



Johann Wolfgang Goethe-Universität  
Frankfurt am Main

---

**Parameter Calibration und Simulation von Pfaden für  
das stochastic-volatility jump diffusion  
Optionsbewertungsmodell von Bakshi**

ANTONIO LASIĆ

© *Draft date 13. November 2002*

Betreuer:  
Prof. Dr. P.E. Kloeden  
Dr. Uwe Wystup



# Inhaltsverzeichnis

<b>Inhaltsverzeichnis</b>	<b>i</b>
<b>Einführung</b>	<b>v</b>
<b>1 Stochastische Prozesse</b>	<b>1</b>
1.1 Martingale . . . . .	3
1.2 Wiener Prozess . . . . .	4
1.3 Poisson Prozess . . . . .	6
<b>2 Stochastische Differentialgleichungen</b>	<b>9</b>
2.1 Stochastische Integrale . . . . .	10
2.2 Numerische Methode zur Lösung der SDGen . . . . .	12
<b>3 Sprung Diffusion SDGen</b>	<b>17</b>
3.1 Numerische Methoden zur Lösung der Sprung-Diffusion SDGen . . . . .	17
<b>4 Bakshi Modell</b>	<b>19</b>
4.1 Modellannahmen . . . . .	19
4.2 Analytische Bewertungsformel . . . . .	21
<b>5 Anwendungen des Black-Scholes Modells</b>	<b>23</b>
5.1 Implizite Volatilität . . . . .	24
5.2 Greeks . . . . .	26
<b>6 Numerische Bewertungsverfahren exotischer Optionen</b>	<b>29</b>
6.1 Exotische Optionen . . . . .	29
6.2 Monte-Carlo-Methode . . . . .	33

<b>7</b>	<b>Numerische Optimierungsverfahren</b>	<b>35</b>
7.1	Lokale Optimierung . . . . .	37
7.2	Globale Optimierung . . . . .	41
<b>8</b>	<b>Parameter Calibration</b>	<b>47</b>
8.1	Parameter Calibration . . . . .	48
8.2	Untersuchungen des Bakshi Modells . . . . .	49
<b>A</b>	<b>Programme</b>	<b>69</b>
	<b>Literaturverzeichnis</b>	<b>75</b>

# Abbildungsverzeichnis

1.1	Wiener Prozess . . . . .	6
1.2	Realisierungen des Poisson-Prozesses mit $\lambda = 3$ , im Zeitintervall $T = [0, 1]$ . . . . .	8
2.1	Ito/Stratonovich Interpretation . . . . .	11
2.2	Eine Trajektorie der SDG, simuliert durch das Euler-Maruyama-Verfahren, $a=2$ , $b=1$ , $Y_0 = 1$ , $\Delta t = 2^{-6}$ , $t = [0, 1]$ . . . . .	14
4.1	Simulierte Trajektorie des Bakshi Modells . . . . .	21
5.1	IV einer Call-option für unterschiedliche Strikes: Smile . . . . .	25
5.2	IV einer Call-Option für unterschiedliche Strikes: Smirk . . . . .	26
6.1	Beispiel einer Range-Option . . . . .	31
7.1	Nichtlineare Zielfunktion mit einem Parameter . . . . .	36
7.2	Struktur eines einfachen evolutionären Algorithmus . . . . .	46
8.1	Calibration Engine . . . . .	48
8.2	Gauss-Laguerre Nodes and Weights . . . . .	51
8.3	Markt <i>smile</i> (EUR/USD) von 05. Nov 2001 mit entsprechend angepassten <i>smile</i> aus dem <i>Bakshi</i> Modell . . . . .	56
8.4	Markt <i>smile</i> (EUR/USD) von 05. Nov 2001 mit entsprechend angepassten <i>smile</i> aus dem <i>Bakshi</i> Modell . . . . .	57
8.5	Markt <i>smile</i> (EUR/USD) von 05. Nov 2001 mit entsprechend angepassten <i>smile</i> aus dem <i>Bakshi</i> Modell . . . . .	57
8.6	Ergebnisse der GEATbx Toolbox . . . . .	60
8.7	Markt <i>smile</i> (EUR/USD) von 03. Sep 2002 mit entsprechend angepassten <i>smile</i> aus dem <i>Bakshi</i> Modell . . . . .	62

8.8	Markt <i>smile</i> (EUR/USD) von 03. Sep 2002 mit entsprechend angepassten <i>smile</i> aus dem <i>Bakshi</i> Modell . . . . .	62
8.9	Simulierter Pfad mit errechneten Parametern unter Verwendung der Monte-Carlo Methode . . . . .	63
8.10	Simulierter Pfad mit errechneten Parametern (initial Vol = 0.12) unter Verwendung der Monte-Carlo Methode . . . . .	63
8.11	Markt <i>smile</i> (EUR/USD) von 05. Nov 2001 mit entsprechend angepassten <i>smile</i> aus dem <i>Bakshi</i> Modell . . . . .	65
8.12	Markt <i>smile</i> (EUR/USD) von 03. Sep 2002 mit entsprechend angepassten <i>smile</i> aus dem <i>Bakshi</i> Modell . . . . .	67
A.1	GLagNodeWt.m . . . . .	69
A.2	impvol.m . . . . .	70
A.3	simann.m . . . . .	71
A.4	simann.m . . . . .	72
A.5	bakshi.m . . . . .	73
A.6	Winccorr.m . . . . .	74
A.7	pois.m . . . . .	74

# Einführung

Die vorliegende Arbeit beschäftigt sich mit der Ermittlung des Preises von Optionen. Optionen sind spezielle Derivate, die wiederum Hull in seinem Buch definiert als: “Ein Derivat ist ein Finanzinstrument, dessen Wert von einem anderen, einfacheren zu Grunde liegenden Finanzinstrument (*underlying*) abhängt“. Ein *underlying* kann unter anderem auch eine Anleihe, eine Aktie oder der Umtauschkurs zweier Währungen sein.

Eine besonders einfache Form von Finanzderivaten sind Terminkontrakte. Ein Terminkontrakt ist ein Vertrag zwischen zwei Parteien, bei der sich der Käufer bzw. der Verkäufer des Kontraktes zu einem Zeitpunkt verpflichtet, zu einem zukünftigen Zeitpunkt  $T > t$ , dem Ende der Laufzeit (*maturity*), ein Objekt zu einem heute vereinbarten Preis  $K$ , dem Terminkurs, zu kaufen bzw. zu verkaufen. Damit sichert sich der Käufer gegen unerwartete Preissteigerungen ab, geht aber das Risiko ein, dass der Preis fällt und er dann zu teuer kauft. Andererseits sichert sich der Verkäufer gegen fallende Preise, vergibt aber die Möglichkeit bei steigenden Preisen mehr Gewinn zu erzielen. Der Lieferpreis wird dabei so festgelegt, dass der Wert des Terminkontraktes am Anfang gerade Null ist.

Bei einem Terminkontrakt wechselt das zugrundeliegende Objekt am Ende der Laufzeit auf jeden Fall den Besitzer. Im Gegensatz dazu gewährt eine Option einem der beiden Vertragspartner das Recht zu entscheiden, ob es überhaupt zu einer Transaktion kommt. Offensichtlich macht es dabei einen Unterschied, ob die Entscheidung beim Käufer oder beim Verkäufer liegt.

Ein *europäischer Call* ist ein Vertrag zwischen zwei Parteien, der dem Käufer das Recht gibt, zu einem festgelegten zukünftigen Zeitpunkt  $T > t$ , dem Ende der Laufzeit oder Verfallszeitpunkt, ein Objekt zu einem heute vereinbarten Preis  $K$ , dem Ausübungskurs (*strike*), zu kaufen. Übt er das Recht nicht aus, verfällt die Option zur Zeit  $T$  ohne weitere Konsequenzen. Ein *europäischer Put* gibt dem Käufer das Recht, das Objekt zur Zeit  $T$  zum vereinbarten Preis  $K$  zu verkaufen. Der Inhaber eines *amerikanischen Calls* bzw. eines *amerikanischen Puts* kann sein Recht, das Objekt zu kaufen bzw. zu verkaufen, jederzeit vom Erwerb bis zum Verkaufszeitpunkt  $T$  ausüben.

Daneben gibt es zahlreiche komplexere Finanzderivate, wobei laufend neue entwickelt werden. Diese Finanzinstrumente werden meist nicht an den Terminbörsen

gehandelt, sondern von Banken für einzelne Kunden maßgeschneidert entworfen. Man spricht dann von *exotischen Optionen*. Die mathematische Behandlung dieser *exotischen Optionen* ist besonders aufwendig.

Schwerpunkt dieser Arbeit liegt in der Anwendungen der Sprung-Diffusion Stochastische Differentialgleichungen (SDGen) auf Bewertung von Optionen. Seit der bahnbrechende Arbeit von Black & Scholes (1973) hat sich das Optiongeschäft rasant entwickelt und schnell große Umsätze erreicht. Die Entwicklung des Optiongeschäft hat im Laufe der Jahre, zur stetigen Entwicklung neuer Optionsbewertungsmodelle geführt (siehe [4]). Speziell wird in dieser Arbeit *stochastic-volatility jump diffusion Modell* von Bakshi (siehe [4])

$$dS(t) = S(t)[R(t) - \lambda\mu_J]dt + \sqrt{V(t)}S(t)dW_S(t) + J(t)S(t)dq(t)$$

$$dV(t) = [\theta_v - \kappa_v V(t)]dt + \sigma_v \sqrt{V(t)}dW_v(t)$$

studiert. Mit dem *Bakshi* Modell lassen sich durch die Zunahme zufälliger Sprünge auch die Ereignisse wie "11. September" (2001) oder die "Öl Krise" (1972) modellieren.

In jüngster Zeit haben sich auf den Märkten außer *standardisierter* Optionen auch viele *exotische* Optionen etabliert. Da sich die *exotischen* Optionen nicht mit herkömmlichen Mitteln bewerten lassen, stellt sich die Frage nach einer geeigneten Methode. Die erste Überlegung ist die Monte-Carlo-Methode. Mit einem numerischen Verfahren für Sprung-Diffusion SDGen wird eine Menge von Trajektorien ( $\geq 10000$ ) simuliert und mittels der Monte-Carlo-Methode, die auch als das Gesetz der großen Zahlen interpretiert werden kann, werden aus der Menge der Trajektorien die Optionspreise bestimmt.

Nun, wie werden die Parameter aus dem Modell gewählt? Die Wahl der Modellparameter  $R_0, \lambda, V_0, \mu_J, \sigma_J, \theta_v, \kappa_v, \sigma_v, \rho$ , die in das Modell von Bakshi eingehen, werden durch numerische Optimierungsverfahren bewältigt, so dass

$$\sum_{i=1}^n [\hat{\sigma}_i - \text{impvolret}(C_i(\text{ModellPar}), \text{Market}, K_i, \text{Put})]^2$$

minimiert wird, wobei  $\hat{\sigma}$  Marktvolatilität und  $\sigma$  Implizite Volatilität bezeichnet.

Die implizite Volatilität (IV) ist nach Black-Scholes die Zahl, die als Volatilitätsparameter in die Black-Scholes Formeln eingegeben wird, so dass der so berechnete Optionspreis mit dem Marktpreis übereinstimmt. Zur Bestimmung der IV wird ein iteratives Optimierungsverfahren (z.B. Newton-Raphson, oder Bisektion Verfahren) herangezogen.

Die vorliegende Arbeit ist 'chronologisch' gegliedert. Der Schwerpunkt in den ersten drei Kapiteln beruht auf den SDGen. Nachdem im ersten Kapitel die Grundlage und benötigte Stochastische Prozesse (Wiener-, Poisson-Prozess) beschrieben werden, werden in den folgenden zwei Kapiteln (Kapiteln zwei und drei) sowohl die



SDGen als auch die Sprung Diffusion SDGen und deren Numerische Lösungsmethoden behandelt.

Im Laufe der Jahre hat die Entwicklung des Optionsgeschäfts zur stetigen Entwicklung neuer Optionsbewertungsmodelle geführt, was auch das vierte Kapitel beschreibt. Besonders wird hier das *stochastic-volatility jump diffusion Modell* von Bakshi untersucht, mit dem sich auch die am Markt sichtbaren zufälligen Sprünge modellieren lassen.

Obwohl das Black-Sholes Modell das älteste Optionsbewertungsmodell darstellt, werden heute noch einige Anwendungen des Modells, wie Implizite Volatilität oder die Sensitivitätskennzahlen, auch Greeks genannt, zu Hilfe genommen, was auch das Kapitel fünf beschreibt.

Das sechste Kapitel beschäftigt sich mit den *exotischen* Optionen und beschreibt im kurzen die Monte-Carlo-Methode.

Im siebten Kapitel werden die später untersuchten Gradienten-, Simulated Annealing- und generationelle Optimierungsverfahren dargestellt.

Den Kern dieser Arbeit bildet das letzte Kapitel. Nachdem zuerst die Parameter Calibration beschrieben wird, untersucht man das Bakshi Modell am Beispiel der Devisen-Optionen (*"Foreign Exchange" FX-Optionen*). Zu einem untersucht man das Verhalten des Modells vom Bakshi im Bezug auf die Parameter Calibration zu verschiedenen Laufzeiten und zu anderen werden auch die im siebten Kapitel beschriebenen Optimierungsverfahren untersucht. Eine mehrmonatige Untersuchung ergab ein sehr gutes Verhalten des Bakshi Modells, nicht nur für mittelfristige (ist der Fall bei Heston Modell, siehe [12]), sondern auch für kurze und langfristige Optionen. Die besten Ergebnisse wurden durch die Kombination des Simulated Annealing und generationellen Optimierungsverfahren erzielt. Zuerst wurde ein grober Punkt mit der Simulated Annealing Methode gesucht, der dann weiter mit dem generationellen Verfahren verfeinert wurde.



# Danksagung

Ich möchte mich bei all jenen bedanken, die direkt oder indirekt zur Entstehung dieser Diplomarbeit in Form fachlicher oder anderweitiger Unterstützung beigetragen haben.

Besonderer Dank gebührt dabei Herrn Prof. Dr. Peter E. Kloeden und Herrn Dr. Uwe Wystup für die Betreuung meiner Arbeit. Ihre zahlreichen wissenschaftlichen Ratschläge waren für mich sehr wertvoll und haben stets zur Verbesserung dieser Arbeit beigetragen.



# Kapitel 1

## Stochastische Prozesse

Das Beispiel der Wasserstandmessung in einem Zeitintervall  $[t_0, T]$ , die Veränderung des Aktienkurses in einem Zeitintervall  $[t_0, T]$  oder die Beschreibung des Ortes eines Partikels unter Brownscher Bewegung als Funktion der Zeit machen die gleichzeitige Betrachtung einer Familie von Zufallsgrößen nötig, die von einem stetigen Parameter abhängt. Die Familie von Zufallsgrößen fasst man unter dem Begriff Stochastischer Prozess zusammen.

**Definition 1.0.1** Sei  $\Omega$  eine nichtleere Menge. Eine Menge  $\mathcal{A}$  von Teilmengen von  $\Omega$  heißt sigma-Algebra (in  $\Omega$ ), falls

- $\Omega \in \mathcal{A}$ ,
- aus  $A \in \mathcal{A}$  auch  $\bar{A} \in \mathcal{A}$  folgt,
- aus  $A_i \in \mathcal{A}$  für  $i \in \mathbb{N}$  auch  $\bigcup_{i=0}^{\infty} A_i \in \mathcal{A}$  folgt.

In einer  $\sigma$ -Algebra sind immer  $\Omega$  (das sichere Ereignis) und  $\emptyset$  (das unmögliche Ereignis) enthalten. Die einfachste  $\sigma$ -Algebra in  $\Omega$  ist offensichtlich

$$\mathcal{A} = (\emptyset, \Omega)$$

und die umfangreichste ist offensichtlich

$$\mathcal{A} = \mathcal{P}(\Omega),$$

wobei  $\mathcal{P}(\Omega)$  die Potenzmenge von  $\Omega$  bezeichnet, also die Menge aller Teilmengen von  $\Omega$ .

**Definition 1.0.2** Sei  $\Omega$  die Ergebnismenge eines Zufallsexperimentes und  $\mathcal{A}$  eine  $\sigma$ -Algebra von Ereignissen, dann heißt die Abbildung (Zuordnung)

$$\begin{aligned} P : \mathcal{A} &\rightarrow [0, 1] \\ A &\rightarrow P(A) \end{aligned}$$

eine Wahrscheinlichkeit, falls gilt:

- $P(A) \geq 0$  (Nichtnegativität),
- $P(\Omega) = 1$  (Normierung)
- $P(\bigcup_{i=0}^{\infty} A_i) = \sum_{i=0}^{\infty} P(A_i)$  ( $\sigma$ -Additivität für eine Folge  $A_1, A_2, \dots$  von paarweise disjunkten Ereignissen).

Das Tripel  $(\Omega, \mathcal{A}, P)$  heist ein Wahrscheinlichkeitsraum.

**Definition 1.0.3** Sei  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum und  $(X_t)_{t \in T}$  eine Familie von Zufallsvariablen, d.h. eine Familie von messbaren Abbildungen  $X_t : \Omega \rightarrow \mathbb{R}$ , wobei

$$\omega \rightarrow X_t(\omega).$$

Dann heißt  $(X_t)_{t \in T}$  stochastischer Prozess.

Einen stochastischen Prozess kann man sich als eine Abbildung

$$\begin{aligned} X : \Omega \times T &\rightarrow \mathbb{R} \\ (\omega, t) &\rightarrow X(\omega, t) \end{aligned}$$

vorstellen. Er hängt sowohl vom Zufall als auch von der Zeit ab. Hält man nun die Zeit  $t$  fest, so erhält man die Zufallsvariable

$$\begin{aligned} X(\cdot, t) = X_t : \Omega &\rightarrow \mathbb{R} \\ \omega &\rightarrow X_t(\omega) \end{aligned}$$

Hält man dagegen  $\omega$  fest, so erhält man einen Pfad (eine sogenannte Trajektorie oder Realisation) des Prozesses, nämlich  $X(\omega, \cdot)$ , d.h. eine Funktion

$$\begin{aligned} T &\rightarrow \mathbb{R} \\ t &\rightarrow X(\omega, t). \end{aligned}$$

**Definition 1.0.4** Eine Folge  $(\varepsilon_t)_{t \in \mathbb{R}}$  von unabhängigen und identisch standardnormalverteilten Zufallsvariablen heißt Gaußscher Prozess.

## 1.1 Martingale

**Definition 1.1.1** Sei  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum. Eine aufsteigende Familie  $(\mathcal{A}_t)_{t \in [0, \infty)}$  von Unter- $\sigma$ -Algebren von  $\mathcal{A}$ , so dass für  $0 \leq s < t$  gilt

$$\mathcal{A}_s \subset \mathcal{A}_t \subset \mathcal{A},$$

heißt Filtration

**Definition 1.1.2** Ein stochastischer Prozess  $(X_t)_{t \in [0, \infty)}$  auf dem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, P)$  heißt adaptiert an  $(\mathcal{A}_t)_{t \in [0, \infty)}$ , wenn  $X_t$  bzgl.  $\mathcal{A}_t$  messbar ist für jedes  $t \in [0, \infty)$ .

Die einfachste Filtration ist durch

$$\mathcal{A}_t := \sigma(X_0, X_1, \dots, X_t),$$

$t \in \mathbb{N}_0$ , gegeben. Hierbei ist  $\mathcal{A}_t$  die durch  $X_0, X_1, \dots, X_t$  erzeugte  $\sigma$ -Algebra, also die kleinste  $\sigma$ -Algebra, bezüglich der  $X_0, X_1, \dots, X_t$  messbar sind. Sie heißt auch natürliche Filtration. Eine Filtration ist einfach zu verstehen, wenn man eine  $\sigma$ -Algebra als Informationsmenge interpretiert. Während die  $\sigma$ -Algebra  $\mathcal{A}$  alle Informationen enthält (bis in alle Ewigkeit), enthält  $\mathcal{A}_t$  nur Informationen bis zum Zeitpunkt  $t$ . Da die  $\sigma$ -Algebren aufsteigend sind, d.h.  $\mathcal{A}_s \subset \mathcal{A}_t$  für  $s < t$ , geht keine Information aus der Vergangenheit verloren.

Der Bezug zu einem Finanzmarkt, auf dem ein Finanzgut mit dem Preis  $X_t$  (zum Zeitpunkt  $t$ ) gehandelt wird, ist auch leicht herzustellen.  $\mathcal{A}_t$  enthält nur solche Informationen, die bis zum Zeitpunkt  $t$  bekannt sind, was sich mathematisch als Messbarkeit von  $X_t$  bezüglich  $\mathcal{A}_t$  ausdrückt.  $X_t$  muss nicht bezüglich  $\mathcal{A}_{t-1}$  messbar sein, denn es könnte sein, dass zum Zeitpunkt  $t$  neue Informationen vorhanden sind, die in  $\mathcal{A}_{t-1}$  noch nicht enthalten sind, aber in  $X_t$  zum Ausdruck kommen sollen. Demgegenüber ist  $X_t$  natürlich bezüglich  $\mathcal{A}_{t+1}$  messbar.  $X_t$  muss nicht sämtliche Informationen in  $\mathcal{A}_t$  widerspiegeln. Dies wäre jedoch der Fall, wenn  $\mathcal{A}_t := \sigma(X_0, X_1, \dots, X_t)$  gelten würde, d.h. wenn die durch  $X_0, X_1, \dots, X_t$  erzeugte  $\sigma$ -Algebra mit  $\mathcal{A}_t$  übereinstimmen würde. Im Allgemeinen gilt jedoch  $\mathcal{A}_t \supset \sigma(X_0, X_1, \dots, X_t)$ , d.h.  $\mathcal{A}_t$  kann zusätzliche Informationen enthalten.

**Definition 1.1.3** Sei  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum. Ein zu  $(\mathcal{A}_t)_{t \in [0, \infty)}$  adaptierter Prozess  $(X_t)_{t \in [0, \infty)}$  mit  $\mathbb{E}(|X_t|) < \infty$  für  $t \in [0, \infty)$  heißt

- Martingal, falls  $\mathbb{E}(X_t | \mathcal{A}_s) = X_s$  für  $s \leq t$ ,
- Super-Martingal, falls  $\mathbb{E}(X_t | \mathcal{A}_s) \leq X_s$  für  $s \leq t$ ,
- Sub-Martingal, falls  $\mathbb{E}(X_t | \mathcal{A}_s) \geq X_s$  für  $s \leq t$ .

Im zeitdiskreten Fall lässt sich die Beziehung  $\mathbb{E}(X_{t+1}|\mathcal{A}_t) = X_t$  leicht interpretieren, falls  $X_t$  den Preis eines Wertpapiers zum Zeitpunkt  $t$  und  $\mathcal{A}_t$  die bis  $t$  verfügbaren Informationen bezeichnen. Dann ist, falls  $(X_t)_{t \in \mathbb{N}_0}$  ein Martingal bildet, der im Zeitpunkt  $t+1$  zu erwartende Preis auf der Basis der Informationen bis zum Zeitpunkt  $t$  gerade  $X_t$ , d.h. der tatsächliche Preis zum Zeitpunkt  $t$ . Der bedingte Erwartungswert  $\mathbb{E}(X_{t+1}|\mathcal{A}_t) = X_t$  ist aber gerade die beste Approximation (im Sinne des quadratischen Fehlers) von  $X_{t+1}$  auf der Basis der Information bis zum Zeitpunkt  $t$ . Also, falls  $(X_t)_{t \in \mathbb{N}_0}$  ein Martingal bildet, so ist  $X_t$  die beste Approximation von  $X_{t+1}$  auf der Basis der Information bis  $t$ .

## 1.2 Wiener Prozess

Für das Beispiel der Brownscher Bewegung hat Norbert Wiener (1930) ein mathematisches Modell, heute Wiener Prozess genannt, vorgeschlagen.

**Definition 1.2.1** *Stochastischer Prozess  $W = (W_t, t \in [0, \infty))$  wird Wiener Prozess genannt (siehe [2]), falls  $W_0 = 0$  mit Wahrscheinlichkeit 1 gilt und die folgende Punkte erfüllt sind:*

- *Wiener Prozess ist ein Gaußsche Prozess, d.h.  $\forall s \geq 0$  und  $\forall t > 0$  Zufallsvariable  $W_{t+s} - W_t$  ist Normalverteilt mit Erwartungswert 0 und der Varianz  $t$ ;*
- *Alle Zuwächse  $\Delta W$  sind unabhängig voneinander; d.h. die  $W_{t_2} - W_{t_1}$  und  $W_{t_4} - W_{t_3}$  sind unabhängig  $\forall 0 \leq t_1 < t_2 < t_3 < t_4$ ;*
- *$W$  ist mit Wahrscheinlichkeit 1 stetig.*

**Bemerkung 1.2.2** *Weiter gilt:*

- $W_t - W_s, t > s$  hat die gleiche Verteilung wie  $W_{t-s}$ , nämlich  $N(0, t-s)$ ;
- $W_{ts}, s > 0$  hat die gleiche Verteilung wie  $\sqrt{t}W_s$ , nämlich  $N(0, ts)$ ;
- $\mathbb{E}(W_t) = 0, \forall t > 0$ ;
- $Cov(W_t, W_s) = \mathbb{E}(W_t W_s) - \mathbb{E}(W_t)\mathbb{E}(W_s) = s$ , falls  $s \leq t$ ,  
(man schreibt  $W_t W_s$  als  $(W_t - W_s)W_s + (W_s)^2$  und nutzt die Unabhängigkeit von  $W_t - W_s$  und  $W_s$ ).

**Theorem 1.2.3** *Die Pfade eines Wiener-Prozesses haben unbeschränkte Variation auf jedem endlichen Intervall, d.h.*

$$\sup \sum_{j=0}^{N_{T-1}} \| W_{t_{j+1}}(\omega) - W_{t_j}(\omega) \| = +\infty,$$



wobei  $t_0 \leq t_1 \leq \dots \leq t_{N_T} \leq T$  Unterteilung des Intervalls  $[t_0, T]$  repräsentiert.

**Theorem 1.2.4** Wiener Prozess  $(W_t)_{t \geq 0}$  bezüglich der natürlicher Filtration  $(\mathcal{A}_t)_{t \geq 0}$ , mit

$$\mathcal{A}_t := \sigma(W_s, 0 \leq s \leq t)$$

ist ein Martingal.

**Beweis:**

Ist ein Wiener Prozess  $(W_t)_{t \in [0, \infty)}$  auf  $(\Omega, \mathcal{A}, P)$  gegeben, so kann man leicht eine Filtration definieren, bezüglich welcher  $(W_t)_{t \in [0, \infty)}$  adaptiert ist. Sei

$$\mathcal{A}_t := \sigma(W_s, 0 \leq s \leq t)$$

d.h.,  $\mathcal{A}_t$  ist die von  $W_s$  mit  $0 \leq s \leq t$  erzeugte Unter- $\sigma$ -Algebra von  $\mathcal{A}$ . Offensichtlich gilt für  $s < t$

$$\mathcal{A}_0 = \{\Omega, \emptyset\} \subset \mathcal{A}_s \subset \mathcal{A}_t \subset \mathcal{A},$$

d.h.  $(\mathcal{A}_t)_{t \in [0, \infty)}$  ist eine aufsteigende Familie von  $\sigma$ -Algebren, also eine Filtration (die sogenannte natürliche Filtration). Per Definition ist  $W_t$  messbar bezüglich  $\mathcal{A}_t$ . Also ist  $(W_t)_{t \geq 0}$  adaptiert an  $(\mathcal{A}_t)_{t \geq 0}$ . Da  $W_t - W_s$  für  $t > s$  von allen  $W_r$  mit  $r \leq s$  unabhängig ist und  $\mathcal{A}_s = \sigma(W_r, r \leq s)$  gilt, folgt aus den Rechenregeln für bedingte Erwartungen, dass

$$\mathbb{E}(W_t - W_s | \mathcal{A}_s) = \mathbb{E}(W_t - W_s) = 0$$

gilt. Weiter gilt dann

$$0 = \mathbb{E}(W_t - W_s | \mathcal{A}_s) = \mathbb{E}(W_t | \mathcal{A}_s) - \mathbb{E}(W_s | \mathcal{A}_s) = \mathbb{E}(W_t | \mathcal{A}_s) - W_s, \text{ d.h. } \mathbb{E}(W_t | \mathcal{A}_s) = W_s.$$

Der Prozess  $(W_t)_{t \in [0, \infty)}$  ist also ein Martingal in Bezug auf die natürliche Filtration.

□

Es soll nun gezeigt werden, wie man einzelne Pfade  $t \rightarrow W_t$  des Wiener-Prozesses für  $t \in [0, T]$  simulieren kann. Man betrachtet hierzu eine Folge  $(\varepsilon_i)_{i \in \mathbb{N}}$  von unabhängigen standardnormalverteilten Zufallsvariablen (also einen Gaußschen Prozess) und konstruiert daraus folgenden zeitstetigen Prozess  $S^n(t)$  für  $0 \leq t \leq T$  und  $n \in \mathbb{N}$

$$S^n(t) = \begin{cases} \sqrt{\frac{T}{n}} \sum_{j=1}^i \varepsilon_j & \text{falls } t = T \frac{i}{n}, i = 0, \dots, n \\ \text{lineare Interpolation} & \text{sonst} \end{cases}$$

Für jedes  $n$  hat der stochastische Prozess  $(S^n(t))_{t \in [0, T]}$  folgende Eigenschaften

- $S^n(0) = 0$ ;

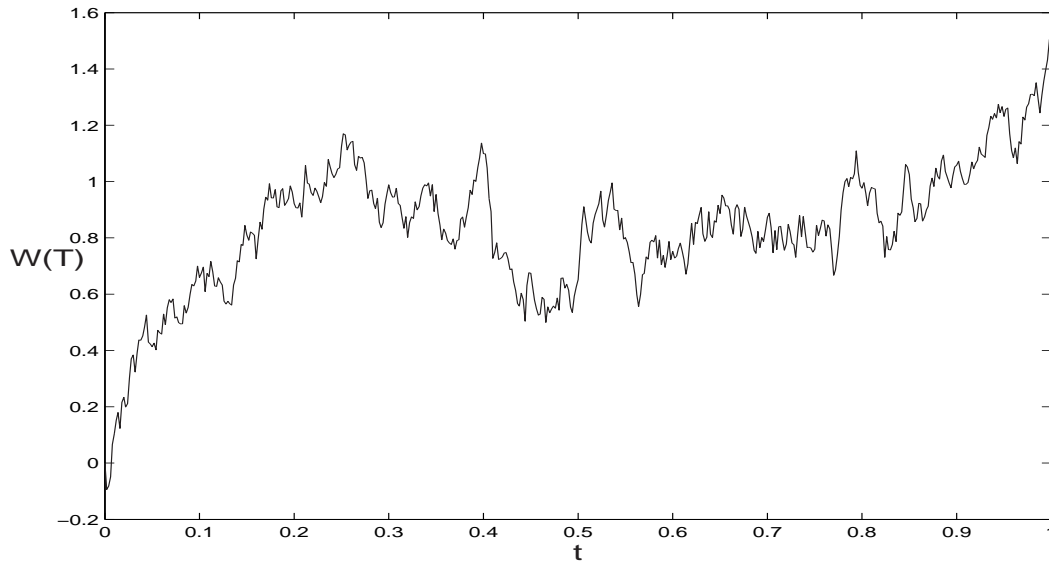


Abbildung 1.1: Wiener Prozess

- Für  $t = \frac{k}{n}T$  und  $s = \frac{r}{n}T$  mit  $r < k$  ist

$$S^n \left( \frac{k}{n}T \right) - S^n \left( \frac{r}{n}T \right)$$

normalverteilt mit Erwartungswert 0 und Varianz  $t - s$ ;

- Für  $0 < i_1 < \dots < i_m$  sind die Zufallsvariablen

$$S^n \left( \frac{i_2}{n}T \right) - S^n \left( \frac{i_1}{n}T \right), \dots, S^n \left( \frac{i_m}{n}T \right) - S^n \left( \frac{i_{m-1}}{n}T \right)$$

stochastisch unabhängig;

- Die Pfade  $t \rightarrow S^n(t)$  sind stetig.

Der Prozess  $(S^n(t))_{t \in [0, T]}$  schaut so ähnlich aus wie der Prozess  $(W_t)_{t \in [0, T]}$ . Das ist kein Zufall, denn Donsker hat gezeigt, dass die Folge der Prozesse  $(S^n(t))_{t \in [0, T]}$  in einem gewissen Sinne gegen  $(W_t)_{t \in [0, T]}$  konvergiert ([16] Seite 375). Den zugehörigen Grenzwertsatz nennt man auch Donskers Invarianzprinzip oder Funktionalen Zentralen Grenzwertsatz. Die Bezeichnung Grenzwertsatz ist sehr berechtigt, denn das Invarianzprinzip bleibt selbst dann gültig, wenn man statt Gaußverteilter  $\varepsilon_i$  andere Zufallsvariablen mit Erwartungswert 0 und Varianz 1 nimmt.

## 1.3 Poisson Prozess

**Definition 1.3.1** Ein stochastischer Prozess  $N_t, t \in [0, \infty)$  mit den Werten  $k = 0, 1, \dots$  auf dem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, P)$  heißt Poisson-Prozess mit der Intensität  $\lambda$ , falls gilt

1.  $P(N_0 = 0) = 1$ ;
2.  $\forall n \in \mathbb{N}$  und für jede Folge  $0 \leq t_0 < \dots < t_n < \infty$  sind die Zuwächse  $N_{t_1} - N_{t_0}, N_{t_2} - N_{t_1}, \dots, N_{t_n} - N_{t_{n-1}}$  stochastisch unabhängig;
3. die Zuwächse  $N_{t+h} - N_t$  sind poissonverteilt mit Parameter  $\lambda h > 0$ , d.h.

$$P(N_{t+h} - N_t = k) = \frac{(\lambda h)^k}{k!} e^{-\lambda h}, \quad \forall k = 0, 1, \dots$$

Die Intensität  $\lambda$  gibt die durchschnittliche Anzahl der eintretenden Ereignisse pro Zeiteinheit an.  $N_t$  zählt die Ereignisse in der Zeit  $t$ , und  $N_t - N_s$  gibt die Zahl der Ereignisse im Zeitintervall  $(s, t]$  an. Für die Parameter eines Poisson Prozesses gilt weiter (siehe [21])

$$\mathbb{E}(X(t)) = \lambda t, \quad \text{Var}(X(t)) = \lambda t$$

**Theorem 1.3.2** Sei  $(N_t, t \in [0, \infty))$  ein adaptierter Poisson Prozess mit Intensität  $\lambda$  bezüglich einer Filtration  $(\mathcal{A}_t)_{t \in [0, \infty)}$ . Dann sind die Prozesse

1.  $N(t) - \lambda t$
2.  $e^{\log(1-u)N(t)+u\lambda t}, \forall 0 < u < 1$ .

*Martingale.*

**Beweis:**

1. Für  $s < t$  gilt

$$\mathbb{E}(N_t - \lambda t \mid \mathcal{A}_s) = \mathbb{E}(N_t - N_s + N_s - \lambda t \mid \mathcal{A}_s) = \lambda(t - s) + N_s - \lambda t = N_s - \lambda s$$

- 2.

$$\begin{aligned} \mathbb{E}(e^{\log(1-u)N(t+s)} \mid \mathcal{A}_t) &= \mathbb{E}(e^{\log(1-u)N(t)+\log(1-u)(N(t+s)-N(t))} \mid \mathcal{A}_t) \\ &= e^{\log(1-u)N(t)} \mathbb{E}(e^{\log(1-u)(N(t+s)-N(t))} \mid \mathcal{A}_t) \\ &\quad (\text{weil } N(t) \text{ bezüglich } \mathcal{A}_t \text{ meßbar ist}) \\ &= e^{\log(1-u)N(t)} \mathbb{E}(e^{\log(1-u)(N(t+s)-N(t))}) \\ &\quad (\text{weil die Zuwächse unabhängig von } \mathcal{A} \text{ sind}) \\ &= e^{\log(1-u)N(t)} e^{-u\lambda s}. \end{aligned}$$

Multipliziert man beide Seiten mit  $e^{\log u\lambda(t+s)}$ , erhält man

$$e^{\log(1-u)N(t)+u\lambda t}.$$

□

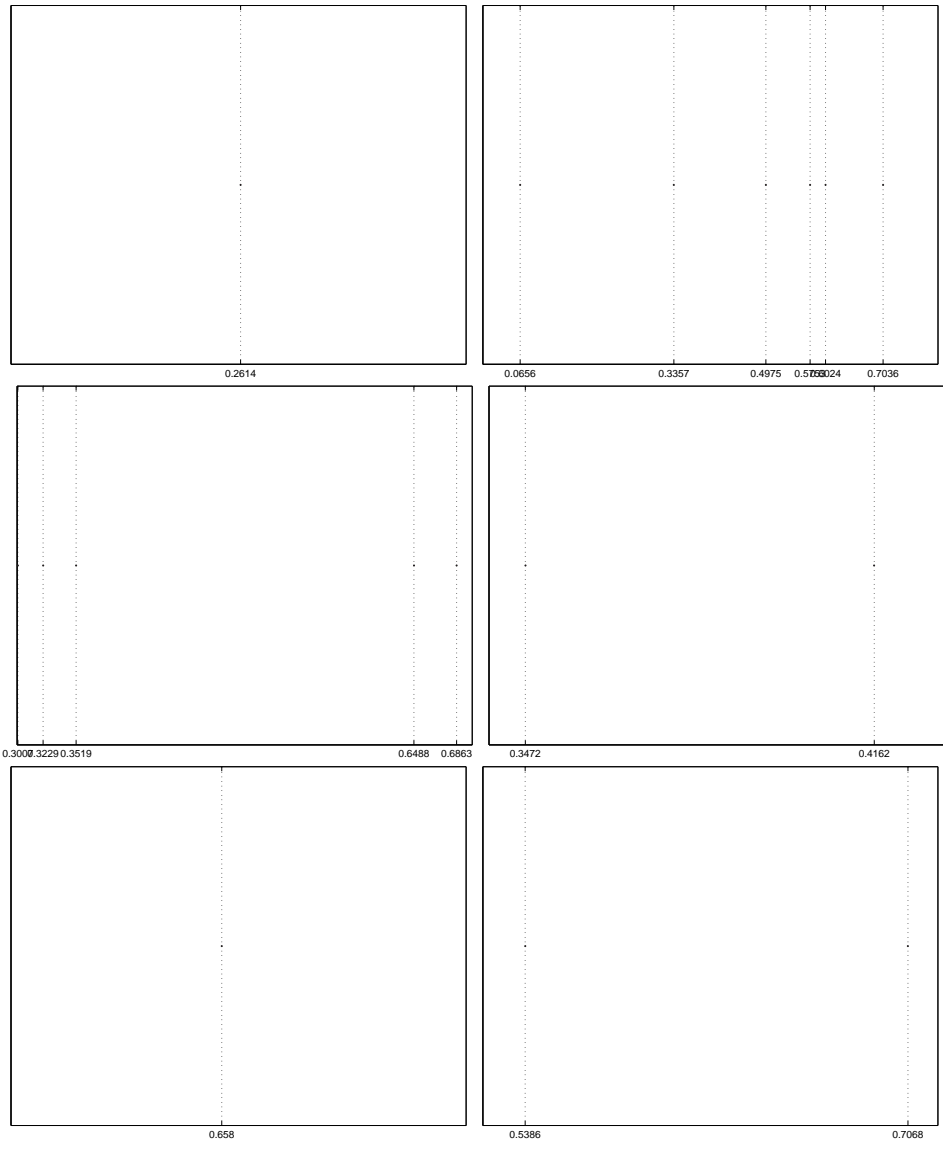


Abbildung 1.2: Realisierungen des Poisson-Prozesses mit  $\lambda = 3$ , im Zeitintervall  $T = [0, 1]$

# Kapitel 2

## Stochastische Differentialgleichungen

Die Aufnahme zufälliger Fluktuationen in die deterministische Differentialgleichungen stellt eine neue Klasse von Differentialgleichungen her, genannt stochastische Differentialgleichungen. Als Beispiel soll die Gleichung

$$\frac{dX}{dt} = a(t, X) + b(t, X)\xi_t$$

angeführt werden, wobei  $\xi_t$  das weiße Rauschen bezeichnet.

**Bemerkung 2.0.3** *Jedem stochastischen Prozess kann man eine verallgemeinerte Version zuordnen (vergleiche dazu [2], Seite 66). Fasst man den Wiener Prozess  $W_t$ , als verallgemeinerten stochastischen Prozess auf, dann kann  $W_t$  differenziert werden, so das gilt*

$$\xi_t = \frac{d}{dt}W_t,$$

oder umgekehrt

$$W_t = \int_0^t \xi_s ds.$$

**Definition 2.0.4** *Sei  $W = W_t, t \geq 0$ , ein Wiener Prozess und seien*

$$a, b : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$$

*reelle Funktionen. Dann wird der Ausdruck*

$$dX = a(t, X)dt + b(t, X)dW_t$$

*stochastische Differentialgleichung genannt, und symbolisch als Integralgleichung interpretiert*

$$X_t = X_{t_0} + \int_{t_0}^t a(s, X_s)ds + \int_{t_0}^t b(s, X_s)dW_s.$$

Dabei ist das erste Integral ein Riemannsches und das zweite ein stochastisches Integral. Fasst man den stochastischen Integral als ein Ito-Integral auf dann nennt man  $X_t$  Ito-Prozess.

Stellt man sich  $X_t$  als Wert einer Aktie zum Zeitpunkt  $t$ ,  $X_{t_0}$  als Wert der Aktie zum heutigen Zeitpunkt vor, so ist durch den deterministischen Term  $a(t, X)dt$  der so genannte Drift-Term und durch den stochastischen Term  $b(t, X)dW_t$  der so genannte Diffusions-Term gegeben.

## 2.1 Stochastische Integrale

Wie schon im letzten Kapitel beschrieben, sind die stetigen Realisierungen von  $W$  fast sicher nicht von beschränkter Variation. Daher kann man das Integral auch nicht pfadweise durch das gewöhnliche Integral definieren. Nahliegend ist zunächst

$$\int_{t_0}^t dW_s = W_t - W_{t_0}.$$

Hiermit ist schon klar, dass das Integral eine Zufallsvariable sein wird. Ist  $t_0 < t_1 < \dots < t$  und  $f$  stochastische Sprungfunktion:

$$f(t, \omega) = \sum_{v=0}^n e_v(\omega) \chi_{(t_v, t_{v+1}]}(t),$$

wobei  $\chi_G(t)$  indicator Funktion ist, so dass

$$\chi_G(t) = \begin{cases} 1 & \text{falls } t \in G \\ 0 & \text{falls } t \notin G \end{cases}$$

und  $\omega$  - Realisierung eines stochastischen Prozesses, definiert man

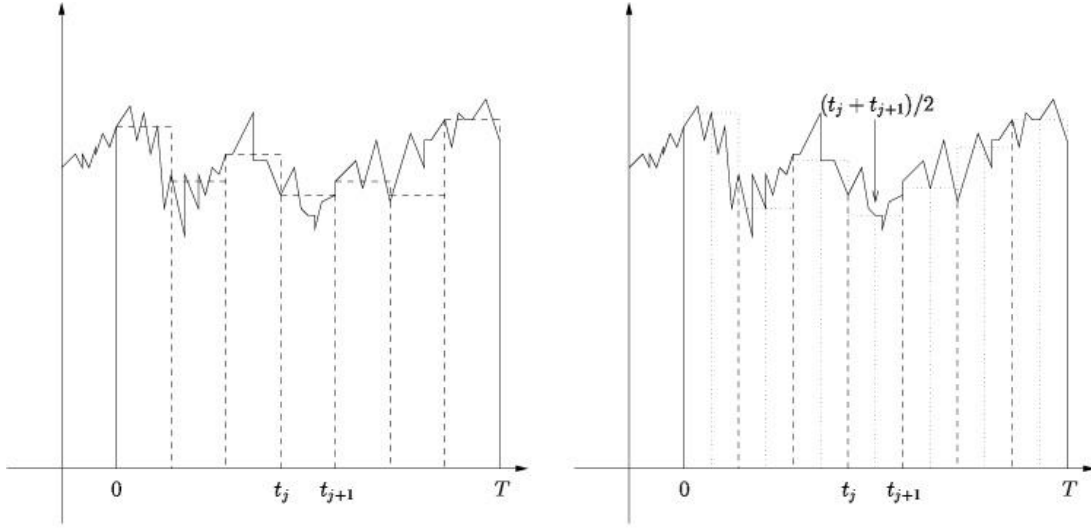
$$I[f](\omega) = \int_{t_0}^t f(s, \omega) dW_s(\omega) := \sum_{v=0}^n e_v(\omega) (W_{t_{v+1}}(\omega) - W_{t_v}(\omega)).$$

Zu gegebenem  $f$  wird in jedem Fall eine stochastische Sprungfunktion

$$f_n^* := \sum_{v=0}^n f(t_v^*, \cdot) \chi_{(t_v, t_{v+1}]}(t), \quad t_v^* \in [t_v, t_{v+1}]$$

zur Definition des Integrals heranziehen. Dabei ergibt sich für

- $t_v^* = t_v$  das Ito-Integral;
- $t_v^* = \frac{t_v + t_{v+1}}{2}$  das Stratonovich-Integral.



(a) Itô interpretation

(b) Stratonovich interpretation

Abbildung 2.1: Ito/Stratonovich Interpretation

Meistens wird eine stochastische Differentialgleichung im Sinne von Ito interpretiert, aber in verschiedenen Zusammenhängen ist es vorteilhaft, diese im Sinne von Stratonovich zu interpretieren. Stochastische Differentiale, die in Termen von Stratonovich Integralen ausgedrückt sind, gehorchen nämlich den Transformationsregeln der klassischen Differential- und Integralrechnung, während das für die Ito Interpretation nicht gilt (insbesondere die gewöhnliche Kettenregel). Die Ito-Formel gibt die Kettenregel für die Ito-Integrale an.

**Theorem 2.1.1 (Ito Formel)** Sei

$$dX_t = a_t dt + b_t dW_t$$

ein stochastisches Differential,

$$a, b : [0, T] \times \Omega \rightarrow \mathbb{R}$$

reelle Funktionen,

$$U : [0, T] \times \Omega \rightarrow \mathbb{R}$$

eine stetige zwei mal differenzierbare Funktion und  $Y_t, 0 \leq t \leq T$  stochastische Prozess, so dass mit Wahrscheinlichkeit 1 gilt

$$Y_t = U(t, X_t).$$

Dann ist das stochastische Differential für  $Y_t$  gegeben durch

$$dY_t = \left\{ \frac{\partial U}{\partial t}(t, X_t) + a \frac{\partial U}{\partial x}(t, X_t) + \frac{1}{2} b^2 \frac{\partial^2 U}{\partial x^2}(t, X_t) \right\} dt + b \frac{\partial U}{\partial x}(t, X_t) dW_t.$$

## 2.2 Numerische Methode zur Lösung der SDGen

Bei der numerischen Integration deterministischer Differentialgleichungen

$$\begin{aligned}x'(t) &= f(x(t), t), & t \in [t_0, T] \\x(t_0) &= x_0\end{aligned}$$

werden häufig bei kurzen Intervallen  $[0, T]$  die Runge-Kutta-Verfahren angewandt. Eine Verallgemeinerung dieses Vorgehens bilden die Einschrittverfahren. Bei diesen wird ein Integrationsintervall  $[0, T]$  durch Unterteilungspunkte  $\{t_i\}_{i=0, \dots, n}$  partitioniert, in jedem Unterteilungspunkt eine (lokale) Zuwachsfunktion berechnet, die das Teilintegral  $x(t_{i+1}) - x(t_i)$  approximiert, und somit aus den lokalen Zuwachsfunktionen das Gesamtintegral über  $[0, T]$  berechnet.

Beim Lösen stochastischer Differentialgleichungen kann analog verfahren werden. Man benötigt Unterteilung über dem Integrationsintervall, über dem Approximationsformeln definiert werden, ferner einen Konsistenzbegriff, eine lokale Konvergenzordnung und eine globale Fehlerabschätzung, Größen, die eine Wertung der Güte von Approximationsverfahren ermöglichen.

Das stochastische Anfangswertproblem, anhand dessen die wesentlichen Definitionen und Aussagen dieses Abschnitts getroffen werden, sei ein eindimensionale stochastische Anfangswertproblem mit einem Rauschterm.

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t \quad t \in [t_0, T]$$

$$X_{t_0} = X_0.$$

Es sei angenommen, dass die Parameter  $a$  und  $b$  derart gewählt seien, dass die Gleichung einen eindeutigen stochastischen Lösungsprozeß  $X = (X_t, [t_0, \infty), \mathcal{F}_{[t_0, \infty)})$  besitzen, dass insbesondere  $a$  und  $b$  den üblichen Lipschitz- und Wachstumsbedingungen genügen

$$a, b \in \mathcal{C}([t_0, T] \times \mathbb{R}, \mathbb{R}), \quad (t_0 < T) \in \mathbb{R}$$

$\exists K_L$  und  $K_G$  mit

$$|a(t, x) - a(t, y)| + |b(t, x) - b(t, y)| \leq K_L |x - y| \quad \text{und}$$

$$|a(t, x)|^2 + |b(t, x)|^2 \leq K_G(1 + |x|^2)$$

für alle  $(t, x), (t, y) \in [t_0, T] \times \mathbb{R}$ .

Nachfolgend wird die stochastische Version einer Approximation, ein Approximationsprozeß, eingeführt. Ist  $Y$  ein zeitdiskreter Approximationsprozeß von  $X$  so wird  $Y_n, n \in \mathbb{N}$  geschrieben. Konsistenz und Konvergenz haben die gleiche Bedeutung wie im deterministischen Fall, wobei jedoch bei der Konvergenz zwischen schwacher und starker Konvergenz unterschieden werden muss.



**Definition 2.2.1** Gegeben sei eine Maximalschrittweite  $\delta \in (0, \delta_0)$ . Unter einer Zeitdiskretisierung

$$(\tau)_\delta = \{\tau_n : n = 0, 1, \dots\}$$

des Intervalls  $[t_0, T]$  versteht man eine zufällige Folge  $\{\tau_n : n = 0, 1, \dots\}$  mit

- $t_0 =: \tau_0 < \tau_1 < \dots < \tau_{n_T} := T$
- $\sup_{n \in \{0, \dots, n_T-1\}} (\tau_{n+1} - \tau_n) \leq \delta$ , und
- $n_T < \infty$

mit Wahrscheinlichkeit 1, wobei  $\tau_{n+1}$  für alle  $n \in \{0, 1, \dots, n_T - 1\}$   $\mathcal{A}_{\tau_n}$ -messbar ist.  $\{\mathcal{A}_t, t \geq 0\}$  ist eine, meistens mit dem Ito-Prozess assoziierte, wachsende Familie von  $\sigma$ -Algebren. Dabei sei  $n_t$  für  $t \in [t_0, T]$  als das größte  $n$  definiert, für welches  $\tau_n$  nicht  $t$  übersteigt, d.h.

$$n_t = \max\{n \in \{0, 1, \dots\} : \tau_n \leq t\}.$$

**Definition 2.2.2**  $Y^\delta$  konvergiert stark gegen  $X$  zur Zeit  $T$ , falls

$$\lim_{\delta \downarrow 0} \mathbb{E}(|X_T - Y^\delta(T)|) = 0.$$

$Y^\delta$  konvergiert stark mit der Ordnung  $\gamma > 0$  gegen  $X$  zur Zeit  $T$ , falls es eine positive Konstante  $C$ , die nicht von  $\delta$  abhängt, und ein  $\delta_0$  gibt, so dass

$$\epsilon(\delta) = \mathbb{E}(|X_T - Y^\delta(T)|) \leq C\delta^\gamma, \quad \forall \delta \in (0, \delta_0).$$

**Definition 2.2.3**  $Y^\delta$  konvergiert schwach gegen  $X$  zur Zeit  $T$  für eine Funktion  $g : \mathbb{R}^d \rightarrow \mathbb{R}$ , falls

$$\lim_{\delta \downarrow 0} |\mathbb{E}(g(X_T)) - \mathbb{E}(g(Y^\delta(T)))| = 0.$$

$Y^\delta$  konvergiert schwach mit der Ordnung  $\beta > 0$  gegen  $X$  zur Zeit  $T$ , wenn es für jedes  $g \in C_P^{2(\beta+1)}(\mathbb{R}^d, \mathbb{R})$  <sup>1</sup> eine positive Konstante  $C_g$ , die nicht von  $\delta$  abhängt, und ein (endliches)  $\delta_0 > 0$  gibt, so dass

$$\mu(\delta) = |\mathbb{E}(g(X_T)) - \mathbb{E}(g(Y^\delta(T)))| \leq C_g \delta^\beta, \quad \forall \delta \in (0, \delta_0).$$

Die Konstruktion eines Verfahrens, das sich auf diskrete Zufallsvariablen stützt und mit einer fixierten Schrittweite arbeitet, erfolgt auf der stochastischen Taylor-Formel. Unter hinreichenden Glattheitsvoraussetzungen an Driftkoeffizient  $a$  und Diffusionskoeffizient  $b$  und durch iteriertes Anwenden der Ito-Formel kann  $X_t$  um  $X_{t_0}$  in eine

<sup>1</sup> $C_P^l(\mathbb{R}^d, \mathbb{R})$  bezeichnet den Raum der  $l$ -mal stetig differenzierbaren Funktionen  $g : \mathbb{R}^d \rightarrow \mathbb{R}$ , die zusammen mit ihren partiellen Ableitungen bis zur und einschließlich der Ordnung  $l$  polynomiales Wachstum haben.

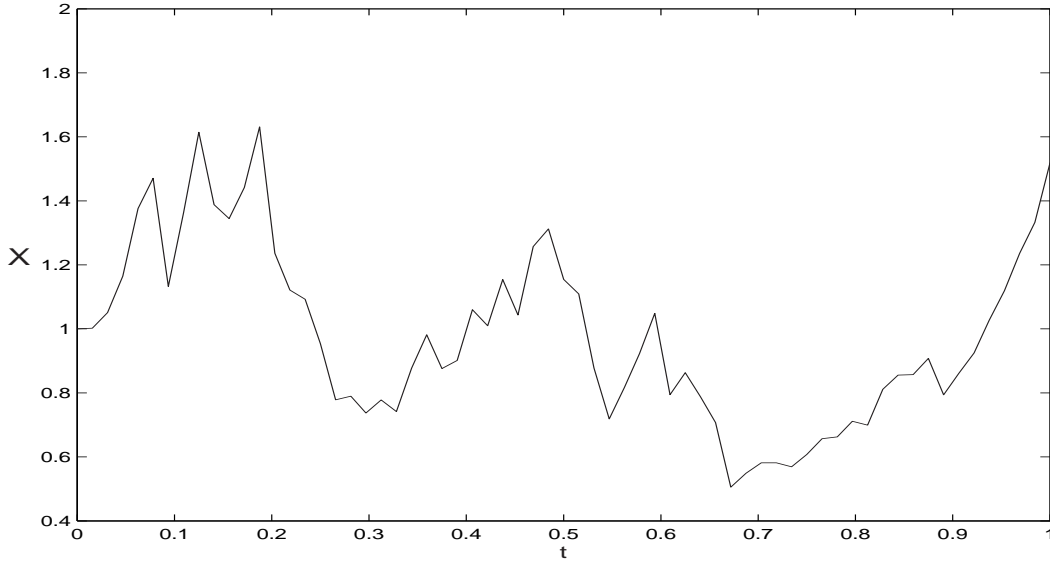


Abbildung 2.2: Eine Trajektorie der SDG, simuliert durch das Euler-Maruyama-Verfahren,  $a=2$ ,  $b=1$ ,  $Y_0 = 1$ ,  $\Delta t = 2^{-6}$ ,  $t = [0, 1]$ .

Reihe von Termen stochastischer Mehrfachintegrale mit bei  $X_{t_0}$  ausgewerteten Koeffizienten entwickelt werden, was auch eine stochastische Verallgemeinerung der deterministischen Taylor-Formel darstellt (siehe [18]).

Soll ein stochastisches numerisches Verfahren schwach mit der Ordnung  $\beta$  konvergieren, dann muß seine Konstruktion auf einer abgeschnittenen Ito-Taylor-Entwicklung beruhen, die gerade alle  $\beta$ -fachen Integrale mit konstanten Integranden enthält. So ergibt sich, wenn  $t_0 = \tau_0 \leq \tau_1 \leq \dots \leq \tau_N = T$  eine Zeitdiskretisierung des Intervalls  $[t_0, T]$  ist und wenn die Koeffizienten  $a$  und  $b$  nur von dem Zustand abhängen, aus der Entwicklung

$$X_{\tau_{n+1}} = X_{\tau_n} + a(\tau_n, X_{\tau_n}) \int_{\tau_n}^{\tau_{n+1}} ds + b(\tau_n, X_{\tau_n}) \int_{\tau_n}^{\tau_{n+1}} dW_s,$$

die lediglich alle einfachen Integrale enthält, die Iteration

$$Y_{n+1} = Y_n + a(t_n, Y_n)\Delta_n + b(t_n, Y_n)\Delta W_n$$

wenn  $\int_{\tau_n}^{\tau_{n+1}} ds$  durch  $\Delta_n := \tau_{n+1} - \tau_n$  und  $\int_{\tau_n}^{\tau_{n+1}} dW_s = W_{\tau_{n+1}} - W_{\tau_n}$  durch eine Zufallsvariable  $\Delta W_n$  mit  $\mathbb{E}(\Delta W_n) = 0$  und  $\mathbb{E}((\Delta W_n)^2) = \Delta_n$  substituiert werden.

Das so erhaltene Verfahren, genannt Euler-Maruyama Verfahren konvergiert schwach mit der Ordnung  $\beta = 1.0$  und stark mit der Ordnung  $\gamma = 0.5$ .

Mit der Zunahme des Terms

$$bb' \int_{\tau_n}^{\tau_{n+1}} \int_{\tau_n}^{\tau_s} dW_u dW_s = \frac{1}{2}b(t_n, Y_n)b'(t_n, Y_n)(\Delta W_n^2 - \Delta t)$$

in das Euler-Maruyama Verfahren erhält man das so genannte Milstein Verfahren

$$Y_{n+1} = Y_n + a(t_n, Y_n)\Delta_n + b(t_n, Y_n)\Delta W_n + \frac{1}{2}b(t_n, Y_n)b'(t_n, Y_n)(\Delta W_n^2 - \Delta t).$$

Das so erhaltene Verfahren konvergiert schwach mit der Ordnung  $\beta = 1.0$  und stark mit der Ordnung  $\gamma = 1.0$ .

Die Entwicklung

$$\begin{aligned} X_{\tau_{n+1}} = X_{\tau_n} &+ a \int_{\tau_n}^{\tau_{n+1}} ds + b \int_{\tau_n}^{\tau_{n+1}} dW_s + (aa' + \frac{1}{2}b^2a'' \int_{\tau_n}^{\tau_{n+1}} \int_{\tau_n}^{\tau_s} dud s \\ &+ (ab' + \frac{1}{2}b^2b'' \int_{\tau_n}^{\tau_{n+1}} \int_{\tau_n}^{\tau_s} dud W_s + ba' \int_{\tau_n}^{\tau_{n+1}} \int_{\tau_n}^{\tau_s} dW_u ds \\ &+ bb' \int_{\tau_n}^{\tau_{n+1}} \int_{\tau_n}^{\tau_s} dW_u dW_s \end{aligned}$$

liefert eine schwache Approximation der Ordnung 2.0.

Die Aufgabe, ein Approximation Verfahren zu konstruieren, reduziert sich also im wesentlichen darauf, die in der entsprechenden Ito-Taylor-Entwicklung vorkommenden stochastischen Mehrfachintegrale geeignet zu substituieren, d.h. unter Verwendung leicht zu generierender Zufallsvariablen bestimmte Momenteigenschaften dieser Mehrfachintegrale zu simulieren.



# Kapitel 3

## Sprung Diffusion SDGen

Durch die Zunahme der Sprungkomponente  $c(X_{t-}, t)dN_t$  in die Ito-SDG

$$dX_t = a(X_t, t)dt + b(X_t, t)dW_t$$

erhält man eine neue Klasse von stochastischen Differentialgleichungen genannt Sprung-Diffusion-SDGen

$$dX_t = a(X_t, t)dt + b(X_t, t)dW_t + c(X_{t-}, t)dN_t,$$

wobei  $W_t$  ein Wiener und  $N_t$  ein inhomogener Poisson Zählprozess ist. Die Existenz und Eindeutigkeit des Prozesses folgt unter der Annahme der üblichen Wachstums, uniformen Lipschitz, und Glattheit Bedingungen an die Koeffizientenfunktionen  $a, b, c$  (siehe dazu [2],[10]).

Symbolische Interpretation der Sprung-Diffusion-SDG stellt eine Integralgleichung in der Form

$$X_t = X_{t_0} + \int_{t_0}^t a(s, X_s)ds + \int_{t_0}^t b(s, X_s)dW_s + \int_{t_0}^t c(s, X_{s-})dN_s$$

dar. Der erste Integral ist ein Riemann-Integral, zweite Integral ein stochastische Ito-Integral und der dritte Integral ist ein stochastisches Integral hinsichtlich des Poisson Zählprozesses. Der Lösungsprozess wird auch Sprung-Diffusion Prozess genannt.

### 3.1 Numerische Methoden zur Lösung der Sprung-Diffusion SDGen

Im folgenden wird auf das numerische Verfahren für Sprung-Diffusion SDG in der Form

$$X_t = X_{t_0} + \int_{t_0}^t a(X_s, s)ds + \int_{t_0}^t b(X_s, s)dW_s + \int_{t_0}^t c(X_{s-}, s)dN_s$$

eingegangen, wobei  $X_{t_0} = X_0$  mit Wahrscheinlichkeit 1 gilt und  $t \in [t_0, T]$ . Seien im folgenden weiterhin  $Y_n$  die Approximation des Prozesses  $X_t$  zur Zeit  $t_n$ , und  $\Delta W_n$  bzw.  $\Delta N_n$  die Zuwächse des Wiener bzw. Poisson Prozesses auf dem Intervall  $[t_n, t_{n+1}]$ . Dann ist das einfachste Verfahren das Euler Verfahren dass durch

$$Y_{n+1} = Y_n + a(t_n, Y_n)\Delta_n + b(t_n, Y_n)\Delta W_n + c(t_n, Y_n)\Delta N_n$$

gegeben ist (siehe dazu [22, 23, 24, 7]). Nach Magshoodi hat das so erhaltene Verfahren, im gewissen Sinne quadratische Varianz der erste Ordnung, d.h.  $O(h)$ , wobei  $h$  die maximale Schrittweite ist, was auch mit der starke Ordnung  $\gamma = \frac{1}{2}$  equivalent ist (vergleiche dazu [18]).

Das Verfahren, das quadratische Varianz zweiter Ordnung, bzw. starke Ordnung  $\gamma = 1$  besitzt und auf der stochastischen Taylor Entwicklung basiert, wird generalisiertes Milstein Verfahren für Sprung-Diffusion SDGen

$$\begin{aligned} Y_{n+1} &= Y_n + \left( a - \frac{1}{2}b \frac{\partial b}{\partial x} \right) \Delta_n + b\Delta W_n + \frac{1}{2}b \frac{\partial b}{\partial x} (\Delta W_n)^2 \\ &+ \frac{1}{2}(3c - c_c)\Delta N_n + (b_c - b)\Delta W_n \Delta N_n \\ &+ \frac{1}{2}(c_c - c)(\Delta N_n)^2 + \left( b \frac{\partial c}{\partial x} - b_c + b \right) \Delta Z_n, \end{aligned}$$

genannt.  $f_c(t, x)$ , für  $f$  entweder  $b$  oder  $c$ , ist definiert durch

$$f_c(t, x) = f(t, x + c(t, x))$$

und die Zufallsvariable  $\Delta Z_n$  durch das stochastische Integral

$$\Delta Z_n = \int_{t_n}^{t_{n+1}} \int_{t_n}^s dW_t dN_s = \int_{t_n}^{t_{n+1}} (W_s - W_{t_n}) dN_s.$$

Das Milstein-Magshoodi Verfahren (siehe [22, 23, 24]), das quadratische Varianz der zweite Ordnung besitzt, ist mittels der Zunahme der Sprung- Zeiten in die Teilung des Intervalls erreichbar. Die Teilung des Intervalls  $[t_0, T]$  ist durch  $t_0 = \tau_0, < \tau_1 \dots < \tau_{N_T} = T$  gegeben, so dass  $\max_{n=0, \dots, N_t-1} (\tau_{n+1} - \tau_n) \leq \Delta t$  mit Wahrscheinlichkeit 1 gilt.

Die Gleichung

$$\begin{aligned} Y_{n+1} &= Y_n + \left( a - \frac{1}{2}b \frac{\partial b}{\partial x} \right) \Delta_{\tau_n} + b\Delta W_{\tau_n} + \frac{1}{2}b \frac{\partial b}{\partial x} (\Delta W_{\tau_n})^2 \\ &+ c\Delta N_{\tau_n} + (b_c - b)\Delta W_{\tau_n} \Delta N_{\tau_n}, \end{aligned}$$

wobei  $Y_n$  die Approximation von  $X_{\tau_n}$  bezeichnet, stellt das Milstein-Magshoodi Verfahren mit adaptierten Sprüngen dar.

# Kapitel 4

## Bakshi Modell

In den letzten Jahrzehnten kam es auf dem Gebiet der derivativen Finanzinstrumenten, insbesondere bei Optionen, zu tiefgreifenden Veränderungen. Einerseits begann mit der Etablierung der Chicago Board Options Exchange im Jahre 1973 eine rasante Entwicklung in der Gründung von Finanzplätzen sowie im Handel mit solchen Finanzinstrumenten. Andererseits wurde die Theorie der Optionen in den siebziger Jahren durch die wegweisende Arbeit von Black und Scholes *revolutioniert*. Im Laufe der Zeit wurden auch andere Modelle, wie

- *stochastic-interest-rate* Modelle von Merton(1973) oder Amin & Jarrow (1992),
- *jump-diffusion / pure jump* Modelle von Bates(1991), Madan & Chang(1996) oder Morton(1996)
- *constant-elasticity-of-variance* Modell von Cox & Ross(1976)
- *Markowsche* Modelle von Rubenstein(1994) oder Ait-Sahalia & Lo(1996)
- *stochastic-volatility* Modelle von Heston(1993), Hull & White(1987), Scott(1987) oder Stein & Stein(1991)
- *stochastic-volatility jump diffusion* Modelle von Bates(1996), Scott(1997) oder Bakshi(1997)

vorgeschlagen. Im folgenden wird das Modell von Bakshi (siehe [4]) vorgestellt.

### 4.1 Modellannahmen

Bakshi verwendet in seinem Modell einen zufälligen Prozess mit Sprüngen

$$dS(t) = S(t)[R(t) - \lambda\mu_J]dt + \sqrt{V(t)}S(t)dW_S(t) + J(t)S(t)dq(t).$$

Dabei ist

- $W_S(t)$  ein Wiener Prozess,
- $R(t)$  der Zinssatz,
- $\lambda$  die durchschnittliche Anzahl der Sprünge in einem Jahr,
- $\mu_J$  der Erwartungswert der Sprünge
- Der Sprung-Term  $J(t)q(t)$  ist ein zusammengesetztes Poisson-Prozess<sup>1</sup>
  - $q(t)$  bezeichnet einen Poisson *counting* Prozess mit Intensität  $\lambda$ , so dass  $Ws\{dq(t) = 1\} = \lambda dt$  bzw.  $Ws\{dq(t) = 0\} = 1 - \lambda dt$ ,
  - Sprunghöhe  $J(t)$

$$\ln[1 + J(t)] \sim N(\ln[1 + \mu_J] - \frac{1}{2}\sigma_J^2, \sigma_J^2),$$

ist immer Lognormal, Identisch und Unabhängig Verteilt mit Erwartungswert  $\mu_J$  und Standardabweichung  $\sigma_J$  (die Volatilität der Sprünge).

Es wird angenommen, dass alle Wertpapiere am betrachteten Markt einem Prozess dieser Form folgen und dass die Volatilität  $V(t)$  einem *square root process*

$$dV(t) = [\theta_v - \kappa_v V(t)]dt + \sigma_v \sqrt{V(t)} dW_v(t)$$

folgt. Die Gleichung beschreibt einen *mean reversion process*  $V$ , dabei ist  $\theta_v$  das Quadrat der langfristigen instantanen Volatilität und  $\kappa_v$  die Stärke der *mean reversion*. Die Größe  $\sigma_v$  bezeichnet die Volatilität der Varianz, oftmals *vol of vol* genannt. Wiener Prozess  $W_v(t)$ , aus der letzten Gleichung, ist mit  $W_S(t)$  korreliert

$$Cov_t[dW_S(t), dW_v(t)] = \rho dt.$$

**Bemerkung 4.1.1** *Zu Erzeugung der korrelierten Zufallsvariablen greift man auf die Cholesky Zerlegung zurück (siehe [42]).*

*Aus zwei unabhängigen Zufallsvariablen  $X_1$  und  $X_2$  lassen sich durch Linearkombination leicht zwei korrelierte Zufallsvariablen erzeugen:*

$$\begin{aligned} Y_1 &= X_1 \\ Y_2 &= \rho X_1 + \sqrt{(1 - \rho^2)} X_2. \end{aligned}$$

*Sind speziell  $X_1$  und  $X_2$  unabhängig und standard normalverteilt, dann sind auch  $Y_1$  und  $Y_2$  standard normalverteilt mit der Korrelation*

$$Cor(Y_1, Y_2) = \rho.$$

---

<sup>1</sup> Einen stochastischen Prozess, der die gleichen Sprungstellen wie ein Poisson-Prozess, aber andere Sprunghöhen besitzt, bezeichnet man als zusammengesetzten Poisson-Prozess.



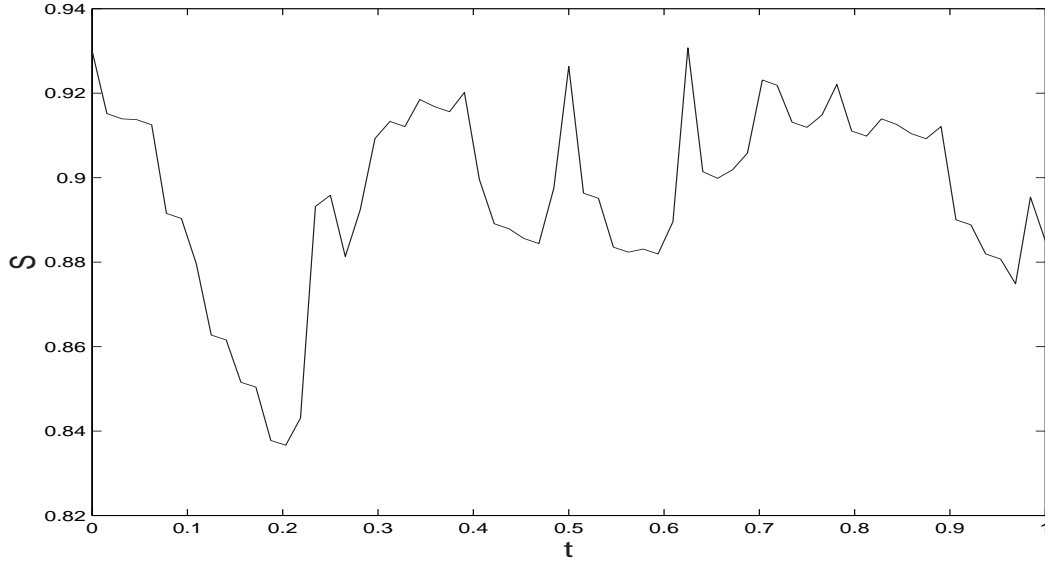


Abbildung 4.1: Simulierte Trajektorie des Bakshi Modells

## 4.2 Analytische Bewertungsformel

Die analytische Formel für den europäischen Call ist durch

$$C(t, \tau) = S(t)\Pi_1(t, \tau; S, R, V) - KB(t, \tau)\Pi_2(t, \tau; S, R, V)$$

gegeben (siehe [4]), wobei für die Risiko-Neutrale Wahrscheinlichkeiten  $\Pi_j (j = 1, 2)$  gilt

$$\Pi_j(t, \tau; S(t), R(t), V(t)) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \operatorname{Re} \left[ \frac{e^{-i\phi \ln[K]} f_j(t, \tau, S(t), R(t), V(t); \phi)}{i\phi} \right] d\phi.$$

Die charakteristische Funktionen  $f_j (j = 1, 2)$ , aus der obige Gleichung, definiert Bakshi als

$$\begin{aligned} f_1(t, \tau) = & \exp \left\{ -\frac{\theta_R}{\sigma_R^2} \left[ 2\ln \left( 1 - \frac{[\xi_R - \kappa_R](1 - e^{-\xi_R \tau})}{2\xi_R} \right) + [\xi_R - \kappa_R]\tau \right] \right. \\ & - \frac{\theta_v}{\sigma_v^2} \left[ 2\ln \left( 1 - \frac{[\xi_v - \kappa_v + (1 + i\phi)\rho\sigma_v](1 - e^{-\xi_v \tau})}{2\xi_v} \right) \right] \\ & - \frac{\theta_v}{\sigma_v^2} [\xi_v - \kappa_v + (1 + i\phi)\rho\sigma_v]\tau + i\phi \ln[S(t)] + \frac{2i\phi(1 - e^{-\xi_R \tau})}{2\xi_R - [\xi_R - \kappa_R](1 - e^{-\xi_R \tau})} R(t) \\ & + \lambda(1 + \mu_J)\tau \left[ (1 + \mu_J)^{i\phi} e^{\frac{i\phi}{2}(1+i\phi)\sigma_J^2} - 1 \right] - \lambda i\phi \mu_J \tau \\ & \left. + \frac{i\phi(i\phi + 1)(1 - e^{-\xi_v \tau})}{2\xi_v - [\xi_v - \kappa_v + (1 + i\phi)\rho\sigma_v](1 - e^{-\xi_v \tau})} V(t) \right\} \end{aligned}$$

$$\begin{aligned}
f_2(t, \tau) = & \exp \left\{ -\frac{\theta_R}{\sigma_R^2} \left[ 2\ln \left( 1 - \frac{[\xi_R^* - \kappa_R](1 - e^{-\xi_R^* \tau})}{2\xi_R^*} \right) + [\xi_R^* - \kappa_R]\tau \right] \right. \\
& - \frac{\theta_v}{\sigma_v^2} \left[ 2\ln \left( 1 - \frac{[\xi_v^* - \kappa_v + i\phi\rho\sigma_v](1 - e^{-\xi_v^* \tau})}{2\xi_v^*} \right) + [\xi_v^* - \kappa_v + i\phi\rho\sigma_v]\tau \right] \\
& + i\phi \ln[S(t)] - \ln[B(t, \tau)] + \frac{2(i\phi - 1)(1 - e^{-\xi_R^* \tau})}{2\xi_R^* - [\xi_R^* - \kappa_R](1 - e^{-\xi_R^* \tau})} R(t) \\
& + \lambda\tau \left[ (1 + \mu_J)^{i\phi} e^{\frac{i\phi}{2}(i\phi-1)\sigma_J^2} - 1 \right] - \lambda i\phi\mu_J\tau \\
& \left. + \frac{i\phi(i\phi - 1)(1 - e^{-\xi_v^* \tau})}{2\xi_v^* - [\xi_v^* - \kappa_v + i\phi\rho\sigma_v](1 - e^{-\xi_v^* \tau})} V(t) \right\},
\end{aligned}$$

mit

$$\begin{aligned}
\xi_R &= \sqrt{\kappa_R^2 - 2\sigma_R^2 i\phi}, \quad \xi_v = \sqrt{[\kappa_v - (1 + i\phi)\rho\sigma_v]^2 - i\phi(i\phi + 1)\sigma_v^2} \\
\xi_R^* &= \sqrt{\kappa_R^2 - 2\sigma_R^2(i\phi - 1)}, \quad \xi_v^* = \sqrt{[\kappa_v - i\phi\rho\sigma_v]^2 - i\phi(i\phi - 1)\sigma_v^2}.
\end{aligned}$$

Detaillierte Herleitung der charakteristischen Funktion ist in [4] beschrieben.

**Bemerkung 4.2.1** *Setzt man  $\lambda = \theta_R = \theta_v = \sigma_R = \sigma_v = \kappa_R = \kappa_v = 0$  in die analytische Formel für den europäischen Call ein, hat man das Black-Scholes Modell und für  $\lambda = \theta_R = \sigma_R = \kappa_R = 0$  das Modell von Heston.*

# Kapitel 5

## Anwendungen des Black-Scholes Modells

Der Black-Scholes Preis für eine europäische Call-Option auf eine Aktie, die während der Laufzeit der Option keine Dividende zahlt, ist gegeben durch

$$C = SN(d) - Xe^{-r(T-t)}N(d - \sigma\sqrt{T-t})$$

mit

$$d = \frac{\ln\left(\frac{S}{X}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}},$$

wobei

- $S$  Kassapreis des Basiswertes
- $X$  Ausübungspreis der Option
- $\sigma$  Annualisierte Volatilität des Basiswertes
- $r$  kontinuierlicher Zinssatz
- $T$  Zeitpunkt des Verfalls der Option
- $t$  heutiger Zeitpunkt
- $N(\cdot)$  Verteilungsfunktion der Normalverteilung.

Eine der wichtigsten Aspekte des Black-Scholes Modells ist die Implizite Volatilität.

## 5.1 Implizite Volatilität

Der Volatilitätsparameter kann nicht von irgendwelchen Marktgrößen abgelesen oder hergeleitet werden, sondern ist eine am Devisenmarkt quotierte eigenständige Größe, die aus einer Schätzung hervorgehen muss. In den Optionspreisen spiegelt sich dann die aggregierte Schätzung aller Marktteilnehmer. Man spricht in diesem Zusammenhang von impliziter Volatilität. Die Marktpreise von Optionen kommen zustande durch die verschiedenen Volatilitätsschätzungen der Marktteilnehmer. Umgekehrt kann aus den Marktpreisen die implizite Volatilität herausgelesen werden.

Die implizite Volatilität nach Black-Scholes ist die Zahl, die als Volatilitätsparameter in die Black-Scholes Formel eingegeben wird, so dass der so berechnete Optionspreis mit dem Marktpreis übereinstimmt (siehe [15, 12, 31]). Wird beispielsweise am Markt für eine einjährige *Call-Option* mit Ausübungspreis 100 auf eine Aktie die \$ 100 kostet, \$ 8 bezahlt, bei einem einjährigen Kassazins von 4%, dann muss in die Black-Scholes Formel neben den entsprechenden Werten der anderen Parameter eine Volatilität von 15% eingesetzt werden, damit das Modell einen Preis von \$ 8 liefert. Die Option hat demnach eine implizite Volatilität von 15%. Die Black-Scholes Formel kann aber nicht einfach nach der Volatilität aufgelöst werden, sondern die implizite Volatilität muss mit einem iterativen Suchverfahren bestimmt werden. Vereinfachend gesagt werden dabei viele Volatilitätswerte ausprobiert, bis derjenige Wert gefunden ist, bei dessen Einsetzen in die Formel der Marktpreis der Option resultiert. *Die Lösung ist eindeutig im Falle ihrer Existenz.*

Schätzt ein Marktteilnehmer die Volatilität des Basiswertes für die Laufzeit einer bestimmten Option als kleiner ein als die beobachtete implizite Volatilität, dann sind die Optionen für diesen Marktteilnehmer teuer. Schätzt umgekehrt ein zweiter Marktteilnehmer die Volatilität höher ein als die implizite Volatilität, dann wäre die Option für diesen Marktteilnehmer "billig". Optionspreise am Markt können also über die implizite Volatilität charakterisiert und eingeschätzt werden.

Zur Gewinnung von Volatilitätswerten und für die Bewertung von Optionen ergeben sich im Wesentlichen drei Möglichkeiten, nämlich historische Volatilitäten, Schätzungen mit statistischen Modellen wie ARCH/GARCH (generalized autoregressive conditional heteroskedasticity) auf der Basis von historischen Volatilitäten, oder aufgrund von impliziten Volatilitäten, möglicherweise von anderen Optionen (siehe [15, 29]).

Wegen der mit historischen Volatilitätsschätzungen verbundenen Problemen wird oft auf implizite Volatilitäten ausgewichen. Die Idee dahinter ist, dass sich in der impliziten Volatilität die Schätzungen aller Marktteilnehmer aggregieren und so eine Art *Consensus Forecast* bilden. Weil statt nur einer einzigen subjektiven Schätzung viele Schätzungen in die implizite Volatilität einfließen, kann argumentiert werden, die implizite Volatilität enthalte mehr Informationen als andere Volatilitätsschätzungen und sei darum im Durchschnitt eine bessere Schätzung.

Eine Berechnung der impliziten Volatilität mit einem iterativen Verfahren er-

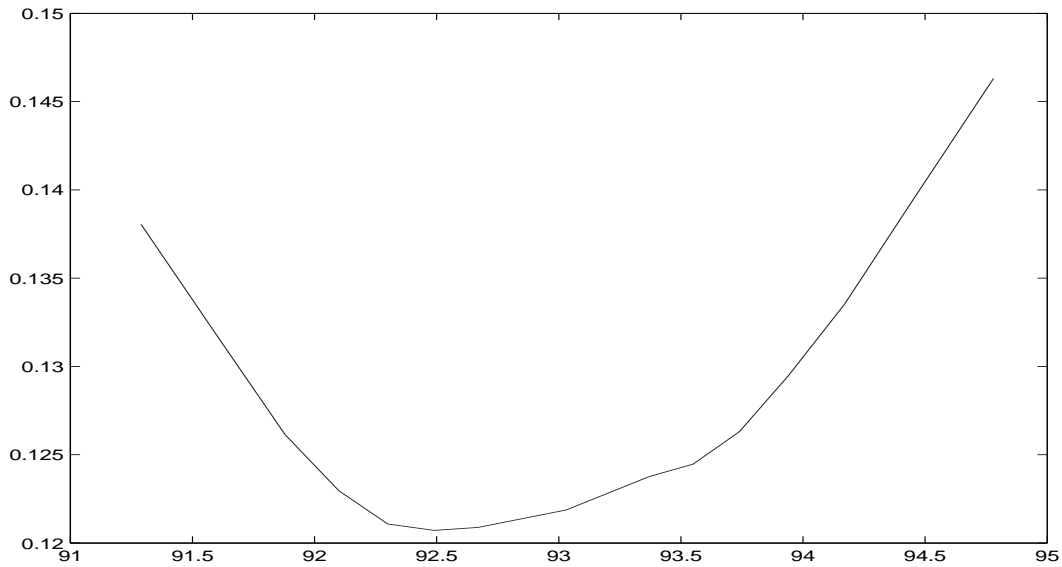


Abbildung 5.1: IV einer Call-option für unterschiedliche Strikes: Smile

gibt für verschiedene Ausübungspreise von Optionen auf denselben Basiswert und mit demselben Verfallsdatum oft verschiedene implizite Volatilitäten. Typische implizite Volatilitäten sind in Abbildung 5.1 und Abbildung 5.2 graphisch dargestellt. Dabei fällt auf, dass die Linie der impliziten Volatilitäten nicht horizontal verläuft, sondern im *at-the-money* Bereich am geringsten ist und ansteigt, je weiter die Option im oder aus dem Geld liegt. Ein solcher Verlauf der impliziten Volatilitäten wird *Smile* genannt. Abbildung 5.2 zeigt ein ähnliches Bild mit dem Unterschied, dass die impliziten Volatilitäten im *in-the-money* Bereich nicht ansteigen. Man spricht in diesem Fall von einem *Smirk*. Die in Abbildung 5.1 und Abbildung 5.2 dargestellten Effekte beweisen selbstverständlich nicht die Systematik von Smile- oder Smirk-Effekten an den Optionsmärkten, sondern sind zufällig ausgewählte Beispiele. Man hat aber die Erfahrung gemacht, dass solche Verläufe der impliziten Volatilitäten bei vielen Optionen typisch sind. Empirische Studien haben bis jetzt nicht nachweisen können, dass eine Methode systematisch besser als die anderen abschneidet. Es kann deshalb nicht eindeutig gesagt werden, dass die implizite Schätzung der historischen überlegen wäre oder umgekehrt. Ein Vorteil der impliziten Schätzung im Zusammenhang mit Optionsbewertung liegt darin, dass eine mögliche Verzerrung der Schätzung aufgrund der Diskrepanz zwischen Markt- und Black-Scholes Bewertung sich in die Zukunft überträgt. In einem solchen Fall sind die historischen Schätzungen im Nachteil, denn es ist nicht offensichtlich, wie die historische, echte Volatilitätsschätzung adjustiert werden muss, damit die modellverursachte Verzerrung in der Schätzung berücksichtigt wird. Empirisch kann festgestellt werden, dass in sehr liquiden Märkten die implizite Volatilität eine recht gute Schätzung für die zukünftige, tatsächliche Volatilität darstellt. Je einfacher Optionsarbitrage mittels einer Replikation der Option möglich ist, desto wahrscheinlicher ist es, dass Abweichungen der Markt-IV von der richtigen IV gewisse Marktbeteiligte dazu veranlasst,

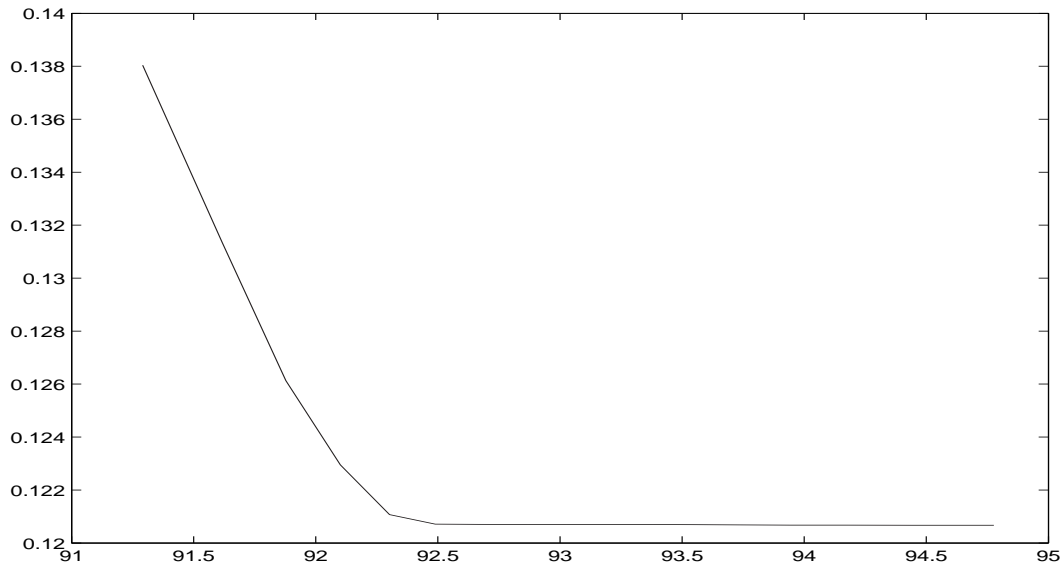


Abbildung 5.2: IV einer Call-Option für unterschiedliche Strikes: Smirk

entsprechende Volatilitäts- Gegenpositionen einzunehmen, um aus der Falschbewertung Profit zu schlagen. In solchen Märkten ist die theoretische Interpretation der Black-Scholes IV, dass die IV die aggregierte Schätzung der zukünftigen Volatilität des Basiswertes darstellt, durchaus gerechtfertigt, allerdings mit der Einschränkung, dass der Markt Optionen anders bewertet als Black und Scholes und so eine Verzerrung der Schätzung entsteht. In weniger liquiden Märkten trifft die Interpretation von IV als aggregierte Marktschätzung aber kaum zu. In diesen Märkten muss davon ausgegangen werden, dass die IV nur ein Maß ist für die gegenwärtige, von Angebot und Nachfrage erzeugte Bewertung der Option. Damit lässt sich zwar abschätzen, wie der Optionspreis auf Veränderung der bewertungsrelevanten Faktoren reagiert, aber eine Volatilitätsschätzung lässt sich daraus nicht ableiten.

## 5.2 Greeks

Die “griechischen Variablen“ (Greeks) bezeichnen die Sensitivitätskennzahlen des Black-Scholes Modells zur Bewertung von Optionen (siehe [40, 41, 8, 37]). Sie zeigen an, wie sich der Optionspreis bei Veränderung einer der Optionspreis bestimmenden Variablen *ceteris paribus* verändert.

Die Sensitivitätskennzahlen sind:

- Das Delta beschreibt die Sensitivität des Optionspreises auf Veränderungen des Kurses des Basiswertes (z. B. Aktien). Es gibt an, wieviele Aktien gebzw. verkauft werden müssen, um die durch die Aktienkursveränderung ausgelöste Wertänderung genau auszugleichen. Delta wird aus dem Quotienten

der Wertänderung der Option und der Wertänderung der Aktie berechnet. Die Kennzahl Delta berechnet sich für eine Call-Option gemäß

$$\Delta(c) = \frac{\partial c}{\partial S} = N(d).$$

Das Delta für eine Put-Option lautet

$$\Delta(p) = \frac{\partial p}{\partial S} = -N(-d).$$

$p$  bezeichnet eine europäische Put-Option,  $c$  eine europäische Call-Option, und  $N$  ist die kumulierte Funktion der Standardnormalverteilung.

- Gamma beschreibt die Sensitivität von Delta auf Aktienkursveränderungen. Gamma wird aus dem Quotienten der Deltaänderung und der Aktienkursänderung berechnet, wobei zwischen Call- und Put-Optionen kein Unterschied besteht. Es kann ebenso als Delta des Deltas betrachtet werden

$$\Gamma = \frac{\partial \Delta}{\partial S} = \frac{\partial^2 c}{\partial S^2}.$$

Sowohl für Call- als auch für Put-Optionen im Black-Scholes Modell berechnet sich Gamma gemäß

$$\Gamma = \frac{N'(d)}{S\sigma\sqrt{T-t}}.$$

$N'(d)$  bezeichnet die Dichtefunktion der Standardnormalverteilung an der Stelle  $d$ . Gamma wird für die Entwicklung von Hedging-Strategien benötigt.

- Omega gibt die Elastizität des Optionspreises bezüglich Aktienkursveränderungen an. Omega ist eine Kennzahl, die das Verhältnis der Wertänderung der Option (in %) und der Wertänderung der Aktie (in %) beschreibt. Daher wird Omega auch als Leverage-Faktor oder Hebel bezeichnet.

$$\Omega = \frac{\partial c/c}{\partial S/S} = \Delta \frac{S}{c}.$$

- Rho beschreibt die Sensitivität des Optionspreises auf eine Änderung des risikolosen Zinssatzes. Es wird durch den Quotienten aus der Optionswertänderung und der Änderung des risikolosen Zinssatzes berechnet

$$\rho = \frac{\partial c}{\partial r}.$$

Für Black-Scholes Call-Optionen ergibt sich Rho als

$$\rho(c) = X\sqrt{T-t}e^{-r(T-t)}N(d - \sigma\sqrt{T-t}),$$

für Put-Optionen

$$\rho(p) = -X\sqrt{T-t}e^{-r(T-t)}N(-d + \sigma\sqrt{T-t}).$$

- Theta beschreibt Sensitivität des Optionspreises auf Veränderungen der Zeit. Es gilt als eine Größe zur Messung des Zeitwertverfalls einer Option. Das Theta einer europäischen Call-Option beträgt

$$\Theta_c = \frac{\partial c}{\partial(T-t)} = -\frac{n(d)S\sigma}{2\sqrt{T-t}} - rXe^{-r(T-t)}N(d-\sigma\sqrt{T-t})$$

und für eine Put-Option ergibt sich

$$\Theta_p = \frac{\partial p}{\partial(T-t)} = -\frac{n(d)S\sigma}{2\sqrt{T-t}} + rXe^{-r(T-t)}N(d-\sigma\sqrt{T-t}).$$



# Kapitel 6

## Numerische Bewertungsverfahren exotischer Optionen

In jüngster Zeit ist ein Trend, weg von standardisierten hin zu individuellen Finanzmarktprodukten zu beobachten. Vor allem im Derivatebereich ist dieser Trend festzustellen. Sogenannte *exotische* Produkte, wie z.B. *Barrier-* oder *Range-Optionen*, spielen eine immer größere Rolle. Diese sind oftmals durch einen hohen Komplexitätsgrad gekennzeichnet. Da für viele dieser Produkte analytische Bewertungsformeln nicht mehr verfügbar sind, ist man auf die Hilfe numerischer Bewertungsverfahren wie Monte-Carlo-Methode angewiesen (siehe [36]).

### 6.1 Exotische Optionen

Die Palette der dem Anleger zur Verfügung stehenden Optionsprodukte ist in jüngerer Zeit durch eine Reihe innovativer Instrumente wesentlich erweitert worden. Neben klassischen abgetrennten Aktien-Optionsscheinen aus Optionsanleihen und selbständigen Optionsscheinen auf unterschiedlichste Basiswerte finden sich zunehmend sogenannte *exotische* Optionsscheine, z.B. in Form von *Digital-, Barrier-* oder *Range-Optionsscheinen* (siehe [8, 37, 25, 30]).

*Exotische* Optionen unterscheiden sich von herkömmlichen (*plain vanilla*) Call- und Put-Optionen durch die Art des Optionsrechts und/oder das Hinzutreten zusätzlicher, den Inhalt des Optionsrechts verändernder Nebenabreden. Im folgenden werden verschiedene Kategorien *exotischer* Optionsscheine bzw. *exotischer* Optionen beispielhaft angeführt.

*Digital-Optionsscheine* verbriefen Optionsrechte, die am Laufzeitende einen im voraus bestimmten festen Betrag auszahlen, wenn der Kurs des Basiswerts den vereinbarten Basispreis über- ) bzw. unterschreitet. Es ist dabei unerheblich, wie weit der Kurs des Basiswerts über bzw. unter dem Basispreis liegt, die Höhe der Auszahlung wird dadurch nicht beeinflusst. Liegt der Kurs des Basiswerts am Fälligkeitstag

unter bzw. über dem Basiskurs des Optionsscheins, verfällt das Optionsrecht wertlos (siehe [8, 37, 25, 30]). *Digital*-Optionsscheine werden in der Regel eingesetzt, wenn eine Prognose über den Kurs des Basiswerts bei Fälligkeit als nicht sehr verlässlich und die Kosten für eine Standardoption als zu hoch eingeschätzt werden (siehe [8, 25]). Gleichzeitig wird das Ertragspotential auf einen festen Betrag begrenzt. Darüber hinaus werden *Digital*-Optionen häufig mit anderen Optionen kombiniert, um spezifische Auszahlungs- und Risikoprofile zu erzeugen (siehe Kombinationen verschiedener Optionen).

*Barrier*-Optionsscheine verbriefen Optionen, die entweder erlöschen (“Knock Out“) oder aber entstehen (“Knock In“), wenn der Basiswert einen im voraus bestimmten Kurs (“Barrier“) erreicht (siehe [8, 37, 25, 30]). *Knock Out*- und *Knock In*-Ausstattungsmerkmale sind Nebenabreden, die jeder Art von Optionsrechten hinzugefügt werden können. Dabei werden unterschiedliche *Barrier*-Optionsscheine angeboten, bei denen nur detaillierte Informationen Aufschluß über das konkrete Chance-Risiko-Profil geben können. So gibt es Call- und Put-Optionen mit *Knock Out*- oder *Knock In*-Barrieren. Im Vergleich zu Standardoptionen ist die Gewinnchance dieser Optionen damit eingeschränkt und das Verlustrisiko erhöht (siehe [8, 37, 25, 30]). Das spiegelt sich im niedrigeren Preis für *Barrier*-Optionen wider. Sie werden häufig dann genutzt, wenn eine sehr präzise Vorstellung über die mögliche Kursentwicklung des Basiswerts besteht.

Bei *Range* (“Bandbreiten“)-Optionsscheinen erhält der Käufer am Ende der Laufzeit einen Betrag, dessen Höhe davon abhängt, ob der Kurs eines oder mehrerer Basiswerte während der Laufzeit der Option über bzw. unter einer oder zwischen mehreren vereinbarten “Grenzwerten“ verläuft. *Range*-Optionen werden eingesetzt, wenn für den Kurs des Basiswerts ein Seitwärtstrend und eine sinkende Volatilität erwartet werden (siehe [8, 37, 25, 30]).

Es gibt verschiedene Typen von *Range*-Optionen, die je nach Ausgestaltung unterschiedliche Chance-Risiko-Profile ausweisen. Beispiele sind:

- Bottom Up/Top Down

Für jeden Tag, an dem der Kurs des Basiswerts über (“Bottom Up“) bzw. unter (“Top Down“) dem festgelegten Grenzwert festgestellt wird, erhält der Anleger einen festen Betrag gutgeschrieben. Die Auszahlung des “angesammelten“ Gesamtbetrags erfolgt am Laufzeitende.

- Single Range

Für jeden Tag, an dem der Kurs des Basiswerts über dem unteren Grenzwert und unter dem oberen Grenzwert festgestellt wird, wird dem Anleger ein fester Betrag gutgeschrieben. Die Auszahlung des “angesammelten“ Gesamtbetrags erfolgt am Laufzeitende.

- Dual Range

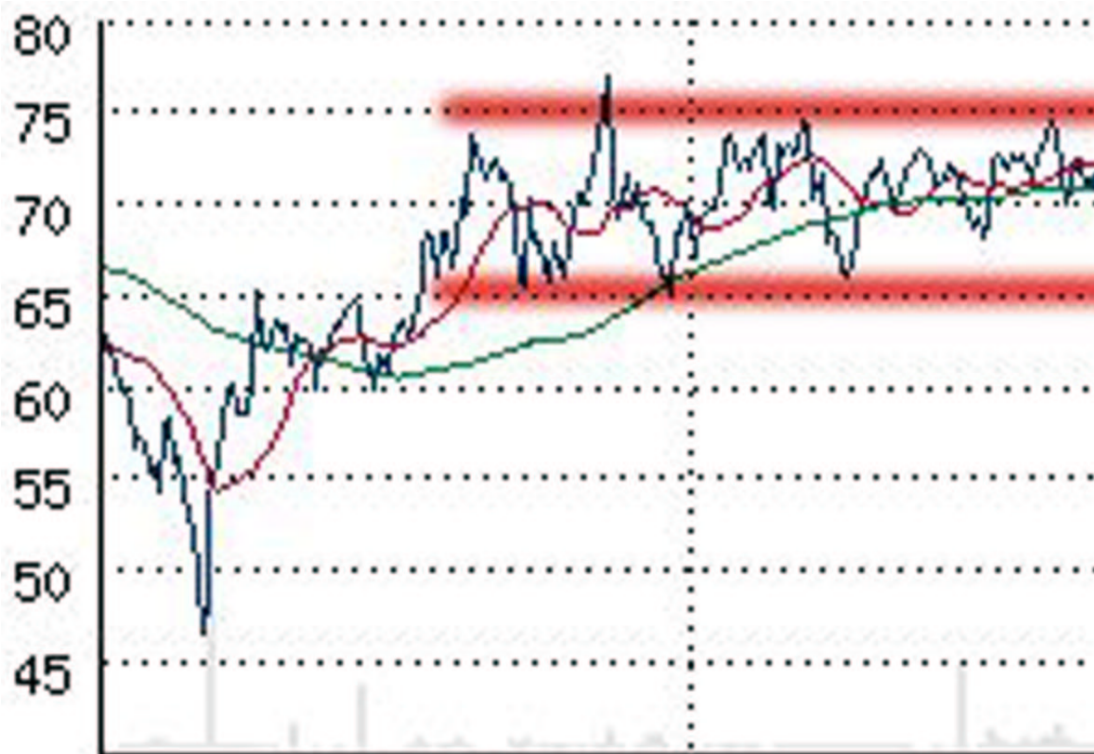


Abbildung 6.1: Beispiel einer Range-Option

Für jeden Tag, an dem der Kurs des Basiswerts über dem unteren Grenzwert und unter dem oberen Grenzwert festgestellt wird, erhält der Anleger einen festen Betrag. Für jeden Tag, an dem der Kurs des Basiswerts außerhalb des Korridors liegt, wird ein entsprechender Betrag abgezogen. Ein positiver Auszahlungswert ergibt sich für den Anleger erst dann, wenn die Anzahl der "Plustage" die Anzahl der "Minustage" übersteigt. Es besteht jedoch keine Nachschußpflicht für den Anleger, wenn bei Laufzeitende die Anzahl der Minustage überwiegt.

- Knock Out-Range (I)

Beim *Knock Out-Range (I)* erhält der Käufer am Ende der Laufzeit einen Betrag, wenn der Kurs des Basiswerts während der gesamten Laufzeit des Optionsscheins zwischen den vereinbarten Grenzwerten verläuft. Das Optionsrecht erlischt, wenn der Kurs des Basiswerts während der Laufzeit einen der Grenzwerte erreicht. Der Anleger erhält in diesem Fall nichts, auch nicht für den Zeitraum, in dem die Grenzwerte eingehalten wurden.

- Knock Out-Range (II)

Für jeden Tag, an dem der Kurs des Basiswerts über dem unteren Grenz-

wert und unter dem oberen Grenzwert festgestellt wird, erhält der Anleger einen festen Betrag. Das Optionsrecht erlischt, wenn der Kurs des Basiswerts einen der Grenzwerte erreicht. Abweichend von der obigen *Knock Out-Range (I)*-Variante erhält der Anleger am Laufzeitende den Gesamtbetrag der bis zum Erreichen des Grenzwerts angesammelten "Tagesbeträge" ausbezahlt.

*Kombinationen von Optionen* werden häufig gewählt, um das Risiko des Totalverlusts zu verringern oder um risikoerhöhend die Ertragsschance zu steigern. Dadurch werden neue Produkte geschaffen, die entweder aus der Kombination von Standardoption und exotischer Option oder aus der Kombination mehrerer exotischer Optionen entstehen. So kann beispielsweise eine Standardoption der Gestalt mit einer *Digital*-Option kombiniert werden, dass unter bestimmten Voraussetzungen am Laufzeitende eine - fest definierte - Auszahlung auch dann erfolgt (siehe [8, 37, 25, 30]), wenn die erwartete Kursentwicklung des Basiswerts nicht eintritt und die "Haupt"-Option wertlos verfällt (Rebate-Struktur I). Ein ähnliches Auszahlungsprofil ergibt sich, wenn ein *Knock Out-Call* mit einem *Digital-Put* kombiniert wird, dessen Basispreis der *Knock Out-Barriere* des *Knock Out-Call* entspricht (Rebate-Struktur II).

Wesentliche Faktoren für die Kursbildung von Optionsscheinen sind die tatsächliche wie auch die von den Marktteilnehmern erwartete zukünftige Kursentwicklung des jeweiligen Basiswerts, die erwartete Häufigkeit und Intensität von Kurschwankungen des Basiswerts (Volatilität) sowie die Laufzeit der verbrieften Option. Der zweite Faktor, die Volatilität, spielt bei der Bewertung exotischer Optionen in der Regel eine größere Rolle als bei den herkömmlichen Optionen (siehe [8, 37, 25, 30]). Sowohl die "Alles oder Nichts-Struktur" digitaler Optionen als auch die Einschränkung des Optionsrechts durch "Nebenabreden" (Knock In- / Knock Out-Barrieren) bedingen eine verstärkte Abhängigkeit des Optionswerts von Veränderungen in der Kursschwankungsbreite und Häufigkeit des Basiswerts.

Eine Kursveränderung des Basiswerts, der dem in einem exotischen Optionsschein verbrieften Optionsrecht zugrunde liegt, kann den Wert des Optionsscheins mindern. Tritt eine Wertminderung ein, so erfolgt sie stets überproportional zur Kursveränderung des Basiswerts (*Hebelwirkung/Leverage-Effekt*), bis hin zur Wertlosigkeit des Optionsscheins. Zu einer Wertminderung kommt es im Fall einer Call-Option generell bei Kursverlusten, im Fall einer Put-Option generell bei Kursgewinnen des Basiswerts. Die Besonderheiten exotischer Optionen können diese Zusammenhänge verstärken. Der Wert einer nach dem "Alles oder Nichts-Prinzip" ausgestalteten *Digital-Option* wird durch Kursschwankungen des Basiswerts um den vereinbarten Basispreis herum stärker beeinflusst als der Wert einer herkömmlichen *Plain Vanilla-Option*. Nähert sich der aktuelle Kurs des Basiswerts einer unterhalb des Basispreises liegenden *Knock Out-Barriere* an, wird dies den Wertverfall eines *Knock Out-Call* drastisch beschleunigen (siehe [8, 37, 25, 30]). Gleiches gilt, wenn sich der Kurs des Basiswerts einer über dem vereinbarten Basispreis liegenden *Knock Out-Barriere* eines *Knock Out-Put* nähert.

Andererseits können die besonderen Ausstattungsmerkmale exotischer Optionen auch abweichende Reaktionen auf Kursveränderungen des Basiswerts hervorrufen. So gewinnt eine Call-Option zwar bei einem Kursanstieg des Basiswerts grundsätzlich an Wert. Ist der fragliche Call aber zugleich mit einer über dem vereinbarten Basispreis liegenden *Knock Out-Barriere* versehen, wird ein weiterer Kursanstieg des Basiswerts in Richtung *Knock Out-Level* den Wert des Optionsscheins beeinträchtigen, bei Annäherung an die Barriere zu einer drastischen Wertminderung und bei Erreichen des *Knock Out-Niveaus* zum Totalverlust des Einsatzes führen. Eine vergleichbare Konstellation ergibt sich bei einer Put-Option, die mit einer oberhalb des Basispreises liegenden *Knock In-Barriere* ausgestattet ist.

Besonderheiten gelten für *Range-Optionsscheine*. Der Wert einer *Range-Option* wird maßgeblich davon beeinflusst, ob der Kurs des Basiswerts innerhalb der vereinbarten Bandbreite notiert (siehe [8, 37, 25, 30]). Theoretisch ist der Preis einer *Range-Option* am größten, wenn sich der Kurs des Basiswerts in der Mitte der Bandbreite befindet. Kursbewegungen des Basiswerts in Richtung "Bandbreiten-Mitte" resultieren dementsprechend tendenziell in einer Wertsteigerung des Optionsrechts. Kursbewegungen des Basiswerts von der Mitte der jeweiligen Bandbreite in die Richtung eines Randes führen demgegenüber regelmäßig zu einer Wertminderung des Optionsrechts. Sind die Ränder der *Range-Option* zugleich als *Knock Out-Barrieren* ausgestattet, verstärkt dies den bei Annäherung an eine der Barrieren eintretenden Wertverfall erheblich. Dabei hängt der Umfang der Wertminderung wiederum von der konkreten Ausgestaltung des Optionsrechts ab. Sie ist am deutlichsten im Fall einer *Double Knock Out Range-Option*, bei der das Erreichen einer der beiden Barrieren zum vollständigen Erlöschen des Optionsrechts, d.h. zu einem Totalverlust des gesamten Optionseinsatzes führt. Sie kann tendenziell geringer sein, wenn die *Knock Out-Option* so konzipiert ist, dass beim Touchieren der Barrieren das Optionsrecht zwar erlischt, die über die Laufzeit bis zu diesem Zeitpunkt "angesammelten" Beträge jedoch ausbezahlt werden (siehe [8, 37, 25, 30]).

## 6.2 Monte-Carlo-Methode

Unter Simulation versteht man allgemein ein Verfahren zur Durchführung von Experimenten mit dem Ziel, Aussagen über das Verhalten eines realen Systems zu gewinnen. Die mittels Simulationen gewonnenen Erkenntnisse sollen Rückschlüsse auf reale Zusammenhänge liefern, wozu das Modellexperiment dem gegebenen System hinreichend ähnlich sein muß. Die Experimente werden zumeist auf einem Digitalrechner unter Benutzung mathematischer Modelle implementiert. Sind gewisse Einflußfaktoren in der betrachteten Problemstellung nicht deterministisch, so enthält die zu simulierende Größe Unsicherheitskomponenten. Zumindest eine Modellvariable wird dabei durch einen Zufallsprozess erzeugt, womit das gesamte System einem stochastischen Einfluß unterliegt; man spricht in diesem Fall von *zufallsbedingten Simulationen*. Synonym werden auch die Begriffe *Monte-Carlo Simulationen* bzw.

*Monte-Carlo-Methoden* benutzt.

Mit *Monte-Carlo-Methoden* lassen sich unterschiedliche Probleme unter Verwendung von Zufallszahlen lösen (siehe [29, 35]). Mögliche Anwendungsbereiche sind die Lösung von Differential- und Integralgleichungen bzw. bestimmter Integrale mit Hilfe statistischer Experimente sowie die Erzeugung von Stichproben von gegebenen stochastischen Prozessen.

Aufgrund der auftretenden Zufallsvariablen handelt sich bei der Simulation eines Kurspfades im Rahmen einer Optionsbewertungsaufgabe um eine zufallsbedingte Simulation (siehe [29]). Ausgangsbasis der Kurssimulation ist die Darstellung der Kursdynamik durch eine aus der postulierten Renditeverteilung abgeleitete stochastische Differentialgleichung, die in einer Rekursionsformel resultiert, mit der sukzessive ein den stochastischen Annahmen genügender Kursverlauf simuliert werden kann. Die benötigten zufälligen Größen erzeugt man mit Hilfe von Zufallsgeneratoren und einschlägigen Algorithmen.

Simulationen und damit auch durch Simulationen geschätzte Optionspreise unterliegen den zufälligen Einflüssen durch entsprechende Zufallsvariable des Simulationsprozesses. Die Ergebnisse einzelner Schätzungen weichen also, eventuell deutlich, voneinander ab. "Stabile Lösungen" unter Angabe einer Konfidenzgröße versucht beispielsweise durch das Mitteln über eine größere Anzahl simulierte Preise zu erhalten (siehe [29, 35]). Verfeinerte, sogenannte varianzreduzierende Monte-Carlo Verfahren können die Stabilität der erhaltenen Lösungen unter Umständen weiter verbessern.

Der Programmaufbau von Simulationen ist bei unterschiedlichen Preismodellen sehr ähnlich. Postuliert man eine andere Kursstochastik, so sind in einem einmal implementierten Simulationsprogramm zuweilen lediglich die Algorithmen zur Erzeugung von Zufallszahlen bzw. die Rekursionsformeln anzupassen. Somit kann die Simulationstechnik als äußerst flexibel bezeichnet werden. Neben diesem Merkmal liegt einer der Hauptvorteile von Simulationen darin, dass der Rechenaufwand bei zunehmender Komplexität der Modelle langsamer ansteigt als dies bei anderen Verfahren der Fall ist. Andererseits ist auch zu beachten, dass der Aufwand bei geringerer Komplexität der Modelle bereits so groß ist, dass sich eine Simulation nur ab einer gewissen Modellgröße empfiehlt.

# Kapitel 7

## Numerische Optimierungsverfahren

Die Lösung vieler technischer Problemstellungen erfordert die Minimierung oder Maximierung mehrdimensionaler Zielfunktionen. Anschaulich betrachtet ist man in der Situation eines Wanderers, der das tiefste Tal oder die höchste Erhebung finden muß. Oftmals treten neben das eigentliche Minimierungsproblem zusätzliche Gleichungs- oder Ungleichungsnebenbedingungen, die eine Einschränkung der Lösungsmenge von Anfang an bedeuten. Man spricht dann von Optimierung unter Nebenbedingungen oder *constrained optimization*. Ein weit verbreitetes Verfahren zur angenäherten Lösung überbestimmter Gleichungssysteme ist die Methode der kleinsten Quadrate (*least squares*). Dieses Verfahren wird häufig zur Kurveninterpolation (*curve fitting*) eingesetzt, ist aber nicht auf diese Anwendung beschränkt. Auch hier existieren Varianten mit Nebenbedingungen sowie lineare und nichtlineare *least squares* Verfahren. Eine Unterproblemstellung bei der Optimumsuche ist häufig die Nullstellenbestimmung von nichtlinearen Gleichungssystemen. Methoden der unrestringierten nichtlinearen Optimierung sind eng verwandt mit Methoden zur Lösung nichtlinearer Gleichungssysteme (siehe [20, 26, 38]) .

**Definition 7.0.1** Eine Optimierungsaufgabe liegt vor, wenn

- Raum  $H$
- Restriktionsbereich  $S \subseteq H$
- Zielfunktion  $f : S \rightarrow \mathbb{R}$

gegeben sind und

- eine Zahl (Optimalwert)  $f^* \in \mathbb{R}$ , sowie
- ein Element  $x^* \in S$  (optimale Lösung),

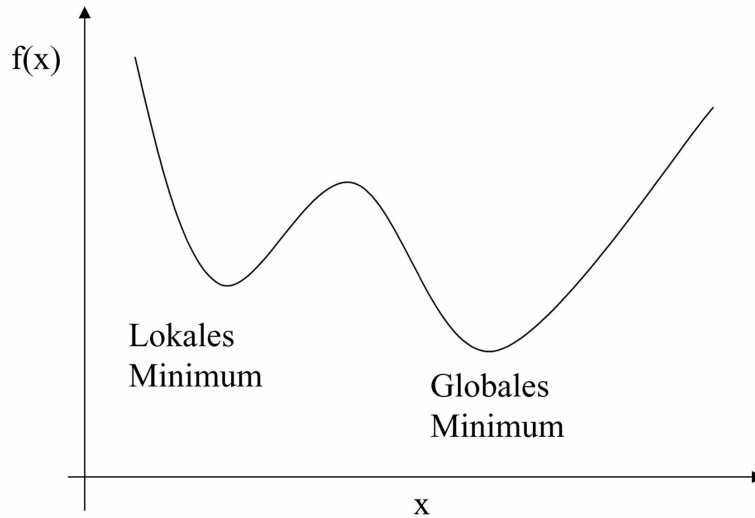


Abbildung 7.1: Nichtlineare Zielfunktion mit einem Parameter

gesucht sind so dass

$$f^* = f(x^*) \leq f(x), \quad \forall x \in S.$$

Häufig erweist sich, dass das Aufsuchen der optimalen Lösung auf Schwierigkeiten stößt und man nur in der Lage ist, Punkte zu berechnen, unter denen die gesuchte optimale Lösung sein kann. Das führt zu dem Begriff der lokal-optimalen Lösung.

**Definition 7.0.2** Gegeben sei die Optimierungsaufgabe  $f^* = \inf\{f(x) : x \in S\}$ . Der Punkt  $x^* \in S$  heißt global-optimale Lösung von der Optimierungsaufgabe, wenn  $f(x^*) \leq f(x), \forall x \in S$ .

Der Punkt  $x^* \in S$  heißt lokal-optimale Lösung von der Optimierungsaufgabe, wenn  $\exists \epsilon > 0 : f(x^*) \leq f(x), \forall x \in S \cap U_\epsilon(x^*)$ .

Wenn im Zusammenhang mit einer Zielfunktion von einem “Punkt“ bzw. einem “Lösungspunkt“ gesprochen wird, so ist darunter ein mit Werten innerhalb des Definitionsbereiches belegter Parametervektor  $x^* \in \mathbb{R}^m$  zu verstehen. Die Zielfunktion bildet die Domäne, d.h. den Definitionsbereich, in eine eindimensionale Ko-Domäne (Funktionswerte) ab, so dass  $\mathbb{R}^m \rightarrow \mathbb{R}^1$  gilt. Da die Dimension des Definitionsbereiches der Qualitätsfunktion durch die Anzahl der unbekannt Parameter  $m$  festgelegt ist, liegt jeder Punkt der Zielfunktion auf einer  $m$ -dimensionalen Hyperfläche in



einem  $(m + 1)$ -dimensionalen Raum. Leicht vorstellbar ist lediglich der eindimensionale Fall (Abb. 7.1). Für jeden Punkt, d.h. für jeden Vektor mit Unbekannten, lässt sich ein Funktionswert der Zielfunktion berechnen. Da der Funktionswert ein wichtiges Kriterium für die Güte eines Punktes ist, wird dieser Wert oft die Qualität oder Fitness bezeichnet. Ist der Funktionswert eines Punktes im Vergleich mit Punkten aus dessen lokaler Umgebung größer (bzw. kleiner) so spricht man von einem lokalen Optimum. Sind alle Funktionswerte des Definitionsbereiches ohne Ausnahme weniger optimal, so nennt man diesen Ort globales Optimum.

Im Hinblick auf die globale Optimierung ist die wichtigste Eigenschaft einer Zielfunktion die Anzahl ihrer lokalen Optima. Wenn gezeigt werden kann, dass ein Problem nur ein lokales Extremum besitzt, d.h. wenn das Problem unimodal ist, dann liefert bereits die lokale Optimierung das globale Optimum (siehe [20]).

## 7.1 Lokale Optimierung

Die lokale Optimierung hat eine lange Tradition. Bereits vor ca. 300 Jahren lieferte die Differentialrechnung ein notwendiges Kriterium für lokale Optima: der Gradient  $\nabla F(x)$  muß gleich null sein. Diese Erkenntnis ermöglichte die Konstruktion von effizienten und im mathematischen Sinne “optimalen“ Optimierungsverfahren, die in “wenigen“ Schritten konvergieren. Dadurch erklärt sich auch deren weite Verbreitung. Voraussetzung ist allerdings das Vorhandensein eines differenzierbaren funktionalen Modells, denn die lokalen Verfahren benötigen Informationen über die internen, differentialgeometrischen Eigenschaften der zu minimierenden Funktion. Einschränkend ist anzumerken, dass das Endergebnis eines lokalen Iterationsverfahrens lediglich ein stationärer Punkt ist und daher nur unimodale Probleme gelöst werden können. Die Entwicklung dieser Verfahren, die mit Namen wie Newton, Euler, Lagrange, oder MacLaurin verknüpft ist, kann aber noch nicht als abgeschlossen gelten, was Publikationen der neunziger Jahre zeigen.

Ausgehend von einem Startwert  $x_0$ , erzeugen die lokalen Optimierungsverfahren eine Punktfolge  $x_1, x_2, x_3, \dots, x_k$ , die gegen ein lokales Minimum von  $f(x)$  konvergiert (siehe [20, 26, 38]). Man spricht auch von lokalen Abstiegsverfahren, weil im Idealfall jeder Punkt der Folge einen kleineren Funktionswert als sein Vorgänger aufweist und somit die Punktfolge in eine Senke des Qualitätsgebirges herabsteigt.

Eine Klassifikation der verschiedenen, lokalen Iterationsverfahren kann durch das Verfahren

$$x_{k+1} = x_k + t_k d_k, \quad \text{mit } k = 0, 1, 2, \dots$$

vorgenommen werden. Die einzelnen lokalen Verfahren unterscheiden sich lediglich in der Wahl des Richtungsvektors  $d_k$  und einer positiven Schrittweite  $t_k$ .

Für ein lokales Abstiegsverfahren müssen die aufeinanderfolgenden Funktions-

werte streng monoton abnehmen. Es muß also gelten (siehe [20])

$$f(x_{k+1}) = f(x_k) + t_k d_k < f(x_k).$$

Entwickelt man  $f(x_{k+1} + t_k d_k)$  in eine Taylorreihe und vernachlässigt die Glieder höherer Ordnung, so erhält man  $f(x_k + t_k d_k) = f(x_k) + t_k \nabla f(x_k) d_k$ . Durch Ersetzung in obige Gleichung ergibt sich für alle Abstiegsrichtungen  $d_k$

$$\nabla f(x_k) d_k < 0.$$

Richtungsvektoren, die die Gleichung erfüllen, führen zu kleineren Funktionswerten in der Iterationsfolge. Die Abstiegsrichtungen  $d_k$  können auch in der allgemeinen Form

$$d_k = -Q(x_k) \nabla f(x_k)$$

geschrieben werden, wobei  $Q(x_k)$  eine beliebige, positiv definite Matrix sei. Durch Einsetzen von  $d_k$  erhält man die Iterationsformel für die gesamte Klasse der Gradientenmethode in Form von (siehe [20, 26])

$$x_{k+1} = x_k - t_k Q(x_k) \nabla f(x_k),$$

wobei sich die Verfahren lediglich in der Wahl der positiven Schrittweite  $t_k$  und der positiv definiten Matrix  $Q(x_k)$  unterscheiden. Die Konvergenz dieser Verfahren ist sowohl von dem funktionalen Zusammenhang als auch von Näherungswerten abhängig und sie kann nicht garantiert werden, da in der Regel nicht vorausgesagt werden kann, ob die Startwerte eine ausreichende Approximation der Lösung darstellen.

Prinzipiell lassen sich beliebig viele Algorithmen entwickeln, die letztendlich eine Gemeinsamkeit haben, mittels Funktionsableitungen das funktionale Modell durch ein einfaches zu ersetzen. Im folgenden werden die meist in der Praxis benutzten Verfahren vorgestellt.

Das Newton-Raphson-Verfahren, auch häufig als Newton-Verfahren bezeichnet, diente ursprünglich als Algorithmus zur Nullstellensuche von Funktionen (siehe [20, 26, 38]). Später erkannte man, dass sich auch Optimierungsprobleme lösen lassen, wenn stationäre Punkte der Zielfunktion gesucht werden. Es müssen also die Nullstellen der ersten Ableitung bestimmt werden, indem das Gleichungssystem  $f(x) = 0$  gelöst wird. Da ein nichtlineares Gleichungssystem oft nicht geschlossen lösbar ist, werden die Gleichungen durch eine Taylorreihe ersten Ordnung approximiert, d.h. sie werden linearisiert. Es entsteht ein lineares Gleichungssystem, das sich geschlossen lösen läßt. Hier wird aber gleichzeitig ein entscheidender Nachteil des Verfahrens deutlich, denn die Frage, inwiefern dieses lineare Gleichungssystem noch mit dem eigentlichen Optimierungsproblem zu tun hat, läßt sich nicht leicht beantworten. Nur noch für hinreichend gute Startwerte konvergiert das Verfahren, andernfalls stellt sich entweder Divergenz oder ein Alternieren ein. Die Grundzüge des Verfahrens werden durch folgende Betrachtungsweise deutlich. Der Zugang zum

Newton-Raphson-Verfahren erfolgt direkt mittels Betrachtung der Zielfunktion. Die ursprüngliche Zielfunktion  $f(x)$  wird durch eine Taylorreihe mit Gliedern der zweiten Ordnung ersetzt

$$\hat{f}(x) := f(x_k) + \frac{\partial f}{\partial x}(x_k)(\Delta x) + \frac{1}{2} \frac{\partial^2 f}{\partial^2 x}(x_k)(\Delta x)^2.$$

Die Ersatzfunktion  $\hat{f}$  ist leichter zu handhaben, aber es darf nicht vergessen werden, dass die Nebeneffekte, die durch diese Ersetzung entstehen, nur schwer abzuschätzen sind. Bildet man mit Hilfe der Ableitungen von  $\hat{f}$  nach den Variablen  $x$  das Gleichungssystem der Ersatzfunktion, so ergibt sich

$$\nabla \hat{f} = \nabla f(x_k) + \frac{\partial^2 f}{\partial x_i \partial x_i}(x_k)(\Delta x) = 0.$$

Berücksichtigt man  $\Delta x = x_{k+1} - x_k$  und löst man die Gleichung nach  $x_{k+1}$  auf, so ergibt sich für das Newton-Raphson-Verfahren (siehe [20]) die Iterationsvorschrift

$$x_{k+1} = x_k - \nabla f(x_k) \left[ \frac{\partial^2 f}{\partial^2 x}(x_k) \right]^{-1}.$$

Gemäß der Einteilung für lokale Iterationsverfahren ergibt sich eine Schrittweite  $t_k = 1$  und eine Abstiegsrichtung  $Q(x_k)$  als Inverse der Matrix der zweiten Ableitungen  $Q(x_k) = [\partial_{xx}^2 f(x)]^{-1}$ . Durch diese  $(m \times m)$ -Matrix mit zweifachen Funktionsableitungen werden Abstiegsrichtung und Schrittweite gemeinsam festgelegt. Die Zielfunktion wird bei jedem Iterationsschritt quadratisch approximiert. Als Konsequenz konvergiert das Newton-Verfahren schnell, bei quadratischen Funktionen sogar in einem Schritt, man spricht hier von quadratischer Konvergenz.

Um den Unzulänglichkeiten des Newton-Raphson-Verfahrens entgegenzuwirken, wurde die Modifizierung von Levenberg (1944) durchgeführt. Für diesen Algorithmus, der auch unter dem Namen *Trust-Region-Method* bekannt ist (siehe [20]), wählt man für  $t_k Q(x_k)$  statt der Matrix  $[\partial_{xx}^2 f(x_k)]^{-1}$  eine erweiterte Formel der Form

$$t_k Q(x_k) = [\partial_{xx}^2 f(x_k)]^{-1} + \alpha_k M.$$

Dabei ist  $\alpha_k$  ein positiver Skalar und  $M$  eine positiv definite Matrix. Es ergibt sich die Iterationsformel für das modifizierte Newton-Raphson-Verfahren

$$x_{k+1} = x_k - \nabla(x_k) \left[ \frac{\partial^2 f}{\partial^2 x}(x_k) + \alpha_k M \right]^{-1}.$$

Da  $M$  eine positiv definite Matrix ist, kann immer ein hinreichend großes  $\alpha_k$  gefunden werden, so dass  $t_k Q(x_k)$  positiv definit ist. Mit der positiv definiten  $Q$ -Matrix ist immer eine Abstiegsrichtung garantiert. Dadurch kann ein Mangel des Newton-Raphson-Verfahrens, nämlich der einer eventuell nicht positiv definiten Matrix  $Q$ , behoben werden. Andererseits zeigt sich, dass das Verfahren ein Kompromiß

zwischen der klassischen Gradientenmethode und dem Newton-Raphson-Verfahren darstellt. Wählt man für  $M$  die Einheitsmatrix  $I$  und für  $\alpha_k = 0$ , so erhält man die Formel des Newton-Raphson-Verfahrens. Für  $M = I$  und sehr große  $\alpha_k$  geht das Verfahren in die klassische Gradientenmethode über. Die möglichen Suchrichtungen liegen also zwischen der des Newton- und der des Gradientenverfahrens. Die Variable  $\alpha_k$  übernimmt dabei die Rolle der Schrittweite  $t_k$ .

Es bleibt aber die Frage offen, welchen Wert  $\alpha_k$  annehmen soll. Da es nicht möglich ist, für alle Optimierungsprobleme eine allgemeine Regel zu formulieren, bleibt nichts anderes übrig, als einen selbst gewählten Wert für  $\alpha_k$  dahingehend zu testen, ob eine Abstiegsrichtung vorliegt oder nicht.

Das Gauß-Newton-Verfahren ist das mit Abstand am häufigsten eingesetzte Verfahren zur Lösung von Optimierungsproblemen. Im Gegensatz zu dem Newton-Raphson-Verfahren wird hier eine Linearisierung nicht außerhalb der Norm (auf der Ebene Normalgleichungen oder Zielfunktionen), sondern innerhalb einer Norm vorgenommen, d.h. die Funktionen der einzelnen Residuen werden an einem Taylorpunkt entwickelt und dann erst einer Norm unterworfen (siehe [20]).

Ausgehend von der Minimierung der Residuenquadrate

$$|v| = \sqrt{\sum_{i=1}^n v_i^2} \rightarrow \min$$

können die Residuen  $v_i$  allgemein als vektorwertige Funktion  $V(x)$  aufgefaßt werden. Es gelte (siehe [20])

$$V(x) = Y - F(x) = y_i - f_i(x) \quad i = 1, \dots, n$$

wobei der Unbekanntenvektor der Form  $x \in \mathbb{R}^m$  und die Dimension des Beobachtungsvektors  $Y \in \mathbb{R}^n$  sei, so dass  $F : \mathbb{R}^m \rightarrow \mathbb{R}^n$  gilt. Werden die Residuen  $V_i(x)$  durch die Funktion  $V(x) = (V_1(x), V_2(x), \dots, V_n(x))^T$  ausgedrückt, so erhält man

$$\frac{1}{2} \sum_{i=1}^n v_i(x)^2 = \frac{1}{2} \|V(x)\|^2 \quad \forall x \in \mathbb{R}^m.$$

Das Gauß-Newton-Verfahren basiert darauf, die Funktionen  $V$  in eine Taylorreihe nur bis zu den Gliedern der erste Ordnung zu entwickeln. Nach dieser häufig als Linearisierung bezeichneten Approximation entsteht dann die Ersatzfunktion

$$V^*(x) := V(x_k) + V'(x_k)(x - x_k), \quad \text{wobei} \quad V'(x_k) = \frac{\partial V(x)}{\partial x}(x_k).$$

Der Entwicklungspunkt muss dabei eine extern bereitgestellte Näherungslösung  $x_k$  sein. Von nun an wird für den folgenden Iterationsschritt  $k + 1$ , statt der ursprünglichen Funktion  $V$ , die lineare Ersatzfunktion  $V^*$  verwendet. Wird die Ersatzfunktion

in die Norm eingesetzt, entsteht das neue Minimierungsproblem

$$Q^*(x) = \frac{1}{2} \sum_{i=1}^n [V(x_k) + V'(x_k)(x - x_k)]^2 \rightarrow \min.$$

Lediglich Vektor  $x$  beinhaltet die unbekannt Parameter, denn  $x_k$ ,  $V(x_k)$  und  $V'(x_k)$  sind durch die Näherung gegeben, liegen also schon numerisch fest. Nun wird deutlich, dass es sich bei  $Q^*$  um ein lineares Problem handelt. Gelöst wird die unimodale, konvexe Minimierungsaufgabe mit Hilfe des Kriteriums für stationäre Punkte, nämlich  $\nabla Q^* = 0$ . Nach Ableitung und Nullsetzung der obige Gleichung erhält man die Gaußschen Normalgleichungen

$$[V(x_k) + V'(x_k)(x - x_k)]V'(x_k) = 0.$$

Löst man die Gleichung unter Berücksichtigung von  $x = x_{k+1}$  nach den Unbekannten  $x_{k+1}$  auf, so ergibt sich die Iterations-formel für das Gauß-Newton-Verfahren (siehe [20, 26])

$$x_{k+1} = x_k - [V'(x_k)^T V'(x_k)]^{-1} V'(x_k)^T V(x_k).$$

Wird das Verfahren als iterative Gradientenmethode interpretiert, so ergibt sich für die Schrittweite  $t_k$  der Wert 1 und für die Matrix der Abstiegsrichtungen  $Q(x_k)$  der Ausdruck  $[V'(x_k)^T V'(x_k)]^{-1}$ . Der Gradient  $\nabla f(x_k)$  entspricht dem Gradienten  $\nabla Q$  der Zielfunktion  $Q$  und wird durch den Ausdruck  $V'(x_k)^T V(x_k)$  beschrieben.

## 7.2 Globale Optimierung

Besitzt eine Zielfunktion eine unbekannt Anzahl lokaler Optima, so handelt es sich um ein echtes globales Optimierungsproblem. Die bisher behandelten Strategien befassten sich lediglich mit der lokalen Optimierung. Globale, multimodale Probleme können nicht allein durch lokale Optimierung gelöst werden. In der Praxis erfordern Probleme dieser Klasse ebenfalls eine akzeptable Lösung in angemessener Rechenzeit. An einem metaphorischen Beispiel, wie im [33] beschrieben, kann dieser Zusammenhang illustriert werden: In einem zerklüfteten Gebirge mit vielen Mulden und Senken soll der tiefste Punkt gefunden werden. Die lokale Optimierung sucht sich nun einen Ort aus, für den vermutet wird, dass er nicht weit von dem tiefsten Punkt entfernt liegt und schüttet dort Wasser aus. Die Stelle, an der sich das Wasser nach dem Herabfließen wieder sammelt, erklärt die lokale Optimierung für den tiefsten Punkt. Was aber, wenn sich das Wasser in einem kleinen, hoch gelegenen Gletschersee sammelt? Was, wenn zwei Punkte an verschiedenen Rändern des Gebietes mehr oder weniger beide den tiefsten Punkt repräsentieren? In einem solchen Fall muß auf global optimierende Verfahren zurückgegriffen werden. Deren wichtigster Unterschied besteht in der Generierung vieler Startpunkte. In der Gebirgslandschaft

würde es regnen. Wäre der Regen stark genug, dann fiel irgendwann ein Tropfen auf den tiefsten Punkt des gesamten Gebietes.

Für die exakte Lösung des globalen Optimierungsproblems gibt es zunächst kein Kriterium, mit dem das globale Minimum aus der offenen Menge lokaler Optima herausgefiltert werden kann. Auch können keine Eigenschaften des globalen Optimums genannt werden, die zu einem effizienten Verfahren zur Gewinnung der exakten Lösung führen. Das multimodale Optimierungsproblem ist daher im allgemeinen unlösbar, denn es können immer Funktionen konstruiert werden, deren lokale Optima beliebig weit auseinander liegen.

Es stellt sich nunmehr die Frage, nach welchem Verfahren der Lösungsraum durchsucht werden soll, d.h. wie die Regentropfen verteilt werden sollen. Zunächst einmal liegt es nahe, den gesamten Lösungsraum systematisch zu durchsuchen. Jedoch steigt die Mächtigkeit des Lösungsraumes mit zunehmender Anzahl der Dimensionen exponentiell an. Bis auf einige "kleinere" Problemstellungen wird die systematische Durchforschung des Lösungsraumes in angemessener Zeit nicht möglich sein. Sicherlich könnte der Lösungsraum mittels Rasterung diskretisiert werden, aber welche Maschenweite ausreichend ist, bleibt offen. In der Komplexitäts- und Algorithmentheorie werden solche Probleme als NP-schwere Probleme bezeichnet. NP-schwere Probleme sind dadurch gekennzeichnet, dass eine Lösung mit exponentiellem Aufwand herbeigeführt werden kann, im Gegensatz zu NP-effizienten Problemen, deren Aufwand mit der Problemgröße polynomial ansteigt. Dieser Definition zufolge müssen globale Optimierungsprobleme als effektiv nicht lösbar angesehen werden. Dennoch besteht die Möglichkeit, heuristische Verfahren auf diese Problemklasse anzusetzen.

Ein Verfahren wird vor allem durch einen Konstruktionsoperator und einen Ersetzungsoperator geprägt. Während der Konstruktionsoperator die Auswahl der zu berechnenden Punkte trifft, entscheidet der Ersetzungsoperator darüber, welche Punkte verworfen und welche gespeichert werden. Die Entscheidung darüber, welche Punkte konstruiert und welche ersetzt werden, wird in Abhängigkeit von vorausgegangenen Auswertungen der Zielfunktion erfolgen. Ein Verfahren muß sich also die Ergebnisse vorangegangener Berechnungen merken. Damit entsteht die Liste

$$[x, f]_H = \{(x_1, f_1), (x_2, f_2), \dots, (x_H, f_H)\} \in \mathbb{R}^m \times \mathbb{R}$$

mit insgesamt  $H$  gespeicherten Elementen bestehend aus Punkten  $x$  und Qualitäten  $f(x)$ . Die Größe  $H$  ist die Anzahl der gespeicherten Punkte.  $H$  kann zwischen 0 und der Anzahl bisheriger Funktionsauswertungen liegen. Nach jeder Iteration  $i$  eines Verfahrens wird der nächste auszuwertende Punkt  $x_i$  in Abhängigkeit der Liste gespeicherter Punkte  $[x, f]_H$  mit Hilfe des Konstruktionsoperators

$$\mathcal{K}_i : (\mathbb{R}^m \times \mathbb{R})_H \rightarrow \mathbb{R}^m$$

ausgewählt. Danach erfolgt die Bewertung des durch  $\mathcal{K}_i$  bestimmten Punktes  $x_i$ , indem  $x_i$  in die Zielfunktion eingesetzt und dessen Funktionswert  $f(x_i)$  berechnet

wird. Anschließend muß entschieden werden, ob der neue Punkt in die Liste der zu speichernden Punkte aufgenommen werden soll. Dies erfolgt mittels des Ersetzungsoperators  $\Psi$  der Form

$$\Psi_i : (\mathbb{R}^m \times \mathbb{R})_{H+1} \rightarrow (\mathbb{R}^m \times \mathbb{R})_H.$$

Der Index  $H$  verdeutlicht, dass  $\epsilon$  von vorangegangenen Auswertungen der Zielfunktion abhängt.

Bei nicht adaptiven Verfahren wird die Auswahl der Auswertepunkte unabhängig von den Erfolgen oder Misserfolgen vorangegangener Auswertungen getroffen. Es gilt  $H = 0$ , d.h. gespeicherte Punkte haben keinen Einfluss auf den Optimierungsvorgang.

Verfahren mit  $H = 1$  werden als Markov-Verfahren bezeichnet. Da  $H$  die Anzahl der gespeicherten Punkte ist, können sich die Markov-Verfahren nur eine Auswertung  $[x^*, f^*]$  der Zielfunktion merken. Die Konstruktion neuer Punkte erfolgt durch einen Konstruktionsoperator  $\mathcal{K}([x^*, f^*])$ , der von dem Punkt  $P^*[u^*, f^*]$  abhängt. Neue Punkte werden in der Regel so konstruiert, dass sie in der Nähe des gespeicherten Punktes liegen. Eine Möglichkeit besteht darin, gleichverteilte Zufallszahlen in einem abgeschlossenen Intervall um den gespeicherten Punkt zu bilden. Es ergibt sich ein Konstruktionsoperator der Form

$$\mathcal{K} = \text{RAND}[x^* - S_i, x^* + S_i].$$

$S$  dient als Schranke für die Unbekannten und ist ein Vektor mit positiven reellen Zahlen. Diese müssen keine Konstanten sein, sondern können beispielsweise von der Anzahl durchlaufener Iterationsschritte abhängen. Sowohl bei einer Gleichverteilung als auch bei einer Normalverteilung läßt sich durch den Parameter  $S$  zwischen globaler und lokaler Suchstrategie umschalten. Je größer  $S$ , desto größer ist der Suchraum um den Punkt  $x^*$  und desto höher ist die Gewichtung der globalen Suche. Für  $S \rightarrow 0$  entartet das Verfahren zu einem lokalen Optimierungsalgorithmus.

Das Simulated Annealing wird im deutschen Sprachgebrauch als *Simuliertes Ausglühen* bezeichnet und ist das bekannteste Markov-Verfahren. Der Begriff steht für ein aus der Natur abgeschautes Verfahren (siehe [3, 39]). Beim langsamen Abkühlen eines Metallstückes ordnen sich die einzelnen Atome so an, dass sie einen Zustand möglichst niedriger Energie einnehmen. Tritt allerdings eine zu schnelle Abkühlung auf, so haben die Atome nicht die Zeit das tatsächliche Minimum zu finden, das System bleibt in einem lokalen Minimum "hängen". Ein niedriger Energiezustand ist in diesem Fall gleichbedeutend mit einem stabilen Endzustand, also einem sehr stabilen Werkstück. Der Ablauf der Grundvariante ist Tabelle 7.1 zu entnehmen.

Eine im Intervall  $[0, 1]$  erzeugte Zufallszahl wird mit dem Ergebnis einer Exponentialfunktion verglichen, die wiederum von der Differenz der Funktionswerte zwischen dem neuen und dem alten Punkt abhängig ist. Ist die Güte des neuen Punktes besser als der gespeicherte Funktionswert, d.h. es gilt  $f_i < f^*$ , so ergibt sich ein positiver Exponent, da  $T \geq 0$  gilt. Die Wahrscheinlichkeit, dass der neue

Initialisierung	$i = 1, x^* = \text{RAND}[x_{\min}, x_{\max}], f^* = f(x^*)$
Konstruktion	$x_i = \mathcal{K}(x^*, S)$
Bewertung	$f_i = f(x_i)$
Ersetzung	$\Psi = \begin{cases} x^* = x_j & \text{falls } \text{RAND}[0, 1] < e^{-\frac{f_i - f^*}{T}} \\ x^* & \text{sonst} \end{cases}$
Abbruchkriterium	wenn Bedingung erfüllt zu <i>Ausgabe</i> sonst $i := i + 1$ , vermindere $(S, T)$ , Rücksprung zu <i>Konstruktion</i>
Ausgabe	$x^*$

Tabelle 7.1: Simulated Annealing Verfahren

Punkt übernommen wird, beträgt also 1, weil ein positiver Exponent in jedem Fall einen Wert größer 1 liefert und die Ungleichung in Tabelle 7.1 in jedem Fall erfüllt. Schlechtere Punkte  $P_i$  werden allerdings nur mit einer Wahrscheinlichkeit akzeptiert, die zwischen 0 und 1 liegt. Über den Steuerparameter  $T$  kann die Akzeptanzwahrscheinlichkeit reguliert werden.

Der Verlauf des Verfahrens ist durch das kontinuierliche Absenken der Steuerparameter und das stufenweise Absacken der Zielfunktion (z.B.  $Q = \sum V^T V$ ) gekennzeichnet.

Die Herabsetzung der Steuerparameter, in der Literatur häufig als Abkühlungsschema bezeichnet, erfolgt in der Regel logarithmisch, um verschiedene Größenordnungen des Suchraumes abzudecken. Die Verminderung von  $T(i)$  erfolgt durch den konstanten Steuerparameter  $R$  mit

$$T(i) = T_0 R^i.$$

$T_0$  ist ein extern übergebener Startwert. Die Konstante  $R$  muss im Bereich zwischen 0 und 1 liegen, beispielsweise 0.9995, um ein langsames, streng monoton Absinken von  $T(i)$  zu ermöglichen. Auch ein monoton fallendes  $T(i)$  ist denkbar. In diesem Fall wird  $T$  nicht in jeder, sondern in jeder  $k$ -ten Iteration einmal herabgesetzt (siehe [3, 39]).

Ist die Anzahl gespeicherter Punkte größer eins ( $H > 1$ ), so ergeben sich viele Möglichkeiten für die Konstruktion von Optimierungsverfahren. Mittels einer Liste von vorangegangenen Auswertungen der Zielfunktion besteht die Möglichkeit, die Zielfunktion zu modellieren. Es ist leicht einzusehen, dass, mit Zuhilfenahme dieser Informationen über die Zielfunktion, effizientere Optimierungsverfahren aufgebaut werden können.

Werden mehrere Punkte gleichzeitig konstruiert und erst anschließend bewert-



tet, spricht man von einem generationellen Verfahren (siehe [27, 9, 11, 33]). Die konstruierten Punkte durchlaufen gemeinsam die Zyklen der Konstruktion, Bewertung und Ersetzung, weshalb sie in Generationen mit der Generationsgröße  $H$  zusammengefasst werden. Die daraus resultierenden Verfahren und Untervarianten zur globalen Optimierung sind sehr vielschichtig. Die Strategien der generationellen Verfahren sind aus den Prinzipien des natürlichen Prozesses der Evolution abgeleitet. Deshalb können die generationellen, bzw. genetischen oder evolutionären Verfahren am treffendsten mit Hilfe von Fachausdrücken aus der Biologie beschrieben werden.

Evolutionäre Algorithmen sind stochastische Suchverfahren, die an die Prinzipien der natürlichen biologischen Evolution angelehnt sind. Sie arbeiten gleichzeitig auf einer Anzahl von potentiellen Lösungen, der Population von Individuen. Auf diese Individuen (Lösungen) wird das Prinzip "Der Starke überlebt" angewandt, um im Sinne einer Zielfunktion immer bessere Individuen zu erzeugen, die am Ende zu einer guten Lösung führen.

Durch einer Initialisierungsphase zu Beginn wird in einem Prozess über mehrere Generationen eine Suche durchgeführt. In jeder Generation wird eine Anzahl neuer Lösungen erstellt. Dies erfolgt durch die Auswahl von Individuen entsprechend ihrer Fitness, aus denen nachfolgend durch die Anwendung evolutionärer Reproduktionsoperatoren neue Individuen erzeugt werden. Diese Individuen werden in die Population eingefügt, wodurch eine neue Population entsteht. Die neue Population dient als Ausgangspunkt für die Erstellung neuer Individuen in der nächsten Generation. Dieser Prozess führt zur Evolution (Entwicklung) von Populationen von Individuen, die immer besser an die jeweilige Zielfunktion angepasst sind, als die Individuen, von denen sie erzeugt wurden. Dies entspricht der natürlichen Anpassung. Da evolutionäre Algorithmen mit Populationen von Individuen arbeiten, wird die Suche im Problemraum in einer parallelen Art und Weise durchgeführt. Evolutionäre Algorithmen modellieren verschiedene natürliche Prozesse, wie Selektion, Reproduktion, Rekombination, Mutation, Migration, Lokalität, Nachbarschaft, Parallelität und Konkurrenz. Jeder dieser Prozesse kann in einer Vielzahl von Varianten auftreten.

Die Abbildung 7.2 zeigt die Struktur eines einfachen evolutionären Algorithmus. Zu Beginn der Arbeit eines evolutionären Algorithmus erfolgt die Initialisierung. Diese beinhaltet neben der Parametrisierung vor allem die Erstellung der Anfangspopulation. Normalerweise werden die Individuen der Anfangspopulation zufällig im Definitionsbereich der Variablen initialisiert. Allerdings können zu diese Initialisierung auch problemspezifische Verfahren verwendet werden. Die erstellte Anfangspopulation wird durch die Zielfunktion bewertet. Damit ist die Population der ersten Generation produziert.

Wenn das definierte Abbruchkriterium noch nicht erreicht ist, beginnt die Erstellung einer neuen Generation und damit der evolutionäre Kreislauf. Jedem Individuum wird entsprechend seines Zielfunktionswertes und im Vergleich zu allen anderen Individuen der Population eine Fitness zugewiesen (Fitnesszuweisung). Entsprechend dieser Fitness werden die Individuen (Eltern) durch die Produktion von

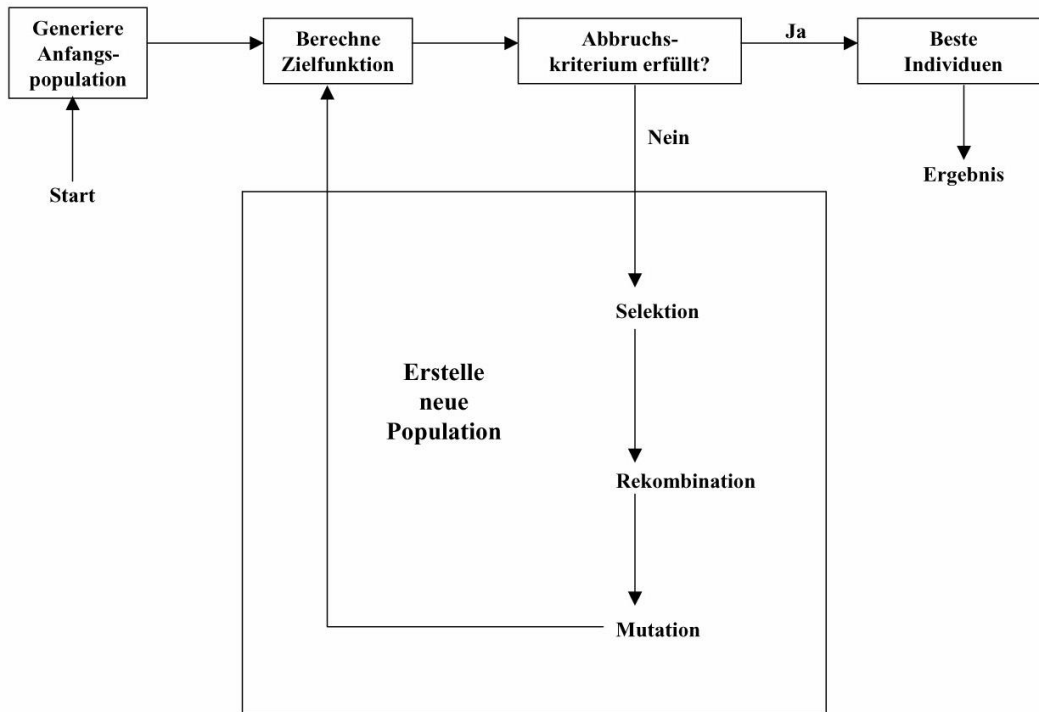


Abbildung 7.2: Struktur eines einfachen evolutionären Algorithmus

neuen Individuen (Nachkommen) ausgewählt (Selektion). Die Eltern produzieren unter Anwendung evolutionärer Operatoren (Rekombination, Mutation) die Nachkommen. Durch Rekombination werden die Nachkommen erstellt und diese werden dann durch Mutation mit einer gewissen Wahrscheinlichkeit verändert. Die neuen Individuen werden durch die Zielfunktion bewertet und in die Population durch die Nachkommen ersetzt werden. Damit ist eine neue Population erstellt. Wenn das Abbruchkriterium noch nicht erfüllt ist, beginnt der Prozess der Erstellung einer neuen Population von vorn.

In der Fitnesszuweisung wird festgelegt, wieviele Nachkommen jedes Individuum produziert. Die Selektion entscheidet, welche Individuen für die Fortpflanzung ausgewählt werden.

Die Konstruktion von Punkten, die hier als Individuen bezeichnet werden, erfolgt mit Hilfe von Paarung und Mutation. Während die Paarung neue Individuen erschafft, indem zwei Eltern (Individuen der alten Generation) rekombiniert werden, erfolgt die Mutation wahllos in Form von zufälligen Abänderungen der frisch erzeugten Kindergeneration.

**Bemerkung 7.2.1** Die Matlab *GEATbx* (Genetic and Evolutionary Algorithm) Toolbox vom Hartmut Pohlheim stellt die implementierte Verfahren zur Verfügung, bei den die Anzahl gespeicherter Punkte größer eins ( $H > 1$ ) ist.

# Kapitel 8

## Parameter Calibration

Mit den herkömmlichen Mitteln, wie analytischen Bewertungsformeln (siehe Abschnitt 4.2), lassen sich leicht die Preise für *standardisierte* Optionen bestimmen. Im Falle *exotischer* Optionen, die wegen ihrer Komplexität nicht mit herkömmlichen Mitteln zu bewerten sind, wird die Monte-Carlo Methode zu Hilfe genommen. Um die Monte-Carlo-Methode auf zahlreiche Bewertungsmodelle, die sich seit der bahnbrechende Arbeit von Black & Scholes rasant entwickelten, sinnvoll anwenden zu können, ist die Bestimmung der Modellparameter notwendig. Speziell die Anwendung der Monte-Carlo-Methode auf das *Bakshi* Modell hat die Bestimmung der Parameter

- $V_0$  initial Varianz,
- $\lambda$  durchschnittliche Anzahl der Sprünge in einem Jahr,
- $\mu_J$  Erwartungswert der Sprünge,
- $\sigma_J$  Varianz der Sprünge,
- $\theta_v$  das Quadrat der langfristigen instantanen Volatilität,
- $\kappa_v$  die Stärke der *mean reversion*,
- $\sigma_v$  vol of vol,
- $\rho$  Korrelation

zu folge.

Die Modellparameter werden mit Hilfe numerischer Optimierungsverfahren, aus impliziten Volatilitäten (vergleiche Kapitel 5) gewonnen, was auch der nächste Abschnitt verdeutlicht.

# Calibration Engine

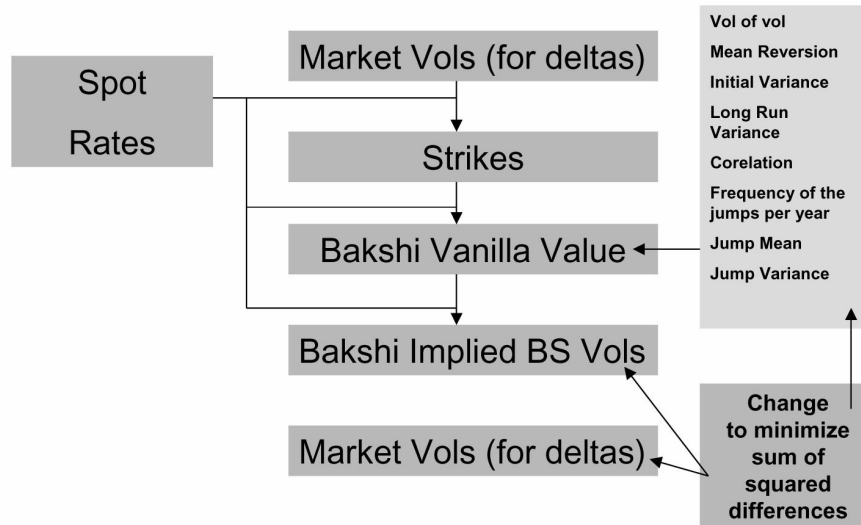


Abbildung 8.1: Calibration Engine

## 8.1 Parameter Calibration

### Ermittlung der Ausübungspreise

Wie man aus der Abbildung 8.1 (siehe [12]) entnehmen kann, ist ein möglicher erster Schritt zur Parameter Calibration die Ermittlung der *Black-Scholes Strikes*. Mittels vorgegebener Marktvolatilitäten  $\hat{\sigma}$  für verschiedener Deltas werden die *Black-Scholes Strikes* ermittelt (siehe Kapitel 5)

$$K_i = \text{BlackScholesStrikeRetriever}(\Delta_i, \hat{\sigma}_i, \text{Market}, \text{Call}).$$

### Berechnung des Call/Put

Der nächste Schritt ist die Berechnung der Call/Put Preise aus dem betrachteten Modell. Die Berechnung erfolgt z.B. wie im Modell von Heston oder im Modell von Bakshi, mittels analytischer Bewertungsformel

$$C_i = C(\text{Modell Parameter}, \text{Market}, K_i, \text{Call}).$$

### Implizite Volatilität

Die implizite Volatilität (IV) ist nach Black- Scholes die Zahl, die als Volatilitätsparameter in die Black-Scholes Formeln eingegeben wird, so dass der so berechnete

Optionspreis mit dem Marktpreis übereinstimmt (siehe Abschnitt 5.1). Zur Bestimmung der IV wird ein iteratives Optimierungsverfahren (z.B. Newton-Raphson, oder Bisektion Verfahren) herangezogen

$$\sigma_i = \text{ImpVolRetriever}(\text{value} = C_i, \text{Market}, K_i, \text{Call}).$$

Der in der Abbildung A.2 aufgelistete Algorithmus zur Bestimmung der impliziten Volatilitäten basiert auf der Newton-Raphson Methode (siehe Kapitel 7.1).

### Parameter Optimierung

Aus den berechneten impliziten Volatilitäten in beiden Modellen (Black-Scholes und Bakshi) werden durch Optimierungsverfahren (siehe Kapitel 7) die Modellparameter *kalibriert*

$$M(\text{Modell Parameter}) \triangleq \sum_{i=1}^n [\hat{\sigma}_i - \sigma_i(\text{Modell Parameter})]^2.$$

## 8.2 Untersuchungen des Bakshi Modells

In der Literatur ist zu beobachten, dass die Parameter Calibration im Heston Modell (siehe [12]), für kurz bis mittelfristige *one-touch*-Optionen, sehr gute Ergebnisse liefert. Im folgenden wird die Verhaltensweise des Bakshi-Modells anhand von Devisen (*FX*)-Optionen untersucht.

Die Devisen (*FX*)-Optionen geben dem Optionsinhaber das Recht, gegen Zahlung einer Prämie (Optionspreis) einen bestimmten Währungsbetrag zu einem vorgegebenen Zeitpunkt zu kaufen oder zu verkaufen. Der Verkäufer der Option erhält die Prämie und hat die Pflicht, den Währungsbetrag zum vereinbarten Preis zu liefern (Call-Option) oder abzunehmen (Put-Option), sofern der Optionsinhaber dies verlangt.

Die Marktdaten <sup>1</sup> *EUR/USD-spot* (0.9300), *domestic (USD) rate* (0.05) und *foreign (EUR) rate* (0.04) werden im folgenden als *Market* bezeichnet. Marktvolatilitäten sind in Tabelle 8.1 aufgeführt.

### Ermittlung der Ausübungspreise

Aus den vorgegebenen Marktvolatilitäten für verschiedene Deltas und unterschiedliche Laufzeiten (siehe Tabelle 8.1), werden die *Black-Scholes Strikes*  $K_i$  ermittelt. Das Delta ist durch die Gleichung

$$\Delta = \phi e^{-r_d T} \mathcal{N}(\phi d_+),$$

wobei

$$d_+ = \frac{\ln \frac{S}{K} + (r_d - r_f + \frac{1}{2} \hat{\sigma}^2) T}{\hat{\sigma} \sqrt{T}},$$

---

<sup>1</sup>Die Daten wurden von Dr. Uwe Wystup (Commerzbank AG) zur Verfügung gestellt.

mit  $r_d$  domestic rate,  $r_f$  foreign rate und entweder  $\phi = +1$  (Call) oder  $\phi = -1$  (Put), gegeben. Auflösung nach  $K$  ergibt

$$K_i = \text{BlackScholesStrikeRetriever}(\Delta_i, \hat{\sigma}_i, \text{Market}, \text{Put}),$$

oder explizit

$$K_i = S \exp \left\{ -\phi \mathcal{N}^{-1}(\phi \Delta_i e^{r_f \tau}) \hat{\sigma}_i \sqrt{\tau} + \hat{\sigma}_i \theta_+ \tau \right\}.$$

Für das  $\hat{\sigma}$  in der Gleichung werden nur die Werte im ATM Bereich berücksichtigt. Das Ergebnis verdeutlicht die Tabelle 8.2.

### Analytische Bewertungsformel

Die Berechnung des Put/Call aus dem *Bakshi Modell*

$$C_i = C(R_0, \lambda, V_0, \mu_J, \sigma_J, \theta_v, \kappa_v, \sigma_v, \rho, \text{Market}, K_i, \text{Put})$$

erfolgt mittels analytische Bewertungsformel (Vergleiche Abschnitt 4.2)

$$C(t, \tau) = S(t)B^*(t, \tau)\Pi_1(t, \tau; S, R, V) - KB(t, \tau)\Pi_2(t, \tau; S, R, V)$$

mit

$$\Pi_j(t, \tau; S(t), R(t), V(t)) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \text{Re} \left[ \frac{e^{-i\phi \ln[K]} f_j(t, \tau, S(t), R(t), V(t); \phi)}{i\phi} \right] d\phi,$$

und

$$B^*(t, \tau) = e^{-r_f(\tau-t)}.$$

**Bemerkung 8.2.1** Zur numerischen Berechnung der Integrale aus der obigen Gleichung wird wegen dem Definitionsbereich des Integrals  $(0, \infty)$  Gauss-Laguerre Quadratur zu Hilfe genommen.

Die Gauss-Laguerre Integration berechnet (siehe [14])

$$\int_0^\infty e^{-x} f(x) dx \approx \sum_{i=1}^n w_i f(x_i),$$

wobei die Stützstellen/Abzissen  $x_i$  die Nullstellen des Laguerre-Polynoms  $n$ -ten Grades  $L_n(x)$  sind und die Gewichtungsfaktoren  $w_i$  durch

$$w_i = \frac{A_{n+1} \gamma_n}{A_n L'_n(x_i) L_{n+1}(x_i)} = \frac{A_n}{A_{n-1}} \frac{\gamma_{n-1}}{L_{n-1}(x_i) L'_n(x_i)},$$

gegeben sind. Dabei ist  $A_n$  Koeffizient  $x^n$  in  $L_n(x)$ . Für Laguerre Polynome gilt

$$A_n = \frac{(-1)^n}{n!},$$

n	$x_i$	$W_i$	n	$x_i$	$W_i$
2	0.585786	0.853553	3	0.415775	0.711093
	3.41421	0.146447		2.29428	0.278518
4	0.322548	0.603154		6.28995	0.0103893
	1.74576	0.357419	5	0.26356	0.521756
	4.53662	0.0388879		1.4134	0.398667
	9.39507	0.000539295		3.59643	0.0759424
		7.08581		0.00361176	
			12.6408	0.00002337	

Abbildung 8.2: Gauss-Laguerre Nodes and Weights

so dass

$$\frac{A_{n+1}}{A_n} = -\frac{1}{n+1}, \quad \frac{A_n}{A_{n-1}} = -\frac{1}{n}.$$

Es gilt

$$\gamma_n = \int_0^\infty e^{-x} [L_n(x)]^2 dx = 1,$$

so dass

$$w_i = \frac{1}{(n+1)L_n'(x_i)L_{n+1}(x_i)} = -\frac{1}{nL_{n-1}(x_i)L_n'(x_i)}.$$

Die Rekursion liefert

$$xL_n'(x) = nL_n(x) - nL_{n-1}(x) = (x - n - 1)L_n(x) + (n + 1)L_{n+1}(x),$$

und da  $x_i$  die Nullstelle des  $L_n(x)$  ist,

$$nL_n(x_i) = (x_i - n - 1)L_n(x_i) = 0,$$

bekommt man

$$xL_n'(x_i) = nL_n(x_i) - nL_{n-1}(x_i) = (n + 1)L_{n+1}(x_i),$$

womit gilt

$$w_i = \frac{1}{x_i[L_n'(x_i)]^2} = \frac{x_i}{(n+1)^2[L_{n+1}(x_i)]^2}.$$

	Volatilities										Call (Vol)				
	10%	15%	20%	25%	30%	35%	40%	ATM	40%	35%	30%	25%	20%	15%	10%
1w	11.05	10.88	10.73	10.62	10.55	10.51	10.49	10.50	10.57	10.63	10.73	10.87	11.06	11.29	11.55
2w	10.95	10.74	10.57	10.45	10.37	10.34	10.33	10.35	10.42	10.48	10.59	10.75	10.96	11.23	11.55
1m	10.85	10.63	10.45	10.32	10.25	10.22	10.21	10.25	10.33	10.41	10.54	10.72	10.97	11.28	11.65
2m	10.95	10.72	10.55	10.42	10.35	10.32	10.31	10.35	10.43	10.51	10.64	10.83	11.07	11.38	11.75
3m	11.55	11.28	11.08	10.93	10.83	10.78	10.77	10.80	10.88	10.96	11.09	11.28	11.53	11.86	12.25
6m	12.30	11.99	11.74	11.58	11.47	11.43	11.42	11.45	11.52	11.60	11.73	11.93	12.20	12.56	13.00
9m	12.45	12.14	11.89	11.72	11.62	11.58	11.57	11.60	11.67	11.75	11.88	12.07	12.35	12.71	13.15
1y	12.75	12.39	12.11	11.93	11.82	11.77	11.77	11.80	11.86	11.94	12.07	12.27	12.57	12.96	13.45
18m	12.80	12.43	12.16	11.98	11.87	11.82	11.82	11.85	11.91	11.99	12.12	12.32	12.62	13.01	13.50
2y	12.85	12.48	12.21	12.02	11.92	11.88	11.87	11.90	11.95	12.04	12.17	12.38	12.67	13.06	13.55

Tabelle 8.1: Marktvolatilitäten (in % ) vom 05. Nov 2001 für EUR/USD

	Put (Strike)					Strikes					Call (Strike)				
	10%	15%	20%	25%	30%	35%	40%	50%	40%	35%	30%	25%	20%	15%	10%
1w	0.91	0.92	0.92	0.92	0.92	0.93	0.93	0.93	0.93	0.94	0.94	0.94	0.94	0.95	0.95
2w	0.91	0.91	0.91	0.92	0.92	0.92	0.93	0.93	0.94	0.94	0.94	0.94	0.95	0.95	0.96
1m	0.90	0.90	0.91	0.91	0.92	0.92	0.93	0.93	0.94	0.94	0.95	0.95	0.96	0.96	0.97
2m	0.88	0.89	0.90	0.91	0.91	0.92	0.92	0.93	0.94	0.95	0.95	0.96	0.97	0.98	0.99
3m	0.87	0.88	0.89	0.90	0.91	0.92	0.92	0.94	0.95	0.96	0.96	0.97	0.98	0.99	1.01
6m	0.84	0.86	0.88	0.89	0.90	0.91	0.92	0.94	0.96	0.97	0.98	0.99	1.01	1.03	1.06
9m	0.83	0.85	0.87	0.89	0.90	0.91	0.93	0.95	0.97	0.98	1.00	1.01	1.03	1.06	1.09
1y	0.81	0.84	0.86	0.88	0.90	0.91	0.93	0.95	0.97	0.99	1.01	1.03	1.05	1.09	1.13
18m	0.79	0.83	0.86	0.88	0.90	0.92	0.94	0.96	0.99	1.01	1.03	1.06	1.09	1.13	1.18
2y	0.78	0.82	0.86	0.88	0.91	0.93	0.95	0.98	1.00	1.03	1.05	1.08	1.12	1.17	1.24

Tabelle 8.2: Zugehörige Ausübungspreise ( $spot=0.93$ ,  $r_d=0.05$ ,  $r_f=0.04$ )



## Parameter Optimierung

Nun sind die Parameter  $R_0, \lambda, V_0, \mu_J, \sigma_J, \theta_v, \kappa_v, \sigma_v, \rho$  aus der analytische Bewertungsformel zu bestimmen, so dass die berechneten Put/Call Preise

$$C_i = C(\text{Modell Parameter, Market, } K_i, \text{Put})$$

mit den gegebenen Marktdaten der Minimierungsaufgabe

$$\sum_{i=1}^n [\hat{\sigma}_i - \sigma_i]^2 =! \min$$

genügen, wobei  $\hat{\sigma}_i$  die Marktvolatilitäten(Black-Scholes Modell) und

$$\sigma_i = \text{ImpVolRetriever}(\text{value} = C_i, \text{Market}, K_i, \text{Put})$$

die aus dem Bakshi Modell implizierte Volatilität für verschiedene Deltas bezeichnet.

Zur Bestimmung der Put/Call Preise wurde zuerst die *Gradienten* Methode (siehe Kapitel 7), die in der Matlab Funktion *fsolve*, aus der *Optimization Toolbox*, implementiert ist, zu Hilfe genommen. Parallel wurde auch die *simulated annealing* Methode eingesetzt. Die Ergebnisse, die in der Tabelle 8.3 sichtbar sind, waren identisch. Die Rechenzeit eines 500 MHz AMD Athlon Prozessors unter der Verwendung der Gradienten- Methode war wesentlich kürzer (160 sec.) als unter der Verwendung der *simulated annealing* Methode (5 min.).

Setzt man die errechneten *Bakshi Preise* zusammen mit den Marktdaten in die analytische Bewertungsformel ein, wird eine globale Optimierungsaufgabe, ein nichtlineares Gleichungssystem sichtbar, wobei die Modellparameter die Unbekannten sind. Zur Lösung der Optimierungsaufgabe wurden sowohl die generationellen Verfahren als auch das *simulated annealing* Verfahren (siehe Abbildungen A.3, A.4) eingesetzt.

Zuerst wurde die *simulated annealing* Methode eingesetzt. Die Ergebnisse wurden mit einem 500 MHz AMD Athlon PC, in ca. 17 min erzielt, und der Abschätzungsfehler lag bei  $10^{-4}$ . Parallel wurden die generationellen Verfahren herangezogen. Die lieferten ein etwas besseres Ergebnis (Abschätzungsfehler lag bei  $10^{-5}$ ), allerdings mit einer Laufzeit von 3 Stunden.

	Put			Call		
	$\Delta = 10\%$	$\Delta = 25\%$	ATM	$\Delta = 25\%$	$\Delta = 10\%$	
1w	4.7916e-004	0.00178858415893	0.00547977515134	0.00201683873027	0.00065774900046	
2w	0.00154428702895	0.00347119672720	0.00768685116676	0.00394900104916	0.00080573061663	
1m	0.00183699200369	0.00322018768440	0.01085557780390	0.00403585282093	0.00144649488003	
2m	0.00166549093800	0.00659854632613	0.01565560447354	0.00556929395712	0.00189215534697	
3m	0.00251994258925	0.00714250157047	0.01549133290519	0.00683728458962	0.00223058739389	
6m	0.00346658874463	0.01145862117050	0.02567160706754	0.01057975953768	0.00293185336873	
9m	0.00528840077521	0.01639341649361	0.02886468211957	0.01226567312661	0.00399587691333	
1y	0.00565375221126	0.01776635893276	0.03549799435775	0.01352614189396	0.00427058201315	
18m	0.00596728884497	0.02213945004832	0.03860402024442	0.01361601669921	0.00424028575707	
2y	0.00909343898567	0.02889516905727	0.04287589500875	0.01923424808154	0.00575266477381	

Tabelle 8.3: *Black-Scholes* Preise

	Put			Call		
	$\Delta = 10\%$	$\Delta = 25\%$	ATM	$\Delta = 25\%$	$\Delta = 10\%$	
1w	0.00047	0.00178	0.00562	0.00231	0.00068	
6m	0.00346	0.01145	0.02567	0.01057	0.00293	
18m	0.00602	0.02214	0.03861	0.013616	0.00424	

Tabelle 8.4: *Bakshi* Preise

initial Varianz ( $V_0$ )	0.046241
die Stärke der <i>mean reversion</i> ( $\kappa_v$ )	1.386352
durchschnittliche Anzahl der Sprünge in einem Jahr ( $\lambda$ )	0.946134
Erwartungswert der Sprünge ( $\mu_J$ )	-0.981478
Korrelation ( $\rho$ )	0.978040
Varianz der Sprünge ( $\sigma_J$ )	0.966891
vol of vol ( $\sigma_v$ )	0.811100
das Quadrat der langfristigen instantanen Volatilität ( $\theta_v$ )	0.339824

Tabelle 8.5: Kalibrierte Parameter aus dem *Bakshi* Modell angepasst für 1 Woche EUR/USD Volatilität *smile* (05. Nov 2001.)

Die besten Ergebnisse, die aus Tabellen 8.5, 8.6, 8.7, 8.8 entnehmen sind, wurden durch die Kombination der zwei eingesetzten Verfahren erzielt. Zuerst wurde ein grober Punkt mit der *simulated annealing* Methode gesucht, der dann weiter mit den generationellen Verfahren verfeinert wurde. Der Abschätzungsfehler ist der Abbildung 8.6 zu entnehmen (Best obj. vals per subpop). Die Graphik ganz rechts unten (Abbildung 8.6, order of subpops) zeigt dass die Optimierung mit fünf Unterpopulationen durchgeführt wurde. Jede der Unterpopulationen verwendete eine andere Strategie, die sich in der Grösse der verwendeten Mutationsschritte unterscheiden. Je kleiner die Mutationsschritte sind, um so lokaler ist die Suche dieser Strategie. Obere rechte Graphik (Obj. vals of all gen. (85 %)) stellt Zielfunktionswerte aller Individuen einer Population über mehrere Generation durch einen 2-D Farbentepich dar, und die mittlere Graphik unten (Obj. vals (85% best)) bildet Zielfunktionswerte aller Individuen einer Generation durch einen 2-D Punktdiagramm ab.

### Bemerkung

Genetische Algorithmen geben keinerlei Garantie für das Finden der optimalen Lösung. Ihre Effektivität ist sehr von der Größe der Population abhängig. Man hat hier also Tradeoff zwischen Qualität der Lösung und der Laufzeit. Für Simulated Annealing wurde ein Beweis der Konvergenz geführt, der auf der Cooling-Strategie basiert. Durch Manipulationen an dieser Strategie erlangt man eine Kontrolle über das Konvergenzverhalten “je langsamer die Abkühlung und somit mehr Iterationen desto besser wird die gefundene Lösung“. Bei genetischen Algorithmen hat man keinerlei Kontrolle über das Konvergenzverhalten. Weil Simulated Annealing zu jedem Zeitpunkt nur eine Lösung bearbeitet, geht beim Akzeptieren der neuen Lösung die alte Information verloren. Es gibt also keine Redundanz und keine Gedächtnis über verarbeitete Strukturen. Demzufolge können gute Lösungen verloren werden und wenn die Abkühlung zu schnell ist, nie mehr gefunden werden. Genetische Algorithmen verlieren auch Lösungen wegen zerstörerischer Wirkung der genetischen Operatoren. Sie machen dies teilweise wieder gut, indem sie verstärkt überdurchschnittlich gute Schemata durchforsten. Wird ein überdurchschnittlich gutes Schema übernommen, geht ein schlechtes Schema verloren, welches eventuell für Konstruktion eines optimalen Schemas unbedingt benötigt wird. Das kann zum Versagen des

initial Varianz ( $V_0$ )	0.077914
die Stärke der <i>mean reversion</i> ( $\kappa_v$ )	0.996879
durchschnittliche Anzahl der Sprünge in einem Jahr ( $\lambda$ )	0.315815
Erwartungswert der Sprünge ( $\mu_J$ )	0.997067
Korrelation ( $\rho$ )	0.759525
Varianz der Sprünge ( $\sigma_J$ )	0.999788
vol of vol ( $\sigma_v$ )	0.961096
das Quadrat der langfristigen instantanen Volatilität ( $\theta_v$ )	0.148484

Tabelle 8.6: Kalibrierte Parameter aus dem *Bakshi* Modell angepasst für 6 Monate EUR/USD Volatilität *smile* (05. Nov 2001.)

initial Varianz ( $V_0$ )	0.131013
die Stärke der <i>mean reversion</i> ( $\kappa_v$ )	2.139713
durchschnittliche Anzahl der Sprünge in einem Jahr ( $\lambda$ )	0.130836
Erwartungswert der Sprünge ( $\mu_J$ )	-0.248598
Korrelation ( $\rho$ )	0.810671
Varianz der Sprünge ( $\sigma_J$ )	0.651131
vol of vol ( $\sigma_v$ )	0.402012
das Quadrat der langfristigen instantanen Volatilität ( $\theta_v$ )	0.364460

Tabelle 8.7: Kalibrierte Parameter aus dem *Bakshi* Modell angepasst für 18 Monate EUR/USD Volatilität *smile* (05. Nov 2001.)

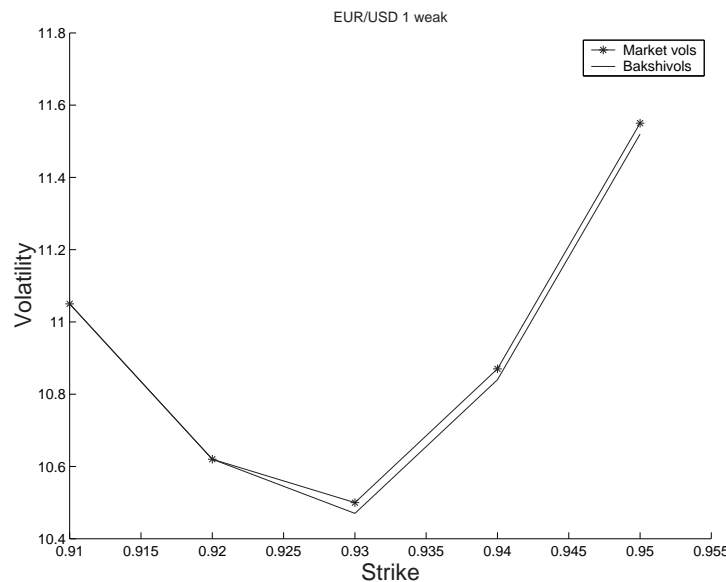


Abbildung 8.3: Markt *smile* (EUR/USD) von 05. Nov 2001 mit entsprechend angepassten *smile* aus dem *Bakshi* Modell

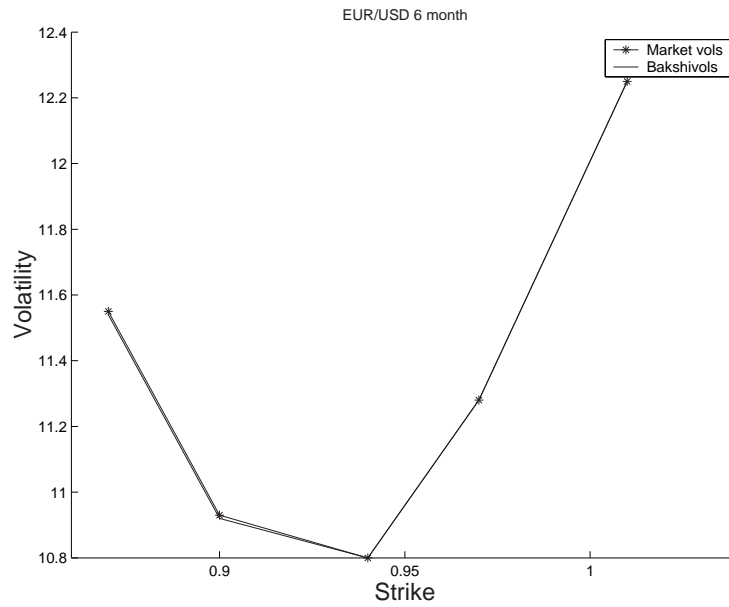


Abbildung 8.4: Markt *smile* (EUR/USD) von 05. Nov 2001 mit entsprechend angepassten *smile* aus dem *Bakshi* Modell

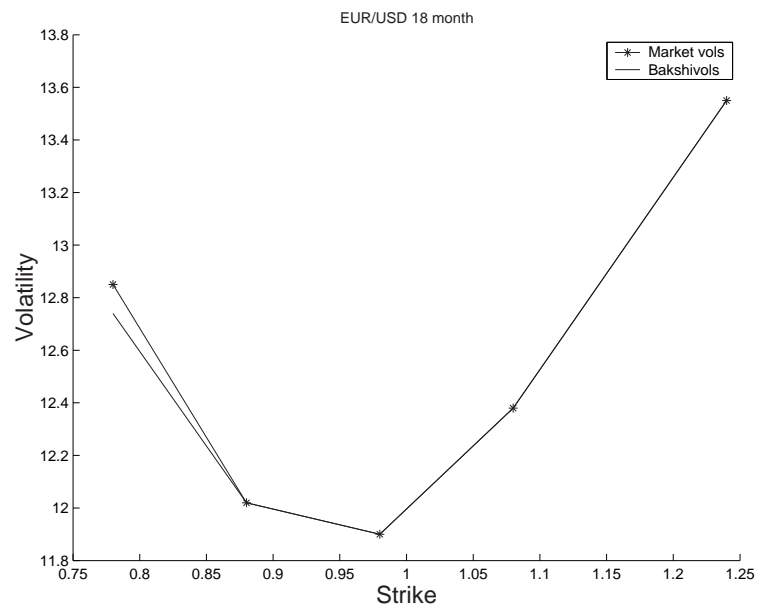


Abbildung 8.5: Markt *smile* (EUR/USD) von 05. Nov 2001 mit entsprechend angepassten *smile* aus dem *Bakshi* Modell

Verfahrens bei solchen täuschenden Problemen führen. Simulated Annealing hat bei dieser Art von Problemen auch Schwierigkeiten, weil es viel Zeit beim Besuchen der lokalen Maxima verlieren wird. Den Erkenntnissen nach ist somit eine Kombination der beiden Methoden zu empfehlen.

### **Bemerkung**

Die Parameter Calibration wurde ohne den Zinskomponenten durchgeführt. Der Zinssatz wurde als konstant angenommen ( $R_0 = r_d - r_f$ ).

Zu weiteren Untersuchungen wurden Marktdaten (EUR/USD) von 03. Sep 2002 genommen ( $EUR/USD\text{-spot}=0.9845$ ,  $domestic (USD) rate=0.0175$ ,  $foreign (EUR) rate=0.033$ ). Die zugehörigen Marktvolatilitäten sind in Tabelle 8.9 aufgeführt, und die errechneten Ausübungspreise mit oben angegebenen Marktdaten sind der Tabelle 8.10 zu entnehmen.

*Wie bisher in der Literatur zu lesen war (siehe [12]), vermutete man, dass die Parameter Calibration für kurz- und langfristige Optionsscheine mit einem stochastic-volatility jump diffusion Modell zu bewältigen ist. Die Untersuchungen, die am Bakshi Modell durchgeführt wurden, mit den Markt Daten von 05. Nov 2001., ergaben sehr gute Ergebnisse, die die Untersuchungen der Markt Daten von 03. Sep 2002. bekräftigten. Sie "ließen" sich für fast alle Laufzeiten gut kalibrieren, was auch Abbildungen 8.7, 8.8 verdeutlichen. Sehr bemerkenswert dabei ist, dass das Modell auch so eine starke Schiefe im Markt, die die Markt Daten von 03. Sep 2002 aufweisen, berücksichtigen kann. Die "Risk Reversals" sind in diesem Fall sehr hoch. Die besten Ergebnisse der Parameter Calibration (Markt Daten von 03. Sep 2002) sind Tabellen 8.11, 8.12, 8.13, 8.14 zu entnehmen.*

*Die Vermutungen aus der Literatur erwiesen sich durch die Untersuchung des Bakshi Modells bestätigt. Dies liegt vor allem daran, dass das Bakshi Modell (8 Parameter) beispielsweise im Vergleich zu Heston Modell (5 Parameter) wesentlich mehr Parameter aufweist, und man viel mehr "Spielraum" zu Calibration hat. Andererseits bedeutet es größeren Rechenaufwand mehrere Parameter zu kalibrieren.*

*Zu einem hat diese Arbeit einige Fragen geklärt, und zum anderen viele neue Fragen aufgeworfen:*

- *Wie kommt es zu solchen Werten?*
- *Wie sinnvoll ist es, überhaupt so viele Parameter zuzulassen und was gewinnt man dadurch?*

	Put (Vol)				Volatilities				Call (Vol)			
	10%	20%	25%	30%	40%	50%	40%	30%	25%	20%	10%	
1w	8.9	8.93	8.92	8.93	9.01	9.14	9.29	9.52	9.67	9.85	10.15	
2w	9.27	9.28	9.28	9.31	9.41	9.54	9.71	9.96	10.13	10.34	10.77	
1m	9.85	9.81	9.8	9.82	9.93	10.07	10.25	10.51	10.7	10.93	11.45	
2m	10.18	10.07	10.05	10.07	10.17	10.31	10.49	10.76	10.95	11.2	11.78	
3m	10.32	10.17	10.13	10.14	10.23	10.37	10.55	10.83	11.03	11.29	11.92	
6m	10.46	10.26	10.22	10.21	10.31	10.44	10.62	10.91	11.12	11.39	12.06	
9m	10.47	10.27	10.23	10.23	10.32	10.45	10.63	10.92	11.13	11.4	12.07	
1y	10.52	10.33	10.29	10.29	10.38	10.5	10.68	10.98	11.19	11.46	12.12	
18m	10.56	10.4	10.36	10.36	10.45	10.54	10.7	11	11.21	11.49	12.16	
2y	10.6	10.42	10.39	10.39	10.49	10.59	10.73	11.03	11.24	11.51	12.2	

Tabelle 8.8: Marktvolatilitäten (in %) vom 03. Sep 2002 für EUR/USD

	Put (Strike)				Strikes				Call (Strike)						
	10%	15%	20%	25%	30%	35%	40%	50%	40%	35%	30%	25%	20%	15%	10%
1w	0.97	0.97	0.97	0.98	0.98	0.98	0.98	0.98	0.98	0.99	0.99	0.99	1.00	1.00	1.00
2w	0.96	0.96	0.97	0.97	0.97	0.98	0.98	0.98	0.98	0.99	0.99	0.99	1.00	1.00	1.01
1m	0.95	0.95	0.96	0.96	0.97	0.97	0.98	0.98	0.98	0.99	0.99	1.00	1.00	1.01	1.02
2m	0.93	0.94	0.95	0.95	0.96	0.97	0.97	0.98	0.98	0.99	1.00	1.00	1.01	1.02	1.04
3m	0.92	0.93	0.94	0.95	0.95	0.96	0.97	0.98	0.98	0.99	1.00	1.01	1.02	1.02	1.05
6m	0.89	0.91	0.92	0.93	0.94	0.95	0.96	0.98	0.98	1.00	1.01	1.02	1.03	1.04	1.07
9m	0.87	0.89	0.90	0.92	0.93	0.94	0.95	0.98	0.98	1.00	1.01	1.02	1.04	1.05	1.09
1y	0.85	0.87	0.89	0.90	0.92	0.93	0.94	0.97	1.00	1.01	1.03	1.04	1.06	1.08	1.11
18m	0.83	0.85	0.87	0.89	0.91	0.92	0.94	0.97	1.00	1.01	1.03	1.05	1.07	1.10	1.13
2y	0.79	0.82	0.84	0.86	0.88	0.90	0.92	0.96	1.00	1.01	1.04	1.06	1.09	1.12	1.16

Tabelle 8.9: Zugehörige Ausübungspreise ( $spot=0.9845$ ,  $r_d=0.0175$ ,  $r_f=0.033$ )

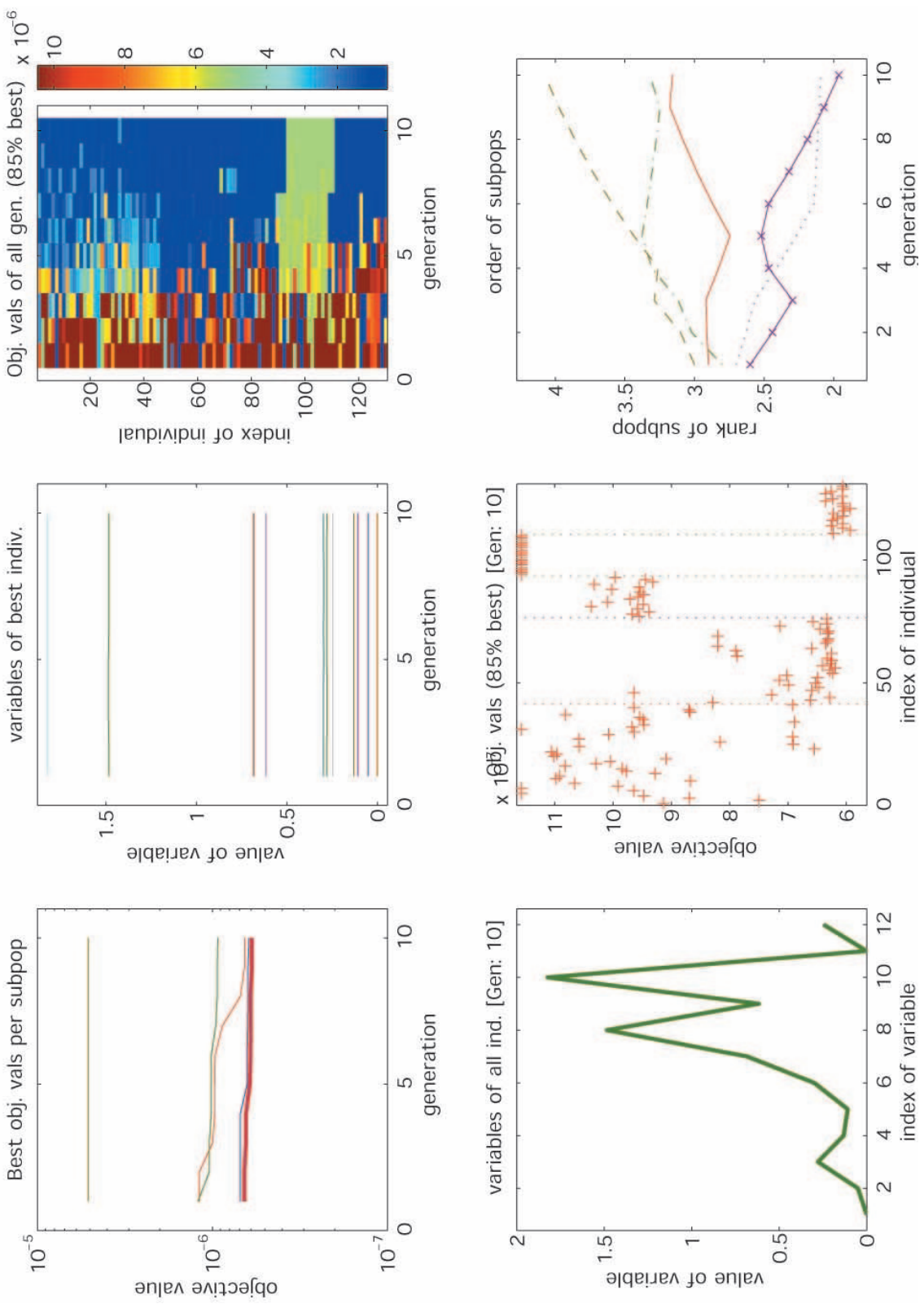


Abbildung 8.6: Ergebnisse der GEATbx Toolbox



initial Varianz ( $V_0$ )	0.060128
die Stärke der <i>mean reversion</i> ( $\kappa_v$ )	2.331390
durchschnittliche Anzahl der Sprünge in einem Jahr ( $\lambda$ )	0.609465
Erwartungswert der Sprünge ( $\mu_J$ )	-0.845508
Korrelation ( $\rho$ )	0.880420
Varianz der Sprünge ( $\sigma_J$ )	0.991871
vol of vol ( $\sigma_v$ )	0.663216
das Quadrat der langfristigen instantanen Volatilität ( $\theta_v$ )	0.262698

Tabelle 8.10: Kalibrierte Parameter aus dem *Bakshi* Modell angepasst für 1 Monat EUR/USD Volatilität *smile* (03. Sep 2002.)

initial Varianz ( $V_0$ )	0.077239
die Stärke der <i>mean reversion</i> ( $\kappa_v$ )	1.111555
durchschnittliche Anzahl der Sprünge in einem Jahr ( $\lambda$ )	0.593811
Erwartungswert der Sprünge ( $\mu_J$ )	0.556571
Korrelation ( $\rho$ )	0.804345
Varianz der Sprünge ( $\sigma_J$ )	0.734862
vol of vol ( $\sigma_v$ )	0.719007
das Quadrat der langfristigen instantanen Volatilität ( $\theta_v$ )	0.174820

Tabelle 8.11: Kalibrierte Parameter aus dem *Bakshi* Modell angepasst für 6 Monate EUR/USD Volatilität *smile* (03. Sep 2002.)

initial Varianz ( $V_0$ )	0.081636
die Stärke der <i>mean reversion</i> ( $\kappa_v$ )	2.450380
durchschnittliche Anzahl der Sprünge in einem Jahr ( $\lambda$ )	0.418159
Erwartungswert der Sprünge ( $\mu_J$ )	0.494872
Korrelation ( $\rho$ )	0.766670
Varianz der Sprünge ( $\sigma_J$ )	0.211674
vol of vol ( $\sigma_v$ )	0.644973
das Quadrat der langfristigen instantanen Volatilität ( $\theta_v$ )	0.140782

Tabelle 8.12: Kalibrierte Parameter aus dem *Bakshi* Modell angepasst für 1 Jahr EUR/USD Volatilität *smile* (03. Sep 2002.)

initial Varianz ( $V_0$ )	0.096623
die Stärke der <i>mean reversion</i> ( $\kappa_v$ )	0.603384
durchschnittliche Anzahl der Sprünge in einem Jahr ( $\lambda$ )	0.184013
Erwartungswert der Sprünge ( $\mu_J$ )	0.971422
Korrelation ( $\rho$ )	-0.098611
Varianz der Sprünge ( $\sigma_J$ )	0.238260
vol of vol ( $\sigma_v$ )	0.821778
das Quadrat der langfristigen instantanen Volatilität ( $\theta_v$ )	0.051608

Tabelle 8.13: Kalibrierte Parameter aus dem *Bakshi* Modell angepasst für 2 Jahre EUR/USD Volatilität *smile* (03. Sep 2002.)

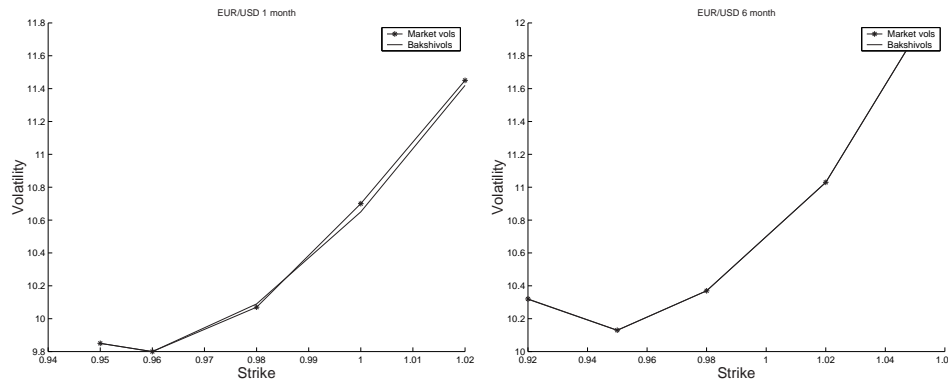


Abbildung 8.7: Markt *smile* (EUR/USD) von 03. Sep 2002 mit entsprechend angepassten *smile* aus dem *Bakshi* Modell

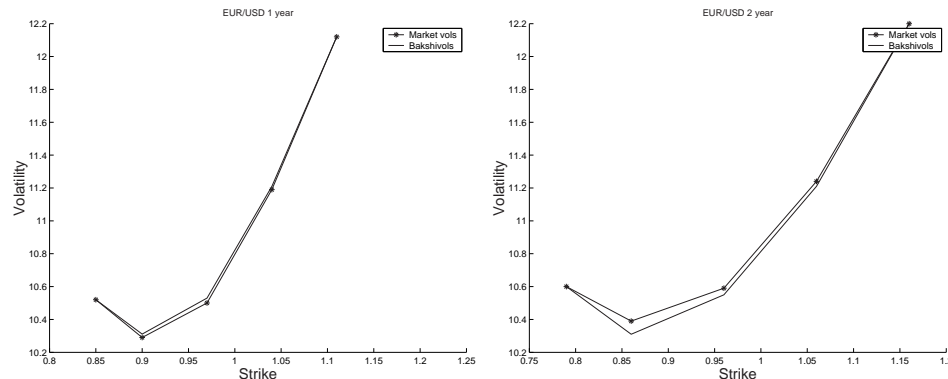


Abbildung 8.8: Markt *smile* (EUR/USD) von 03. Sep 2002 mit entsprechend angepassten *smile* aus dem *Bakshi* Modell

## Simulationen

Setzt man so errechnete Modellparameter in *stochastic-volatility jump diffusion Modell* von Bakshi (siehe [4])

$$dS(t) = S(t)[R(t) - \lambda\mu_J]dt + \sqrt{V(t)}S(t)dW_S(t) + J(t)S(t)dq(t)$$

$$dV(t) = [\theta_v - \kappa_v V(t)]dt + \sigma_v \sqrt{V(t)}dW_v(t),$$

das mittels Euler Verfahren implementiert wurde (siehe Abbildungen A.5,A.6,A.7), ein, bekommt man eine Trajektorie, einen simulierten *Bakshi* Pfad. Die Anwendung der Monte-Carlo Methode (1000 Pfade)

$$\frac{1}{1000} e^{-r_d T} \sum_{j=1}^{1000} (S_t^{(j)} - K)^+,$$

mit  $T = 2$  und Schrittweite  $\Delta = 1/360$  liefert Ergebnisse, die in Abbildung 8.9 zu sehen sind. Die Abbildung 8.9 sind tägliche Veränderungen des Optionspreises zu entnehmen. Somit kann man auch die Preise für exotische Optionen bestimmen.

Desweiteren wurden auch die Trajektorien mit höhere initial Volatilität ( $V_0 = 0.12$ ) simuliert (Abbildung 8.10). Vergleicht man die Abbildungen 8.9 und 8.10 wird einsichtig, dass die Einsetzung einer höherer initial Volatilität in das *stochastic-volatility jump diffusion Modell* von Bakshi, das mittels Euler Verfahren implementiert wurde, auch höheren Volatilität der simulierten Pfade zur Folge hat.

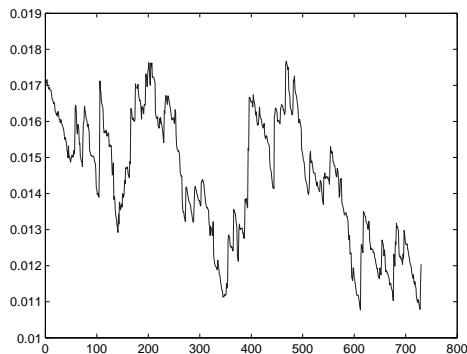


Abbildung 8.9: Simulierter Pfad mit errechneten Parametern unter Verwendung der Monte-Carlo Methode

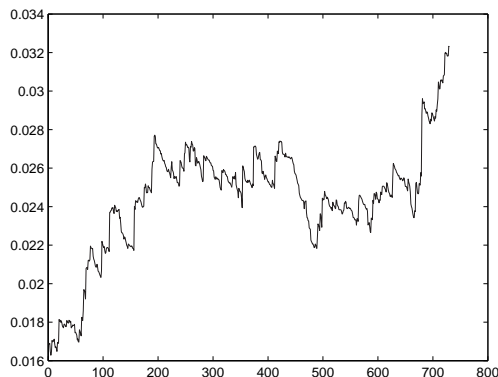


Abbildung 8.10: Simulierter Pfad mit errechneten Parametern (initial Vol = 0.12) unter Verwendung der Monte-Carlo Methode

## Weitere Untersuchungen

Die bisherigen Untersuchungen wurden mit einem konstanten Zinssatz durchgeführt, was konkret bedeutet dass alle Zinskomponenten aus der analytische Bewertungsformel gleich Null gesetzt waren. Die Versuchung auch die Zinskomponenten

- Zinssatz ( $R_0$ ),
- Rücktriebrate *rate of mean reversion* ( $\kappa_R$ ),
- Volatilität der Varianz ( $\sigma_R$ ),
- langfristige instantane Volatilität ( $\theta_R$ ),

in die Calibration mitreinzunehmen ergab erstaunliche Ergebnisse. Die Untersuchungen mit den Markt Daten von 05. Nov 2001., lieferten sehr gute Ergebnisse, allerdings nur für eine Laufzeit von zwei Jahren. Die Markt Daten von 03. Sep 2002. "ließen" sich auch für kürzere Laufzeiten, als zwei Jahren, gut kalibrieren.

Im folgenden sind die besten Ergebnisse der Parameter Calibration (siehe Tabelle 8.15, 8.16, 8.17, 8.18) angegeben und mit dazu gehörenden Graphiken belegt (siehe Abbildung 8.11, 8.12).

Zinssatz ( $R_0$ )	0.00162
initial Varianz ( $V_0$ )	0.0507713
Rücktriebrate <i>rate of mean reversion</i> ( $\kappa_R$ )	0.2819509
die Stärke der <i>mean reversion</i> ( $\kappa_v$ )	0.1349423
durchschnittliche Anzahl der Sprünge in einem Jahr ( $\lambda$ )	0.108597
Erwartungswert der Sprünge ( $\mu_J$ )	0.294545
Korrelation ( $\rho$ )	0.684248
Volatilität der Varianz ( $\sigma_R$ )	1.48316
Varianz der Sprünge ( $\sigma_J$ )	0.618109
vol of vol ( $\sigma_v$ )	1.825962
langfristige instantane Volatilität ( $\theta_R$ )	0.0028246
das Quadrat der langfristigen instantanen Volatilität ( $\theta_v$ )	0.245418

Tabelle 8.14: Kalibrierte Parameter aus dem *Bakshi* Modell angepasst für 2 Jahre EUR/USD Volatilität *smile*(05. Nov 2001)

*Obwohl die Parameter Calibration ziemlich gute Ergebnisse ohne die Zinskomponenten ergab, versuchte man, durch die Zunahme zusätzlichen Komponenten mit denen man mehr "Spielraum" zur Kalibrierung schuf, Verbesserungen zu erzielen. Als offene Fragen dieser Untersuchung blieben zu klären, weshalb die Hinzunahme einer stochastischen Zinskomponente die Kalibrierungsergebnisse eher verschlechtert haben und weshalb in diesem Fall die Marktdaten vom 5. November 2001 für einige Laufzeiten nicht kalibriert werden konnten.*

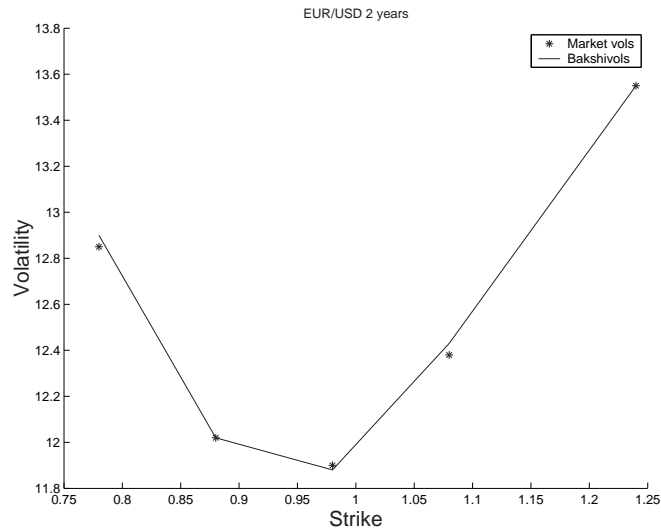


Abbildung 8.11: Markt *smile* (EUR/USD) von 05. Nov 2001 mit entsprechend angepassten *smile* aus dem *Bakshi* Modell

Zinssatz ( $R_0$ )	0.000428
initial Varianz ( $V_0$ )	0.017946
Rücktriebrate <i>rate of mean reversion</i> ( $\kappa_R$ )	0.504053
die Stärke der <i>mean reversion</i> ( $\kappa_v$ )	0.984509
durchschnittliche Anzahl der Sprünge in einem Jahr ( $\lambda$ )	0.559925
Erwartungswert der Sprünge ( $\mu_J$ )	1.086306
Korrelation ( $\rho$ )	0.862373
Volatilität der Varianz ( $\sigma_R$ )	1.647644
Varianz der Sprünge ( $\sigma_J$ )	1.148369
vol of vol ( $\sigma_v$ )	0.786723
langfristige instantane Volatilität ( $\theta_R$ )	0.939859
das Quadrat der langfristigen instantanen Volatilität ( $\theta_v$ )	0.219532

Tabelle 8.15: Kalibrierte Parameter aus dem *Bakshi* Modell angepasst für 2 Wochen EUR/USD Volatilität *smile* (03. Sep 2002)

Zinssatz ( $R_0$ )	0.017009
initial Varianz ( $V_0$ )	0.262286
Rücktriebrate <i>rate of mean reversion</i> ( $\kappa_R$ )	0.372711
die Stärke der <i>mean reversion</i> ( $\kappa_v$ )	1.922335
durchschnittliche Anzahl der Sprünge in einem Jahr ( $\lambda$ )	0.151429
Erwartungswert der Sprünge ( $\mu_J$ )	-0.107493
Korrelation ( $\rho$ )	0.797217
Volatilität der Varianz ( $\sigma_R$ )	0.345949
Varianz der Sprünge ( $\sigma_J$ )	0.778453
vol of vol ( $\sigma_v$ )	1.457889
langfristige instantane Volatilität ( $\theta_R$ )	0.335067
das Quadrat der langfristigen instantanen Volatilität ( $\theta_v$ )	0.113034

Tabelle 8.16: Kalibrierte Parameter aus dem *Bakshi* Modell angepasst für 9 Monate EUR/USD Volatilität *smile* (03. Sep 2002)

Zinssatz ( $R_0$ )	0.008915
initial Varianz ( $V_0$ )	0.035901
Rücktribrate <i>rate of mean reversion</i> ( $\kappa_R$ )	0.022184
die Stärke der <i>mean reversion</i> ( $\kappa_v$ )	0.764251
durchschnittliche Anzahl der Sprünge in einem Jahr ( $\lambda$ )	0.369916
Erwartungswert der Sprünge ( $\mu_J$ )	0.311224
Korrelation ( $\rho$ )	0.718154
Volatilität der Varianz ( $\sigma_R$ )	1.062600
Varianz der Sprünge ( $\sigma_J$ )	0.290927
vol of vol ( $\sigma_v$ )	1.327819
langfristige instantane Volatilität ( $\theta_R$ )	0.021845
das Quadrat der langfristigen instantanen Volatilität ( $\theta_v$ )	0.149454

Tabelle 8.17: Kalibrierte Parameter aus dem *Bakshi* Modell angepasst für 2 Jahre EUR/USD Volatilität *smile*(03. Sep 2002)

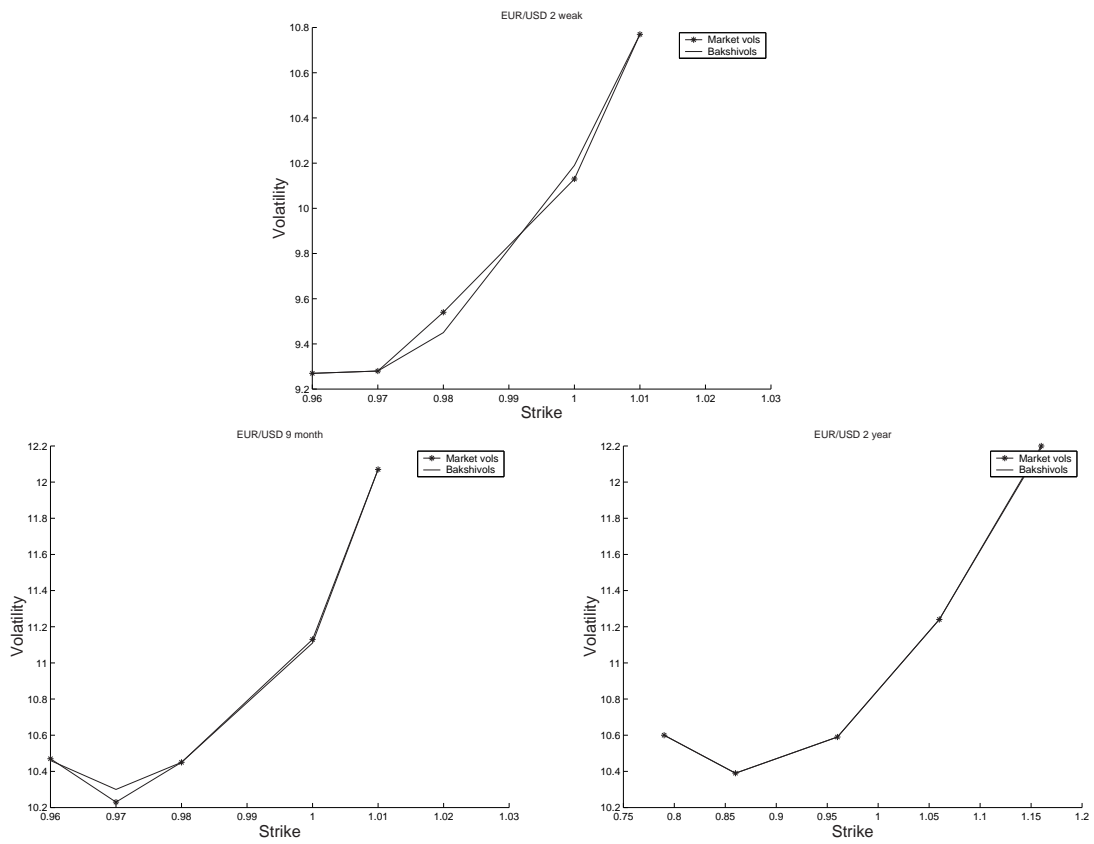


Abbildung 8.12: Markt *smile* (EUR/USD) von 03. Sep 2002 mit entsprechend angepassten *smile* aus dem *Bakshi* Modell





# Anhang A

## Programme

Folgende Programme sind im Matlab implementiert. Der Abbildung A.1 ist das Programm GLagNodeWt.m zu entnehmen. Das Programm berechnet Gauss-Laguere *nodes* und *weights*. In der Abbildung A.2 ist das Programm impvol.m dargestellt, was mittels Newton Methode die Implizite Vol's berechnet. Das Programm simann.m, was der Seite "<http://univie.ac.at/>" entnommen wurde, ist in den Abbildungen A.3 und A.4 aufgeführt. In den Abbildungen A.5, A.6 und A.7 ist die notwendige Apparatur zur Simulation des Bakshi Modells dargestellt.

14. Juni 2002

Page 1  
13:22:40

```
function [x,w] = GLagNodeWt(n)
% GLagNodeWt  Nodes and weights for Gauss-Laguerre quadrature
%
% Synopsis:   [x,w] = GLagNodeWt(n)
%
% Input:     n = order of quadrature rule
%
% Output:    x = vector of nodes
%            w = vector of weights

J = diag(1:2:2*n-1) - diag(1:n-1,1) - diag([1:n-1],-1); % Jacobi matrix
[V,D] = eig(J);
[x,ix] = sort(diag(D)); % nodes are eigenvalues, which are on diagonal of D
w      = 1*V(1,ix)'.^2; % V(1,ix)' is column vector of first row of sorted V
```

Abbildung A.1: GLagNodeWt.m

---

```

% IMPVOL Computes option implied volatilities
% USAGE
%   sigma=impvol(V,S,K,rd,rf,T,put,maxiters);
% INPUTS
%   V   : market value of option
%   S   : price of underlying asset
%   K   : strike price
%   rd  : domestic rate
%   rf  : foreign rate
%   T   : time-to-maturity
%   put : 0=call, 1=put
%
% OUTPUT
%   sigma : the implied volatility
%
% European options are priced via the Black-Scholes formula
% Newton's method used to find implied volatility

function sigma=impvol2(V,S,K,rd,rf,T,put,maxiters)
epsilon=sqrt(eps);
    if (V < 0)
        sigma=0;
    else
        saddle=sqrt(2/T*abs(log(S./K)+(rd-rf)*T));
        if saddle > 0
            sigma=saddle*0.9;
        else
            sigma=0.1;
        end
    end
tol=epsilon*V;
factor=S.*exp(-rf.*T).*sqrt(T/(2*pi));
S=exp(-rf.*T).*S;
K=exp(-rd.*T).*K;
logSK=log(S./K);
sT=sqrt(T);
for i=1:maxiters
    sig=sigma*sT;
    d1=logSK./sig + sig/2;
    Vhat=S.*normcdf(d1)-K.*normcdf(d1-sig);
    Vhat=(-put).*Vhat + (put).*(Vhat-S+K);
    vega=factor.*exp(-0.5*d1.*d1);
    res=Vhat-V;
    sigma0=sigma;
    sigma=sigma-res./vega;
    if all(abs(res))<=tol, break; end
end
end

```

Abbildung A.2: impvol.m

```
function min=simann(NVAR,NCHT,NT,NS,FACT,NINI,a,b)

% NCHT ... Number of Temperature Points
% NT ... Number of Configurations at constant Temperature
% NS ... Number of Configurations at constant Steplength
% FACT ... Factor controlling the cooling Process
% NINI ... Number of initial Configurations
% a,b ... Box-constraints

CK=1;
dummy=1;
% Evaluation of initial step-lengths:

v=(b-a)/10.0d0;

nquenall=0;
flagquol=0;
flag=0;
NITER=0;

% NINI random configurations:
breakcnt=1;
NITER=NITER+1;
Itcnt=0;
flagini=1;
flagmax=0;
flagmin=0;
nquenall=0;

x0= rancon(a,b,NVAR);

FOBJ=costfunc(NVAR,x0);
FOP=FOBJ;
FMAX=FOBJ;
XOP=x0;
FMAX=FOBJ;
flagini=1;
flagmax=1;
flagmin=1;

for I=1:NINI
    NITER=NITER+1;
    XT=rancon(a,b,NVAR);
    FOBJ= costfunc(NVAR,XT);

    if (FOBJ<FOP)
        XOP=XT;
        FOP=FOBJ;
    end
end
```

```
    if (FOBJ>FMAX)
        FMAX=FOBJ;
    end
end

T=(FMAX-FOP)^2/(CK*FMAX^2);
TMAX=T;
TMIN=T*FACT^NCHT;

% Optimization

XT=XOP;
XI=XOP;
FI=FOP;

% Temperature-Point-Loop

for K=1:NCHT
    disp('Temperature Point:');
    disp(K);
    FSUM=0.0;
    NTOT_T=0;
    NMBTR=0;

    for L=1:NT
        NSUCC=zeros(1,NVAR);
        for M=1:NS
            for J=1:NVAR
                NITER=NITER+1;
                ISUCC=0;
                R=rini02(dummy);
                R=R^2-1;
                XT(J)=XT(J)+(R*v(J));
                XT=CUT(NVAR,XT,a,b,K,L,M);
                FOBJ= costfunc(NVAR,XT);

                if (FI>FOBJ) % Improvement:
                    FI=FOBJ;
                    XI=XT;
                    NSUCC(J)=NSUCC(J)+1;
                    ISUCC=1;

                    if (FOP>FOBJ)
                        XOP=XT;
                        FOP=FOBJ;
                    end
                else % Metropolis:

```

Abbildung A.3: simann.m

```

P=exp((F1-FOBJ)/T);
R=rini02(dummy) ;

if (R<P)
F1=FOBJ;
XI=XT;
NSUCC(J)=NSUCC(J)+1;
NMETR=NMETR+1;
ISUCC=1;
else
XT=XI;
end
end

if (ISUCC==1)
FSUM=FSUM+FOBJ;
end

end % J-Loop
end % M-Loop

% Steplength-Update:
for I=1:NVAR
CHI=NSUCC(I)/NS;

if (CHI<0.4)
v(I)=v(I)/(1.0+2.0*(0.4-CHI)/0.4);
elseif (CHI>0.6)
v(I)=v(I)*(1.0+2.0*(CHI-0.6)/0.4);
end

% upper limit for steplength:
d=((b-a)/5.0);

if (v(I)>d(I))
v=d;
end
end

flagstep=0;

for i=1:NVAR

if (v(i)<1.0e-12)
flagstep=flagstep+1;
end
end

```

```

if (flagstep==NVAR)
min=XOP;
return
end

NTOTS=0;

for I=1:NVAR
NTOTS=NTOTS+NSUCC(I);
end

NTOT_T=NTOT_T+NTOTS;

end % L-Loop

if (NTOT_T == 0)
FMED=FSUM/NTOT_T;
else
FMED=0.0;
end

flagquol=nquenall;

XT=XOP;
XI=XOP;
FI=POP;
T=T*FACT;

end % K-Loop
min=XOP
return

```

Abbildung A.4: simann.m

```

% Bakshi Simulated Bakshi model
% USAGE
% rtz = bakshi(T,N,lambda,sigmav,Xzero,Bzero,R,muj,thetav,kv,sigmaj,rho);
% INPUTS
% T      : time to maturity
% lambda : frequency of the jumps per year
% sigmav : vol of vol
% Xzero  : spot
% Bzero  : inital variance
% R      : risk rate
% muj    : jump mean
% thetav : long run variance
% kv     : mean reversion
% sigmaj : jump variance
% rho    : corelation
%
% Euler's method used to simulated Bakshi model

function rtz = bakshi(T,N,lambda,sigmav,Xzero,Bzero,R,muj,thetav,kv,sigmaj,rho)
clf
normrnd('state',sum(100*clock));
Delta = T/N;
Xem = zeros(1,N+1);
Bum = zeros(1,N+1);
Bum(1) = Bzero;
Xem(1) = Xzero;
a=pois(T,lambda);b=a;
b(length(a)+1)=0;
b=sort(b);
u=[0:Delta:T];
ua=cell(length(a)+1,1);
bua=[];
for j = 1:length(a)
    ua{j}=[b(j):Delta:b(j+1)];
    bua=union(bua,ua{j});
    c(j)=a(j)-max(ua{j});
    d(j)=find(bua == max(ua{j}));
end
if length(bua) > N + 1
    while length(bua) > N
        k=find(c == min(c) & c > 0);
        bua(d(k))=[];
        c(k)=[];
    end
end
d(find(d == max(d)))=[];
bua(N+1)=T;
for j = 1:N
    Winc = sqrt(Delta)*normrnd(0,1);
    Xem(j+1) = (Xem(j) + Delta*(R-lambda*muj)*Xem(j) * sqrt(Bum(j))*Winc);
    if find(d == j+1)
        Xem(j+1)=real(Xem(j+1)+(exp(normrnd(log(1+muj)-1/2*sigmaj*sigmaj,sigmaj*sigmaj))-1));
    else
        Xem(j+1)=Xem(j+1);
    end
    Bum(j+1) = real(Bum(j)+(thetav-kv*Bum(j))*Delta+sqrt(Bum(j))*Winc*corr(Delta,rho,sigmav,Winc));
end

plot(real(bua),real(Xem),'r--')
xlabel('t','FontSize',16), ylabel('X','FontSize',16)
return

```

Abbildung A.5: bakshi.m

```

17. Juni 2002
Page 1
00:00:05
%Winccorr erzeugt korrelierte standardnormalverteilte ZV
%
%
function y = Winccorr(Delta,rho,sigmav,Winc)
    y= sigmav*rho*Winc + sigmav*sqrt(1 - rho*rho)*sqrt(Delta)*randn;

```

Abbildung A.6: Winccorr.m

```

17. Juni 2002
Page 1
00:00:24
% pois simulated a Poisson process
%
% USAGE
%   y = pois(Tmax,lambda);
% INPUTS
%   T       : time to maturity
%   lambda  : frequency of the jumps per year

function y = pois(Tmax,lambda)

clear T;
T(1)=random('Exponential',1/lambda);
i=1;

while T(i) < Tmax
    T(i+1)=T(i)+random('Exponential',1/lambda);
    i=i+1;
end

T(i)=Tmax;
y = T(1:i);

```

Abbildung A.7: pois.m

# Literaturverzeichnis

- [1] Abramowitz, M. and Stegun, C. A. **Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables** Dover New York, 1972.
- [2] L. Arnold, **Stochastische Differentialgleichungen** R. Oldenburg Verlag Munchen Wien, 1973.
- [3] R. Azencott, **Simulated annealing, parallelization techniques** Wiley New York, 1992.
- [4] G. Bakshi, C. Cao, Z. Chen, **Empirical Performance of Alternative Option Pricing Models** Journal of Finance, 1997.
- [5] H. Braun, **Neuronale Netze, Optimierung durch Lernen und Evolution** Springer Verlag Berlin, 1997.
- [6] I.N. Bronstein, K.A. Smendjajew **Taschenbuch der Mathematik** Harri Deutsch Verlag Thun und Frankfurt/Main, 1991.
- [7] S. Cyganowski, P.E. Kloeden, L. Grüne, **MAPLE for Jump-Diffusion Stochastic Differential Equations in Finance, Programming Languages and Systems in Computational Economics and Finance**, Kluwer Academic Publishers Amsterdam, 2002.
- [8] Deutsche Bank AG, **Exotische Optionsscheine**, Deutsche Bank AG, Frankfurt, 2000.
- [9] V.J. Frost, **Kalibrationsoptimierung mittels genetischer Algorithmen**, Univ. Diss., Duisburg, 2000.
- [10] I.I. Gichman, A.W. Skorochod, **Stochastische Differentialgleichungen** Akademie Verlag Berlin, 1971.
- [11] D.E. Goldberg, **Genetic algorithms in search, optimization, and machine learning** Addison-Wesley, 1989.

- [12] J. Hakala, U. Wystup, **Heston's Stochastic Volatility Model, Applied to Foreign Exchange Option** Foreign Exchange Risk, Risk Publications, London, 2002.
- [13] D.J. Higham, P.E. Kloeden, **Maple and Matlab for Stochastic Differential Equations in Finance** SIAM Review, 2001.
- [14] F.B. Hildebrand, **Introduction to Numerical Analysis** McGraw-Hill New York, 1956.
- [15] J.C. Hull, **Options, Futures and other Derivatives** Prentice Hall International Editions, 2000.
- [16] J. Jacod, A.N. Shiryaev, **Limit Theorem for Stochastic Processes** Springer Verlag Berlin , 1987.
- [17] F. Klebaner, **Introduction to Stochastic Calculus with Applications** Imperial College Press, 1999.
- [18] P.E. Kloeden, E. Platen, **Numerical Solution of SDE** Springer Verlag Berlin Heidelberg, 1992.
- [19] P.E. Kloeden, E. Platen, H. Schurz, **Numerical Solution of SDE Through Computer** Springer Verlag Berlin, 1994.
- [20] P. Kosmol, **Methoden zur numerischen Behandlung nichtlinearer Gleichungen und Optimierungsaufgaben** B.G. Teubner Stuttgart, 1993.
- [21] D. Lamberton, B. Lapeyre, **Introduction to Stochastic Calculus Applied to Finance** Chapman and Hall London, 1996.
- [22] Y. Magshoodi, **Exact Solutionen and Doubly Efficient Approximation of Jump-Diffusion Ito Equations** Stochastic Analysis and Applications, 1998.
- [23] Y. Magshoodi, C.J. Harris, **In-Probability Approximation and Simulation of Nonlinear Jump-Diffusion SDE** IMA Journal of Mathematical Control and Information, 1987.
- [24] Y. Magshoodi, **Mean Square Efficient Numerical Solution of Jump-Diffusion SDE** The Indian Journal of Statistics, 1996.
- [25] S.R. Mayer, **Exotische Optionen** IFA, Abt. Unternehmensplanung Ulm, 1996.
- [26] E. Polak, **Computational Methods in Optimization** Academic Press New York London, 1971.
- [27] H. Pohlheim, **Evolutionaere Algorithmen, Verfahren, Operatoren und Hinweise fuer die Praxis** Springer Verlag Berlin Heidelberg, 2000.



- [28] R. Seydel, **Einführung in die numerische Berechnung von Finanz-Derivaten** Springer Verlag Berlin Heidelberg New York, 2000.
- [29] K. Schäfer, **Optionsbewertung mit Monte-Carlo-Methoden** Verlag Josef Eul Bergisch Gladbach, 1994.
- [30] K. Schäfer, F.A. Adam-Müller, **Exotische Optionen** Wirtschaftswissenschaftliches Studium, Heft 11 , 27. Jg. 1998.
- [31] K. Schäfer, **Implizite Volatilität** Wirtschaftswissenschaftliches Studium, Heft 6, 26. Jg. 1997.
- [32] K. Schäfer, **Zur Bewertung von Devisen-Optionen** Wirtschaftswissenschaftliches Studium, Heft 3, 20. Jg. 1991.
- [33] M. Schwehm, **Globale Optimierung durch massiv parallele genetische Algorithmen** Dissertation, Universität Erlangen-Nürnberg, 1996.
- [34] D.L. Snyder, M.I. Miller, **Random Point Processes in Time und Space** Springer Verlag New York, 1991.
- [35] I.M. Sobol, **Die Monte-Carlo-Methode** Deutscher Verlag der Wissenschaften Berlin, 1991.
- [36] K. Spanderen, **Monte-Carlo-Simulationen im Risikomanagement für derivative Finanzinstrumente** WestLB Frankfurt, 2002.
- [37] Sparkassen RatgeberService, **Basisinformationen über Termingeschäfte** Bank Verlag Köln, 2001.
- [38] P. Spellucci, **Numerische Verfahren der nichtlinearen Optimierung** Birkhäuser Verlag Basel Boston Berlin, 1993.
- [39] P.J.M. van Laarhoven, E.H.L. Aarts, **Simulated annealing : theory and applications** Kluwer Dordrecht, 1992.
- [40] U. Wystup, **Numerical Methods for Derivatives Pricing** Vorlesungsskript, Goethe-Universität, 2000.
- [41] U. Wystup, **Vanilla Options** Foreign Exchange Risk, Risk Publications, London, 2002.
- [42] U. Wystup, **Cholesky Decomposition** Formula Catalogue, Mathfinance, <http://www.mathfinance.de>, 1999.



# Ehrenwörtliche Erklärung

Ich versichere ehrenwörtlich durch meine Unterschrift, dass ich die vorstehende Arbeit selbständig und ohne fremde Hilfe angefertigt und alle Stellen, die ich wörtlich oder annähernd wörtlich aus Veröffentlichungen entnommen habe, als solche kenntlich gemacht habe, mich auch keiner anderen als der angegebenen Literatur oder sonstigen Hilfsmittel bedient habe. Die Arbeit hat in dieser oder ähnlichen Form noch keiner anderen Prüfungsbehörde vorgelegen.

Frankfurt am Main