    lastPoint = nextPoint;
    lastValue = value;

    super.setValueAndDerivative(value, derivative);
}

return;
}
}

```

25.3.3.2. Polymorphie

An der Klasse `SecantMethod` zeigt sich die Polymorphie. Objekte der Klasse `SecantMethod` sind gleichzeitig Objekt der Klasse `NewtonMethod`, da `SecantMethod` von `NewtonMethod` erbt und somit das zugehörige Interface anbietet. Das heißt insbesondere das `SecantMethod` nicht nur das Interface `RootFinder` implementiert, sondern (dann als `Newton Methode`) auch das Interface `RootFinderWithDerivative`. Das sich `SecantMethod` bezüglich des Interface `RootFinderWithDerivative` exakt wie `NewtonMethod` verhält ist nicht verwunderliche: Wir haben keine Methode dieses Interfaces geändert. Dies zeigt auch das Ergebnis des Tests 25.10 aller Nullstellenlöser in 25.11.

Remark 223 (Vererbung: Spezialisierung und Erweiterung): Meist wird das Konstrukt „*B erbt von A*“ sprachlich als „*B ist ein A*“ gefasst. Zum Beispiel „*eine diskrete Zufallsvariable ist eine Zufallsvariable*“ oder „*eine binomialverteilte Zufallsvariable ist eine diskrete Zufallsvariable*“. Die Motivation für diese „sprachliche Krücke“ ist die Vorstellung, dass eine geerbte Klasse *B* eine Spezialisierung der Basisklasse *A*

ist. Diese Vorstellung hilft bei Erstellung einer Vererbungfolge (Vererbungshierarchie, Klassenhierarchie), sie ist aber nicht universell. So mutet z.B. „das Sekantenverfahren ist ein Newtonverfahren“ als falsch an. Die Verwendung von „... erweitert ...“ statt „... ist ein ...“ ist hier universeller. Beispielsweise gilt: „Das Sekantenverfahren erweitert das Newtonverfahren um einen Ableitungsschätzer“. Und in Java™ heisst das Schlüsselwort eben auch *extends*.

Wir testen die implementierten Nullstellenlöser mit der Klasse `TestRootFinders`. Da die Klassen nur zwei verschiedenen Interfaces besitzen genügt die Implementierung von zwei Testfunktionen:

Listing 25.10: Testfunktionen für *RootFinder* und *RootFinderWithDerivative* Klassen

```
/*
 * Created on 02.12.2004
 *
 * (c) Copyright Christian P. Fries, Germany. All rights reserved. Contact: email@christian-fries.de.
 */
package net.finmath.fdm1.rootFinder;

import java.text.DecimalFormat;

/**
 * @author Christian Fries
 */
public class TestRootFinders {
    public static void main(String[] args) {
        System.out.println("Applying_root_finders_to_x^3+_2*y^2+_x+_1=_0\n");

        System.out.println("Root_finders_without_derivative:");
        System.out.println("-----");

        RootFinder rootFinder;

        rootFinder = new BisectionSearch(-10.0,10.0);
        testRootFinder(rootFinder);

        rootFinder = new RiddersMethod(-10.0,10.0);
        testRootFinder(rootFinder);

        rootFinder = new SecantMethod(2.0,10.0);
        testRootFinder(rootFinder);

        System.out.println("");

        System.out.println("Root_finders_with_derivative:");
        System.out.println("-----");

        RootFinderWithDerivative rootFinderWithDerivative;

        rootFinderWithDerivative = new NewtonsMethod(2.0);
        testRootFinderWithDerivative(rootFinderWithDerivative);

        rootFinderWithDerivative = new SecantMethod(2.0,10.0);
        testRootFinderWithDerivative(rootFinderWithDerivative);
    }

    public static void testRootFinder(RootFinder rootFinder) {
        System.out.println("Testing_" + rootFinder.getClass().getName() + ":");

        // Find a solution to x^3 + x^2 + x + 1 = 0
        while(rootFinder.getAccuracy() > 1E-11 && !rootFinder.isDone()) {
            double x = rootFinder.getNextPoint();

            double y = x*x*x + x*x + x + 1;    // Our test function. Analytic solution is -1.0.

            rootFinder.setValue(y);
        }

        // Print result:
        DecimalFormat formatter = new DecimalFormat("0.00E00");
    }
}
```

```

        System.out.print("Root.....:_" + formatter.format(rootFinder.getBestPoint()) + " _");
        System.out.print("Accuracy.:_" + formatter.format(rootFinder.getAccuracy()) + " _");
        System.out.print("Iterations:_" + rootFinder.getNumberOfIterations() + "\n");
    }

    public static void testRootFinderWithDerivative(RootFinderWithDerivative rootFinder) {
        System.out.println("Testing_" + rootFinder.getClass().getName() + ":");

        // Find a solution to x^3 + x^2 + x + 1 = 0
        while(rootFinder.getAccuracy() > 1E-11 && !rootFinder.isDone()) {
            double x = rootFinder.getNextPoint();

            double y = x*x*x + x*x + x + 1;
            double p = 3*x*x + 2*x + 1;

            rootFinder.setValueAndDerivative(y,p);
        }

        // Print result:
        DecimalFormat formatter = new DecimalFormat("0.00E00");
        System.out.print("Root.....:_" + formatter.format(rootFinder.getBestPoint()) + " _");
        System.out.print("Accuracy.:_" + formatter.format(rootFinder.getAccuracy()) + " _");
        System.out.print("Iterations:_" + rootFinder.getNumberOfIterations() + "\n");
    }
}

```

Listing 25.11: Ausgabe des Tests der *RootFinder* und *RootFinderWithDerivative* Klassen

```

Applying root finders to x^3 + 2*y^2 + x + 1 = 0

Root finders without derivative:
-----
Testing net.finmath.fdm1.rootFinder.BisectionSearch:
Root.....: -1,00E00 Accuracy..: 9,09E-12 Iterations: 43
Testing net.finmath.fdm1.rootFinder.RiddersMethod:
Root.....: -1,00E00 Accuracy..: 3,33E-16 Iterations: 22
Testing net.finmath.fdm1.rootFinder.SecantMethod:
Root.....: -1,00E00 Accuracy..: 0,00E00 Iterations: 15

Root finders with derivative:
-----
Testing net.finmath.fdm1.rootFinder.NewtonsMethod:
Root.....: -1,00E00 Accuracy..: 1,33E-15 Iterations: 10
Testing net.finmath.fdm1.rootFinder.SecantMethod:
Root.....: -1,00E00 Accuracy..: 1,33E-15 Iterations: 10

```

25.4. Anatomy of a Java™ Class

Wir geben in Abbildung 25.1 auf Seite 384 (den Ausschnitt) einer Java™ Klasse mit den wichtigsten Elementen.

Vor der Deklaration der Klasse steht die Angabe des Paketnamens zu dem die Klasse gehört. Der volle Klassenname setzt sich aus dem Paketnamen und dem einzelnen Klassennamen zusammen und sollte weltweit eindeutig sein. Dazu wird der Paketname aus der Internetadresse des Herstellers der Klasse abgeleitet.

Es folgt die Angabe der im Folgenden verwendeten Klassen über das Schlüsselwort `import` gefolgt vom vollem Klassennamen.

Die Deklaration der Klasse beginnt mit dem Schlüsselwort `class` gefolgt vom Klassennamen und, eingeleitet durch die Schlüsselwörter `extends` und `implements` (optional) die Angabe der Basisklasse und der implementierten *Interfaces*. Fehlt die Angabe der Basisklasse so wird `java.lang.Object` als Ba-

sisklasse verwendet. Die bedeutet das alle Klassen mittelbar oder unmittelbar von `java.lang.Object` erben.

Es folgt die Deklaration des Datenmodells in Form von Attributen. Ein Attribut wird definiert durch die Angabe eines Typs (primitiver Typ, wie `double`, `int`, etc., oder einer Klasse) und seinem Namen. Zur Festlegung der Sichtbarkeit (**Kapselung**) kann `private`, `protected` oder `public` vorangestellt werden, keine Angabe deklariert das Attribut als `private`.

Der Rest der Klassendeklaration besteht aus der Deklaration und Implementierung der **Konstruktoren** und **Methoden**. Dem Methodenname wird der Typ Rückgabewert vorangestellt (sowie `void` für eine Methode ohne Rückgabewert). Diesem können weiter Schlüsselwörter vorangestellt werden (Sichtbarkeit: `public`, `private`, Deklaration als **Klassenmethode**: `static`, Unterbindung des Überschreibens: `final`). Ein Konstruktor ist eine Methode deren Name dem Klassennamen entspricht, die `public` ist und keinen Rückgabewert besitzt (die Angabe von `void` fehlt).¹¹



Further Reading: Die Java™ Schlüsselwörter `privat`, `protected`, `public`, `void`, `static`, `final`, `implements`, `extends`, `package` und `import` in [11, 34]. <|

25.5. Libraries

Ein grosser Vorteil an Java™ ist der Reichtum an Klassenbibliotheken, die, aufgrund klar definierter Interfaces und meist vorliegender JavaDoc Dokumentation leicht benutzt werden können. Nicht nur Datenverwaltung wie *Collections* sondern auch numerische Bibliotheken mit Verfahren aus Lineare Algebra, Statistik und Stochastik sind verfügbar.

Um bei der Vielzahl der vorhandenen Klassen Namenskonflikte zu vermeiden, gleichzeitig aber einfach Klassennamen zu erlauben werden die Klassen sogenannten Paketen (`package`) zugeordnet. Obige Klasse

25.5.1. Java™ 2 Platform, Standard Edition (j2se)

Die Pakete `java.*` der *Java™ 2 Platform, Standard Edition* bieten Basisfunktionalitäten, insbesondere die Verwaltung von Listen, Zeichenketten, Dateien, etc.

```
Paket(e) . . . : java.*
Hersteller . . : Sun Microsystems
Bezugsquelle: http://java.sun.com/j2se/
Lizenz . . . . : frei, siehe http://java.sun.com/j2se/
```

¹¹ Im Grunde könnte ein Objekt der betreffenden Klasse selbst als Rückgabewert des Konstruktors gesehen werden.

25.5.2. Java™ 2 Platform, Enterprise Edition (j2ee)

Die Pakete `javax.*` der *Java™ 2 Platform, Enterprise Edition* decken spezielle Aspekte ab, wie die graphische Benutzeroberfläche Swing und Internet Kommunikation.

```
Paket(e) . . . : javax.*
Hersteller . . : Sun Microsystems
Bezugsquelle:  http://java.sun.com/j2ee/
Lizenz . . . . : frei, siehe http://java.sun.com/j2ee/
```

25.5.3. Colt

Die Colt Bibliothek stellt in den Paketen `cern.colt.*` Funktionalitäten aus dem Bereich der Linearen Algebra zur Verfügung (Matrix Multiplikation, Invertierung, Eigenvektorzerlegung) und stellt in den Paketen `cern.jet.*` Funktionalität aus dem Bereich der Stochastik zur Verfügung (Zufallszahlengeneratoren, Verteilungsfunktionen).

```
Paket(e) . . . : cern.colt.*, cern.jet.*, cern.clhep
Hersteller . . : Wolfgang Hoschek.
                  Copyright (c) 1999 CERN - European Organization for Nuclear
                  Research.
Bezugsquelle:   http://dsd.lbl.gov/~hoschek/colt/
Lizenz . . . . : Die Pakete cern.colt*, cern.jet*, cern.clhep sind frei
                  für die kommerzielle Nutzung, siehe http://dsd.lbl.gov/~hoschek/colt/license.html.
```

25.5.4. Commons-Math: The Jakarta Mathematics Library

```
Hersteller . . : Diverse (Open Source), The Apache Software Foundation.
Bezugsquelle:   http://jakarta.apache.org/commons/math/
Lizenz . . . . : Apache License, siehe http://jakarta.apache.org/commons/license.html.
```

25.6. Some Final Remarks

25.6.1. Object Oriented Design (OOD) / Unified Modeling Language (UML)

Zwei wesentliche Stärken der objektorientierten Programmierung liegen in der Modularisierung der Lösung eines Problems, wie sie durch Kapselung und Abstraktion, unterstützt wird, sowie die Wiederverwendbarkeit und Erweiterbarkeit der Lösung eines Problems, wie sie durch Vererbung und Polymorphie unterstützt wird. Klar definierte

Interfaces ermöglichen ein (zeitlich wie personell) unabhängiges Entwickeln, Verfeinern und Optimieren von Teilkomponenten.

Bei der Erarbeitung einer objektorientierten Lösung muss somit zunächst eine sehr sorgfältige Festlegung der Interfaces erfolgen. Diese sollten eine effiziente Kommunikation der Objekte gewährleisten. Der Entwurf der Interfaces (und dann weiter der Klassen) wird als objektorientiertes Design (*objectoriented design*, *OOD*) bezeichnet, [13]. Für den Entwurf von komplexeren Lösungen existiert eine graphische Entwurfssprache (eine Konvention von graphischen) Symbolen (die *unified modeling language*, *UML*), [26].

Die Implementierung, Verfeinerung und Optimierung des Codes kann nach diesem Schritt bei grösseren Projekten leichter von mehreren Personen realisiert werden. Spätere Verbesserungen der Implementierung können lokal vorgenommen werden und dennoch profitiert die Gesamtimplementierung auf globaler Ebene. Benutzt beispielsweise eine Gesamtlösung in exzessiver Weise einen Nullstellenlöser, kommuniziert jedoch mit diesem nur über ein wohldefiniertes Interface, so lässt sich das Lösungsverfahren schnell gegen ein alternatives oder besseres austauschen.

Die Möglichkeit durch die Verwendung von gleichen Interfaces den Code auszutauschen ist nicht nur für die Weiterentwicklung und Optimierung der Gesamtlösung von Vorteil, sie bietet eine hervorragende Möglichkeit verschiedene lokale Lösungsansätze in ihrer Gesamtwirkung gegeneinander antreten zu lassen und zu testen.



Further Reading: Zum objektorientierten Design: Entwurfsmuster (*Design Patterns*) in [13], UML in [26].

<|

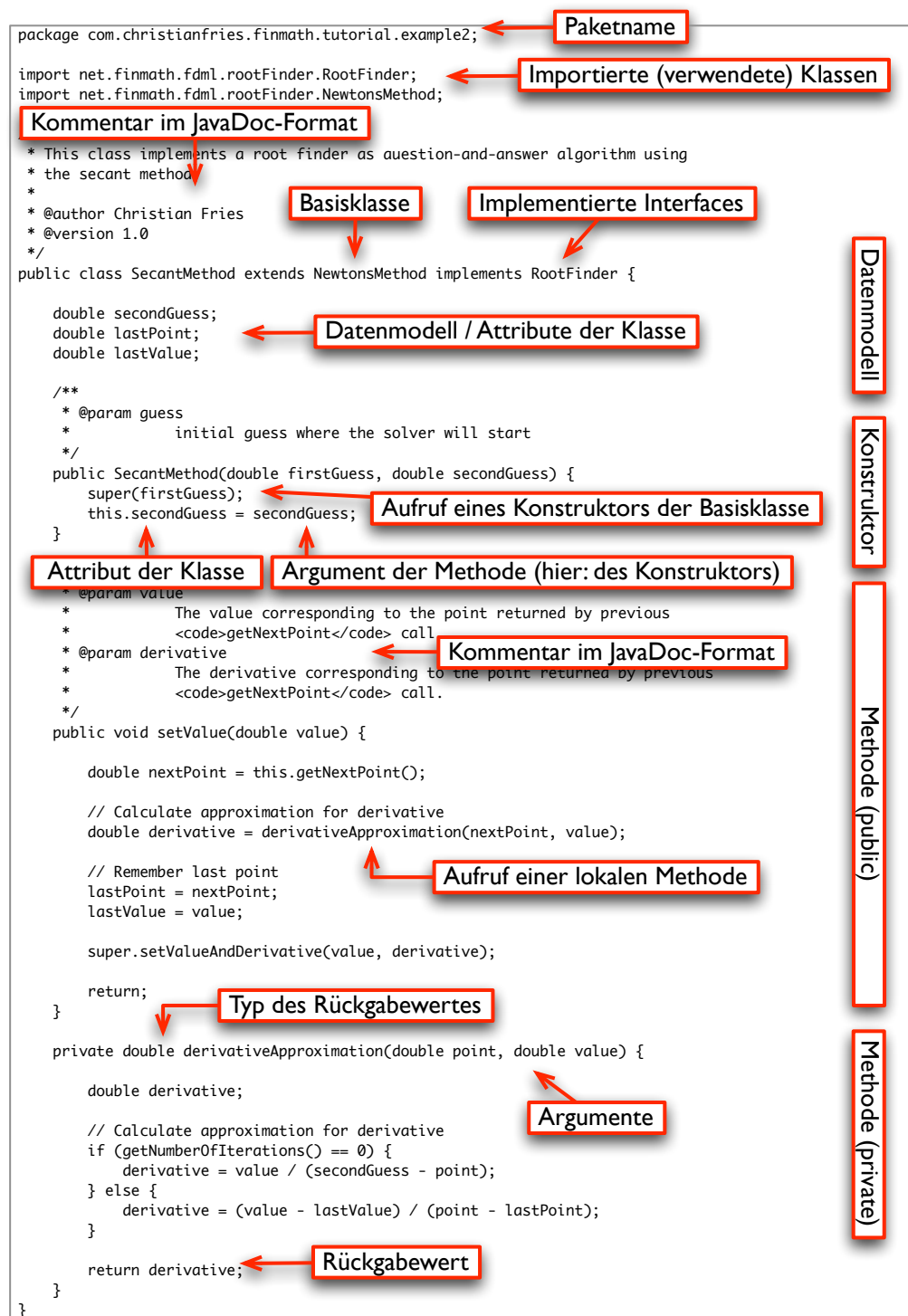


Figure 25.1.: Anatomie einer Java™ Klasse.

Part VIII.

Appendix

APPENDIX A

Tools (Selection)

An English version of this part is currently not available. What follows is the German version.

A.1. Cholesky Decomposition

Lemma 224 (Cholesky Zerlegung): Es sei A eine positiv definite, symmetrische Matrix. Dann existiert eine obere Dreiecksmatrix U mit

$$A = U^T U.$$

U heißt *Cholesky Zerlegung* von A und wird manchmal mit $\sqrt{A}$ bezeichnet.

A.2. Linear Regression

Lemma 225 (Linear Regression): Let $\Omega^* = \{\omega_1, \dots, \omega_n\}$ be a given sample space, $V : \Omega^* \rightarrow \mathbb{R}$ and $Y := (Y_1, \dots, Y_p) : \Omega^* \rightarrow \mathbb{R}^p$ given random variables. Furthermore let

$$f(y_1, \dots, y_p, \alpha_1, \dots, \alpha_p) := \sum \alpha_i y_i.$$

Then we have for any α^* with $X^T X \alpha^* = X^T v$

$$\|V - f(Y, \alpha^*)\|_{L_2(\Omega^*)} = \min_{\alpha} \|V - f(Y, \alpha)\|_{L_2(\Omega^*)},$$

where

$$X := \begin{pmatrix} Y_1(\omega_1) & \dots & Y_p(\omega_1) \\ \vdots & & \vdots \\ Y_1(\omega_n) & \dots & Y_p(\omega_n) \end{pmatrix}, \quad v := \begin{pmatrix} V(\omega_1) \\ \vdots \\ V(\omega_n) \end{pmatrix}.$$

If $(X^T X)^{-1}$ then $\alpha^* := (X^T X)^{-1} X^T v$. Dabei heißen $Y_1, \dots, Y_p$ *Basisfunktionen (explanatory variables)*.

Proof: Zu lösen ist das Minimierungsproblem

$$g(\alpha) := \|V - f(Y, \alpha)\|_{L_2(\Omega^*)}^2 = (v - X \cdot \alpha)^T \cdot (v - X \cdot \alpha) \rightarrow \min.$$

Die quadratische Funktion rechter Hand nimmt ihr Minimum dort an, wo die partiellen Ableitungen nach α_i verschwinden. Es ist

$$\frac{\partial g}{\partial \alpha} = 2X^T \cdot (v - X \cdot \alpha) = 2(X^T v - X^T X \cdot \alpha)$$

und somit

$$\frac{\partial g}{\partial \alpha} = 0 \Leftrightarrow X^T v = X^T X \cdot \alpha.$$

□

A.3. Generation of Random Numbers

Dieser Abschnitt soll eine kurze Starthilfe zur Erzeugung von (Pseudo-)Zufallszahlen darstellen und zeigen wie damit eigene Monte-Carlo Simulation erstellt werden können. Die zahlreichen Verfahren zur Erzeugung und die verschiedenen Aspekte hinsichtlich der Qualität von Zufallszahlen werden nicht adressiert. Die beschriebenen Methoden genügen jedoch für die meisten Anwendungen.

A.3.1. Uniform Distributed Random Variables

A.3.1.1. Mersenne Twister

Ein sehr populärer (aber auch sehr guter) Zufallszahlengenerator für $[0, 1]$ -gleichverteilte Zufallszahlen ist der Mersenne Twister (MT 19937). Der Zufallszahlengenerator hat eine Periodenlänge von $2^{19937} - 1$, d.h. die generierte Zahlenfolge wiederholt sich erst ab der 2^{19937} -ten Stelle. Er ist auch in sehr hohen Dimensionen (bis 623) gleichverteilt, siehe [77]. Basierend auf dem MT 19937 kann daher ein n -dimensionaler stochastischer Prozess erzeugt werden, indem in jedem Zeitschritt n aufeinanderfolgende Zufallszahlen zur Berechnung der jeweiligen Inkremente benutzt werden.¹

Viele Bibliotheken enthalten Implementierungen des MT 19937. Für die populärsten Sprachen ist er auch als Quellcode erhältlich.

A.3.2. Transformation of the Random Number Distribution via the Inverse Distributionfunction

Ist Z eine $[0, 1]$ -gleichverteilte Zufallsvariable und Φ eine Verteilungsfunktion, so ist $X := \Phi^{-1}(Z)$ eine Zufallsvariable die entsprechend der durch Φ beschriebenen Verteilung verteilt ist. Liegt ein Zufallszahlengenerator für gleichverteilte Zufallszahlen vor (z.B. der [Mersenne Twister](#)), so zieht man damit Realisierungen $Z(\omega_i)$, $i = 1, 2, \dots$ und erhält mit $\Phi^{-1}(Z(\omega_i))$ die Realisierungen von X . Es wird also lediglich eine Inverse der Verteilungsfunktion benötigt.

A.3.3. Normal Distributed Random Variables

A.3.3.1. Inverse Distribution Function

Die Dichte der Standardnormalverteilung ist $\phi(x) := \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$, die (kumulative) Verteilungsfunktion ist $\Phi(x) := \int_{-\infty}^x \phi(\xi) d\xi$. Der in [85] beschriebene Algorithmus liefert eine Approximation $\tilde{\Phi}^{-1}$ von Φ^{-1} mit einem relativen Fehler $\frac{|\tilde{\Phi}^{-1} - \Phi^{-1}|}{1 + |\Phi^{-1}|} < 10^{-15}$.

¹ Man beachte die Bemerkung am Ende von Abschnitt [A.3.5](#).

A.3.3.2. Box-Muller Transformation

Die Box-Muller Transformation vermag aus zwei unabhängigen $[0, 1]$ -gleichverteilten Zufallszahlen zwei unabhängige normalverteilte Zufallszahlen zu erzeugen. Sie lautet:

Lemma 226 (Box-Muller Transformation): Sind z_1 und z_2 zwei unabhängige $[0, 1]$ -gleichverteilte Zufallsvariablen, so sind

$$x_1 := \rho \cos(\theta), \quad x_2 := \rho \sin(\theta)$$

mit $\rho := \sqrt{-2 \log(z_1)}$ und $\theta := 2\pi z_2$ zwei unabhängige normalverteilte Zufallsvariablen mit Mittelwert 0 und Standardabweichung 1.

A.3.4. Poisson Distributed Random Variables

A.3.4.1. Inverse Distribution Function

Die (kumulative) Verteilungsfunktion der Poissonverteilung lautet $\Phi(\tau) := 1 - \exp(-\int_0^\tau \lambda(t)dt)$, wobei λ die *Intensität* bezeichnet. Ist λ konstant so gilt

$$\Phi^{-1}(z) = -\frac{\log(1-z)}{\lambda}.$$

Ist q die Wahrscheinlichkeit, dass ein Ereignis im Intervall $[T_1, T_2]$ liegt und λ konstant, so ist

$$\lambda = -\frac{\log(1-q)}{T_2 - T_1}.$$

A.3.5. Generation of Paths of an n -dimensional Brownian Motion

Es sei $T_0 < T_1 < \dots < T_m$ eine gegebene Zeitdiskretisierung. Es sollen an diesen Zeitpunkten die Realisierungen einer n -dimensionalen Brownschen Bewegung $W := (W_1, \dots, W_n)$ auf Pfaden $\omega_1, \dots, \omega_k$ erzeugt werden. Für einen Pfad müssen also $n \cdot m$ Zufallszahlen gezogen werden. Wir benutzen zur Generierung des $n \cdot m$ -Tupels zunächst den Mersenne Twister und transformieren dann geeignet.

Es sei $\{z_i\}_{i=1,2,\dots}$ die vom Mersenne Twister generierte Folge $[0, 1]$ -gleichverteilter Zufallszahlen. Dann ist $\{\Phi^{-1}(z_i)\}_{i=1,2,\dots}$ eine Folge von standardnormalverteilten Zufallszahlen. Ist nun `normalDistribution.nextDouble()` eine Methode welche bei jedem Aufruf das nächste Element der Folge $\{\Phi^{-1}(z_i)\}_{i=1,2,\dots}$ liefert, so wird die zeitdiskrete Brownsche Bewegung durch den Code in Listing A.1 erzeugt.

Listing A.1: Erzeugung einer n -dimensionalen Brownschen Bewegung

```
/**
 * This class represents a multidimensional brownian motion W = (W(1), ..., W(n))
 * where W(i), W(j) are uncorrelated for i not equal j.
 * Here the dimension n is called factors since this brownian motion is used to
 * generate multi-dimensional multi-factor Ito processes and there one might use
 * a different number of factors to generate Ito processes of different dimension.
```

```

*
* @author Christian Fries
* @version 1.1
*/
public class BrownianMotion {
    double[][][] brownianIncrement;

    double[]    timeDiscretization;
    int         numberOfFactors;
    int         numberOfPaths;
    int         seed;

    :

    private void doGenerateBrownianMotion() {
        // Create random number sequence generator
        AbstractDistribution normalDistribution = new Normal( 0,1, new MersenneTwister64(seed) );

        // Allocate memory
        brownianIncrement = new double[(timeDiscretization.length-1)][numberOfFactors][numberOfPaths];

        // Precalculate square roots of deltaT
        double[] sqrtOfTimeStep = new double[timeDiscretization.length-1];
        for(int timeIndex=0; timeIndex<timeDiscretization.length-1; timeIndex++) {
            sqrtOfTimeStep[timeIndex] = Math.sqrt(timeDiscretization[timeIndex+1]-timeDiscretization[timeIndex]);
        }

        // Set increments
        for(int path=0; path<numberOfPaths; path++) {
            for(int timeIndex=0; timeIndex<timeDiscretization.length-1; timeIndex++) {
                double sqrtDeltaT = sqrtOfTimeStep[timeIndex];
                // Generate uncorrelated Brownian increment
                for(int factor=0; factor<numberOfFactors; factor++) {
                    brownianIncrement[timeIndex][factor][path] = normalDistribution.nextDouble() * sqrtDeltaT;
                }
            }
        }
    }
}

```



Tip (Erzeugung von Pfaden zeitdiskreter stochastischer Prozesse): Der Mersenne Twister ist gleichverteilt in 623 Dimensionen. Wie weit diese Eigenschaft in höheren Dimensionen erhalten bleibt ist nicht sicher. Aus diesem Grund ist es ratsam die Zufallszahlen zur

Erzeugung der n Inkremente eines n -dimensionalen Prozesses nacheinander (also als n -Tupel) generieren zu lassen. Desweiteren besteht eine Anforderung an die temporale Unabhängigkeit der m Inkremente des stochastischen Prozesses, hier die Inkremente $\Delta W(T_i) := W(T_{i+1}) - W(T_i)$ der Brownschen Bewegung. Daher entspricht ein Pfad ω der Realisation einer $n \cdot m$ -dimensionalen Zufallsvariablen, hier

$$(\Delta W_j(T_i))_{j=1,\dots,n, i=1,\dots,m}.$$

Mit anderen Worten: Sind Pfade eines n -dimensionalen Prozesses mit m Zeitschritten zu generieren, so wird für jeden Pfad ein $n \cdot m$ -Tupel gezogen. Dies verlangt, dass der Zufallszahlengenerator die Verteilung in $n \cdot m$ Dimensionen erzeugt. Die hohe Dimensionsanzahl des Mersenne-Twister ist für unsere Anwendung also keines Falls übertrieben. Er eignet sich damit z.B. für einen 7-dimensionalen Prozess mit 89 Zeitschritten.²

² Es müssen nicht alle 623 Dimensionen ausgenutzt werden, in diesem Beispiel wäre es so.

Die Reihenfolge nach der im Listing A.1 die Schleifen geschachtelt sind ist also bewusst so gewählt. <|



Further Reading: Zur Erzeugung von Zufallszahlen insbesondere im Kontext von Monte-Carlo Simulationen und Derivatebewertung siehe [16]. Hintergrundinformationen zum Mersenne Twister sowie Verweise auf Quellcode und Bibliotheken finden sich im Wikipedia Artikel „Mersenne Twister“, http://de.wikipedia.org/wiki/Mersenne_Twister. <|

A.4. Generation of Correlated Brownian Motion

Lemma 227 (Faktorzerlegung): Sei $R = (\rho_{i,j})_{i,j=1\dots n}$ eine gegebene Korrelationsmatrix. Somit ist R symmetrisch und positiv semi-definit, woraus folgt, dass R nur reelle Eigenwerte $\lambda_1 \geq \dots \geq \lambda_n \geq 0$ besitzt und eine ortho-normalbasis aus zugehörigen Eigenvektoren $v_1, \dots, v_n$ von R existiert, d.h.

$$\exists V : \quad V^T R V = D := \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \quad \text{wobei } V = (v_1, \dots, v_n).$$

und $R = V D V^T$ sowie $V^T V = I$.

Seien $U_1, \dots, U_n$ unabhängige Brownsche Bewegungen, $U := (U_1, \dots, U_n)^T$. Dann ist W mit

$$dW := (dW_1, \dots, dW_n)^T := V \sqrt{D} dU$$

eine n -dimensionale Brownsche Bewegung mit

$$\langle dW_i, dW_j \rangle = \rho_{i,j} dt.$$

Mit $F := (v_1 \sqrt{\lambda_1}, \dots, v_n \sqrt{\lambda_n})$ ist also $dW = F dU$.

Proof: Eine Korrelationsmatrix ist offensichtlich symmetrisch, somit besitzt die Matrix nur reelle Eigenwerte. Ist R die Korrelationsmatrix des Zufallsvektors $X = (X_1, \dots, X_n)^T$ mit $\text{Var}(X_i) = 1$ und $E(X_i) = 0$, so ist $R = E(X \cdot X^T)$. If v_i der Eigenvektor zum Eigenwert λ_i , so ist

$$E(\|X^T v_i\|_{l_2}^2) = E(v_i^T X \cdot X^T v_i) = v_i^T R v_i = \lambda_i \|v_i\|_{l_2}^2$$

und damit

$$\lambda_i = \frac{E(\|X^T v_i\|_{l_2}^2)}{\|v_i\|_{l_2}^2} \geq 0.$$

Damit ist R also positiv semi-definit und $dW := V \sqrt{D} dU$ wohl definiert. $\square$

A.5. Factor Reduction

Ausgehend von der in A.4 besprochenen Konstruktion korrelierter Brownscher Bewegungen ist es möglich, die Anzahl der relevanten Faktoren (d.h. der von Null verschiedenen Eigenwerte) zu reduzieren, ohne die ursprüngliche Korrelationsstruktur allzu stark zu verändern. Es seien R, V, D entsprechend Abschnitt A.4, sowie $m < n$. Mit

$$(f_1, \dots, f_n) = F = V \sqrt{D}, \quad f_i = (f_{j,i})_{j=1}^n$$

sei nun

$$F^r = (f_i^r)_{i=1, \dots, m} = (f_{j,i}^r)_{\substack{j=1, \dots, n \\ i=1, \dots, m}}, \quad f_{j,i}^r := \frac{f_{j,i}}{(\sum_k f_{j,k}^2)^{1/2}},$$

d.h. die $n \times m$ -Matrix F^r entsteht aus der $n \times m$ -Matrix $(v_1 \sqrt{\lambda_1}, \dots, v_m \sqrt{\lambda_m})$ durch Normierung der n Zeilen.

Seien $U_1, \dots, U_m$ unabhängige Brownsche Bewegungen, $U := (U_1, \dots, U_m)^T$. Dann ist W mit

$$dW := (dW_1, \dots, dW_n)^T := F^r dU$$

eine n -dimensionale m -faktorielle Brownsche Bewegung.

Die vorgestellte *Faktorreduktion* entspricht einer Hauptkomponentenanalyse (*Principal Component Analysis*) mit nachfolgender Renormierung der Komponenten.

Remark 228 (Faktorreduktion): Die Größe des Absolutbetrages des Eigenwertes λ_i gibt das Gewicht bezüglich der Bedeutung des zugehörigen Faktors f_i an und kann benutzt werden, um über die Anzahl der verbleibenden Faktoren zu entscheiden. Ein Beispiel gibt hier der Grenzfall perfekter Korrelation $\rho_{i,j} = 1$. Die entsprechende Korrelationsmatrix besitzt einen einfachen Eigenwert n mit Eigenvektor $(1, \dots, 1)$ und einen $n - 1$ -fachen Eigenwert 0 , was dahingehend zu interpretieren ist, dass lediglich eine eindimensionale Brownsche Bewegung zur Erklärung der Dynamik benötigt wird (die Parallelbewegung).

In Abbildung A.1 ist für einen nicht-perfekt, aber dennoch hoch korrelierten Fall $\rho_{i,j} = \exp(-0.005 * |i - j|)$ eine Reduktion auf die ersten drei Faktoren gezeigt. Werden hingegen viele Faktoren mit relativ hohem Gewicht vernachlässigt, so führt die Faktorreduktion auch zu einer beachtlichen Änderung der Korrelationsstruktur (vergleiche Abbildung A.3), sowie der verbleibenden Faktoren (vergleiche Abbildung A.2).



Experiment: Die Auswirkung einer Faktorreduktion auf die Korrelationsmatrix kann für verschiedene Korrelationsstrukturen unter <http://www.christian-fries.de/finmath/applets/FactorReduction.html> nachvollzogen werden. <|

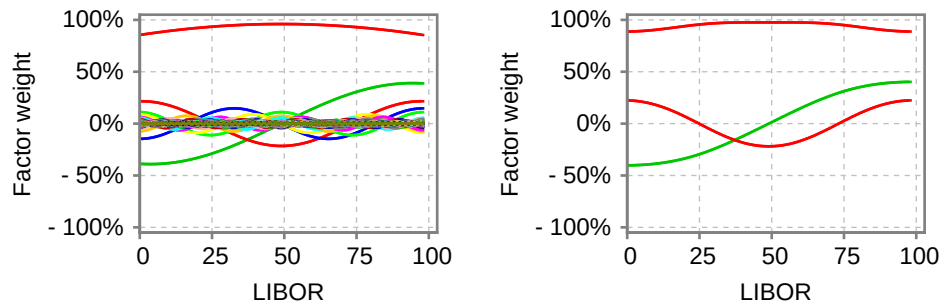


Figure A.1.: Faktorreduktion im Falle hoher Korrelation: Die Faktoren f_i (Eigenvektoren) der Korrelationsmatrix $\rho_{i,j} = \exp(-0.005 * |i - j|)$ (links) und eine Reduktion auf die drei Faktoren mit den größten Eigenwerten (rechts).

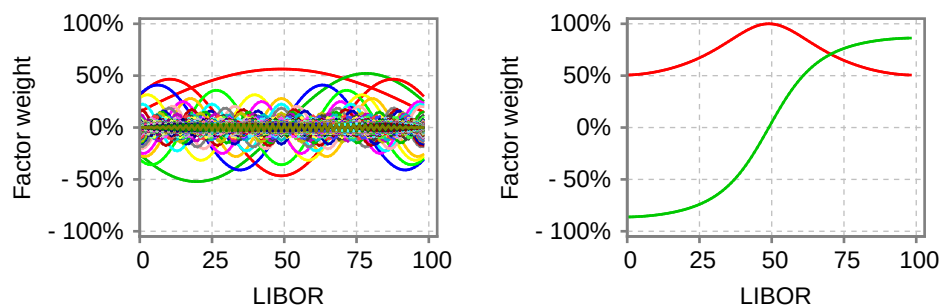


Figure A.2.: Faktorreduktion im Falle niedriger Korrelation: Die Faktoren f_i (Eigenvektoren) der Korrelationsmatrix $\rho_{i,j} = \exp(-0.1 * |i - j|)$ (links) und eine Reduktion auf die zwei Faktoren mit den größten Eigenwerten (rechts).

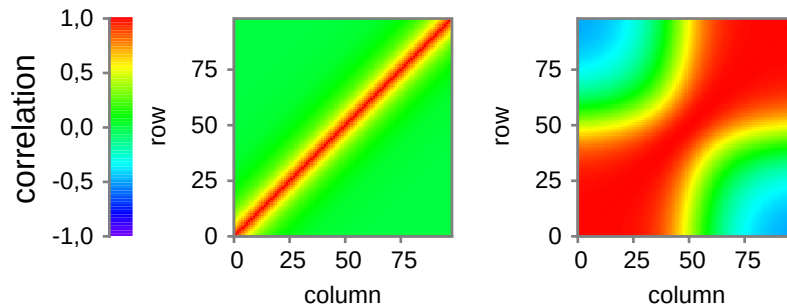


Figure A.3.: Faktorreduktion im Falle niedriger Korrelation: Die ursprüngliche Korrelationsmatrix $\rho_{i,j} = \exp(-0.1 * |i - j|)$ (links) und die der Reduktion auf zwei Faktoren entsprechende Korrelationsmatrix (rechts). Dies entspricht der in Abbildung A.2 dargestellten Faktorreduktion.

A.6. Optimization (one dimensional): Golden Section Search

Gegeben sei eine Funktion $f : [a, b] \rightarrow \mathbb{R}$. Es sei ferner $\lambda \in (0, 1)$ und für $m_0 = \lambda a + (1 - \lambda)b$ sei $f(m_0) < \min\{f(a), f(b)\}$. Dann konvergiert die durch folgenden Algorithmus definierte Folge $\{m_i\}_{i=0}^{\infty}$ gegen ein lokales Minimum von f (und somit das globale Minimum, falls f auf $[a, b]$ strikt konvex ist):

Iterationsanfang:

$$a_0 := a, \quad m_0 := \lambda a + (1 - \lambda)b, \quad b_0 := b$$

Iterationsschritt:

- Falls $b_i - m_i > m_i - a_i$ so setze $z := \lambda m_i + (1 - \lambda)b_i$ und
 Falls $f(m_i) < f(z)$: $\begin{cases} a_{i+1} := a_i \\ m_{i+1} := m_i \\ b_{i+1} := z \end{cases}$ sonst: $\begin{cases} a_{i+1} := m_i \\ m_{i+1} := z \\ b_{i+1} := b_i \end{cases}$.
- Falls $b_i - m_i \leq m_i - a_i$ so setze $z := \lambda a_i + (1 - \lambda)m_i$ und
 falls $f(z) < f(m_i)$: $\begin{cases} a_{i+1} := a_i \\ m_{i+1} := z \\ b_{i+1} := m_i \end{cases}$ sonst: $\begin{cases} a_{i+1} := z \\ m_{i+1} := m_i \\ b_{i+1} := b_i \end{cases}$.

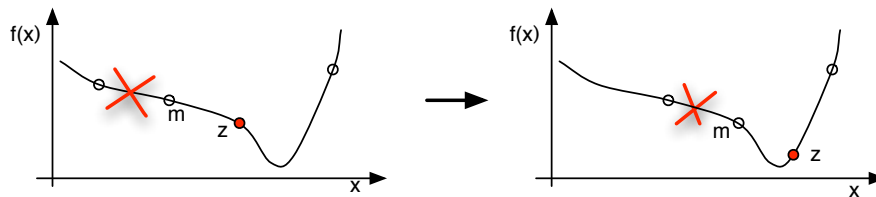


Figure A.4.: Golden Section Search

Der Algorithmus setzt einen Punkt (z) in das grössere der beiden Intervalle $[a, m]$, $[m, b]$ und verwirft von den drei entstandenen Intervallen dann jenes Intervall, welches an den grösseren der Werte $f(m)$, $f(z)$ grenzt.

Für das Teilungsverhältnis ist der Wert

$$\lambda = \frac{3 - \sqrt{5}}{2}$$

im folgenden Sinne optimal: Im ungünstigsten Fall, in welchem der Algorithmus in jedem Schritt das jeweils kleinere Randintervall verwirft, führt der Wert $\lambda = \frac{3 - \sqrt{5}}{2}$ zu schnellsten Konvergenzrate. Da dieses Teilungsverhältnis dem *Goldenen Schnitt* entspricht, wird der Algorithmus als *Golden Section Search* bezeichnet.

A.7. Convolution with the Normal Density

Lemma 229 (Integration von $\exp(a \cdot X)$, X normalverteilt): Es ist

$$\int_{h_1}^{h_2} e^{a \cdot x} \cdot \phi(x; y, \sigma) dx = \left[\Phi\left(\frac{h_2 - (y + a\sigma^2)}{\sigma}\right) - \Phi\left(\frac{h_1 - (y + a\sigma^2)}{\sigma}\right) \right] e^{ay + \frac{a^2\sigma^2}{2}},$$

wobei $\phi(\xi; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(\xi-\mu)^2}{2\sigma^2}}$ die Dichte und $\Phi(x) := \int_{-\infty}^x \phi(\xi; 0, 1) d\xi$ die Verteilungsfunktion der Normalverteilung bezeichne. Insbesondere gilt:

$$\int_{-\infty}^{\infty} e^{a \cdot x} \cdot \phi(x; y, \sigma) dx = e^{ay + \frac{a^2\sigma^2}{2}}.$$

Proof: Es ist

$$\begin{aligned} \int_{h_1}^{h_2} e^{a \cdot x} \cdot \phi(x; y, \sigma) dx &= \frac{1}{\sqrt{2\pi}\sigma} \int_{h_1}^{h_2} e^{a \cdot x} \cdot e^{-\frac{1}{2}\left(\frac{x-y}{\sigma}\right)^2} dx \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{h_1}^{h_2} \exp\left(a \cdot x - \frac{1}{2}\left(\frac{x-y}{\sigma}\right)^2\right) dx. \end{aligned}$$

Da

$$\begin{aligned} &\exp\left(a \cdot x - \frac{1}{2}\left(\frac{x-y}{\sigma}\right)^2\right) \\ &= \exp\left(a \cdot x - \frac{1}{2\sigma^2}(x^2 - 2x \cdot y + y^2)\right) \\ &= \exp\left(-\frac{1}{2\sigma^2}(x^2 - 2x \cdot y + y^2 - 2a\sigma^2 \cdot x)\right) \\ &= \exp\left(-\frac{1}{2\sigma^2}(x^2 - 2x(y + a\sigma^2) + y^2)\right) \\ &= \exp\left(-\frac{1}{2\sigma^2}(x^2 - 2x(y + a\sigma^2) + (y + a\sigma^2)^2 - 2a\sigma^2 y - a^2\sigma^4)\right) \\ &= \exp\left(-\frac{1}{2\sigma^2}(x^2 - 2x(y + a\sigma^2) + (y + a\sigma^2)^2)\right) \cdot \exp\left(ay + \frac{a^2\sigma^2}{2}\right) \\ &= \exp\left(-\frac{(x - (y + a\sigma^2))^2}{2\sigma^2}\right) \cdot \exp\left(ay + \frac{a^2\sigma^2}{2}\right), \end{aligned}$$

folgt

$$\begin{aligned} &\int_{h_1}^{h_2} e^{a \cdot x} \cdot \phi(x; y, \sigma) dx \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{h_1}^{h_2} \exp\left(-\frac{(x - (y + a\sigma^2))^2}{2\sigma^2}\right) dx \cdot \exp\left(ay + \frac{a^2\sigma^2}{2}\right) \\ &= \int_{h_1}^{h_2} \phi(x; y + a\sigma^2, \sigma) dx \cdot \exp\left(ay + \frac{a^2\sigma^2}{2}\right). \end{aligned}$$



APPENDIX B

Exercises

An English version of this part is currently not available. What follows is the German version.

Exercise 1 (Wahrscheinlichkeitsraum, Zufallsvariable [15 Punkte]): Es sei $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum und $X : \Omega \rightarrow \mathbb{R}$ eine Zufallsvariable.

1. Geben sie ein Beispiel für eine (falsche) Modellierung $\Omega, \mathcal{F}, X$ eines Zufallsexperimentes, so dass X nicht $(\mathcal{F}, \mathcal{B}(\mathbb{R}))$ -messbar ist (geben Sie die Definition der mathematischen Objekte und deren Interpretation).
2. Es sei X nun $(\mathcal{F}, \mathcal{B}(\mathbb{R}))$ -messbar.

- a) Zeigen Sie, dass

$$\{X^{-1}(A) \mid A \in \mathcal{B}(\mathbb{R})\}$$

eine σ -Algebra und Teilmenge von $\mathcal{F}$ ist (Sub- σ -Algebra von $\mathcal{F}$). Wie kann diese interpretiert werden?

- b) Zeigen Sie, dass

$$P_X(A) := P(X^{-1}(A)) \quad \forall A \in \mathcal{B}(\mathbb{R})$$

ein Wahrscheinlichkeitsmaß definiert (Bildmaß).

Exercise 2 (Bedingte Erwartung [20 Punkte])¹: Es sei X eine $\mathcal{F}$ -messbare (numerische) Zufallsvariable und $\mathcal{G}$ eine σ -Algebra mit $\mathcal{G} \subset \mathcal{F}$ (d.h. $\mathcal{G}$ ist Sub- σ -Algebra von $\mathcal{F}$). Beweisen Sie die folgenden Eigenschaften der bedingten Erwartung:

1. Ist X eine $\mathcal{G}$ -messbare Zufallsvariable, so ist $E(X|\mathcal{G}) = E(X)$ (P -fast sicher).
2. Ist $X \geq 0$, so ist $E(X|\mathcal{G}) \geq 0$.

¹ Siehe [39].

3. (*Tower Law*) Ist $\mathcal{H}$ eine σ -Algebra mit $\mathcal{H} \subset \mathcal{G}$, so gilt $E(E(X|\mathcal{G})|\mathcal{H}) = E(X|\mathcal{H})$.
4. (*Taking out what is known*) Ist Z eine beschränkte, $\mathcal{G}$ -messbare Zufallsvariable, so ist $E(ZX|\mathcal{G}) = ZE(X|\mathcal{G})$.

Exercise 3 (Verteilungsfunktion [10 Punkte])²: Es sei $X : \Omega \rightarrow \mathbb{R}$ eine Zufallsvariable und F die Verteilungsfunktion von X , d.h. $F(x) := P_X((-\infty, x])$. Zeigen Sie:

1. F ist links-stetig, d.h. $F(x) = \lim_{h \searrow 0} F(x+h)$.
2. Ist $g : \mathbb{R} \rightarrow \mathbb{R}$ messbar mit $E(|g(X)|) < \infty$, so ist

$$E(g(X)) = \int_{-\infty}^{\infty} g(x) dF(x),$$

wobei das Integral als Lebesgue-Stieltjes Integral interpretiert wird (siehe [25]).

Exercise 4 (Brownian scaling [10 Punkte])³: Es sei W eine Brownsche Bewegung und $c > 0$. Zeigen Sie, dass mit

$$\tilde{W}(t) := \frac{1}{c} W(c^2 t)$$

$\tilde{W}$ ebenfalls eine Brownsche Bewegung ist. Sind die beiden Prozesse W und $\tilde{W}$ in irgendeinem Sinne (vergleiche Definition ??) gleich?

Exercise 5 (Quadratische Variation [10 Punkte])⁴: Es sei X ein stetiger⁵ stochastischer Prozess. Für $p > 0$ sei der p -te Variations Prozess definiert als

$$\langle X, X \rangle^{(p)}(t, \omega) := \lim_{\Delta t \rightarrow 0} \sum_{t_k \leq t} |X(t_{k+1}, \omega) - X(t_k, \omega)|^p,$$

wobei $\{t_i\}_{i=1}^{\infty}$ eine streng monotone Folge mit $t_0 = 0$, $\lim_{k \rightarrow \infty} t_k = \infty$, $\Delta t := \sup_k |t_{k+1} - t_k|$ ist. Speziell heißt der Prozess $\langle X, X \rangle^{(1)}$ die *Totale Variation* von X und der Prozess $\langle X, X \rangle^{(2)}$ die *Quadratische Variation* von X .

Es sei W eine (eindimensionale) Brownsche Bewegung.

1. Zeigen Sie, dass

$$\langle W, W \rangle(t, \omega) = t \quad P\text{-fast sicher}$$

gilt.

² Siehe [25].

³ Siehe [25].

⁴ Siehe [25].

⁵ Unter einem stetigen stochastischen Prozess verstehen wir einen stochastischen Prozess, dessen Pfade stetige Funktionen sind.

2. Zeigen Sie, dass

$$\langle W, W \rangle^1(t, \omega) = \infty \quad P\text{-fast sicher}$$

gilt.

Beachten Sie, dass diese Eigenschaften nicht etwa im Mittel über alle Pfade gelten, sondern („punktweise“) für fast alle Pfade!

Exercise 6 (Itô Integral [10 Punkte])⁶: Zeigen Sie durch direkte Anwendung der Definition des Itô Integrals, dass

1. $\int_0^T t \, dW(t) = T \cdot W(T) - \int_0^T W(t) \, dt,$
2. $\int_0^T W(t)^2 \, dW(t) = \frac{1}{3} W(T)^3 - \int_0^T W(t) \, dW(t).$

Exercise 7 (Stratonovich Integral [10 Punkte])⁷: Es sei $T^{(n)} := \{t_0, \dots, t_n\}$ mit $0 = t_0 < t_1 < \dots < t_n = T$ eine Zerlegung des Intervalls $[0, T]$ und $\Delta T^{(n)} := \sup_i |t_i - t_{i-1}|$ die Feinheit der Zerlegung. Es sei ferner $f : [0, T] \times \Omega \rightarrow \mathbb{R}$ aus der Klasse der Integranden des Itô Integrals (auf $[0, T]$) und $t \mapsto f(t, \omega)$ stetig für (fast) alle ω . Dann ist

$$\int_0^T f(t, \omega) \, dW(t, \omega) = \lim_{\Delta T^{(n)} \rightarrow 0} \sum_{j=1}^n f(t_{j-1}, \omega) (W(t_j, \omega) - W(t_{j-1}, \omega))$$

in $L_2(P)$ (Beweis?). Das Stratonovich Integral wird entsprechend definiert durch

$$\int_0^T f(t, \omega) \circ dW(t, \omega) = \lim_{\Delta T^{(n)} \rightarrow 0} \sum_{j=1}^n f(t_{j-\frac{1}{2}}, \omega) (W(t_j, \omega) - W(t_{j-1}, \omega))$$

mit $t_{j-\frac{1}{2}} := \frac{t_{j-1} + t_j}{2}$. Berechnen Sie

1. $\int_0^T W(t, \omega) \circ dW(t, \omega)$ und
2. $\int_0^T W(t, \omega) \circ dW(t, \omega) - \int_0^T W(t, \omega) \, dW(t, \omega).$

Lösung: Es ist

$$\int_0^T W(t, \omega) \circ dW(t, \omega) = \lim_{\Delta T^{(n)} \rightarrow 0} \sum_{j=1}^n W(t_{j-\frac{1}{2}}, \omega) (W(t_j, \omega) - W(t_{j-1}, \omega))$$

⁶ Siehe [25].

⁷ Siehe [25].

und

$$\begin{aligned}
& W(t_{j-\frac{1}{2}}, \omega)(W(t_j, \omega) - W(t_{j-1}, \omega)) \\
= & W(t_{j-\frac{1}{2}}, \omega)(W(t_j, \omega) - W(t_{j-\frac{1}{2}}, \omega)) + W(t_{j-\frac{1}{2}}, \omega)(W(t_{j-\frac{1}{2}}, \omega) - W(t_{j-1}, \omega)) \\
= & \frac{1}{2}(W(t_j, \omega) + W(t_{j-\frac{1}{2}}, \omega))(W(t_j, \omega) - W(t_{j-\frac{1}{2}}, \omega)) \\
& - \frac{1}{2}(W(t_j, \omega) - W(t_{j-\frac{1}{2}}, \omega))(W(t_j, \omega) - W(t_{j-\frac{1}{2}}, \omega)) \\
& + \frac{1}{2}(W(t_{j-\frac{1}{2}}, \omega) - W(t_{j-1}, \omega))(W(t_{j-\frac{1}{2}}, \omega) - W(t_{j-1}, \omega)) \\
& + \frac{1}{2}(W(t_{j-\frac{1}{2}}, \omega) + W(t_{j-1}, \omega))(W(t_{j-\frac{1}{2}}, \omega) - W(t_{j-1}, \omega)) \\
= & \frac{1}{2}(W(t_j, \omega)^2 - \cancel{W(t_{j-\frac{1}{2}}, \omega)^2}) + \frac{1}{2}(\cancel{W(t_{j-\frac{1}{2}}, \omega)^2} - W(t_{j-1}, \omega)^2) \\
& - \frac{1}{2}(W(t_j, \omega) - W(t_{j-\frac{1}{2}}, \omega))(W(t_j, \omega) - W(t_{j-\frac{1}{2}}, \omega)) \\
& + \frac{1}{2}(W(t_{j-\frac{1}{2}}, \omega) - W(t_{j-1}, \omega))(W(t_{j-\frac{1}{2}}, \omega) - W(t_{j-1}, \omega)) \\
= & \frac{1}{2}(W(t_j, \omega)^2 - W(t_{j-1}, \omega)^2) \\
& - \frac{1}{2}(W(t_j, \omega) - W(t_{j-\frac{1}{2}}, \omega))^2 + \frac{1}{2}(W(t_{j-\frac{1}{2}}, \omega) - W(t_{j-1}, \omega))^2.
\end{aligned}$$

Und da

$$\lim_{\Delta T^{(n)} \rightarrow 0} \sum_{j=1}^n (W(t_j, \omega) - W(t_{j-\frac{1}{2}}, \omega))^2 = \lim_{\Delta T^{(n)} \rightarrow 0} \sum_{j=1}^n (W(t_{j-\frac{1}{2}}, \omega) - W(t_{j-1}, \omega))^2$$

folgt

$$\int_0^T W(t, \omega) \circ dW(t, \omega) = \frac{1}{2} W(t_n, \omega)^2.$$

Exercise 8 Itô Produktregel, Itô Quotientenregel [15 Punkte]: Benutzen Sie die Itô Formel und beweisen Sie

1. Die Produktregel: Es seien X und Y Itô Prozesse. Dann gilt

$$d(X \cdot Y) = Y \cdot dX + X \cdot dY + dX \cdot dY$$

2. Die Quotientenregel: Es seien X und Y Itô Prozesse, $Y > c$ für ein $c \in \mathbb{R}$. Dann gilt

$$d\left(\frac{X}{Y}\right) = \frac{X}{Y} \left(\frac{dX}{X} - \frac{dY}{Y} - \frac{dX}{X} \cdot \frac{dY}{Y} + \left(\frac{dY}{Y}\right)^2 \right)$$

3. Das *Drift-adjustment* eines lognormal Prozesses: Es sei $S(t) > 0$ ein Itô Prozess der Form

$$dS(t) = \mu(t)S(t)dt + \sigma(t)S(t)dW(t),$$

sowie $Y(t) := \log(S(t))$. Dann ist

$$dY(t) = (\mu(t) - \frac{1}{2}\sigma^2(t))dt + \sigma(t)dW(t).$$

Exercise 9 (Martingal: Itô Formel [15 Punkte])⁸: Benutzen Sie die Itô Formel und zeigen Sie, dass folgende Prozesse $\mathcal{F}_t$ -martingale sind:

1. $X(t) = \exp(\frac{1}{2}t) \cos(W(t))$
2. $X(t) = \exp(\frac{1}{2}t) \sin(W(t))$
3. $X(t) = (t + W(t)) \exp(-\frac{1}{2}t - W(t))$

($W(t)$ sei eine eindimensionale Brownsche Bewegung).

Exercise 10 (Martingal: Girsanov Theorem [x Punkte])⁹:

Exercise 11 (Black-Scholes Partielle Differentialgleichung in den Koordinaten $t, N(t), S(t)$ [20 Punkte]): Man zeige, dass die Funktion

$$V(t, n, s) = s\Phi(d_+) - n\frac{K}{N(T)}\Phi(d_-)$$

mit

$$d_{\pm} = \frac{1}{\sigma \sqrt{T-t}} \left[\log\left(\frac{s}{n} \cdot \frac{N(T)}{K}\right) \pm \frac{\sigma^2(T-t)}{2} \right],$$

und $\Phi'(x) = \phi(x) = C \exp(-x^2/2)$ die partielle Differentialgleichung

$$\frac{\partial V(t)}{\partial t} dt + \frac{1}{2} \frac{\partial^2 V(t)}{\partial s \partial s} \sigma^2 = 0$$

löst und die *Endbedingung*

$$V(T, N(T), s) = \max(s - K, 0)$$

erfüllt.

Exercise 12 (Black 76 Formula for Swaption [30 Punkte]): Gegeben seien Zeiten $T_1 < \dots < T_n$. Es sei V_{swap} ein Swap entsprechend Definition 106 mit konstanten

⁸ Siehe [25].

⁹ Siehe [25].

Swaprates $S_i = K$, $i = 1, \dots, n-1$. ($T_1, \dots, T_{n-1}$ seien *Fixing-Zeiten* und $T_2, \dots, T_n$ seien *Payment-Zeiten*). Es sei S die zugehörige *par swap rate* entsprechend Definition 111 (vergleiche Bemerkung 110). Leiten Sie eine Bewertungsformel für den Wert V_{swaption} einer (*europäischen*) Option (mit Ausübezeitpunkt T_1) auf V_{swap} her, unter der Annahme, dass S einer lognormalen Dynamik

$$dS(t) = \mu(t)S(t)dt + \sigma(t)S(t)dW(t) \quad \text{unter } \mathbb{P}$$

folgt.

Hinweis: Schreiben Sie zunächst den Wert des Swaps V_{swap} in Abhängigkeit von S und K , durch Umformung der Zahlungsströme: $L(T_i, T_{i+1}) - K = (L(T_i, T_{i+1}) - S) + (S - K)$, beachten Sie die Definition von S . Betrachten Sie dann den Wert des Swaps zum Ausübezeitpunkt der Option, d.h. $V_{\text{swap}}(T_1)$ und überlegen Sie sich eine geeignete Wahl eines Numéraires (vergleiche Bewertung eines Caplets). Begründen Sie kurz, dass der von Ihnen gewählte Numéraire ein gehandeltes Produkt ist.

Lösung: Gemäß Definition 106 zahlt ein Swap

$$(L(T_i, T_{i+1}; T_i) - S_i) \cdot (T_{i+1} - T_i) \quad \text{in } T_{i+1}. \quad (\text{B.1})$$

Es sei nun $t \leq T_1$. Der Wert der Zahlung (B.1) in t ist

$$(L(T_i, T_{i+1}; T_i) - S_i) \cdot (T_{i+1} - T_i) \cdot P(T_{i+1}; t)$$

und somit ist der Wert des Swaps in t gegeben durch

$$V_{\text{swap}}(t) = \sum_{i=1}^{n-1} (L(T_i, T_{i+1}; T_i) - S_i) \cdot (T_{i+1} - T_i) \cdot P(T_{i+1}; t).$$

Nach Definition 111 ist die *Par Swap Rate* gegeben durch

$$S_{\text{par}}(t) := \frac{P(T_1; t) - P(T_n; t)}{\sum_{i=1}^{n-1} (T_{i+1} - T_i) \cdot P(T_{i+1}; t)},$$

so dass

$$\sum_{i=1}^{n-1} (L(T_i, T_{i+1}; T_i) - S_{\text{par}}(t)) \cdot (T_{i+1} - T_i) \cdot P(T_{i+1}; t) = 0.$$

Damit folgt

$$\begin{aligned}
 V_{\text{swap}}(t) &= \sum_{i=1}^{n-1} (L(T_i, T_{i+1}; T_i) - S_i) \cdot (T_{i+1} - T_i) \cdot P(T_{i+1}; t) \\
 &= \sum_{i=1}^{n-1} (L(T_i, T_{i+1}; T_i) - S_{\text{par}}(t)) \cdot (T_{i+1} - T_i) \cdot P(T_{i+1}; t) \\
 &\quad + \sum_{i=1}^{n-1} (S_{\text{par}}(t) - S_i) \cdot (T_{i+1} - T_i) \cdot P(T_{i+1}; t) \\
 &= \sum_{i=1}^{n-1} (S_{\text{par}}(t) - S_i) \cdot (T_{i+1} - T_i) \cdot P(T_{i+1}; t).
 \end{aligned}$$

Für einen Swap mit $S_i = K$ gilt also

$$\begin{aligned}
 V_{\text{swap}}(t) &= (S_{\text{par}}(t) - K) \cdot \sum_{i=1}^{n-1} (T_{i+1} - T_i) \cdot P(T_{i+1}; t) \\
 &= (S_{\text{par}}(t) - K) \cdot A(t),
 \end{aligned}$$

wobei $A(t) := \sum_{i=1}^{n-1} (T_{i+1} - T_i) \cdot P(T_{i+1}; t)$ als *Swap Annuity* bezeichnet wird. Da $A(t) > 0$ folgt für den Wert der Option auf diesen Swap

$$V_{\text{swaption}}(t) = \max(S_{\text{par}}(t) - K, 0) \cdot A(t).$$

Man wähle nun $A(t)$ als Numeraire. Unter dem entsprechenden Martingalmaß Q^A ist S ein A -relativer Preis, also ein Martingal und folglich $dS(t) = \sigma(t)S(t)dW^{Q^A}(t)$. Analog zur Black Formel für ein Caplet leitet sich nun eine entsprechende Bewertungsformel für eine Swaption ab.

APPENDIX C

List of Symbols

An English version of this part is currently not available. What follows is the German version.

section under construction.

Symbol	Interpretation
$\mathbf{1}$	Indikatorfunktion. $\mathbf{1}(x) = \begin{cases} 1 & \text{für } x > 0, \\ 0 & \text{sonst.} \end{cases}$
$\mathbf{1}_{(a,b]}(x)$	Indikatorfunktion. $\mathbf{1}_{(a,b]}(x) = \begin{cases} 1 & \text{für } x \in (a, b], \\ 0 & \text{sonst.} \end{cases}$
$\mathcal{N}(\mu, \sigma^2)$	Normalverteilung mit Mittelwert μ und Varianz σ^2 .
$\mathbb{P}$	Reales Maß.
$\mathbb{Q}^N$	Martingalmaß zum Numéraire N . N -relative Preisprozesse $\frac{V}{N}$ gehandelter Produkte V sind $\mathbb{Q}^N$ -Martingale. Existiert (als zu $\mathbb{P}$ äquivalentes Maß) unter Voraussetzungen.
$[x]$	Gauß-Klammer. Größte ganze Zahl, die kleiner oder gleich x ist. $[x] := \max\{n \in \mathbb{N} \mid n \leq x\}$

APPENDIX D

Java™ Source Code (Selection)

D.1. Java™ Classes for Chapter 25

```
package com.christianfries.finmath.tutorial.example1;

/**
 * @author Christian Fries
 * @version 1.0
 */
class BinomialDistributedRandomVariable {

    private double value1;
    private double value2;
    private double probabilityOfState1;

    /**
     * This class implements a binominal distributed random variable
     * @param value1 The value in state 1.
     * @param value2 The value in state 2.
     * @param probabilityOfState1 The probability of state 1.
     */
    public BinomialDistributedRandomVariable(double value1, double value2,
        double probabilityOfState1) {
        this.value1 = value1;
        this.value2 = value2;
        this.probabilityOfState1 = probabilityOfState1;
    }

    public static void main( String args[] ) {
        /**
         * Test of class BinominalDistributedRandomVariable
         */

        System.out.println("Creating_random_variable.");
        BinomialDistributedRandomVariable randomVariable =
            new BinomialDistributedRandomVariable(1.0, -1.0, 0.5);    // Create object

        double expectation = randomVariable.getExpectation();
        System.out.println("Expectation_is:_" + expectation);

        System.out.println("Changing_value_of_state_1_to_0.4.");
        randomVariable.setValue1(0.4);

        expectation = randomVariable.getExpectation();
        System.out.println("Expectation_is:_" + expectation);
    }

    /**
     * @param value1 The value of state 1.
     */
    public void setValue1(double value1) {
        this.value1 = value1;
    }
}
```

```

    */@return Returns the probability of state 1.
    */
    public double getProbabilityOfState1() {
        return probabilityOfState1;
    }

    /**
     * @param probabilityOfState1 The probability of state 1.
     */
    public void setProbabilityOfState1(double probabilityOfState1) {
        this.probabilityOfState1 = probabilityOfState1;
    }

    /**
     * @return Returns the probability of state 2.
     */
    public double getProbabilityOfState2() {
        return 1.0 - probabilityOfState1;
    }

    /**
     * @param probabilityOfState2 The probability of state 2.
     */
    public void setProbabilityOfState2(double probabilityOfState2) {
        this.probabilityOfState1 = 1.0 - probabilityOfState2;
    }

    /**
     * @return The expectation of the random variable
     */
    public double getExpectation() {
        double expectation = probabilityOfState1 * value1
            + (1-probabilityOfState1) * value2;
        return expectation;
    }
}

```

Listing D.1: BisectionSearch: Nullstellenlöser nach dem Bisektionsverfahren der dem *RootFinder*-Interface genügt.

```

/*
 * Created on 16.02.2004
 *
 * (c) Copyright Christian P. Fries, Germany. All rights reserved. Contact: email@christian-fries.de.
 */
package net.finmath.fdm1.rootFinder;

/**
 * This class implements a Bisection search algorithm,
 * implemented as a question-and-answer search algorithm.
 *
 * @author Christian Fries - http://www.christian-fries.de
 * @version 1.0
 */
public class BisectionSearch implements RootFinder {

    // We store the left and right end point of the intervall
    double[] points = new double[2]; // left, right
    double[] values = new double[2]; // left, right

    /*
     * State of solver
     */

    double nextPoint; // Stores the next point to return by getPoint()
    boolean expectingValue = false; // Stores the state (true, if next call should be setValue(), false for getPoint())

    int numberOfIterations = 0; // Number of numberOfIterations
    double accuracy; // Current accuracy of solution
    boolean isDone = false; // Will be true if machine accuracy has been reached

    public static void main(String[] args) {
        // Test
        BisectionSearch search = new BisectionSearch(-1.0, 5.0);

        while(search.getAccuracy() > 1E-13 && !search.isDone()) {
            double x = search.getNextPoint();

```



```

        double y = (x - 0.656);
        search.setValue(y);
    }

    System.out.println(search.getBestPoint());
}

/**
 * @param leftPoint left point of search interval
 * @param rightPoint right point of search interval
 */
public BisectionSearch(double leftPoint, double rightPoint) {
    super();
    points[0] = leftPoint;
    points[1] = rightPoint;

    nextPoint = points[0];
    accuracy = points[1]-points[0];
}

/**
 * @return Best point obtained so far
 */
public double getBestPoint() {
    // Lazy: we always return the middle point as best point
    return (points[1] + points[0]) / 2.0;
}

/**
 * @return Next point for which a value should be set using <code>setValue</code>.
 */
public double getNextPoint() {
    expectingValue = true;
    return nextPoint;
}

/**
 * @param value Value corresponding to point returned by previous <code>getNextPoint</code> call.
 */
public void setValue(double value) {
    if (numberOfIterations < 2) {
        /**
         * Initially fill values
         */
        values[numberOfIterations] = value;

        if (numberOfIterations < 1) {
            nextPoint = points[numberOfIterations + 1];
        } else {
            nextPoint = (points[1] + points[0]) / 2.0;
            /**
             * @todo Check if values[0]*values[1] < 0 here
             */
        }
    }
    else {
        /**
         * Bisection search update rule
         */

        if (values[1] * value > 0) {
            /**
             * Throw away right point
             */
            points[1] = nextPoint;    // This is not yet the nextPoint, it is the point corresponding to value.
            values[1] = value;
        } else {
            /**
             * Throw away left point
             */
            points[0] = nextPoint;    // This is not yet the nextPoint, it is the point corresponding to value.
            values[0] = value;
        }

        // Calculate next point (bisection)
        nextPoint = (points[1] + points[0]) / 2.0;

        // Safety belt: check if still improve or if we have reached machine accuracy
    }
}

```

```

        if(points[1]-points[0] >= accuracy) isDone = true;

        // Update accuracy
        accuracy = points[1]-points[0];
    }

    numberOfIterations++;
    expectingValue = false;
    return;
}

/**
 * @return Returns the numberOfIterations.
 */
public int getNumberOfIterations() {
    return numberOfIterations;
}

/**
 * @return Returns the accuracy.
 */
public double getAccuracy() {
    return accuracy;
}

/**
 * @return Returns the isDone.
 */
public boolean isDone() {
    return isDone;
}
}

```

APPENDIX E

Technical Terms (Selection)

An English version of this part is currently not available. What follows is the German version.

section under construction.

Englisch	Deutsch	Interpretation
Notional	Nennwert	Betrag mit dem eine Zinsrate (dimensionslos) multipliziert wird.
Strike, Strike price	Ausübungspreis	Wert des Underlyings (z.B. $S(T)$) der einen qualitativen Wechsel im Auszahlungsprofil markiert (z.B. im Sinne einer Fallunterscheidung). Beispiele: K in $\max(S(T) - K, 0)$, c und f in $\min(\max(S(T) - K, f), c)$.
Strike rate	Ausübungsrate	Die Rate (Zinsrate), die bei einem Derivat zu einer Veränderung der Grundlage der Auszahlungsfunktion führt (Fallunterscheidung).
Hedge	(Hedge)	Eine Investition, welche das Risiko (unerwünschter) Preisschwankungen in einem Finanzprodukt reduzieren soll, z.B. durch Beziehen einer gegenstelligen Position in relevanten Finanzprodukten. Vergleiche auch Replikationsportfolio.
Short	(Short)	Ein Underlying $X_i(t)$ ist <i>short</i> , falls es zu negativen Anteilen Teil eines Portfolios (z.B. des Replikationsportfolios) ist ($\phi_i(t) < 0$). <i>short selling</i> : Leerverkauf.
Long	(Long)	Ein Underlying $X_i(t)$ ist <i>long</i> , falls es zu positiven Anteilen Teil eines Portfolios (z.B. des Replikationsportfolios) ist ($\phi_i(t) > 0$).

Englisch	Deutsch	Interpretation
risk neutral measure	Risikoneutrales Maß	Martingalmaß $\mathbb{Q}^N$ zum Numéraire $N(t) := B(t) := \exp(\int_0^t r(\tau) d\tau)$ (siehe 4) - Bemerkung: Der Name wird streng nur für dieses Martingalmaß verwendet, wegen der sprachlichen Nähe zu <i>risk neutral evaluation</i> aber unter Umständen auch einfach synonym zu Martingalmaß.
spot measure	Spot Maß	Martingalmaß $\mathbb{Q}^N$ zum Numéraire $N(t) := P(T_{m(t)+1}; t) \prod_{j=0}^{m(t)} (1 + L_j(T_j) \cdot \delta_j)$ (siehe Abschnitt 17.1.1.2). Diskrete Variante des <i>risk neutral measure</i> .
terminal measure	(Terminal Measure)	Martingalmaß $\mathbb{Q}^N$ zum Numéraire $N(t) := P(T_N; t)$, wobei T_N der Zeithorizont des betrachteten Modells sei. (siehe Abschnitt 17.1.1.1). Kann auch als T_N -forward measure bezeichnet werden.
T_k -forward measure	(T_k -Forward Measure)	Martingalmaß $\mathbb{Q}^N$ zum Numéraire $N(t) := P(T_k; t)$ für $t \leq T_k$ und $N(t) := P(T_{m(t)+1}; t) \prod_{j=k}^{m(t)} (1 + L_j(T_j) \cdot \delta_j)$ für $t \geq T_k$. Siehe Abschnitt 17.1.1.1.
fixing date	(Fixing Date)	Siehe Definition 103
payment date	(Payment Date)	Siehe Definition 103
accrued interest	Aufgelaufene Stückzinsen	Der mit dem Anteil abgelaufenen Periode gewichtete Coupon einer angebrochenen Periode. Siehe Bemerkung 99
dirty price	(Dirty Price)	Der Wert eine Coupon Bonds (ungemindert um evtl. aufgelaufene Stückzinsen). Siehe Bemerkung 99
clean price	(Clean Price)	Der Wert eine Coupon Bonds gemindert um evtl. aufgelaufene Stückzinsen. Siehe Bemerkung 99

List of Figures

2.1. Illustration of Measurability	30
2.2. Lebesgue integral versus Riemann integral	32
2.3. Conditional Expectation	35
2.4. Illustration of a filtration and an adapted process	37
2.5. Brownian motion: Time discretization	39
2.6. Brownian motion: Paths	41
2.7. Brownian motions with time dependent instantaneous volatility	41
2.8. Brownian motion with drift	42
2.9. Drift as a consequence of a non-linear function of a stochastic process	50
2.10. Integration of stochastic processes	57
3.1. Replication: The two-times-two-states-two-assets example	65
3.2. Replication: Generalisation to multiple states	67
3.3. Real Measure versus Martingale Measure	70
6.1. Arbitrage-free option prices	92
6.2. Linear interpolation of option prices	93
6.3. Linear interpolation of implied volatilities	94
6.4. Spline interpolation of option prices and implied volatilities	95
6.5. Linear interpolation for decreasing implied volatility	96
6.6. Linear interpolation for increasing implied volatility	96
7.2. Delta Hedge	112
7.1. Streuung des Wertes des Replikationsportfolios bei wöchentlichem Delta-Hedging	113
7.3. Streuung des Wertes des Replikationsportfolios bei monatlichem Re- hedging: Delta versus Delta-Gamma Hedge	116
7.4. Streuung des Wertes des Replikationsportfolios bei wöchentlichem Rehedging mit falscher Zinsrate	117

7.5. Wert des Replikationsportfolios ohne Rehedging	118
8.1. Modeling an interest rate curve through a family of stochastic processes.	128
8.2. Cashflows for a <i>forward bond</i>	129
8.3. UML Diagram: The class <code>DiscountFactors</code>	133
8.4. UML Diagram: The class <code>DiscountFactors</code> with bootstrapper	135
9.1. Cashflow of a floater with exchange of notional N	140
9.2. Fixing date, payment date and evaluation date	141
9.3. Call spread approximation of a digital option	147
12.1. Cashflows for a <i>coupon bond</i> . Left with imaginary exchange of notionals at the end of each period, right with effective cashflows only.	159
12.2. Cashflows for a <i>swap</i> who's fixed leg corresponds to the coupon bond in Figure 12.1. Left with imaginary exchange of notionals at the end of each period, right with effective cashflows only.	159
12.3. Cashflows for a <i>zero coupon bond</i> . Left with imaginary exchange of notionals at the end of each period, right with effective cashflows only.	162
12.4. Cashflows for a <i>swap</i> who's fixed leg corresponds to the zero coupon bond in Figure 12.3. Left with imaginary exchange of notionals at the end of each period, right with effective cashflows only.	162
12.5. Bermudan Swaption	164
13.1. Discretization and Implementation of Itô Processes	180
13.2. Monte Carlo Simulation	185
13.3. UML Diagram: Monte Carlo Simulation / Lognormal Process	186
13.4. UML Diagram: Black-Scholes Model / Monte Carlo Simulation	187
13.5. Monte Carlo Simulation	189
13.6. Lattice	190
13.7. Binomial Tree	191
13.8. Pfade des Binomial Tree	191
13.9. Lattice with overlaid Monte Carlo Simulation	192
14.1. Bermudan Option	196
14.2. Calculation of the conditional expectation by re-simulation	201
14.3. Perfect Foresight	202
14.4. Predictor variable versus realized value (continuation value)	203
14.5. Calculation of the conditional expectation by binning	204
14.6. <i>Continuation Value</i> and Binning	206
14.7. Regression of the conditional expectation estimator with and without restriction of the regression domain	211

14.8. Regression of the conditional expectation estimator - polynomial of fourth and eighth order	212
14.9. UML Diagram: Conditional Expectation Estimator	214
14.10 Binning using the linear regression algorithm with piecewise constant basis functions	215
14.11 Bermudan Option: Example of a successful optimization of the exercise criteria	218
14.12 Bermudan Option: Example of a failing optimization of the exercise criteria	218
15.1. Payoff of an auto cap under a parallel shift of the interest rate curve . .	233
15.2. Value of an auto cap under a parallel shift of the interest rate curve . .	233
16.1. Object Oriented Design of the Monte Carlo Pricing Engine: We depict the impact of a change of different market data scenarios $\theta + h$ and $\theta - h$ on the pricing code of a standard Monte-Carlo simulation and a proxy scheme simulation.	250
16.2. Importance sampling using a drift adjusted proxy scheme	251
17.1. Swaption as a function of the forward rates	267
17.2. UML Diagramm: LIBOR Market Model / Monte Carlo Simulation . .	275
17.3. UML Diagram: LIBOR Market Model: Abstraction of model parameters	276
17.4. UML Diagram: LIBOR Market Model: Abstraction of model parameters as volatility and correlation	277
17.5. UML Diagram: LIBOR Market Model: Abstraction of model parameters: Parametric covariance models	278
18.1. Correlation in a one- and multi-factor model	287
18.2. Correlation: Full Decorrelation in a One-Factor Model	288
18.3. The terminal distribution function of a forward rate under different martingale measures	293
18.4. The terminal distribution function of a forward rate under different martingale measures	294
22.1. Shape of the fixed rates $L_i(T_i)$ and the interest rate curve for different instantaneous volatilities	317
22.2. Shape of the interest rate curve under different factor configurations .	318
22.3. Shape of the fixed rates $L_i(T_i)$ and the interest rate curve under different instantaneous volatilities	319
22.4. Shape of the fixed rates $L_i(T_i)$ and the interest rate curve under different instantaneous correlations	321

23.1. Markov Functional Model: Calibration of the LIBOR Functionals . . .	339
23.2. Bedingte Erwartung im Lattice	341
23.3. State space discretization for the Markov functional model	344
25.1. Anatomie einer Java™ Klasse	384
A.1. Faktorreduktion im Falle hoher Korrelation	395
A.2. Faktorreduktion im Falle niedriger Korrelation - Faktoren	395
A.3. Faktorreduktion im Falle niedriger Korrelation - Korrelationen	395
A.4. Golden Section Search	396

List of Tables

7.1. Greeks im Black-Scholes Modell	108
8.1. Interest Rate Models.	132
12.1. Prototypical product properties and corresponding model requirements and implementation techniques.	156
12.2. Markov dimension of some models.	156
12.3. Product Toolbox	173
17.1. co-sliding and co-terminal swap rates	279
18.1. Caplet and swaption prices for different instantaneous correlations and volatilities	289
20.1. Selection of Short Rate Models	304
22.1. Free parameters of the LIBOR Market Model considered	316

Listings

25.1. BinomialDistributedRandomVariable: Binomialverteilte Zufallsvariable	364
25.2. Getter und Setter	367
25.3. DiscreteRandomVariableInterface: Interfacebeschreibung für eine diskrete Zufallsvariable	370
25.4. DiscreteRandomVariable: Abstrakte Basisklasse für diskrete Zufallsvariable	371
25.5. BinomialDistributedRandomVariable: Binomialverteilte Zufallsvariable	371
25.6. RootFinder: Interface eines eindimensionalen Nullstelllösers . . .	374
25.7. Testfunktion für Nullstellenlöser die dem Interface RootFinder genügen	375
25.8. NewtonsMethod: Nullstellenlöser nach dem Newtonverfahren der dem RootFinderWithDerivative-Interface genügt.	376
25.9. SecantMethod: Nullstellenlöser nach dem Sekantenverfahren der dem RootFinderWithDerivative-Interface genügt.	377
25.10. Testfunktionen für RootFinder und RootFinderWithDerivative Klassen	379
25.11. Ausgabe des Tests der RootFinder und RootFinderWithDerivative Klassen	380
A.1. Erzeugung einer n -dimensionalen Brownschen Bewegung	390
D.1. BisectionSearch: Nullstellenlöser nach dem Bisektionsverfahren der dem RootFinder-Interface genügt.	410

Bibliography

Books

- [1] BAUER, HEINZ: Maßtheorie und Integrationstheorie. 2. Auflage. de Gruyter, Berlin, 1992. ISBN 3-11-013625-2.
- [2] BAUER, HEINZ: Wahrscheinlichkeitstheorie. de Gruyter, Berlin 2001. ISBN 3-11-017236-4.
- [3] BAXTER, MARTIN W.; RENNIE, ANDREW J.O.: Financial Calculus: An introduction to derivative pricing. Cambridge University Press, Cambridge, 2001. ISBN 0-521-55289-3.
- [4] BRANGER, NICOLE; SCHLAG, CHRISTIAN: Zinsderivate - Modelle und Bewertung. Springer Verlag, 2004. 199 Seiten. ISBN 3-540-21228-0.
- [5] BIERMANN, BERND: Die Mathematik von Zinsinstrumenten. Oldenbourg Verlag, 2002. ISBN 3-486-25976-8.
- [6] BINGHAM, NICHOLAS H.; KIESEL, RÜDIGER: Risk-neutral valuation: pricing and hedging of financial derivatives. (Springer Finance). Springer Verlag, London, 1998. ISBN 1-852-33458-4. 296 Seiten.
- [7] BJÖRK, THOMAS: Arbitrage Theory in Continuous Time. Oxford University Press, 1999. ISBN 0-198-77518-0. 328 Seiten.
- [8] BRIGO, DAMIANO; MERCURIO, FABIO: Interest Rate Models - Theory and Practice. Springer-Verlag, Berlin, 2001. ISBN 3-540-41772-9.
- [9] CONT, RAMA; TANKOV, PETER: Financial Modelling with Jump Processes. CRC Press, 2003. ISBN 1-584-88413-4.
- [10] DRAPER, NORMAN R.; SMITH, HARRY: Applied Regression Analysis. Third Edition. Wiley-Interscience, 1998. ISBN 0-471-02995-5.

- [11] ECKEL, BRUCE: Thinking in Java (3rd Edition). Prentice Hall PTR, 2002. ISBN 0-131-00287-2.
- [12] HUNT, PHIL J.; KENNEDY, JOANNE E.: Financial Derivatives in Theory and Practice. John Wiley & Sons, 2000. ISBN 0-471-96717-3.
- [13] GAMMA, ERICH; HELM, RICHARD; JOHNSON, RALPH E.: Design Patterns. Addison-Wesley Professional, 1997. ISBN 0-2-016-3361-2.
- [14] GLASSERMAN, PAUL: Monte Carlo Methods in Financial Engineering. 596 Seiten. Springer, 2003. ISBN 0-387-00451-3.
- [15] GÜNTHER, MICHAEL; JÜNGEL, ANSGAR: Finanzderivate mit MATLAB. Mathematische Modellierung und numerische Simulation. Vieweg, 2003. ISBN 3-528-03204-9.
- [16] JÄCKEL, PETER: Monte Carlo Methods in Finance. 238 Seiten. John Wiley and Sons Ltd., 2002. ISBN 0-471-49741-X.
- [17] JOSHI, MARK S.: The Concepts and Practice of Mathematical Finance. Cambridge University Press, 2003. ISBN 0-521-82355-2.
- [18] KARATZAS, IOANNIS; SHREVE, STEVEN E.: Brownian Motion and Stochastic Calculus. Second Edition. Springer Verlag, 1991. ISBN 0-387-97655-8.
- [19] KLOEDEN, PETER E.; PLATEN, ECKHARD: Numerical Solution of Stochastic Differential Equations (Applications of Mathematics. Stochastic Modelling and Applied Probability Vol. 23). Springer Verlag, 1999. ISBN 3-540-54062-8.
- [20] MALLIAVIN, PAUL: Stochastic Analysis (Grundlehren Der Mathematischen Wissenschaften). Springer Verlag, 1997. ISBN 3-540-57024-1 .
- [21] MEISTER, MARKUS: Smile Modeling in the LIBOR Market Model. Diploma Thesis. University of Karlsruhe, 2004.
- [22] MUSIELA, MAREK; RUTKOWSKI, MAREK: Martingale methods in financial modelling: theory and applications. Springer-Verlag, 1997. ISBN 3-540-61477-X.
- [23] PAUL, WOLFGANG; BASCHNAGEL JÖRG: Stochastic Processes. From Physics to Finance. Springer-Verlag, 2000. ISBN 3-540-66560-9.
- [24] PELSSER, ANTOON: Efficient Methods for Valuing Interest Rate Derivatives. Springer-Verlag, 2000. ISBN 1-852-33304-9.
- [25] ØKSENDAL, BERNT K.: Stochastic differential equations: an introduction with applications. Springer-Verlag, 2000. ISBN 3-540-63720-6.

- [26] OESTEREICH, BERND: Objektorientierte Softwareentwicklung. Oldenburg Verlag, 2004. ISBN 3-486-27266-7.
- [27] PROTTER, PHILIP E.: Stochastic Integration and Differential Equations. Springer-Verlag, 2003. ISBN 3-540-00313-4. 415 Seiten.
- [28] REBONATO, RICCARDO: Modern Pricing of Interest-Rate Derivatives: The Libor Market Model and Beyond. Princeton University Press, 2002. ISBN 0-691-08973-6.
- [29] ROGERS, L. C. G.; WILLIAMS, DAVID: Diffusions, Markov Processes and Martingales: Volume 2, Ito Calculus (2nd Edition). Cambridge University Press, 2000. ISBN 0-521-77593-0.
- [30] SEYDEL, RÜDIGER: Tools for Computational Finance. Springer Verlag, Berlin, 2003. 240 Seiten. ISBN 3-540-40604-2.
- [31] SHIRYAEV, ALBERT N.: Essentials of Stochastic Finance: Facts, Models, Theory. World Scientific Publishing Company, 1999. ISBN 9-810-23605-0.
- [32] J. MICHAEL STEELE: Stochastic Calculus and Financial Applications. Springer-Verlag, 2001. 344 pages. ISBN 0-387-95016-8
- [33] TAVELLA, DOMINGO; RANDALL CURT: Pricing Financial Instruments: The Finite Difference Method. Wiley, 2000. 237 pages. ISBN 0-471-19760-2
- [34] ULLENBOOM, CHRISTIAN: Java ist auch eine Insel, mit CD ROM. 5. Auflage. Galileo Press, 2004. ISBN 3-898-42526-6.
- [35] YOUNG, JEFFREY S.; SIMON, WILLIAM L.: iCon Steve Jobs. John Wiley & Sons, 2005. ISBN 0-471-72083-6.
- [36] WILDE, OSCAR: The Picture of Dorian Gray. ISBN 0-679-60001-9.
- [37] WILDE, OSCAR: The Importance of Being Earnest and Other Plays. Oxford University Press (Reprint). ISBN 0-198-12167-9.
- [38] WILMOTT, PAUL: Paul Wilmott on Quantitative Finance. John Wiley, 2000. ISBN 0-471-87438-8.
- [39] WILLIAMS, DAVID: Probability with Martingales. Cambridge University Press, 2002. ISBN 0-521-40605-6. (First published 1991)
- [40] WILLIAMS, DAVID: Weighing the Odds: A Course in Probability and Statistics. Cambridge University Press, 2001. ISBN 0-521-00618-X.
- [41] ZHANG, PETER: Exotic Options: A Guide to Second Generation Options, 1998. ISBN 9-810-23521-6.

Papers

- [42] ANDERSEN, LEIF: A simple approach to the pricing of Bermudan swaptions in the multi-factor Libor market model. Working paper. General Re Financial Products, 1999.
- [43] ANDERSEN, LEIF; BROADIE, MARK: A Primal-Dual Simulation Algorithm for Pricing Multi-Dimensional American Options. Working paper. General Re Financial Products, 1999.
- [44] ANDERSEN, LEIF; SIDENIUS, JAKOB; BASU, SUSANTA: All your hedges in one basket. *Risk Magazine* 11/2003, 67-72 (2003).
- [45] BENHAMOU, ERIC: Optimal Malliavin Weighting Function for the Computation of the Greeks. 2001.
- [46] BLACK, FISCHER: The Pricing of Commodity Contracts. *Journal of Financial Economics* 3, 167-179, 1976. North-Holland Publishing Company, 1976.
- [47] BOUCHAUD, JEAN-PHILIPPE; SORNETTE, DIDIER: The Black-Scholes Option Pricing Problem in Mathematical Finance: Generalizations and Extensions for a Large Class of Stochastic Processes. *Journal de Physique I* 4, 863, 1994.
- [48] BRACE, ALAN; GATAREK, DARIUSZ; MUSIELA, MAREK: The Market Model of Interest Rate Dynamics. *Mathematical Finance* 7, p. 127. Blackwell 1997.
- [49] BRASCH, HANS-JÜRGEN: A Note on Efficient Pricing and Risk Calculation of Credit Basket Products. Preprint, 2005. http://defaultrisk.com/pp_crdrv_54.htm.
- [50] BREEDEN, D. T.; LITZENBERGER, R. H.: Prices of state-contingent claims implicit in option prices. *Journal of Business* 51(4), 621-651, 1978.
- [51] BOYLE, PHELIM; BOADIE, MARK; GLASSERMAN, PAUL: Monte Carlo methods for security pricing. *Journal of Economic Dynamics and Control*, 21, 1267-1321 (1997).
- [52] BRIGO, DAMIANO; MERCURIO, FABIO; RAPISARDA, FRANCESCO: Lognormal-mixture dynamics and calibration to market volatility smiles. 2000. <http://www.damianobrigo.it>
- [53] BROADIE, MARK; GLASSERMAN, PAUL: Estimating Security Price Derivatives using Simulation. *Management Science*, 1996, Vol. 42, No. 2, 269-285.
- [54] BROADIE, MARK; GLASSERMAN, PAUL: Pricing American-Style Securities by Simulation. *J. Econom. Dynam. Control*, 1997, Vol. 21, 1323-1352.

- [55] CARRIERE, JACQUES F.: Valuation of Early-Exercise Price of Options Using Simulations and Nonparametric Regression. *Insurance: Mathematics and Economics* 19, 19-30, 1996.
- [56] CLÉMENT, EMMANUELLE; LAMBERTON, DAMIEN; PROTTER, PHILIP: An analysis of a least squares regression method for American option pricing. *Finance and Stochastics* 6, 449-471, 2002.
- [57] DAVIS, MARK; KARATZAS, IOANNIS: A Deterministic Approach to Optimal Stopping, with Applications. In: Whittle, Peter (Ed.): Probability, Statistics and Optimization: A Tribute to Peter Whittle, 455-466, 1994. John Wiley & Sons, New York and Chichester, 1994.
- [58] FOURNIÉ, ERIC; LASRY JEAN-MICHEL; LEBUCHOUX, JÉRÔME; LIONS, PIERRE-LOUIS; TOUZI, NIZAR: Applications of Malliavin calculus to Monte Carlo methods in finance. *Finance Stochastics*. 3, 391-412 (1999). Springer- Verlag 1999.
- [59] FRIES, CHRISTIAN P.: The Foresight Bias in Monte-Carlo Pricing of Options with Early Exercise: Classification, Calculation and Removal. 2005. <http://www.christian-fries.de/finmath/foresightbias>.
- [60] FRIES, CHRISTIAN P.; JOSHI, MARK S.: Partial Proxy Simulation Schemes for Generic and Robust Monte-Carlo Greeks. 2006. <http://www.christian-fries.de/finmath/proxyscheme>
- [61] FRIES, CHRISTIAN P.; KAMPEN, JÖRG: Proxy Simulation Schemes for generic robust Monte-Carlo sensitivities, process oriented importance sampling and high accuracy drift approximation (with applications to the LIBOR Market Model). 2005. <http://www.christian-fries.de/finmath/proxyscheme>
- [62] FRIES, CHRISTIAN P.; ROTT, MARIUS G.: Cross Currency and Hybrid Markov Functional Models. Preprint, 2004. <http://www.christian-fries.de/finmath/markovfunctional>
- [63] GLASSERMAN, PAUL; LI, JINGYI: Importance Sampling for Portfolio Credit Risk. Working paper, Columbia University (2003). http://www2.gsb.columbia.edu/faculty/pglasserman/other/is_credit.pdf.
- [64] GLASSERMAN, PAUL; ZHAO, XIAOLIANG: Arbitrage-free discretization of lognormal forward Libor and swap rate models. *Finance and Stochastics* 4, 35-68, 2000. Springer-Verlag, 2000.
- [65] GLASSERMAN, PAUL; ZHAO, XIAOLING: Fast greeks in forward LIBOR models. *Journal of Computational Finance*. 3, 5-39 (1999).

- [66] HAGAN, PATRIK S.; KUMAR, DEEP; LESNIEWSKI, ANDREW S.; WOODWARD, DIANA E.: Managing Smile Risk (SABR Model), Wilmot Magazin, September 2002.
- [67] HAGAN, PATRIK S.; WEST, GRAEME: Interpolation Methods for Curve Construction. Preprint. 2005.
- [68] HAUGH, MARTIN; KOGAN, LEONIS: Pricing American Options: A Duality Approach, MIT Sloan Working Paper No. 4340-01, 2001.
- [69] HEATH, DAVID; JARROW, ROBERT; MORTON, ANDREW: Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation. *Econometrica* 60, No. 1 (January), 77-105, 1992.
- [70] HUNT, PHIL J.; KENNEDY, JOANNE E.; PELSSER, ANTOON: Markov-Functional Interest Rate Models. *Finance and Stochastics*, Volume 4 (4), 391-408, 2000. Springer-Verlag, 2000. Reprinted in: Hughston, Lane (ed.): *The New Interest Rate Models*, RISK Publications.
- [71] HUNTER, CHRISTOPHER J.; JÄCKEL, PETER; JOSHI, MARK S.: Drift Approximations in a Forward-Rate-Based LIBOR Market Model. *Getting the Drift*. *Risk* 14, 81-84 (July, 2001).
- [72] JAMSHIDIAN, FARSHID: LIBOR and swap market models and measures. *Finance and Stochastics* 1, 293-330 (1997). Springer-Verlag, 1997.
- [73] JOSHI, MARK S.: Applying Importance Sampling to Pricing Single Tranches of CDOs in a One-Factor Li Model. QUARC, Group Risk Management, Royal Bank of Scotland. Working paper, 2004.
- [74] JOSHI, MARK S.; KAINTH, DHERMINDER: Rapid computation of prices and deltas of n^{th} to default swaps in the Li Model. *Quantitative Finance*, volume 4, issue 3, (June 04), p. 266- 275.
<http://www.quarchome.org/>.
- [75] LI, D.: On Default Correlation: a Copula Approach (*Journal of Fixed Income*, 9, 43-45)
- [76] LONGSTAFFE, FRANCIS A.; SCHWARTZ, EDUARDO S.: Valuing American Options by Simulation: A Simple Least-Square Approach. *Review of Financial Studies* 14:1, 113-147, 2001.
- [77] MATSUMOTO, M.; NISHIMURA, T.: Mersenne twister: A 623-dimensionally equidistributed uniform pseudorandom number generator. *ACM Transactions on Modeling and Computer Simulations*, 1998.

- [78] MILTERSEN, KRISTIAN R.; SANDMANN, KLAUS; SONDERMANN, DIETER: Closed Form Solutions for Term Structure Derivatives with Lognormal Interest Rates. *Journal of Finance* 52, 409-430 (1997).
- [79] PITERBARG, VLADIMIR V.: A practitioner's guide to pricing and hedging callable LIBOR exotics in forward LIBOR models, *Preprint*. 2003.
- [80] PITERBARG, VLADIMIR V.: Computing deltas of callable LIBOR exotics in forward LIBOR models. *Journal of Computational Finance*. 7 (2003).
- [81] ROGERS, L. C. G.: Monte Carlo valuation of American options, *Preprint*. 2001.
- [82] ROTT, MARIUS G.; FRIES, CHRISTIAN P.: Fast and Robust Monte Carlo CDO Sensitivities and their Efficient Object Oriented Implementation. 2005. <http://www.christian-fries.de/finmath/cdogreeks>
- [83] SCHLÖGL, ERIK: A Multicurrency Extension of the Lognormal Interest Rate Market Models. *Finance and Stochastics* 6 (2), 173-196 (2002). Springer-Verlag, 2002.
- [84] SCHLÖGL, ERIK: Arbitrage-Free Interpolation in Models of Market Observable Interest Rates. *Preprint* (2002).
- [85] WICHURA, MICHAEL J.: Algorithm AS 241: The percentage points of the normal distribution. *Applied Statistics* 37, 477-484, 1988.

Index

A			
abgeleitete Klasse	373	Bermudan Swaption	163
Abstrakte Klasse	370	Bermudan swaption	164
Abstrakte Methode	370	Binning	203
Accrued Interest	137	Binomial Tree	190, 191
Annuity	405	– Pfade	191
Arbitragefreie Optionspreise	91	Black volatility	150
Auszahlungsfunktion		Black-Derman-Toy Model	304
– replizierbare	77	Black-Karasinski Model	304
Auto Cap	170	Black-Scholes	
		– Partielle Differentialgle-	
		ichung	103, 106
		– PDE	103, 106
		Black-Scholes Modell	
		– Delta-Gamma Hedge	115
		– Delta-Hedge	111
		– Evaluation via Monte-Carlo Simula-	
		tion	186
		– Greeks	108
		Black-Scholes volatility	85, 88
		Black-Scholes-Merton Formel	104
		Black-Scholes-Merton formula	87
		Black-Scholes-Merton Model	85
		Black-Scholes-Merton model	88
		Bond	128
		– Forward Bond	129
		– interpolation of prices	349
		– Zero Bond	128
		bond volatility	298
		Bootstrapping	
B			
Backward Algorithm	191		
Backward algorithm			
– for short rate models	305		
backward algorithm	192		
Backward Algorithmus	199		
Basisklasse	373		
Bedingte Erwartung			
– Eigenschaften	399		
– Prozess	37		
Bermudan	163		
Bermudan Callable	172, 196, 197		
Bermudan Callable Inverse CMS			
Floater	169		
Bermudan Callable Structured Swap	164		
Bermudan Cancelable	165		
Bermudan Option	196, 198		

- terminal 293, 294
 - distribution 32
 - distribution function 32
 - n -dimensional 33
 - Donsker Invariance Principle 38
 - Doob-Meyer Zerlegung 222
 - Dothan Model 304
 - drift 41, 50, 52
- E**
- Entwurfsmuster 383
 - Equity Hybrid LIBOR Market Model 352
 - Equity-Hybrid Cross-Currency LIBOR Market Model
 - Implementation 357
 - Equity-Hybrid LIBOR Market Model
 - Implementation 354
 - equivalent martingale measure see Martingal-Maß
 - equivalent measure 53
 - Erweiterung (Vererbung) 378
 - Euler Scheme 179
 - Europäische Option
 - Interpolation der Preise 91
 - Europäische Option
 - Arbitragefrei Preise 92
 - European option
 - Probability density of the underlying 89
 - Exercise Boundary 198
 - Exercise Strategy 198
 - Exotische Derivate 155
 - Expectation 33
 - Extended Vasicek model 304
- F**
- factor loading 186
 - factor matrix 51
 - factor reduction 315
 - factors 315
 - Filtration 36
 - generierte 36
 - filtration
 - generated filtration 43
 - Fixed Leg 142
 - Fixing Date 140
 - Flexi Cap 170
 - Floater 138
 - Cashflow Diagram 140
 - Floating Leg 142
 - Floating Rate Bond 139
 - Floor 145
 - floor 168
 - Floored 172
 - floored 146
 - Floored floater 168
 - Floorlet 145
 - Foreign-Caplet 147
 - Foresight Bias 207
 - Foresight Fehler
 - Eliminierung 224, 225
 - forward algorithm 192
 - Forward Bond
 - Cashflow Diagram 129
 - Forward forward 131
 - Forward FX Rate 152
 - Forward Rate 130
 - Instantaneous 130
 - forward rate
 - instantane 295
 - forward volatility 131
 - FX Forward 152
- G**
- Gamma 107
 - Gestoppter Prozess 222
 - Getter 365
 - Girsanov, Cameron, Martin Theorem 53
 - Goldener Schnitt 396
 - Greeks 107
 - Delta 107

- Gamma 107
- im Black-Scholes Modell 108
- Rho 107
- Theta 107
- Vega 107
- Griechen *see* Greeks

H

- Handelsstrategie
 - zulässige 77
- Heath, Jarrow, Morton
 - Drift Condition 296
- Heath-Jarrow-Morton 295, 307
- Hedge 99, 100, 413
- Hedging 99
- hedging 100
- HJM 295, 296, 307
- Ho-Lee Model 304
- Hull-White Model 304

I

- image measure 29
- Implementation
 - Black-Scholes Modell (via Monte Carlo Simulation) 186
 - Cross-Currency LIBOR Market Model 352
 - Equity-Hybrid Cross-Currency LIBOR Market Model 357
 - Equity-Hybrid LIBOR Market Model 354
 - LIBOR Market Model 263, 275
- Implementierung 362
- Implied Black volatility 150
- Implied Black-Scholes volatility 88
- implied volatility 97
- import (Java™ Schlüsselwort) 380
- Importance sampling
 - using a proxy scheme 251
- independence
 - of events 28
 - of random variables 33
- information (filtration) 37
- Instantaneous Correlation
 - Definition 286
- instantaneous correlation 50
 - in the LIBOR Market Model 256, 265
 - within the model for a quanto caplet 153
- Instantaneous Covariance
 - Definition 286
- Instantaneous Forward Rate 130
- instantaneous forward rate 295
- integrable (with respect to a measure) 31
- Integral
 - Stratonovich 401
- integral 57
 - Itô integral 43
 - Lebesgue integral 30, 32
 - Riemann integral 32
 - stochastic integral 57
- integrated covariance 293
- integrator 46
- Interest Rate Curve
 - Interpolation 134
- Interface 362, 366, 369
- Interface (Class) 188
- Interpolation
 - of bond prices 349
 - von Optionspreisen 91
- Interpolation (Interest Rates) 134
- Inverse 172
- Inverse floater 168
- iPod™ 361
- Itô Calculus 40–50
- Itô Integral 44
- Itô integral 43–45
 - integrand 44
- Itô Isometry 44
- Itô Lemma 47–50
 - multi dimensional 47
 - one dimensional 47

M

- MT 19937 389
- N**
- Neville algorithm 343
- Neville Algorithmus 343
- Newtonverfahren 376
- Notional 413
- notional 140, 141
- Numéraire 67, 74
- O**
- Oberklasse 373
- Objective-C 374
- objectoriented design 383
- Objekt 362
- objektorientiertes Design 383
- OOD 383
- Optimal Exercise 198
- Optimal Stopping* Problem 223
- Over fitting 268
- over fitting 268
- P**
- P*-almost surely (footnote) 34
- Par swap rate 143, 404
- path 36
- path dependent produkt 170
- Payer Swap 142
- Payment Date 140
- Payoff 77
- PCA . *see* Principal Component Analysis
- Perfect Foresight 201, 202
- Polymorph 373
- Polymorphie 373
- Portfolioprozess 72
- Predictor-Corrector Scheme 182
- previsible process 58
- Price of Risk 303, *see*
- Market Price of Risk
- Principal Component Analysis 315, 394
- probability
- conditional 28
- Probability density
- of the underlying of an European call option 89
- probability density 33
- probability space 27
- Products
- Auto Cap 170
- Bermudan 163
- Bermudan Callable Inverse CMS Floater 169
- Bermudan Callable Structured Swap 164
- Bermudan Cancelable 165
- Bermudan Swaption 163
- Bermudan swaption 164
- cancelable 165
- Cap 145
- Caplet 145
- Capped CMS Floater 168
- Capped Floater 168
- Chooser Cap 171
- Compound Option 166
- Coupon Bond 137
- Digital Caplet 146
- Flexi Cap 170
- Floater 138
- Floating Rate Bond 139
- Floor 145
- Floored Floater 168
- Floorlet 145
- Floreign-Caplet 147
- Inverse Floater 168
- Money-Market Account 145
- path dependent 170
- Product toolbox 172
- Quanto Caplet 147
- Range Accrual 169
- Rolling Bond 145
- Savings Account 145

Q

R

Comments welcome.
©2004, 2005, 2006 Christian Fries
Version 1.3.19 [build 20061210]- 13th December 2006
<http://www.christian-fries.de/finmath/>

- Smile
 – Caplet-Smile 266
 Snell Envelope 222
 Snowball 170
 Späte Bindung 373
 Spezialisierung (Vererbung) 378
 spot measure 259, 348
 Spread 172
 static (Java™ Schlüsselwort) ... 366
 statische Methode 365
 Stochastic Integral
 – Itô Process as Integrator 46
 stochastic integral 57
 – semi-martingale as integrator 58
 – with Brownian motion as integrator 44
 stochastic process 35
 – lognormal 49
 – previsible 58
 – previsible process 58
 Stochastische Differentialgleichung
 – Diskretisierung 177
 Stochastischer Prozess 35
 – Gestoppter Prozess 222
 – quadratische Variation 400
 – totale Variation 400
 Stopping Time 221
 Stoppzeit 221
 Stratonovich Integral 401
 Strike 413
 Structured Bond 144, 157
 – definition 160
 structured coupon 144
 Structured Swap 144, 157
 – definition 160
 super class 373
 Superklasse 373
 Supermartingal 221
 Swap 141
 – Cashflow Diagram 159
 – Swap rate 142, 289
 Swap (Payer) 142
 Swap (Receiver) 142
 Swap Annuity 289, 405
 Swap annuity 280
 Swap measure 280
 Swap rate 142, 143, 289
 – constant maturity 169
 Swaption 146, 267
 – as function of the forward rates .. 267
- T**
- Taking out what is know 400
 Target Redemption Note 166
 Tenor structure 128
 tenor structure 256
 term sheet 174
 Terminal Correlation
 – Definition 286
 Terminal Covariance
 – Definition 286
 terminal measure 152, 172, 258, 335
 Theta 107
 time to maturity 309
 T_k -forward measure 261
 Totale Variation 400
 Tower Law 400
 Trading Strategy 72, *see* Handelsstrategie
 Typ
 – eines Objekts 362, 363
 Typsicherheit 374
- U**
- Überladen (einer Methode) 365
 UML 185
 – Digram 186
 – Digramm 133, 135, 187, 214, 276–278
 – – LIBOR Market Model 275
 underlying 146
 Unified Modeling Language 185
- V**

Variable Maturity Inverse Floater ...	167	Zufallszahlen	389
Vasicek Model	304		
Vasicek model			
– Extended Vasicek model	304		
Vega	107		
Vega Hedge	117		
Vererbung	378		
void (Java™ Schlüsselwort)	365		
Volatility			
– implied volatility	88		
– time structure	309		
volatility			
– forward volatility	131		
– implied volatility	97		
volatility bootstrapping	268		
volatility surface	97		
– definition	97		

W

weak convergence	39
weather derivative	61
weighted Monte-Carlo	185
Wiener measure	40
Wiener process	38

Y

Yield	130

Z

Zerlegung	401
– Feinheit	401
Zero Coupon Bond	
– Cashflow Diagram	162
Zero Structure	157
zero structure	
– definition	160
Zero Swap	
– Cashflow Diagram	162
Zinsstrukturen	127

440 pages. 80 figures. 9 tables.